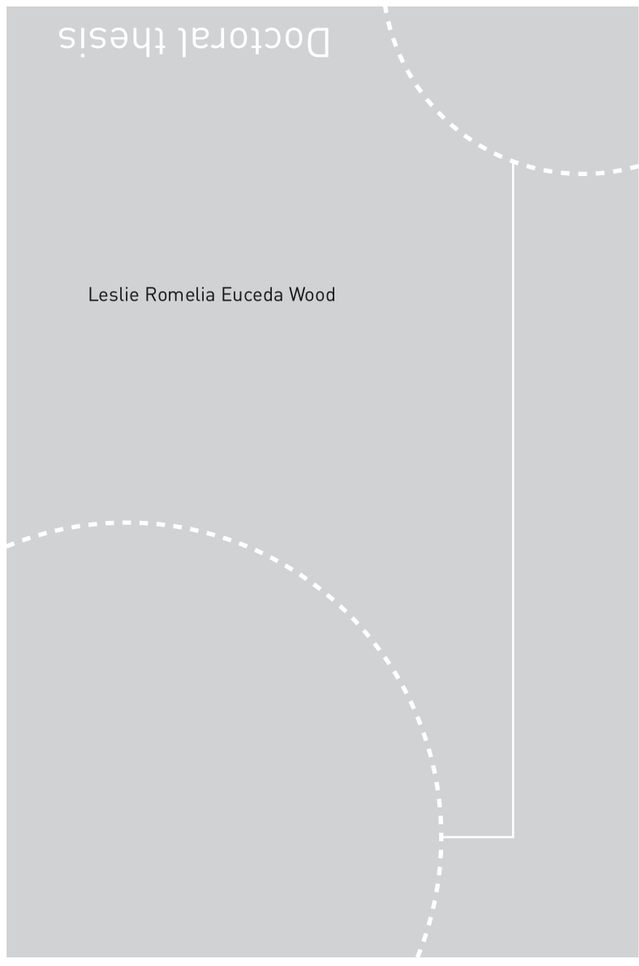
Doctoral theses at NTNU, 2016:373

Leslie Romelia Euceda Wood

Metabolic characterization of breast cancer heterogeneity and response to treatment



ISBN 978-82-326-2090-6 (printed ver.)
ISBN 978-82-326-2091-3 (electronic ver.)
ISSN 1503-8181

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Medicine
Department of Circulation and Medical Imaging

Doctoral theses at NTNU, 2016:373

NTNU

Leslie Romelia Euceda Wood

Metabolic characterization of breast cancer heterogeneity and response to treatment

Thesis for the Degree of Philosophiae Doctor

Trondheim, December 2016

Norwegian University of Science and Technology
Faculty of Medicine
Department of Circulation and Medical Imaging

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Medicine
Department of Circulation and Medical Imaging

ISBN 978-82-326-2090-6 (printed ver.)
ISBN 978-82-326-2091-3 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2016:373

Printed by NTNU Grafisk senter

Metabolsk karakterisering av brystkreft – heterogenitet og respons på behandling

På verdensbasis er brystkreft den mest vanlige kreftformen som rammer kvinner. Heterogenitet og kompleksitet i brystkreftens biologi gjør at pasienter med samme kliniske diagnose opplever ulik prognose og behandlingsrespons. Det er derfor nødvendig å finne mer effektive og sikrere behandlingsregimer basert på brystkreftens individuelle egenskaper. Endret metabolisme (stoffskifte) er ett av de nyeste beskrevne kjennemerkene på kreft. Denne endringen i metabolismen er nødvendig for å tilfredsstille kreftcellenes krav til energi og biomasse, slik at de kan fortsette å vokse (proliferere). Metabolsk profilering eller metabolomikk involverer storskalastudier av cellulære metabolitter og deres mellomprodukt som er representert i levende system ved et gitt tidspunkt. Metabolsk profilering kan blant annet gjøres ved hjelp av magnetisk resonans spektroskopi (MRS), og denne teknikken har blitt mye brukt for å måle metabolitter i intakte vevsprøver fra ulike kreftformer.

Hovedmålet i denne avhandlingen har vært å bruke metabolomikk for å øke kunnskapen om heterogenitet i brystkreft og for å identifisere muligheter for å optimalisere behandlingen av brystkreft. Som en del av forberedelsene til metabolomikk-studiene ble det utført et omfattende litteratursøk på pre-prosessering av MRS-data. Dette resulterte i en oversiktsartikkel som er inkludert i denne avhandlingen (artikkel I). Artikkelen peker på de mest vanlige pre-prosesseringstrinnene som benyttes for å analysere MR-baserte metabolske data. Gode metoder for preprosessering er nødvendig for å skille viktig biologisk informasjon fra irrelevant variasjon forårsaket av for eksempel eksperimentelle og instrumentelle artefakter.

I artikkel II er tre metabolske undergrupper av brystkreft etablert og metabolomikk-dataene kombinert med gen- og proteinutrykk fra samme tumor. Signifikante forskjeller i både gen- og proteinnivå ble observert mellom de metabolske gruppene. Den mest ulike gruppen viste nedregulering av gener som relateres til kollagener og ekstracellulær matriks. Dette er i samsvar med mangelen av mikromiljørelaterte sub-typer av protein i denne gruppen.

I artikkel III ble de metabolske effektene av neoadjuvant kjemoterapi med eller uten angiogenesehemmeren bevacizumab i brystkreftpasienter, undersøkt. Det ble observert klare metabolske forandringer som en effekt av kjemoterapi. Ved operasjonstidspunktet var de metabolske profilene i tumorene forskjellige i pasienter som hadde god respons på behandlingen, og de som ikke responderte. En liten nedgang i antioksidanten glutation ble

observert som en effekt av bevacizumab, noe som kan være knyttet til en redox-destabiliserende, apoptose-fremmende rolle for legemiddelet i kreftceller.

I artikkel IV ble det brukt dyremodeller av trippel-negativ brystkreft for å evaluere behandling med mTOR inhibitoren everolimus. Signifikante forskjeller i det metabolske uttrykket ble påvist med behandling. Stor variasjon i det metabolske uttrykket mellom de ulike dyremodellene gjorde det imidlertid vanskelig å påvise klare forskjeller knyttet til god eller dårlig respons på behandlingen, men den metabolske effekten etter behandling var imidlertid mer uttrykt ved respons.

Samlet sett viser avhandlingen at MRS metabolomikk kan ha fremtidig nytte for stratifisering av pasienter til målrettet behandling. Metoden utfyller dessuten andre molekylære analyser slik at man får et mer komplett bilde på underliggende mekanismer som regulerer behandlingsrespons.

Kandidat: Leslie Romelia Euceda Wood

Institutt: Institutt for sirkulasjon og bildediagnostikk

Veiledere: Prof. Tone F. Bathen, Dr. Guro F. Giskeødegård og Prof. Lutgarde M.C. Buydens

Finansieringskilde: Norges forskningsråd

Ovennevnte avhandling er funnet verdig til å forsvares offentlig for graden

Philosophiae Doctor i medisinsk teknologi.

Disputas finner sted i Auditoriet MTA, Medisinsk-Teknisk Forksningssenter,

Onsdag 21. Desember 2016 kl. 12:15.

Acknowledgement

The work presented in this thesis has been carried out at the MR Cancer group, Department of Circulation and Medical Imaging, the Norwegian University of Science and Technology (NTNU), between June 2013 and September 2016. Financial support was provided by the Research Council of Norway.

I would like to thank the women who have participated in the research studies included in this thesis, without whom this work would not have been possible.

So many people have helped and supported me to achieve this. I would like to thank my supervisors, Prof. Tone Bathen and Dr. Guro Giskeødegård. Tone, no matter how challenging things got at times, knowing you were there, I never once doubted that a solution could be found. You always remained accessible to provide guidance all throughout my project with an encouraging demeanor. You lead in a way that motivates me to do things not just to fulfill a requirement, but to give it my very best because I do not get any less from you. Guro, I admire the way you can make the most complicated things so simple and understandable. I have really enjoyed learning from you; you have helped me improve thinking critically whether when performing statistical analyses or writing (although I still have to work on my wordiness…). Thank you both so much!

I would like to express my gratitude to Professor Lutgarde Buydens, for welcoming me for a research stay at her group at Radboud University Nijmegen, The Netherlands. Thank you to Dr. Geert Postma and Dr. Jasper Engel for all the help during and after my stay, and to all the members of the group there for not only making my stay fruitful with regards to how much I learned, but also for showing me a fun time both in and out of office.

I feel fortunate to work at the MR Center with such talented people every day who create such a positive work environment. Thanks to all of you, especially to past and present colleagues at the MR Cancer group. Exchanging knowledge, experiences, cultures, and laughs with all of you has been so valuable, and has played an important role in my enjoying my PhD as much as I have. Thank you Siver for guidance and interesting discussions, especially regarding xenografts and biological pathways. Thank you Trygve, Santosh, and Øystein for providing training, guidance, and troubleshooting assistance during HR MAS MRS analyses. Special thanks to Tonje for her continuous support, from guiding me on my first days in the group to helping me with thesis submission and defense preparations. Thank you also to

Debbie and Neil for useful feedback, discussions, and pep talks. To my office mates: Ailin, Elise, Maria Karoline, Marie, Tonje, Torfinn, and Saurabh, it has been so much fun sharing an office with you, discussing things both related to work and not, and also celebrating each other's successes, and encouraging each other when we needed it.

Thank you to my co-authors and collaborators, especially to Dr. Olav Engebråten, Dr. Anne-Lise Børresen-Dale, Dr. Laxmi Silwal-Pandit, and Dr. Elisabetta Marangoni, for taking time to discuss my work and give suggestions to improve it.

I would also like to thank my friends and family for their support. Special thanks go to Gustavo, Miriam, Glenda, Allan, and Ana Romelia, for always encouraging me to go for what I want, and for believing that I can do anything I set my mind to. Donde quiera que yo vaya, ahí siempre están ustedes. Los quiero!

Last but not least, thanks to my unsung hero, my dear Ion, for your unconditional support and understanding, for doing everything you could to make things easier for me during this period. You balance out all the stress and make me happy every day. Thanks for always reminding me everything is going to be OK. Te iubesc guapo!

Leslie Romelia Euceda Wood
Trondheim, December 2016

Summary

Breast cancer is the most common cancer in women worldwide. Due to the heterogeneity and complexity of breast cancer tumor biology, patients with similar clinical diagnosis might have different prognosis and response to treatment. More efficient and safer breast cancer therapeutic regimens based on the individual tumor properties are necessary. In parallel, defining new strategies for patient stratification and early detection of response/resistance to treatment is warranted.

Altered metabolism is one of the most recently described hallmarks of cancer, and provides cancer cells the means to satisfy their high energy requirements, to synthesize macromolecule precursors, and to ensure proper redox stability, allowing them to survive, grow, and proliferate. The aberrant metabolic characteristics observed in cancer cells include increased glucose uptake and glycolytic lactate production even when enough oxygen is available for oxidative phosphorylation, the more efficient energy-generating pathway predominantly occurring in normal, non-proliferating cells. In addition, cancer cells become dependent on the normally non-essential amino acid glutamine as an alternative source of energy and macromolecular biosynthesis.

Metabolic profiling, or metabolomics, involves the large scale study of cellular metabolites and their intermediates present within a living system at a given moment. Metabolites represent the downstream output of genes, RNA, and proteins, and so metabolomics handles the final step in the omics cascade, conveying information closer to the phenotype. Metabolic profiles can be acquired using magnetic resonance spectroscopy (MRS). When analyzing tumor material, high resolution (HR) magic angle spinning (MAS) MRS generates highly resolved spectra from intact tissue. This technique has been widely applied to study metabolites involved in important pathways in different cancers.

This thesis aimed to use HR MAS MRS metabolomics to gain insight into the heterogeneity of breast cancer and to identify possibilities for optimizing treatment of the disease. Data from different omics levels and advanced statistical techniques were employed for this purpose. In preparation of the metabolomics studies carried out as part of this thesis, an extensive literature search on preprocessing of MRS metabolomics data was performed. This resulted in a review article which has been included as the first paper in this thesis. Whatever the approach used to extract biological meaning from data resulting from metabolomics studies, proper preprocessing is essential to remove irrelevant variation such as experimental and instrumental artifacts. This first paper outlines the most common

preprocessing steps when analyzing MRS metabolomics data, and compares different methods to achieve these.

In the second paper, three metabolic clusters of breast cancer were established and combined with gene and protein expression data from the same tumor. Significant differences at the genetic and protein levels were detected between the metabolic clusters. The metabolic cluster most different from the rest exhibited downregulation of genes related to collagens and extracellular matrix, which was in accordance with the lack of microenvironment-related protein subtypes within this cluster. Importantly, therapies that target different metabolic aberrations observed in separate clusters have already been developed.

The third paper investigated the metabolic effects of neoadjuvant chemotherapy with and without the antiangiogenic drug bevacizumab in HER2 negative breast cancer patients. Clear metabolic changes as an effect of chemotherapy, independent of bevacizumab administration, were observed, indicating a decline in glucose consumption and a transition to normal breast adipose tissue. Metabolic profiles significantly discriminated pathological responders from non-responders after treatment, but not before or during treatment. A subtle decrease in the antioxidant glutathione was detected as an effect of bevacizumab, suggesting a redox-destabilizing, apoptosis-promoting role for the drug in cancer cells.

In the fourth paper, breast tumor tissue from triple negative breast cancer patient-derived xenografts (PDX) was analyzed to evaluate treatment with the mTOR inhibitor everolimus. Treated animals could be significantly discriminated from untreated controls based on metabolite information, indicating reduced glycolytic lactate production and glutaminolysis after treatment, consistent with PI3K/AKT/mTOR pathway inhibition. Innate metabolic heterogeneity between different PDX models was detected, and appeared to impair prediction of treatment response. However, the metabolic effects following treatment were more prominent in responding xenografts compared to non-responders.

In conclusion, MRS metabolomics showed promise for improving breast cancer patient stratification to targeted therapy and for complementing other molecular assays to gain further insight into the underlying mechanisms regulating treatment response. The metabolic signatures identified in this thesis related to long-term clinical outcomes should be further investigated when follow-up data for the studied cohorts is available.

Symbols and Abbreviations

Symbol	Description	Page
^{13}C	Carbon-13 nucleus	42
^{1}H	Hydrogen-1 nucleus	15
α	Lower energy state	15
β	Higher energy state	15
δ	Chemical shift	17
μ	Magnetic moment	15
ω_0	Larmor frequency	15
γ	Gyromagnetic ratio	16
Acetyl CoA	Acetyl-coenzyme A	9
AsLS	Asymmetric least squares smoothing	19
ATP	Adenosine triphosphate	8
$\boldsymbol{b}$	Vector of regression coefficients	23
B_0	External magnetic field	15
B_1	Radio frequency pulse	16
cDNA	Complementary DNA	44
Cho	Choline	12
CK-α	Choline kinase alpha	13
cpmg	Carr-Purcell-Meiboom-Gill	60
CR	Complete response	7
CT	Computerized tomography	7
CV	Cross validation	26
DAB	3,3′-diaminobenzidine	45
DAVID	Database for Annotation, Visualization, and Integrated Discovery	33
DNA	Deoxyribonucleic acid	1
DSS	2,2-dimethyl-2-silapentane-5-sulfonate	18
$\boldsymbol{e}$	Residual vector	23
$\boldsymbol{E}$	Residual variance matrix	21
ER	Estrogen receptor	4
ERETIC	Electronic reference to access *in vivo* concentration	29
$\boldsymbol{F}$	Matrix of score residuals	24
$FADH_2$	Reduced form of flavin adenine dinucleotide	9
FDA	Food and Drug Administration	6
FDG	^{18}fluorodeoxyglucose	80
FDR	False discovery rate	31
FEC100	5-fluorouracil 600 mg/m2, epirubicin 100 mg/m2, cyclophosphamide 600 mg/m2	39
FID	Free induction decay	16
FN	False negative	28
FNAC	Fine needle aspiration cytology	2
FP	False positive	28
FWER	Family-wise error rate	31
GC	Gas chromatography	60
GLS	Glutaminase	11
GO	Gene ontology	33
GPC	Glycerophosphocholine	12
GR	Good response	40
GSEA	Gene set enrichment analysis	33

GSH	Glutathione	11
HCA	Hierarchical cluster analysis	20
HER2	Human epidermal growth factor receptor 2	4
HES	Hematoxylin, erythrosine, and saffron	62
HIF	Hypoxia-inducible factor	73
HR	High resolution	18
HRP	Horseradish peroxidase	45
I	Spin	14
IL-8	Interleukin 8	77
IR	Intermediate response	40
K	Kelvin	15
KEGG	Kyoto Encyclopedia of Genes and Genomes	34
LABC	Locally advanced breast cancer	4
LDA	Linear discriminant analysis	25
LDH	Lactate dehydrogenase	9
LMM	Linear mixed-effects model	30
LOO	Leave-one-out	26
LV	Latent variable	24
M_0	Net magnetization vector	15
MAS	Magic angle spinning	18
Mc	Metabolic cluster	48
MICE	Multivariate imputation by chained equations	43
MLR	Multiple linear regression	23
MR	Magnetic resonance	19
MRI	Magnetic resonance imaging	2
MRS	Magnetic resonance spectroscopy	14
MS	Mass spectrometry	14
mTOR	Mammalian target of rapamycin	6
mTORC1	Mammalian target of rapamycin complex 1	6
NADH	Nicotinamide adenine dinucleotide	9
NADPH	Reduced form of nicotinamide adenine dinucleotide phosphate	9
NBCG	Norwegian Breast Cancer Group	6
NIPALS	Non-linear iterative partial least squares	25
NMR	Nuclear magnetic resonance	14
non-ox-PPP	Non-oxidative pentose phosphate pathway	9
NR	No response	40
OS	Overall survival	79
OSBREAC	The Oslo Breast Cancer Consortium	48
ox-PPP	Oxidative pentose phosphate pathway	9
$\boldsymbol{P}$	Loadings matrix	21
pAKT	Phosphorylated Akt	45
PAM50	Prediction analysis of microarray 50	32
PC	Principal component	21
PCA	Principal component analysis	20
PCh	Phosphocholine	12
pCR	Pathological complete response	7
PD	Progressive disease	7
PDC	Pyruvate dehydrogenase complex	9
PDX	Patient-derived xenograft	8
PET	Positron emission tomography	80

PFS	Progression-free survival	7
PgR	Progesterone receptor	4
PI3K	Phosphoinositide 3-kinase	6
PLS	Partial least squares	24
PLS-DA	Partial least squares-discriminant analysis	25
pMRD	Pathological minimal residual disease	40
pNRs	Pathological non-responders	40
ppm	Parts per million	18
PPP	Pentose phosphate pathway	8
PQN	Probabilistic quotient normalization	20
PR	Partial response	7
PtdCho	Phosphatidylcholine	12
PTEN	Phosphatase and tensin homolog	45
$\boldsymbol{Q}$	Y-loadings matrix	24
q-q	Quantile-quantile	30
R5P	Ribose-5-phosphate	9
RECIST	Response Evaluation Criteria in Solid Tumors	7
REK	Regional Ethics Committee for Medical and Health Research	38
RF	Radio frequency	16
RIN	RNA integrity	45
RNA	Ribonucleic acid	5
ROS	Reactive oxygen species	11
RPPA	Reverse-phase protein array	5
RT-PCR	Real-time polymerase chain reaction	44
RTV	Relative tumor volume	7
SAM	Significance analysis of microarray	33
SD	Stable disease	7
SDH	Succinate dehydrogenase	73
SIMPLS	Statistically inspired modification of the partial least squares method	25
$\boldsymbol{T}$	Scores matrix	21
T1	Longitudinal relaxation	16
T2	Transverse relaxation	16
T2*	Effective transverse relaxation	17
TCA	Tricarboxylic acid	9
tCho	Total choline	12
TE	Echo time	61
TMA	Tissue microarray	45
TMS	Tetramethylsilane	18
TN	True negative	28
TNBC	Triple negative breast cancer	4
TNM	Tumor size (T), degree of lymph node involvement (N), distant metastasis (M)	3
TP	True positive	28
TP1	Time point 1, at baseline	38
TP2	Time point 2, 12 weeks into treatment	38
TP3	Time point 3, at surgery, 24 weeks after treatment start	38
TP_E	Earlier time point	65
TP_L	Later time point	65
TSP	Trimethylsilyl propionic acid	29
VEGF	Vascular endothelial growth factor	6

ν_{nuc}	Resonance frequency of a nucleus	17
ν_{ref}	Resonance frequency of a reference compound	17
$\boldsymbol{W}$	Weight matrix	24
$\boldsymbol{X}$	Data matrix containing explanatory or independent variables	19
$\boldsymbol{y}$	Vector containing the response or dependent variable	20
$\boldsymbol{Y}$	Matrix containing multiple response or dependent variables	68

List of Papers

Paper I

Preprocessing of NMR metabolomics data

Leslie R. Euceda, Guro F. Giskeødegård, and Tone F. Bathen

Scandinavian Journal of Clinical and Laboratory Investigation 2015 March 4;75:3, 193-203, DOI: 10.3109/00365513.2014.1003593

Paper II

Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes

Tonje H. Haukaas, Leslie R. Euceda, Guro F. Giskeødegård, Santosh Lamichhane, Marit Krohn, Sandra Jernström, Miriam R. Aure, Ole C. Lingjærde, Ellen Schlichting, Øystein Garred, Eldri U. Due, Gordon B. Mills, Kristine K. Sahlberg, Anne-Lise Børresen-Dale, Tone F. Bathen, and The Oslo Breast Cancer Consortium (OSBREAC)

Cancer & Metabolism 2016 June 27; 4:12. DOI: 10.1186/s40170-016-0152-x

Paper III

Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Submitted to Metabolomics 2016

Paper IV

Metabolic response to everolimus in patient-derived xenografts of triple negative breast cancer

Leslie R. Euceda, Deborah K. Hill, Endre Stokke, Rana Hatem, Rania El Botty, Ivan Bièche, Elisabetta Marangoni, Tone F. Bathen, Siver A. Moestue

Submitted to the Journal of Proteome Research 2016

Table of Contents

List of Figures

List of Tables

1 Introduction

The human body is made up of cells, which perform all essential functions inside an organism. Hence, cells are commonly referred to as the basic units of life. Most types of cells in the human body are able to divide and grow. This is necessary for the proper maintenance of the body, as it allows processes such as wound repair and replacement of worn out cells. The division of a cell is governed by a series of tightly regulated events known as the cell cycle [1]. This process involves the replication of deoxyribonucleic acid (DNA), which stores the biological instructions of a cell. To ensure its proper development, when the cell cycle is impaired, a series of safety mechanisms are put in motion to either stop its progression or, if the damage is irreparable, to program cell death. Cancer cells are capable of avoiding these mechanisms, as they exhibit uncontrolled growth and proliferation, or multiplication. This is caused by the mutation, or permanent alteration, and subsequent malfunction of different genes, which are regions of DNA. Somatic mutations can occur spontaneously, e.g. due to errors during DNA replication, but can also be induced by exposure to certain physical or chemical mutagenic agents, such as ultraviolet light and chemical carcinogens [2]. Germline mutations are those occurring in sperm or egg germ cells and are thus inherited from a parent. Possessing inherited mutations is referred to as having a genetic predisposition to cancer [3].

As there are more than 100 distinct types of cancer, the term refers to a collection of diseases, typically classified according to the organ and type of cell in which they originate. Nevertheless, cancer cells exhibit common traits or capabilities known as hallmarks of cancer, first described by Hanahan and Weinberg [4]. The first four of these traits refer to uncontrolled growth and multiplication, as cancer cells can 1) stimulate their own growth, 2) bypass anti-growth signals, 3) resist programmed cell death, or apoptosis, and 4) become immortal, i.e. able to multiply endlessly. In addition, cancer cells are able to 5) induce the formation of new blood vessels, or angiogenesis, in an attempt to increase oxygen and nutrient delivery, and 6) invade adjacent tissues and henceforth spread, or metastasize, to distant sites via the blood and lymph systems. The aforementioned hallmarks were later extended [5] with two additional capabilities being added, specifically 7) the reprogramming of energy metabolism and 8) evasion of immune destruction. The acquisition of both original and more recently-described hallmarks is promoted by two enabling characteristics, namely genomic instability and mutation, and inflammation, which drive tumorigenesis.

1.1 Breast cancer

Breast cancer is by far the most common cancer in women worldwide [6]. In Norway, there were 3 324 diagnosed breast cancer cases among women in 2014 [7], making it the most frequent neoplasm in this population group. Although the estimated five year survival in breast cancer patients in Norway is almost 90% [7], predicting patient outcome is still a major challenge.

In addition to previously mentioned germline mutations, whose existence in the BRCA1 and BRCA2 genes account for most hereditary breast cancer cases [8], a number of risk factors that are non-genetic have been associated with breast cancer risk [9]. Reproductive risk factors include early menarche, older age (>30 years) at first full-term pregnancy, lower number of childbirths, and later menopause, as well as use of oral contraceptives and menopausal hormone therapy. Other risk factors include postmenopausal obesity, physical inactivity, and tobacco and alcohol consumption.

1.1.1 Anatomy of the breast

The breast contains a complex network of lobes, or groups of lobules, and ducts, as well as adipose tissue (Figure 1.1A). Milk is produced in the lobules, and from there, it is transported to the nipple via the lactiferous ducts. The epithelium of these two structures is made up of two main cellular lineages positioned in two layers. The cells in the under or basal layer adjacent to the basement membrane are called myoepithelial cells, and are highly elongated, while the surface layer is made up of luminal epithelial cells surrounding the central lumen, or tubular cavity [10] (Figure 1.1B). The majority of breast cancers are either lobular or ductal, depending on whether they originate in the lobules or lactiferous ducts, respectively. The cancer is invasive, or infiltrating, if it spreads from its origin to surrounding tissue, and non-invasive in the case of *in situ* carcinomas. Invasive ductal carcinoma is by far the most frequent invasive breast cancer (65-80%), followed by invasive lobular carcinoma (5-10%) [11].

In Norway, an approach known as triple diagnostics is recommended for breast cancer diagnosis [12]. For this, three diagnostic modalities are employed: clinical examination, imaging tests (mammogram, ultrasound, and/or magnetic resonance imaging (MRI)), and histological examination of needle biopsies or fine needle aspiration cytology (FNAC). This approach has been found to be 100% accurate when all three modalities are in agreement (i.e. all indicating either a benign or malignant lesion) [13].

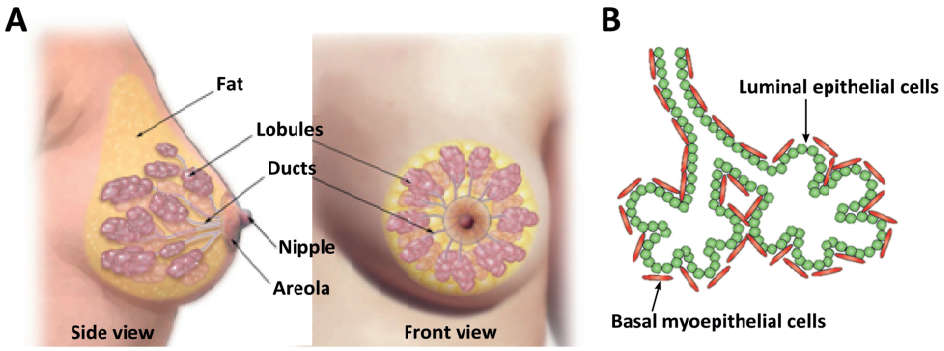


Figure 1.1. Anatomy of the female breast. A) The breast contains a network of lobes and ducts, to produce milk and transport it to the nipple. Reprinted with permission, Cleveland Clinic Center for Medical Art & Photography © 2016. All Rights Reserved. B) The lobes and ducts are made up of two main cellular lineages: luminal epithelial cells and basal myoepithelial cells, located at the surface layer and underlayer, respectively. Reprinted with permission from Macmillan Publishers Ltd: [Nature Cell Biology] (Blanpain C, *Nat Cell Biol*, 2013, 15:126–134), Copyright 2013.

1.1.2 Diagnosis and staging of breast cancer

Following diagnosis, breast cancer stage is determined. The stage describes the extent of cancer, particularly the size of a tumor and its dissemination. The TNM classification system, which is the most widely used, stages cancer based on primary tumor size (T), the degree of regional lymph node involvement (N), and presence of distant metastasis (M) [14]. Increasing stage indicates a more aggressive disease, with Stage 0 referring to non-invasive breast cancer, Stages I through III describing invasive breast cancer, i.e. spread to the skin or chest wall, with or without lymph node metastasis, and Stage IV indicating dissemination to other organs.

1.1.3 Predictive and prognostic factors

Prognostic factors are those capable of predicting clinical outcome at the time of diagnosis, regardless of the administered therapy. Predictive factors, on the other hand, provide information regarding the probability of response to a specific therapy. Axillary lymph node status is the most powerful prognostic factor for early stage breast cancer, followed by tumor size which is the most significant prognostic factor in node-negative patients [15]. The TNM staging system is essentially a grouped prognostic factor. Histological grade, which increases as the degree of tumor differentiation decreases, also provides prognostic information, with higher grades being associated with lower long-term survival [16].

Hormone (estrogen and progesterone) receptor status is the main indicator of the response to endocrine therapy [17]. The presence of estrogen receptors (ER) and/or progesterone receptors (PgR) indicates that tumor progression is driven by these and other related reproductive hormones, and hence benefit of endocrine therapy is likely. Although 70% of human breast cancers are ER positive [18], the mechanisms governing the abnormal expression of hormone receptors in breast cancer are yet to be completely elucidated. Presence of ER and PgR has been widely considered an indication of good prognosis [19, 20]. This is still under debate, however, particularly after long-term follow-up [21, 22]; thus, they are more powerful as predictive markers and less commonly used as prognostic markers.

HER2 positive tumors refer to those with amplification of the gene for human epidermal growth factor receptor 2 (HER2), also known as ERBB2, and/or overexpression of its product, the HER2 protein. HER2 promotes cellular growth and proliferation and its overexpression has been associated with poor prognosis of patients with lymph node-positive breast cancer [23-25]. Novel targeted therapies that inhibit HER2, e.g. trastuzumab, have been developed and approved for clinical use. HER2 status is thus a predictive marker for anti-HER2 treatment and in addition, has been suggested to play a role for predicting response to chemotherapy [26] and endocrine therapy [27].

Breast tumors not expressing ER, PgR, and HER2 are classified as triple negative breast cancer (TNBC). This breast cancer subtype is biologically aggressive, and is a poor prognostic and predictive factor for disease-free survival [28].

1.1.4 Locally advanced breast cancer

Locally advanced breast cancer (LABC) refers to non-metastatic tumors greater than 5 cm, tumors extending to the skin or chest wall, and/or tumors with extensive spread to lymph nodes (N status = N2 or N3) [29]. LABC can be considered the most advanced stage of the disease prior to metastasis. LABC prognosis varies greatly among patients depending mostly on tumor size and lymph node status; this makes it a clinically challenging disease requiring combination therapy strategies. These tumors are generally considered inoperable and are therefore treated with neoadjuvant or preoperative chemotherapy, to reduce them to an extent where surgical removal is possible. In addition to systemic chemotherapy and surgery, it is recommended that LABC patients receive subsequent radiotherapy to reduce the risk of relapse and death [12].

1.1.5 Breast cancer heterogeneity

Due to the heterogeneity and complexity of breast cancer tumor biology, patients with similar clinical diagnosis might have different prognosis and response to treatment. Breast cancer heterogeneity is manifested at different molecular levels. At the genetic level, germline mutations in the BRCA1 and BRCA2 genes account for most hereditary breast cancer cases (Section 1.1).

At the transcriptomic or ribonucleic acid (RNA) level, five intrinsic subtypes of breast cancer based on gene expression profiles have been identified: luminal A, luminal B, basal-like, HER2 enriched, and normal-like [30, 31]. These subtypes have been found to correlate with survival, with the luminal A subtype having the best prognosis while the basal-like subtype has the worst prognosis and is considered the most aggressive [31].

At the proteomic level, differences in the expression of estrogen and progesterone receptors and HER2 protein play an important role in current breast cancer clinical decision-making in terms of prognosis and optimal treatment plan. For TNBC patients, chemotherapy is often a mandatory inclusion in the treatment plan, since these tumors are unresponsive to endocrine and anti-HER2 therapy. TNBC is considered a clinically, histologically, and molecularly heterogeneous disease [32]. Furthermore, six protein breast cancer subtypes based on protein expression by reverse-phase protein array (RPPA) have been identified [33, 34]. These protein subtypes display considerable overlap with the gene intrinsic luminal A, luminal B, HER2, and basal-like subtypes. In addition, two novel protein-defined subtypes were identified and termed *reactive* due to many of their characteristic proteins probably being produced by the tumor microenvironment. The reactive I protein subtype appeared as a subset of the luminal A intrinsic subtype while the reactive II protein subtype was a combination of intrinsic subtypes.

On the metabolic level, the luminal A gene expression subtype has been further divided into three subclasses based on metabolic profiles [35]. One of the subclasses, luminal A2, exhibited a seemingly more malignant metabolic profile compared to the others. This suggests that metabolic variability within gene expression intrinsic subtypes could potentially be useful in further characterizing breast cancer heterogeneity.

1.1.6 Breast cancer treatment strategies

Breast cancer treatment regimens depend on tumor size, histological characterization, grading and receptor status, axillary lymph node status, and age of the patients. Surgery and

radiation therapy are local treatments and are directed towards the tumor. Systemic medical breast cancer therapies include endocrine therapy and novel targeted therapy, as well as chemotherapy. Chemotherapeutic agents such as anthracyclines and taxanes are currently the most potent drugs for use in breast cancer [36]. Taxanes disrupt microtubule function, inhibiting cell division, while anthracyclines damage DNA, causing cell death. Chemotherapy is not tumor specific, and so is more toxic than endocrine or targeted therapy. Figure 1.2 provides a simplified overview of breast cancer treatment in Norway.

Angiogenesis, a hallmark of cancer, can also be therapeutically targeted. To fulfill high nutrient and oxygen demands, tumors secrete growth factors, the most important of which is the vascular endothelial growth factor (VEGF), which stimulate the growth of new blood vessels from existing vasculature into the tumor [37]. Angiogenesis also promotes invasion and metastasis as it creates direct access into the circulatory system through which tumor cells can spread. One clinically used antiangiogenic agent is bevacizumab, a humanized monoclonal antibody, which blocks VEGF signals by binding to it, impeding its interaction with its receptors [38]. Bevacizumab was first approved by the United States Food and Drug Administration (FDA) for metastatic colorectal cancer [38]. However, FDA accelerated approval for the use of bevacizumab in metastatic breast cancer was revoked due to the drug not inhibiting tumor growth nor prolonging patient survival markedly enough to justify its risks [39], which include cardiovascular complications such as hypertension and congestive heart failure, hemorrhaging, and nasal and gastrointestinal perforations.

Treatment for TNBC is challenging, as there are currently no targeted drugs for this group of patients. Other anticancer drugs target signaling pathways known to be deregulated in cancer. Everolimus inhibits the phosphoinositide 3-kinase (PI3K)/Akt signaling pathway, a pathway which contributes to cell survival, growth and proliferation [40], by inhibiting the

		Operable Tumor	
	Inoperable tumor		
	Before Surgery	**Surgery**	**After Surgery**
Treatment	Neoadjuvant therapy	Mastectomy or breast conserving surgery	Radiation Therapy Endocrine Treatment Chemotherapy Anti-Her2 treatment
Purpose	Tumor shrinkage and downstaging	Remove primary tumor and lymph nodes	Reduce recurrence

Figure 1.2. Overview of breast cancer treatment strategies in Norway. Based on treatment guidelines from the Norwegian Breast Cancer Group (NBCG) [12].

mammalian target of rapamycin (mTOR), a regulatory protein that is a key effector downstream of Akt. As a rapamycin analog, everolimus has high affinity for the FKBP-12 protein, which directly targets the mTOR complex 1 (mTORC1). Everolimus displaces mTORC1 in binding to FKBP-12, blocking the interaction between the two [41]. This interrupts further activation of the PI3K/Akt pathway downstream of mTORC1. The addition of everolimus into the treatment regime of hormone receptor positive [42, 43] and HER2 positive [44] breast cancer patients has been found to significantly prolong progression-free survival (PFS).

1.1.7 Breast cancer treatment response criteria

The efficacy of breast cancer treatment strategies can be measured according to different response criteria. Pathological complete response (pCR) has been considered the gold standard treatment outcome and refers to complete eradication of invasive cancer cells at treatment completion. pCR has been associated with improved survival [45]. However, pCR association to long-term outcome varies between different breast cancer subtypes [46].

The most traditional approaches to measure treatment response are based on anatomical measurements of tumors. The Response Evaluation Criteria in Solid Tumors (RECIST criteria), updated to RECIST 1.1 in 2009 [47], were created in an attempt to standardize treatment assessment, and reflect change in tumor size following treatment. The four RECIST response categories are: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). Briefly, CR is defined as the disappearance of all tumor lesions, PR refers to >30% tumor shrinkage, PD indicates >20% increase in tumor size and/or the appearance of new lesions, and SD describes tumors whose size has not decreased or increased enough to qualify as PR or PD, respectively. Computerized tomography (CT) and MRI are the recommended modalities to measure tumor size at both baseline and follow-up [47].

Relative tumor volume (RTV), defined as the tumor volume after treatment divided by baseline tumor volume, is an important measure to assess cancer treatment response in preclinical animal studies. The current standard non-invasive method to measure volume of subcutaneous animal tumors *in vivo* is using external calipers [48]. However, variability of caliper volume measurements, due to e.g. observer subjectivity, differences in tumor shape, or skin and fat layer thickness can affect RTV accuracy and reproducibility.

1.1.8 Patient-derived xenograft models

Animal models, in particular mouse models, are widely used to study breast cancer and other diseases. Since breast cancer onset occurs as an effect of several genetic events, transgenic mouse models, which are animals that have been genetically modified to develop tumors, have been highly valuable to study the disease [49]. Breast cancer xenograft models have been used for research purposes for decades; these models are established by implanting *in vitro* cultivated breast cancer cells of human origin into immunodeficient animals. Cancer cells cultured *in vitro*, however, have been found to lose many traits of the tumors they were derived from, such as their complexity of gene expression patterns, tumor microenvironment interactions, and heterogeneity in general [50]. This has led to poor clinical translation of findings in mouse xenografts of cancer cell lines. The need for more clinically predictive animal models of cancer has led to the use of patient-derived xenografts (PDX), which are established by implanting primary tumor tissue directly from a patient into an immunodeficient animal. PDX models preserve more similarities to human tumors than traditional cell line xenografts, including histology [51] and gene expression profiles [52]. Therefore, PDX provide an important means to study breast tumor biology and for the assessment of breast cancer treatment strategies.

1.2 Tumor metabolism

Reprogramming of energy metabolism is necessary for tumor survival, growth, and proliferation [53]. This field has gained increasing interest as it has been found that signaling pathways controlled by three gene types of particular relevance to cancer (oncogenes, tumor suppressors, and DNA repair genes) orchestrate this metabolic switch [53-55]. Metabolic deregulation provides cancer cells with three basic needs: 1) rapid generation of adenosine triphosphate (ATP) as a source of energy, 2) increased synthesis of the four types of macromolecules: lipids, carbohydrates, proteins, and nucleic acids, and 3) proper redox stability [53].

1.2.1 Glucose metabolism

Glucose is the major source of energy and carbon for mammalian cells and has three major metabolic fates: glycogen synthesis, glycolysis, or the pentose phosphate pathway (PPP) [56]. Although glycogen has been found to be accumulated in cancerous cells [57], this pathway is not believed to be fundamental for glucose metabolism in cancer.

Glycolysis, or the breakdown of glucose, is a linear pathway whose end product is pyruvate and also yields ATP and the reduced form of nicotinamide adenine dinucleotide (NADH) [54]. Pyruvate then follows one of two pathways depending on the presence or absence of oxygen. Under aerobic conditions, pyruvate enters the mitochondria where it is converted to Acetyl-coenzyme A (Acetyl CoA) by the pyruvate dehydrogenase complex (PDC). The Acetyl CoA then enters the tricarboxylic acid (TCA) cycle which yields CO_2 as a byproduct, as well as NADH and the reduced form of flavin adenine dinucleotide ($FADH_2$). The NADH and $FADH_2$ that are produced during glycolysis and the TCA cycle are high-energy electron carriers that enter oxidative phosphorylation, also known as the electron transport chain. During oxidation of NADH and $FADH_2$, electrons are transferred from these coenzymes to oxygen and free energy is released (Equation 1.1). Glycolysis, the TCA cycle, and oxidative phosphorylation yield two, two, and thirty-two molecules of ATP, respectively (Figure 1.3), and together comprise cellular respiration.

$$NADH/FADH_2 \xrightarrow{electron\ transfer} O_2 \quad (1.1)$$
$$\downarrow$$
$$\text{ATP}$$

When little or no oxygen is available to oxidize pyruvate and NADH produced during glycolysis, pyruvate is quickly reduced to lactate via the action of the enzyme lactate dehydrogenase (LDH). The oxidized form of NAD produced through this reaction fuels glycolysis creating a positive feedback loop (Figure 1.3). Although the production of ATP via anaerobic glycolysis is 100 times faster than oxidative phosphorylation [58], it is much less efficient as it yields only two molecules of ATP from glycolysis per glucose molecule. In both cancer cells and normal proliferating cells, most of the pyruvate produced during glycolysis is converted to lactate whether oxygen is present or not; hence, their metabolism is termed *aerobic glycolysis* [54]. This metabolic switch is known as the Warburg effect, and was first observed by Otto Warburg [59, 60].

The PPP, found to be elevated in cancer [61], is a major source of the reduced form of nicotinamide adenine dinucleotide phosphate (NADPH), which is an electron carrier or antioxidant. The PPP involves an oxidative (ox-PPP) and non-oxidative (non-ox-PPP) phase. The non-ox-PPP provides a link to glycolysis as it converts ribose-5-phosphate (R5P) into the glycolytic intermediates fructose-6-phosphate and glyceraldehyde-3-phosphate to promote cellular energy metabolism. This pathway is reversible, being able to redirect the glycolysis intermediates to produce R5P to increase biosynthesis of ribonucleotides when required for proliferation (Figure 1.3).

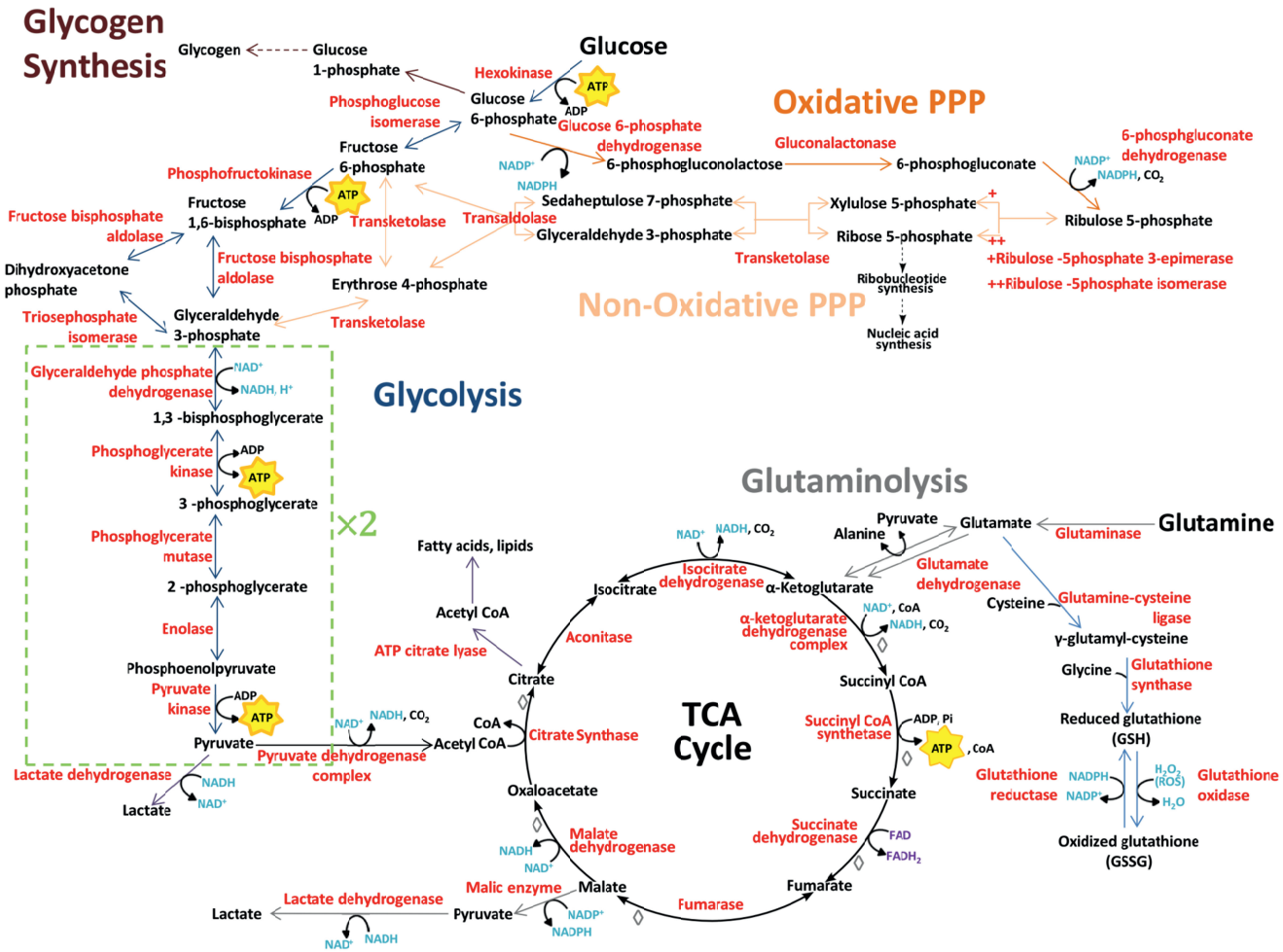


Figure 1.3. Overview of cellular glucose and glutamine metabolism. Glucose can undergo three fates: glycolysis (dark blue), glycogen synthesis (brown), or the pentose phosphate pathway (PPP), which has an oxidative (dark orange) and non-oxidative (light orange) phase. The TCA cycle (black) reactions that are recruited by glutaminolysis (gray) are marked ◊. Enzymes are shown in red. ADP: adenosine diphosphate; ATP: adenosine triphosphate; CoA: coenzyme A; FAD: oxidized flavin adenine dinucleotide; $FADH_2$: reduced flavin adenine dinucleotide; NAD^+: oxidized nicotinamide adenine dinucleotide; NADH: reduced nicotinamide adenine dinucleotide; $NADP^+$: oxidized nicotinamide adenine dinucleotide phosphate; NADPH: reduced nicotinamide adenine dinucleotide phosphate; Pi: inorganic phosphate; ROS: reactive oxygen species.

1.2.2 Amino acid metabolism

In culture, cancer cells exhibiting the Warburg effect have been found to require exogenous glutamine for survival, a phenomenon known as glutamine addiction [53, 62, 63]. This normally non-essential amino acid becomes essential as it plays an important role in different pathways that promote cancer cell growth and proliferation. Particularly, it becomes a key source of carbon and nitrogen for nucleotide, lipid, and protein biosynthesis.

Glutamine can be imported into the cell by different transporters, where it is hydrolyzed to glutamate and ammonium by the enzyme glutaminase (GLS). Increased glutamine transporter activity and GLS expression have been found in cancer [55, 62]. Glutamate provides the amino group for non-essential amino acids necessary for macromolecular synthesis such as alanine, aspartate, serine, and glycine [64]. Glycine has been associated with poor prognosis [65] and larger tumor size [66] in breast cancer. Glutamate is also the precursor of glutathione (GSH) and the TCA cycle intermediate α-ketoglutarate. GSH is a major cellular antioxidant [67], and as such provides protection from reactive oxygen species (ROS) that oxidize and damage cellular proteins, lipids, and nucleic acids and may ultimately cause cellular dysfunction or death [68].

With most glucose being diverted to the production of the waste product lactate via the Warburg effect, glutamine becomes an important anaplerotic and source of carbon by replenishing TCA cycle intermediates. Metabolism of α-ketoglutarate generates NADPH and thus energy through oxidative phosphorylation as mentioned before [64]. Oxaloacetate is also generated as a TCA cycle intermediate, and is necessary for the formation of citrate which in turn provides Acetyl CoA groups for cholesterol and fatty acid synthesis [62, 64].

Significant amounts of glutamine carbon have been found to be converted to lactate and secreted from cancer cells *in vitro*, in a similar manner as glucose [53, 62]. This reaction involves the malic enzyme and leads to synthesis of CO_2, pyruvate, and NADPH, the latter being a reducing equivalent, and as such is important for lipid and nucleotide synthesis. Through this pathway, increased glutaminolysis in proliferative cells can fulfill an important proportion of their NADPH demands [53, 62]. An overview of glutamine metabolism is shown in Figure 1.3.

Taurine is another amino acid whose relevance in cancer has been investigated. However, findings for this metabolite vary with different types of cancer [69], and the involved mechanisms are unknown. In gliomas, taurine has been suggested as a marker of apoptosis [70], while in prostate [71] and colon cancer [72, 73] higher taurine levels have been found in

tumor compared to normal tissue. In breast cancer, taurine has been negatively associated with axillary lymph node spread [66] and poor prognosis [65], and has been found to be elevated in ER positive compared to ER negative tumors [74].

1.2.3 Choline phospholipid metabolism

Abnormal choline metabolism has become widely accepted as a cancer metabolic trait over the past years. Choline-containing compounds of particular interest include free choline (Cho), phosphocholine (PCh), and glycerophosphocholine (GPC). These compounds are involved in the metabolism of phosphatidylcholine (PtdCho), the major phospholipid constituent of eukaryote cell membranes. After being transported into the cell, Cho is converted to PCh in the first step of the Kennedy pathway, which is the major source of *de novo* PtdCho. GPC is a product of PtdCho catabolism that is subsequently converted to choline, thereby completing the choline cycle (Figure 1.4). MR spectral peaks for Cho, PCh, and GPC are detected *in vivo* as a single peak termed total choline (tCho), which is elevated in different cancers when compared to normal tissue [69].

Choline-containing compounds are considered essential to sustain cell proliferation in terms of being substrates for membrane multiplication for new cells [75]. In addition, since lipid second messengers are synthesized via choline phospholipid metabolism, this pathway

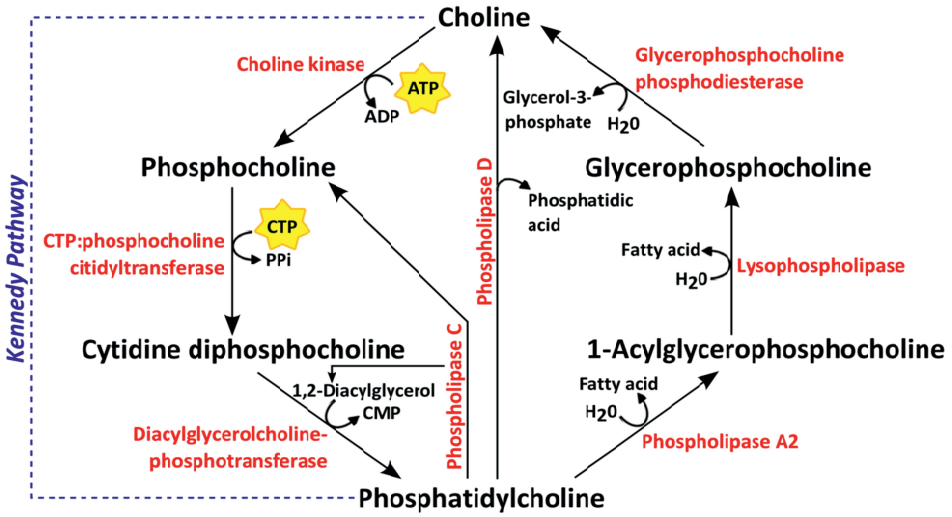


Figure 1.4. Choline phospholipid metabolism. Enzymes are shown in red. ADP: adenosine diphosphate; ATP: adenosine triphosphate; CMP: cytidine monophosphate; CTP: cytidine triphosphate; PPi: inorganic pyrophosphate.

also plays a role in lipid-based signal transduction. Increased tCho and PCh in cancer has been attributed to a higher choline uptake rate [76, 77], as well as increased expression and activity of the enzymes choline kinase alpha (CK-α) [77, 78] , phospholipases C [77, 79] and D [77, 80], and glycerophosphocholine phosphodiesterases [81, 82]. The involvement of these enzymes in PtdCho metabolism is illustrated in Figure 1.4. Nevertheless, much of the mechanisms governing altered choline phospholipid metabolism in cancer are still not fully understood. Malignant transformation, rather than rapid proliferation, has been suggested as the main driver for abnormal choline metabolism, as non-cancer proliferating cells were found to sustain low PCh, GPC, and tCho levels [83]. In the same study, the ratio between GPC and PCh was inverted from higher GPC/PCh to higher PCh/GPC with immortalization of cells, suggesting the latter ratio to be a marker for aggressiveness. In contradiction, *ex vivo* studies have found GPC/PCh to be elevated in breast cancer subtypes with worse prognosis [74, 84]. The discrepancy may be due to the association of higher PCh/GPC to aggressiveness being based on *in vitro* studies that do not take the effects of the tumor microenvironment into account. Acidic extracellular pH, for example, has been found to significantly increase GPC and decrease PCh levels [85]. In spite of the lack of consensus regarding choline metabolite ratios, increased levels of one or more of the individual metabolites constituting tCho are consistently observed in tumors and are associated with cancer aggressiveness. Decreased tCho signal has therefore been suggested as a marker of tumor response to treatment [86, 87].

1.2.4 Metabolomics

Metabolic profiling, or metabolomics, involves the large scale study of small (size < 1000 Da) cellular metabolites and their intermediates present within a living system at a given moment. The metabolome comprises the total sum of metabolites within a living system and is at the last level of the omics cascade, a flow of biochemical information that is exchanged through the interaction of different molecular levels (Figure 1.5) [88]. The remaining three

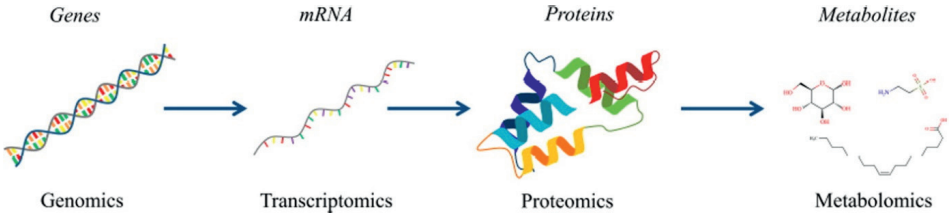


Figure 1.5. The omics cascade. Metabolomics is the final step in the cascade, and thus studies information closer to the phenotype. Reprinted from Paper I, with permission.

levels of the cascade are the genome, transcriptome, and proteome. Because metabolites are downstream of genes, RNA, and proteins, metabolite levels can be considered the ultimate response to genetic modifications or environmental changes including disease, drug, and dietary influences. Metabolomics therefore studies information closer to the observable biological endpoints, i.e. the phenotype, and has been described as "the link between genotypes and phenotypes" [89]. The two main analytical techniques within metabolomics are magnetic resonance spectroscopy (MRS) and mass spectrometry (MS), which are complementary techniques. This thesis concerns metabolic profiling of breast tumor tissue using MRS. MRS metabolomics of breast tumor tissue has revealed significantly different metabolic profiles between breast cancer patients and healthy controls [90], as well as distinct metabolic profiles linked to breast cancer prognostic factors [74, 91] and prognosis [92].

1.3 Magnetic Resonance Spectroscopy

1.3.1 Principles of Nuclear Magnetic Resonance

The theory behind nuclear magnetic resonance (NMR) comes from the concept of a quantum mechanical property called spin (*I*) that atomic nuclei can possess. Spin can be compared to the rotation of a planet on its axis (Figure 1.6A). This property depends on the number of protons and neutrons in a nucleus, with different combinations of these subatomic particles giving rise to distinct spin configurations. In nuclei with even numbers of both protons and neutrons, antiparallel spin pairs, which cancel each other out, are formed, and thus they experience no spin (*I*=0). Nuclei with odd numbers of both protons and neutrons possess integer spins (*I*=1,3,5, etc.), while nuclei whose sum of protons and neutrons is odd have half-integer spins (*I*=1/2, 3/2, 5/2, etc.). Nuclei having these two last configurations, i.e.

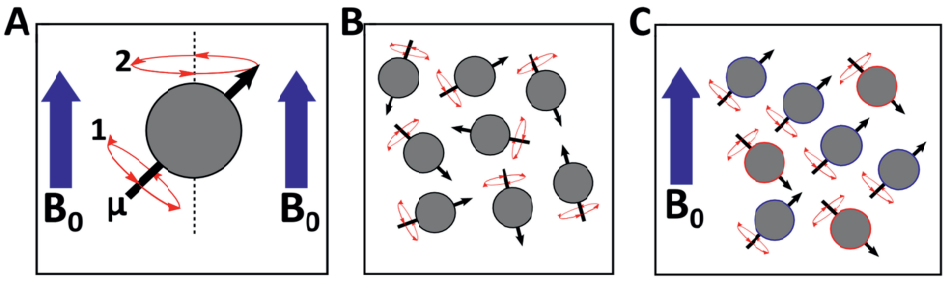


Figure 1.6. Depiction of the classical view of the magnetic resonance of a nucleus. A) The magnetic moment vector (μ) is the axis of spin (1), while the external magnetic field (B_0) is the axis of precession (2). B) Under normal circumstances, μ is oriented at random. C) If B_0 is applied, nuclear magnetic moments align either parallel (lower energy state, circled in blue) or anti-parallel (higher energy state, circled in red) to B_0.

$I \neq 0$, generate their own magnetic field, called the magnetic moment (μ), which is proportional to its spin, and thus give rise to an NMR signal. According to quantum mechanics, nuclei have $2I+1$ possible orientations, meaning that nuclei with spin of ½ have two orientations, or energy states, and are thus the simplest systems observable by NMR. The hydrogen-1 nucleus (^{1}H) is most commonly used for MRS biomedical applications due to its high abundance in the body and high sensitivity [93].

The magnetic moment μ of protons, which can be represented schematically as a vector with direction and magnitude, is usually oriented at random (Figure 1.6B). If an external magnetic field (B_0) is applied, μ will align either parallel or antiparallel to B_0 (Figure 1.6C) in a low (I=+1/2) or high (I=-1/2) energy state, respectively. The sum of magnetic moments from two protons at different energy states yields a net magnetization of zero. However, at a state of thermal equilibrium, a very slight majority of protons are found at the lower energy state (α) than at the higher energy state (β), as described by the Boltzmann probability distribution (Equation 1.2):

$$\frac{N_\beta}{N_\alpha} = e^{-\frac{\Delta E}{k_B T}} \qquad (1.2)$$

where N is the population of β or α spin states, ΔE is the energy difference between the spin states, k_B is the Boltzmann constant, and T is the temperature in Kelvin (K). The energy state population difference gives rise to a net magnetization vector (M_0) aligned with B_0 from the sum of magnetic moments of all protons (Figure 1.7A). This residual component of the magnetization causes the magnetic moment of each nuclei to move circularly, or precess, about the B_0 rotational axis at a constant rate (Figure 1.6A). This frequency of precession is the resonance frequency, called the Larmor frequency (ω_0), and is the number of complete

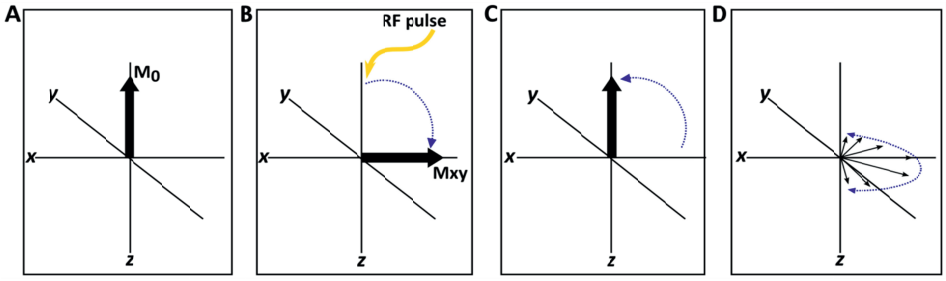


Figure 1.7. NMR excitation and relaxation. A) The sum of magnetic moments of all protons in different energy states (α or β) gives rise to a residual component of the magnetization (M_0), or net magnetization vector, parallel to the external magnetic field. B) By applying a radio frequency (RF) pulse, M_0 can be tilted from the longitudinal z-axis to the transverse x-y plane. C) Interrupting the RF pulse causes M_0 (i.e. the precessing nuclei) to gradually return from the x-y plane to the z-axis, through longitudinal (T1) relaxation. D) In addition, the nuclei spins, which become synchronized with the application of the RF pulse, gradually dephase or fan out chaotically, through transverse (T2) relaxation.

rotations of a given nucleus per second. The resonance frequency corresponds to ΔE and is proportional to the strength of B_0 and to the gyromagnetic ratio (γ), which is a fundamental constant unique for each nuclei (Equation 1.3).

$$\omega_0 = \gamma B_0 / 2\pi \quad (1.3)$$

The magnetization vector can be manipulated by applying a radio frequency (RF) pulse (B_1), at the same frequency as the Larmor frequency of the nuclei of interest, to excite said nuclei to a higher energy state. This tilts the magnetization vector away from the B_0 axis at an angle dependent on the pulse magnitude and duration. A 90° pulse rotates M_0 90 degrees, from the longitudinal z-axis to the transverse x-y plane (Figure 1.7B). In addition, the pulse brings the excess α state spins comprising the residual magnetization to a temporary state of phase coherence, meaning they remain fixed with each other instead of fanning out chaotically in a cone-shaped distribution as in Figure 1.7D. When the RF pulse is interrupted, the excited nuclei gradually return to the equilibrium state, precessing around the z-axis with characteristic longitudinal (T1) and transverse (T2) relaxation times, releasing the absorbed energies. The precession of the magnetic vector induces a RF current in a receiver coil, and this is observed and recorded as a free induction decay (FID) or time domain signal, which when converted to the frequency domain via Fourier transform [94] gives the NMR spectrum.

T1 and T2 relaxation are illustrated in Figure 1.7C and D. T1 or spin-lattice relaxation is an enthalpy process and refers to a nuclei's loss of thermal energy to its surroundings, or lattice, to return to a state of equilibrium, i.e. affecting the magnetization components that are parallel to B_0. T2 or spin-spin relaxation refers to the exponential decay of the transverse component of the net magnetization as a result of the loss of coherence of the individual precessing spins of each nuclei. In other words, spins become desynchronized, affecting the magnetization components perpendicular to B_0. Spin decoherence occurs due to interactions between neighboring spins that lead to random fluctuations of the local magnetic field of protons and ultimately causes the precession frequency of their magnetic moments to vary at random. The slower a molecule moves, the more its protons are affected by these inhomogeneities in its local magnetic field. Because the mobility of small molecules, such as water, is much higher than that of macromolecules, such as lipids, the latter exhibit shorter T2. The difference in T2 can thus be exploited to suppress lipid signals by employing T2-weighted pulse sequences which generate NMR spectra where the signal intensity of protons is directly proportional to its T2.

T2 relaxation is not only affected by motion-dependent inhomogeneities in the local magnetic field of each proton. Perturbations in the static, external magnetic field B_0, arising

e.g. from defects in the magnet, sample inhomogeneity, or from other materials placed within the field, also contribute to the loss of coherence, therefore accelerating T2. The observed or effective T2 is called T2*, and is always shorter or equal to T2 (T2≥T2*).

The resonance frequency is dependent of the external magnetic field, and nuclei of the same type, e.g. protons, that are in different chemical environments will resonate at slightly different frequencies. The resonance frequency deviation due to electron shielding is known as the chemical shift and is what allows signals from protons situated in compounds of different molecular structure, e.g. metabolites, to be distinguished separately in an NMR spectrum, a plot of the signal intensity as a function of the resonance frequency (Figure 1.8). Other factors, such as pH and temperature, also alter chemical shift. In addition, rather than appearing as sharp resonance lines, peaks are split into multiplets as a result of spin-spin interactions, or couplings, which are transmitted between neighboring nuclei through chemical bonds.

The chemical shift (δ) of a specific nucleus is calculated as the difference between its resonance frequency (ν_{nuc}) and that of a reference compound (ν_{ref}), followed by division of the reference compound resonance frequency (Equation 1.4).

$$\delta = \frac{\nu_{nuc} - \nu_{ref}}{\nu_{ref}} \qquad (1.4)$$

The reference compound will thus have a chemical shift of zero. The most common

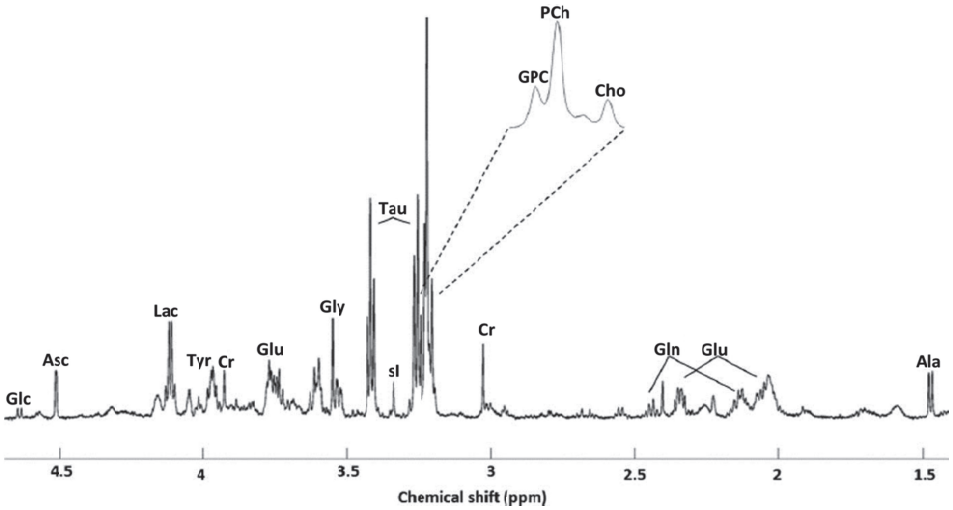


Figure 1.8. Proton high resolution (HR) magic angle spinning (MAS) magnetic resonance (MR) spectrum of breast tumor tissue. Metabolite peaks observed include glucose (Glc), ascorbate (Asc), lactate (Lac), tyrosine (Tyr), creatine (Cr), glutamate (Glu), glycine (Gly), taurine (Tau), scyllo-inositol (sI), glycerophosphocholine (GPC), phosphocholine (PCh), choline (Cho), glutamine (Gln), and alanine (Ala). Adapted from Paper I, with permission.

reference compound employed in ^{1}H NMR is tetramethylsilane (TMS). Sodium salts of 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) are used as references in inorganic solvents, where TMS is not soluble. Because the resonance frequency differences are typically around one million times smaller than the actual resonance frequency values, the chemical shift is commonly expressed in terms of parts per million (ppm), which are independent of the magnetic field strength. Thus, the chemical shift, characteristic of the chemical arrangement a nucleus is in, together with the splitting, reflecting the bonding of a nucleus, provides information about molecular structure. Furthermore, the signal intensity in an NMR spectrum is proportional to the molecular concentration of nuclei producing the signal, and thus allows for quantification of detected compounds.

1.3.2 High Resolution Magic Angle Spinning MRS

NMR of solids and semisolids, such as biological tissue, is characterized by the presence of nuclear interactions that are anisotropic, i.e. dependent on the orientation or direction of the molecules which the involved nuclei are part of with respect to the external magnetic field. Anisotropic interactions arise due to the lack of molecular mobility in these samples, essentially being magnetic and electronic interactions between spatially proximal nuclei in static positions [95]. This gives broad signals in the spectra, increasing the risk of significant overlap of metabolite peaks. The same does not occur in liquid solution state, as the rapid movement of the molecules averages out the anisotropic interactions. The orientation dependence of these interactions can be described by the angular function $3cos^{2}\beta-1$, which can be exploited to substantially reduce the line broadening effect. This angular function equals 0 when β, corresponding to the angle of a spinning axis with respect to the external magnetic field, is 54.7°. Rapid spinning of a solid sample on its axis at this angle, referred to as the magic angle, mimics a liquid solution state in which the anisotropy of the interactions is averaged to zero. This method, called high resolution (HR) magic angle spinning (MAS) MRS [96, 97], therefore produces spectra from tissue samples with resolution comparable to that of conventional liquid solution NMR (Figure 1.8).

HR MAS MRS requires minimal sample preparation and can detect over 30 metabolites in breast cancer tissue [98], many of which are involved in the previously described cancer-relevant pathways of glycolysis (Section 1.2.1), glutaminolysis (Section 1.2.2), and choline phospholipid metabolism (Section 1.2.3). The sample remains intact after spectral acquisition, which allows for subsequent analysis, such as histological, gene, and protein analysis, on the same sample.

1.4 Analysis of MRS metabolomics data

Statistical analysis approaches for MRS metabolomics data should be selected taking into consideration the purpose of the study and the structure of the data to be analyzed. Prior to statistical data analysis, however, preprocessing of the data must be performed to avoid interference from unwanted effects such as experimental and instrumental artifacts during subsequent extraction of biologically relevant information.

1.4.1 Preprocessing of MRS metabolomics data

For the purposes of multivariate statistics, one can think of metabolomics data as being organized in a two-dimensional, rectangular structure called a data matrix (***X***), where the rows represent each sample and the columns represent each variable or point in a spectrum. Alternatively, variables can correspond to individual, quantified metabolites. Raw NMR metabolic spectra are inadequate for statistical analysis as they may reflect differences not related to the biological traits of interest. For instance, lower or higher metabolite peak intensities may arise due to differences in analyzed sample amounts rather than a metabolic response to a certain stimulus. Peaks may shift left or right during NMR acquisition due to e.g. pH or temperature differences, resulting in metabolites appearing at slightly different positions at the ppm scale in raw spectra of different samples and therefore making these not comparable. To remove these and other unwanted effects, a series of computational procedures, collectively referred to as data preprocessing, are applied to the raw data, transforming it into a proper format for subsequent statistical analysis. This decreases sources of error, misleading results, and inaccurate biological interpretations.

Typical magnetic resonance (MR) spectral preprocessing procedures include correction of distorted baselines, alignment of shifted peaks, normalization to correct for variations in amount of sample, and mean centering and scaling to balance differences in abundance of metabolites. Baseline offset can be corrected by shifting the lowest spectral point to zero. Another baseline correction method is asymmetric least squares smoothing (AsLS) [99], which fits the baseline using a vector of asymmetric weights to minimize the residuals and adds a smoothing penalty to the estimated baseline. Spectral alignment can be performed using icoshift [100], for which spectra are divided into segments. Each segment is aligned individually in a way that maximizes the correlation between each spectrum and a reference, without stretching or compression of the signal. Area normalization is widely used for MRS-acquired metabolomics data of tissue, consisting of dividing each spectral point by the total

spectral area. Probabilistic quotient normalization (PQN) calculates a most probable dilution factor based on the distribution of quotients derived from dividing all points of the spectrum to be normalized by the corresponding points in a reference spectrum [101]. Mean centering is typically preformed prior to scaling, by subtracting the column mean from each individual spectral value. The purpose of this is so that the maximum variation is not the deviation from zero. Autoscaling consists of dividing each value by the column standard deviation after mean centering. This makes all metabolites or spectral points have unit variance, and thus contribute equally to subsequent statistical model building.

Paper I of this thesis is a review pertinent to preprocessing of MRS metabolomics data and provides a more detailed description of methods to perform the aforementioned procedures.

1.4.2 Multivariate analysis

MR spectra consist of thousands of points, making the datasets they constitute complex and high dimensional. Because many spectral points make up one metabolite peak, and a single metabolite can be represented by many peaks, MRS metabolomics datasets are typically highly collinear, i.e. an approximate linear relationship will hold between at least some of the variables in the $\boldsymbol{X}$-matrix. In addition, the high number of variables (p) often greatly exceeds the number of samples (m) in metabolomics datasets. Multivariate analysis methods discover structure in the data and can be divided into two types based on the purposes they are typically used for [102]. Unsupervised methods aim to describe naturally-occurring, underlying patterns, usually employing visualization techniques to identify outliers and sample clusters or groupings. Supervised methods are used to relate the metabolomics data to a response variable ($\boldsymbol{y}$) of interest for prediction purposes. Response or outcome variables are factors of interest, such as clinical and biological characteristics, that can be continuous, e.g. percentage of proliferating cancer cells, or categorical, e.g. labels indicating response to treatment or not. The main difference between unsupervised and supervised multivariate methods is that the latter makes use of the response variable for modelling while the former searches for patterns in the explanatory data only [102, 103]. Common unsupervised methods are principal component analysis (PCA) and hierarchical cluster analysis (HCA), while supervised methods include partial least squares methods.

Principal component analysis

Plotting all variables in a metabolomics dataset simultaneously would result in an abstract visual representation of p-dimensions, essentially uninterpretable. Plotting all possible paired combinations of the p variables is unfeasible and potentially uninformative. Principal component analysis simplifies the data by calculating new variables, or latent variables, as linear combinations from the original variables. PCA latent variables are called principal components (PCs), and they form a new coordinate system of reduced dimensionality. PCs are mutually orthogonal, or independent from each other, which is graphically translated as their being perpendicular to one another, completely uncorrelated. Each new PC explains less variation than the previous one, with the first PC thereby describing the direction of the maximum variation [104]. Thus, the first two or three PCs typically encompass most of the relevant data variation, and plotting them alone eases visualization and allows the identification of hidden trends within the original multidimensional data.

PCA decomposes the original ***X***-matrix into scores (***T***) and loadings (***P***) matrices as shown in Equation 1.5, where ^{T} indicates the transpose of a matrix [104]:

$$X = TP^T + E \qquad (1.5)$$

The scores (***T***) define the position of the samples in the new PC coordinate system. Similarly, the loadings (***P***) show the contribution of each original variable on the PCs (Figure 1.9A and B). The variation not explained by the PCA model, or residual variance (***E***), decreases as more PCs are included. The optimal number of PCs to incorporate can be determined using scree plots, i.e. plots of the variance explained by each PC [105]. The PC number positioned where the plotted curve bends suddenly before it starts to straighten horizontally should be selected (Figure 1.9C). The remaining PCs are likely to explain only

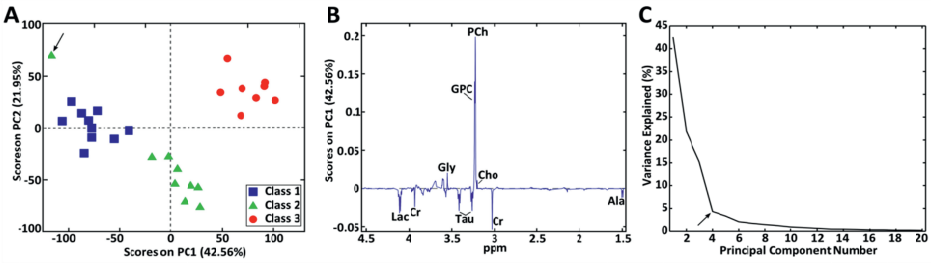


Figure 1.9. Hypothetical example of principal component (PC) analysis. A) The scores plot shows the three sample classes clearly separated, with scores on PC1 increasing with increasing class number. A potential outlier is indicated by the arrow. B) The loadings plot shows lactate (Lac), creatine (Cr), taurine (Tau), and alanine (Ala) having lower values for PC1 (i.e. higher in Class 1 samples) while glycine (Gly), glycerophosphocholine (GPC), phosphocholine (PCh), and choline (Cho), have higher values for PC1 (i.e. higher in Class 3 samples). C) Scree plot with arrow indicating the bend in the curve (optimal dimensionality) at four principal components.

unimportant variation or noise that should not be modelled.

The scores plot can reveal important patterns and sample groupings of physiological importance. Furthermore, using the established PCA model, scores can be calculated for new samples. The position of new sample scores can be used to infer unknown information about new samples, e.g. a class they belong to based on the class of the surrounding samples. Another application is the identification of potential outliers, as these are typically observed separated from the general sample population in the scores plot. PCA can also be used to select important variables, as their being positioned further away from the origin in the loadings plot indicates a higher influence on the model and vice versa. The combined use of scores and loadings plots, which can be plotted together as what is known as a biplot [105], allows the positive or negative association of samples with variables, i.e. metabolites, observed in similar or opposite positions, respectively, in the new coordinate system.

Hierarchical cluster analysis

Hierarchical cluster analysis (HCA) finds natural sample groupings based solely on the variation of the ***X***-matrix. These groupings or clusters are organized in a multilevel hierarchy as visualized in a dendogram or cluster tree (Figure 1.10). The clustering is performed according to determined similarity measures, with samples with similar variable measurements being grouped closer together. At the bottom of the dendogram, each sample is considered a cluster on its own. A similarity or distance measure is calculated for every pair of clusters. The most similar pair of clusters is merged into a single cluster, and distance measures between the new cluster and the rest are calculated to consider every possible cluster combination when determining the subsequent pair to be merged. This process is repeated so that clusters are linked progressively as one moves up the dendogram until only one cluster comprised of all samples remains. The height of the dendogram indicates the

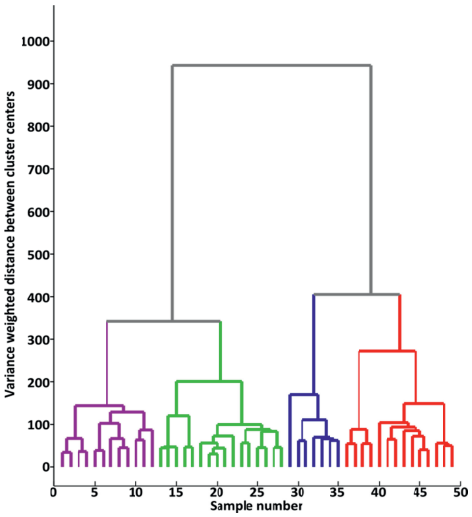


Figure 1.10. Dendogram or cluster tree example. Cutting the dendoram at height 450 partitions the data into four groups or clusters, shown in different colors.

distance between clusters. Thus, as the height where clusters are linked increases, so does their level of dissimilarity. The approach described above is a bottom up approach, termed agglomerative. Alternatively, a top down approach, termed divisive, can be adopted, where a single cluster containing all samples is repeatedly split until each cluster is represented by a single sample [106].

Similarity between samples or clusters is defined using distance measures, as samples which are closer together in the space spanning the ***X***-matrix variable measurements are more similar. Several distance measures exist. The Euclidean distance is the most commonly used distance measure for individual samples [106], and it describes the length of the straight line connecting two samples.

Other distance measures include the Mahalanobis distance [107], which takes into consideration the covariance of the population of samples, and the Manhattan distance [108], which is computed following a grid-like path restricted to vertical and horizontal movements.

HCA can be performed using several linkage methods that define the way the inter-cluster distance is measured [106, 109]. For instance, using single linkage, the distance for a given pair of clusters *A* and *B* is computed as the minimum distance found between one sample in *A* and one sample in *B*. On the contrary, with complete linkage, the distance between individual samples in *A* and *B* that are farthest away from each other is defined as the measure of similarity between both clusters. Ward's method, also called the minimum variance method, aims to minimize the within-cluster variation with each cluster merger [110].

Partial least squares methods

Supervised multivariate methods model the relationship between the metabolomics data matrix (***X***) and a response variable of interest (***y***), represented as shown in Equation 1.6:

$$y = Xb + e \qquad (1.6)$$

where ***b*** is a vector of regression coefficients, one for each variable in ***X***, and ***e*** is a vector of residuals, i.e. variation not explained by the model. To regress ***y*** on ***X*** essentially means to predict the value of ***y*** using the values in ***X***. An unknown response for a new sample ($\boldsymbol{y}_{new}$) can be predicted by multiplying the vector containing the new sample's measurements for the explanatory variables included in ***X*** with the vector of determined regression coefficients (***b***), provided that the new sample comes from a similar population as the samples used to fit the model. Classical multiple linear regression (MLR) fits the model, i.e. calculates the regression coefficients, using the least squares approach (Equation 1.7):

$$b = (X^T X)^{-1} X^T y \quad (1.7)$$

where ^{T} and $^{-1}$ indicate the transpose and the inverse of a matrix, respectively. However, MLR is unable to handle cases where the explanatory variables in $\boldsymbol{X}$ are highly collinear and/or there is a higher number of these variables than samples, features typical of metabolomics datasets. For such datasets, $\boldsymbol{X}$ is said to not be of full rank [111], with the inverse of $\boldsymbol{X}^T\boldsymbol{X}$ in Equation 1.7 being either incalculable or ill-conditioned [112] due to correlated variables in $\boldsymbol{X}$ cancelling each other out. In these cases, partial least squares (PLS) regression methods are an appropriate alternative to establish linear relationships between $\boldsymbol{X}$ and $\boldsymbol{y}$.

Like PCA, PLS regression is a dimensionality reduction technique that finds orthogonal latent variables (LVs), here called PLS LVs, as linear combinations of the original variables. It is therefore suitable that the PLS abbreviation alternatively stands for "projection to latent structures" [113]. Unlike PCs, however, PLS LVs are determined using $\boldsymbol{y}$ to guide the decomposition of $\boldsymbol{X}$ into scores and loadings, thereby capturing the most relevant information for predicting $\boldsymbol{y}$.

The PLS regression model is established by producing a weight matrix $\boldsymbol{W}$ with columns corresponding to weight vectors for each of the p columns in $\boldsymbol{X}$. The number of rows (c) in $\boldsymbol{W}$ is the number of LVs included in the model. The weights are calculated so that the covariance between $\boldsymbol{X}$ and $\boldsymbol{y}$ is maximized [114], and together with $\boldsymbol{X}$ produce the scores matrix $\boldsymbol{T}$ as shown in Equation 1.8:

$$T = XW \quad (1.8)$$

$\boldsymbol{Y}$ weights or loadings ($\boldsymbol{Q}$) are then calculated using $\boldsymbol{y}$ and $\boldsymbol{T}$ according to Equation 1.9:

$$y = TQ + e \quad (1.9)$$

From Equation 1.6 and Equation 1.9, Equation 1.10 is formulated, in which subsequent replacement of $\boldsymbol{T}$ according to Equation 1.8 leads to Equation 1.11:

$$Xb = TQ \quad (1.10)$$

$$Xb = XWQ \quad (1.11)$$

$\boldsymbol{X}$ on both sides of Equation 1.11 is cancelled out, and thus, for the fitted PLS model, the regression coefficeints ($\boldsymbol{b}$) in Equation 1.6 result from $\boldsymbol{W}$ and $\boldsymbol{Q}$ (Equation 1.12), thereby completing the prediction model:

$$b = WQ \quad (1.12)$$

The loadings ($\boldsymbol{P}$) for the PLS model are calculated from the $\boldsymbol{X}$-matrix and the scores, (Equation 1.13), where $\boldsymbol{F}$ contains the residuals for the scores:

$$X = TP + F \quad (1.13)$$

Scores and loadings can be visualized and interpreted using plots similarly as for PCA (Figure 1.9).

The two most common algorithms for performing PLS regression are the non-linear iterative partial least squares (NIPALS) [115, 116] and the statistically inspired modification of the PLS method (SIMPLS) [117]. The NIPALS algorithm was first developed for PCA but was later extended for PLS. Although both algorithms basically produce the same result, SIMPLS is gaining increasing popularity over NIPALS due to its higher speed [118, 119].

PLS regression was originally developed for the prediction of continuous response variables. However, it can be applied to classification problems, i.e. when the response variable is categorical, using the variant partial least squares-discriminant analysis (PLS-DA). For this, the categories or classes corresponding to each sample are represented using dummy variables, e.g. 1 and 2, to indicate a responder and non-responder, respectively, to a certain treatment [120]. Linear discriminant analysis (LDA) can be performed on the scores matrix (***T***) extracted by PLS [121], using e.g. Fisher's algorithm which maximizes the between-class variation while minimizing the within-class variation [122]. Often, however, PLS regression is used to relate the metabolomics ***X***-matrix and the response comprised of dummy variables in the same way as when applied to continuous responses, and the model built will assign new samples of unknown origin a numeric value used to determine the class it most closely resembles. This is done by setting an optimal classification threshold, which in the example above would be between 1 and 2, with samples that have predicted responses below that threshold being assigned the class *responder* and those with responses above it are assigned the class *non-responder*. The PLS models in the studies included in this thesis have been built for classification purposes, and thus the following descriptions will be generalized for PLS-DA.

1.4.3 Validation of multivariate models

Because supervised methods use a response variable to fit a model, a good prediction performance may be achieved by chance; phenomena unique to the data used for building or training the model may be described, with the model not being applicable to new samples within the same population. This is known as overfitting, and occurs when irrelevant LVs describing only noise are included in the model [123]. In contrast, underfitting results from important phenomena not being considered by the model, i.e. too few LVs are used, therefore excluding relevant variation when modelling. Overfitting and underfitting are the two main causes for poor performance of PLS models. Underfitting is easy to detect as when it occurs

the model performs poorly on both training and new samples. With overfitted models, however, successful prediction of the training samples will be achieved. This may lead to underlying phenomena pertinent only to the training samples being erroneously generalized to the entire population they represent. Overfitting is therefore a more serious problem, as model prediction results are overestimated, leading to false discoveries and interpretations. To avoid this, proper validation of multivariate models using techniques such as cross validation and permutation testing is crucial [124].

Cross validation

Model quality and robustness is commonly assessed using a test set approach, where the said test set is comprised of samples with available explanatory metabolomics data for which the corresponding response values or classes are known. The training samples, constituting the training set, and the test set should be processed in the same way, and should represent a similar population that the new samples intended to be predicted by the model also belong to. The test set can be a large group of unseen samples completely excluded from the training procedure. However, a sufficient amount of samples is needed for a proper division of the data for training and testing, as training sets with low sample size may lead to model underfitting. Technical and time constraints limit sample availability, often to an extent where the use of an independent test set is inappropriate.

A widely used alternative is cross validation (CV), a resampling technique that evaluates how model performance will translate to new samples. This can be carried out by dividing the dataset into a number (k) of subsets, called folds. Each subset is temporarily excluded at a time to serve as a test set, while the rest are used collectively for training the model. The process is repeated until all samples have been used for testing only once. Model diagnostic statistics, such as a misclassification error, of the k models built are combined, e.g. by calculating the mean.

The described procedure is a type of CV designated k-fold CV, according to the number of folds (k) defined. Sample allocation to each fold can be carried out in different ways [125], one of which is random sampling. Because randomly varying the composition of each test set when repeating the CV procedure will likely produce different results, the random CV error is less biased. Instead of establishing the number of k folds, a percentage of samples comprising each fold can be defined, such that leaving out 20% of samples at each repetition will result in the data being divided into five non-overlapping folds. For datasets including too few samples to allow for large test sets, leave-one-out (LOO) CV can be used. With this procedure, each

alternating test set is comprised of a single sample, i.e. *k* corresponds to the total number of samples. Implementing LOO CV will produce the same total error if repeated, as the test sets consisting of only one sample will not change with each repetition. Thus, random sampling is unnecessary. For datasets containing repeated measurements, leave-one or *n*- patients-out CV should be employed, so that all samples from the same patient are included in either the training or test set. Additionally, centering and scaling factors should be calculated from training data and applied to test set data to avoid overfitting.

CV is commonly used to determine model parameters, such as model dimensionality for PLS-DA. For this, PLS-DA models are built with each training set including a different number of LVs at a time. The models are used to predict the class of the corresponding test set samples. The number of LVs with which fewer samples are misclassified is selected as optimal. However, assessing both the number of LVs to include and the final model performance using the same CV-generated test sets is likely to result in overestimating the utility of the model. It is therefore advisable to perform a separate validation using an independent sample set after determining the optimal number of LVs [124]. Alternatively, if enough samples are available, a double CV procedure can be employed, where an inner loop nested in an outer loop allows for separate dimensionality optimization and prediction performance assessment as illustrated in Figure 1.11. The validation set samples in the outer loop are completely unseen to the inner loop LV optimization procedure, making their use for testing the resulting model less prone to overfitting. For interpreting the scores and loadings, a

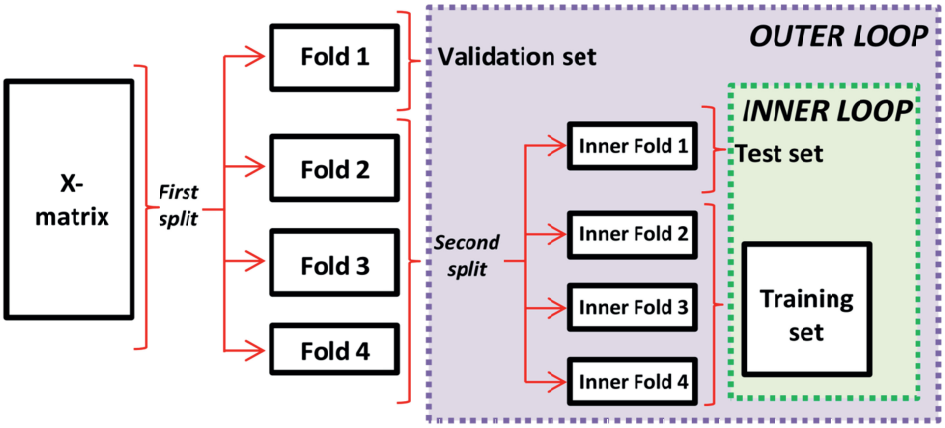


Figure 1.11. Schematic representation of a four-fold cross validation procedure. The number of latent variables (LVs) is optimized in the inner loop while model quality is assessed in the outer loop. A new fold is alternated to be left out as a validation or test set, at each repetition for the outer and inner loop, respectively, until all folds have been left out once. The risk of overfitting is reduced since the validation set samples are unseen to the LV optimization procedure.

final model can be built from the entire dataset using the most frequently selected number, i.e. the mode, of LVs in the inner loop. When performing variable selection, which is a type of model parameter optimization, validation should be performed similarly as for LV optimization, using test set data completely unseen to the variable selection procedure to validate models built with the selected variables [126].

Permutation testing

The statistical significance of a classification model can be assessed through permutation testing. Essentially, permutation tests provide information on whether classification results are better than random. They do so by rearranging, i.e. permuting, the class labels in the response variable $\boldsymbol{y}$ in random order. This destroys the true relationship between the $\boldsymbol{X}$-matrix and the $\boldsymbol{y}$ response vector, as samples will now be assigned random class labels. A model is built using the permuted $\boldsymbol{y}$-vector employing CV to obtain a misclassification error. A permuted model is expected to perform much worse than the model that has been built using the true class labels. The procedure is repeated a number of times (N), to obtain N permuted misclassification errors. A p-value can be calculated as the proportion of permuted models achieving an equal or lower error than the non-permuted model (Equation 1.14).

$$p = \frac{numberof\ permuted\ errors\ equal\ or\ lower\ than\ the\ nonpermuted\ error}{N} \quad (1.14)$$

In order to be able to achieve p-values in the thousandth order of magnitude (10^{-3}) or lower, the number of permutations (N) should be set to at least 1000. As is common practice, a p-value of 0.05 or 0.001 (outside the 95% or 99% bounds, respectively, of the permuted misclassification error distribution), should be considered statistically significant [124].

Classification model diagnostic statistics

Common model diagnostic statistics to assess the performance of a classification model include the aforementioned misclassification error, as well as the sensitivity and specificity [120]. To calculate these, the classes, hereby referred to as cases (*positives*) and controls (*negatives*), assigned by the classification model to each sample are compared to the corresponding true class. The sums for correctly classified samples of the case class and control class (true positives (TP) and true negatives (TN), respectively) and incorrectly classified samples of the case and control class (false negatives (FN) and false positives (FP), respectively) are used to summarize the predictive ability of the classification model as described below.

The misclassification error is the ratio of incorrectly classified samples by the total number of classified samples (Equation 1.15). In contrast, the accuracy can be calculated as the ratio of correctly classified samples by the total number of classified samples, or simply one minus the misclassification error.

$$Error = \frac{number\ of\ incorrectly\ classified\ samples}{total\ number\ of\ samples} = \frac{FN + FP}{TP + TN + FN + FP} \quad (1.15)$$

The ability to correctly classify the case classes is measured by the sensitivity (Equation 1.16) while the specificity indicates how well the model classifies the control cases (Equation 1.17).

$$Sensitivity = \frac{number\ of\ correctly\ classified\ case\ samples}{total\ number\ of\ case\ samples} = \frac{TP}{TP + FN} \quad (1.16)$$

$$Sensitivity = \frac{number\ of\ correctly\ classified\ control\ samples}{total\ number\ of\ control\ samples} = \frac{TN}{TN + FP} \quad (1.17)$$

A highly sensitive model is one that correctly associates the case samples with the presence of the condition being predicted, while a highly specific model correctly associates the control samples with the absence of said condition.

1.4.4 Analysis of individual metabolites

As has previously been mentioned (Section 1.3.1), the proportionality of the NMR signal intensity to the number of nuclei producing that signal makes MRS highly quantitative. Absolute quantification of spectra is possible but requires a reference signal with known concentration. Trimethylsilyl propionic acid (TSP) [127] and the electronic reference to access *in vivo* concentration (ERETIC) [128] can be used as internal references for the absolute quantification of HR MAS MR spectra. However, TSP may bind to proteins and other membrane components in tissue, which may lead to overestimation of metabolite concentrations [95]. ERETIC is an electronically-simulated reference which can be positioned at any region of the spectrum, therefore avoiding reference-analyte interactions and peak overlap [129]. Relative concentrations can be calculated by integrating peaks corresponding to each metabolite, provided that an appropriate sample normalization strategy is implemented to correct for differences in amounts of sample analyzed. The concentrations can then be used for univariate hypothesis testing to evaluate differences between groups or classes.

Univariate tests can be parametric or non-parametric, and for the proper selection between these, three assumptions should be checked: normality, homoscedasticity, and

independency of samples [130]. Normality refers to variables, in this case metabolite concentrations, following a normal, also called Gaussian, distribution. This can be evaluated visually using histograms or quantile-quantile (q-q) plots. A symmetric, bell-shaped curve or an ascending diagonal line in a histogram or q-q plot, respectively, indicates that the data is normally distributed. Normality can also be checked employing statistical tests such as Shapiro-Wilk or Kolmogorov-Smirnov [131]. Figure 1.12 illustrates how data distribution and number of groups tested should be considered for the selection of an appropriate univariate hypothesis test for group differences. Homoscedasticity refers to homogeneity of variances, and heteroscedastic data does not allow for parametric tests. However, data transformation, e.g. a logarithmic transformation, might make the data suitable for such tests, which also applies for non-normally distributed data. Independency of samples is lost in datasets where for instance, multiple metabolite measurements from the same patient are included. This is referred to as repeated measurements, and can be handled by methods such as linear mixed-effects models (LMM).

Linear mixed-effects models

Linear mixed-effects models (LMM) [132] are extensions of linear regression models, whose name derives from the inclusion of two types of effects when modelling: fixed effects, which are somewhat systematic, and are the explanatory variables, and random effects, which are a component of the unsystematic variation that arises e.g. due to individual differences among patients [133]. LMM can be employed for multilevel analysis of data grouped according to multiple classification factors, which are essentially the fixed effects. These

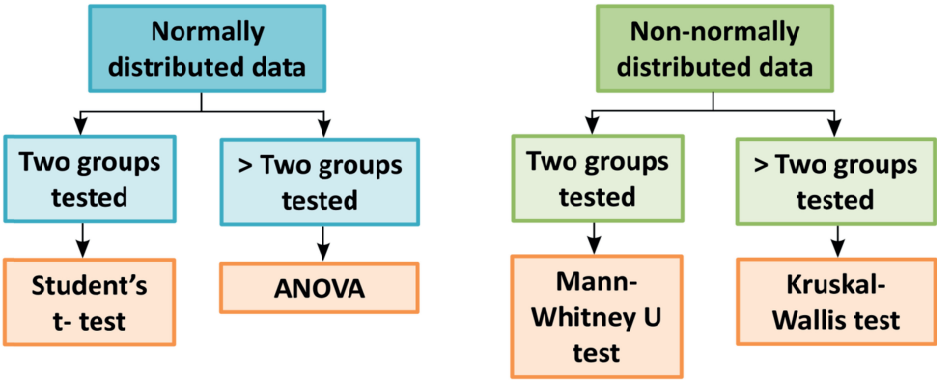


Figure 1.12. Selection of univariate hypothesis tests according to data distribution and number of classes tested. Adapted with permission from Tonje H. Haukaas.

effects, which one has modelling control over, originate from variations between the different factor levels, while the random effects originate from the between-patient variation that cannot be accounted for, such as unrecorded diet and exercise habits. Incorporating a random effect variable that contains labels for the patient each sample was collected from allows for the correction of repeated measurements when relating the fixed effects to the response variable ($\boldsymbol{y}$), e.g. a metabolite concentration. This is especially useful for longitudinal studies where sample collection is performed on the same patients over time for follow-up purposes. In addition, LMM can handle incomplete data, which refers to measurements from certain patients not being available at all time points. This is not uncommon in clinical studies, and for other methods requires the exclusion of patients lacking data at any given time point from the analysis, or the use of estimated values for the missing data.

1.4.5 Multiple testing correction

Statistical tests which associate a p-value to individual metabolites are often repeated in parallel for all metabolites under investigation. As the number of tests performed increases, so does the probability of obtaining a significant p-value by chance, ultimately leading to the accumulation of false positive findings. Multiple testing corrections aim to control the number of false positives, also known as type I errors. The Bonferroni adjustment [134] is a widely used approach which is carried out simply by multiplying each p-value by the number of tests performed. This adjustment is the standard approach for controlling the family-wise error rate (FWER), which is the probability of producing at least one false positive. However, for omics studies, especially those which are untargeted, this adjustment proves to be too strict, risking potentially important compounds to be overlooked, i.e. the generation of type II errors or false negatives.

Correcting for the false discovery rate (FDR) is an alternative approach that provides a more suitable balance between the dangers of committing type I and type II errors. This can be achieved through the Benjamini-Hochberg adjustment [135, 136], which produces a corrected p-value (q-value) for each tested metabolite. The q-values are calculated based on the metabolite p-value and on the distribution of the entire population of p-values generated with multiple parallel tests for all metabolites being considered. The q-value specifies the smallest FDR at which the tested metabolite would be considered important. For instance, a q-value of 0.05 implies that for the corresponding test to be considered significant, of all tests achieving p-values equal to or smaller as such test, i.e. those that would be considered

significant, 5% would be false positives [130]. In this example, the FDR threshold, which should be defined prior to testing, would have to be set to 5% or above for the test to be considered significant.

1.5 Gene and protein expression analysis

While profiling and statistical analysis of metabolomics data has been described, data from other omics levels is complementary and can be used to increase the validity of results. Combining omics data provides a powerful approach for the global understanding of pathway regulation. Several methods exist to extract biological meaning from omics data, with applications ranging from the determination of differentially expressed proteins and genes, to functional annotation of gene sets, some of which are described below.

1.5.1 Gene expression profiling

Gene expression levels can be measured based on the detection of total RNA. This procedure includes sample RNA extraction and purification. Gene expression microarrays are employed to measure the expression levels of thousands of genes in one sample simultaneously [137]. DNA hybridization, which occurs between two DNA strands that are complimentary to each other, is central to this procedure. Gene expression microarrays contain thousands of DNA probes, i.e. fragments of single-stranded DNA of variable length that can be fluorescently labeled. The ribonucleic acid sequence of each probe uniquely identifies a gene, and the extracted RNA will bind to each probe according to the base pairs of nucleobases (guanine binds with cytosine and adenine with thymine). Fluorescence levels at each sequence-specific microarray location are subsequently measured, and are considered directly proportional to the amount of complementary mRNA sequence present. Although gene expression microarrays do not provide absolute quantification of expression, they are useful to compare expression levels between two groups (e.g. normal vs. tumor tissue).

Prediction analysis of microarray 50

Prediction analysis of microarray 50 (PAM50) [138] is an algorithm with which a nearest centroid classifier for breast cancer gene expression intrinsic subtypes (Section 1.1.5) is built based on the expression levels of 50 genes. New samples are assigned the subtype of training samples whose shrunken mean, i.e. centroid [139], they are closest to. The distances to centroids are calculated using Spearman rank correlation [140].

Significance analysis of microarray

Significance analysis of microarray (SAM) [141] determines significant changes in gene expression and at the same time accounts for the problem of chance discoveries with the immense numbers of genes tested. Genes are assigned a score based on differences in gene expression relative to the standard deviation of repeated expression measurements [141]. A threshold for the score is defined to determine statistical significance. In addition, SAM employs permutations on expression measurements to test which potentially significant genes are identified by chance, with which the FDR is calculated.

Gene set enrichments analysis

While SAM focuses on differences in expression of individual genes, gene set enrichment analysis (GSEA) [142] tests for collective expression changes in groups of genes between two classes, biological states, or phenotypes. Prior knowledge is used to define the gene sets, e.g. based on cellular location of gene products, molecular function, associated biological processes, or gene ontology (GO) [143] category. Each gene is ranked according to their correlation with the classes. The ranked list is examined to determine whether genes are randomly distributed within the list or accumulated at the top and bottom extremes, with the latter indicating that the tested gene set is related to the distinction between classes.

DAVID

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) [144] consists of a set of tools for functional annotation of gene lists, where annotation terms describe the biological roles of specific genes. The inputted gene lists can be, for instance, a compilation of genes found to be significantly differently expressed between two groups. Functional annotation clustering within DAVID makes groups of annotation terms taking into account the extent of common genes involved, considering group members to have similar biological meaning. In addition to related annotation terms being grouped together, an enrichment score is provided for each group or cluster [145]. This eliminates redundant annotation terms within a gene list, organizing them into a clean summary easy to visualize, along with an associated measure of relevance in the enrichment score, thereby easing biological interpretation.

1.5.2 Reverse-phase protein array

Reverse-phase protein array is a high-throughput method for the quantitative measurement of protein expression levels using antibodies [34]. The procedure consists of fractionating cells to produce a cell lysate rich in released proteins. The lysate is then spotted on a slide that can contain hundreds to thousands of different immobilized samples. The slide is then incubated with a single primary antibody that binds specifically to a protein of interest. Secondary antibodies that are labeled, e.g. with a fluorescent label, bind to the primary antibody and the intensity of the fluorescence signal provides a measure of the protein expression in each of the samples in the array. Protein expression data determined by RPPA can be used to classify samples into the previously mentioned RPPA breast cancer subtypes (Section 1.1.5) defined by Hennessy *et al.* [34].

1.5.3 Integrated pathway analysis

Gene expression and metabolomics data can be combined through integrated pathway analysis using the online data analysis tool suite MetaboAnalyst 3.0 [146]. Metabolites and genes whose concentration and expression, respectively, exhibit significant changes are linked together based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database [147, 148] to define potential molecular pathways regulating biological processes as enriched. Topology analysis [149] is subsequently performed to determine the relative importance of metabolites and genes based on their relative locations in the pathway. Maps of the significantly enriched pathways can be generated, highlighting the involved genes and metabolites identified as important.

2 Thesis Objectives

2.1 Overall aim

The main aim of this thesis was to explore the potential of HR MAS MRS metabolomics of breast tumor tissue combined with advanced statistical techniques for complementing transcriptomics and proteomics in the characterization of tumor heterogeneity, optimization of treatment stratification, and for the evaluation of current available breast cancer therapies.

2.2 Specific Objectives

- To review the available methods for the main preprocessing steps of MRS metabolomics data.
- To establish and characterize metabolic subgroups of breast cancer patients and further examine the molecular heterogeneity of the disease by combining metabolite, gene, and protein expression data from human breast tumors.
- To explore the effect of neoadjuvant therapy and the antiangiogenic drug bevacizumab on metabolic profiles and to examine the potential of MR metabolomics in discriminating responders from non-responders.
- To investigate the tumor response to everolimus treatment in a heterogeneous panel of triple negative breast cancer xenografts.

3 Materials and methods

This thesis includes four papers pertaining to MRS metabolomics. Paper I is a review article that summarizes the preprocessing steps for MRS metabolomics and compares the most common methods to achieve these. Papers II-IV present MRS metabolomics data from human and PDX breast tumor tissue. Materials and methods used in Papers II-IV are summarized in Table 3.1.

Table 3.1. Materials and methods used in Papers II-IV.

		Paper II	Paper III	Paper IV
Materials	**Human Samples**	228 breast tumor tissue samples from 228 patients	270 breast tumor tissue samples from 122 patients	----
	Xenograft samples	----	----	103 breast tumor tissue samples (13 PDX models)
Data acquisition	**Metabolic profiling**	HR MAS MRS	HR MAS MRS	HR MAS MRS
	Gene expression	Microarray profiling	Microarray profiling	----
	Protein expression	RPPA	---	IHC
	Other	---	---	Mutation screening
Data analysis	**Analysis of individual metabolites**	-Relative quantification -Kruskal-Wallis test -Benjamini-Hochberg multiple testing correction	-Relative quantification and imputation -LMM -Benjamini-Hochberg multiple testing correction	-Relative quantification -LMM -Benjamini-Hochberg multiple testing correction
	Multivariate modelling	-HCA -PLS-DA	-PCA -PLS-DA	-PCA -PLS-DA
	Gene and protein expression	-PAM50 intrinsic subtyping -RPPA subtyping -SAM analysis -DAVID -GSEA	-PAM50 intrinsic subtyping	----
	Combining omics	Integrated pathway analysis	----	----

DAVID: Database for Annotation, Visualization, and Integrated Discovery; GSEA: gene set enrichment analysis; HCA: hierarchical cluster analysis; HR MAS MRS: high resolution magic angle spinning magnetic resonance spectroscopy; IHC: immunohistochemistry; LMM: linear mixed-effects model; PAM50: prediction analysis of microarray 50; PCA: principal component analysis; PDX: patient-derived xenografts; PLS-DA: partial least squares-discriminant analysis; RPPA: reverse-phase protein array; SAM: significance analysis of microarray.

3.1 Patients and xenograft models

3.1.1 Patient cohorts

Papers II and III included patients enrolled in large clinical breast cancer studies in Norway.

In Paper II, surgical biopsies from human untreated, primary breast tumors included in the Oslo 2 study were examined. Patients underwent surgery at the Oslo University Hospital (Radium Hospital and Ullevål Hospital) between 2006 and 2009. A piece of the surgically removed tumor was collected for the Oslo 2 study. Samples were divided into three sections for analysis; histological analysis, including ER status and tumor cell percentage, was performed on the edges of the two outer sections. An adequate amount of the middle section was used for metabolic profiling using HR MAS MRS. The remaining material from all three sections were pooled for DNA, RNA, and protein analysis. Of the 228 samples analyzed by MRS, gene and protein expression was successfully measured for 201 and 217, respectively, with a total of 191 samples being analyzed using the three omics modalities (Figure 3.1).

Paper III included tumor tissue samples (n=270) from 122 patients enrolled in a phase 2 randomized clinical trial: the NeoAva study. This study recruited women with non-metastatic, HER2 negative breast tumors with a diameter larger than 2.5 cm, between 2008 and 2012. The purpose of the study was to evaluate neoadjuvant antiangiogenesis treatment with bevacizumab. Recruitment was carried out at St. Olavs University Hospital in Trondheim in addition to Oslo University Hospital (Radium Hospital and Ullevål Hospital). The patients were treated as described in Section 3.1.3 and sampled at three different time points: prior to treatment (TP1), 12 weeks into treatment (TP2), and at surgery (TP3), i.e. 24 weeks after treatment start. Ultrasound-guided needle biopsies were obtained at TP1 and TP2, while a piece from the surgically removed tumor was sampled at TP3.

Samples from both Papers II and III were snap-frozen after collection and subsequently stored at -80 °C. Both the Oslo 2 and NeoAva studies were approved by the Regional Ethics Committee for Medical and Health Research (REK) and the Norwegian Medical Agency, and carried out in accordance with the Declaration of Helsinki, International Conference on Harmonization Good Clinical practice. Informed written consent was obtained

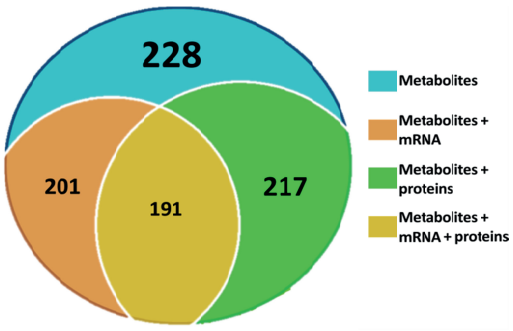


Figure 3.1. Representation of Paper II sample proportions analyzed by three different omics modalities. Of the 228 samples included for which metabolic profiles were acquired using magnetic resonance spectroscopy, 201 and 217 were measured for gene and protein expression, respectively. Data from the three omics levels was measured for a total of 191 samples. The correspondence of the shaded areas with the sample number has been estimated for illustrative purposes and should not be considered exact.

from all patients.

3.1.2 Patient-derived xenograft models

Paper IV included thirteen TNBC PDX that were previously established as described in [150]. In short, breast tumor fragments from primary surgical specimens obtained from patients with informed written consent were transplanted subcutaneously into the interscapular fat pad of female Swiss nude mice. After xenograft emergence 2-8 months after transplantation, serial grafting was performed from mouse to mouse. Animals were handled at the Translational Research Department, Institut Curie in Paris, France. The animals were divided to either be treated as described in Section 3.1.3 or to be untreated controls, so that approximately four animals per PDX model were included in each group. This resulted in a total of 50 and 53 animals in the treated and untreated groups, respectively, giving tissue samples from a total of 103 animals. The harvested tumor tissue was stored at -80 °C after snap-freezing. The tumor material was divided for mutation screening, immunohistochemistry analysis, western blotting, and metabolic profiling. All procedures and experiments involving animals were carried out according to the institutional guidelines of the French Ethical Committee and the European Convention for the Protection of Vertebrates used for Scientific Purposes.

3.1.3 Treatment protocols

All patients in Paper III were administered neoadjuvant chemotherapy according to current Norwegian treatment guidelines. This included four cycles of anthracyclines in the form of FEC100 (5-fluorouracil 600 mg/m2, epirubicin 100 mg/m2, cyclophosphamide 600 mg/m2) every three weeks and subsequent therapy with taxanes for 12 weeks (weekly infusion with paclitaxel 80 mg/m2 or four cycles of docetaxel 100 mg/m2 every three weeks). Patients were randomized to receive bevacizumab in a dose of 15 mg/kg every three weeks. Most patients were started on docetaxel as the taxane therapy, with those treated with bevacizumab continuing to receive the same dose of the antiangiogenic agent as with the previous 12 weeks of anthracycline-based therapy. The taxane therapy was changed to paclitaxel for a majority of patients due to toxicity, with those treated with bevacizumab subsequently receiving it at a dose of 10 mg/kg every other week. The treatment protocol for Paper III is summarized in Figure 3.2.

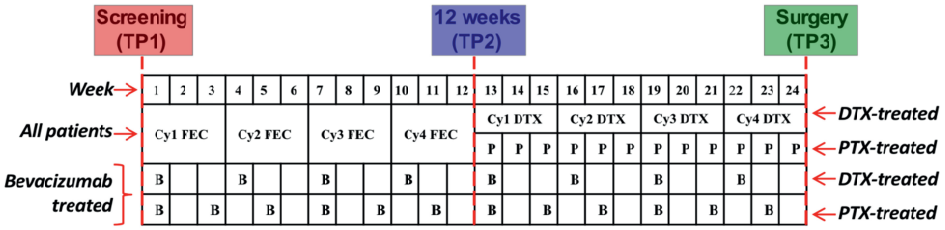


Figure 3.2. Treatment protocol for Paper III. All patients received four three-weekly cycles (Cy) of FEC followed by 12 weeks of taxane-based treatment. For those patients receiving bevacizumab (B), this was administered once every three weeks and every other week for docetaxel (DTX) receivers and paclitaxel (PTX) receivers, respectively. P:Paclitaxel administration.

Animals treated with everolimus in Paper IV received it orally for 4-5 weeks at a dose of 2.5 mg/kg three times a week.

3.1.4 Response criteria

An overview of response criteria used in the present thesis is given in Table 3.2.

Three treatment response criteria were used in Paper III. First, pCR, defined as complete eradication of all invasive cancer cells in both breast and axillary lymph nodes, was evaluated. However, since only 20/122 patients achieved pCR, a second response criterion that considered patients experiencing a good response but not achieving pCR was employed. For this, a 1 cm cut-off value for the tumor diameter measured pathologically at the time of surgery (TP3) was used to classify patients as having pathological minimal residual disease (pMRD, tumor diameter < 1 cm) or as pathological non-responders (pNRs, tumor diameter ≥ 1 cm). Finally, the change in tumor size with treatment was evaluated by calculating a response ratio as tumor diameter at TP3/tumor diameter at TP1 (mmTP3/mmTP1), i.e. inversely proportional to a good response. This ratio was used to classify patients as having a good, intermediate, or no response (GR, IR, NR) as shown in Table 3.2.

In Paper IV, everolimus treatment response was evaluated based on RTV. Significant differences between RTVs of treated and untreated tumors from the same PDX model were calculated at the end of the treatment period using Student's t-tests. PDX models achieving $p \leq 0.05$ were classified as responders, while those with $p > 0.05$ were classified as non-responders.

Table 3.2. Overview of treatment response criteria employed in Papers III and IV.

	Metric	Definition	Groups	Criteria
Paper III	Pathological complete response (pCR)	Complete eradication of all invasive cancer cells in both breast and axillary lymph nodes.	pCR+	Patients achieving pCR
			pCR-	Patients not achieving pCR
	Pathological minimal residual disease (pMRD) response	1 cm cut-off value for the tumor diameter measured pathologically at the time of surgery.	pMRD	Diameter < 1 cm
			Pathological non-responders (pNRs)	Diameter ≥ 1 cm
	Response ratio (RR)	$\frac{\text{Tumor diameter at TP3}}{\text{Tumor diameter at TP1}}$	Good response (GR)	RR≤0.10
			Intermediate response (IR)	0.10<RR<0.90
			No response (NR)	RR ≥0.90
Paper IV	Relative tumor volume (RTV)	$\frac{\text{Tumor volume at time of treatment end}}{\text{Tumor volume at baseline}}$	Responders	Significantly different RTVs (p≤0.05) between treated and control animals
			Non-responders	No significant different RTVS (p>0.05) between treated and control animals

TP: time point.

3.2 HR MAS MRS protocol

Metabolic profiles for Papers II-IV were acquired from breast tumor tissue using HR MAS MRS as described below.

3.2.1 Sample preparation

Samples were prepared on a cooling work station filled with liquid nitrogen so that they were kept cold to minimize tissue degradation (Figure 3.3). The preparation time did not exceed five minutes, and consisted in cutting the samples to fit into 30 µL, leak-proof, disposable inserts containing 3.0 µL cold solution of sodium formate in D20 at a concentration of 24.29 mM. The insert was then transferred into a MAS zirconium rotor with 4 mm outer diameter and kept at -20 °C until

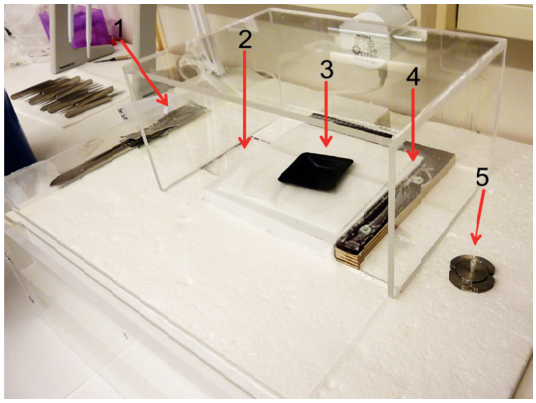


Figure 3.3. Cooling work station for tissue sample preparation. The work station, designed to minimize sample degradation, is filled with liquid nitrogen to keep samples cold during preparation. 1: disposable scalpel; 2: opening for filling of liquid nitrogen covered in plexi glass to avoid evaporation; 3: disposable weighing boat on which samples are prepared; 4: metal cooling block; 5: insert filled with sodium formate solution. Photo: Deborah K. Hill.

spectral acquisition.

3.2.2 Spectral acquisition

HR MAS MR spectra were acquired on a Bruker Avance DRX600 spectrometer (Bruker Biospin GmbH, Germany) equipped with a $^{1}H/^{13}C$ MAS probe. A spin-echo one dimensional experiment with presaturation (cpmgpr1d, Bruker BioSpin, Germany) for water and lipid suppression was recorded for all samples using the acquisition parameters summarized in Table 3.3. Two additional experiments, namely noesy and two-dimensional J-resolved pulse sequences, were acquired, but were not used for the purposes of this thesis (total acquisition time for all experiments < 50 minutes).

Table 3.3. HR MAS MR spectral acquistion parameters for Papers II-IV.

Temperature	Spin rate	Effective Echo time	Number of scans	Relaxation delay	Delay (τ)	Spectral width
5 °C	5 kHz	77 ms	256	4s	0.3 ms	−5 to 15 ppm

3.3 Spectral preprocessing and data analysis

The acquired FIDs were Fourier transformed into 64K real points after modification by an exponential line broadening factor of 0.30 Hz. Automatic phase correction for each spectrum was implemented using TopSpin 3.1 (Bruker BioSpin GmbH). Subsequent preprocessing was performed in Matlab R2013b (The Mathworks, Inc., USA). Chemical shifts were referenced to the creatine peak at 3.03 ppm. The spectral region between 1.40 – 4.70 ppm, containing the majority of low-molecular weight metabolites, was selected as the region of interest. The spectra were baseline corrected using AsLS with parameters λ = 1e7 and p = 0.0001. In Papers II and IV, the baseline offset for each spectrum was corrected prior to AsLS by setting the lowest point to zero. In Paper III, the baseline was set to vary around zero by subtracting the mean value of the noise regions at 4.52-4.49, 3.10-3.06, 2.63-2.58, and 1.84-1.75 ppm from each spectrum after AsLS. The spectra were peak aligned using icoshift and normalized using either area normalization (Papers II and IV) or PQN (Paper III). An overview of the spectral preprocessing performed in this thesis is shown in Table 3.4.

Table 3.4. Spectral preprocessing for Papers II-IV.

	Paper II	Paper III	Paper IV
Chemical shift reference	Creatine 3.03 ppm	Creatine 3.03 ppm	Creatine 3.03 ppm
Baseline correction	Lowest point shifted to 0 and AsLS	AsLS and baseline set to vary around 0	Lowest point shifted to 0 and AsLS
Peak alignment	icoshift	icoshift	icoshift
Normalization	Area normalization	PQN	Area normalization

AsLS: Asymmetric least squares smoothing; PQN: probabilistic quotient normalization.

3.3.1 Analysis of individual metabolites

Metabolite peaks were assigned based on previous identification in HR MAS MR spectra of human breast tumor tissue [98]. Relative quantification was performed by integrating fixed spectral regions corresponding to the metabolite of interest in Matlab R2013b. In Paper III, the integral of the metabolite lactate could not be calculated for 116 samples due to an overlapping lipid peak at 4.13 ppm. The imputation procedure of multivariate imputation by chained equations (MICE) [151] was thus employed to calculate lactate integrals for these samples in R 3.1.1 (R Foundation for Statistical Computing) [152]. In Paper IV, the dataset of individual metabolite integrals included the ratios of lactate/glucose, taurine/creatine, and glycerophosphocholine/phosphocholine. Metabolite levels between the metabolic clusters in Paper II were compared using the Kruskal-Wallis test.

LMM were built in R 3.1.1 (R Foundation for Statistical Computing) [152] for multilevel analysis of data grouped according to different classification factors in Papers III and IV. For this, the method of restricted maximum likelihood was employed using the function *lme* from the 'nlme' package [153]. Furthermore, LMM was applied in Paper IV to correct for repeated PDX model measurements while determining individual metabolite differences based on treatment group only.

The resulting p-values for tests repeated for multiple metabolites were adjusted for FDR by the Benjamini Hochberg method. Adjusted p (q-value) ≤ 0.05 was considered statistically significant.

3.3.2 Multivariate analysis

Multivariate analysis on spectra (Papers II and III) and integrals (Paper IV) was performed in Matlab R2013b (The Mathworks, Inc., USA). Spectra or integrals were mean-centered or autoscaled, respectively, prior to multivariate model building. HCA (Paper II) was performed on spectra of untreated primary tumors using Matlab's Statistical toolbox to identify metabolic clusters. For this, Euclidean distance and Ward's method were set as the distance measure and linkage criterion, respectively. PCA and PLS-DA were employed using PLS Toolbox 7.8.2 (Eigenvector Research Inc., U.S.A) to explore naturally occurring differences and to build classification models, respectively. The number of PCs for PCA was selected based on visual inspection of residual explained variance plots.

PLS-DA models were validated using double CV, where the number of PLS LVs was optimized in the inner loop and model performance was assessed in the outer loop. At each

loop, samples were split into a training and test set by randomly selecting a percentage of samples (20% for Papers II and IV, and 10% for Paper III) to be left out for testing. Since Paper IV included multiple samples of the same PDX type, spectra from samples derived from the same patient were all included in either the training or test set. The maximum number of LVs tested was 10. The final classification assessment parameters of accuracy, sensitivity, and specificity resulted from the mean of 20 repetitions of the double CV procedure, while the optimal number of LVs was the value most often selected as optimal during the repetitions, i.e. the mode. Additionally, validation with permutation testing was performed, building PLS-DA models using the final optimal number of LVs from double CV and randomly shuffled class labels ($\boldsymbol{y}$) for each sample. The permutation procedure was repeated 1000 times, resulting in a permuted accuracy distribution. Models were considered significant if the final accuracy obtained from non-permuted double CV was higher than 99% (Paper II) or 95% (Papers III and IV) of the permuted accuracy values ($p \leq 0.01$ or $p \leq 0.05$, respectively).

3.4 Gene and protein analysis

3.4.1 Mutation screening

Mutation status for the tumor suppressor p53 was determined in Paper IV from complementary DNA (cDNA) amplified by real-time polymerase chain reaction (RT-PCR) as described in [154]. The BigDye Terminator kit was used for cDNA sequencing on an ABI PRISM 3130 automated DNA sequencer (Applied Biosystems, Courtaboeuf, France), with 10% sensitivity for mutated cell detection.

3.4.2 Gene expression measurement and subtyping

In Papers II and III, total RNA extraction from human breast tumor tissue was performed using TRIzol® reagent (Invitrogen, Carlsbad, CA, USA). Gene expression profiles were acquired using one color Sureprint G3 Human GE 8x60k Microarrays (Agilent Technologies) according to the manufacturer's protocol (One-Color Microarray-Based Gene expression Analysis, Low Input Quick Amp Labeling, v.6.5, May 2010). Preprocessing of the microarray data included log2 transformation, quantile normalization, and adjustment for hospital differences. In Paper II, values for probes corresponding to the same gene were averaged resulting in a single expression measurement for each gene. In Paper III, gene expression values were additionally adjusted for batch effect from array design and correlations to RNA

integrity (RIN) value and background signal. Microarray data in Paper III was also mean centered on both rows and columns. The PAM50 algorithm (Section 1.5.1) was employed in both Papers II and III to assign samples the intrinsic subtype labels of luminal A, luminal B, HER2-enriched, basal-like, or normal-like.

3.4.3 Protein expression measurement and subtyping

Protein expression measurements in Paper II were acquired from tissue lysates with RPPA (Section 1.5.2) using 151 primary antibodies. Subsequent primary antibody detection was carried out using a DakoCytomation-catalyzed system (Dako, Glostrup, Denmark) for signal amplification with a biotin secondary antibody. Spotted slides were scanned to determine signal intensities using MicroVigene software (VigeneTech Inc., Carlise, MA, USA). Quantification was achieved through curve fitting, employing a standard curve relating signal intensity to protein concentration (R package SuperCurve /version 1.01). Protein concentrations were log2 transformed and median centered. RPPA subtypes were assigned by consensus clustering [155], with which the number of groups that best fit the data was identified to be five: luminal, HER2, basal, reactive I, and reactive II as defined by the Cancer Genome Atlas Network dataset [33].

3.4.4 Immunohistochemical staining

Tumor material from the PDX models in Paper IV was fixed in 10% neutral buffered formalin and paraffin embedded. Two tissue cores from the paraffin blocks of two or three tumors per PDX model type were transferred into tissue microarray (TMA) [156] blocks. TMA were analyzed by immunohistochemistry for the expression of phosphatase and tensin homolog (PTEN) and phosphorylated Akt (pAkt) as described in [154]. Immunostaining was performed using a Discovery XT Platform (Ventana Medical Systems, Inc., Tucson, Arizona, USA, part of Roche Diagnostics). Monoclonal rabbit antibodies were used as primary antibodies, and horseradish peroxidase (HRP) conjugated with 3,3′-diaminobenzidine (DAB) tetrahydrochloride as substrate was used as secondary antibody for color development.

3.4.5 Gene expression analysis

In Paper II, differential gene expression between metabolic clusters was determined in R 2.15.2 using SAM analysis, considering differences to be statistically significant for adjusted $p < 0.01$. For this, 21 851 genes from 42 405 mRNA probes were inputted and the number of permutations was set to 100 for validation. Functional annotation of gene sets determined

based on SAM results was performed using DAVID, selecting the official gene symbol as gene identifier. Furthermore, GSEA was employed for the identification of enriched gene sets in the metabolic clusters, setting the number of permutations on phenotypes to 1000 and an FDR cutoff of 25%.

3.4.6 Integrated Pathway Analysis

Transcriptomics and metabolomics data were combined in Paper II using the *Integrated pathway analysis* tool in the MetaboAnalyst 3.0 online data analysis suite. Genes achieving a SAM analysis p-value < 0.05 and significantly differently expressed metabolites between clusters were selected to be inputted in the analysis. Pathways achieving a $p \leq 0.05$ were considered to be significantly enriched.

4 Summary of Papers

4.1 Paper I

Preprocessing of NMR metabolomics data

Leslie R. Euceda, Guro F. Giskeødegård, and Tone F. Bathen

Scandinavian Journal of Clinical and Laboratory Investigation 2015 March 4;75:3, 193-203, DOI: 10.3109/00365513.2014.1003593

Metabolomics involves metabolite profiling carried out on a large scale to study cellular processes in a biological sample. It is the youngest discipline within the omics sciences, emerging only when technological advancements enabled the measurement of many biochemical compounds within a sample simultaneously. The two main analytical techniques used to acquire metabolomics data are mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy.

Independently of the analytical technique used, metabolomics studies produce vast amounts of information which result in complex datasets with a large number of variables. Therefore, multivariate statistical techniques are typically used for the proper biological interpretation of this type of data. Prior to multivariate analysis, irrelevant variation such as that which arises due to baseline distortion, peak shifting, or differences in amount of sample analyzed or abundance of metabolites within the sample, must be removed. This is achieved through a series of computational procedures collectively referred to as data preprocessing.

This paper is a review article which aims to outline the most common steps in the preprocessing of NMR metabolomics data. In addition, some of the methods to perform these steps are described and compared. Since data analysis results may vary when using different preprocessing methods, an appropriate pipeline for the selection of the optimal combination of methods in the preprocessing workflow should be established.

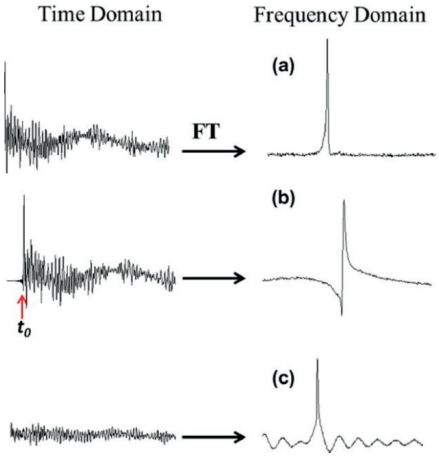


Figure 4.1. Paper I: Baseline correction in the time domain based on proper data acquisition timing. (A) Absence of dead time leads to no baseline distortion. (B) Phase errors are observed with both dead time and signal detection starting at time zero. (C) Baseline distortion observed when signal detection starts after a dead time of *t0*. FT: Fourier transform.

4.2 Paper II

Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes

Tonje H. Haukaas, Leslie R. Euceda, Guro F. Giskeødegård, Santosh Lamichhane, Marit Krohn, Sandra Jernström, Miriam R. Aure, Ole C. Lingjærde, Ellen Schlichting, Øystein Garred, Eldri U. Due, Gordon B. Mills, Kristine K. Sahlberg, Anne-Lise Børresen-Dale, Tone F. Bathen, and The Oslo Breast Cancer Consortium (OSBREAC)

Cancer & Metabolism 2016 June 27; 4:12. DOI: 10.1186/s40170-016-0152-x

This paper aimed to combine metabolite, protein, and gene expression data to study the molecular heterogeneity of breast cancer. Because patients with similar clinical diagnosis, histological characteristics, and disease stage may experience different outcomes, with regards to e.g. response to treatment and survival, identifying mechanisms contributing to breast cancer heterogeneity at a molecular level may reveal new targets for therapy or clinically relevant subgroups useful to stratify patients for treatment management.

Primary tumor samples from untreated breast cancer patients (n=228 tumor tissue samples) were analyzed using high-resolution magic-angle spinning magnetic resonance spectroscopy (HR MAS MRS). Hierarchical cluster analysis was performed on the acquired metabolic profiles, resulting in three significantly different metabolic clusters (Mc1, Mc2, and Mc3) being derived. Gene and protein expression data from the same tumors were used to further characterize the clusters.

The three metabolic clusters displayed distinct metabolic characteristics. Among these, Mc1 displayed the highest levels of glycerophosphocholine and phosphocholine, Mc2 was characterized by significantly higher levels of glucose, and Mc3 exhibited higher lactate and alanine levels. Combining metabolite and gene expression data using integrated pathway analysis revealed differences in glycolysis/gluconeogenesis and glycerophospholipid metabolism between the clusters.

Although significant differences in the distribution of reverse-phase protein array (RPPA) breast cancer subtypes were determined for all three clusters, the gene expression intrinsic subtypes were evenly distributed among them. The metabolic clusters may thus contribute additional information for the understanding of breast cancer heterogeneity. As for the differences in protein subtypes, these were predominantly due to only 9% of the samples classified as RPPA-reactive belonging to Mc1, while 44% of Mc2 tumors were of the reactive subtypes. The reactive RPPA-subtypes have a protein expression pattern believed to be

produced by the tumor microenvironment. In accordance with this, genes related to collagens and extracellular matrix were downregulated in Mc1 compared to Mc2 and Mc3.

In conclusion, metabolic profiles combined with gene and protein expression data provide new information about the heterogeneity of breast tumors. Furthermore, the distinct metabolic traits identified for each cluster infer that they may respond to different metabolically targeted drugs.

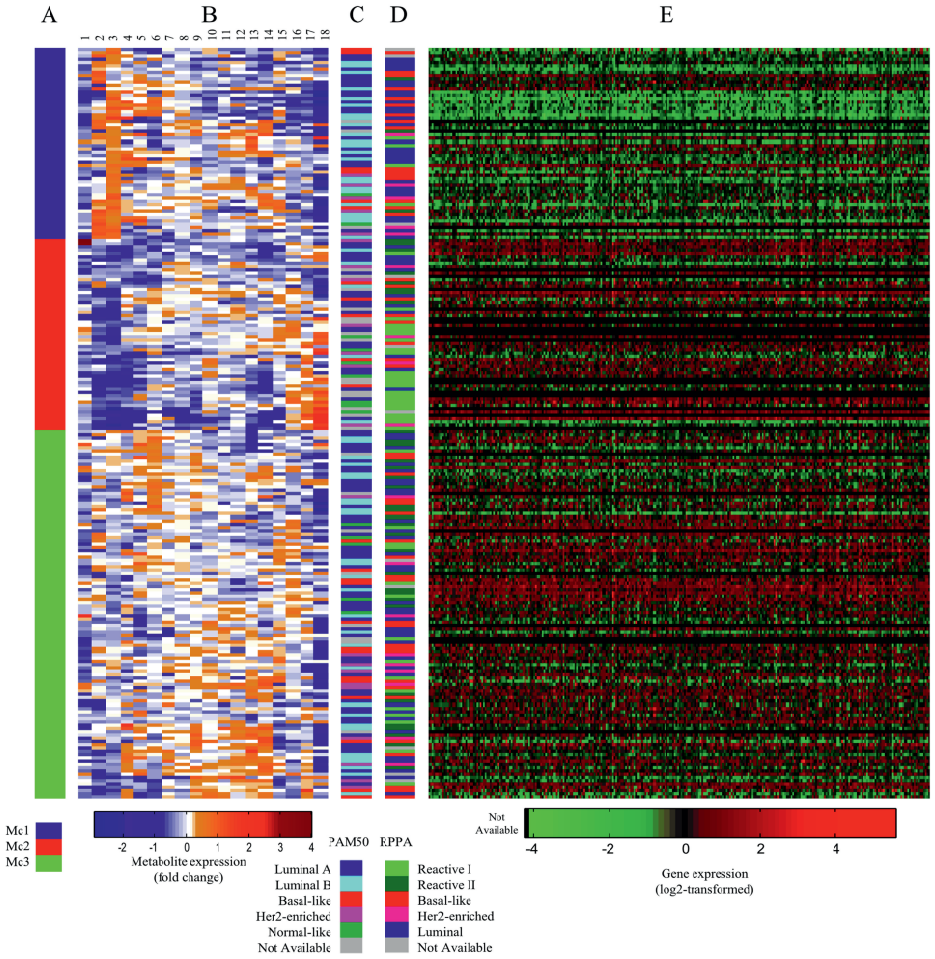


Figure 4.2. Paper II: Main differences between metabolic clusters. (A) Metabolic cluster label from hierarchical clustering of HR MAS MR spectra of samples. The samples clustered in three groups called Mc1, Mc2 and Mc3. (B) Fold change in expression levels of (1) scyllo-inositol, (2) glycerophosphocholine, (3) phosphocholilne, (4) creatine, (5) ascorbate, (6) taurine, (7) glutathione, (8) tyrosine, (9) lactate, (10) glutamate, (11) succinate, (12) glutamine, (13) glycine, (14) alanine, (15) choline, (16) myo-inositol, (17) acetate, (18) glucose. Blue regions in the heat map represent decreased levels while red levels represent increased metabolite levels. (C) PAM50-subtype (D) RPPA-subtype (E) Gene expression levels (quantile normalized, log 2 transformed) for the 277 overlapping significant genes (SAM, adjusted $p < 0.01$) between Mc1 and Mc3. The genes have been clustered.

4.3 Paper III

Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Submitted to Metabolomics 2016

This paper aimed to use magnetic resonance (MR) spectroscopy to explore the metabolic changes in breast tumor tissue occurring as an effect of overall neoadjuvant therapy, to investigate metabolic differences between therapy responders and non-responders, and to determine the metabolic effects of antiangiogenic treatment with bevacizumab.

The metabolic profiles of 122 tumors (n=270 tissue samples) from breast cancer patients were acquired using high resolution (HR) magic angle spinning (MAS) MR spectroscopy. All patients received neoadjuvant chemotherapy, while they were randomized to receive bevacizumab or not. Biopsies were sampled at three different time points: at baseline, 12 weeks into treatment, and after treatment (at surgery). Principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA) were used to analyze the metabolic profiles. The levels of 16 metabolites were calculated by integrating fixed regions of preprocessed spectra, and analyzed by linear mixed-effects models (LMM). For this, the response variable was the metabolite level while the fixed effects were *time*, *pathological minimal residual disease* (pMRD) *response*, and *bevacizumab*, and the random effect was the patient ID.

PCA showed clear metabolic changes with increasing time of treatment, indicating a decline in glucose consumption and a transition to normal breast adipose tissue as an effect of chemotherapy. PLS-DA revealed metabolic differences between pMRD patients and pathological non-responders (pNRs) after treatment, but not before or during treatment, with an accuracy of 77% ($p<0.001$). Furthermore, patients exhibiting a good response (≥90% tumor reduction) could be discriminated from those with no response (≤10% tumor reduction) based on metabolic profiles before and after treatment, with a classification accuracy of 76% ($p=0.001$) and 75% ($p=0.002$), respectively. Lower glucose and higher lactate was observed in the good response group before treatment, while the opposite was observed after treatment. Bevacizumab-receiving and chemotherapy-only patients could not be discriminated at any time point. LMM detected significant differences during and/or after treatment for 11 out of

16 metabolites, while eight metabolites differed between pMRD and pNRs. LMM additionally revealed a significant interaction between *time* and *bevacizumab* for glutathione, indicating higher levels of this antioxidant in chemotherapy-only patients than in bevacizumab receivers after treatment.

In conclusion, MR-based metabolic profiles reflected changes as an effect of chemotherapy and showed potential in detecting metabolic response to treatment and complementing other molecular assays for the elucidation of underlying mechanisms affecting pathological response.

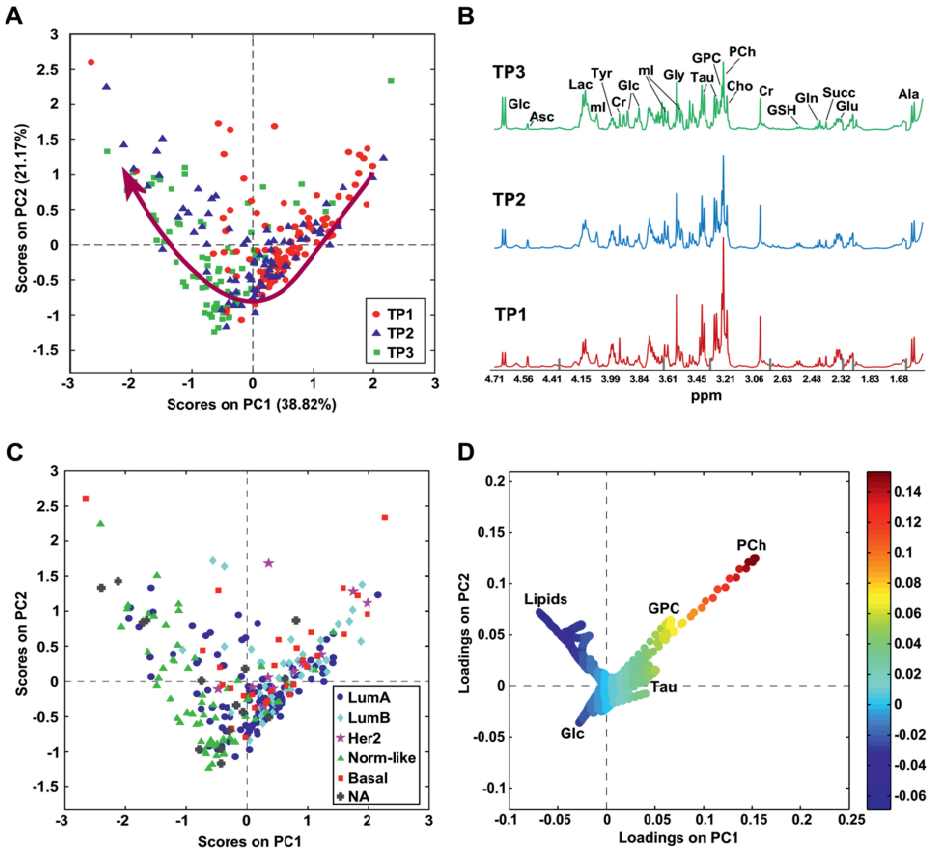


Figure 4.3. Paper III: Principal component analysis including all samples and mean spectra for each time point. (A) The scores plot shows a trend in the direction of the arrow with increasing time point. (B) PQN-normalized mean spectra at each time point. Gray bars indicate removed spectral regions. (C) The normal-like intrinsic subtype is most clearly separated from the rest in the scores plot, showing a similar distribution as TP3 in A. (D) The loadings plot indicates higher phosphocholine, glycerophosphocholine, and taurine at TP1 and increasing glucose and lipids with increasing time of treatment and in normal-like samples. Loadings are colored according to LV1. Ala:alanine; Asc:ascorbate; Cho:choline; Cr:creatine; Glc:glucose; Gln:glutamine; Glu:glutamate; Gly:glycine; GPC:glycerophosphocholine; GSH:glutathione; LumA:luminal A; LumB:luminal B; NA:not available; Norm-like:normal-like; Lac:lactate; mI:myo-inositol; PCh:phosphoscholine; Succ:succinate; Tau:taurine; Tyr:tyrosine.

4.4 Paper IV

Metabolic response to everolimus in patient-derived xenografts of triple negative breast cancer

Leslie R. Euceda, Deborah K. Hill, Endre Stokke, Rana Hatem, Rania El Botty, Ivan Bièche, Elisabetta Marangoni, Tone F. Bathen, Siver A. Moestue

Submitted to the Journal of Proteome Research 2016

Since triple negative breast cancer (TNBC) lacks expression of estrogen receptors (ER), progesterone receptors (PgR), and HER2 protein, these patients are unresponsive to both endocrine and anti-HER2 therapy, limiting the therapeutic option to chemotherapy. Furthermore, the heterogeneous nature of TNBC contributes greatly to the variability in treatment response. TNBC displays the most overlap with the basal-like gene expression subtype, which is associated with the highest activation of the PI3K/AKT/mTOR signaling pathway. Thus, the mTOR inhibitor everolimus is a potential candidate for TNBC targeted therapy.

In this study, we assessed the metabolic responses to treatment with everolimus in a heterogeneous panel of TNBC patient-derived xenografts (PDX) using high resolution (HR) magic angle spinning (MAS) magnetic resonance spectroscopy (MRS), aiming to identify metabolic biomarkers for response/resistance. Breast tumor tissue from 103 animals representing 13 different TNBC PDX models that were divided to be treated with everolimus or not were analyzed by HR MAS MRS. The relative concentrations of 17 metabolites in addition to three metabolite ratios were combined to form a single dataset for multivariate analysis. Multilevel linear mixed-effects models (LMM) were built using individual metabolite relative concentrations as the response variable, with *treatment group* (treated or untreated) and *response group* (responding or non-responding PDX) as fixed effects and the PDX model as the random effect.

Partial least squares-discriminant analysis (PLS-DA) discriminated treated animals and untreated controls with an accuracy of 67% ($p=0.003$). LMM revealed similar treatment group trends observed by PLS-DA, strongly indicating reduced glycolytic lactate production and glutaminolysis after treatment, consistent with PI3K/AKT signaling pathway inhibition. Although inherent metabolic heterogeneity between different PDX models seemed to hinder prediction of treatment response, the metabolic effects following everolimus treatment were more pronounced in responding xenografts compared to non-responders. Furthermore, mutation status for the tumor suppressor p53 could be predicted using MR-based metabolite

levels, which may provide complimentary insight into the interplay between PI3K signaling and other drivers of disease progression.

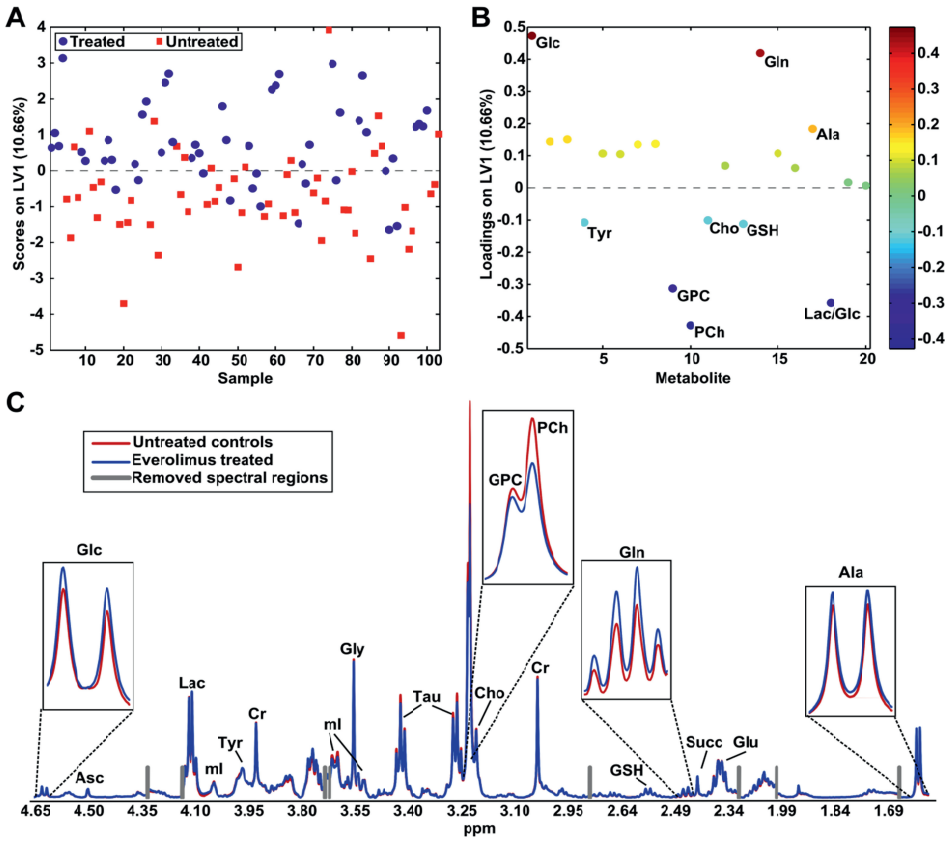


Figure 4.4. Paper IV: PLS-DA scores (A) and loadings plots (B) of treated animals vs untreated controls (n=103), and mean spectra (C) of treated animals and untreated controls. Loadings are colored according to latent variable (LV) 1. Ala: alanine; Asc: ascorbate; Cho: choline; Cr: creatine; Glc: glucose; Gln: glutamine; Glu: glutamate; Gly: glycine; GPC: glycerophosphocholine; GSH: glutathione; Lac: lactate; mI: myo-inositol; PCh: phosphocholine; Succ: succinate; Tau: taurine; Tyr: tyrosine.

5 Discussion

This thesis aimed to use *ex vivo* HR MAS MRS metabolomics to gain insight into the heterogeneity of breast cancer and to identify possibilities for optimizing treatment of the disease. Experimental data were acquired from breast tumor tissue samples of both human subjects (Papers II and III), and preclinical mouse models implanted with human cancers or patient-derived tumor xenografts (Paper IV).

Preprocessing has been found to be a critical step when analyzing the complex datasets generated from metabolomics studies. Paper I resulted from an extensive literature search on preprocessing of MRS metabolomics data, performed in preparation of the subsequent studies carried out in Papers II-IV. In Paper II, three significantly different metabolic clusters were derived from MR spectral data from a large cohort of untreated human breast tumors, and were characterized at the proteomic and transcriptomic levels. The metabolic clusters did not exhibit significant differences in intrinsic subtype distribution, indicating that the metabolic level may contribute additional biological information to further elucidate the heterogeneity of breast cancer. Papers III and IV investigated the effect of different available systemic breast cancer therapies. Paper III was carried out as part of the NeoAva phase 2 clinical trial, for which HER2 negative patients were administered neoadjuvant chemotherapy and were randomized to receive antiangiogenic bevacizumab in parallel, or not. The study design was longitudinal, including three sampling time points over the treatment course. In Paper IV, a heterogeneous panel of patient-derived triple negative breast tumor xenografts were divided to either receive treatment with the mTOR inhibitor everolimus, or not. In both Papers III and IV, significant metabolic changes as an effect of treatment were detected. Regarding treatment response, patients achieving pathological minimal residual disease (pMRD) were significantly discriminated from pathological non-responders (pNRs) after neoadjuvant chemotherapy using metabolic profiles in Paper III, independent of bevacizumab administration, while in Paper IV, the metabolic effects following everolimus treatment were more pronounced in responding xenografts compared to non-responders.

5.1 Methodological considerations

5.1.1 Patients and animal models

Tumor material from Papers II and III was obtained from large cohorts ($n > 122$ patients) of breast cancer patients. This represents a strength of this thesis, as small sample sizes may have a negative impact on statistical power, i.e. in the instance that statistical significance

cannot be achieved despite of there being a true effect, or that a statistically significant finding does not reflect a true effect [157]. Recruiting clinical cohorts, such as these, is challenging and usually takes several years (≥3 years in Papers II and III), and typically even more to collect relevant clinical follow-up data. Although pilot studies are hindered by imprecision from small sample sizes, they are often useful indicators of feasibility for larger, full-scale studies, as they give insight into study design and how a particular biological question should be addressed [158]. Pilot studies specifically designed for the studies in Papers II and III were not performed; however, similar studies of a smaller scale have previously been carried out by our research group. HCA on HR MAS MRS metabolic profiles (n=43 patients) was employed by Borgan *et al.* [35] to determine three subgroups of the luminal A intrinsic subtype, with one of them exhibiting significant metabolic differences and enriched genes related to cell cycle and DNA repair. The potential of HR MAS MRS metabolic profiles (n=33 and 89 patients, respectively) to predict survival and monitor response to neoadjuvant chemotherapy was evaluated by Cao *et al.* [159, 160]. In these studies, metabolic profiles reflected 5-year survival, while a subtle metabolic association to treatment response was detected in [159]. This gave the motivation for further investigation and result validation in larger cohorts, which was made possible through the inclusion of metabolomics for the global molecular characterization of breast tumors in the Oslo2 and NeoAva clinical studies.

Animal models are living experimental systems used to study human diseases. The advantages of employing animal models in cancer research include avoiding the risk of harming human subjects, and facilitating subject monitoring and control of experimental conditions [161]. Animal models of cancer are a more faithful representation of human tumor systems than cell lines grown *in vitro*, as *in vitro* models lack the presence of tumor microenvironment, i.e. stroma composed of the extracellular matrix, fibroblasts, blood vessels, and multiple cell types. The absence of such structures and their mutual interactions potentially hamper the translation of *in vitro* findings to the clinic [162]. It is important, however, to be aware that most of the components of the tumor microenvironment in all preclinical animal models are of the host and not of human origin, and as such, findings should be interpreted with caution.

Tumor material from PDX models was examined in Paper IV; patient-derived breast cancer cells were transplanted subcutaneously into the intrascapular backs of immunosuppressed mice. These xenograft models have been found to be more clinically predictive than those of *in vitro*-cultured cell lines [51, 52]. However, the general structure of the tumor tissue, which includes the tumor microenvironment, and perfusion may not be

entirely representative of the clinical site of interest, i.e. the breast. While orthotopic transplantation of cells into the mammary fat pad would provide a more similar microenvironment for proliferating cancer cells, a more complicated surgical procedure is required, representing a higher risk of infection or death for the animals. In addition, subcutaneous tumors are easily visualized and measured, which reduces the need to perform *in vivo* imaging to monitor cancer growth. This was the case for the animals in Paper IV, where the RTV was determined by external caliper measurement.

A general limitation of xenograft models is that the emergence of the tumor does not occur spontaneously, as it would clinically. The tumor development process is therefore not accurately portrayed in these preclinical models. To circumvent this, transgenic mouse models can be used. Another advantage of transgenic models is that, unlike xenograft models, the immune system of the host is intact. Therefore, the tumor microenvironment and structural composition will be influenced by the presence of immune cells and their interactions [163]. However, transgenic models also have limitations, with one of the most important being the large variability in the tumor development rates [163]. In these models, *in vivo* imaging is advantageous to determine appropriate recruitment criteria for mice into studies, which is often time consuming and expensive. Without such criteria, results can be misleading, for example, RTVs may differ due to differences in tumor onset time, rather than as a result of treatment. In addition, as opposed to human cells in PDX models, tumor cells in transgenic models are animal cells, which prevents direct studies of human disease. Finally, an important consideration made for Paper IV was that the study should take into account the heterogeneity of TNBC, which was achieved by examining PDX models derived from 13 TNBC patients who demonstrated variability in genetic and protein characteristics.

In the metabolomics workflow, sample collection, storage, and transport are the initial steps, and it is of the utmost importance that these steps are carried out according to strict protocol; inappropriate sample handling can lead to modification of the metabolic profile such that it no longer represents the physiological state of the tissue when it was harvested. Samples should be kept at low temperatures to avoid further enzymatic reactions occurring *ex vivo* that would lead to confounding metabolic alterations. It has been found that HR MAS MRS metabolic profiles of breast tumor samples are not significantly altered for tissues snap-frozen within 30 minutes post-surgery [164]. In accordance with this, all tissue samples analyzed as part of this thesis were kept at room temperature for no more than 30 minutes after collection, before being snap-frozen in liquid nitrogen. In addition, samples were transported and stored at -80°C (Section 3.1), and were kept cold during preparation and MRS

acquisition as described in Sections 3.2.1 and 3.2.2, respectively, to minimize tissue degradation.

5.1.2 Criteria for treatment evaluation

Strategies for determining metabolic changes as an effect of treatment in Papers III and IV differed slightly due to their respective study designs. In Paper III, the use of needle biopsies for sampling at baseline (TP1) and during treatment (TP2), in addition to sampling on the surgically removed tumor after treatment (TP3), facilitated a longitudinal study design. This provides the possibility to compare unpaired pre and post treatment samples. In addition, paired analysis as well as repeated measures analysis, e.g. LMM, can be performed, with which the between-patient variability can be disregarded when investigating the treatment effect, which is a component of the within-patient variability. In Paper IV, the animals were divided to receive treatment or not. Samples were collected from both groups at the same time point following completion of treatment in the treated group. The treatment effect was therefore evaluated by comparing the treated and untreated groups. In this case, since only one sample per animal was measured, paired analyses are not possible, and the effect of interest is a component of the between-animal variability, i.e. variability between treated and untreated animals. The use of mice as experimental subjects in Paper IV make longitudinal, *ex vivo* measurements challenging to acquire. Although longitudinal preclinical studies using non-invasive techniques such as *in vivo* MRS or MRI are common practice, the same is not the case for the *ex vivo* setting. Repeated biopsy collection in mouse models is challenging owing to the much smaller size of mice compared to humans and the higher risk of infection following tissue removal. Furthermore, the smaller mouse tumors could be more severely affected by the loss of a portion of a couple of millimeters in diameter, which could even alter further tumor development, e.g. as a result of stimulating wound repair processes, such as angiogenesis. Still, subcutaneous tumors, such as the ones grafted in the mice in Paper IV, are the optimal candidates for longitudinal needle biopsy sampling, as they are easier to locate and would require a less invasive procedure. However, a detailed protocol should be well established and ethical approval granted before applying this in preclinical studies.

An important aspect when evaluating cancer therapeutics is to define the clinical endpoint of response. In Paper III, three different response criteria were considered to assess response to neoadjuvant chemotherapy with or without bevacizumab: pCR, pMRD response, and response ratio. pCR is widely used as a primary endpoint in clinical studies and has been considered a standard outcome surrogate for survival [165]. A small number of patients in

Paper III achieved pCR (20/122 patients), which is in accordance with the reported rates of pCR after neoadjuvant chemotherapy [166, 167]. This gave the motivation to introduce a less strict metric for pathological response, i.e. pMRD response, as having minimal residual tumor burden after treatment (tumor diameter < 1 cm measured pathologically), so that metabolic information from patients not achievieng pCR but experiencing a good response, would not be lost. A similar metric was previously employed in [168] to evaluate dynamic contrast-enhanced MRI for predicting LABC patients' response to neoadjuvant chemotherapy. In said study, however, negative lymph node status was an additional criterion for pMRD. The pMRD criterion may be considered misleading since it does not take the initial tumor diameter into account. Tumors whose size increased with treatment progression might be classified as pMRD if the diameter at TP3 was below the cutoff value. However, all patients included in Paper III had tumors measuring at least 2.5 cm in diameter; thus, all pMRD cases experienced a decrease in tumor size of at least 1.5 cm with treatment. Nevertheless, the response ratio was included as a third response criterion to investigate the effect of tumor reduction on metabolic profiles, rather than just considering size at TP3. One must take into account, however, that the tumor sizes used for the calculation of this ratio were measured using imaging techniques at TP1 and pathologically at TP3. Ideally, the same method for measuring size should be used at both time points to ensure comparability. Consequently, pMRD response was used as the response fixed effect for multilevel analysis with LMM.

In Paper IV, the response criteria were based on changes in RTV, which is typically employed in preclinical studies to assess antitumor activity of a therapy under investigation [169-171]. Comparing raw tumor volumes for treated animals and untreated controls would be biased, as tumor volumes at baseline are typically heterogeneous.

5.1.3 Metabolic profiling using HR MAS MRS

The most widely used analytical techniques in metabolomics research are MRS and MS. A brief overview of the most important advantages and limitations when employing these techniques for metabolomics studies was provided in Paper I. Despite MS being a more sensitive technique, advantages of MRS of relevance to this thesis include the simple sample preparation procedure and the possibility for analyzing intact tissue. Papers II-IV included a high number of samples (n = 601 tumor tissue samples), and simpler sampling preparation is less-time consuming, enabling higher-throughput analysis. Furthermore, complex procedures involving a higher degree of sample manipulation introduce more potential sources of variability, which can reduce reproducibility and comparability between samples. Unlike HR

MAS MRS for intact tissue, MS analysis requires sample pretreatment e.g. derivatization and extraction, which can affect the metabolic profile. The global metabolic profile has been found to sustain significant alterations when exposed to elevated temperatures commonly used in gas chromatography (GC)/MS for processes such as derivatization, vaporization, and ionization [172]. The metabolites detected in intact tissue can therefore be considered less likely to be a product of changes other than the innate biological processes of the tumor.

A possible source of degradation when performing HR MAS MRS is prolonged spectral acquisition; during this time, the sample is spinning rapidly, which can cause the release of bound metabolites, and is kept at temperatures not considered low enough to completely interrupt enzymatic activity [173]. This has been investigated in human breast tumor tissue samples by our research group [164], with glucose, glycine, choline, and GPC exhibiting significant differences after being kept spinning inside the magnet for 1.5 hours. However, the acquisition protocols used in this thesis involved a shorter time (< 50 minutes, Section 3.2.2). Therefore, this consideration is more applicable for studies requiring a longer acquisition time, such as quantitative two dimensional HR MAS MRS studies.

One of the disadvantages of MR spectroscopy compared to MS is its low sensitivity, and thus the relatively few detectable metabolites. Nevertheless, metabolites involved in the pathways of glycolysis, glutaminolysis, and choline phospholipid metabolism, can be measured using this technique. More metabolites can be detected using MRS when analyzing tissue extracts in the same way as biofluids [98]. However, the degree of chemical degradation of the tissue that this entails, which gives rise to issues similar to the effect of thermal degradation, is unknown. One of the benefits of HR MAS MRS that was not exploited in this thesis was to use the same sample for histological, gene, or protein analyses after metabolic profiling. Rather, the clinical studies of which Papers II and III were a part of were designed so that the tumor material was divided for analysis with the different molecular platforms. However, post HR MAS MRS tissue has been allocated for large scale proteomic analysis (Paper II) or single cell whole genome sequencing (Paper III), of which the resulting data can be used in future follow-up studies.

Metabolic profiles for Papers II-IV were acquired using a Carr-Purcell-Meiboom-Gill (cpmg) pulse sequence. Breast cancer cells are surrounded by many different cell types, the most abundant of which are adipocytes, i.e. fat-storing cells [174]. It is therefore typical to observe broad lipid signals in breast tumor tissue spectra, which can mask low-molecular weight metabolites. The difference in mobility between macromolecules, such as lipids, and small molecules leads to the former experiencing a quicker T2 relaxation (Section 1.3.1). The

cpmg pulse sequence acts as a T2 filter; as this is a T2-weighted pulse sequence, the signal intensity of the faster-relaxing lipids is suppressed, while signals from the smaller metabolites are enhanced. However, quantification using cpmg spectra can be misleading unless proper action is taken, as metabolite peak intensities can differ slightly due to variations in T2 relaxation times rather than just reflecting differences in concentrations [175]. In order to ensure accurate absolute quantification, the T2 relaxation effect should be corrected for. This process involves calculating T2 for each metabolite to be quantified, by repeating MRS experiments with different echo times (TE), and relating each resulting metabolite peak area with the corresponding TE to fit a monoexponential decay function representing T2 [176]. This can be time-consuming, especially for untargeted studies like those included in this thesis, in which only relative quantification was employed, as discussed in Section 5.1.4.

5.1.4 Data analysis

Paper I concludes that in order to optimize MRS metabolomics data analysis, the most appropriate combination of preprocessing procedures and methods, which depends on the types of samples analyzed, should be selected. The previously mentioned lipid signals typically observed in spectra of breast tumor tissue make their preprocessing challenging, often requiring modifications to the initially proposed strategy. Of the three experimental papers included in this thesis, preprocessing for Paper III differed slightly from Papers II and IV (Table 3.4). A considerable number of spectra in Paper III exhibited lipid signals of substantially higher intensity compared to Papers II and IV. This was attributed to tumor reduction as an effect of chemotherapy, leading to much less cancer cells in TP2 and TP3 tissue samples compared to samples from untreated tumors.

Preprocessing includes the removal of unwanted peaks, such as lipids or contaminants, e.g. ethanol. Normalization after the removal of lipid peaks not only corrects for differences in sample size, but also tumor cell content, as it is assumed that most of the lipid signals originate from adjacent normal adipose cells. However, the noise level of spectra with high lipid signals became more pronounced after normalization, due to the originally high lipid-to-metabolite ratio.

To address the elevated noise level after normalization in Paper III, three strategies were implemented. First, instead of shifting the lowest point of the baseline to zero, the baseline was set to vary around zero so that positive and negative noise signals would cancel each other out as much as possible. Second, PQN was attempted instead of area normalization. Although the latter is the *de facto* standard for normalizing MRS metabolomics data, it has

been found to be greatly affected by massive amounts of single metabolites in samples [101], such as those observed for lipids in Paper III. PQN has been widely applied for the normalization of data from biofluids analysis, particularly urine, and has been reported to be more robust and accurate than area normalization in the aforementioned case of extreme concentrations of single analytes [101]. Finally, spectra with high lipid content were excluded from data analysis. The criterion for exclusion was based on a lipid ratio, calculated as the integral of lipid peaks (2.50 – 0.60 ppm) divided by the integral of the total spectral intensity. Samples with lipid ratio ≥ 0.90 were excluded from further analysis. This cut-off was based on previous experience where staining with hematoxylin, erythrosine, and saffron (HES) was performed after HR MAS MRS on the same breast tissue sample [90], revealing that samples with spectra consisting of 90% lipid signals are not faithful representations of tumor tissue. Employing the three aforementioned strategies in Paper III reduced the noise levels considerably. The lipid ratio strategy was implemented for Paper II as well, as presence of high-intensity lipid signals was also an issue, albeit with much fewer spectra affected and lipid-to-metabolites ratio generally much lower. Although tumors in Paper IV were also treated with chemotherapy, the spectral lipid signals were not as elevated as in Paper III. This may be due to availability of a larger amount of tumor tissue from which to cut the sample to be analyzed when using animal models, and thus less chance of incorporating adipose cells during sample preparation.

The two main approaches for extracting biological meaning from MRS metabolomics data are analyzing quantified metabolites or metabolic profiles. With regards to the former, there are several methods available to perform absolute quantification of metabolites from MR spectra. Curve fitting routines, such as Peakfit, previously implemented in our research group to determine the area of metabolite peaks, can handle overlapping signals. However, since this approach can be time consuming as each metabolite peak of interest must be fitted manually, it was considered unfeasible for the high number of samples and untargeted approach employed in Papers II-IV.

A more automatic quantification method that has been successfully employed for the quantification of *ex vivo* spectra of brain [175] and prostate [177, 178] tumor tissue is LCModel [179]. This method uses linear combinations of spectra of individual metabolites to fit the spectra to be quantified, which will be referred to as the target spectra. Attempts were made to establish this method for HR MAS MR spectra of breast tumor tissue, including the generation of spectra of individual metabolites, collectively referred to as a basis set, through computer simulation using NMR-SIM (Bruker BioSpin, Germany) based on chemical shifts

and coupling constants as in [175]. However, LCModel requires all peaks observed in the target spectra, including those considered irrelevant, e.g. lipids or contaminants, to be comprised in the basis set; a signal that is unaccounted for will negatively affect the area calculation for the rest of the peaks in the spectrum. Simulating the lipid peaks in the breast tissue spectra proved to be the biggest challenge, achieving fits of acceptable quality for no more than half of the target spectra (Figure 5.1). This consistent, poor success rate, together with the need to invest time for T2 relaxation correction, led to a re-evaluation of the feasibility of LCModel for the automatic quantification of HR MAS MR spectra of breast tumor tissue, and the establishment of the method was ultimately deemed to be not worth pursuing as one of the goals of this thesis.

Relative quantification using integration was ultimately employed for Papers II-IV. This can be performed automatically once the spectra are appropriately aligned and the metabolite regions of interest are defined. Although relative concentrations are not comparable with other studies, they allow comparisons between two groups within the same study, similarly as for the gene expression microarray measurements (Section 1.5.1) acquired in Papers II and III.

An important drawback of peak integration is that areas of overlapping peaks cannot be distinguished. This can lead to peak areas being influenced by signals other than those belonging to the metabolite of interest. To avoid this, only metabolite regions known to be isolated, i.e. not having overlap, were selected for integration. The procedure was evaluated manually by visually inspecting the regions selected for integration for each spectrum.

As lipid signals overlapping with the lactate peak at 4.11 ppm in Paper III would not allow accurate relative quantification for this metabolite, imputation was employed to determine the missing lactate level values. Although including only the measurable lactate values for LMM analysis was a possible solution, this may introduce another type of bias with regards to the results not being representative of the cohort. Imputation is not usually used in metabolomics studies; however, it is quite frequently applied to gene expression microarray data [180], which can contain 1-11% missing values affecting up to 95% of the measured genes [181]. In fact, imputation was employed on the gene expression microarray data for Papers II and III. Validation for imputation procedures is not typically performed. Nevertheless, we validated the imputation of missing lactate level values in Paper III using a resampling technique, with results indicating high confidence for the procedure (RMSE=0.006, R^2=0.9996).

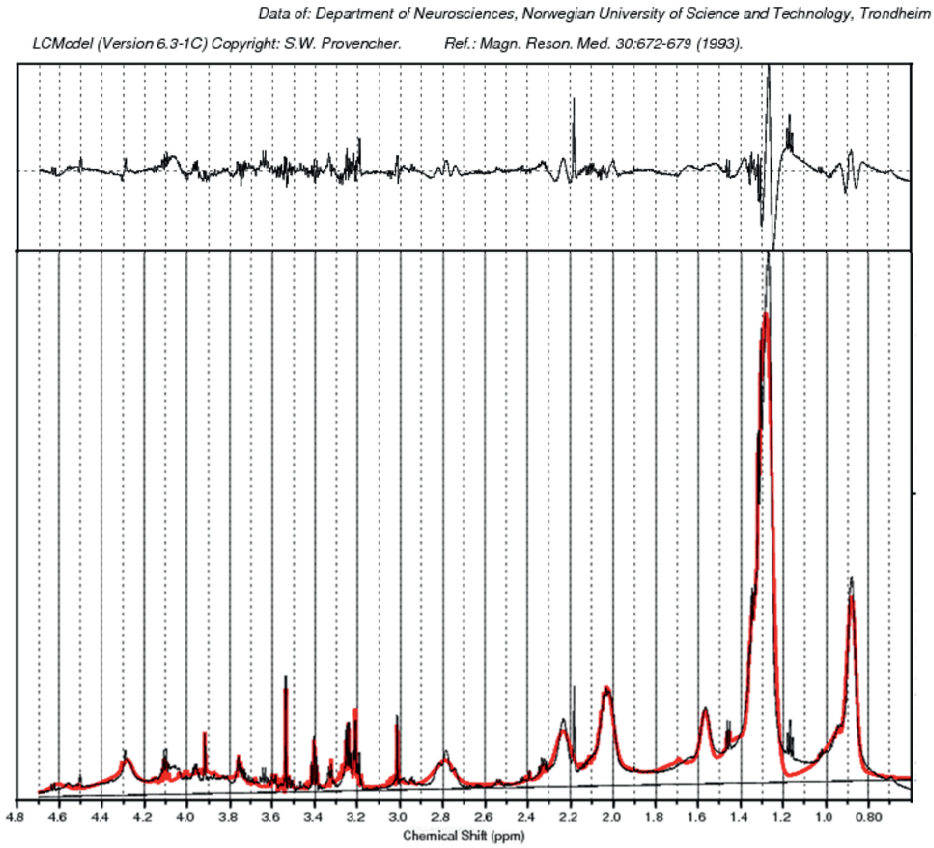


Figure 5.1. LCModel fit for a breast tumor tissue HR MAS MR spectrum with high-intensity lipid signals employing a basis set composed of twenty-three metabolite spectra simulations. The lower panel shows the target spectrum (black) and the fit (red) automatically generated by LCModel. The broad peaks observed from 2.30 to 0.80 ppm correspond to lipid signals. The upper panel shows the LCModel residuals, i.e. unfitted signals. Optimal residuals would be of minimal intensity and would vary randomly around the horizontal dotted line at the center of the upper panel. Residuals show inadequate fit for the lipid peak at 1.30 ppm, as well as for the peak originating at around 2.20 ppm, corresponding to the contaminant lidocaine, a local anesthetic. Cramér-Rao lower bounds [182] vary between 2% and 49% for the metabolite/lipid fits.

Biological pathways involve many molecular compounds that interact with one another. Multivariate analyses can model information from multiple metabolites simultaneously, taking into account the interactions between them while identifying metabolic patterns associated with the biological question of interest. However, supervised multivariate methods, such as PLS-DA used in Papers II-IV, are highly prone to provide over optimistic results. Proper validation of PLS-DA models is therefore of the utmost importance. All studies in this thesis included enough samples to perform a double cross validation, with which the assessment of the quality of the model is performed using samples completely independent of

the optimization of the model, i.e. the number of LVs. It is generally considered that the best case scenario for supervised model validation is to use a completely independent validation set, to assess the quality of the model after optimization with CV. An independent validation set comprises samples that are not alternated to serve for training or testing, but rather excluded so they are solely used for the latter. Although the number of samples in Papers II-IV was probably sufficient enough to implement this, it can be argued that in the independent validation set situation, a single model is built from one combination of samples, while with double CV, different sample combinations from the same population are used for modelling and the quality assessment results are pooled from the different models built. Results from double CV are therefore likely to be more robust than when employing one independent validation set. In addition to double CV, PLS-DA results were validated using permutation testing, to ensure CV results were better than random, giving high confidence to the multivariate analyses.

The study design for Papers III and IV allowed for multilevel analyses. Multiple LMMs were built for this purpose, where the individual metabolite levels were alternated to serve as the response, the fixed effects were treatment group and response group, and the random effect was the different individuals (Paper III) or PDX model types (Paper IV). Since Paper III had a longitudinal design, this study additionally included a fixed effect for time. Multilevel approaches where the information from all metabolites can be analyzed simultaneously are also currently available. One such approach implemented previously in our research group [160] is multilevel PLS-DA [183], with which the within-subject variation is separated from the between-subject variation of spectra paired by time point. This method was attempted in Paper III to maximize the differences between earlier (TP_E) and later time points (TP_L) by subtracting TP_E-TP_L for the former and TP_L-TP_E for the latter. This was carried out after separating samples into either pMRD or pNRs and bevacizumab-receiving or chemotherapy-only patients for each of the three different time point combinations: TP1 vs TP2, TP2 vs TP3, and TP1 vs TP3. Further PLS-DA classifications of the resulting within-subject variation to discriminate between TP_E and TP_L were performed using double CV including both spectra from the same patient in either the training or test set to avoid overfitting.

A significant separation of TP_E from TP_L was achieved with multilevel PLS-DA in all cases ($p \leq 0.003$), indicating a strong metabolic effect due to treatment in accordance with PCA. When compared to unilevel PLS-DA of all time points (not shown in Paper III), multilevel PLS-DA improved classification results in terms of accuracy, sensitivity, and

specificity (Table 5.1), reflecting how time point differences are enhanced when between-subject variations are removed. However, multilevel PLS-DA loadings plots of earlier vs later time points were similar when comparing pMRD and pNRs and bevacizumab-receiving patients and chemotherapy-only patients; essentially, the metabolic effect of treatment appeared to be similar in all groups. Figure 5.2 shows TP1 vs TP2 for pMRD and pNRs as an example.

Table 5.1. PLS-DA and multilevel PLS-DA classification results not shown in Paper III.

Method	Patients/samples included in the model	Discriminated Classes	n	No. of LVs	Class. Accuracy (%)	Sensitivity/ Specifcity (%)	Perm. p-value
Unpaired PLS-DA	All[#]	TP1 vs TP2 vs TP3	122	6	73	73/73	<0.001*
Paired Multilevel PLS-DA	pMRD	TP1 vs TP2	21	1	82	82/82	0.001*
	pNRs	TP1 vs TP2	45	1	73	73/73	<0.001*
	pMRD	TP2 vs TP3	15	2	81	81/81	<0.001*
	pNRs	TP2 vs TP3	40	2	95	95/95	<0.001*
	pMRD	TP1 vs TP3	27	1	100	100/100	<0.001*
	pNRs	TP1 vs TP3	46	2	86	86/86	<0.001*
	Bev-treated	TP1 vs TP2	31	1	84	84/84	<0.001*
	Chemo-only	TP1 vs TP2	35	1	68	68/68	0.003*
	Bev-treated	TP2 vs TP3	29	1	83	83/83	<0.001*
	Chemo-only	TP2 vs TP3	26	2	88	88/88	<0.001*
	Bev-treated	TP1 vs TP3	43	1	88	88/88	<0.001*
	Chemo-only	TP1 vs TP3	30	1	90	90/90	<0.001*

Bev: Bevacizumab; Chemo:chemotherapy; Class.:classification; n: number of patients; Perm.:permutation; pMRD:pathological minimal residual disease; pNRs:pathological non-responders.
[#]Sensitivity/Specificity are reported as the mean of the three class labels. * indicates significant p-values ≤0.05).

When comparing LMM and multilevel PLS-DA, the advantages of the former over the latter include its capability to handle unbalanced data; patients not having samples from both time points available must be excluded when performing multilevel PLS-DA. Furthermore, LMM can analyze the relationship between various classification factors at once. In addition, multilevel PLS-DA can handle samples from only two time points at a time, while a single LMM can incorporate as many time points as there were measured. However, the benefit of simultaneously analyzing all metabolites in multilevel PLS-DA is that interactions between metabolites are taken into account. Multilevel PLS-DA detected metabolic differences due to treatment time only, while LMM could detect differences between pMRD response groups

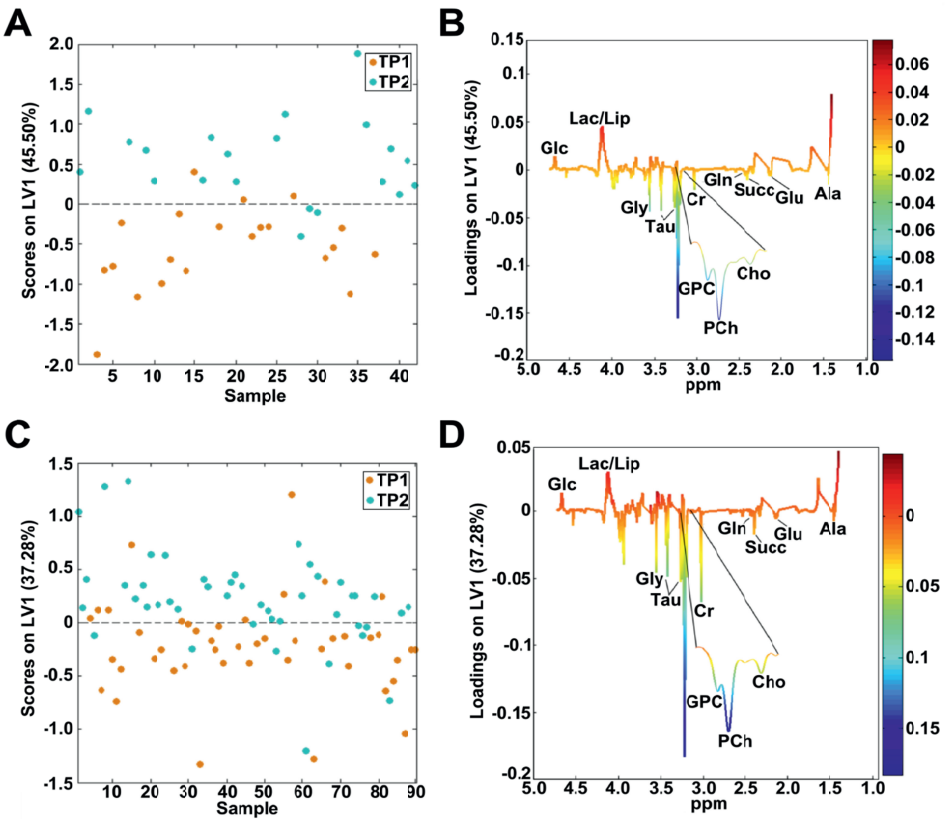


Figure 5.2. Multilevel PLS-DA scores for A) pathological minimal residual disease (pMRD) patients and C) pathological non-responders (pNRs) and loadings for B) pMRD and D) pNRs when discriminating TP1 and TP2 (results not shown in Paper III). Ala:alanine; Cho:choline; Cr:creatine; Glc:glucose; Gln:glutamine; Glu:glutamate; Gly:glycine; GPC:glycerophosphocholine; Lac:lactate; Lip:lipids; LV:latent variable; PCh:phosphocholine, Succ:succinate; Tau:taurine.

and bevacizumab randomization groups in addition. Multilevel PLS-DA could not be employed in Paper IV since paired samples were not available, as discussed in Section 5.1.2.

5.1.5 Combining omics

Biological processes occur as a result of the interaction between molecules from different molecular levels in the omics cascade (Figure 1.5). Different analytical platforms have seen advances that have facilitated the high-throughput collection of large amounts of molecular data from a specific level. The availability of multi-layer omics data provides a unique opportunity to uncover a more comprehensive picture of biological systems. As the different molecular levels are complementary, integrating omics data, rather than focusing on data from a single omics layer, has the potential to improve models for clinical applications. Previous

studies have integrated metabolomics and transcriptomics data to explore possibilities to identify further subgroups in breast cancer using HCA and multidimensional scaling [184] within intrinsic gene expression subtypes [35] and to investigate the underlying mechanisms behind abnormal levels of citrate and choline-containing compounds in prostate cancer using a PLS model relating gene expression (***X***-matrix) and metabolite levels (***Y***-matrix) [185].

In this thesis, particularly in Paper II, metabolite, protein, and gene expression data were merged in the search for a more comprehensive characterization of the molecular heterogeneity of breast cancer. In Paper IV, a targeted approach was employed, where key proteins involved in the PI3K/Akt pathway, as well as the well-known tumor suppressor p53, were selected to investigate the effect of their expression or mutation status, respectively, on metabolite levels. Thus, only a reduced number of molecules were measured. In contrast, an untargeted approach was employed for Paper II, where expression differences within the established metabolic clusters were investigated for 21 851 genes using SAM analysis. Gene expression microarray data was further exploited in Paper II for functional analysis of lists of genes identified by SAM using DAVID, and to determine sets of enriched or over-represented genes within the metabolic clusters using GSEA. Finally, metabolite and gene expression levels were combined using integrative pathway analysis to determine significantly different metabolic pathways between the metabolic clusters.

5.2 Biological and clinical interpretation

In the following sections, the metabolic findings from Papers II-IV are integrated and interpreted by pathway. Metabolomics results are subsequently put into the context of the clinical aspects of breast cancer heterogeneity, treatment effect, and response assessment.

5.2.1 Synthesis of metabolic findings by pathway

Glucose metabolism

To satisfy their high energy demands, cancer cells and normal proliferating cells exhibit the Warburg effect: higher rate of glycolysis and shunting of pyruvate away from the TCA cycle and into lactate production independently of oxygen availability [54]. Of the three metabolic clusters (Mcs) established in Paper II, Mc2 exhibited significantly higher levels of glucose compared to the other metabolic clusters, while Mc3 displayed higher lactate levels from the rest. This suggests lower glucose consumption and hence, lower energy requirements for Mc2 tumors, while the opposite seems to be the case for Mc3 tumors. Based on significant

differences for two metabolites and five genes, integrated pathway analysis detected glycolysis/gluconeogenesis as a significantly different pathway between Mc1 and Mc2, with the latter consistently displaying higher glucose levels. Although the five genes detected could not be directly related to differences in glucose levels, they seem to be involved in directing pyruvate into the TCA cycle instead of lactate production in Mc1. This was inferred from alanine and lactate, the main end products of aerobic glycolysis formed from pyruvate, respectively displaying significantly or close to significant ($p=0.056$) higher levels in Mc1 compared to Mc2.

In both Papers III and IV, the clear metabolic changes observed as an effect of breast cancer therapy included a decreased glycolytic phenotype. In Paper III, glucose was elevated with increasing time of treatment with neoadjuvant chemotherapy, and was significantly higher in pMRD patients compared to pNRs by the end of treatment (TP3). This indicates an even greater drop in glucose consumption rate with treatment response. Similarly, in Paper IV, glucose was significantly higher in everolimus treated animals than untreated controls. PI3K inhibition has been found to lead to an increase in glucose and decrease in lactate concentrations in basal-like xenografts [186]. No significant differences in lactate, however, could be detected between the treatment groups in Paper IV. Nevertheless, the everolimus treated animals displayed a significantly lower lactate-to-glucose ratio, in accordance with a decline in aerobic glycolysis. Furthermore, in Paper IV, p53 mutation was associated with higher lactate levels. This is consistent with wild-type p53 promoting the conversion of pyruvate to Acetyl CoA by the pyruvate dehydrogenase complex (PDC) for its subsequent entrance into the TCA cycle (Figure 1.3). It does this by negatively regulating pyruvate dehydrogenase kinase-2, which inactivates PDC [187].

Interestingly, in Paper III, lactate did not behave as expected with respect to treatment progression and response groups. Levels for this metabolite significantly increased with treatment progression, and were significantly higher in pMRD patients compared to pNRs at TP3, i.e. displaying similar trends as glucose. Elevated lactate is not only usually an indicator of a glycolytic phenotype, but also of poor prognosis, as observed for ER positive breast cancer [92] as well as other cancer types [188-190]. In spite of this, the increase in lactate as an effect of neoadjuvant chemotherapy has been observed previously by our research group [160]. However, when dividing the patient cohort according to 5-year survival in the same study, a trend of higher lactate levels post-treatment in non-survivors compared to survivors was detected. Nevertheless, not all cancer cells exhibit the glycolytic phenotype; a better oxygenated sub-population of cancer cells recycles the lactate secreted by cells undergoing

aerobic glycolysis and preferentially utilizes it to fuel oxidative phosphorylation [5, 191, 192]. It has been suggested that the dominant phenotype of cancer cells depends on the environment, conferring these cells a dual metabolic nature [193]. Under conditions of lactic acidosis as an effect of the Warburg effect, a transition to a non-glycolytic phenotype is promoted. In this setting, a more economic use of glucose stimulated by higher lactate levels may result in glucose and lactate not behaving oppositely as in Paper III. Another possible explanation is accumulation of lactate due to morphological changes in tissue arising as an effect of chemotherapy and consequently interfering with cellular lactate excretion. One must also consider that in Paper III, the relative levels of lactate for 116 out of 270 samples could not be calculated by integration due to the presence of an overlapping lipid peak at 4.13 ppm. Imputation was therefore employed to calculate lactate values for these samples, and this can be considered a limitation of the study. Results for the validation of the imputation process, however, attest to the high certainty of the predicted values.

Glutamine metabolism

As glutamine can be synthesized from glucose, it is a non-essential amino acid for normal, non-proliferating cells. Cancer cells, however, require exogenous glutamine to survive, a condition which is referred to as glutamine addiction [63]. With glucose being shunted away from the TCA cycle in cancer cells, glutamine becomes a critical anaplerotic and source of carbon and nitrogen to replenish the TCA cycle intermediates and for nucleotide, lipid, and protein biosynthesis [64]. Increased glutaminolysis is therefore characteristic in cancer cells.

In Paper II, integrated pathway analysis identified D-glutamine and D-glutamate metabolism to be significantly different between Mc1 and Mc2. However, this was based on significant differences in only two molecules involved in the pathway: glutamate and glutaminase (GLS). Achieving significance with this low number of hits is unusual, but not unexpected, as there are only nine metabolites and genes within this pathway. GLS is an enzyme which catalyzes the first step in glutaminolysis, hydrolyzing glutamine into glutamate (Figure 1.3). Glutamate, in turn, serves as a precursor of non-essential amino acids and glutathione. Mc1 exhibited downregulated GLS and the lowest levels of glutamate. Since no significant differences in glutamine levels were identified between Mc1 and Mc2, these findings in Mc1 tumors may be a result of glutamine being diverted into other metabolic pathways. Alternatively, glutamate in these tumors may be metabolized more rapidly than the rest; this however, would not be consistent with GLS downregulation.

As for the effect of breast cancer therapy on glutamine metabolism, levels of glutamine were found to rise significantly in tumors during neoadjuvant chemotherapy in Paper III. Similarly, PDX models treated with everolimus in Paper IV displayed significantly higher levels of glutamine compared to untreated controls. These findings from both Papers III and IV are consistent with a decline in glutamine consumption/addiction as an effect of treatment. In addition, in Paper III, multilevel LMM identified a significant interaction between *time* and *bevacizumab* for glutathione. Essentially, this indicates that differences in glutathione levels between patients receiving and not receiving bevacizumab vary at different time points. Being one of the most important cellular antioxidants, glutathione provides cancer cells with yet another mean to avoid apoptosis which could ultimately be caused by ROS [67]. We found that chemotherapy-only patients displayed significantly higher glutathione levels than bevacizumab-receivers at TP3. This is consistent with findings from a previous study in glioblastoma [194], where reduction of glutathione levels with bevacizumab treatment was observed. This suggests that bevacizumab may promote cancer cell apoptosis through the disruption of redox homeostasis.

In Paper IV, however, while significant changes in glutamine were observed with treatment progression in both responding and non-responding PDX models, the same was not the case for metabolites involved in glycolysis for the non-responding group. This is in agreement with findings from a previous study [195] investigating the metabolic dependencies of cells with mutated PIK3CA, one of the genes driving PI3K signaling. While glutamine-dependent growth was observed in both mutant and wild-type cells, only the former exhibited glucose-dependent growth. In Paper IV, metabolic dependencies in responding PDX models were similar to those of the PIK3CA-mutated cells in the aforementioned study, while non-responding PDX models behaved similarly to the wild-type cells in this respect. This may reflect a higher dependency for PI3K/AKT/mTOR signaling in tumors from responding PDX models, consequently making these models more sensitive to treatment targeting this pathway. Nevertheless, when investigating the effect of everolimus on genes and proteins in the same cohort of animals studied in Paper IV [154], PIK3CA mutations were found to be rare and not predictive of everolimus treatment response.

Choline phospholipid metabolism

Abnormal choline metabolism can be considered an emerging hallmark of cancer, and HR MAS MRS detects the metabolites choline, phosphocholine, and glycerophosphocholine

(GPC). These choline-containing compounds have been associated with malignant transformation [66, 83].

In Paper II, the most different metabolic cluster, Mc1, was defined by significantly higher levels of GPC and phosphocholine. The expression of *CHKA*, the gene encoding for the enzyme choline kinase alpha that converts choline to phosphocholine (Figure 1.4), has been found to correlate with GPC and phosphocholine expression [81, 196]. Although differences in *CHKA* expression comparing Mc1 with the other two metabolic clusters could not be identified by SAM analysis, significantly higher expression levels for the gene in Mc1 were confirmed through univariate analysis. Furthermore, glycerophospholipid metabolism was identified as the most significantly different pathway between Mc1 and Mc2 by integrated pathway analysis. This was based on significantly higher levels of GPC and phosphocholine, as well as significantly downregulated genes *LCAT*, *LPCAT2*, *PPAP2A*, *PPAP2B*, *PLD1*, and *AGPAT4* in Mc1 compared to Mc2. These genes act upon different steps of the phosphatidylcholine cycle, of which GPC and phosphocholine are intermediates (Figure 1.4). This cycle is linked to lipid turnover at the cell membrane, which is crucial for cell proliferation [197]. GPC and phosphocholine may therefore accumulate in Mc1 tumors due to lower phosphatidylcholine cycle activity. Integrated pathway analysis comparing Mc1 and Mc3 revealed the same trends in metabolites and genes in Mc1 when compared to Mc2, with additional downregulation of *PLA2G5*. This gene encodes for the phospholipase A2 enzyme that hydrolyzes phosphatidylcholine to acyl-GPC, the precursor of GPC (Figure 1.4). Altogether, these findings suggest that Mc1 tumors experience a higher degree of alteration in phosphatidylcholine metabolism than the rest.

The choline-containing compounds were also found to be affected by treatment in Papers III and IV. GPC, phosphocholine, and choline, i.e. all the constituents of tCho (Section 1.2.3), significantly decreased as an effect of neoadjuvant chemotherapy in Paper III. Of these, however, only choline was significantly lower in pMRD patients compared to pNRs. In Paper IV, while phosphocholine was found to be significantly lower in treated animals compared to untreated controls, the opposite was observed for GPC/PCh. GPC on its own was not significantly different between the treatment groups, and phosphocholine is thus assumed to be the main contributor of variation between treated animals and controls for the GPC/PCh ratio. In addition, phosphocholine was different between treated animals and controls within both responding and non-responding PDX, while GPC/PCh was different between treatment groups within responders only. Phosphocholine therefore seems to be the most affected tCho-constituent by everolimus treatment. All in all, these findings point to a decrease in choline

phospholipid metabolism as an effect of breast cancer therapy, which appears more prominent in responders. This decrease may reflect an attenuation of malignancy and aggressiveness in the treated tumors.

Other metabolic findings

In Paper III, other metabolites that were significantly lower in pMRD patients compared to pNRs at TP3 included succinate, tyrosine, and creatine. The latter two also decreased with time on therapy, while succinate was also significantly lower in pMRD patients compared to pNRs at TP2 and TP3. In addition to playing a key role in neurotransmitter and thyroid hormone synthesis, the amino acid tyrosine is present in the intracellular domain of receptor tyrosine kinases (RTKs). However, it is difficult to attribute the tyrosine pool within a cell to a single metabolic pathway. Thus, more studies are needed to investigate the relevance of this metabolite in breast cancer treatment.

Creatine promotes the rapid formation of ATP, thus aiding in supplying energy to all cells [198]. High levels of creatine before treatment and in pNRs may reflect an additional attempt by cancer cells in meeting their high energy requirements. In Paper IV, creatine was found to be associated with p53 mutation. The tumor suppressor p53 activates the expression of the enzyme that catalyzes creatine synthesis [199]. Dysregulated creatine synthesis in cancer cells with p53 mutation may result as a means of compensation for inefficient glycolytic energy metabolism. In breast cancer, creatine has been found to be higher in treatment naïve ER positive tumors when compared to ER negative [74], suggesting it may be associated with a better prognosis before treatment.

Succinate is a TCA cycle metabolite that is oxidized by succinate dehydrogenase (SDH) to form fumarate (Figure 1.3). SDH has been identified as a tumor suppressor; when this enzyme is inactivated, succinate is accumulated in the mitochondria and then leaks into the cytosol [200]. Cytosolic succinate subsequently inhibits prolyl hydroxylases involved in the degradation of the alpha subunit of the hypoxia-inducible factor (HIF alpha) [201]. HIF alpha is a transcription factor that activates mechanisms that promote survival, migration, and the glycolytic phenotype in cancer cells [57, 200]. Decreased succinate in pMRD patients compared to pNRs at TP2 and TP3 suggests a less malignant phenotype in responders after neoadjuvant chemotherapy.

Taurine is an osmolyte involved in cell volume and fluid balance maintenance. Findings regarding taurine vary with different types of cancer [69]. In breast cancer, it has been found to be increased in ER positive tumors and five year survivors compared to ER negative and

non-survivors, respectively [65, 74]. This suggests taurine to be associated with better prognosis in breast cancer. In Paper III, a significant increase in taurine in pMRD patients compared to pNRs was observed at TP3. However, multivariate analyses could not confirm this.

The increase of lipids with treatment progression observed in Paper III suggests a normalization to breast adipose tissue. Although lipid metabolism is known to mediate cancer cell apoptosis and necrosis induced by the administered treatment [202], most lipid signals in HR MAS MR spectra are assumed to originate from normal breast adipose cells. In accordance with this, TP3 and normal-like samples from Paper III corresponded well in the PCA scores plot of all samples. The normal-like subtype displays high expression of basal epithelial and adipose cell genes and is characterized as adipose-enriched [30].

5.2.2 Characterization of metabolic heterogeneity in breast cancer

Because heterogeneity influences disease progression and therapeutic response, its study is one of the most clinically-relevant areas of cancer research. The heterogeneous nature of breast cancer has been observed at different molecular levels within a tumor (Section 1.1.5). Metabolomics is downstream from the more familiar omics, and metabolites, being the end products of cellular processes, can be considered the closest link to the phenotype. However, the biological and clinical implications of metabolic heterogeneity in breast cancer are yet to be established.

Paper II aimed to derive subgroups of breast cancer based on metabolic differences, and to characterize them on the transcriptomic and proteomic levels. The division into metabolic subgroups was performed using HCA in a similar approach as performed by Perou *et al.* [30] to define the intrinsic subtypes based on gene expression, which have been found to correlate with clinical outcome. Three metabolic clusters were identified in Paper II: Mc1, Mc2, and Mc3. Metabolic diversity was confirmed between the Mcs by means of univariate comparisons of metabolite levels that revealed that 15 out of 18 metabolites measured were significantly different between at least two clusters. Furthermore, permutation testing after multivariate PLS-DA model building using metabolic profiles to predict the Mcs showed that all three of these clusters exhibited significantly different metabolic profiles ($p<0.001$).

To further characterize the metabolic clusters, their association with RPPA and intrinsic subtypes classified according to protein and gene expression, respectively, was investigated. The distribution of RPPA subtypes was found to be significantly different between all three clusters. This suggests that the metabolome reflects the expression of cancer-related proteins,

in this case those used for RPPA classification. In line with this, distinct metabolic signatures among protein expression subtypes have previously been reported by our research group; metabolic profiles have been found to be predictive of ER and PgR status [74] and to be capable of differentiating TNBC from ER+/PgR+/HER2+ tumors [91]. In Paper IV, however, the metabolic heterogeneity of TNBC was manifested, as PCA using metabolite information revealed clear groupings in the scores plot by PDX model type. Although findings from [91] indicate that metabolic variability within the TNBC subtype is lower than that between TNBC and ER+/PgR+/HER2+ breast cancer, the metabolic heterogeneity within TNBC is an important factor to be aware of when translating metabolomics findings related to this subtype to the clinic. Findings from studying one PDX model type may not be applicable to the whole TNBC subtype. This was exemplified in Paper IV, with initial metabolic differences between subjects hindering prediction of treatment response. Furthermore, the metabolic heterogeneity observed indicates that treatment stratification within TNBC may be beneficial. A variety of gene expression subtypes was represented in the cohort of PDX models examined in Paper IV, and a previous study in our research group found that basal-like breast cancer PDX responded to PI3K inhibitors, while this was not the case for luminal-like PDX [186]. However, sources of heterogeneity other than gene expression, e.g. at the metabolic level, are worth investigating for treatment optimization. This is supported by both responding and non-responding groups of PDX in Paper IV consisting of a combination of basal-like, luminal-like, and HER2-enriched tumors.

With regards to the relation between metabolic and transcriptomics molecular levels, no significant difference in distribution of intrinsic subtypes between the metabolic clusters in Paper II was detected. Similarly, in Paper III, PCA on metabolic profiles revealed no groupings by intrinsic subtype in the scores plot. This is an indication that the metabolic level contributes with additional information beyond the intrinsic subtypes as an innate component of breast cancer heterogeneity. Metabolite information could therefore be complimentary to improve breast cancer diagnosis and to further stratify patients to therapy. With regards to the genomics level, however, in Paper IV, p53 mutation status for PDX TNBC models could be discriminated based on metabolite information. In contrast, significant discrimination for the expression or no expression of key proteins involved in PI3K pathway regulation could not be achieved. This, together with the metabolic heterogeneity within intrinsic subtypes observed in Papers II and III, seems counterintuitive; the genome is at a level farther from the metabolome than the proteome and the transcriptome. One would therefore expect metabolite levels to be more dependent on proteins or mRNA than genes, with biological information

following a continuous flow between genomics, transcriptomics, proteomics, and metabolomics, and vice versa. However, since p53 is the most important tumor suppressor [203] and the most frequently mutated gene in human cancers (50% of all cases) [204], it should not come as a surprise that the information it holds is generally translated all the way to the metabolic level. In fact, p53 mutation has been found to influence energy metabolism at multiple levels [205].

In spite of each omic level providing different information, the benefit of combining this information for the elucidation of biological mechanisms was demonstrated in Paper II. Of the metabolically-derived Mcs, Mc1 was the most phenotypically different based on it being connected with Mc2 and Mc3 at a greater height in the HCA dendogram. At the protein level, the majority of Mc1 tumors (49%) were classified as RPPA-luminal, while only 12% were classified as RPPA-reactive I or II. The proteins characteristic of these reactive RPPA subtypes are thought to be a product of the tumor microenvironment [33], suggesting a less active tumor microenvironment in Mc1. At the gene expression level, gene sets related to the extracellular matrix, a component of the tumor microenvironment, were found to be downregulated in Mc1 compared to Mc2 and Mc3. Furthermore, the prognostic potential of the identified differences in extracellular matrix activity was investigated. Breast carcinomas have previously been stratified into four groups based on differences in expression of genes related to the extracellular matrix, termed ECM classification [206]. Although this classification was not performed for the cohort in Paper II, downregulation of several extracellular matrix-related genes in Mc1 matched those downregulated in the ECM2 group, which is associated with a better prognosis than ECM1 tumors [206].

5.2.3 Investigating the effect of breast cancer therapy on metabolic profiles

The metabolome is a highly dynamic entity that is not only influenced by physiological processes within a biological system, but also by external environmental stimuli. The genome, transcriptome, and proteome are much less prone to undergo changes due to environmental influences. Among these external environmental factors exerting modifications in metabolic profiles is medication. Following the administration of a certain therapy, changes in the metabolome occur as a result of the biological activity of said therapy. Pinpointing specific drug-related alterations in metabolite levels and pathways as an effect of a specific therapy can provide information related to mechanism of action and/or resistance, toxicity, interactions with other drugs, or even predict clinical outcome.

The effect of different breast cancer therapies on metabolic profiles was investigated in Papers III and IV. In Paper III, the effect of neoadjuvant chemotherapy with or without bevacizumab in patients was investigated, while that of the mTOR inhibitor everolimus was assessed in PDX models in Paper IV. In both Papers III and IV, the strongest finding was a conversion to a less malignant metabolic profile (Section 5.2.1) as an effect of the corresponding therapies.

In Paper III, all patients received FEC from TP1 to TP2 and taxane treatment from TP2 to TP3. Multilevel analyses of individual metabolites detected more significantly different metabolites between TP3 and TP2 than between TP2 and TP1, suggesting a higher metabolic effect due to taxanes compared to FEC. However, samples taken at TP2 had been exposed to FEC, which could influence the impact of the subsequent taxane treatment. A more unbiased comparison of the drug types would require a study design where patients were administered both types independently, without possible interaction effects from prior treatment.

The changes observed due to neoadjuvant chemotherapy in Paper III were independent of bevacizumab administration. Other than a subtle change in glutathione levels detected due to an interaction between the fixed effects of *time* and *bevacizumab*, changes in metabolic profiles due to the antiangiogenic drug were not detected. Nonetheless, the bevacizumab-treated group exhibited a slightly higher percentage of pCR than the control group (22 % vs 11 %, respectively). Both bevacizumab-treated and chemotherapy-only patients underwent pronounced metabolic changes as an effect of chemotherapy. Metabolic changes due to bevacizumab may be very slight compared to the potent chemotherapeutic effect, resulting in them being concealed, even to the multilevel approach employed in Paper III. Furthermore, an unexpected increase in lactate was observed with therapy progression. Lactate has been found to stimulate angiogenesis dependent on interleukin 8 (IL-8) [207], potentially counteracting the inhibitory effects of bevacizumab on VEGF-dependent angiogenesis. In fact, increased IL-8 signaling has been identified *in vitro* as a proangiogenic compensatory mechanism enabling bevacizumab resistance in head and neck squamous cell carcinoma [208]. Altered lactate metabolism may therefore play a role in the development of tumor resistance to bevacizumab.

In Paper IV, multivariate PLS-DA could discriminate everolimus treated animals and untreated controls, with six metabolites identified to have the most discriminatory power. Of these, the levels of five metabolites significantly differed between the treatment groups when assessed individually using multilevel LMM. Agreeing trends among the treatment groups for these metabolites were detected with both methods, pointing to reduced glycolytic lactate production and glutaminolysis after treatment (Section 5.2.1). Both techniques therefore

revealed the same metabolic differences between treated animals and controls, which were in accordance with PI3K/Akt signaling pathway inhibition.

As the metabolic clusters in Paper II were established from untreated tumors, the distinct metabolic traits observed for each of these clusters could provide a means to further stratify patients for treatment targeting different metabolic pathways. Mc1 tumors, exhibiting a more pronounced dysregulation in choline phospholipid metabolism, may benefit from choline kinase-targeted therapy, which has been shown to cause growth arrest and apoptosis in preclinical models [209]. Mc3 displayed significantly higher lactate than the rest and lower glucose along with Mc1, suggesting a higher sensitivity for treatment targeting glycolytic enzymes. The addition of glycolytic-targeted therapy to chemotherapy has been found to re-sensitize chemo-resistant cancer cells [210].

5.2.4 Assessment of breast cancer therapy response using HR MAS MRS metabolomics

Paper III included three different criteria to assess response to neoadjuvant chemotherapy with and without bevacizumab (Section 5.1.2). Metabolic profiles were unable to predict pathological response before treatment (TP1). By the end of treatment (TP3), however, a significant difference in metabolic profiles between pCR+ patients and pCR-, which was even clearer between pMRD patients and pNRs, was identified. pCR+ and pMRD patients exhibited similar metabolite trends compared to pCR- and pNRs, respectively. Metabolic profiles therefore show potential in providing complementary information to other omics levels for understanding the mechanism of treatment response in breast cancer. Nevertheless, the observed metabolic differences between pathological responders and non-responders may arise due to differences between tumor and normal tissue, since for the former most of the cancer cells will have responded to treatment. Still, in a previous study by our research group assessing metabolic response to neoadjuvant chemotherapy [160], HR MAS MRS metabolic profiles were more strongly associated to 5-year survival than to treatment response. The relation between metabolic profiles and the clinical outcome of 5-year survival will be followed-up on when this data is available for this cohort.

With regards to the response ratio criteria, significant metabolic differences between the good response and no response groups were detected at both TP1 and TP3. At TP1, the good response group exhibited higher levels of metabolites associated with a more malignant phenotype as described in Section 5.2.1, including the choline-containing metabolites, glutathione, glutamate, and succinate. At TP3, an apparent metabolic switch had occurred,

with the no response group now displaying a more malignant metabolic profile. Since a considerable number of patients, i.e. the intermediate response group, were excluded to build the model to discriminate the extreme response ratio groups at TP1, this model is not adequate for prediction of treatment response. Nonetheless, our findings suggest that patients with metabolic profiles that would be considered more unfavorable are more likely to experience tumor reduction with chemotherapy.

As for the utility of metabolite information predicting response to bevacizumab, this may be worth studying in the adjuvant setting. The randomized phase 3 trial NASBP B-40 investigated the benefit of adding neoadjuvant bevacizumab continued post-operatively to neoadjuvant chemotherapy in female patients with large, HER2 negative breast tumors [211]. A significant increase in overall survival (OS) was observed in bevacizumab receivers compared to those receiving chemotherapy only. The significant metabolic differences between pathological responders and non-responders we observed at surgery (TP3) in Paper III, may prove to be useful for clinical decision-making in terms of interrupting or continuing bevacizumab treatment after surgery.

In Paper IV, multivariate PLS-DA could not discriminate between metabolic information from responding and non-responding PDX models in either treated animals or untreated controls. This was attributed to metabolic heterogeneity among the 13 different PDX types used interfering with the prediction of treatment response based on metabolite information (Section 5.2.2). The between-PDX model heterogeneity should be accounted for when building LMM including the random effect *PDX model*; however, this technique could not detect differences between the response groups either. As an alternative approach, differences in treated animals and untreated controls were studied separately in responding and non-responding PDX models. This revealed a more prominent metabolic effect of everolimus in responding PDX models, with significant changes in metabolites involved in glycolysis being only detected in these models and not in the non-responders.

In an attempt to determine the clinical utility of the metabolic clusters of breast cancer established in Paper II, samples from Paper III were classified as Mc1, Mc2, or Mc3 based on metabolic profiles using the PLS-DA model established in Paper II. The purpose of this was to compare the rate of response to therapy among clusters. However, a large proportion of the samples (79%) were classified as Mc2 (Figure 5.3). This considerable imbalance in sample size among clusters did not allow for an appropriate comparison in response rate. For instance, at TP1 (Figure 5.3A), four, thirty-four, and three samples classified as Mc1, Mc2, and Mc3, respectively, belonged to pMRD patients. This results in a similar pMRD response

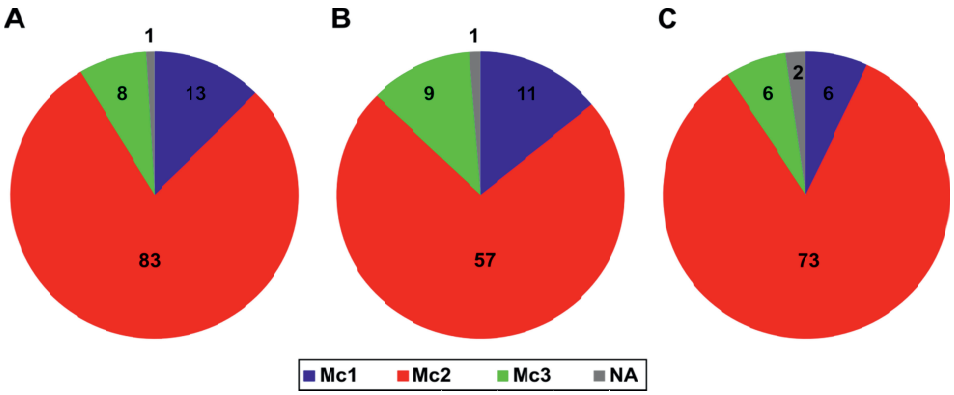


Figure 5.3. Number of samples in Paper III classified as Mc1, Mc2, and Mc3, at TP1 (A), TP2 (B), and TP3 (C). Mc: metabolic cluster; NA: not available.

rate (31 - 41 %) for all clusters at baseline, but cannot be considered representative for Mc1 and Mc3, due to the low number of samples in these subgroups. The association of the clusters to 5-year survival will therefore be investigated when this data is available for the cohort examined in Paper II.

Response criteria based on anatomical measurements of tumors, such as pMRD response and response ratio (Paper III) and RTV (Paper IV), as well as the more standardized RECIST criteria, are traditionally employed to assess cancer therapy response. In spite of significant metabolic changes as an effect of treatment being detected in both Papers III and IV, metabolic association to actual treatment response or resistance was more subtle in both studies. For some more recently developed treatment strategies, such as those targeting the immune system, anatomical criteria have been found inappropriate for measuring long-term treatment benefit [212].

The potential role of imaging biological and physiological changes in tissue for predicting cancer treatment response is of increasing interest in research [213, 214]. One such imaging modality is positron emission tomography (PET), a highly sensitive and specific clinical tool used to identify primary and metastatic lesions [215, 216]. This technique is employed to observe metabolic processes using positron-emitting radionuclide tracers. The glucose analogue tracer ^{18}fluorodeoxyglucose (FDG) is commonly used in PET to image and quantify glucose uptake, which has been found to be correlated with poor prognosis and highly dependent on the glycolytic rate [215]. Detection of pathological response to neoadjuvant chemotherapy in stage II and III breast cancer patients has been shown to be more accurate based on glucose uptake measured by FDG PET than by assessment of tumor size [217]. A major disadvantage of PET is that it can only provide information on one

metabolite at a time. *In vivo* MRS, on the other hand, can measure several metabolites simultaneously. A significant decrease of the tCho peak (Section 1.2.3) measured *in vivo* using MRS has been suggested as an indicator of response to neoadjuvant chemotherapy in locally advanced breast cancer [86, 218]. The clinical use of *in vivo* MRS, however, is limited by its low sensitivity, and tCho cannot be detected in all tumors.

Hyperpolarization of atomic nuclei can increase MR sensitivity more than 10 000 fold [219]. Hyperpolarized MRS is a technique that facilitates the measurement of metabolism in real time [214]. This technique was employed in [220] for the assessment of glioma tumor radiotherapy in rats, where a positive response determined based on 2-month survival, renewed weight gain, and increased appetite and activity could be detected by reduced hyperpolarized pyruvate to lactate conversion in spite of an increase in tumor size. Such an increase, which simulates progressive disease and is thus termed pseudo-progression, can be observed in brain tumors after combined chemotherapy and radiotherapy. This technique, however, requires the injection of hyperpolarized tracers into the subject. The short lifetime of these tracers leads to a fast decay of the enhanced signal and is arguably the major limitation of hyperpolarized MRS. Nevertheless, a first clinical study has shown promise for the safety and feasibility of this technique to non-invasively characterize metabolic alterations in cancer [221].

The promising findings from metabolic imaging techniques strengthen the argument that changes in tumor metabolism are of clinical significance independent of changes in size. The metabolic changes as an effect of treatment in Papers III and IV may therefore be correlated with clinical outcomes other than anatomical measures of response. The association between metabolic response and 5-year PFS and OS from Paper III should be investigated once this data is available.

6 Concluding remarks and future perspectives

In this thesis, MR-based metabolomics was used to identify metabolic features contributing to breast cancer heterogeneity and to assess currently available therapies for the disease. For this, advanced statistical techniques were employed, including optimization of spectral preprocessing, and information from the four molecular levels of the omics cascade was utilized.

Metabolic heterogeneity was manifested as an additional component in the complex biology of breast cancer, and gave way to the establishment of three metabolic subgroups or clusters. These clusters displayed diversity not only at the metabolic level, but also in the expression of breast cancer-related proteins and genes linked to extracellular matrix. Notably, metabolic heterogeneity was found to affect the prediction of response to treatment in terms of reduction of tumor size. Moreover, therapies that target the metabolic attributes characteristic of the established clusters are currently available, attesting to the potential benefit of stratifying patients based on metabolic information. This, along with the prognostic value of the clusters in terms of predicting clinical outcomes such as 5-year survival, should be further investigated.

Metabolites involved in pathways known to be aberrant in cancer were found to be affected by breast cancer treatment in both patient-derived xenografts (PDX) and human subjects. Among the clear metabolic changes observed in patients, a decline in glucose consumption and a transition to normal adipose tissue seemed to occur as neoadjuvant treatment with chemotherapy progressed. Metabolic profiles could not predict pathological response before treatment was completed. Nevertheless, the significant discrimination between pathological responders and non-responders at surgery suggests potential for MRS metabolomics in complementing other molecular assays for the elucidation of the underlying mechanisms affecting pathological response. Importantly, metabolic profiles from Paper III may be correlated with survival, and will be evaluated once this data is available. Furthermore, treatment with the mTOR inhibitor everolimus appeared to decrease glycolytic lactate production and glutaminolysis in PDX, in accordance with PI3K/AKT/mTOR inhibition. This treatment-dependent decrease was more prominent in responding PDX models than in non-responders.

Although only a subtle change in glutathione as an effect of bevacizumab was detected, which indicated impairment in redox stability when treating with the antiangiogenic drug, metabolite information may be useful in predicting response to bevacizumab in the adjuvant

setting. The predictive potential of the distinct metabolic signatures identified for pathological responders and non-responders at surgery is worth exploring, as new evidence has emerged indicating that neoadjuvant bevacizumab continued post-operatively can increase overall survival in breast cancer patients. Furthermore, a possible resistance mechanism to bevacizumab was observed in the increase in lactate, as this promotes new vessel formation mediated by pro-angiogenic factors other than vascular endothelial growth factor (VEGF), the target of bevacizumab.

The clinical utility of imaging biochemical and physiological function in cancer is gaining increasing interest. Assessment of cancer treatment based on features of tumor metabolism has been shown to outperform anatomical or morphological criteria. The metabolic response to treatment observed in the studies included in this thesis may be indicative of other clinical outcomes, ultimately being useful for clinical decision-making during treatment monitoring. Metabolic signatures related to survival rates should be investigated, once five-year follow-up data has been collected for the clinical cohorts.

Overall, MRS metabolomics proved to be a valuable tool for further characterizing breast tumor heterogeneity and for identifying opportunities for improving treatment management. Furthermore, this thesis employed novel approaches for extracting biological meaning from clinical and preclinical data from different molecular levels. The findings of this thesis may contribute to a more refined stratification of breast cancer patients, as well as to a better understanding of the underlying mechanisms affecting response to treatment.

References

1. Cooper GM. The Cell: A Molecular Approach. Second ed. Sunderland (MA): Sinauer Associates; 2000.

2. Lodish H, Berk A, Zipursky L, Matsudaira P, Baltimore D, Darnell J. Section 8.1, Mutations: Types and Causes. Molecular Cell Biology. 4th Edition. New York, NY, USA: W.H. Freeman; 2000.

3. Turnbull C, Hodgson S. Genetic predisposition to cancer. Clin Med. 2005;5(5):491-8. doi:10.7861/clinmedicine.5-5-491.

4. Hanahan D, Weinberg RA. The Hallmarks of Cancer. Cell. 2000;100(1):57-70. doi:http://dx.doi.org/10.1016/S0092-8674(00)81683-9.

5. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell. 2011;144(5):646-74. doi:http://dx.doi.org/10.1016/j.cell.2011.02.013.

6. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359-E86. doi:10.1002/ijc.29210.

7. Cancer Registry of Norway. Cancer in Norway 2014 - Cancer incidence, mortality, survival and prevalence in Norway Oslo, Norway: Cancer Registry of Norway; 2015.

8. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P et al. Genetic Heterogeneity and Penetrance Analysis of the BRCA1 and BRCA2 Genes in Breast Cancer Families. Am J Hum Genet. 1998;62(3):676-89. doi:http://dx.doi.org/10.1086/301749.

9. American Cancer Society. Special Section: Female Breast Cancer. Global Cancer Facts & Figures. Third ed. Atlanta, GA, USA: American Cancer Society; 2015.

10. Visvader JE, Stingl J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. Gene Dev. 2014;28(11):1143-58. doi:10.1101/gad.242511.114.

11. Suryadevara A, Paruchuri LP, Banisaeed N, Dunnington G, Rao KA. The clinical behavior of mixed ductal/lobular carcinoma of the breast: a clinicopathologic analysis. World J Surg Oncol. 2010;8:51-. doi:10.1186/1477-7819-8-51.

12. Norsk Bryst Cancer Gruppe (NBCG). Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av pasienter med brystkreft. Oslo, Norway: Helsedirektoratet; 2016.

13. Vetto J, Pommier R, Schmidt W, Wachtel M, DuBois P, Jones M et al. Use of the "Triple Test" for palpable breast lesions yields high diagnostic accuracy and cost savings. Am J Surg. 1995;169(5):519-22. doi:http://dx.doi.org/10.1016/S0002-9610(99)80209-8.

14. AJCC Cancer Staging Manual. Sixth ed. Chicago, IL, USA: American Joint Committee of Cancer; 2002.

15. Cianfrocca M, Goldstein LJ. Prognostic and Predictive Factors in Early-Stage Breast Cancer. Oncologist. 2004;9(6):606-16. doi:10.1634/theoncologist.9-6-606.

16. Soerjomataram I, Louwman MWJ, Ribot JG, Roukema JA, Coebergh JWW. An overview of prognostic factors for long-term survivors of breast cancer. Breast Cancer Res Treat. 2008;107(3):309-30. doi:10.1007/s10549-007-9556-1.

17. Lumachi F, Brunello A, Maruzzo M, Basso U, Basso SMM. Treatment of Estrogen Receptor-Positive Breast Cancer. Curr Med Chem. 2013;20(5):596-604. doi:http://dx.doi.org/10.2174/092986713804999303.

18. Tokunaga E, Hisamatsu Y, Tanaka K, Yamashita N, Saeki H, Oki E et al. Molecular mechanisms regulating the hormone sensitivity of breast cancer. Cancer Sci. 2014;105(11):1377-83. doi:10.1111/cas.12521.

19. Anderson WF, Chatterjee N, Ershler WB, Brawley OW. Estrogen Receptor Breast Cancer Phenotypes in the Surveillance, Epidemiology, and End Results Database. Breast Cancer Res Treat. 2002;76(1):27-36. doi:10.1023/a:1020299707510.

20. Liu S, Chia SK, Mehl E, Leung S, Rajput A, Cheang MC et al. Progesterone receptor is a significant factor associated with clinical outcomes and effect of adjuvant tamoxifen therapy in breast cancer patients. Breast Cancer Res Treat. 2010;119(1):53-61. doi:10.1007/s10549-009-0318-0.

21. Natarajan L, Pu M, Parker BA, Thomson CA, Caan BJ, Flatt SW et al. Time-Varying Effects of Prognostic Factors Associated With Disease-Free Survival in Breast Cancer. Am J Epidemiol. 2009;169(12):1463-70. doi:10.1093/aje/kwp077.

22. Costa SD, Lange S, Klinga K, Merkle E, Kaufmann M. Factors influencing the prognostic role of oestrogen and progesterone receptor levels in breast cancer—results of the analysis of 670 patients with 11 years of follow-up. Eur J Cancer. 2002;38(10):1329-34. doi:http://dx.doi.org/10.1016/S0959-8049(02)00067-9.

23. Borg Å, Tandon AK, Sigurdsson H, Clark GM, Fernö M, Fuqua SAW et al. HER-2/neu Amplification Predicts Poor Survival in Node-positive Breast Cancer. Cancer Res. 1990;50(14):4332-7.

24. Ménard S, Balsari A, Casalini P, Tagliabue E, Campiglio M, Bufalino R et al. HER-2-positive Breast Carcinomas As a Particular Subset with Peculiar Clinical Behaviors. Clin Cancer Res. 2002;8(2):520-5.

25. Winstanley J, Cooke T, Murray GD, Platt-Higgins A, George WD, Holt S et al. The long term prognostic significance of c-erbB-2 in primary breast cancer. Br J Cancer, BJC. 1991;63(3):447-50.

26. Paik S, Bryant J, Park C, Fisher B, Tan-Chiu E, Hyams D et al. erbB-2 and Response to Doxorubicin in Patients With Axillary Lymph Node-Positive, Hormone Receptor- Negative Breast Cancer. Journal of the National Cancer Institute. 1998;90(18):1361-70. doi:10.1093/jnci/90.18.1361.

27. Carlomagno C, Perrone F, Gallo C, De Laurentiis M, Lauria R, Morabito A et al. c-erb B2 overexpression decreases the benefit of adjuvant tamoxifen in early-stage breast cancer without axillary lymph node metastases. J Clin Oncol. 1996;14(10):2702-8.

28. Ismail-Khan R, Bui MM. A review of triple-negative breast cancer. Cancer Control. 2010;17(3):173-6.

29. Giordano SH. Update on Locally Advanced Breast Cancer. Oncologist. 2003;8(6):521-30. doi:10.1634/theoncologist.8-6-521.

30. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA et al. Molecular portraits of human breast tumours. Nature. 2000;406. doi:10.1038/35021093.

31. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA. 2001;98(19):10869-74. doi:10.1073/pnas.191367098.

32. Metzger-Filho O, Tutt A, de Azambuja E, Saini KS, Viale G, Loi S et al. Dissecting the Heterogeneity of Triple-Negative Breast Cancer. J Clin Oncol. 2012. doi:10.1200/jco.2011.38.2010.

33. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61-70. doi:http://www.nature.com/nature/journal/v490/n7418/abs/nature11412.html#supplementary-information.

34. Hennessy BT, Lu Y, Gonzalez-Angulo AM, Carey MS, Myhre S, Ju Z et al. A Technical Assessment of the Utility of Reverse Phase Protein Arrays for the Study of the Functional Proteome in Non-microdissected Human Breast Cancers. Clin Proteomics. 2010;6(4):129-51. doi:10.1007/s12014-010-9055-y.

35. Borgan E, Sitter B, Lingjærde OC, Johnsen H, Lundgren S, Bathen TF et al. Merging transcriptomics and metabolomics - advances in breast cancer profiling. BMC Cancer. 2010;10(1):1-14. doi:10.1186/1471-2407-10-628.

36. Yiannakopoulou EC. Breast Cancer Therapy-Classical Therapy, Drug Targets, and Targeted Therapy. In: Barh D, editor. Omics Approaches in Breast Cancer. New Delhi, India: Springer; 2014.

37. Ninck S, Reisser C, Dyckhoff G, Helmke B, Bauer H, Herold-Mende C. Expression profiles of angiogenic growth factors in squamous cell carcinomas of the head and neck. Int J Cancer. 2003;106(1):34-44. doi:10.1002/ijc.11188.

38. Ferrara N, Hillan KJ, Gerber H-P, Novotny W. Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. Nat Rev Drug Discov. 2004;3(5):391-400.

39. U.S. Food and Drug Administration. FDA Commissioner Removes Breast Cancer Indication from Avastin Label. U.S. Food and Drug Administration, Silver Spring, MD, USA. 2011. http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm280536.htm. Accessed June 20 2016.

40. Vara JÁF, Casado E, de Castro J, Cejas P, Belda-Iniesta C, González-Barón M. PI3K/Akt signalling pathway and cancer. Cancer Treat Rev. 2004;30(2):193-204. doi:http://dx.doi.org/10.1016/j.ctrv.2003.07.007.

41. Houghton PJ. Everolimus. Clin Cancer Res. 2010;16(5):1368-72. doi:10.1158/1078-0432.CCR-09-1314.

42. Baselga J, Campone M, Piccart M, Burris HA, Rugo HS, Sahmoud T et al. Everolimus in Postmenopausal Hormone-Receptor–Positive Advanced Breast Cancer. N Engl J Med. 2012;366(6):520-9. doi:doi:10.1056/NEJMoa1109653.

43. Ejlertsen B, Heinrich G, Jerusalem M, Hurvitz SA, De Boer RH, Taran T et al. BOLERO-6: Phase II study of everolimus plus exemestane versus everolimus or capecitabine monotherapy in HR+, HER2- advanced breast cancer. J Clin Oncol; 2013.

44. André F, O'Regan R, Ozguroglu M, Toi M, Xu B, Jerusalem G et al. Everolimus for women with trastuzumab-resistant, HER2-positive, advanced breast cancer (BOLERO-3): a randomised,

double-blind, placebo-controlled phase 3 trial. Lancet Oncol. 2014;15(6):580-91. doi:http://dx.doi.org/10.1016/S1470-2045(14)70138-X.

45. Kuerer HM, Newman LA, Smith TL, Ames FC, Hunt KK, Dhingra K et al. Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy. J Clin Oncol. 1999;17(2):460-9.

46. von Minckwitz G, Untch M, Blohmer J-U, Costa SD, Eidtmann H, Fasching PA et al. Definition and Impact of Pathologic Complete Response on Prognosis After Neoadjuvant Chemotherapy in Various Intrinsic Breast Cancer Subtypes. J Clin Oncol. 2012;30(15):1796-804. doi:10.1200/jco.2011.38.8595.

47. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47. doi:10.1016/j.ejca.2008.10.026.

48. Jensen MM, Jørgensen JT, Binderup T, Kjær A. Tumor volume in subcutaneous mouse xenografts measured by microCT is more accurate and reproducible than determined by 18F-FDG-microPET or external caliper. BMC Med Imaging. 2008;8(1):1-9. doi:10.1186/1471-2342-8-16.

49. Amundadottir LT, Merlino G, Dickson RB. Transgenic mouse models of breast cancer. Breast Cancer Res Treat. 1996;39(1):119-35. doi:10.1007/bf01806083.

50. Siolas D, Hannon GJ. Patient-Derived Tumor Xenografts: Transforming Clinical Samples into Mouse Models. Cancer Res. 2013;73(17):5315-9. doi:10.1158/0008-5472.can-13-1069.

51. DeRose YS, Wang G, Lin YC, Bernard PS, Buys SS, Ebbert MT et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nat Med. 2011;17(11):1514-20. doi:10.1038/nm.2454.

52. Zhao X, Liu Z, Yu L, Zhang Y, Baxter P, Voicu H et al. Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. Neuro Oncol. 2012;14(5):574-83. doi:10.1093/neuonc/nos061.

53. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nat Rev Cancer. 2011;11(2):85-95.

54. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science (New York, NY). 2009;324(5930):1029-33. doi:10.1126/science.1160809.

55. Chen J-Q, Russo J. Dysregulation of glucose transport, glycolysis, TCA cycle and glutaminolysis by oncogenes and tumor suppressors in cancer cells. BBA-Rev Cancer. 2012;1826(2):370-84. doi:http://dx.doi.org/10.1016/j.bbcan.2012.06.004.

56. Gottlieb E. p53 guards the metabolic pathway less travelled. Nat Cell Biol. 2011;13(3):195-7.

57. Brahimi-Horn MC, Bellot G, Pouyssegur J. Hypoxia and energetic tumour metabolism. Curr Opin Genet Dev. 2011;21(1):67-72. doi:10.1016/j.gde.2010.10.006.

58. Koukourakis MI, Giatromanolaki A, Sivridis E. Colorectal Cancer: Lactate Dehydrogenase (LDH) activity as a Prognostic Marker. In: Hayat MA, editor. Methods of Cancer Diagnosis, Therapy, and Prognosis. Union, NJ, USA: Springer Science+Business Media B.V; 2009.

59. Warburg O. On the origin of cancer cells. Science. 1956;123(3191):309-14.

60. The metabolism of tumours: Investigations from the kaiser wilhelm institute for biology, berlin-dahlem. Journal of the American Medical Association. 1931;96(23):1982-. doi:10.1001/jama.1931.02720490062043.

61. Patra KC, Hay N. The pentose phosphate pathway and cancer. Trends Biochem Sci. 2014;39(8):347-54. doi:http://dx.doi.org/10.1016/j.tibs.2014.06.005.

62. Wise DR, Thompson CB. Glutamine addiction: a new therapeutic target in cancer. Trends Biochem Sci. 2010;35(8):427-33. doi:10.1016/j.tibs.2010.05.003.

63. Lukey MJ, Wilson KF, Cerione RA. Therapeutic strategies impacting cancer cell glutamine metabolism. Future Med Chem. 2013;5(14):1685-700. doi:10.4155/fmc.13.130.

64. Hensley CT, Wasti AT, DeBerardinis RJ. Glutamine and cancer: cell biology, physiology, and clinical opportunities. J Clin Invest. 2013;123(9):3678-84. doi:10.1172/jci69600.

65. Sitter B, Bathen TF, Singstad TE, Fjosne HE, Lundgren S, Halgunset J et al. Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy. NMR Biomed. 2010;23.

66. Sitter B, Lundgren S, Bathen TF, Halgunset J, Fjosne HE, Gribbestad IS. Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters. NMR Biomed. 2006;19. doi:10.1002/nbm.992.

67. Franco R, Cidlowski JA. Apoptosis and glutathione: beyond an antioxidant. Cell Death Differ. 2009;16(10):1303-14. doi:10.1038/cdd.2009.107.

68. Fukuda R, Zhang H, Kim JW, Shimoda L, Dang CV, Semenza GL. HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. Cell. 2007;129(1):111-22. doi:10.1016/j.cell.2007.01.047.

69. Moestue S, Sitter B, Bathen TF, Tessem MB, Gribbestad IS. HR MAS MR spectroscopy in metabolic characterization of cancer. Curr Top Med Chem. 2011;11(1):2-26.

70. Opstad KS, Bell BA, Griffiths JR, Howe FA. Taurine: a potential marker of apoptosis in gliomas. Br J Cancer, BJC. 2009;100(5):789-94.

71. Swanson MG, Vigneron DB, Tabatabai ZL, Males RG, Schmitt L, Carroll PR et al. Proton HR-MAS spectroscopy and quantitative pathologic analysis of MRI/3D-MRSI-targeted postsurgical prostate tissues. Magn Reson Med. 2003;50(5):944-54. doi:10.1002/mrm.10614.

72. Chan ECY, Koh PK, Mal M, Cheah PY, Eu KW, Backshall A et al. Metabolic Profiling of Human Colorectal Cancer Using High-Resolution Magic Angle Spinning Nuclear Magnetic Resonance (HR-MAS NMR) Spectroscopy and Gas Chromatography Mass Spectrometry (GC/MS). J Proteome Res. 2009;8(1):352-61. doi:10.1021/pr8006232.

73. Righi V, Durante C, Cocchi M, Calabrese C, Di Febo G, Lecce F et al. Discrimination of Healthy and Neoplastic Human Colon Tissues by ex Vivo HR-MAS NMR Spectroscopy and Chemometric Analyses. J Proteome Res. 2009;8(4):1859-69. doi:10.1021/pr801094b.

74. Giskeødegård GF, Grinde MT, Sitter B, Axelson DE, Lundgren S, Fjosne HE et al. Multivariate modeling and prediction of breast cancer prognostic factors using MR metabolomics. J Proteome Res. 2010;9(2):972-9. doi:10.1021/pr9008783.

75. Ridgway ND. The role of phosphatidylcholine and choline metabolites to cell proliferation and survival. Crit Rev Biochem Mol Biol. 2013;48(1):20-38. doi:10.3109/10409238.2012.735643.

76. Katz-Brull R, Degani H. Kinetics of choline transport and phosphorylation in human breast cancer cells; NMR application of the zero trans method. Anticancer Res. 1996;16(3b):1375-80.

77. Iorio E, Mezzanzanica D, Alberti P, Spadaro F, Ramoni C, D'Ascenzo S et al. Alterations of Choline Phospholipid Metabolism in Ovarian Tumor Progression. Cancer Res. 2005;65(20):9369-76. doi:10.1158/0008-5472.can-05-1146.

78. Ramírez de Molina A, Rodríguez-González An, Gutiérrez R, Martínez-Piñeiro L, Sánchez JJ, Bonilla F et al. Overexpression of choline kinase is a frequent feature in human tumor-derived cell lines and in lung, prostate, and colorectal human cancers. Biochem Biophys Res Commun. 2002;296(3):580-3. doi:http://dx.doi.org/10.1016/S0006-291X(02)00920-8.

79. Glunde K, Jie C, Bhujwalla ZM. Molecular Causes of the Aberrant Choline Phospholipid Metabolism in Breast Cancer. Cancer Res. 2004;64(12):4270-6. doi:10.1158/0008-5472.can-03-3829.

80. Noh DY, Ahn SJ, Lee RA, Park IA, Kim JH, Suh PG et al. Overexpression of phospholipase D1 in human breast cancer tissues. Cancer Lett. 2000;161(2):207-14.

81. Cao MD, Dopkens M, Krishnamachary B, Vesuna F, Gadiya MM, Lonning PE et al. Glycerophosphodiester phosphodiesterase domain containing 5 (GDPD5) expression correlates with malignant choline phospholipid metabolite profiles in human breast cancer. NMR Biomed. 2012;25(9):1033-42. doi:10.1002/nbm.2766.

82. Stewart JD, Marchan R, Lesjak MS, Lambert J, Hergenroeder R, Ellis JK et al. Choline-releasing glycerophosphodiesterase EDI3 drives tumor cell migration and metastasis. Proc Natl Acad Sci USA. 2012;109(21):8155-60. doi:10.1073/pnas.1117654109.

83. Aboagye EO, Bhujwalla ZM. Malignant transformation alters membrane choline phospholipid metabolism of human mammary epithelial cells. Cancer Res. 1999;59(1):80-4.

84. Moestue SA, Borgan E, Huuse EM, Lindholm EM, Sitter B, Borresen-Dale AL et al. Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models. BMC Cancer. 2010;10:433. doi:10.1186/1471-2407-10-433.

85. Galons JP, Job C, Gillies RJ. Increase of GPC levels in cultured mammalian cells during acidosis. A 31P MR spectroscopy study using a continuous bioreactor system. Magn Reson Med. 1995;33(3):422-6.

86. Jagannathan NR, Kumar M, Seenu V, Coshic O, Dwivedi SN, Julka PK et al. Evaluation of total choline from in-vivo volume localized proton MR spectroscopy and its response to neoadjuvant chemotherapy in locally advanced breast cancer. Br J Cancer, BJC. 2001;84(8):1016-22. doi:10.1054/bjoc.2000.1711.

87. Baek HM, Chen JH, Nalcioglu O, Su MY. Proton MR spectroscopy for monitoring early treatment response of breast cancer to neo-adjuvant chemotherapy. Ann Oncol. 2008;19(5):1022-4. doi:10.1093/annonc/mdn121.

88. Bujak R, Struck-Lewicka W, Markuszewski MJ, Kaliszan R. Metabolomics for laboratory diagnostics. J Pharm Biomed Anal. 2015;113:108-20. doi:http://dx.doi.org/10.1016/j.jpba.2014.12.017.

89. Fiehn O. Metabolomics--the link between genotypes and phenotypes. Plant Mol Biol. 2002;48(1-2):155-71.

90. Bathen TF, Geurts B, Sitter B, Fjøsne HE, Lundgren S, Buydens LM et al. Feasibility of MR Metabolomics for Immediate Analysis of Resection Margins during Breast Cancer Surgery. PLoS ONE. 2013;8(4):e61578. doi:10.1371/journal.pone.0061578.

91. Cao MD, Lamichhane S, Lundgren S, Bofin A, Fjøsne H, Giskeødegård GF et al. Metabolic characterization of triple negative breast cancer. BMC Cancer. 2014;14(1):1-12. doi:10.1186/1471-2407-14-941.

92. Giskeødegård GF, Lundgren S, Sitter B, Fjosne HE, Postma G, Buydens LM et al. Lactate and glycine-potential MR biomarkers of prognosis in estrogen receptor-positive breast cancers. NMR Biomed. 2012;25(11):1271-9. doi:10.1002/nbm.2798.

93. Vande Velde G, Himmelreich U. Magnetic Resonance-Based Cell Imaging Using Contrast Media and Reporter Genes. In: Garrido L, Beckmann N, editors. New Applications of NMR in Drug Discovery and Development. Cambridge, UK: The Royal Society of Chemistry; 2013.

94. Keeler J. Understanding NMR spectroscopy. Second ed. West Sussex, UK: John Wiley & Sons, Ltd; 2010.

95. Sitter B, Bathen TF, Tessem M-B, Gribbestad IS. High resolution magic angle spinning (HR MAS) MR spectroscopy in metabolic characterization of human cancer. Prog Nucl Magn Reson Spectrosc. 2009;54:239-54.

96. Andrew ER. The narrowing of NMR spectra of solids by high-speed specimen rotation and the resolution of chemical shift and spin multiplet structures for solids. Prog Nucl Magn Reson Spectrosc. 1971;8(1):1-39. doi:http://dx.doi.org/10.1016/0079-6565(71)80001-8.

97. Lowe IJ. Free Induction Decays of Rotating Solids. Phys Rev Lett. 1959;2(7):285-7.

98. Sitter B, Sonnewald U, Spraul M, Fjosne HE, Gribbestad IS. High-resolution magic angle spinning MRS of breast cancer tissue. NMR Biomed. 2002;15(5):327-37. doi:10.1002/nbm.775.

99. Eilers PHC. Parametric Time Warping. Anal Chem. 2004;76(2):404-11. doi:10.1021/ac034800e.

100. Savorani F, Tomasi G, Engelsen SB. icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. J Magn Reson. 2010;202(2):190-202. doi:http://dx.doi.org/10.1016/j.jmr.2009.11.012.

101. Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. Anal Chem. 2006;78(13):4281-90. doi:10.1021/ac051632c.

102. Anzanello MJ, Ortiz RS, Limberger R, Mariotti K. Performance of some supervised and unsupervised multivariate techniques for grouping authentic and unauthentic Viagra and Cialis. EJFS. 2014;4(3):83-9. doi:http://dx.doi.org/10.1016/j.ejfs.2014.03.004.

103. Livingstone DJ. Building QSAR Models: A Practical Guide. In: Cronin MTD, Livingstone DJ, editors. Predicting Chemical Toxicity and Fate. Boca Raton, FL, USA: CRC Press; 2004.

104. Wold S, Esbensen K, Geladi P. Principal Component Analysis. Chemometr Intell Lab. 1987;2:37-52.

105. Vandeginste BGM, Massart DL, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J. Handbook of Chemometrics and Qualimetrics: Part B. Amsterdam, The Netherlands: Elsevier Science B.V.; 1998.

106. Yim O, Ramdeen KT. Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. Tutor Quant Methods Psychol. 2015;11(1):8-21.

107. Mahalanobis PC, editor. On the generalised distance in statistics. Proceedings National Institute of Science, India; 1936.

108. Craw S. Manhattan Distance. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning. Boston, MA: Springer US; 2010. p. 639-.

109. Cormack RM. A Review of Classification. J Roy Stat Soc A Sta. 1971;134(3):321-67. doi:citeulike-article-id:3039214 doi: 10.2307/2344237.

110. Ward JH. Hierarchical Grouping to Optimize an Objective Function. J Am Stat Assoc. 1963;58(301):236-44. doi:10.1080/01621459.1963.10500845.

111. Schafer W, Lin Z. The Laboratory Use of Computers. In: Cazes J, editor. Ewing's Analytical Instrumentation Handbook. Third ed. Boca Raton, FL, USA: Taylor & Francis Group; 2004.

112. Jackson JE. Linear Models I: Regression; PCA of Predictor Variables. A User's Guide to Principal Components. John Wiley & Sons, Inc.; 2004. p. 263-300.

113. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometr Intell Lab Syst. 2001;58(2):109-30. doi:http://dx.doi.org/10.1016/S0169-7439(01)00155-1.

114. Hill T, Lewicki P. Partial Least Squares (PLS). Statistics Methods and Applications, A Comprehensive Reference for Science Industry and Data Mining. First ed. Tulsa, OK, USA: StatSoft, Inc.; 2006.

115. Wold H. Estimation of Principal Components and Related Models by Iterative Least squares. Multivariate Analysis.: Academic Press; 1966. p. 391-420.

116. Wold H. Path Models with latent variables: The NIPALS approach. Academic Press; 1975.

117. de Jong S. SIMPLS: An alternative approach to partial least squares regression. Chemometr Intell Lab Syst. 1993;18(3):251-63. doi:http://dx.doi.org/10.1016/0169-7439(93)85002-X.

118. Andersson M. A comparison of nine PLS1 algorithms. J Chemometrics. 2009;23(10):518-29. doi:10.1002/cem.1248.

119. Alin A. Comparison of PLS algorithms when number of objects is much larger than number of variables. Stat Pap. 2009;50(4):711-20. doi:10.1007/s00362-009-0251-7.

120. Szymańska E, Saccenti E, Smilde AK, Westerhuis JA. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. Metabolomics. 2012;8(Suppl 1):3-16. doi:10.1007/s11306-011-0330-3.

121. Barker M, Rayens W. Partial least squares for discrimination. J Chemometrics. 2003;17(3):166-73. doi:10.1002/cem.785.

122. Fisher RA. The Use of Multiple Measurements in Taxonomic Problems. A Eug. 1936;7(2):179-88. doi:10.1111/j.1469-1809.1936.tb02137.x.

123. Hawkins DM. The Problem of Overfitting. J Chem Inf Comput Sci. 2004;44(1):1-12. doi:10.1021/ci0342472.

124. Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, van Velzen EJJ et al. Assessment of PLSDA cross validation. Metabolomics. 2008;4(1):81-9. doi:10.1007/s11306-007-0099-6.

125. Eigenvector Research Inc. Using Cross-Validation. http://wiki.eigenvector.com/index.php?title=Using_Cross-Validation. Accessed June 27 2016.

126. Anderssen E, Dyrstad K, Westad F, Martens H. Reducing over-optimism in variable selection by cross-model validation. Chemometr Intell Lab Syst. 2006;84(1–2):69-74. doi:http://dx.doi.org/10.1016/j.chemolab.2006.04.021.

127. Swanson MG, Zektzer AS, Tabatabai ZL, Simko J, Jarso S, Keshari KR et al. Quantitative analysis of prostate metabolites using 1H HR-MAS spectroscopy. Magn Reson Med. 2006;55(6):1257-64. doi:10.1002/mrm.20909.

128. Akoka S, Barantin L, Trierweiler M. Concentration Measurement by Proton NMR Using the ERETIC Method. Anal Chem. 1999;71(13):2554-7. doi:10.1021/ac981422i.

129. Albers MJ, Butler TN, Rahwa I, Bao N, Keshari KR, Swanson MG et al. Evaluation of the ERETIC method as an improved quantitative reference for 1H HR-MAS spectroscopy of prostate tissue. Magn Reson Med. 2009;61(3):525-32. doi:10.1002/mrm.21808.

130. Vinaixa M, Samino S, Saez I, Duran J, Guinovart JJ, Yanes O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. Metabolites. 2012;2(4):775-95. doi:10.3390/metabo2040775.

131. Ghasemi A, Zahediasl S. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. Int J Endocrinol Metab. 2012;10(2):486-9. doi:10.5812/ijem.3505.

132. Pinheiro JC, Bates DM. Linear Mixed-Effects Models: Basic Concepts and Examples. Mixed-Effects Models in S and S-PLUS. Statistics and Computing. New York, NY, USA: Springer New York; 2000. p. 3-56.

133. Winter B. A very basic tutorial for performing linear mixed effects analyses (Tutorial 2). Merced, CA, USA. 2013. http://www.bodowinter.com/tutorial/bw_LME_tutorial2.pdf.

134. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze. 1936;8:3-62. doi:citeulike-article-id:1778138.

135. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met. 1995;57(1):289-300.

136. Benjamini Y, Yekutieli D. The Control of the False Discovery Rate in Multiple Testing under Dependency. Ann Stat. 2001;29(4):1165-88.

137. Tarca AL, Romero R, Draghici S. Analysis of microarray experiments of gene expression profiling. Am J Obstet Gynecol. 2006;195(2):373-88. doi:10.1016/j.ajog.2006.07.001.

138. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T et al. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. J Clin Oncol. 2009;27(8):1160-7. doi:10.1200/jco.2008.18.1370.

139. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci USA. 2002;99(10):6567-72. doi:10.1073/pnas.082099299.

140. Prasad TV. Gene Expression Data Analysis Suite (GEDAS). 2013. http://gedas.bizhat.com/dist.htm. Accessed June 28 2016.

141. Tusher VG, Tibshirani R, Chu G. Significance Analysis of Microarrays Applied to the Ionizing Radiation Response. Proc Natl Acad Sci USA. 2001;98(9):5116-21.

142. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545-50. doi:10.1073/pnas.0506580102.

143. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9. doi:10.1038/75556.

144. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57. doi:10.1038/nprot.2008.211.

145. David Bioinformatic Team. Functional Annotation Tool. Laboratory of Human Retrovirology and Immunoinformatics, Frederick, MD, USA. 2003. https://david.ncifcrf.gov/content.jsp?file=functional_annotation.html#E4. Accessed June 28 2016.

146. Xia J, Sinelnikov IV, Han B, Wishart DS. MetaboAnalyst 3.0—making metabolomics more meaningful. Nucleic Acids Res. 2015. doi:10.1093/nar/gkv380.

147. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30.

148. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44(D1):D457-62. doi:10.1093/nar/gkv1070.

149. Xia J, Wishart DS. MetPA: a web-based metabolomics tool for pathway analysis and visualization. Bioinformatics. 2010;26(18):2342-4. doi:10.1093/bioinformatics/btq418.

150. Marangoni E, Vincent-Salomon A, Auger N, Degeorges A, Assayag F, de Cremoux P et al. A New Model of Patient Tumor-Derived Breast Cancer Xenografts for Preclinical Assays. Clin Cancer Res. 2007;13(13):3989.

151. van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate Imputation by Chained Equations in R. J Stat Softw. 2011;45(3). doi:10.18637/jss.v045.i03.

152. R Core Team (2014). R: A language and environment for statistical computing. Vienna, Austria. http://www.R-project.org/: R Foundation for Statistical Computing.

153. Pinheiro JC, Bates DM, DebRoy S, Sarkar D, R Core Team (2014). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-117. http://CRAN.R-project.org/package=nlme.

154. Hatem R, Botty RE, Chateau-Joubert S, Servely J-L, Labiod D, Plater Ld et al. Targeting mTOR pathway inhibits tumor growth in different molecular subtypes of triple-negative breast cancers. Oncotarget. 2016. doi:10.18632/oncotarget.10195.

155. Monti S, Tamayo P, Mesirov J, Golub T. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. Mach Learn. 2003;52(1):91-118. doi:10.1023/a:1023949509487.

156. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med. 1998;4(7):844-7.

157. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14(5):365-76. doi:10.1038/nrn3475.

158. Leon AC, Davis LL, Kraemer HC. The Role and Interpretation of Pilot Studies in Clinical Research. J Psychiat Res. 2011;45(5):626-9. doi:10.1016/j.jpsychires.2010.10.008.

159. Cao MD, Sitter B, Bathen TF, Bofin A, Lonning PE, Lundgren S et al. Predicting long-term survival and treatment response in breast cancer patients receiving neoadjuvant chemotherapy by MR metabolic profiling. NMR Biomed. 2011;25. doi:10.1002/nbm.1762.

160. Cao MD, Giskeødegård GF, Bathen TF, Sitter B, Bofin A, Lonning PE et al. Prognostic value of metabolic response in breast cancer patients receiving neoadjuvant chemotherapy. BMC Cancer. 2012;12. doi:10.1186/1471-2407-12-39.

161. van der Staay FJ. Animal models of behavioral dysfunctions: Basic concepts and classifications, and an evaluation strategy. Brain Res Rev. 2006;52(1):131-59. doi:http://dx.doi.org/10.1016/j.brainresrev.2006.01.006.

162. Mueller MM, Fusenig NE. Friends or foes - bipolar effects of the tumour stroma in cancer. Nat Rev Cancer. 2004;4(11):839-49.

163. Ruggeri BA, Camp F, Miknyoczki S. Animal models of disease: Pre-clinical animal models of cancer and their applications and utility in drug discovery. Biochem Pharmacol. 2014;87(1):150-61. doi:http://dx.doi.org/10.1016/j.bcp.2013.06.020.

164. Haukaas TH, Moestue SA, Vettukattil R, Sitter B, Lamichhane S, Segura R et al. Impact of Freezing Delay Time on Tissue Samples for Metabolomic Studies. Front Oncol. 2016;6:17. doi:10.3389/fonc.2016.00017.

165. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet.384(9938):164-72. doi:10.1016/S0140-6736(13)62422-8.

166. Loibl S, Volz C, Mau C, Blohmer JU, Costa SD, Eidtmann H et al. Response and prognosis after neoadjuvant chemotherapy in 1,051 patients with infiltrating lobular breast carcinoma. Breast Cancer Res Treat. 2014;144(1):153-62. doi:10.1007/s10549-014-2861-6.

167. Earl H, Provenzano E, Abraham J, Dunn J, Vallier A-L, Gounaris I et al. Neoadjuvant trials in early breast cancer: pathological response at surgery and correlation to longer term outcomes – what does it all mean? BMC Med. 2015;13(1):1-11. doi:10.1186/s12916-015-0472-7.

168. Teruel JR, Heldahl MG, Goa PE, Pickles M, Lundgren S, Bathen TF et al. Dynamic contrast-enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. NMR Biomed. 2014;27(8):887-96. doi:10.1002/nbm.3132.

169. Wu J. Statistical inference for tumor growth inhibition T/C ratio. J Biopharm Stat. 2010;20(5):954-64. doi:10.1080/10543401003618983.

170. Sanceau J, Poupon MF, Delattre O, Sastre-Garau X, Wietzerbin J. Strong inhibition of Ewing tumor xenograft growth by combination of human interferon-alpha or interferon-beta with ifosfamide. Oncogene. 2002;21(50):7700-9. doi:10.1038/sj.onc.1205881.

171. Tsukihara H, Nakagawa F, Sakamoto K, Ishida K, Tanaka N, Okabe H et al. Efficacy of combination chemotherapy using a novel oral chemotherapeutic agent, TAS-102, together with bevacizumab, cetuximab, or panitumumab on human colorectal cancer xenografts. Oncol Rep. 2015;33(5):2135-42. doi:10.3892/or.2015.3876.

172. Fang M, Ivanisevic J, Benton HP, Johnson CH, Patti GJ, Hoang LT et al. Thermal Degradation of Small Molecules: A Global Metabolomic Investigation. Anal Chem. 2015;87(21):10935-41. doi:10.1021/acs.analchem.5b03003.

173. Rocha CM, Barros AS, Gil AM, Goodfellow BJ, Humpfer E, Spraul M et al. Metabolic Profiling of Human Lung Cancer Tissue by 1H High Resolution Magic Angle Spinning (HRMAS) NMR Spectroscopy. J Proteome Res. 2010;9(1):319-32. doi:10.1021/pr9006574.

174. Wang Y-Y, Lehuédé C, Laurent V, Dirat B, Dauvillier S, Bochet L et al. Adipose tissue and breast epithelial cells: A dangerous dynamic duo in breast cancer. Cancer Lett. 2012;324(2):142-51. doi:http://dx.doi.org/10.1016/j.canlet.2012.05.019.

175. Wright AJ, Fellows GA, Griffiths JR, Wilson M, Bell BA, Howe FA. Ex-vivo HRMAS of adult brain tumours: metabolite quantification and assignment of tumour biomarkers. Mol Cancer. 2010;9:66-. doi:10.1186/1476-4598-9-66.

176. Emir UE, Deelchand D, Henry P-G, Terpstra M. Noninvasive quantification of T2 and concentrations of ascorbate and glutathione in the human brain from the same double-edited spectra. NMR Biomed. 2011;24(3):263-9. doi:10.1002/nbm.1583.

177. Giskeødegård GF, Bertilsson H, Selnæs KM, Wright AJ, Bathen TF, Viset T et al. Spermine and Citrate as Metabolic Biomarkers for Assessing Prostate Cancer Aggressiveness. PLoS ONE. 2013;8(4):e62375. doi:10.1371/journal.pone.0062375.

178. Hansen AF, Sandsmark E, Rye MB, Wright AJ, Bertilsson H, Richardsen E et al. Presence of TMPRSS2-ERG is associated with alterations of the metabolic profile in human prostate cancer. Oncotarget. 2016. doi:10.18632/oncotarget.9817.

179. Provencher SW. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. Magn Reson Med. 1993;30(6):672-9.

180. Chiu C-C, Chan S-Y, Wang C-C, Wu W-S. Missing value imputation for microarray data: a comprehensive comparison study and a web tool. BMC Syst Biol. 2013;7(Suppl 6):S12-S. doi:10.1186/1752-0509-7-S6-S12.

181. de Brevern AG, Hazout S, Malpertuy A. Influence of microarrays experiments missing values on the stability of gene groups by hierarchical clustering. BMC Bioinform. 2004;5:114-. doi:10.1186/1471-2105-5-114.

182. Provencher S. LCModel & LCMgui User's Manual. May 2016. http://s-provencher.com/pub/LCModel/manual/manual.pdf. Accessed September 9 2016.

183. van Velzen EJJ, Westerhuis JA, van Duynhoven JPM, van Dorsten FA, Hoefsloot HCJ, Jacobs DM et al. Multilevel Data Analysis of a Crossover Designed Human Nutritional Intervention Study. J Proteome Res. 2008;7(10):4483-91. doi:10.1021/pr800145j.

184. Rabinowitz GB. An Introduction to Nonmetric Multidimensional Scaling. Am J Polit Sci. 1975;19(2):343-90. doi:10.2307/2110441.

185. Bertilsson H, Tessem MB, Flatberg A, Viset T, Gribbestad I, Angelsen A et al. Changes in gene transcription underlying the aberrant citrate and choline metabolism in human prostate cancer samples. Clin Cancer Res. 2012;18. doi:10.1158/1078-0432.ccr-11-2929.

186. Moestue SA, Dam CG, Gorad SS, Kristian A, Bofin A, Maelandsmo GM et al. Metabolic biomarkers for response to PI3K inhibition in basal-like breast cancer. Breast Cancer Res. 2013;15(1):R16. doi:10.1186/bcr3391.

187. Contractor T, Harris CR. p53 negatively regulates transcription of the pyruvate dehydrogenase kinase Pdk2. Cancer Res. 2012;72(2):560-7. doi:10.1158/0008-5472.can-11-1215.

188. Walenta S, Wetterling M, Lehrke M, Schwickert G, Sundfor K, Rofstad EK et al. High lactate levels predict likelihood of metastases, tumor recurrence, and restricted patient survival in human cervical cancers. Cancer Res. 2000;60(4):916-21.

189. Walenta S, Mueller-Klieser WF. Lactate: mirror and motor of tumor malignancy. Semin Radiat Oncol. 2004;14(3):267-74. doi:10.1016/j.semradonc.2004.04.004.

190. Saraswathy S, Crawford FW, Lamborn KR, Pirzkall A, Chang S, Cha S et al. Evaluation of MR markers that predict survival in patients with newly diagnosed GBM prior to adjuvant therapy. J Neurooncol. 2009;91(1):69-81. doi:10.1007/s11060-008-9685-3.

191. Semenza GL. Tumor metabolism: cancer cells give and take lactate. J Clin Invest. 2008;118(12):3835-7. doi:10.1172/jci37373.

192. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiother Oncol. 2009;92(3):329-33. doi:10.1016/j.radonc.2009.06.025.

193. Xie J, Wu H, Dai C, Pan Q, Ding Z, Hu D et al. Beyond Warburg effect – dual metabolic nature of cancer cells. Sci Rep. 2014;4:4927. doi:10.1038/srep04927 http://www.nature.com/articles/srep04927#supplementary-information.

194. Fack F, Espedal H, Keunen O, Golebiewska A, Obad N, Harter PN et al. Bevacizumab treatment induces metabolic adaptation toward anaerobic metabolism in glioblastomas. Acta Neuropathol. 2015;129(1):115-31. doi:10.1007/s00401-014-1352-5.

195. Foster R, Griffin S, Grooby S, Feltell R, Christopherson C, Chang M et al. Multiple Metabolic Alterations Exist in Mutant PI3K Cancers, but Only Glucose Is Essential as a Nutrient Source. PLoS ONE. 2012;7(9):e45061. doi:10.1371/journal.pone.0045061.

196. Grinde MT, Skrbo N, Moestue SA, Rødland EA, Borgan E, Kristian A et al. Interplay of choline metabolites and genes in patient-derived breast cancer xenografts. Breast Cancer Res. 2014;16(1):1-16. doi:10.1186/bcr3597.

197. Fagone P, Jackowski S. Phosphatidylcholine and the CDP–choline cycle. BBA-Mol Cell Biol L. 2013;1831(3):523-32. doi:http://dx.doi.org/10.1016/j.bbalip.2012.09.009.

198. Wyss M, Kaddurah-Daouk R. Creatine and Creatinine Metabolism. Physiol Rev. 2000;80(3):1107.

199. Ide T, Chu K, Aaronson SA, Lee SW. GAMT joins the p53 network: Branching into metabolism. Cell Cycle. 2010;9(9):1706-10. doi:10.4161/cc.9.9.11473.

200. King A, Selak MA, Gottlieb E. Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. Oncogene. 2006;25(34):4675-82.

201. Selak MA, Armour SM, MacKenzie ED, Boulahbel H, Watson DG, Mansfield KD et al. Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase. Cancer Cell. 2005;7(1):77-85. doi:http://dx.doi.org/10.1016/j.ccr.2004.11.022.

202. Huang C, Freter C. Lipid Metabolism, Apoptosis and Cancer Therapy. Int J Mol Sci. 2015;16(1):924.

203. Sun W, Yang J. Functional Mechanisms for Human Tumor Suppressors. J Cancer. 2010;1:136-40.

204. Cadwell C, Zambetti GP. The effects of wild-type p53 tumor suppressor activity and mutant p53 gain-of-function on cell growth. Gene. 2001;277(1–2):15-30. doi:http://dx.doi.org/10.1016/S0378-1119(01)00696-5.

205. Harami-Papp H, Pongor LS, Munkacsy G, Horvath G, Nagy AM, Ambrus A et al. TP53 mutation hits energy metabolism and increases glycolysis in breast cancer. Oncotarget. 2016. doi:10.18632/oncotarget.11594.

206. Bergamaschi A, Tagliabue E, Sorlie T, Naume B, Triulzi T, Orlandi R et al. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. J Pathol. 2008;214(3):357-67. doi:10.1002/path.2278.

207. Vegran F, Boidot R, Michiels C, Sonveaux P, Feron O. Lactate influx through the endothelial cell monocarboxylate transporter MCT1 supports an NF-kappaB/IL-8 pathway that drives tumor angiogenesis. Cancer Res. 2011;71(7):2550-60. doi:10.1158/0008-5472.can-10-2828.

208. Gyanchandani R, Sano D, Ortega Alves MV, Klein JD, Knapick BA, Oh S et al. Interleukin-8 as a modulator of response to bevacizumab in preclinical models of head and neck squamous cell carcinoma. Oral Oncol. 2013;49(8):761-70. doi:10.1016/j.oraloncology.2013.03.452.

209. Glunde K, Jiang L, Moestue SA, Gribbestad IS. MRS and MRSI guidance in molecular medicine: targeting and monitoring of choline and glucose metabolism in cancer. NMR Biomed. 2011;24(6):673-90. doi:10.1002/nbm.1751.

210. Zhao Y, Butler EB, Tan M. Targeting cellular metabolism to improve cancer therapeutics. Cell Death Dis. 2013;4:e532.

211. Bear HD, Tang G, Rastogi P, Geyer CE, Jr., Liu Q, Robidoux A et al. Neoadjuvant plus adjuvant bevacizumab in early breast cancer (NSABP B-40 [NRG Oncology]): secondary outcomes of a phase 3, randomised controlled trial. Lancet Oncol.16(9):1037-48. doi:10.1016/S1470-2045(15)00041-8.

212. Ascierto PA, Marincola FM. What have we learned from cancer immunotherapy in the last 3 years? J Transl Med. 2014;12(1):1-11. doi:10.1186/1479-5876-12-141.

213. Juweid ME, Cheson BD. Positron-Emission Tomography and Assessment of Cancer Therapy. N Engl J Med. 2006;354(5):496-507. doi:doi:10.1056/NEJMra050276.

214. Gutte H, Hansen AE, Johannesen HH, Clemmensen AE, Ardenkjaer-Larsen JH, Nielsen CH et al. The use of dynamic nuclear polarization (13)C-pyruvate MRS in cancer. Am J Nucl Med Mol Imaging. 2015;5(5):548-60.

215. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? Nat Rev Cancer. 2004;4(11):891-9. doi:10.1038/nrc1478.

216. Stokkel MP, Draisma A, Pauwels EK. Positron emission tomography with 2-[18F]-fluoro-2-deoxy-D-glucose in oncology. Part IIIb: Therapy response monitoring in colorectal and lung tumours, head and neck cancer, hepatocellular carcinoma and sarcoma. J Cancer Res Clin Oncol. 2001;127(5):278-85.

217. Rousseau C, Devillers A, Sagan C, Ferrer L, Bridji B, Campion L et al. Monitoring of early response to neoadjuvant chemotherapy in stage II and III breast cancer by [18F]fluorodeoxyglucose positron emission tomography. J Clin Oncol. 2006;24(34):5366-72. doi:10.1200/jco.2006.05.7406.

218. Danishad KK, Sharma U, Sah RG, Seenu V, Parshad R, Jagannathan NR. Assessment of therapeutic response of locally advanced breast cancer (LABC) patients undergoing neoadjuvant chemotherapy (NACT) monitored using sequential magnetic resonance spectroscopic imaging (MRSI). NMR Biomed. 2010;23(3):233-41. doi:10.1002/nbm.1436.

219. Dzien P, Fages A, Jona G, Brindle KM, Schwaiger M, Frydman L. Following Metabolism in Living Microorganisms by Hyperpolarized 1H NMR. J Am Chem Soc. 2016. doi:10.1021/jacs.6b07483.

220. Day SE, Kettunen MI, Cherukuri MK, Mitchell JB, Lizak MJ, Morris HD et al. Detecting response of rat C6 glioma tumors to radiotherapy using hyperpolarized [1-13C]pyruvate and 13C magnetic resonance spectroscopic imaging. Magn Reson Med. 2011;65(2):557-63. doi:10.1002/mrm.22698.

221. Nelson SJ, Kurhanewicz J, Vigneron DB, Larson PEZ, Harzstark AL, Ferrone M et al. Metabolic Imaging of Patients with Prostate Cancer Using Hyperpolarized [1-^{13}C]Pyruvate. Sci Transl Med. 2013;5(198):198ra08.

Paper I

Preprocessing of NMR metabolomics data

Leslie R. Euceda, Guro F. Giskeødegård & Tone F. Bathen

To cite this article: Leslie R. Euceda, Guro F. Giskeødegård & Tone F. Bathen (2015) Preprocessing of NMR metabolomics data, Scandinavian Journal of Clinical and Laboratory Investigation, 75:3, 193-203, DOI: 10.3109/00365513.2014.1003593

To link to this article: http://dx.doi.org/10.3109/00365513.2014.1003593

Is not included due to copyright

Paper II

Haukaas et al. Cancer & Metabolism (2016) 4:12
DOI 10.1186/s40170-016-0152-x

Cancer & Metabolism

RESEARCH Open Access

Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes

Tonje H. Haukaas[1,2], Leslie R. Euceda[1], Guro F. Giskeødegård[1,3], Santosh Lamichhane[1,4], Marit Krohn[2,5], Sandra Jernström[2,5], Miriam R. Aure[2,5], Ole C. Lingjærde[2,6,7], Ellen Schlichting[8], Øystein Garred[9], Eldri U. Due[2,5], Gordon B. Mills[10], Kristine K. Sahlberg[2,11], Anne-Lise Børresen-Dale[2,5*], Tone F. Bathen[1,2*] and The Oslo Breast Cancer Consortium (OSBREAC)

Abstract

Background: The heterogeneous biology of breast cancer leads to high diversity in prognosis and response to treatment, even for patients with similar clinical diagnosis, histology, and stage of disease. Identifying mechanisms contributing to this heterogeneity may reveal new cancer targets or clinically relevant subgroups for treatment stratification. In this study, we have merged metabolite, protein, and gene expression data from breast cancer patients to examine the heterogeneity at a molecular level.

Methods: The study included primary tumor samples from 228 non-treated breast cancer patients. High-resolution magic-angle spinning magnetic resonance spectroscopy (HR MAS MRS) was performed to extract the tumors metabolic profiles further used for hierarchical cluster analysis resulting in three significantly different metabolic clusters (Mc1, Mc2, and Mc3). The clusters were further combined with gene and protein expression data.

Results: Our result revealed distinct differences in the metabolic profile of the three metabolic clusters. Among the most interesting differences, Mc1 had the highest levels of glycerophosphocholine (GPC) and phosphocholine (PCho), Mc2 had the highest levels of glucose, and Mc3 had the highest levels of lactate and alanine. Integrated pathway analysis of metabolite and gene expression data uncovered differences in glycolysis/gluconeogenesis and glycerophospholipid metabolism between the clusters. All three clusters had significant differences in the distribution of protein subtypes classified by the expression of breast cancer-related proteins. Genes related to collagens and extracellular matrix were downregulated in Mc1 and consequently upregulated in Mc2 and Mc3, underpinning the differences in protein subtypes within the metabolic clusters. Genetic subtypes were evenly distributed among the three metabolic clusters and could therefore contribute to additional explanation of breast cancer heterogeneity.

Conclusions: Three naturally occurring metabolic clusters of breast cancer were detected among primary tumors from non-treated breast cancer patients. The clusters expressed differences in breast cancer-related protein as well as genes related to extracellular matrix and metabolic pathways known to be aberrant in cancer. Analyses of metabolic activity combined with gene and protein expression provide new information about the heterogeneity of breast tumors and, importantly, the metabolic differences infer that the clusters may be susceptible to different metabolically targeted drugs.

Keywords: Metabolomics, HR MAS MRS, Breast cancer subgroups, Metabolic cluster, Extracellular matrix

* Correspondence: a.l.borresen-dale@medisin.uio.no; tone.f.bathen@ntnu.no
[2]K.G. Jebsen Center for Breast Cancer Research, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
[1]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
Full list of author information is available at the end of the article

Background

Breast cancer accounts for 25 % of newly diagnosed cancers and 15 % of cancer deaths among women worldwide [1]. It is a heterogeneous disease [2] with high diversity in prognosis and response to treatment. Identification of underlying mechanisms contributing to this heterogeneity may reveal new cancer targets and clinically relevant subgroups and has thus been the focus of many recent studies [3–5].

Searching for genetic features causing the variation in breast cancers, Perou et al. used gene expression analyses followed by hierarchical clustering and defined naturally occurring molecular subtypes [4, 6]. These subtypes are named basal-like, luminal A, luminal B, Erb-B2+ (Her2 enriched), and normal-like, and are found to be associated with tumor characteristics and clinical outcome; patients with basal-like tumors having the shortest and luminal A the longest relapse-free survival [6]. A centroid-based method called prediction analysis of microarrays 50 (PAM50), which uses the expression of 50 genes to classify breast cancer into these five intrinsic subtypes was later established and is now broadly implemented [7].

Proteins are the ultimate cellular effectors of pathways and networks within cells, tissues, and organisms. Although protein levels are dependent on mRNA expression, not all mRNA will be translated into protein and further protein levels are also influenced by protein stability. In a study by Myhre et al. only 22 of 52 quantified breast cancer-related proteins were found to correlate with mRNA expression levels [8] and similar low levels of correlation have been seen in large scale studies [9, 10]. Protein expression subtypes of breast cancer could give further understanding of underlying mechanisms causing heterogeneity [11]. Based on the expression of 171 breast cancer-associated proteins detected by reverse phase protein array (RPPA), six breast cancer subtypes, called RPPA subtypes, have been defined [5]. Four of these subgroups were in high accordance with the gene expression profiles of the PAM50 subtypes and named accordingly; Basal, Her2, luminal A, and luminal A/B. In addition, two new subgroups were defined; reactive I and reactive II, based on expression of proteins possibly produced by the surrounding microenvironment.

The chemical processes controlled by proteins involve metabolites as intermediates or end-products. In metabolomics, metabolite levels are measured to gather the final downstream information of ongoing cellular processes. Which processes are active at a specific time point is strongly influenced by environmental factors like diet and drugs as well as disease state. Well-established metabolic differences have been observed when comparing cancer cells to normal cells. Cancer cell energy production frequently depends on increased glycolysis and production of lactate from glucose regardless of access to oxygen, in contrast to normal cells which produce pyruvate and lactate in aerobic conditions [12]. Also, to produce macromolecules/biomass, mitochondrial metabolism is reprogrammed [13]. Altered metabolism has therefore been included as one of the emerging hallmarks of cancer [14]. In breast cancer, metabolic differences between cancer tissue and normal adjacent tissue have been studied by the magnetic resonance spectroscopy (MRS) method high-resolution magic-angle spinning (HRMAS) MRS [15]. Using this technique, metabolic profiles and biomarkers predicting long-term survival for locally advanced breast cancer [16], node involvement of patients with infiltrating ductal carcinoma [17], and 5-year survival for ER positive patients [18] have been identified.

Merging transcriptomics and metabolomics led to the discovery of three luminal A subgroups with distinct metabolic profiles and significant differences within gene set expression in a study by Borgan et al. [19]. The aim of the current study was to establish clusters of breast cancer based on the metabolic expression using an approach similar to Borgan et al., but in a larger cohort of patients including all PAM50 subgroups. This approach reveals the main metabolic differences between untreated breast tumors. In addition, the combination of the metabolic clusters with transcriptomics and protein expression data provide an opportunity for information gain from each -omics technology, giving further characterization of the defined metabolic clusters.

Methods

Patients and tissue samples

Primary breast carcinoma samples from 228 patients at the Oslo University Hospital (Radium Hospital and Ullevål Hospital) were collected in the time period 2006–2009 as part of the Oslo2 study. The samples were fresh frozen after surgery and stored at −80 °C. The tumors were divided into smaller pieces depending on their size, and one of them was selected for this study. The samples were cut into three sections where the edges of the two outer pieces were used for histological evaluation (including estrogen receptor (ER) status and tumor cell percentage), and an adequate part of the mid pieces were used for HR MAS MRS experiments to obtain metabolic profiles. The remnants of all three pieces were pooled and cut into smaller pieces with scalpel, depending on the size of the tumor and divided into fractions used for extraction of DNA, RNA, and protein. Due to high lipid content, HR MAS MRS was performed on a second piece from the same tumor for 13 of the samples. A total of 228 samples were analyzed by MR spectroscopy, of which 201 and 217 were analyzed for gene expression by arrays and protein expression using

RPPA, respectively, leaving a total of 191 samples analyzed by all three methods. Patient and tumor characteristics are shown in Table 1.

HR MAS MRS spectra

HR MAS MRS spectra were acquired from tissue samples (mean sample weight 7.3 mg ± 2.6 mg) on a Bruker Avance III 600 MHz/54 mm US (Bruker, Biospin GmbH, Germany) equipped with a 1H/13C MAS probe with gradient aligned with the magic angle (Bruker, Biospin GmbH, Germany). Spin-echo spectra were recorded using a Carr-Purcell-Meiboom-Gill (cpmg) pulse sequence (cpmgpr1d; Bruker). For experimental details and information about data processing, see Additional file 1.

Forty-three samples were excluded from the original sample cohort of 271 samples due to large lipid content. The spectral region between 1.40 and 4.70 ppm was chosen for further analysis excluding lipid peaks at 4.36–4.27, 2.88–2.70, 2.30–2.20, 2.09–1.93, and 1.67–1.50 ppm. After removal of the lipid residuals, the spectra were mean normalized by dividing each spectral variable to the average spectral intensity. This is done to account for differences in tumor cell percentage and sample weight, as it can be assumed that most of the lipid signals from breast samples do not originate from cancer cells.

Protein experiments and protein expression subtyping

Protein levels were determined using reverse phase protein array (RPPA), a platform where single protein levels can be measured across a series of samples simultaneously [20]. One hundred fifty primary antibodies were used to detect breast cancer-related proteins (Additional file 2: Table S1). For analytical details, see Additional file 1.

The samples underwent consensus clustering with an option for four or five groups. The best fit on consensus clustering identified five groups, luminal, HER2, basal, and reactive I and II subsets as defined in The Cancer Genome Atlas Network data set [5].

mRNA expression profiling and gene expression subtyping

Total RNA was isolated with TRIzol (Invitrogen, Carlsbad, CA, USA). Expression of mRNA was measured using SurePrint G3 Human GE 8x60K (Aglient Technologies) according to the manufactory's protocol (one-color microarray-based gene expression analysis, low input Quick Amp Labeling, v.6.5, May 2010) and 100 ng RNA was used as input for labeling. Arrays were log2-transformed, quantile normalized, and hospital adjusted [21]. Values corresponding to probes with identical Entrez ID were averaged to form a single expression value per gene.

The PAM50 subtype algorithm [7] was used to assign a subtype label to each sample as previously described [22].

Statistical analysis

Subgrouping with cluster analysis of metabolic data

Hierarchical cluster analysis (HCA) was performed with Euclidean distance as the distance parameter and Ward's method (furthest inner square distance) as the clustering distance (Statistical toolbox, Matlab R2013b, The Mathworks, Inc., USA) on the preprocessed metabolic spectra. Similar spectra based on the distance measures cluster together. The dendrogram was cut to give three clusters. To evaluate the robustness of the three HCA clusters, partial least square discriminant analysis (PLS-DA) model, using the cluster group for classification was carried out and classification accuracy was evaluated. For details, see Additional file 1.

Analysis of metabolic profiles

Metabolite assignments were performed based on literature values [23]. Relative metabolite quantification was performed by peak integration of fixed regions corresponding to the metabolite of interest. In total, the level of 18 metabolites were calculated. Kruskal-Wallis test was performed to compare metabolite levels between clusters. Calculated p values were corrected for multiple testing by The Benjamini Hochberg false discovery rate (FDR) in Matlab, and the differences were considered statistically significant for adjusted $p \leq 0.05$.

Analysis of subtype and clinical distributions

Differences in the distributions of RPPA and PAM50 subtype as well as that of other clinical characteristics of the tumors between the different metabolic clusters were tested for significance using Fisher's exact test for count data (R 2.15.2). Calculated p values were corrected for multiple testing by The Benjamini Hochberg FDR, and the differences were considered statistically significant for adjusted $p < 0.05$.

Analysis of gene expression data

Significance analysis of microarrays (SAM) was used to identify differentially expressed gene between the metabolic clusters [24]. SAM analysis was performed using 21851 genes from 42405 mRNA probes. The expression analysis was performed in R 2.15.2 [25] with the cluster group as the dependent variable and a total of 100 permutations. T statistics/Wilcoxon statistics were calculated using multiclass comparisons and two-class unpaired tests while comparing two clusters. The differences were considered statistically significant for adjusted $p < 0.01$.

DAVID, an online network analysis tool [26], was used to search for biological functions within gene sets. DAVID was performed on the gene list over for each of

Table 1 Patient and tumor characteristics

	Total	Mc1	Mc2	Mc3
Number of patients	228	58	58	112
Age (years), mean (range)	55.5 (31.8–81.1)	58.0 (33.2–80.8)	58.6 (40.9–81.1)	52.7 (31.8–73.9)
Clinical classification				
Histology				
Ductal	186	52	37	97
Lobular	21	4	11	6
Medullary	0	0	0	0
Ductal carsinoma in situ (DCIS)	4	0	4	0
Metaplastic	1	0	1	0
Mucinous	4	0	2	2
Tubular	4	1	1	2
Mixed	2	1	0	1
Papillary	0	0	0	0
NA	6	0	2	4
Primary tumor				
Tx or *NA*	9	1	3	5
T0	0	0	0	0
pTis	4	0	4	0
T1	113	31	28	54
T2	93	24	21	48
T3	9	2	2	5
T4	0	0	0	0
Grade				
I	31	8	10	13
II	93	20	24	49
III	97	30	21	46
NA	7	0	3	4
Node status				
N0	133	34	36	63
N1(mi)	8	3	3	2
N1	59	17	13	29
N2	14	2	3	9
N3	8	2	1	5
NA	6	0	2	4
Receptor status				
HER2+	26	7	7	12
HER2–	192	51	45	96
ER+	178	49	42	87
ER–	40	9	10	21
PR+	155	39	36	80
PR–	63	19	16	28
NA	10	0	6	4

NA, not available

the class comparisons produced by SAM. Official gene symbol was selected as gene identifier. The functional annotation clustering report of this software reports similar annotations together, where the member of a cluster have similar biological meaning due to sharing of similar gene members.

Gene set enrichment analysis (GSEA) was used to identifying sets of genes that were enriched in the metabolic clusters [27, 28]. During each cluster comparison, genes were ranked depending on calculated absolute signal-to-noise ratio (Eq. 1), where μ and σ are the mean and standard deviation, respectively.

$$\mathrm{abs}\left(\frac{\mu_A - \mu_B}{\sigma_A + \sigma_B}\right) \quad (1)$$

High absolute signal-to-noise ratio will represent genes that are more likely to be "class markers" in the comparison because of high difference in expression.

The gene set C5 (gene ontology (GO) gene sets) available from the Molecular Signatures Database (MSigDB) [29] from The Broad Institute was chosen for evaluation of enrichment. One thousand four (of 1454) gene sets from this data base passed the filtering of lacking any gene from the expression data followed by minimum and maximum size of 15 and 500 genes, respectively. For each comparison, 1000 permutations on phenotypes were performed and FDR cutoff was set to 25 % (recommended in the manual).

Integrated pathway analysis

To combine transcriptomics and metabolic data the "Integrated pathway analysis" tool in MetaboAnalyst 3.0 software was used [30]. Genes with adjusted $p < 0.05$ from SAM analysis and metabolites differently expressed between the clusters were used as input. Pathways with p values ≤0.05 were interpreted as significant.

Results

Three main metabolic clusters of breast cancer

From the spectral data of 228 breast tumors, hierarchical clustering gave a dendrogram divided in three metabolic clusters (Mc) (Fig. 1a) Mc1, Mc2, and Mc3. The mean spectra of the clusters are illustrated in Fig. 1b.

The prediction of the metabolic clusters by PLS-DA resulted in a model with two valid latent variables LVs (Fig. 2a). The clusters Mc1 and Mc2 were well separated in the score plot of LV1 and LV2, while most Mc3 samples had low values of LV2. Classification accuracy was found to be 91.1, 88.7, and 69.9 %, respectively, for the three clusters. Permutation testing showed that all three clusters had significantly different metabolic profiles ($p < 0.001$). The regression vectors for each of the clusters (Fig. 2b) indicate each metabolite's influence on the cluster prediction. The regression vector for Mc1 showed that high levels of glycerophosphocholine (GPC) and phosphocholine (PCho) and low levels of lactate (Lac), taurine (Tau), and alanine (Ala) were important for the class prediction result. For Mc2, high levels of β-glucose (β-Glc) were important as well as low levels of Lac, creatine (Cr), glycine (Gly), Tau, GPC, PCho, and Ala. Mc3 had a regression profile with low β-Glc, GPC, and PCho levels, and high Lac, Gly, Tau, Cr, and Ala levels. Univariate comparison of metabolite levels (Additional file 2: Table S2) between the three clusters revealed that 15 out of 18 metabolites analyzed were found to be significantly different (adjusted $p < 0.05$) between at least two of the clusters (Table 2). A combination of metabolic cluster labels and heatmap of metabolite fold change further illustrate this (Fig. 3).

Clinical parameters (tumor size, histology, grade, node status, hormone receptor status) were analyzed for differences in distribution among the metabolic clusters. Only histology was found to be significantly different between the clusters (adjusted $p = 0.0144$), where 11 of 21 lobular tumors and all ductal carcinoma in situ (DCIS) ($n = 4$) were classified as Mc2 (Table 1).

Protein expression subtype (RPPA) distribution differs between the three metabolic clusters

The metabolic clusters were investigated for differences in distribution of PAM50 and RPPA subtypes. While PAM50 subtypes did not show increased frequency of occurrence in any of the metabolic clusters, (Fig. 3c, adjusted $p = 0.138$), RPPA distribution was significantly different (Fig. 3d, adjusted $p = 1.43\text{E-}04$) with only 9 % of the RPPA reactive I and II samples being classified as Mc1, and 44 % of Mc2 samples subtyped as reactive I. The complete distribution of PAM50 and RPPA subtypes is listed in Table 3.

SAM reveals only one metabolic cluster to have differences in gene expression

SAM was performed to identify expression differences between the metabolic clusters. Of the 21,851 genes, multiclass SAM showed that 696 were differently expressed between the metabolic clusters with adjusted $p < 0.01$ (Fig. 3e, Additional file 2: Table S3). Further investigation through two-class SAM revealed that Mc2 and Mc3 did not have significant differences in mRNA expression, while they had 413 and 617 genes upregulated, respectively, compared to Mc1 (Additional file 2: Tables S4 and S5, respectively). Out of these, 277 genes were found in both comparisons and upregulated compared to Mc1. DAVID software was used to investigate the biological interactions between genes that were found to be significantly differentially expressed between the metabolic clusters.

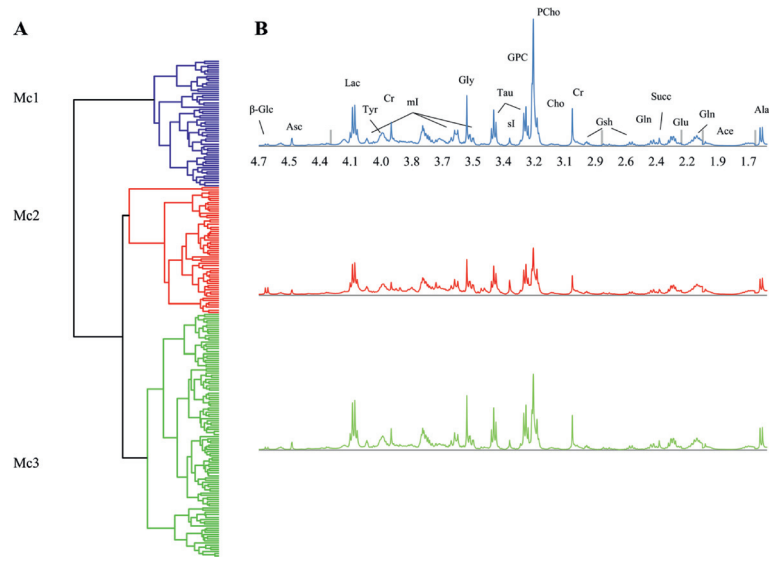


Fig. 1 Metabolic subtyping of breast cancer tissue samples using HCA. **a** The HR MAS MRS spectra for 228 samples was clustered using Euclidean distance and Wards linkage as similarity measure which separated the samples into three metabolic clusters (Mc); Mc1, Mc2, and Mc3. **b** Mean spectra for the three metabolic clusters. *β-Glc* β-glucose, *Asc* ascorbate, *Lac* lactate, *Tyr* tyrosine, *Cr* creatine, *mI* myoinositol, *Gly* glycine, *Tau* taurine, *sI* scylloinositol, *GPC* glycerophosphocholine, *PCho* phosphocholine, *Cho* choline, *Gsh* glutathione, *Gln* glutamine, *Succ* succinate, *Glu* glutamate, *Ace* acetate, *Ala* alanine. *Grey bars* indicate removed spectral regions (containing lipid peaks)

A total of 404 of the 413 significant genes from SAM between Mc1 and Mc2 were identified by DAVID. Functional Annotation Clustering resulted in 117 clusters (Top 10 in Additional file 2: Table S6), where the clusters with the highest enrichment scores were linked to signaling, extracellular region, and cell adhesion.

A total of 653 of the 671 significant genes from SAM between Mc1 and Mc3 were identified by DAVID. Functional Annotation Clustering resulted in 236 clusters (Top 10 in Additional file 2: Table S7), where the clusters with the highest enrichment scores were linked to extracellular matrix (ECM), cell adhesion, and basement membrane.

Enrichment analysis shows gene expression differences to be related to ECM activity

Since Mc1 was found to have a gene expression pattern different from both Mc2 and Mc3 and these two clusters lacked statistically significant gene expression differences, Mc1 was compared to Mc2 and Mc3 combined in GSEA. This resulted in 146 of the gene ontology gene sets altered in Mc1 compared to Mc2 and Mc3 (Additional file 2: Table S8). Gene sets with the highest significance were classified with functions within collagen, ECM, and integrin binding. None of the gene ontology sets were significantly different when comparing Mc2 to Mc1 combined with Mc3, but 44 gene sets were significantly enriched when comparing Mc2 to Mc1 alone, with gene ontology terms relevant to ECM dominating the result (Additional file 2: Table S9). Eleven gene sets were significantly altered between Mc3 and Mc1 combined with Mc2 (Additional file 2: Table S10) and also here ECM-related findings were reported. One hundred fourteen gene sets were significantly different between Mc1 and Mc3, while none were significant between Mc2 and Mc3 (results not shown).

Joint analysis of gene and metabolite expression shows differences in metabolic pathways

Integrated pathway analysis resulted in 12 significantly different metabolic pathways (p value <0.05) between Mc1 and Mc2 (Additional file 2: Table S11). The most significant pathway was "tyrosine metabolism' with eight hits of genes and metabolites, but also "D-glutamine and D-glutamate metabolism," "glycolysis/gluconeogenesis" (Fig. 4a), and "glycerophospholipid metabolism" (Fig. 4b) were among the significant pathways. Integrated pathway analysis resulted in four significantly different metabolic pathways (p value <0.05) between Mc1 and Mc3 (Additional file 2: Table S12). The most significant pathway was glycerophospholipid metabolism with nine hits, succeeded by D-glutamine and D-glutamate metabolism.

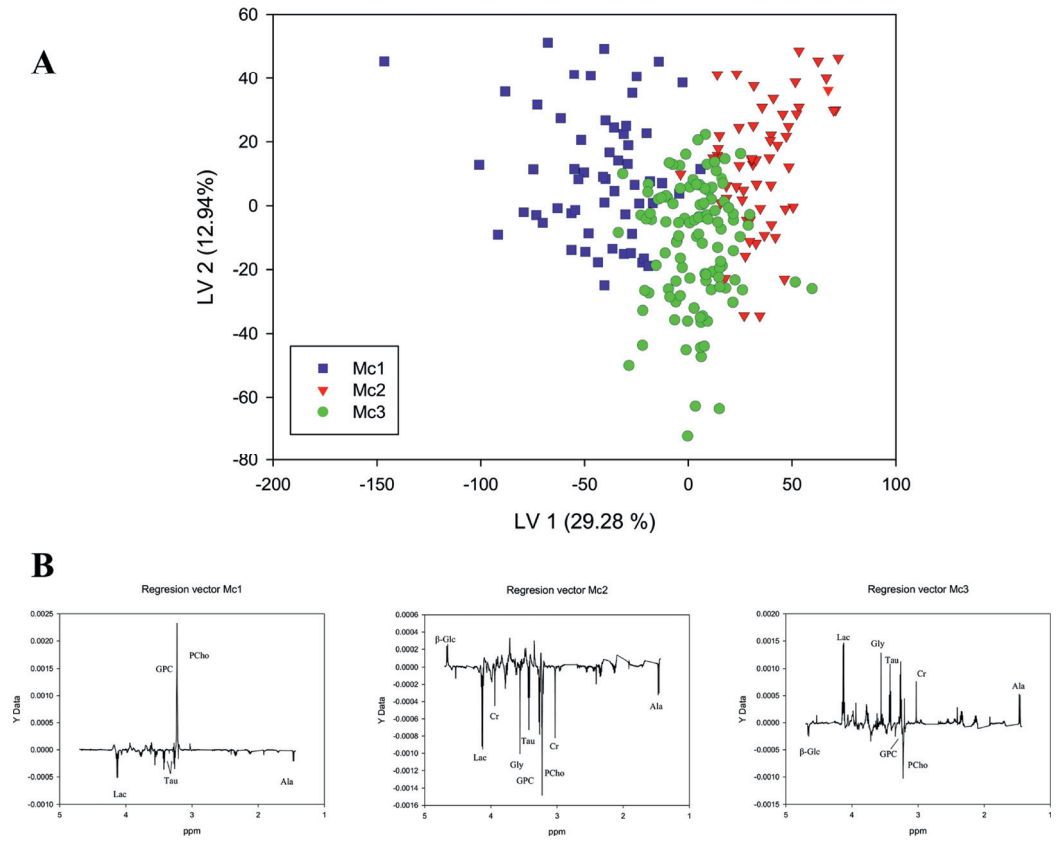


Fig. 2 Results from PLS-DA of metabolic clusters. **a** Score plot of the two first latent variables explaining 42.2 % of the *X*-variance and 28.2 % of the *Y*-variance. (**b**) Regression vectors for the three metabolic clusters (Mc)

Discussion

In the present work, metabolite, protein, and gene expression data from 228 breast tumors were combined to search for new insight into the heterogeneity of breast cancer. MR metabolite data was used to derive naturally occurring metabolic clusters, which were further combined with data from the proteomics and transcriptomics levels. We identified three significantly different metabolic clusters, Mc1, Mc2, and Mc3, with significant differences in gene expression and protein expression profiles, but not within PAM50 subgroups. The metabolic clusters could therefore contribute with additional information beyond the intrinsic gene sets for understanding breast cancer heterogeneity.

Of the three metabolic clusters, Mc1 was on a separate branch in the dendrogram indicating that the metabolic profile of this cluster was the most different. This cluster is defined by significantly higher levels of GPC and PCho, two choline-containing metabolites involved in the synthesis and degradation of phosphatidylcholine (PtdCho), a major component of cell membranes [31]. Altered choline metabolism has been considered an emerging hallmark for malignant transformations and has been detected in several cancer types including breast cancer [32]. PCho in particular has been suggested a biomarker of breast cancer [33]. Both GPC and PCho are confirmed elevated in tumor tissue compared to adjacent non-involved tissue from breast cancer patients [17], and a higher GPC/PCho-ratio has been reported in ER negative tumors [34, 35]. The latter was also observed for our cohort (results not shown); however, there was no significant difference in ER status between the three metabolic clusters. Thus, the high level of GPC and PCho is not resulting from differences in the distribution of estrogen receptor (ER) status. Interestingly, integrated pathway analysis showed that

Table 2 Metabolite levels for the metabolic clusters

	Mc1 ($n = 58$)		Mc2 ($n = 58$)		Mc3 ($n = 112$)		Adjusted p value	Significant between
Metabolite name	Mean	SE	Mean	SE	Mean	SE		
Beta-D-glucose	30.0	27.7	71.7	55.2	32.3	16.3	3.62E-09	Mc2 vs rest
Ascorbate	40.0	17.1	28.8	8.6	38.3	13.6	1.02E-05	Mc2 vs rest
Lactate	259.5	73.0	229.4	57.6	303.6	76.7	4.98E-09	Mc3 vs rest
L-tyrosine	407.5	56.5	352.8	82.6	405.9	62.4	1.22E-04	Mc2 vs rest
Glycine	187.0	80.8	152.3	41.9	195.7	68.8	1.04E-04	Mc2 vs Mc3
Myoinositol	163.7	47.0	217.7	53.5	196.1	54.3	9.44E-07	all
Taurine	332.2	122.7	330.2	84.0	369.3	99.3	0.017	Mc1 vs Mc3
Scylloinositol	55.0	16.2	94.7	186.5	62.5	32.1	0.138	NS
Glycerophosphocholine	210.0	91.6	107.9	33.6	151.2	48.4	4.44E-12	all
Phosphocholine	552.0	131.1	216.8	66.8	327.2	69.9	9.59E-33	all
Choline	135.2	44.6	120.3	37.7	132.9	42.2	0.128	NS
Creatine	149.9	64.2	93.2	33.7	136.0	52.1	1.41E-09	Mc2 vs rest
Glutathione	57.5	13.8	50.9	13.9	58.1	14.5	0.011	Mc2 vs Mc3
Glutamine	134.4	41.3	134.3	30.2	145.4	43.6	0.223	NS
Succinate	58.0	15.7	53.6	10.6	62.2	15.7	0.003	Mc2 vs Mc3
Glutamate	237.9	61.3	266.2	63.3	277.5	61.2	1.95E-04	Mc1 vs rest
Acetate	32.7	9.0	48.4	17.2	40.3	13.1	7.89E-08	all
Alanine	82.6	36.6	66.0	24.9	95.1	33.8	6.56E-07	all

The values are calculated by integrated peak areas from normalized spectra to equal total areas. Kruskal-Wallis test was performed to compare metabolite levels between clusters, and p values were adjusted for multiple testing by The Benjamini Hochberg false discovery rate
NS not significant (adjusted $p > 0.05$)

glycerophospholipid metabolism was the most significant pathway, when comparing Mc1 to Mc2. This metabolic pathway had eight hits including the metabolites GPC and PCho and genes *LCAT, LPCAT2*, *PPAP2A*, *PPAP2B*, *PLD1*, and *AGPAT4*. Downregulation of the expression of these genes in Mc1 indicate a less active degradation of PtdCho causing an accumulation of GPC and PCho, thus explaining the higher levels of GPC and PCho in Mc1. Furthermore, *LPCAT2* is involved in the reaction where the GPC precursor (acyl-GPC) is converted into PtdCho. Lower expression of this gene may explain why the GPC precursor is directed to the production of GPC instead of PtdCho. The same hits were obtained when Mc1 was compared to Mc3. In addition, *PLA2G5*, one of the enzymes degrading PtdCho to acyl-GPC, is downregulated in Mc1 compared to Mc3, further supporting that Mc1 has an altered PtdCho metabolism.

The levels of PCho and GPC were higher in Mc1 compared to the two other clusters, but no significant difference in the expression of choline kinase alpha (*CHKA)* could be detected in the SAM analysis. However, univariate analysis confirmed that *CHKA* expression was significantly higher in Mc1. This is in agreement with previous findings revealing a positive correlation between levels of PCho and GPC and expression of *CHKA* [34, 36].

For Mc1 compared to Mc2 through integrated pathway analysis, D-glutamine and D-glutamate metabolism has only two hits, but comes out as significant because of the small number of genes and metabolites within this pathway. Interestingly, the gene *GLS* which catalyzes the conversion of glutamine to glutamate is downregulated in Mc1, the cluster with lowest levels of glutamate. Glutamine metabolism is considered a therapeutic target as some cancer cells exhibit high uptake and addiction to this nonessential amino acid [37]. Since there were no differences in glutamine levels of Mc1 and Mc2, less glutamate in Mc1 could indicate that more glutamine is directed towards other metabolic pathways necessary for proliferation, glutathione needed for reducing power or further that glutamate is rapidly metabolized in cells through the TCA cycle or other mechanisms.

The distribution of protein subtypes (RPPA) was significantly different between the metabolic clusters, whereas no significant differences in the distribution of PAM50 subtypes were found. Thus, the metabolic difference between Mc1, Mc2, and Mc3 is not a result of intrinsic subtypes and might therefore contain additional information for understanding breast cancer heterogeneity. Among the tumors clustered in Mc1, 12 % were classified as RPPA-reactive (either I or II) while 49 % were classified as RPPA-luminal. The reactive RPPA

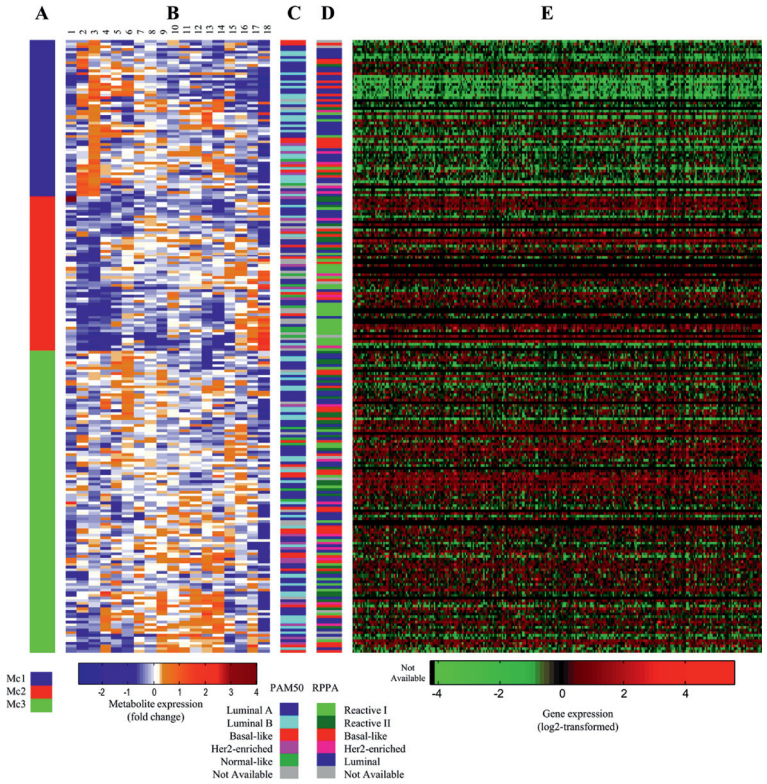


Fig. 3 Main differences between metabolic subtypes. **a** Metabolic cluster label from HCA with Euclidean distance and Wards linkage of HR MAS MR spectra of samples. The samples clustered in three groups called Mc1, Mc2, and Mc3. **b** Fold change in expression levels of (1) scylloinositol, (2) GPC, (3) PCho, (4) creatine, (5) ascorbate, (6) taurine, (7) GSH, (8) tyrosine, (9) lactate, (10) glutamate, (11) succinate, (12) glutamine, (13) glycine, (14) alanine, (15) choline, (16) myoinositol, (17) acetate, and (18) glucose. *Blue regions* in the heat map represent decreased levels while *red levels* represent increased metabolite levels. **c** PAM50-subtypes. **d** RPPA-subtype. **e** Gene expression levels (quantile normalized, log2-transformed) for the 277 overlapping significant genes (SAM, adjusted $p < 0.01$) between Mc1 and Mc3. The genes have been clustered

subtypes have a characteristic protein expression pattern probably produced by the microenvironment [5], indicating less microenvironmental activity within Mc1. Mc1 also had downregulation of several genes involved in processes within the ECM of the stroma compared to both Mc2 and Mc3. As ECM changes can drive cancer behavior [38], these genetic differences between Mc1 and Mc2 might be of prognostic relevance. In fact, differences in expression of ECM-related genes have been used to stratify breast carcinomas into four groups, where the subgroup ECM1 have the worst prognosis [39]. ECM classification was not performed on this cohort. However, 34 of 43 genes that clustered with a tendency of being downregulated in ECM1 and ECM2 were also found to be downregulated in Mc1. In addition, only 5 of 46 genes reported to be downregulated in ECM2 compared to ECM1 were downregulated in Mc1 (results extracted from SAM analyses, Additional file 2: Table S6–S7). These results support the contention that Mc1 tumors have an ECM signature similar to the reported ECM2 tumors. ECM2 did not show significant difference in disease outcome compared to ECM3 and ECM4, but had better prognosis than ECM1 tumors [39].

Mc2 has a metabolic profile with significant higher glucose level and at the same time lower levels of most of the other metabolites compared to one or both of the remaining clusters. High glucose level could reflect lower glucose consumption, inferring a lower demand for energy within these tumors. Glycolysis/gluconeogenesis came out as a significant pathway when Mc1 was compared to Mc2 during integrated pathway analysis with two metabolite

Table 3 Distribution of PAM50 and RPPA subtype among the metabolic clusters

		Metabolic cluster		
	Total	Mc1	Mc2	Mc3
PAM50 subtype				
Luminal A	85	19 (35)	18 (43)	48 (46)
Luminal B	56	23 (42)	5 (12)	28 (27)
Basal	24	6 (11)	5 (12)	13 (13)
Her2 enriched	22	5 (9)	7 (17)	10 (10)
Normal-like	14	2 (4)	7 (17)	5 (5)
NA	27	4	15	8
Total	201	55	42	104
RPPA subtype				
Reactive I	43	4 (7)	24 (44)	15 (14)
Reactive II	36	3 (5)	8 (15)	25 (23)
Basal	47	16 (29)	8 (15)	23 (21)
Her2	18	5 (9)	4 (7)	9 (8)
Luminal	73	27 (49)	11 (20)	35 (33)
NA	11	3	3	5
Total	217	55	55	107

Values in brackets are each subtype's percentage distribution within the metabolic clusters
NA not available

hits and five gene hits. For the most significant metabolite, glucose, the levels are higher in Mc2 compared to Mc1. Glucose is the main source of energy for mammalian cells, either through aerobic glycolysis (production of lactate even in the presence of oxygen) or tricarboxylic acid (TCA) cycle and oxidative phosphorylation. For normal proliferating cells and cancer cells, which both have an increased energy demand, a glycolytic switch is often observed (higher glycolytic rate) [12]. The increased glycolysis is followed by fermentation of the pyruvate to lactate (Warburg effect), in contrast to the conversion of acetyl CoA through the TCA cycle that occurs in normal non-proliferating cells. Increased glucose consumption is commonly used in tumor detection by using a glucose analogue and positron emission tomography (PET) [40] and has shown to correlate with poor prognosis and tumor aggressiveness [12]. However, not all breast cancers are detected by PET. Here, we expect lower sensitivity in detection of Mc2 tumors due to the possible difference in glycolytic rate. None of the genes with hits in glycolysis/gluconeogenesis for the comparison of Mc1 and Mc2 could directly explain the high glucose levels of Mc2 tumors, but altered expression of the genes indicates pyruvate being guided towards the TCA cycle rather than lactate production. Two of the alternative fates of pyruvate showed significantly higher levels (alanine) or levels approaching significance (lactate, adjusted $p = 0.056$), supporting a higher glycolytic rate in Mc1 and that the pyruvate produced is not directed to metabolism in the TCA cycle. The significantly lower acetate levels in Mc1 compared to Mc2 could be linked to *ALDH1A3* and *ALDH2* downregulation, since the enzymatic product of these genes catalyzes the reversible reaction where acetaldehyde is converted to acetate.

Both DAVID and GSEA showed that many of the genes found to be downregulated in Mc1 and consequently upregulated in Mc2 were related to ECM activity. Mc2 had the highest percentage of RPPA-reactive I with 44 % of Mc2 tumors classified as this protein subtype, also related to stromal changes. Together with the metabolic finding, this implies that Mc2 tumors have cancer cells with low proliferating rate and at the same time ongoing changes within the ECM of the stroma. Mc2 tumors also had a higher frequency of lobular and ductal carcinoma in situ, indicating metabolic differences between histological subtypes of breast cancer which should be further investigated.

Mc3 has the highest lactate levels of all three clusters and higher glycine level than Mc2. These metabolites have been related to poor prognosis in ER positive patients [18], and higher levels of glycine is also associated with poor prognosis in a study irrespective of ER status [41]. Although the ER-positive patients are equally distributed among our reported metabolic clusters, Mc3 expressed higher levels of both of these metabolites compared to Mc2. Moestue et al. detected differences in the expression of genes involved in choline degradation that could explain higher glycine concentrations in the poor-prognosis basal-like breast cancer xenograft model compared to luminal-like [42]. Five of the genes described by Moestue et al. were significantly upregulated in Mc3 compared to Mc1; *AGPAT4*, *PPAP2B*, *PPAP2A*, *LCAT*, and *PLD1*. Of these, *LCAT* and *PLD1* are directly involved in choline metabolism. *LCAT* catalyze the conversion of PtdCho to acyl-GPC while *PLD1* catalyzes the conversion of PtdCho to choline. Higher GPC levels, but no difference in choline levels in Mc3 compared to Mc1 indicates that a higher amount of GPC is converted to choline in Mc3, and further contributing to higher glycine levels through choline degradation.

Mc3 shares similarities with a previously reported metabolic subgroup of luminal A tumors with significantly lower levels of glucose, higher levels of alanine, and nearly significantly higher lactate levels [19]. In Mc3, we also see a significant higher level of lactate. Since one of the main sources of alanine is pyruvate, which also is the source for lactate, it appears that Mc3 is a cluster with a switch in glycolytic activity.

The majority of Mc3 tumors were classified as RPPA-luminal, similar to Mc1. In contrast to Mc1, Mc3 had a higher percentage of RPPA–reactive II tumors, probably linked to changes in stromal content. Also,

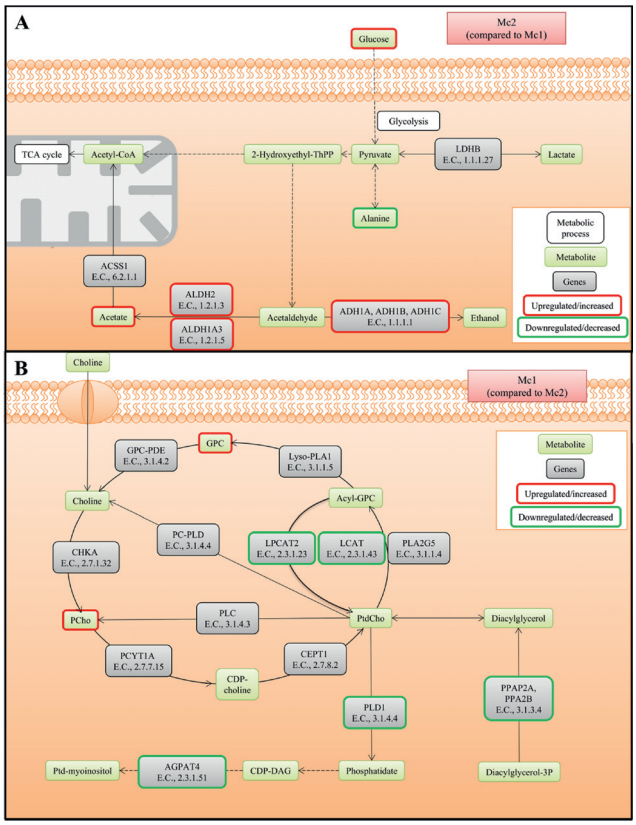


Fig. 4 Illustration of metabolic pathways reported to have altered gene and metabolite expression by integrated pathway analysis (MetaboAnalyst). **a** Result within "glycolysis/glutaminolysis" genes and metabolites differently expressed in metabolic cluster (Mc) Mc2 compared to Mc1. Adapted from KEGG ID: hsa00010. *LDHB* lactate dehydrogenase B; *ADH1A/ADH1B/ADH1C* alcohol dehydrogenase 1A/1B/1C; *ALDH1A3*: aldehyde dehydrogenase 1 family member A3; *ALDH2* aldehyde dehydrogenase 2 family; *ACSS1* acetate CoA ligase; *TCA cycle* trucarboxylic acid cycle. **b** Result within "glycerophospholipid metabolism" of genes and metabolites differently expressed in Mc1 compared to Mc2. Adapted from KEGG ID: hsa00564. *CHKA* choline kinase alpha; *PCYT1A* phosphate cytidylytransferase 1; *CEPT1* choline/ethanolamine phosphotransferase 1; *PLA2G5* phospholipase A2; *LCAT* lecithin-cholesterol acyltransferase; *LPCAT2* lysophosphatidyl-choline acyltransferase; *PC-PLD* phospholipase D; *Lyso-PLA1* lysophospholipase I; *GPC-PDE* glycerophos-phocholine phosphodiesterase; *PLC* phospholipase C; *PLD1* phospholipidase D1; *PPAP2A, PPAP2B* phosphatidate phosphatase LPIN; *AGPAT4* 1-acylglycerol-3-phosphate O-acyltransferase

gene expression wise, this was observed by significantly different gene expressions linked to ECM activity and the gene expression profile of Mc3 was found similar to the previously reported ECM3 or ECM4 subtypes [39].

In this study, information flow between the transcriptomics, proteomics, and metabolomics levels is illustrated; at the transcriptomics level, only one of the metabolic clusters shows difference in gene expression compared to the two others, while at the proteomics level, there is difference between all three clusters. Combining these findings, Mc1 is expected to have the worst prognosis due to the distinct gene expression profile and the alterations in both glycerophospholipid metabolism and evidence of increased glycolytic rate. However, this has to be validated when 5-year follow-up of this cohort is available. The main metabolic characteristics, especially of Mc1 and Mc3, have been proposed as treatment targets that could improve the therapeutic effect [43]. Cancer therapy targeting *CHKA*, the enzyme responsible for PCho production from choline, causes tumor growth arrest and apoptosis in preclinical models [44], while treatment targeting glycolytic enzymes in combination with chemotherapy has been shown to re-sensitize

cancer cells that had become resistant to treatment [43]. Metabolic classification as illustrated here could therefore be relevant for developing a more targeted treatment plan. Importantly, the prognostic value of the clusters should be evaluated once 5-year follow-up is available.

Conclusions

We have here identified three metabolic clusters of breast cancer, also characterized with differences at the proteomic and transcriptomic level. The metabolic clusters are not reflecting the intrinsic genetic subtypes and may give important additional information for understanding breast cancer heterogeneity. Gene enrichment analysis revealed diverse ECM characteristics among these clusters in accordance with RPPA-subtyping. The approach of combining information from several -omics levels in the same tumor shows promise in improving the understanding of breast cancer heterogeneity potentially leading to more patient specific treatment.

Additional files

Additional file 1: Additional methods and references. Information of additional methods. (DOCX 22 kb)

Additional file 2: Table S1. Antibodies used for reverse phase protein array (RPPA). **Table S2.** Metabolite integrals for each of the samples included in the study. **Table S3.** Significantly different expressed genes between the three metabolic clusters. **Table S4.** Significantly different expressed genes between the metabolic clusters Mc1 and Mc2. **Table S5.** Significantly different expressed genes between the metabolic clusters Mc1 and Mc3. **Table 6.** Statistically over-represented annotation terms, according to DAVID, of differently expressed genes between metabolic cluster Mc1 and Mc2. **Table S7.** Statistically over-represented annotation terms, according to DAVID, of differently expressed genes between metabolic cluster Mc1 and Mc3. **Table S8.** Gene set enrichment analysis (GSEA) result for gene ontology (GO) gene sets. Metabolic cluster Mc1 was compared with Mc2 and Mc3. **Table S9.** Gene set enrichment analysis (GSEA) result for gene ontology (GO) gene sets. Metabolic cluster Mc1 was compared with Mc2. **Table S10.** Gene set enrichment analysis (GSEA) result for gene ontology (GO) gene sets. Metabolic cluster Mc3 was compared with Mc1 and Mc2. **Table S11.** Integrated pathway analysis result from the comparison of the metabolic clusters Mc1 compared to Mc2. **Table S12.** Integrated pathway analysis result from the comparison of the metabolic clusters Mc1 compared to Mc3. (XLSX 337 kb)

Abbreviations

Ala, alanine; Cr, creatine; DCIS, ductal carcinoma in situ; ECM, extracellular matrix; ER, estrogen receptor; FDR, false discovery rate; Gly, glycine; GPC, glycerophosphocholine; GSEA, gene set enrichment analysis; HCA, hierarchical cluster analysis; HR MAS MRS, high resolution magic angle spinning magnetic resonance spectroscopy; Lac, lactate; Mc, metabolic cluster; MRS, magnetic resonance spectroscopy; MSigDB, molecular signatures database; PAM50, prediction analysis of microarrays 50; PCho, phosphocholine; PET, positron emission tomography; PLS-DA, partial least square discriminant analysis; PtdCho, phosphatidylcholine; RPPA, reverse phase protein array; SAM, significance analysis of microarrays; Tau, taurine; TCA, tricarboxylic acid; β-Glc, β-glucose

Acknowledgements

The HR MAS MRS analysis was performed at the MR Core Facility, Norwegian University of Science and Technology (NTNU). MR core facility is funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority.

The Oslo Breast Cancer Research Consortium (OSBREAC).
Vessela N Kristensen[1,2,3], Torill Sauer[4,5], Elin Borgen[6], Olav Engebråten[7,8,9], Øystein Fodstad[7,9], Rolf Kåresen[9,10], Bjørn Naume[2,8], Gunhild Mari Mælandsmo[2,7,11], Hege G Russnes[1,2,12], Therese Sørlie[1,2], Helle Kristine Skjerven[13], Britt Fritzman[14].
[1]Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway. [2]K.G. Jebsen Centre for Breast Cancer Research, Institute for Clinical Medicine, University of Oslo, Oslo, Norway. [3]Department of Clinical Molecular Biology and Laboratory Science (EpiGen), Division of Medicine, Akershus University Hospital, Lørenskog, Norway. [4]Department of Pathology, Akershus University Hospital, Lørenskog, Norway. [5]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [6]Department of Pathology, Division of Diagnostics and Intervention, Oslo University Hospital, Oslo, Norway. [7]Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway. [8]Department of Oncology, Division of Surgery and Cancer and Transplantation Medicine, Oslo University Hospital, Oslo, Norway. [9]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [10]Department of Breast- and Endocrine Surgery, Division of Surgery, Cancer and Transplantation, Oslo University Hospital, Oslo, Norway. [11]Department of Pharmacy, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway. [12]Department of Pathology, Oslo University Hospital, Oslo, Norway. [13]Breast and Endocrine Surgery, Department of Breast and Endocrine Surgery, Vestre Viken Hospital, Drammen, Norway. [14]Østfold Hospital, Østfold, Norway.
Email: Vessela N Kristensen v.n.kristensen@medisin.uio.no, Torill Sauer Torill.sauer@medisin.uio.no, Elin Borgen ebg@ous-hf.no, Olav Engebråten Olav.engebraten@medisin.uio.no, Øystein Fodstad Oystein.Fodstad@rr-research.no, Rolf Kåresen rolf.karesen@medisin.uio.no, Bjørn Naume bjorn.naume@medisin.uio.no, Gunhild Mari Mælandsmo Gunhild.Mari.Malandsmo@rr-research.no, Hege G Russnes Hege.russnes@rr-research.no, Therese Sørlie therese.sorlie@rr-research.no, Helle Kristine Skjerven Helle.skjerven@vestreviken.no, Britt Fritzman Britt.Fritzman@so-hf.no

Funding

This work was supported by
(1) K. G. Jebsen Center for Breast Cancer Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript
(2) Cancer Society of Norway
(3) South-Eastern Norway Regional Health Authority (Project no 2011042, OSBREAC: Towards personalized therapy for breast cancer)
(4) The Research Council of Norway, Imaging the breast cancer metabolome, Project no 221879
(5) Liaison Committee between the Central Norway Regional Health Authority and NTNU Project no 46056615
(6) NCI grants P01CA0099031 and P30 CA016672 to GBM

Availability of data and materials

The datasets supporting the conclusions of this article are included within the Additional files and included in the article; "Integrated analysis reveals microRNA networks coordinately expressed with key proteins in breast cancer" Aure MR, Jernström S, Krohn M, Vollan HK, Due EU, Rødland E, Kåresen R; Oslo Breast Cancer Research Consortium (OSBREAC), Ram P, Lu Y, Mills GB, Sahlberg KK, Børresen-Dale AL, Lingjærde OC, Kristensen VN. Genome Med. 2015 Feb 2; where mRNA expression data have been submitted to the Gene Expression Omnibus (GEO) database as a SuperSeries record with accession number GSE58215.

Authors' contributions

THH, LRE, GFG, OLC, ES, ØG, OSBREAC, KKS, ALBD, and TFB participated in the design of the study. KKS, ALBD, and TFB conceived the study. THH, LRE, GFG, KKS, ALBD, and TFB interpreted the data. SL and THH performed the HR MAS MRS acquisition. THH performed the statistical analysis and drafted the manuscript. MK, SJ, MRA, OCL, ES, ØG, EUD, OSBREAC, GBM, KKS, ALBD, and TFB participated in acquisition of the data. All authors have read and helped to revise the manuscript. The final manuscript is approved by all the authors.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate
The study is approved by the Norwegian Regional Committee for Medical Research Ethics (Biobank approval 1.2006.1607), and all patients have given written consent for the use of material for research purposes.

Author details
[1]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [2]K.G. Jebsen Center for Breast Cancer Research, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [3]St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [4]Department of Food Science, Faculty of Science and Technology, Aarhus University, Årslev, Denmark. [5]Department of Cancer Genetics, Institute for Cancer Research Oslo University Hospital, The Norwegian Radium Hospital, Oslo, Norway. [6]Department of Computer Science, University of Oslo, Oslo, Norway. [7]Centre for Cancer Biomedicine, University of Oslo, Oslo, Norway. [8]Section for Breast and Endocrine Surgery, Oslo University Hospital, Ullevål, Oslo, Norway. [9]Department of Pathology, Oslo University Hospital, Oslo, Norway. [10]Department of Systems Biology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA. [11]Department of Research, Vestre Viken, Drammen, Norway.

Received: 15 February 2016 Accepted: 6 June 2016
Published online: 27 June 2016

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0: Cancer incidence and mortality worldwide: IARC CancerBase No. 11. 2013. http://globocan.iarc.fr. Accessed: 04 Jan 2015.
2. Polyak K. Heterogeneity in breast cancer. J Clin Invest. 2011;121(10):3786–8.
3. Wang Y, Crampin E. Biclustering reveals breast cancer tumour subgroups with common clinical features and improves prediction of disease recurrence. BMC Genomics. 2013;14(1):102.
4. Perou C, Sørlie T, Eisen M, van de Rijn M, Jeffrey S, Rees C, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747–52.
5. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70.
6. Sørlie T, Perou C, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869–74.
7. Parker J, Mullins M, Cheang M, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol. 2009;27(8):1160–7.
8. Myhre S, Lingjærde O-C, Hennessy B, Aure M, Carey M, Alsner J, et al. Influence of DNA copy number and mRNA levels on the expression of breast cancer related proteins. Mol Oncol. 2013;7(3):704–18.
9. Akbani R, Ng P, Werner H, Shahmoradgoli M, Zhang F, Ju Z, et al. A pan-cancer proteomic perspective on The Cancer Genome Atlas. Nat Commun. 2014;5.
10. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, et al. Proteogenomic characterization of human colon and rectal cancer. Nature. 2014; 513(7518):382–7.
11. Hennessy B, Lu Y, Gonzalez-Angulo A, Carey M, Myhre S, Ju Z, et al. A technical assessment of the utility of reverse phase protein arrays for the study of the functional proteome in non-microdissected human breast cancers. Clin Proteomics. 2010;6(4):129–51.
12. Gatenby R, Gillies R. Why do cancers have high aerobic glycolysis? Nat Rev Cancer. 2004;4(11):891–9.
13. Ward P, Thompson C. Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. Cancer Cell. 2012;21(3):297–308.
14. Hanahan D, Weinberg R. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74.
15. Bathen T, Geurts B, Sitter B, Fjøsne H, Lundgren S, Buydens L, et al. Feasibility of MR metabolomics for immediate analysis of resection margins during breast cancer surgery. PLoS One. 2013;8(4):e61578.
16. Cao M, Sitter B, Bathen T, Bofin A, Lønning P, Lundgren S, et al. Predicting long-term survival and treatment response in breast cancer patients receiving neoadjuvant chemotherapy by MR metabolic profiling. NMR Biomed. 2012;25(2):369–78.
17. Sitter B, Lundgren S, Bathen T, Halgunset J, Fjosne H, Gribbestad I. Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters. NMR Biomed. 2006;19(1):30–40.
18. Giskeødegård G, Lundgren S, Sitter B, Fjøsne HE, Postma G, Buydens L, et al. Lactate and glycine—potential MR biomarkers of prognosis in estrogen receptor-positive breast cancers. NMR Biomed. 2012;25(11):1271–9.
19. Borgan E, Sitter B, Lingjærde O, Johnsen H, Lundgren S, Bathen T, et al. Merging transcriptomics and metabolomics-advances in breast cancer profiling. BMC Cancer. 2010;10(1):628.
20. Tibes R, Qiu Y, Lu Y, Hennessy B, Andreeff M, Mills G, et al. Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. Mol Cancer Ther. 2006;5(10):2512–21.
21. Aure M, Jernström S, Krohn M, Vollan H, Due E, Rødland E, et al. Integrated analysis reveals microRNA networks coordinately expressed with key proteins in breast cancer. Genome Med. 2015;7(1):1-17.
22. Nilsen G, Vollan H, Pladsen A, Borgan Ø, Liestøl K, Vitelli V, et al. Dissecting genome aberration patterns in tumors, Submitted. 2015.
23. Sitter B, Sonnewald U, Spraul M, Fjösne H, Gribbestad I. High-resolution magic angle spinning MRS of breast cancer tissue. NMR Biomed. 2002;15(5):327–37.
24. Tusher V, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001;98(9):5116–21.
25. R Core Team. R: a language and environment for statistical computing. Viena: R Foundation for Statistical Computing; 2012.
26. da Huang W, Sherman B, Lempicki R. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2008;4(1):44–57.
27. Subramanian A, Tamayo P, Mootha V, Mukherjee S, Ebert B, Gillette M, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.
28. Mootha V, Lindgren C, Eriksson K-F, Subramanian A, Sihag S, Lehar J, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet. 2003;34(3):267–73.
29. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, and Mesirov J. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011; http://www.broadinstitute.org/gsea/msigdb/index.jsp. Accessed: 08 Oct 14.
30. Xia J, Mandal R, Sinelnikov I, Broadhurst D, Wishart D. MetaboAnalyst 2.0—a comprehensive server for metabolomic data analysis. Nucleic Acids Res. 2012;40(Web Server issue):W127–33.
31. Gibellini F, Smith T. The Kennedy pathway—de novo synthesis of phosphatidylethanolamine and phosphatidylcholine. IUBMB Life. 2010;62(6):414–28.
32. Glunde K, Bhujwalla Z, Ronen S. Choline metabolism in malignant transformation. Nat Rev Cancer. 2011;11(12):835–48.
33. Eliyahu G, Kreizman T, Degani H. Phosphocholine as a biomarker of breast cancer: molecular and biochemical studies. Int J Cancer. 2007;120(8):1721–30.
34. Grinde M, Skrbo N, Moestue S, Rødland E, Borgan E, Kristian A, et al. Interplay of choline metabolites and genes in patient-derived breast cancer xenografts. Breast Cancer Res. 2014;16:1-16.
35. Giskeødegård G, Grinde M, Sitter B, Axelson D, Lundgren S, Fjøsne H, et al. Multivariate modeling and prediction of breast cancer prognostic factors using MR metabolomics. J Proteome Res. 2010;9(2):972–9.
36. Cao MD, Döpkens M, Krishnamachary B, Vesuna F, Gadiya M, Lønning P, et al. Glycerophosphodiester phosphodiesterase domain containing 5 (GDPD5) expression correlates with malignant choline phospholipid metabolite profiles in human breast cancer. NMR Biomed. 2012;25(9):1033–42.
37. Wise D, Thompson C. Glutamine addiction: a new therapeutic target in cancer. Trends Biochem Sci. 2010;35(8):427–33.
38. Tlsty T, Hein P. Know thy neighbor: stromal cells can contribute oncogenic signals. Curr Opin Genet Dev. 2001;11(1):54–9.
39. Bergamaschi A, Tagliabue E, Sørlie T, Naume B, Triulzi T, Orlandi R, et al. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. J Pathol. 2008;214(3):357–67.
40. Mankoff D, Eary J, Link J, Muzi M, Rajendran J, Spence A, et al. Tumor-specific positron emission tomography imaging in patients:[18F] fluorodeoxyglucose and beyond. Clin Cancer Res. 2007;13(12):3460–9.
41. Sitter B, Bathen T, Singstad T, Fjøsne H, Lundgren S, Halgunset J, et al. Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy. NMR Biomed. 2010;23(4):424–31.
42. Moestue S, Borgan E, Huuse E, Lindholm E, Sitter B, Børresen-Dale A-L, et al. Distinct choline metabolic profiles are associated with differences in gene

expression for basal-like and luminal-like breast cancer xenograft models. BMC Cancer. 2010;10(1):433.
43. Zhao Y, Butler E, Tan M. Targeting cellular metabolism to improve cancer therapeutics. Cell Death Dis. 2013;4(3):e532.
44. Glunde K, Jiang L, Moestue S, Gribbestad I. MRS and MRSI guidance in molecular medicine: targeting and monitoring of choline and glucose metabolism in cancer. NMR Biomed. 2011;24(6):673–90.

Supplementary Information

Article Title: Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes

Cancer & Metabolism 2016 June 27; 4:12. DOI: 10.1186/s40170-016-0152-x

Additional File 1: Additional methods and references

Information on additional methods.

Additional File 2: Additional Tables

Additional Tables 1-12.

Additional Information:

Additional Methods

HR MAS MRS acquisition and data processing. Before HR MAS MRS experiments, 3 µL cold sodium formate in D2O (24.29mM) was added to a leak-proof disposable 30µL insert (Bruker, Biospin GmbH, Germany) as a chemical shift reference. Tissue samples were cut to fit the insert on a dedicated work station designed to keep the samples frozen [1]. The insert containing the frozen sample was placed in a 4-mm diameter zirconium rotor (Bruker, Biospin GmbH, Germany) and kept at -20 ˚C for maximum 8 hours before the experiments. Samples were spun at 5000 Hz and experiments run at 5 ˚C. The samples were allowed 5 minutes temperature acclimatization before shimming and spectral acquisition. Spin-echo spectra were recorded using a Carr-Purcell-Meiboom-Gill (cpmg) pulse sequence (cpmgpr1d; Bruker) with 4s water suppression prior to a 90° excitation pulse. T2 filtering was obtained using a delay of 0.6 ms between each 180˚ pulse to suppress macromolecules and lipid signals and enhance signal from small molecules. This resulted in an effective TE of 77 ms. A total of 256 scans over a spectral region of 12 kHz was collected into 72k complex data points with an acquisition time of 3.07 s. The FIDs were multiplied by a 0.30 Hz exponential weighting function and Fourier transformed into 64k real points. Phase correction was performed automatically for each spectrum using TopSpin 3.1 (Bruker). Further preprocessing of the HR MAS spectra were performed in Matlab R2013b (The Mathworks, Inc., USA). Chemical shifts were referenced to the creatine peak at 3.92 ppm. Baseline correction was performed using asymmetric least squares [2] with parameters λ = 1e7 and p = 0.0001, and baseline offset was adjusted by setting the minimum value of each spectrum to zero by subtracting the lowest value. Peak alignment was performed using icoshift [3].

Reverse Phase Protein Array (RPPA). Tumor tissue was lysed by homogenization in lysis buffer containing proteinase inhibitors and phosphatase inhibitors. The tumor lysates were diluted in 1.33 mg/ml concentration as assessed by bicinchonic acid assay (BCA) and boiled in 1% SDS and 2-mercaptoethanol.

Supernatants were manually diluted in five serial 2-fold dilutions with lysis buffer. The samples were spotted onto and immobilized on nitrocellulose-coated FAST slides. The slides were probed with 151 primary antibodies (Supplementary Table 1) in appropriate dilutions. The signal intensity was captured by a biotin conjugated secondary antibody and was amplified by Dako Cytomation-catalysed system (Dako, Glostrup, Denmark). Slides were scanned, analyzed and quantitated using MicroVigene software (VigeneTech Inc., Carlise, MA, USA) to generate spot signal intensities. These were then processed by the R package SuperCurve /version 1.01. The protein concentrations were derived from the supercurve for each sample by curve fitting, log2-transformed, and the relative concentrations were normalized by median centering of the samples for each of the antibodies [4].

Statistical analysis. PLS-DA was performed on mean centered spectra using double cross validation [5]. In the outer loop, the model was built on randomly chosen training samples (80 % of the spectra) and used to predict the class of the remaining independent test samples (20 % of the spectra). The outer loop was repeated 20 times in total. For each outer loop, 20 repetition of the inner loop was performed. Here, the training set was again divided into an inner set of training (80 % of the samples) and test (20 % of the samples) samples. The average classification result is calculated based on the outer loop, while the inner loop was used to determine the optimal number of LVs. To validate that the result is not achieved simply by random predictions, permutation testing was performed. Here Y-data (class labels for the samples) are permutated to resemble random classification. For each permutation 20 random training and test sets are chosen as described for the PLS-DA model. This was repeated 1000 times before the error distribution was compared with the classification error for the original data. P values ≤ 0.01 were considered significant. PCA and PLS-DA were performed in Matlab using PLS_Toolbox 7.5.2 (Eigenvector Research, Inc., Wenatchee, USA).

Additional References

1. Giskeødegård, G.F., M.D. Cao, and T.F. Bathen, *High-Resolution Magic-Angle-Spinning NMR Spectroscopy of Intact Tissue*, in *Metabonomics: Methods and Protocols*, J.T. Bjerrum, Editor 2015, Springer New York. p. 37-50.
2. Eilers, P.H., *Parametric time warping.* Anal Chem, 2004. **76**(2): p. 404-411.
3. Savorani, F., G. Tomasi, and S.B. Engelsen, *icoshift: A versatile tool for the rapid alignment of 1D NMR spectra.* J Magn Reson, 2010. **202**(2): p. 190-202.
4. Aure, M.R., et al., *Integrated analysis reveals microRNA networks coordinately expressed with key proteins in breast cancer.* Genome Med, 2015. **7**(1).
5. Smilde, A.K., et al., *Data processing in metabolomics*, in *Metabolomics in practice: successful strategies to generate and analyze metabolic data*, M. Lämmerhofer and W. Weckwerth, Editors. 2013, John Wiley & Sons.

Add. Table 1: Antibodies used for reverse phase protein array (RPPA)

RPPA Antibody	Company	Catalog #	Dilution used	Validation status	Species
14-3-3_epsilon	Santa Cruz	sc-2395	250	C- use with caution	Mouse
4E-BP1	CST	9452	100	V - validated	Rabbit
4E-BP1_pS65	CST	9456	250	V - validated	Rabbit
53BP1	CST	4937	400	C- use with caution	Rabbit
A-Raf_pS299	CST	4431	100	NA	Rabbit
ACC_pS79	CST	3661	200	V - validated	Rabbit
ACC1	Epitomics	1768-1	300	C- use with caution	Rabbit
AIB1	BD	611105	100	V - validated	Mouse
Akt	CST	9272	250	V - validated	Rabbit
Akt_pS473	CST	9271	100	V - validated	Rabbit
Akt_pT308	CST	9275	50	V - validated	Rabbit
alpha-Catenin	Calbiochem	CA1030	750	V - validated	Mouse
AMPK_alpha	CST	2532	200	C- use with caution	Rabbit
AMPK_pT172	CST	2535	100	V - validated	Rabbit
Annexin_I	Invitrogen	71-3400	90000	V - validated	Rabbit
AR	Epitomics	1852-1	50	V - validated	Rabbit
B-Raf	Santa Cruz	sc-5284	500	NA	Mouse
Bak	Epitomics	1542-1	50	C- use with caution	Rabbit
Bax	CST	2772	300	V - validated	Rabbit
Bcl-2	Dako	M0887	50	V - validated	Mouse
Bcl-X	Epitomics	1018-1	100	C- use with caution	Rabbit
Bcl-xL	CST	2762	50	V - validated	Rabbit
Beclin	Santa Cruz	sc-10086	200	V - validated	Goat
beta-Catenin	CST	9562	800	V - validated	Rabbit
Bid	Epitomics	1008-1	50	C- use with caution	Rabbit
Bim	Epitomics	1036-1	50	V - validated	Rabbit
c-Jun_pS73	CST	9164	50	C- use with caution	Rabbit
c-Kit	Epitomics	1522	50	V - validated	Rabbit
c-Met	CST	3127	500	C- use with caution	Mouse
c-Met_pY1235	CST	3129	500	C- use with caution	Rabbit
c-Myc	CST	9402	200	C- use with caution	Rabbit
C-Raf	Millipore	05-739	1000	V - validated	Rabbit
C-Raf_pS338	CST	9427	300	C- use with caution	Rabbit
Caspase-3_active	Epitomics	1476-1	200	C- use with caution	Rabbit
Caspase-7_cleavedD198	CST	9491	50	C- use with caution	Rabbit
Caspase-8	CST	9746	250	C- use with caution	Mouse
Caspase-9_cleavedD330	CST	9501	250	C- use with caution	Rabbit
Caveolin-1	CST	3238	500	V - validated	Rabbit
CD31	Dako	M0823	50	V - validated	Mouse
CDK1	CST	9112	250	V - validated	Rabbit
Chk1	CST	2345	100	C- use with caution	Rabbit
Chk1_pS345	CST	2348	50	C- use with caution	Rabbit
Chk2	CST	3440	300	C- use with caution	Mouse
Chk2_pT68	CST	2197	150	C- use with caution	Rabbit
cIAP	Millipore	07-759	250	V - validated	Rabbit
Claudin-7	Novus	NB100-91714	1000	V - validated	Rabbit
Collagen_VI	Santa Cruz	sc-20649	250	V - validated	Rabbit
COX-2	Epitomics	2169-1	150	C- use with caution	Rabbit
Cyclin_B1	Epitomics	1495-1	500	V - validated	Rabbit
Cyclin_D1	Santa Cruz	sc-718	150	V - validated	Rabbit
Cyclin_E1	Santa Cruz	sc-247	150	V - validated	Mouse
DJ-1	Abcam	ab76008	10000	C- use with caution	Rabbit
Dvl3	CST	3218	1000	V - validated	Rabbit
E-Cadherin	CST	4065	750	V - validated	Rabbit
eEF2	CST	2332	100	V - validated	Rabbit
eEF2K	CST	3692	250	V - validated	Rabbit
EGFR	Santa Cruz	sc-03	350	C- use with caution	Rabbit
EGFR_pY1068	CST	2234	50	V - validated	Rabbit
EGFR_pY1173	Epitomics	1124	50	C- use with caution	Rabbit
EGFR_pY992	CST	2235	50	V - validated	Rabbit
eIF4E	CST	9742	250	V - validated	Rabbit
ER-alpha	Lab Vision	RM-9101-S	200	V - validated	Rabbit
ER-alpha_pS118	Epitomics	1091-1	150	V - validated	Rabbit
ERCC1	Lab Vision	MS-671-PO	100	C- use with caution	Mouse
FAK	Epitomics	1700-1	75	C- use with caution	Rabbit
Fibronectin	Epitomics	1574-1	5000	C- use with caution	Rabbit
FOXO3a	CST	9467	50	C- use with caution	Rabbit
FOXO3a_pS318_S321	CST	9465	75	C- use with caution	Rabbit
GAB2	CST	3239	250	V - validated	Rabbit
GATA3	BD	558686	100	V - validated	Mouse
GSK3_pS9	CST	9336	250	V - validated	Rabbit
GSK3-alpha-beta	Santa Cruz	sc-7291	2000	V - validated	Mouse
GSK3-alpha-beta_pS21_S9	CST	9331	250	V - validated	Rabbit
HER2	Lab Vision	MS-325-P1	250	V - validated	Mouse
HER2_pY1248	R&D	AF1768	350	NA	Rabbit

RPPA Antibody	Company	Catalog #	Dilution used	Validation status	Species
HER3	Santa Cruz	sc-285	250	V - validated	Rabbit
HER3_pY1298	CST	4791	250	C- use with caution	Rabbit
IGF-1R-beta	CST	3027	250	C- use with caution	Rabbit
IGFBP2	CST	3922	50	V - validated	Rabbit
INPP4B	Santa Cruz	sc-12318	250	C- use with caution	Goat
IRS1	Millipore	06-248	2000	V - validated	Rabbit
JNK2	CST	4672	50	C- use with caution	Rabbit
K-Ras	Santa Cruz	sc-30	75	C- use with caution	Mouse
MAPK_pT202_Y204	CST	4377	100	V - validated	Rabbit
MEK1	Epitomics	1235-1	5000	V - validated	Rabbit
MEK1_pS217_S221	CST	9121	500	V - validated	Rabbit
MIG-6	Sigma	WH0054206M1	100	V - validated	Mouse
Mre11	CST	4847	250	C- use with caution	Rabbit
MSH2	CST	2850	50	C- use with caution	Mouse
MSH6	SDI	2203.00.02	1000	C- use with caution	Rabbit
N-Cadherin	CST	4061	50	V - validated	Rabbit
NF-kB-p65_pS536	CST	3033	100	C- use with caution	Rabbit
NF2	SDI	2271.00.02	500	C- use with caution	Rabbit
Notch1	CST	3268	50	V - validated	Rabbit
Notch3	Santa Cruz	sc-5593	600	C- use with caution	Rabbit
P-Cadherin	CST	2130	50	C- use with caution	Rabbit
p21	Santa Cruz	sc-397	100	C- use with caution	Rabbit
p27	Epitomics	1591-1	50	V - validated	Rabbit
p27_pT157	R&D	AF1555	500	C- use with caution	Rabbit
p27_pT198	Abcam	ab64949	75	V - validated	Rabbit
p38_MAPK	CST	9212	100	C- use with caution	Rabbit
p38_pT180_Y182	CST	9211	50	V - validated	Rabbit
p53	CST	9282	2500	V - validated	Rabbit
p70S6K	Epitomics	1494-1	750	V - validated	Rabbit
p70S6K_pT389	CST	9205	50	V - validated	Rabbit
p90RSK_pT359_S363	CST	9344	50	C- use with caution	Rabbit
PARP_cleaved	CST	9546	50	C- use with caution	Mouse
Paxillin	Epitomics	1500-1	500	V - validated	Rabbit
PCNA	Abcam	ab29	500	V - validated	Mouse
PDK1_pS241	CST	3061	500	V - validated	Rabbit
PI3K-p110-alpha	CST	4255	150	C- use with caution	Rabbit
PI3K-p85	Millipore	06-195	10000	V - validated	Rabbit
PKC-alpha	Millipore	05-154	1000	V - validated	Mouse
PKC-alpha_pS657	Millipore	06-822	750	V - validated	Rabbit
PR	Epitomics	1483-1	200	V - validated	Rabbit
PRAS40_pT246	Biosource	441100G	1000	V - validated	Rabbit
PTCH	SDI	2113.00.02	1000	C- use with caution	Rabbit
PTEN	CST	9552	750	V - validated	Rabbit
Rab11	CST	3539	250	V - validated	Rabbit
Rab25	Covance	Custom	2500	C- use with caution	Rabbit
Rad50	Millipore	05-525	100	C- use with caution	Mouse
Rad51	Chem Biotech	na 71	125	C- use with caution	Mouse
Rb	CST	9309	200	V - validated	Mouse
Rb_pS807_S811	CST	9308	250	V - validated	Rabbit
S6_pS235_S236	CST	2211	3000	V - validated	Rabbit
S6_pS240_S244	CST	2215	2000	V - validated	Rabbit
Shc_pY317	CST	2431	50	NA	Rabbit
Smac	CST	2954	250	V - validated	Mouse
Smad1	Epitomics	1649-1	250	V - validated	Rabbit
Smad3	Epitomics	1735-1	750	V - validated	Rabbit
Smad4	Santa Cruz	sc-7866	50	V - validated	Mouse
Snail	CST	3895	100	C- use with caution	Mouse
Src	Millipore	05-184	200	V - validated	Mouse
Src_pY416	CST	2101	125	C- use with caution	Rabbit
Src_pY527	CST	2105	250	V - validated	Rabbit
STAT3_pY705	CST	9131	50	V - validated	Rabbit
STAT5-alpha	Epitomics	1289-1	250	V - validated	Rabbit
Stathmin	Epitomics	1972-1	150	V - validated	Rabbit
Syk	Santa Cruz	sc-1240	250	V - validated	Mouse
Tau	Millipore	05-348	100	C- use with caution	Mouse
TAZ_pS89	Santa Cruz	sc-17610	75	C- use with caution	Rabbit
Tuberin	Epitomics	1613-1	500	C- use with caution	Rabbit
VASP	CST	3112	50	C- use with caution	Rabbit
VEGFR2	CST	2479	5000	V - validated	Rabbit
XIAP	CST	2042	50	C- use with caution	Rabbit
XRCC1	CST	2735	50	C- use with caution	Rabbit
YAP	Santa Cruz	sc-15407	300	V - validated	Rabbit
YAP_pS127	CST	4911	350	C- use with caution	Rabbit
YB-1	SDI	1725.00.02	200	V - validated	Rabbit
YB-1_pS102	CST	2900	150	V - validated	Rabbit

Add. Table 2: Metabolite integrals for each of the samples included in the study

Sample number	Metabolic cluster	Beta-D-glucose	Ascorbate	Lactate	L-Tyrosine	Glycine	Myoinositol	Taurine	Scylloinositol	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Acetate	Alanine
MCOS-1	Mc3	65.0	25.6	238.7	349.7	130.7	241.8	314.1	60.1	121.1	308.4	88.8	121.7	48.9	201.4	79.6	281.0	39.9	75.7
MCOS-2	Mc1	39.4	26.3	186.7	374.2	145.4	135.3	186.8	36.5	274.4	624.2	149.2	116.7	36.1	139.1	56.4	221.9	42.5	70.7
MCOS-3	Mc3	12.5	50.1	337.1	396.7	544.6	143.0	484.9	49.8	94.0	220.5	121.5	165.8	93.7	131.8	57.0	228.2	50.6	105.0
MCOS-4	Mc2	160.4	20.4	172.3	245.8	108.5	238.4	324.7	54.2	76.9	114.9	79.4	54.4	30.8	127.2	36.4	335.8	67.0	54.3
MCOS-5	Mc3	8.6	57.4	367.0	401.6	128.5	144.2	536.0	53.4	251.3	391.5	119.4	156.9	52.5	145.1	61.9	281.0	47.6	88.7
MCOS-6	Mc2	122.7	17.3	220.2	183.5	156.1	169.6	217.8	53.4	54.2	125.9	45.8	62.1	30.0	109.4	44.7	345.9	56.8	49.9
MCOS-7	Mc3	48.8	38.9	436.3	349.7	360.3	133.3	465.5	38.0	191.7	263.8	56.0	162.6	53.8	181.0	74.8	270.9	36.4	120.5
MCOS-8	Mc1	43.8	31.8	242.2	371.8	136.7	153.4	408.9	77.8	310.5	721.9	62.7	79.8	42.7	151.4	43.9	133.0	29.5	50.4
MCOS-9	Mc1	22.9	15.6	414.1	374.3	228.8	154.1	207.5	51.6	324.5	619.1	149.1	112.9	55.5	133.9	63.7	242.2	53.8	105.7
MCOS-10	Mc3	46.5	25.4	281.1	388.9	175.2	303.3	376.7	64.2	178.3	261.5	129.9	148.0	47.3	183.4	71.1	277.7	45.3	65.8
MCOS-11	Mc3	37.2	58.0	284.5	350.9	109.4	189.7	315.0	58.1	133.9	418.7	83.9	86.8	49.6	101.0	50.7	251.2	59.3	67.9
MCOS-12	Mc3	29.2	45.5	259.1	473.9	167.2	153.7	418.4	47.8	213.1	355.4	172.4	112.9	60.6	140.1	55.6	313.2	40.3	116.4
MCOS-13	Mc2	81.0	31.7	166.6	334.1	190.2	205.0	303.5	63.7	110.2	243.4	118.5	109.8	43.6	175.4	65.7	315.6	55.4	103.7
MCOS-14	Mc1	34.8	41.8	242.2	375.5	267.8	167.7	365.6	54.4	150.8	481.6	115.9	118.9	50.6	132.1	56.0	246.6	57.3	73.9
MCOS-15	Mc3	42.6	23.1	281.2	408.3	185.8	299.6	367.5	62.4	151.1	245.7	146.6	147.6	39.0	180.1	75.3	279.7	51.3	74.9
MCOS-16	Mc2	237.9	17.2	127.1	243.2	208.2	251.1	494.4	41.3	105.6	295.5	183.2	79.8	23.4	168.7	69.5	190.2	44.3	56.4
MCOS-17	Mc3	9.5	19.8	309.3	404.9	222.2	181.6	479.0	63.9	179.4	290.3	112.0	88.0	52.7	138.6	33.0	284.8	33.0	84.6
MCOS-18	Mc1	12.5	19.6	198.6	330.7	117.1	89.7	227.1	31.3	211.4	1040.4	115.1	370.7	59.0	144.5	59.6	181.6	15.9	51.3
MCOS-19	Mc2	11.4	26.9	263.7	475.2	139.0	105.6	245.3	34.6	198.2	256.0	111.6	102.7	55.5	138.5	48.6	266.7	36.2	93.8
MCOS-20	Mc3	25.2	31.2	290.4	418.6	311.8	217.4	281.4	68.9	192.7	356.7	79.8	109.5	75.3	102.6	56.5	359.9	28.3	102.5
MCOS-21	Mc1	32.5	55.3	203.0	345.7	186.5	280.9	390.6	100.6	188.3	625.4	143.9	218.3	62.9	129.3	52.6	185.2	36.8	59.4
MCOS-22	Mc3	38.8	69.0	303.7	419.8	224.1	325.0	283.0	97.2	56.7	234.4	123.4	132.8	70.1	166.0	70.7	290.5	38.3	119.6
MCOS-23	Mc1	23.1	46.0	283.3	435.5	346.2	222.2	267.9	87.4	98.3	450.3	121.5	133.9	30.1	134.3	48.6	203.9	37.6	86.2
MCOS-24	Mc2	100.1	33.4	281.5	307.7	164.7	208.9	397.4	77.8	119.0	191.1	125.1	110.1	54.9	158.5	54.9	238.8	59.2	90.3
MCOS-25	Mc3	40.9	31.5	251.3	304.7	106.3	156.0	389.9	67.3	152.1	323.1	136.0	114.7	60.0	111.4	71.3	246.6	56.0	83.7
MCOS-26	Mc1	46.8	55.7	249.8	365.9	169.9	215.1	293.9	69.9	367.6	548.4	217.5	92.1	53.7	143.5	65.9	252.9	40.6	82.0
MCOS-27	Mc1	71.1	34.7	271.4	386.5	408.8	185.1	214.0	52.9	108.3	463.2	118.7	107.8	45.1	166.8	78.1	200.3	37.6	80.4
MCOS-28	Mc1	191.5	31.6	202.7	321.8	448.3	217.7	262.0	53.3	261.3	525.1	131.5	135.3	45.7	163.6	83.5	225.5	30.0	84.0
MCOS-29	Mc3	11.9	32.9	546.0	480.3	246.9	157.1	374.9	66.3	211.4	267.4	84.9	182.5	54.4	195.8	61.7	278.1	47.9	153.8
MCOS-30	Mc1	26.5	89.5	267.3	356.8	138.9	141.9	263.0	58.6	308.4	474.3	136.0	278.3	67.2	122.8	52.8	213.0	33.0	68.6
MCOS-31	Mc1	15.6	83.1	434.3	453.0	216.8	123.2	145.1	56.2	220.5	479.5	59.2	167.5	92.8	223.6	97.7	222.3	22.0	175.7
MCOS-32	Mc3	63.2	30.1	166.8	447.8	159.4	199.6	411.5	54.5	250.2	294.2	161.6	79.3	42.4	134.8	63.3	229.5	58.4	74.8
MCOS-33	Mc1	20.2	40.6	318.8	380.5	256.4	133.0	348.7	58.3	321.9	556.9	102.7	108.5	63.4	119.5	58.0	195.7	22.2	83.5
MCOS-34	Mc3	42.9	50.0	296.2	442.0	291.9	233.7	375.7	88.2	171.6	329.2	181.8	95.9	41.5	105.7	58.8	237.7	35.4	75.4
MCOS-35	Mc3	33.0	50.7	347.6	377.4	157.0	279.8	382.4	101.8	189.2	364.5	149.9	118.9	51.8	170.8	57.6	200.5	43.3	119.8
MCOS-36	Mc3	13.6	42.7	345.0	330.0	234.3	191.7	407.7	110.7	136.3	324.1	60.6	124.3	48.2	265.9	84.6	398.0	29.6	172.0
MCOS-37	Mc3	16.7	43.8	180.6	404.9	103.0	195.0	369.9	47.5	174.5	313.0	123.9	88.8	53.7	79.6	45.1	299.0	45.8	50.2
MCOS-38	Mc2	48.6	41.8	238.8	430.5	136.2	189.8	455.1	39.5	142.2	221.4	182.1	120.4	52.1	99.0	33.8	277.0	54.9	63.7
MCOS-39	Mc2	34.1	23.7	299.1	405.8	134.8	200.6	218.1	67.2	142.4	308.0	101.4	97.3	54.1	135.6	44.2	184.1	24.9	59.8
MCOS-40	Mc3	46.0	37.4	373.6	365.6	201.4	129.9	466.5	65.0	135.4	431.9	123.2	156.8	47.4	122.7	51.9	299.0	33.6	77.3
MCOS-41	Mc3	40.4	41.8	227.5	392.1	176.9	217.4	401.2	58.1	178.0	367.0	213.7	142.7	51.5	151.8	59.9	239.2	51.5	85.1
MCOS-42	Mc1	14.3	65.8	223.6	544.3	117.6	106.8	637.1	38.9	219.2	683.9	200.8	117.6	57.2	86.3	28.6	229.6	28.8	60.4
MCOS-43	Mc3	21.7	40.1	329.4	374.5	123.0	139.8	673.9	61.4	185.7	276.5	162.3	125.3	51.9	125.7	59.6	293.7	36.2	100.3
MCOS-44	Mc2	31.2	26.0	268.8	358.7	192.8	149.5	210.7	47.8	62.9	223.5	123.8	67.7	35.5	103.3	44.5	437.3	90.3	48.9
MCOS-45	Mc1	23.4	34.2	311.7	512.2	143.6	170.0	260.9	49.1	304.6	492.4	75.8	145.9	61.4	134.6	37.6	208.2	19.1	144.4
MCOS-46	Mc3	33.7	33.1	201.3	453.1	92.1	200.6	300.2	59.1	194.8	297.1	121.3	99.6	48.4	97.0	57.8	204.9	33.4	56.7
MCOS-47	Mc3	35.0	35.1	305.7	457.6	144.4	276.0	242.1	80.8	129.6	222.4	105.6	103.9	47.7	127.0	52.6	245.0	37.9	78.8
MCOS-48	Mc3	14.2	26.9	304.3	343.1	151.4	127.4	338.6	32.8	49.3	388.9	51.4	171.7	95.8	126.2	34.0	224.7	33.6	78.7
MCOS-49	Mc3	12.9	29.2	527.1	431.3	187.1	111.4	289.9	49.6	56.2	351.8	49.7	73.3	58.9	136.3	50.3	449.3	30.5	197.6
MCOS-50	Mc1	15.9	35.6	244.7	408.3	141.6	130.8	351.7	39.1	193.7	717.7	239.9	157.4	82.6	87.1	37.0	349.2	34.8	103.5
MCOS-51	Mc3	11.8	45.3	420.4	359.6	312.7	141.6	345.7	43.3	162.6	429.8	89.3	297.0	61.6	220.0	128.9	274.8	39.0	173.2
MCOS-52	Mc3	51.2	43.4	275.5	413.0	187.2	186.6	606.7	78.2	212.9	247.9	146.4	139.4	69.3	114.4	44.0	285.9	25.4	89.4
MCOS-53	Mc2	72.5	41.0	214.7	440.7	123.5	220.2	256.2	58.4	75.7	167.9	87.9	89.9	38.8	107.7	42.8	182.2	36.0	53.5
MCOS-54	Mc1	17.5	43.6	180.2	419.9	186.5	126.7	468.9	61.5	194.8	811.3	147.3	283.9	65.4	100.4	47.8	157.4	27.0	41.9
MCOS-55	Mc2	88.2	39.1	171.1	365.9	155.9	259.9	396.3	88.0	110.7	221.6	181.6	144.6	59.4	148.4	65.0	313.8	37.1	81.0
MCOS-56	Mc1	11.1	31.5	242.9	468.7	193.3	130.0	401.8	41.7	195.7	596.7	226.0	146.5	59.1	144.8	47.8	318.3	44.2	158.9
MCOS-57	Mc3	26.8	30.7	281.2	523.3	234.0	162.1	245.0	76.9	129.4	274.2	233.2	169.1	49.2	214.1	72.4	327.3	53.0	159.1
MCOS-58	Mc3	15.8	43.3	391.1	412.5	175.0	126.7	478.5	41.6	159.0	398.6	76.4	390.3	66.5	178.4	48.4	263.2	28.6	123.7
MCOS-59	Mc3	22.9	39.0	416.4	410.9	183.4	182.5	462.7	43.8	131.6	335.1	78.5	204.4	45.6	133.8	55.4	302.9	31.5	168.0
MCOS-60	Mc3	48.3	38.7	185.0	442.5	149.9	286.7	407.9	37.2	172.2	280.4	165.4	76.1	41.1	122.4	58.3	317.6	33.9	53.6
MCOS-61	Mc3	19.5	20.7	338.9	382.1	275.4	221.8	302.8	111.9	180.1	271.4	158.5	127.6	30.2	200.9	99.8	365.2	42.1	141.2
MCOS-62	Mc3	19.0	40.7	285.4	423.9	262.5	176.8	329.6	51.9	163.7	430.1	167.0	120.8	26.7	152.9	72.6	305.4	38.8	91.8
MCOS-63	Mc3	35.0	43.5	240.6	440.3	184.0	293.6	492.9	89.7	151.5	328.9	181.9	106.4	57.2	126.8	69.2	210.5	29.0	55.2
MCOS-64	Mc2	112.1	50.3	175.4	306.2	219.8	214.8	424.2	63.6	123.6	282.6	132.5	100.3	46.9	134.5	54.3	219.4	41.8	80.3
MCOS-65	Mc3	30.6	76.6	281.3	421.6	185.9	226.7	391.4	76.0	152.2	281.1	178.7	231.0	59.0	204.6	62.3	276.4	25.6	115.8
MCOS-66	Mc1	12.8	62.9	334.8	410.2	153.4	163.2	331.1	57.4	282.9	456.6	156.9	229.5	52.1	209.7	57.1	216.6	26.6	103.4
MCOS-67	Mc3	60.9	29.5	271.7	326.1	128.5	305.2	323.5	91.4	196.3	316.1	140.5	110.2	51.0	115.4	47.2	206.6	51.5	87.4
MCOS-68	Mc2	27.6	21.8	340.5	537.7	177.6	253.9	144.0	83.0	119.7	405.1	89.0	64.7	39.0	124.1	54.5	313.0	29.7	111.2
MCOS-69	Mc2	28.0	32.3	371.4	372.8	187.8	277.6	420.2	72.4	88.9	185.5	130.6	130.7	52.6	183.6	74.4	359.6	50.7	106.7
MCOS-70	Mc3	73.0	19.9	538.6	236.5	134.6	133.5	289.0	61.9	65.3	91.5	87.9	84.3	23.4	349.6	123.9	389.9	66.1	82.5
MCOS-71	Mc1	5.9	51.9	162.5	359.1	153.8	140.8	248.0	27.5	96.8	601.8	107.4	123.6	85.8	149.5	48.8	207.1	28.2	86.2
MCOS-72	Mc3	26.7	45.0	306.2	426.6	155.2	296.8	444.0	124.4	157.8	367.9	118.7	120.9	61.3	135.2	55.5	338.3	29.6	118.1
MCOS-73	Mc3	18.0	59.4	424.4	448.1	303.3	178.5	231.7	36.5	182.3	352.9	56.1	100.5	69.6	232.1	79.6	285.2	39.3	190.6
MCOS-74	Mc2	83.8	42.8	236.8	366.7	181.8	217.6	417.0	85.2	92.2	209.2	191.4	146.7	48.4	139.4	54.2	240.1	45.9	135.5
MCOS-75	Mc2	119.7	30.1	261.0	275.0	123.7	164.1	374.9	44.9	96.7	133.9	94.2	115.5	54.2	163.0	57.2	207.8	61.0	89.3
MCOS-76	Mc3	15.8	37.6	539.9	410.7	370.0	179.8	245.6	68.7	151.7	320.1	95.3	141.2	59.5	207.2	70.5	250.1	39.4	110.0
MCOS-77	Mc3	38.8	48.1	335.8	336.7	149.2	224.8	358.6	61.5	176.4	245.7	139.7	118.2	52.9	105.9	45.3	298.8	47.7	79.7
MCOS-78	Mc1	12.5	42.0	288.9	430.2	260.0	227.5	375.6	46.5	384.6	422.0	181.8	145.6	57.5	196.6	66.2	252.2	28.8	103.5
MCOS-79	Mc3	27.9	37.9	298.2	417.3	153.7	179.3	363.4	51.5	177.0	462.2	112.7	108.1	72.8	151.9	56.8	309.6	23.0	100.4
MCOS-80	Mc2	199.7	20.5	100.2	217.8	123.3	242.5	410.9	66.4	97.0	175.6	135.1	66.4	42.0	153.1	43.6	165.3	42.1	29.3
MCOS-81	Mc2	39.0	25.9	222.1	382.0	108.6	244.7	416.5	60.8	92.1	183.2	119.5	64.5	70.0	104.7	45.6	259.6	33.1	31.6
MCOS-82	Mc2	30.0	31.2	242.0	438.6	227.0	169.4	352.8	44.8	70.4	172.0	129.5	167.8	72.1	160.9	73.6	363.2	39.6	80.4

Sample number	Metabolic cluster	Beta-D-glucose	Ascorbate	Lactate	L-Tyrosine	Glycine	Myoinositol	Taurine	Scylloinositol	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Acetate	Alanine
MCOS-83	Mc2	50.4	24.5	233.5	440.3	146.7	281.0	358.6	70.6	133.4	241.9	128.9	80.9	63.4	128.4	54.5	219.3	27.7	81.6
MCOS-84	Mc2	49.4	43.1	259.0	373.0	169.8	338.6	431.8	103.3	104.8	227.0	183.2	119.4	61.0	119.3	56.5	247.6	43.1	79.3
MCOS-85	Mc3	29.4	31.5	326.1	403.5	214.4	163.1	227.2	40.3	47.4	257.6	90.6	84.9	68.8	149.0	64.6	316.4	46.9	105.7
MCOS-86	Mc3	20.1	51.8	340.2	528.6	236.3	156.9	289.0	50.2	82.1	292.2	168.5	172.2	68.0	184.7	59.1	248.9	50.5	155.2
MCOS-87	Mc3	19.2	23.4	342.7	381.0	179.6	208.4	333.0	43.5	148.9	362.9	155.8	91.8	50.2	174.5	78.3	294.0	35.3	137.3
MCOS-88	Mc3	13.1	51.0	256.1	549.7	198.9	147.9	331.3	44.7	150.1	324.9	110.2	150.4	63.3	166.1	63.7	252.3	41.6	94.8
MCOS-89	Mc3	39.9	49.9	295.5	403.2	169.6	257.7	347.4	66.2	199.4	333.6	159.1	81.2	48.3	128.9	54.4	227.5	44.6	78.2
MCOS-90	Mc2	46.9	24.0	257.9	457.2	169.6	264.1	327.0	85.4	76.3	209.5	88.0	101.3	80.7	179.4	58.9	259.4	21.9	58.0
MCOS-91	Mc3	44.4	28.9	233.9	368.5	170.6	237.5	313.7	61.7	189.9	267.0	130.5	197.7	57.5	148.1	59.1	308.7	33.6	107.9
MCOS-92	Mc3	23.1	41.3	324.6	408.9	188.8	211.0	383.7	76.8	139.1	265.2	116.3	124.8	52.9	178.9	81.0	246.4	45.7	98.5
MCOS-93	Mc3	18.0	34.6	218.9	390.8	264.3	113.4	299.5	41.1	96.0	326.1	135.3	199.7	112.5	197.5	72.7	259.4	43.1	110.2
MCOS-94	Mc3	42.3	55.8	173.1	407.6	130.7	157.0	363.3	40.2	66.7	411.5	122.8	107.5	62.0	167.2	79.3	274.3	47.3	122.5
MCOS-95	Mc3	95.8	75.7	206.6	426.1	84.1	196.4	466.0	149.7	161.3	363.3	108.1	96.0	44.9	88.5	45.2	212.5	28.3	45.0
MCOS-96	Mc3	52.4	41.8	235.5	435.6	104.1	214.3	667.0	58.3	80.0	377.6	157.9	96.7	87.0	110.6	48.0	239.1	25.2	59.4
MCOS-97	Mc2	115.6	27.0	168.9	330.9	175.9	351.9	332.8	91.3	106.8	215.1	133.2	90.2	40.0	127.5	48.1	247.2	61.4	62.7
MCOS-98	Mc1	24.4	40.6	329.3	399.9	153.9	109.1	481.0	58.9	87.7	481.2	92.0	120.0	64.8	105.6	39.2	195.5	26.9	77.5
MCOS-99	Mc1	20.9	50.6	234.5	402.7	113.2	258.1	449.8	81.6	359.7	609.1	185.8	151.8	41.4	88.8	33.1	178.7	25.8	80.5
MCOS-100	Mc3	31.1	21.2	353.9	331.6	212.5	125.2	226.4	51.3	116.2	274.8	102.7	182.2	67.3	129.4	84.1	445.6	54.3	124.3
MCOS-101	Mc3	34.6	28.1	279.4	329.9	218.0	192.3	232.5	37.7	129.3	291.3	148.4	98.7	57.0	107.0	70.5	371.0	48.6	112.4
MCOS-102	Mc3	64.9	41.3	361.2	390.3	225.8	145.7	491.2	42.2	144.2	233.1	151.3	122.0	41.6	142.9	56.0	205.2	56.3	95.9
MCOS-103	Mc3	48.6	54.5	281.2	404.3	218.4	160.8	327.8	42.5	243.8	429.4	236.4	125.9	44.8	165.4	46.2	192.9	42.1	65.5
MCOS-104	Mc1	46.3	49.6	224.0	365.9	111.8	156.3	281.9	47.2	279.0	291.8	127.9	128.1	59.8	125.5	63.7	236.8	54.9	71.3
MCOS-105	Mc3	38.1	20.6	321.7	338.0	210.1	132.0	222.5	49.6	96.7	255.2	107.4	175.5	61.5	131.8	82.0	427.1	56.8	123.2
MCOS-106	Mc3	43.2	69.8	243.3	391.5	207.3	205.1	230.1	80.7	121.1	389.7	104.3	151.2	62.0	127.8	39.7	188.9	46.2	79.7
MCOS-107	Mc1	40.1	94.2	161.4	487.8	128.6	217.6	209.9	70.6	163.9	497.8	125.6	290.4	84.7	130.1	50.7	184.6	30.6	45.1
MCOS-108	Mc3	27.4	40.8	423.6	356.2	170.6	181.4	401.3	49.8	157.1	424.4	132.7	132.3	52.4	173.3	59.2	238.3	37.6	63.8
MCOS-109	Mc3	13.1	56.2	430.5	488.6	287.5	219.0	366.6	79.8	160.5	277.5	126.5	102.5	70.5	185.0	58.2	285.3	36.3	122.7
MCOS-110	Mc1	44.3	40.7	280.4	420.9	219.4	154.8	372.3	61.3	114.6	474.1	161.8	249.7	58.0	178.5	75.3	239.5	36.2	126.5
MCOS-111	Mc3	49.3	39.8	314.8	422.2	158.7	302.8	259.1	108.7	174.3	312.4	135.9	132.4	69.3	58.4	50.9	263.6	20.7	149.3
MCOS-112	Mc3	50.0	50.1	228.9	461.0	191.0	178.5	392.4	43.4	232.4	424.0	146.0	162.1	60.8	132.4	62.6	254.4	41.8	91.8
MCOS-113	Mc1	17.8	28.2	306.4	475.4	167.6	164.1	266.8	58.8	88.9	624.1	119.6	162.9	70.9	109.9	61.2	256.4	26.8	53.7
MCOS-114	Mc3	19.8	68.2	276.0	460.9	320.5	193.1	298.7	73.7	126.4	294.6	159.6	227.1	71.5	211.3	53.2	290.5	49.7	117.7
MCOS-115	Mc1	29.6	42.6	249.8	431.6	288.4	137.1	399.8	49.1	148.0	543.0	93.8	192.0	79.8	210.3	79.2	224.3	30.7	143.2
MCOS-116	Mc2	89.6	45.3	191.1	300.0	123.1	249.0	323.5	80.9	58.0	255.0	77.5	84.4	61.2	148.7	61.4	281.3	47.7	71.6
MCOS-117	Mc3	30.3	41.7	277.4	320.9	88.4	189.7	542.2	78.6	166.0	403.0	131.6	107.0	55.7	119.6	54.9	223.5	37.4	82.3
MCOS-118	Mc1	11.4	29.6	328.8	419.5	168.2	181.9	321.4	54.7	293.1	538.8	222.4	149.2	56.8	83.1	50.1	388.6	31.5	136.3
MCOS-119	Mc1	15.3	28.3	284.1	441.5	137.4	165.5	345.9	74.4	154.0	579.6	129.2	104.8	73.8	185.4	82.8	355.4	29.3	105.4
MCOS-120	Mc1	33.6	31.8	234.9	412.3	178.3	173.4	263.1	54.2	338.0	335.6	169.1	149.0	79.6	140.7	42.8	247.2	34.8	64.3
MCOS-121	Mc3	25.3	70.9	262.7	464.2	146.5	123.0	454.6	58.8	175.3	341.6	169.4	132.8	85.0	159.4	64.1	225.9	36.0	73.1
MCOS-122	Mc1	11.1	69.4	301.5	383.8	126.4	75.0	606.4	50.2	160.2	615.1	63.6	178.4	58.5	200.5	54.2	199.9	24.2	73.2
MCOS-123	Mc3	41.5	44.0	242.2	430.3	102.1	151.4	408.0	37.8	223.1	379.8	93.7	114.6	55.5	134.2	60.4	206.5	39.4	68.5
MCOS-124	Mc3	11.7	26.3	353.2	327.6	182.2	122.6	553.7	49.2	155.7	339.8	153.0	273.6	67.3	128.1	52.0	254.4	42.1	70.8
MCOS-125	Mc3	44.2	31.4	222.7	385.9	218.5	299.3	380.7	43.1	88.5	328.0	193.1	180.8	58.4	135.2	50.2	282.1	36.3	80.3
MCOS-126	Mc3	42.5	44.1	264.7	477.1	96.4	139.6	473.9	62.8	153.9	498.5	104.8	241.1	54.7	114.3	54.0	149.0	26.9	89.4
MCOS-127	Mc3	26.5	35.7	335.8	601.5	244.8	181.2	327.4	51.1	239.1	298.5	245.6	51.3	64.0	137.6	52.3	297.2	29.9	163.0
MCOS-128	Mc3	41.0	33.1	271.4	405.5	148.7	206.2	336.4	57.1	98.3	272.3	146.5	78.0	62.4	131.8	61.4	242.0	38.4	97.5
MCOS-129	Mc2	9.9	31.0	285.1	424.2	171.5	202.7	422.1	49.9	94.7	213.0	187.8	173.9	61.2	172.9	68.4	248.8	36.8	53.3
MCOS-130	Mc2	46.7	28.0	243.9	321.1	188.6	168.4	381.2	51.1	97.8	142.7	123.6	81.4	57.5	113.5	47.5	426.3	68.7	120.6
MCOS-131	Mc2	201.1	23.5	179.6	188.9	151.4	225.7	378.3	56.3	99.8	80.9	40.1	49.7	31.6	139.1	40.8	243.3	61.6	84.2
MCOS-132	Mc1	39.5	45.0	200.5	359.8	126.2	210.2	339.7	101.9	320.8	309.9	158.3	122.6	47.1	99.4	37.1	284.1	44.2	61.8
MCOS-133	Mc3	17.1	32.0	455.9	404.3	190.5	170.9	266.4	59.6	129.8	300.9	143.3	167.0	71.7	165.6	68.3	335.8	49.2	165.9
MCOS-134	Mc2	49.6	27.3	270.0	293.1	174.3	182.8	212.2	59.1	109.3	210.8	66.4	55.8	59.1	108.5	43.3	232.1	35.3	32.2
MCOS-135	Mc3	58.6	45.3	307.8	359.0	192.7	156.8	547.9	66.2	136.2	276.0	137.5	193.6	70.4	137.5	71.3	228.8	58.9	94.5
MCOS-136	Mc1	11.8	37.5	355.7	532.7	209.1	121.4	379.3	52.5	178.5	798.7	111.4	110.6	70.2	129.0	76.7	267.2	26.9	155.0
MCOS-137	Mc3	22.9	51.2	323.4	458.0	271.4	148.3	471.7	45.7	144.6	310.0	85.1	97.0	70.6	121.9	43.4	312.6	31.5	102.7
MCOS-138	Mc2	69.3	25.9	177.6	382.3	132.8	239.3	307.9	77.5	133.0	233.0	104.8	62.5	30.1	114.4	47.3	275.7	40.6	42.8
MCOS-139	Mc3	63.8	29.3	218.3	293.8	205.7	204.4	314.3	52.4	84.8	359.4	128.4	95.1	54.6	165.7	51.1	273.7	57.3	95.5
MCOS-140	Mc1	26.7	38.9	222.7	387.4	343.8	237.2	179.0	90.4	185.6	502.4	105.0	79.6	53.6	142.9	57.7	249.6	31.9	89.3
MCOS-141	Mc3	19.2	25.0	273.3	496.5	127.3	172.6	355.6	45.1	224.1	332.2	213.2	102.4	72.0	86.5	51.3	243.4	37.9	59.0
MCOS-142	Mc1	26.9	61.0	252.6	497.4	149.1	222.3	436.9	61.2	379.6	380.0	151.8	147.4	54.5	102.5	54.5	212.8	28.8	58.9
MCOS-143	Mc3	28.7	27.7	265.9	400.6	243.0	199.9	281.8	42.9	154.2	301.4	91.1	157.5	68.6	133.2	59.7	337.5	21.9	112.0
MCOS-144	Mc1	21.5	35.7	224.0	342.1	86.3	117.3	509.3	37.5	105.4	805.9	75.6	139.6	53.9	57.4	60.4	119.3	12.7	50.9
MCOS-145	Mc3	17.1	37.3	276.1	489.7	152.5	200.0	368.3	47.3	164.6	422.1	213.9	187.9	71.4	147.4	63.8	215.7	33.9	68.1
MCOS-146	Mc2	23.2	32.9	265.0	476.7	73.1	107.6	341.4	40.4	127.7	306.2	97.9	109.9	70.2	109.4	53.6	235.9	26.3	53.2
MCOS-147	Mc3	20.8	28.4	249.7	531.5	191.3	211.9	344.6	62.0	110.1	434.9	154.5	99.3	66.9	157.6	66.4	325.4	23.6	87.4
MCOS-148	Mc3	41.6	43.4	272.6	320.9	317.3	204.0	382.5	47.9	81.1	272.6	104.0	199.1	70.5	159.9	63.5	277.2	41.0	81.4
MCOS-149	Mc2	95.8	31.9	196.9	319.0	257.3	294.0	339.2	64.2	72.7	158.5	191.7	113.6	44.4	154.0	53.5	291.9	36.9	79.9
MCOS-150	Mc2	116.3	30.2	149.2	310.9	121.1	289.0	313.5	42.8	88.6	202.5	100.4	83.0	42.9	119.7	45.8	271.3	56.9	47.4
MCOS-151	Mc1	9.6	37.2	219.4	347.5	138.2	213.6	189.1	37.4	62.8	510.8	66.7	253.9	76.8	83.9	35.4	173.6	18.2	54.4
MCOS-152	Mc2	17.5	17.1	239.5	347.7	174.3	219.4	232.3	51.0	231.0	278.5	124.4	81.9	50.5	92.1	38.2	320.9	31.1	77.8
MCOS-153	Mc3	11.8	77.9	339.4	470.4	153.3	175.5	507.8	45.1	131.5	432.6	136.3	106.7	69.4	129.1	51.7	255.8	22.2	93.9
MCOS-154	Mc1	55.4	19.4	343.6	398.4	145.6	157.0	242.7	52.5	74.5	498.0	85.0	59.3	40.4	89.3	50.0	236.0	34.6	67.1
MCOS-155	Mc1	18.5	23.4	187.3	422.2	135.2	210.4	308.6	46.8	355.8	635.9	141.8	87.9	29.6	128.1	73.5	405.0	29.7	67.4
MCOS-156	Mc1	48.0	52.2	182.7	343.9	147.1	232.5	535.7	35.0	118.5	490.9	131.3	181.9	65.7	104.5	45.7	224.4	33.4	37.2
MCOS-157	Mc2	220.3	19.2	132.1	204.4	125.4	233.0	395.7	61.7	114.5	131.3	61.8	50.5	70.2	120.8	32.7	139.5	77.7	32.8
MCOS-158	Mc1	9.2	30.0	240.0	424.7	355.4	157.7	384.2	52.3	147.4	504.6	167.6	137.6	48.4	181.2	87.4	390.6	36.5	100.0
MCOS-159	Mc1	17.5	56.2	198.7	479.8	94.3	146.3	594.8	59.4	144.9	611.7	228.0	136.7	41.3	93.5	51.5	171.0	33.4	52.8
MCOS-160	Mc3	20.6	22.8	345.7	428.0	181.3	161.4	270.0	43.5	209.2	394.7	153.2	124.5	58.6	90.1	66.8	419.8	38.1	120.1
MCOS-161	Mc2	77.1	26.6	190.2	351.5	102.6	150.8	313.2	29.5	106.1	207.4	91.5	62.9	40.4	114.7	63.7	169.1	64.7	39.8
MCOS-162	Mc1	14.0	14.2	330.9	547.8	202.6	142.0	144.7	40.2	304.5	515.2	168.3	78.6	49.7	59.5	51.1	283.6	26.0	159.5
MCOS-163	Mc2	56.9	29.4	278.8	310.6	100.7	199.0	211.3	47.6	122.7	147.2	66.3	55.4	41.3	90.7	46.9	206.3	48.4	40.8
MCOS-164	Mc3	23.9	46.5	229.2	398.9	163.2	89.9	451.9	31.9	164.0	365.7	121.5	137.4	71.8	168.9	68.3	187.4	31.9	46.8
MCOS-165	Mc3	68.9	32.0	218.4	347.5	195.2	257.8	396.5	45.7	94.3	316.1	107.0	138.2	45.7	135.2	68.6	177.7	31.1	59.5

Sample number	Metabolic cluster	Beta-D-glucose	Ascorbate	Lactate	L-Tyrosine	Glycine	Myoinositol	Taurine	Scylloinositol	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Acetate	Alanine
MCOS-166	Mc2	13.9	41.1	344.4	429.7	211.8	272.4	492.5	50.5	84.4	206.4	144.5	131.6	67.2	154.1	72.0	294.8	33.5	71.3
MCOS-167	Mc2	19.0	22.0	332.1	468.3	190.5	186.8	284.9	39.6	128.4	288.6	108.7	93.8	63.6	125.6	52.6	274.0	27.4	62.0
MCOS-168	Mc3	26.4	41.5	303.0	527.6	240.6	190.4	347.8	44.8	172.7	362.7	154.6	64.0	78.2	108.3	47.1	378.1	22.3	61.4
MCOS-169	Mc2	64.6	25.4	314.0	362.9	148.6	232.3	379.7	65.8	81.7	171.6	103.1	144.0	66.0	219.8	65.9	226.9	35.0	101.4
MCOS-170	Mc2	54.6	35.3	204.8	368.6	127.2	202.9	354.3	42.9	82.1	165.1	125.5	53.6	46.2	88.6	40.6	357.9	63.3	39.9
MCOS-171	Mc1	20.6	19.8	492.2	367.9	369.4	137.3	170.8	56.4	82.7	481.6	126.1	89.8	57.9	192.6	66.1	299.0	39.7	132.1
MCOS-172	Mc3	19.5	36.6	245.3	456.9	198.0	182.2	512.8	47.5	141.0	242.7	98.4	264.3	63.8	122.5	51.9	257.3	25.3	54.7
MCOS-173	Mc2	30.6	47.4	272.3	408.7	141.0	208.7	524.1	52.7	127.0	254.6	184.6	132.8	63.0	91.7	43.5	350.8	38.9	91.1
MCOS-174	Mc3	27.4	24.1	237.8	490.0	163.5	310.5	403.7	49.8	250.7	434.1	127.4	103.3	50.7	103.4	51.0	219.6	26.4	68.0
MCOS-175	Mc1	87.6	37.8	389.6	455.4	235.8	73.1	286.6	27.3	271.8	395.9	76.3	190.1	75.1	199.2	85.2	281.3	39.9	133.9
MCOS-176	Mc3	42.1	51.7	205.5	418.1	115.4	320.0	429.9	71.7	122.9	331.6	170.5	101.0	57.6	125.5	80.0	240.4	30.1	72.1
MCOS-177	Mc2	55.2	27.9	260.7	481.4	209.0	175.7	209.8	50.6	63.3	232.9	127.6	102.7	50.1	174.0	72.3	252.5	48.0	77.4
MCOS-178	Mc2	200.5	14.1	158.6	172.3	93.5	223.7	350.6	60.3	55.5	107.4	63.6	47.1	34.0	144.5	44.4	212.9	102.8	38.1
MCOS-179	Mc3	47.8	27.6	446.9	268.3	230.7	133.8	210.7	60.6	68.6	107.6	129.4	168.0	41.1	162.2	66.9	361.4	124.8	91.6
MCOS-180	Mc1	25.0	22.6	384.6	393.1	166.3	169.8	266.8	50.0	128.6	551.1	151.8	96.6	55.1	96.0	50.9	255.7	38.0	80.1
MCOS-181	Mc3	26.7	36.4	225.4	435.3	258.1	308.2	347.2	53.0	79.7	315.2	109.6	121.9	74.5	237.7	64.8	259.1	30.3	71.9
MCOS-182	Mc2	35.2	33.9	231.9	496.8	167.8	139.2	224.5	34.3	167.5	304.5	95.2	99.6	78.0	218.9	64.3	205.6	30.4	63.1
MCOS-183	Mc3	37.0	32.6	213.4	385.7	144.4	274.5	294.4	51.9	265.3	376.7	184.1	115.0	50.4	143.1	62.0	232.6	47.9	70.4
MCOS-184	Mc3	50.7	15.9	294.9	535.2	238.9	209.0	177.9	50.6	118.4	301.8	201.8	181.4	23.2	126.7	69.1	246.5	45.3	64.6
MCOS-185	Mc3	17.1	23.3	391.5	448.0	212.7	171.4	359.9	47.2	127.1	235.2	258.7	95.8	54.4	80.4	68.0	288.3	54.8	143.5
MCOS-186	Mc2	55.5	31.8	250.4	370.7	187.8	259.2	397.7	74.8	103.9	255.9	171.6	83.5	27.6	114.6	45.3	328.8	36.9	89.2
MCOS-187	Mc3	34.3	47.3	249.3	388.8	95.1	140.0	512.0	67.6	220.0	510.9	111.5	162.4	50.1	94.0	62.5	223.7	34.2	46.0
MCOS-188	Mc2	64.4	48.2	138.1	375.0	141.3	350.2	270.3	73.8	107.4	283.7	145.1	95.5	49.1	138.0	69.4	288.0	44.3	55.8
MCOS-189	Mc3	18.9	30.9	272.6	363.1	213.7	185.7	500.5	52.6	105.1	266.7	76.1	212.5	53.2	98.5	47.5	269.1	39.1	57.7
MCOS-190	Mc1	6.4	36.0	245.1	447.8	152.2	119.5	645.5	40.7	151.2	552.2	112.2	290.5	63.8	107.6	66.7	191.4	28.4	52.1
MCOS-191	Mc3	58.0	31.0	342.9	344.8	217.0	184.8	231.4	76.9	131.0	231.4	70.1	113.7	57.8	167.3	59.8	265.7	48.7	101.4
MCOS-192	Mc3	14.0	29.2	395.8	383.3	168.8	225.2	469.9	52.2	178.2	366.6	165.2	104.0	47.3	94.6	33.3	259.5	26.9	83.4
MCOS-193	Mc1	7.1	25.3	172.8	366.4	190.8	150.5	275.9	32.3	179.4	608.6	184.4	85.8	40.8	200.6	70.3	291.6	48.4	43.0
MCOS-194	Mc1	75.0	29.9	177.4	335.0	97.5	129.1	381.9	45.1	191.5	389.9	131.6	125.8	51.3	131.3	73.2	167.7	30.2	37.4
MCOS-195	Mc2	32.8	21.0	257.7	329.5	73.4	120.1	238.1	57.5	147.7	181.5	112.9	58.4	42.8	91.8	57.6	259.0	55.4	70.4
MCOS-196	Mc1	31.4	32.2	160.4	320.9	144.7	167.8	247.2	46.9	243.1	453.8	106.7	87.4	58.0	138.4	46.2	212.0	44.0	41.6
MCOS-197	Mc2	48.1	29.0	213.4	431.4	124.4	151.9	327.5	64.9	146.4	239.0	137.9	64.6	76.9	122.2	64.6	238.7	53.9	56.4
MCOS-198	Mc3	24.9	22.7	351.9	317.6	252.1	166.6	377.5	52.1	129.3	311.5	107.1	107.3	13.8	150.0	88.5	255.2	61.6	69.8
MCOS-199	Mc2	20.5	22.3	215.6	372.2	152.0	257.8	338.0	59.1	156.6	256.1	148.7	82.3	58.1	129.9	54.4	346.1	40.1	61.0
MCOS-200	Mc1	51.9	45.8	203.7	349.1	156.3	103.2	246.9	65.6	237.5	511.5	51.4	104.5	48.5	160.3	67.7	236.6	26.1	57.8
MCOS-201	Mc3	54.6	49.2	329.4	356.7	112.2	172.3	442.7	83.7	138.6	427.4	96.9	111.3	48.7	102.3	63.1	167.3	42.2	82.7
MCOS-202	Mc2	70.5	19.9	247.7	324.1	150.1	212.7	277.9	55.8	91.1	160.1	136.7	67.7	46.8	116.3	50.9	284.5	54.8	44.2
MCOS-203	Mc3	11.7	33.4	273.6	352.9	198.5	178.9	311.2	37.1	146.9	315.3	157.8	214.3	57.2	204.9	84.6	224.1	33.5	90.0
MCOS-204	Mc3	23.0	22.5	240.7	349.1	156.2	191.5	325.1	53.9	166.1	343.2	152.8	87.1	67.0	110.2	49.9	401.1	34.7	51.6
MCOS-205	Mc3	21.7	20.2	288.6	388.4	155.4	225.3	314.7	47.5	102.9	360.6	107.9	143.5	52.4	113.8	42.5	220.1	24.5	53.9
MCOS-206	Mc3	14.1	21.2	200.0	393.5	206.1	201.1	275.8	55.3	107.7	253.5	155.1	122.3	81.0	130.7	62.6	466.6	52.5	48.7
MCOS-207	Mc2	21.9	21.8	295.3	354.4	118.8	197.0	259.3	42.4	105.9	174.0	112.7	79.7	39.3	128.5	65.0	369.8	73.0	33.2
MCOS-208	Mc2	14.4	18.1	175.1	280.1	121.2	243.9	272.1	1001.3	143.5	315.8	124.7	100.4	51.7	101.0	43.9	194.5	37.2	37.2
MCOS-209	Mc3	36.8	21.4	349.0	318.9	307.7	142.4	160.8	73.4	108.3	299.6	49.6	113.9	55.4	207.7	48.3	336.0	47.8	138.9
MCOS-210	Mc2	96.0	22.6	158.5	268.5	120.1	215.8	278.5	66.0	90.2	173.0	104.3	65.2	53.5	163.5	59.8	208.7	43.1	38.0
MCOS-211	Mc2	60.7	21.6	208.2	306.7	92.5	211.7	423.7	71.7	98.5	257.5	119.1	84.1	55.8	88.7	45.4	234.5	65.5	30.3
MCOS-212	Mc2	42.1	27.5	242.1	387.4	142.9	248.9	378.7	78.1	111.0	216.0	129.2	207.0	61.8	160.1	56.8	270.4	35.4	74.2
MCOS-213	Mc3	29.1	29.2	253.6	419.6	141.8	290.4	429.9	98.5	119.0	361.8	133.2	88.6	59.5	77.8	43.3	273.7	25.4	56.6
MCOS-214	Mc1	13.3	25.5	183.9	376.2	162.3	277.2	185.4	65.8	206.4	691.6	161.6	139.5	50.6	72.9	52.7	314.1	21.9	65.5
MCOS-215	Mc3	17.5	24.6	337.8	418.1	185.8	147.4	292.1	48.3	231.5	399.9	154.5	129.0	61.7	92.6	54.0	323.9	43.1	86.0
MCOS-216	Mc2	31.1	19.3	235.7	283.2	128.3	152.3	240.0	1133.8	136.1	387.5	99.8	93.7	32.5	138.3	56.2	241.7	38.0	71.0
MCOS-217	Mc3	18.9	21.6	282.6	331.9	171.4	211.6	274.0	83.1	102.7	295.0	116.8	150.3	62.1	172.6	84.0	248.9	44.9	68.3
MCOS-218	Mc1	26.0	25.3	167.8	428.4	87.4	205.8	466.8	45.1	90.1	491.2	165.8	129.0	50.1	76.0	30.4	153.2	36.9	31.4
MCOS-219	Mc1	16.5	17.9	245.9	327.6	158.1	157.1	187.5	60.0	133.1	595.6	108.8	90.3	60.6	95.1	42.5	316.6	29.2	48.6
MCOS-220	Mc3	31.1	20.2	286.7	313.8	246.2	255.1	263.6	323.4	141.9	378.2	124.4	65.2	62.5	157.9	81.1	227.7	29.6	76.7
MCOS-221	Mc2	124.9	33.2	195.6	267.2	106.0	191.3	244.2	53.6	63.5	86.2	78.6	58.0	52.2	139.4	51.4	201.3	64.1	45.8
MCOS-222	Mc1	15.5	34.0	283.7	466.7	182.2	162.2	345.7	67.1	164.4	540.7	129.3	205.8	52.9	185.8	71.1	248.5	37.5	67.6
MCOS-223	Mc2	37.0	26.2	258.5	401.8	270.1	190.6	284.6	50.3	113.3	281.4	174.1	106.8	29.2	150.3	63.9	269.0	91.6	98.0
MCOS-224	Mc3	23.7	30.5	299.3	423.4	212.1	207.7	282.5	57.9	173.6	316.7	199.7	131.4	28.0	197.3	102.8	340.8	57.2	79.3
MCOS-225	Mc1	31.2	34.7	246.8	406.5	139.5	138.2	369.9	62.2	342.3	472.4	153.3	173.8	35.1	143.2	84.4	233.8	37.7	49.0
MCOS-226	Mc1	45.0	36.7	303.2	441.3	260.5	127.2	290.2	48.3	172.4	735.4	141.3	85.1	50.2	152.0	58.0	209.7	31.6	147.5
MCOS-227	Mc3	37.0	33.1	300.5	516.4	141.9	277.2	424.3	67.8	189.0	277.7	113.2	68.3	54.9	72.8	46.9	325.8	34.5	84.7
MCOS-228	Mc3	29.0	35.4	285.7	421.6	279.9	149.0	417.9	37.4	208.5	323.3	140.0	139.4	62.7	127.1	54.0	280.9	37.6	144.4

Add. Table 3: Significantly different expressed genes between the three metabolic clusters

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	contrast 1	contrast 2	contrast 3	adjusted *P* value (%)	Direction
CALD1	2825	0.8	0.138	0.173	-3.427	1.454	1.225	0	Up
TAGLN	18949	0.774	0.168	0.218	-3.239	1.689	1.031	0	Up
FERMT2	6321	0.772	0.1	0.13	-3.384	1.624	1.134	0	Up
NNMT	13555	0.764	0.163	0.214	-3.207	0.811	1.369	0	Up
CTGF	4323	0.749	0.168	0.224	-3.148	1.414	1.094	0	Up
GPX8	7474	0.74	0.138	0.186	-3.146	0.877	1.309	0	Up
MRC2	12641	0.74	0.146	0.198	-3.081	0.466	1.441	0	Up
RUNX2	17000	0.739	0.086	0.116	-2.988	-0.457	1.765	0	Up
CLMP	3830	0.737	0.136	0.184	-3.137	0.946	1.277	0	Up
DPYSL3	5150	0.737	0.16	0.217	-3.108	1.192	1.162	0	Up
GEM	7035	0.735	0.146	0.199	-3.047	1.907	0.841	0	Up
COL18A1	3970	0.729	0.128	0.175	-3.082	0.576	1.397	0	Up
TSHZ3	20086	0.715	0.118	0.165	-3.019	0.472	1.406	0	Up
NEXN	13395	0.712	0.098	0.137	-3.114	1.352	1.101	0	Up
PTRF	15990	0.712	0.078	0.109	-3.201	1.178	1.217	0	Up
MSRB3	12776	0.705	0.141	0.2	-2.98	1.426	1	0	Up
CYR61	4528	0.701	0.164	0.234	-2.891	1.727	0.831	0	Up
PLS3	15215	0.7	0.136	0.195	-2.866	1.989	0.712	0	Up
ACTA2	202	0.699	0.163	0.233	-2.862	1.832	0.774	0	Up
TIMP2	19344	0.698	0.141	0.202	-2.948	0.796	1.238	0	Up
SRPX2	18547	0.698	0.142	0.203	-2.897	0.414	1.365	0	Up
GFPT2	7054	0.697	0.123	0.177	-2.877	0.188	1.446	0	Up
CNN2	3901	0.691	0.141	0.204	-2.92	1.308	1.016	0	Up
TCF4	19112	0.689	0.108	0.157	-2.975	1.219	1.081	0	Up
VCAN	20645	0.686	0.157	0.229	-2.846	0.518	1.296	0	Up
TGFB1I1	19244	0.684	0.104	0.152	-2.958	1.35	1.019	0	Up
FAP	6172	0.684	0.161	0.236	-2.795	0.311	1.353	0	Up
LMOD1	9739	0.68	0.075	0.111	-3	1.753	0.879	0	Up
FBLN2	6199	0.678	0.163	0.241	-2.843	1.027	1.089	0	Up
LOXL2	11683	0.677	0.138	0.204	-2.746	0.096	1.413	0	Up
LIMS2	9684	0.676	0.093	0.138	-2.851	1.956	0.718	0	Up
COL3A1	3983	0.675	0.18	0.266	-2.742	0.3	1.329	0	Up
S100A10	17025	0.674	0.117	0.173	-2.872	1.458	0.93	0	Up
DACT3	4555	0.674	0.139	0.206	-2.852	1.184	1.03	0	Up
GSN	7558	0.673	0.104	0.155	-2.808	1.934	0.704	0	Up
NID1	13456	0.672	0.129	0.192	-2.847	0.754	1.201	0	Up
ANGPTL2	633	0.671	0.134	0.199	-2.799	0.461	1.294	0	Up
SLIT3	18036	0.666	0.142	0.213	-2.8	0.736	1.183	0	Up
PDLIM4	14775	0.666	0.087	0.13	-2.831	1.921	0.721	0	Up
ISM1	8696	0.666	0.152	0.228	-2.763	1.56	0.831	0	Up
FSTL1	6776	0.662	0.143	0.217	-2.789	0.862	1.127	0	Up
SNAI2	18133	0.662	0.096	0.145	-2.879	0.945	1.141	0	Up
MIR100HG	12463	0.661	0.13	0.197	-2.804	1.207	0.995	0	Up
PDGFRB	14755	0.66	0.115	0.174	-2.782	0.46	1.285	0	Up
MRVI1	12739	0.659	0.114	0.173	-2.806	1.449	0.899	0	Up
LOX	11680	0.658	0.141	0.214	-2.678	0.176	1.345	0	Up
PLK3	15203	0.656	0.085	0.129	-2.877	1.408	0.952	0	Up
TPM4	19855	0.656	0.101	0.153	-2.787	0.43	1.301	0	Up
SPARC	18362	0.652	0.155	0.238	-2.665	0.3	1.288	0	Up
C5orf62	2509	0.648	0.12	0.185	-2.742	1.445	0.867	0	Up
PODN	15296	0.648	0.099	0.153	-2.807	1.128	1.029	0	Up
SKAP2	17634	0.648	0.095	0.147	-2.714	1.876	0.677	0	Up
RARRES2	16240	0.646	0.152	0.236	-2.698	1.339	0.886	0	Up
GLT8D2	7182	0.646	0.079	0.123	-2.809	0.442	1.307	0	Up
SHOX2	17566	0.644	0.092	0.143	-2.807	0.931	1.109	0	Up
EMILIN1	5520	0.644	0.108	0.168	-2.719	0.45	1.256	0	Up
PODNL1	15297	0.644	0.098	0.152	-2.79	1.059	1.048	0	Up
MGC24103	12397	0.643	0.142	0.22	-2.654	0.376	1.252	0	Up
EGR1	5396	0.64	0.177	0.276	-2.557	1.817	0.618	0	Up
ANXA1	754	0.639	0.129	0.203	-2.627	1.735	0.688	0	Up
TMEM200A	19554	0.638	0.086	0.135	-2.793	0.878	1.123	0	Up
MXRA5	12901	0.637	0.092	0.145	-2.737	0.52	1.237	0	Up
DKK3	4932	0.636	0.1	0.158	-2.746	1.188	0.972	0	Up
SERPINF1	17400	0.635	0.129	0.203	-2.692	1.101	0.979	0	Up
PDLIM7	14777	0.635	0.092	0.145	-2.763	1.15	0.997	0	Up
GPR124	7367	0.635	0.104	0.163	-2.735	0.947	1.064	0	Up
RNF144A	16653	0.634	0.089	0.14	-2.705	0.352	1.288	0	Up
PRKCDBP	15652	0.631	0.11	0.175	-2.701	1.204	0.942	0	Up
COL5A1	3991	0.63	0.102	0.162	-2.595	0.076	1.342	0	Up
CNN1	3900	0.629	0.185	0.295	-2.424	2.015	0.468	0	Up
VIM	20673	0.629	0.099	0.157	-2.711	0.82	1.102	0	Up
ZEB1	21166	0.628	0.107	0.171	-2.675	0.629	1.161	0	Up
HTRA3	8272	0.628	0.119	0.189	-2.43	-0.352	1.427	0	Up
IGFBP3	8405	0.628	0.115	0.184	-2.676	0.879	1.06	0	Up
POSTN	15383	0.627	0.179	0.285	-2.532	0.246	1.239	0	Up
ITPRIP	8761	0.627	0.094	0.151	-2.719	0.975	1.044	0	Up
NDN	13282	0.626	0.111	0.177	-2.63	1.566	0.759	0	Up
MYL9	12956	0.624	0.125	0.201	-2.618	1.423	0.81	0	Up
NOX4	13603	0.624	0.093	0.149	-2.514	-0.195	1.408	0	Up
COL1A2	3973	0.623	0.17	0.273	-2.522	0.25	1.233	0	Up
HTRA1	8270	0.623	0.156	0.251	-2.577	0.524	1.151	0	Up
SCG5	17165	0.623	0.11	0.177	-2.643	0.597	1.157	0	Up
PALLD	14490	0.623	0.078	0.126	-2.732	0.661	1.178	0	Up
COL5A2	3992	0.621	0.167	0.269	-2.453	-0.024	1.307	0	Up
LTBP2	11867	0.621	0.093	0.15	-2.691	1.199	0.939	0	Up
DSE	5180	0.62	0.088	0.142	-2.657	0.425	1.234	0	Up
DZIP1L	5277	0.618	0.071	0.115	-2.708	0.406	1.268	0	Up
FEZ1	6326	0.618	0.075	0.121	-2.688	1.629	0.764	0	Up
ARHGAP28	926	0.616	0.082	0.133	-2.665	0.486	1.213	0	Up
TUBB6	20251	0.615	0.122	0.198	-2.613	0.979	0.986	0	Up
TRPC1	20038	0.613	0.09	0.147	-2.659	0.822	1.074	0	Up
SERPING1	17402	0.611	0.08	0.131	-2.669	1.367	0.859	0	Up
C1S	2133	0.611	0.118	0.193	-2.595	1.088	0.933	0	Up
THY1	19311	0.61	0.121	0.198	-2.477	0.081	1.277	0	Up
CHADL	3589	0.609	0.127	0.209	2.555	-1.338	-0.811	0	Up
RHOJ	16542	0.609	0.092	0.152	-2.578	1.622	0.709	0	Up
FOXO1	6712	0.609	0.077	0.127	-2.47	2.137	0.443	0	Up
TNS1	19782	0.609	0.074	0.122	-2.69	1.263	0.912	0	Up
DACT1	4553	0.608	0.126	0.208	-2.559	0.64	1.095	0	Up
GAS1	6942	0.608	0.13	0.215	-2.554	1.265	0.84	0	Up
AEBP1	364	0.608	0.154	0.254	-2.502	0.431	1.149	0	Up
DCN	4634	0.606	0.157	0.26	-2.53	0.849	0.995	0	Up
TNFAIP6	19715	0.606	0.13	0.214	-2.484	0.235	1.219	0	Up
GADD45B	6883	0.605	0.099	0.163	-2.597	1.238	0.873	0	Up
AKAP12	465	0.605	0.104	0.172	-2.435	1.928	0.509	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	contrast 1	contrast 2	contrast 3	adjusted *P* value (%)	Direction
LAMA4	9489	0.605	0.078	0.129	-2.628	0.497	1.189	0	Up
PRICKLE1	15627	0.604	0.137	0.226	-2.538	1.167	0.871	0	Up
MYLK	12958	0.604	0.138	0.229	-2.473	1.599	0.662	0	Up
PCSK5	14697	0.604	0.054	0.089	-2.788	0.801	1.151	0	Up
COL1A1	3972	0.604	0.165	0.273	-2.415	0.113	1.231	0	Up
TPM2	19853	0.602	0.109	0.181	-2.573	0.928	0.986	0	Up
EGR2	5397	0.602	0.127	0.21	-2.449	1.711	0.604	0	Up
BGN	1500	0.601	0.103	0.172	-2.562	0.633	1.099	0	Up
FBN1	6202	0.601	0.108	0.179	-2.44	0.01	1.286	0	Up
BHLHE41	1507	0.6	0.133	0.222	-2.455	1.614	0.647	0	Up
CDH11	3348	0.6	0.089	0.148	-2.533	0.262	1.234	0	Up
UHMK1	20465	0.599	0.061	0.102	2.512	-2.131	-0.468	0	Up
INHBA	8573	0.599	0.148	0.246	-2.303	-0.261	1.324	0	Up
CFH	3562	0.598	0.101	0.17	-2.565	1.098	0.913	0	Up
FIBIN	6389	0.598	0.12	0.2	-2.536	0.914	0.972	0	Up
GRASP	7484	0.597	0.094	0.157	-2.527	1.556	0.708	0	Up
KCNMB4	8924	0.597	0.122	0.205	-2.496	0.478	1.127	0	Up
ANTXR1	750	0.597	0.108	0.181	-2.537	0.663	1.074	0	Up
ISLR	8694	0.596	0.133	0.223	-2.506	0.765	1.016	0	Up
SRPX	18546	0.596	0.123	0.206	-2.52	0.847	0.991	0	Up
MMP19	12536	0.594	0.062	0.104	-2.687	0.932	1.045	0	Up
SEMA5A	17317	0.593	0.115	0.194	-2.511	1.193	0.846	0	Up
PPIC	15449	0.593	0.098	0.166	-2.527	0.556	1.112	0	Up
C1R	2131	0.592	0.114	0.192	-2.516	1.092	0.89	0	Up
PRRX1	15774	0.592	0.121	0.204	-2.445	0.285	1.178	0	Up
ARL4C	996	0.592	0.106	0.179	-2.462	1.594	0.658	0	Up
RPRD2	16856	0.591	0.05	0.085	2.704	-1.656	-0.761	0	Up
DLC1	4936	0.59	0.059	0.1	-2.664	1.387	0.849	0	Up
MMP2	12537	0.589	0.126	0.213	-2.49	0.915	0.947	0	Up
THBS2	19280	0.588	0.146	0.248	-2.399	0.274	1.158	0	Up
NAP1L3	13066	0.588	0.074	0.127	-2.447	1.901	0.526	0	Up
FBLN5	6200	0.586	0.12	0.205	-2.45	1.386	0.736	0	Up
PPAPDC3	15424	0.586	0.06	0.102	-2.609	0.418	1.211	0	Up
CCDC8	3099	0.585	0.114	0.194	-2.387	1.693	0.579	0	Up
TEAD1	19159	0.585	0.072	0.123	-2.294	-0.496	1.414	0	Up
PDGFRA	14754	0.582	0.128	0.219	-2.453	0.919	0.926	0	Up
MEG3	12275	0.581	0.083	0.143	-2.516	0.611	1.084	0	Up
COL6A3	3996	0.581	0.103	0.178	-2.283	-0.253	1.31	0	Up
RBMS1	16338	0.581	0.088	0.152	-2.516	0.868	0.98	0	Up
CXorf57	4437	0.58	0.09	0.154	-1.891	2.587	-0.045	0	Up
SFRP2	17448	0.58	0.164	0.283	-2.412	0.85	0.932	0	Up
MARVELD1	12120	0.58	0.103	0.179	-2.362	0.038	1.234	0	Up
EMP1	5529	0.579	0.105	0.182	-2.25	2.007	0.379	0	Up
CYS1	4529	0.578	0.118	0.204	-2.414	0.447	1.096	0	Up
LHFP	9642	0.577	0.112	0.194	-2.431	1.293	0.764	0	Up
ADAMTS6	284	0.577	0.055	0.095	-2.571	0.27	1.251	0	Up
MMP14	12532	0.577	0.125	0.216	-2.36	0.218	1.16	0	Up
PTPN21	15959	0.576	0.098	0.171	-2.366	1.706	0.562	0	Up
PDLIM3	14774	0.576	0.103	0.179	-2.317	1.817	0.492	0	Up
EHD2	5403	0.575	0.087	0.151	-2.486	1.178	0.839	0	Up
PRKD1	15660	0.575	0.073	0.127	-2.467	1.61	0.655	0	Up
RPUSD1	16914	0.575	0.065	0.113	2.397	-1.956	-0.478	0	Up
CAV2	2944	0.575	0.064	0.111	-2.24	2.289	0.26	0	Up
SPON1	18456	0.574	0.165	0.287	-2.382	0.714	0.971	0	Up
MXRA8	12903	0.574	0.121	0.211	-2.402	0.517	1.061	0	Up
APCDD1	809	0.573	0.125	0.218	-2.314	1.67	0.549	0	Up
CYTH3	4535	0.572	0.098	0.171	-2.45	1.117	0.845	0	Up
COL6A1	3994	0.572	0.13	0.228	-2.271	-0.048	1.22	0	Up
EVC2	5727	0.571	0.045	0.078	-2.686	0.6	1.178	0	Up
FN1	6649	0.571	0.157	0.276	-2.064	-0.558	1.317	0	Up
C9orf125	2684	0.571	0.081	0.143	-2.449	1.422	0.721	0	Up
AGPAT4	404	0.57	0.065	0.113	-2.542	0.752	1.04	0	Up
FGF1	6341	0.568	0.043	0.075	-2.691	0.561	1.196	0	Up
PROS1	15712	0.568	0.071	0.125	-2.481	1.387	0.752	0	Up
RGS2	16502	0.568	0.136	0.24	-2.071	2.112	0.242	0	Up
MGC4294	12410	0.567	0.101	0.179	-2.14	-0.485	1.327	0	Up
MDFIC	12218	0.567	0.071	0.124	-2.492	0.635	1.062	0	Up
sep.11	17342	0.566	0.09	0.159	-2.443	0.794	0.971	0	Up
CCDC80	3100	0.566	0.114	0.201	-2.394	1.102	0.821	0	Up
EFHA2	5366	0.565	0.071	0.126	-2.496	0.891	0.96	0	Up
PSMD4	15868	0.565	0.051	0.09	2.515	-1.752	-0.623	0	Up
C1orf15-NBL1	2163	0.564	0.115	0.205	-2.375	0.605	1.011	0	Up
COL6A2	3995	0.562	0.118	0.21	-2.373	0.745	0.954	0	Up
PLXDC1	15224	0.561	0.066	0.118	-2.454	0.392	1.14	0	Up
LUM	11879	0.56	0.155	0.276	-2.294	0.379	1.06	0	Up
ADAM12	247	0.56	0.072	0.128	-2.384	0.16	1.196	0	Up
SVEP1	18824	0.56	0.066	0.118	-2.433	1.539	0.665	0	Up
FOXF2	6694	0.559	0.059	0.105	-2.446	1.638	0.632	0	Up
LCAT	9536	0.558	0.069	0.124	-2.469	0.878	0.951	0	Up
SERPINH1	17403	0.558	0.113	0.203	-2.207	-0.128	1.219	0	Up
LAYN	9528	0.557	0.084	0.151	-2.322	1.628	0.57	0	Up
RFX8	16471	0.556	0.065	0.117	-2.298	-0.177	1.287	0	Up
CD200	3228	0.556	0.067	0.12	-2.42	1.464	0.689	0	Up
DNAJB5	5018	0.555	0.065	0.116	-2.473	0.775	0.995	0	Up
SLC2A3	17831	0.555	0.045	0.08	-2.562	0.361	1.209	0	Up
PDPN	14781	0.555	0.107	0.193	-2.356	0.772	0.934	0	Up
HSPB2	8233	0.554	0.123	0.222	-2.212	1.686	0.489	0	Up
NID2	13457	0.552	0.063	0.114	-2.25	-0.303	1.312	0	Up
ZFP36	21191	0.552	0.107	0.193	-2.298	1.412	0.645	0	Up
SGCD	17470	0.552	0.082	0.148	-2.314	0.16	1.159	0	Up
DAB2	4549	0.551	0.082	0.149	-2.393	0.966	0.875	0	Up
VEGFC	20657	0.551	0.078	0.142	-2.337	1.517	0.624	0	Up
FOSB	6676	0.55	0.182	0.331	-2.048	1.891	0.319	0	Up
TBX15	19068	0.55	0.052	0.094	-2.512	0.635	1.072	0	Up
SSPN	18583	0.55	0.1	0.182	-2.333	1.175	0.759	0	Up
PMEPA1	15241	0.549	0.103	0.188	-2.303	0.393	1.059	0	Up
PDGFRL	14756	0.549	0.145	0.264	-2.281	0.61	0.96	0	Up
CAV1	2943	0.549	0.11	0.2	-2.093	1.944	0.322	0	Up
SH3PXD2A	17528	0.548	0.085	0.154	-2.374	0.876	0.901	0	Up
HIF1A	7874	0.548	0.083	0.151	-2.362	0.583	1.014	0	Up
C12orf34	1798	0.547	0.097	0.178	2.336	-1.068	-0.804	0	Up
SLC2A14	17829	0.547	0.08	0.147	-2.358	0.54	1.029	0	Up
VGLL3	20667	0.547	0.05	0.092	-2.484	0.462	1.127	0	Up
CTSK	4353	0.546	0.133	0.243	-2.288	0.767	0.901	0	Up
TWIST2	20278	0.546	0.113	0.207	-2.304	1.061	0.79	0	Up
PRRX2	15775	0.546	0.102	0.186	-2.319	0.667	0.957	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	contrast 1	contrast 2	contrast 3	adjusted *P* value (%)	Direction
COPZ2	4039	0.545	0.108	0.198	-2.267	0.319	1.07	0	Up
C10orf10	1681	0.545	0.121	0.222	-2.207	1.553	0.54	0	Up
C14orf37	1892	0.545	0.076	0.139	-2.359	0.505	1.044	0	Up
CLIP3	3821	0.545	0.067	0.123	-2.407	0.697	0.991	0	Up
HABP4	7685	0.545	0.073	0.133	-2.355	1.36	0.696	0	Up
IKBIP	8438	0.545	0.05	0.092	-2.273	-0.35	1.343	0	Up
C11orf96	1787	0.544	0.106	0.196	-2.288	1.218	0.718	0	Up
CACNA2D1	2793	0.543	0.051	0.094	-2.438	0.311	1.164	0	Up
KIAA1462	9087	0.542	0.085	0.156	-2.328	0.584	0.995	0	Up
DDR2	4674	0.542	0.076	0.14	-2.366	0.907	0.885	0	Up
PCOLCE	14687	0.542	0.1	0.185	-2.276	0.409	1.038	0	Up
SFXN4	17466	0.541	0.056	0.104	2.418	-1.324	-0.744	0	Up
MARCKS	12111	0.54	0.082	0.152	-2.315	0.463	1.037	0	Up
PVRL3	16018	0.54	0.062	0.114	-2.355	0.292	1.128	0	Up
MAP4K4	12065	0.54	0.058	0.108	-2.425	0.771	0.971	0	Up
SLC12A4	17666	0.538	0.06	0.112	-2.34	1.55	0.612	0	Up
CMTM3	3874	0.538	0.089	0.166	-2.3	1.17	0.744	0	Up
BNC2	1562	0.538	0.071	0.132	-2.327	0.388	1.074	0	Up
GRP	7537	0.537	0.177	0.329	-2.214	1.026	0.756	0	Up
PPFIBP1	15442	0.537	0.065	0.122	-2.329	0.302	1.109	0	Up
KLF12	9198	0.537	0.069	0.128	-2.308	1.462	0.63	0	Up
C7orf10	2582	0.537	0.077	0.144	-2.289	0.325	1.079	0	Up
MYO1B	12972	0.536	0.062	0.116	-2.35	1.394	0.68	0	Up
ADCK3	304	0.536	0.052	0.097	2.436	-1.237	-0.788	0	Up
CELF2	3478	0.536	0.069	0.129	-2.316	1.397	0.661	0	Up
SOBP	18276	0.536	0.054	0.1	-2.419	1.244	0.777	0	Up
PLXDC2	15225	0.536	0.09	0.169	-2.183	-0.003	1.156	0	Up
PXDN	16027	0.535	0.116	0.217	-2.2	0.255	1.061	0	Up
ARAP3	887	0.535	0.077	0.144	-2.252	1.526	0.575	0	Up
LIMS1	9683	0.534	0.06	0.112	-2.361	0.434	1.073	0	Up
CFHR3	3565	0.534	0.086	0.162	-2.291	1.115	0.762	0	Up
RAB23	16092	0.534	0.086	0.161	-2.26	0.36	1.05	0	Up
FOS	6675	0.534	0.162	0.303	-1.938	1.941	0.241	0	Up
MT1M	12797	0.533	0.143	0.268	-2.222	0.953	0.79	0	Up
LRP1	11735	0.533	0.051	0.096	-2.424	0.588	1.045	0	Up
MFAP5	12348	0.533	0.127	0.238	-2.155	0.142	1.082	0	Up
ETV1	5718	0.532	0.05	0.094	-2.436	0.683	1.012	0	Up
PSMB4	15847	0.532	0.054	0.101	2.241	-1.867	-0.431	0	Up
COL16A1	3968	0.532	0.076	0.143	-2.318	0.886	0.868	0	Up
CSRP1	4276	0.532	0.059	0.11	-2.094	2.087	0.264	0	Up
COL8A2	4002	0.531	0.121	0.227	-2.158	0.151	1.081	0	Up
FOXQ1	6719	0.531	0.146	0.275	-2.197	1.097	0.719	0	Up
PPAPDC1A	15421	0.53	0.149	0.281	-2.113	0.084	1.083	0	Up
KIRREL	9182	0.53	0.072	0.135	-2.318	0.694	0.945	0	Up
ARSI	1056	0.53	0.066	0.124	-2.294	0.316	1.086	0	Up
SPON2	18457	0.529	0.065	0.124	-2.343	0.789	0.92	0	Up
PRKCA	15649	0.528	0.1	0.189	-2.103	1.697	0.427	0	Up
RASD1	16251	0.528	0.155	0.293	-1.846	2.036	0.154	0	Up
SCAMP3	17128	0.528	0.055	0.104	2.362	-1.284	-0.731	0	Up
SCARNA17	17147	0.527	0.107	0.202	-2.2	1.27	0.651	0	Up
PPP1R16A	15494	0.527	0.074	0.141	2.214	-1.543	-0.548	0	Up
CLDN11	3758	0.527	0.131	0.25	-1.985	1.833	0.31	0	Up
NIT1	13476	0.526	0.055	0.104	2.313	-1.522	-0.609	0	Up
MT1X	12798	0.526	0.113	0.214	-2.199	1.189	0.683	0	Up
ARRDC3	1046	0.525	0.082	0.157	-2.113	1.74	0.415	0	Up
ADAMTS5	283	0.525	0.11	0.21	-2.22	0.803	0.85	0	Up
KLF2	9204	0.525	0.093	0.178	-2.208	1.295	0.645	0	Up
COL12A1	3964	0.524	0.141	0.269	-2.073	-0.001	1.097	0	Up
GXYLT2	7646	0.524	0.137	0.261	-2.165	0.458	0.96	0	Up
FNDC4	6658	0.524	0.072	0.137	-2.131	1.763	0.415	0	Up
MB21D2	12148	0.524	0.067	0.127	-1.767	2.352	-0.016	0	Up
SLFN11	18026	0.523	0.084	0.161	-2.253	0.756	0.886	0	Up
FNDC1	6655	0.523	0.119	0.229	-2.19	0.613	0.911	0	Up
JDP2	8787	0.523	0.077	0.148	-2.265	0.668	0.928	0	Up
DOCK11	5084	0.522	0.076	0.146	-2.169	1.591	0.504	0	Up
AIM1	446	0.522	0.116	0.223	-1.118	2.569	-0.446	0	Up
H19	7663	0.522	0.125	0.24	-2.038	1.676	0.401	0	Up
YY1AP1	21044	0.521	0.055	0.105	2.35	-1.031	-0.826	0	Up
C1QTNF5	2125	0.521	0.099	0.19	-2.218	0.848	0.831	0	Up
FLNA	6617	0.521	0.062	0.12	-2.278	1.329	0.668	0	Up
LRIG3	11730	0.521	0.094	0.181	-2.196	1.235	0.662	0	Up
TMEM45A	19610	0.521	0.131	0.252	-2.17	0.607	0.902	0	Up
UACA	20323	0.521	0.069	0.132	-2.282	0.674	0.935	0	Up
FMO2	6639	0.52	0.174	0.335	-1.91	1.835	0.269	0	Up
LPCAT2	11695	0.52	0.054	0.103	-2.32	1.347	0.683	0	Up
DPYSL2	5149	0.519	0.083	0.16	-1.976	1.941	0.261	0	Up
KDELC2	8968	0.519	0.062	0.119	-2.261	0.313	1.069	0	Up
ARHGAP39	934	0.519	0.072	0.139	2.265	-0.722	-0.906	0	Up
SGK1	17475	0.519	0.091	0.176	-2.109	1.583	0.476	0	Up
SGIP1	17474	0.518	0.072	0.14	-2.098	-0.165	1.176	0	Up
CLEC11A	3782	0.517	0.105	0.203	-2.141	0.274	1.022	0	Up
CES1	3545	0.517	0.084	0.163	-2.226	0.948	0.794	0	Up
CFI	3568	0.517	0.086	0.166	-2.074	1.691	0.414	0	Up
KLF7	9209	0.517	0.07	0.136	-2.263	0.805	0.872	0	Up
PHLDB1	14949	0.517	0.059	0.115	-2.243	0.226	1.095	0	Up
ACTN1	214	0.517	0.087	0.168	-2.22	0.9	0.81	0	Up
SFRP4	17449	0.516	0.129	0.25	-2.024	1.62	0.416	0	Up
NHS	13452	0.516	0.049	0.095	-2.341	1.263	0.728	0	Up
RECK	16395	0.515	0.065	0.125	-2.265	0.575	0.965	0	Up
GGT5	7072	0.515	0.089	0.172	-2.187	1.182	0.679	0	Up
LRCH2	11717	0.515	0.046	0.09	-2.366	1.213	0.762	0	Up
PRDM1	15591	0.515	0.084	0.164	-2.11	0.025	1.105	0	Up
HOXA11-AS1	8086	0.514	0.094	0.183	-2.138	0.234	1.036	0	Up
HEG1	7809	0.514	0.081	0.158	-2.174	1.341	0.608	0	Up
RUNX1T1	16999	0.514	0.1	0.194	-2.159	1.201	0.657	0	Up
BACH2	1366	0.514	0.043	0.084	-2.236	1.853	0.434	0	Up
LSP1	11857	0.513	0.07	0.136	-2.238	1.085	0.745	0	Up
LAMB1	9491	0.513	0.04	0.078	-2.429	0.696	1.003	0	Up
F2RL1	5787	0.512	0.078	0.153	-2.033	1.769	0.361	0	Up
TSPO2	20120	0.512	0.058	0.114	-2.115	-0.194	1.197	0	Up
ZFHX4	21181	0.512	0.093	0.181	-2.184	0.936	0.777	0	Up
WDR86	20856	0.511	0.136	0.266	-2.115	1.108	0.671	0	Up
PKD2	15071	0.51	0.071	0.139	-2.186	1.324	0.621	0	Up
HOXA3	8089	0.51	0.106	0.207	-2.022	1.616	0.416	0	Up
LOC645676	11405	0.51	0.054	0.107	2.295	-1.013	-0.805	0	Up
MALL	12005	0.51	0.13	0.256	-1.653	2.146	0.008	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	contrast 1	contrast 2	contrast 3	adjusted *P* value (%)	Direction
FAM89A	6139	0.51	0.086	0.169	-1.907	1.935	0.227	0	Up
HIC1	7872	0.509	0.072	0.142	-2.15	1.438	0.556	0	Up
FAS	6180	0.508	0.057	0.113	-2.239	1.303	0.658	0	Up
NCRNA00241	13204	0.508	0.078	0.153	-2.195	0.619	0.911	0	Up
FAM43A	6040	0.508	0.076	0.15	-2.146	1.388	0.574	0	Up
ASPN	1122	0.508	0.184	0.362	-1.976	-0.027	1.056	0	Up
DIO1	4878	0.507	0.109	0.215	2.139	-0.895	-0.77	0	Up
TGFBR2	19249	0.507	0.08	0.157	-1.869	1.999	0.181	0	Up
HMCN1	8005	0.507	0.107	0.21	-2.079	0.199	1.019	0	Up
COL10A1	3961	0.506	0.174	0.344	-1.768	-0.595	1.175	0	Up
C21orf34	2296	0.506	0.051	0.101	-2.296	1.042	0.793	0	Up
KRTCAP3	9459	0.505	0.097	0.193	2.046	-1.518	-0.469	0.214	Up
ITGA11	8711	0.505	0.127	0.252	-1.952	-0.188	1.108	0.214	Up
ADAMTS14	273	0.505	0.123	0.244	-2.013	0.011	1.06	0.214	Up
NEFH	13351	0.505	0.107	0.212	-2.131	0.642	0.868	0.214	Up
DST	5189	0.505	0.094	0.187	-2.075	1.433	0.519	0.214	Up
SH3PXD2B	17529	0.504	0.083	0.165	-2.17	0.72	0.856	0.214	Up
E4F1	5287	0.504	0.059	0.117	2.201	-1.336	-0.624	0.214	Up
ETS2	5717	0.504	0.067	0.133	-1.812	2.119	0.103	0.214	Up
SCHIP1	17177	0.503	0.1	0.198	-2.109	1.188	0.635	0.214	Up
OLFML2B	13917	0.503	0.084	0.168	-2.129	0.366	0.978	0.214	Up
DZIP1	5276	0.503	0.07	0.14	-2.187	1.045	0.734	0.214	Up
CCDC3	3053	0.502	0.106	0.21	-1.973	1.637	0.382	0.214	Up
MPZL2	12634	0.502	0.102	0.204	-1.668	2.126	0.024	0.214	Up
ZNHIT2	21794	0.502	0.052	0.104	2.229	-1.334	-0.64	0.214	Up
CNRIP1	3925	0.502	0.045	0.089	-2.294	1.295	0.69	0.214	Up
JAG1	8774	0.502	0.074	0.147	-2.181	0.773	0.841	0.214	Up
ZC3H12B	21103	0.502	0.04	0.079	-2.311	1.478	0.625	0.214	Up
SPARCL1	18363	0.501	0.129	0.257	-2.023	1.38	0.513	0.214	Up
SSC5D	18577	0.501	0.117	0.234	-2.083	0.45	0.92	0.214	Up
PRKD3	15662	0.501	0.077	0.154	-2.025	1.639	0.409	0.214	Up
PTPRD	15972	0.5	0.103	0.206	-2.049	0.169	1.016	0.214	Up
CD99	3298	0.499	0.053	0.107	-2.227	0.464	0.99	0.214	Up
SPRY2	18483	0.499	0.1	0.2	-1.991	1.565	0.421	0.214	Up
LOC100130876	10110	0.499	0.104	0.208	-2.107	0.64	0.856	0.214	Up
TMPRSS6	19690	0.499	0.084	0.167	2.025	-1.564	-0.439	0.214	Up
GREM1	7492	0.499	0.097	0.195	-2.072	0.272	0.986	0.214	Up
COL15A1	3967	0.498	0.082	0.164	-2.13	1.106	0.679	0.214	Up
C17orf51	2000	0.498	0.078	0.157	-2.155	0.808	0.813	0.214	Up
GNG11	7241	0.498	0.093	0.187	-2.058	1.374	0.534	0.214	Up
TMEM2	19552	0.498	0.056	0.113	-2.029	-0.29	1.19	0.214	Up
ID3	8312	0.498	0.068	0.136	-2.137	1.303	0.604	0.214	Up
ARHGEF40	965	0.498	0.074	0.149	-2.163	0.749	0.841	0.214	Up
FAM20B	6010	0.497	0.057	0.114	2.072	-1.685	-0.415	0.214	Up
JAZF1	8786	0.497	0.06	0.12	-2.191	1.12	0.706	0.214	Up
ZNF469	21487	0.497	0.116	0.234	-2.061	0.414	0.923	0.214	Up
RBMS3	16340	0.496	0.042	0.085	-2.304	0.586	0.982	0.214	Up
PYCR2	16036	0.496	0.048	0.097	2.105	-1.751	-0.406	0.214	Up
BEND6	1483	0.496	0.04	0.08	-2.314	0.465	1.036	0.214	Up
PYCRL	16037	0.496	0.056	0.113	2.215	-1.015	-0.762	0.214	Up
TARS2	18972	0.495	0.04	0.08	2.228	-1.64	-0.516	0.214	Up
MOXD1	12603	0.495	0.106	0.215	-1.923	1.652	0.35	0.214	Up
GLIPR2	7156	0.495	0.089	0.18	-2.109	0.61	0.869	0.214	Up
ECM2	5316	0.495	0.056	0.113	-2.201	0.536	0.947	0.214	Up
ACSL4	189	0.495	0.053	0.108	-2.204	0.469	0.976	0.214	Up
RAB7L1	16138	0.494	0.078	0.158	-2.118	0.488	0.923	0.214	Up
FXYD6	6823	0.494	0.078	0.157	-1.718	2.082	0.068	0.214	Up
C1orf182	2177	0.494	0.061	0.124	1.707	-2.192	-0.018	0.214	Up
WTIP	20938	0.494	0.083	0.168	-2	1.557	0.429	0.214	Up
ITGB1	8727	0.493	0.059	0.12	-2.162	1.215	0.653	0.214	Up
FZD7	6836	0.493	0.101	0.205	-2.04	1.276	0.563	0.214	Up
ANKRD35	695	0.493	0.123	0.25	-1.977	1.412	0.475	0.214	Up
ZEB2	21168	0.492	0.063	0.127	-2.173	0.831	0.814	0.214	Up
LRRC32	11774	0.492	0.089	0.182	-2.102	0.885	0.754	0.214	Up
ANTXR2	751	0.492	0.048	0.098	-2.177	1.441	0.569	0.214	Up
HEYL	7849	0.492	0.086	0.175	-2.057	0.262	0.982	0.214	Up
LEPRE1	9607	0.491	0.048	0.097	-2.104	-0.098	1.152	0.214	Up
VCL	20646	0.491	0.065	0.132	-1.865	1.938	0.204	0.214	Up
PELI2	14817	0.491	0.081	0.164	-2.107	1.02	0.702	0.214	Up
DUSP1	5217	0.491	0.104	0.211	-2.007	1.361	0.511	0.214	Up
SALL4	17067	0.491	0.115	0.235	-1.935	-0.073	1.053	0.214	Up
CPXM1	4137	0.49	0.085	0.174	-2.034	1.348	0.532	0.363	Up
ACVR2A	234	0.49	0.049	0.1	-2.227	0.932	0.801	0.363	Up
EDN1	5334	0.49	0.112	0.228	-1.874	1.673	0.316	0.363	Up
CTSB	4346	0.489	0.076	0.156	-1.904	-0.331	1.141	0.363	Up
C9orf53	2729	0.489	0.121	0.246	-2.048	0.664	0.815	0.363	Up
ATF3	1147	0.489	0.087	0.177	-2.092	0.844	0.766	0.363	Up
FGF7	6360	0.489	0.062	0.127	-2.143	1.087	0.694	0.363	Up
NTM	13753	0.488	0.057	0.118	-1.978	-0.28	1.159	0.363	Up
LOC100129940	9999	0.488	0.067	0.138	-2.059	0.13	1.036	0.363	Up
CD36	3255	0.488	0.167	0.342	-1.492	2.094	-0.057	0.363	Up
RASA3	16246	0.488	0.069	0.141	-2.125	0.627	0.871	0.363	Up
SNRNP35	18218	0.488	0.048	0.097	2.04	-1.786	-0.358	0.363	Up
MME	12524	0.488	0.086	0.176	-1.901	1.697	0.32	0.363	Up
NXN	13843	0.488	0.061	0.125	-2.14	0.506	0.927	0.363	Up
WNT2	20911	0.487	0.075	0.153	-2.019	0.063	1.042	0.363	Up
ITGA1	8709	0.487	0.064	0.132	-2.058	0.1	1.048	0.363	Up
MAP1B	12033	0.487	0.066	0.136	-2.113	1.14	0.657	0.363	Up
PEX5L	14845	0.487	0.07	0.144	2.116	-0.653	-0.855	0.363	Up
COL5A3	3993	0.486	0.092	0.188	-1.893	-0.24	1.098	0.363	Up
SULF1	18774	0.486	0.118	0.244	-1.76	-0.501	1.133	0.363	Up
FAT1	6188	0.485	0.103	0.212	-2	1.267	0.546	0.363	Up
SPOCK1	18453	0.485	0.121	0.25	-1.847	-0.26	1.082	0.363	Up
JAM3	8784	0.485	0.06	0.124	-2.144	0.937	0.756	0.363	Up
FST	6775	0.484	0.135	0.279	-1.957	0.183	0.961	0.363	Up
SPRY1	18482	0.484	0.086	0.179	-2.009	1.332	0.524	0.363	Up
LIMS3L	9686	0.484	0.067	0.139	-2.102	1.076	0.677	0.363	Up
ZNF300P1	21390	0.484	0.088	0.182	-2.031	1.203	0.589	0.363	Up
RNF175	16670	0.484	0.046	0.096	-2.089	1.629	0.447	0.363	Up
USP21	20558	0.484	0.048	0.099	2.18	-1.18	-0.676	0.363	Up
PRDM6	15602	0.484	0.061	0.125	-2.016	-0.048	1.085	0.363	Up
SELENBP1	17293	0.484	0.122	0.253	2.019	-0.891	-0.708	0.363	Up
DPT	5137	0.483	0.133	0.276	-2.003	0.554	0.835	0.363	Up
PID1	14982	0.483	0.093	0.192	-1.816	1.768	0.247	0.363	Up
MAGEL2	11991	0.482	0.07	0.146	-2.078	1.118	0.647	0.363	Up
FAM195A	5990	0.482	0.056	0.117	2.066	-1.453	-0.506	0.363	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	contrast 1	contrast 2	contrast 3	adjusted *P* value (%)	Direction
AOC3	770	0.482	0.109	0.227	-1.872	1.592	0.347	0.363	Up
LAMA1	9486	0.482	0.056	0.117	-2.147	0.846	0.794	0.363	Up
C5orf13	2477	0.482	0.065	0.135	-2.099	0.52	0.9	0.363	Up
KANK2	8818	0.481	0.067	0.14	-2.068	1.216	0.602	0.363	Up
AKAP13	466	0.481	0.049	0.101	-2.176	1.061	0.722	0.363	Up
TNFRSF10D	19723	0.481	0.075	0.155	-1.726	1.967	0.119	0.363	Up
PTGIR	15922	0.48	0.063	0.131	-2.096	0.495	0.909	0.363	Up
TSHZ2	20085	0.48	0.074	0.155	-1.944	1.556	0.4	0.363	Up
MAMDC2	12007	0.48	0.062	0.13	-2.112	0.75	0.814	0.363	Up
FABP4	5804	0.479	0.125	0.261	-1.372	2.169	-0.15	0.363	Up
WISP1	20897	0.479	0.055	0.114	-2.067	0.131	1.04	0.363	Up
SWAP70	18829	0.479	0.054	0.112	-1.574	2.249	-0.076	0.363	Up
ABCA6	42	0.479	0.091	0.19	-2.03	0.988	0.675	0.363	Up
ADAMTS1	269	0.479	0.101	0.211	-1.933	1.413	0.452	0.363	Up
VSNL1	20740	0.479	0.108	0.226	-1.956	1.289	0.514	0.363	Up
LOC541471	11281	0.478	0.083	0.174	-1.619	-0.816	1.186	0.363	Up
KRT17	9319	0.478	0.228	0.477	-1.774	1.583	0.299	0.363	Up
DAP3	4567	0.478	0.049	0.104	2.162	-0.88	-0.788	0.363	Up
CR2	4142	0.477	0.098	0.206	-0.932	2.384	-0.47	0.363	Up
D4S234E	4545	0.477	0.091	0.19	-1.441	2.165	-0.112	0.363	Up
PECAM1	14810	0.477	0.067	0.14	-1.884	1.721	0.301	0.363	Up
KRT16P2	9318	0.477	0.217	0.455	-1.729	1.671	0.24	0.363	Up
STARD8	18661	0.477	0.051	0.106	-2.149	0.913	0.768	0.363	Up
ADIPOR1	333	0.477	0.054	0.114	2.04	-1.473	-0.484	0.363	Up
JAM2	8783	0.477	0.069	0.144	-2.007	1.357	0.513	0.363	Up
MRPL9	12705	0.477	0.047	0.098	2.174	-0.737	-0.852	0.363	Up
COMP	4020	0.476	0.153	0.32	-1.894	0.096	0.963	0.363	Up
MT1E	12791	0.476	0.107	0.224	-2.006	0.751	0.758	0.363	Up
PARVA	14557	0.476	0.06	0.126	-2.073	1.185	0.618	0.363	Up
ST6GAL2	18617	0.476	0.083	0.174	-1.883	-0.183	1.069	0.363	Up
FAM20A	6009	0.476	0.073	0.154	-2.062	0.778	0.776	0.363	Up
C12orf70	1830	0.475	0.072	0.152	-1.987	0.129	0.999	0.363	Up
ROR1	16743	0.475	0.055	0.117	-2.118	0.753	0.816	0.363	Up
MXRA7	12902	0.475	0.058	0.121	-2.109	0.818	0.785	0.363	Up
FAT4	6191	0.475	0.037	0.077	-2.236	1.206	0.695	0.363	Up
RND3	16628	0.475	0.114	0.241	-1.494	2.053	-0.039	0.363	Up
ITGAV	8725	0.474	0.072	0.153	-2.036	1.075	0.643	0.381	Up
C20orf103	2234	0.474	0.129	0.272	-1.975	0.716	0.755	0.381	Up
RUSC1	17002	0.474	0.051	0.107	2.111	-1.161	-0.647	0.381	Up
ITGA5	8716	0.474	0.075	0.159	-2.012	0.343	0.926	0.381	Up
EIF5A2	5468	0.474	0.049	0.104	-2.097	0.29	0.992	0.381	Up
AKR1B15	488	0.473	0.118	0.249	-0.978	2.321	-0.42	0.381	Up
CADM3	2813	0.473	0.058	0.123	-1.255	2.373	-0.295	0.381	Up
IGFBP7	8409	0.473	0.082	0.175	-2.007	1.06	0.633	0.381	Up
MAGIX	11996	0.472	0.04	0.084	2.193	-1.162	-0.69	0.381	Up
MT1L	12796	0.472	0.096	0.203	-1.999	0.869	0.706	0.381	Up
CCL14	3132	0.472	0.142	0.3	-1.391	2.087	-0.107	0.381	Up
FAM7A1	6115	0.472	0.046	0.097	-1.947	-0.324	1.161	0.381	Up
TROVE2	20036	0.472	0.045	0.095	2.149	-1.092	-0.696	0.381	Up
PDE2A	14730	0.472	0.088	0.187	-1.685	1.88	0.132	0.381	Up
CLIP4	3822	0.472	0.07	0.149	-2.041	0.972	0.687	0.381	Up
SPHK2	18419	0.472	0.039	0.083	2.207	-0.97	-0.775	0.381	Up
MMP23B	12540	0.471	0.092	0.196	-1.875	1.504	0.384	0.381	Up
CLDN5	3774	0.471	0.107	0.227	-1.721	1.767	0.196	0.381	Up
PRNP	15688	0.471	0.095	0.202	-1.986	1.004	0.645	0.381	Up
C20orf194	2258	0.471	0.054	0.115	-2.101	0.661	0.844	0.381	Up
ARHGAP20	918	0.471	0.054	0.116	-2.073	1.16	0.628	0.381	Up
F2R	5786	0.47	0.057	0.121	-2.089	0.772	0.793	0.381	Up
KRT14	9315	0.47	0.235	0.5	-1.738	1.571	0.285	0.381	Up
USP13	20548	0.47	0.062	0.131	2.06	-0.998	-0.686	0.381	Up
ANXA3	761	0.47	0.143	0.304	-1.289	2.143	-0.184	0.381	Up
NPR2	13642	0.47	0.071	0.151	-2.039	0.762	0.77	0.381	Up
LMCD1	9727	0.47	0.087	0.184	-1.93	0.111	0.976	0.381	Up
CECR5-AS1	3467	0.47	0.071	0.151	2.015	-1.095	-0.624	0.381	Up
KLF6	9208	0.469	0.05	0.107	-2.062	1.318	0.558	0.381	Up
LOC150622	10676	0.469	0.06	0.127	-1.046	2.42	-0.424	0.381	Up
HOXA4	8090	0.469	0.057	0.12	-2.024	1.326	0.535	0.381	Up
ITGBL1	8739	0.469	0.117	0.249	-1.903	0.166	0.939	0.381	Up
EFNB2	5378	0.469	0.078	0.167	-2.007	0.989	0.662	0.381	Up
ZFPM2	21206	0.469	0.07	0.15	-2.031	0.922	0.702	0.381	Up
ECI1	5313	0.469	0.072	0.153	2	-1.158	-0.59	0.381	Up
C1orf27	2201	0.469	0.056	0.119	2.055	-1.168	-0.615	0.381	Up
CXCL12	4398	0.468	0.072	0.155	-2.013	0.996	0.662	0.381	Up
VPS72	20728	0.468	0.048	0.103	2.096	-1.14	-0.648	0.381	Up
KIAA0408	9027	0.468	0.053	0.113	-2.003	1.442	0.477	0.381	Up
F3	5790	0.468	0.119	0.254	-1.895	1.258	0.494	0.381	Up
ZNF646	21624	0.467	0.026	0.055	2.23	-1.886	-0.418	0.381	Up
IGF2	8396	0.467	0.066	0.141	-1.99	1.25	0.548	0.381	Up
PEAK1	14806	0.467	0.053	0.114	-2.078	1.01	0.691	0.381	Up
ADH1C	324	0.466	0.153	0.328	-1.265	2.126	-0.19	0.381	Up
C6orf145	2533	0.466	0.075	0.16	-1.789	1.716	0.253	0.381	Up
PTGFRN	15921	0.466	0.07	0.149	-1.983	0.298	0.928	0.381	Up
KIF26A	9151	0.466	0.054	0.116	-1.609	2.095	0.005	0.381	Up
KCTD12	8949	0.465	0.074	0.158	-1.984	1.108	0.602	0.381	Up
ADD3	321	0.465	0.104	0.223	-1.869	1.372	0.434	0.381	Up
TSNAX	20094	0.465	0.047	0.102	2.014	-1.467	-0.472	0.381	Up
SPSB1	18490	0.464	0.067	0.144	-2.018	0.604	0.823	0.381	Up
TMEM43	19608	0.464	0.05	0.108	-2.088	0.698	0.822	0.381	Up
MFAP4	12347	0.463	0.163	0.353	-1.788	1.434	0.366	0.381	Up
GLIS1	7157	0.463	0.075	0.161	-1.827	-0.228	1.058	0.381	Up
SYNPO	18864	0.463	0.054	0.117	-2.022	1.226	0.574	0.381	Up
KIAA0907	9049	0.463	0.044	0.096	2.062	-1.337	-0.551	0.381	Up
CERCAM	3536	0.463	0.061	0.131	-1.812	-0.37	1.108	0.381	Up
NKD2	13486	0.462	0.068	0.147	-1.985	0.383	0.895	0.381	Up
ETHE1	5713	0.462	0.056	0.121	-1.899	1.594	0.36	0.381	Up
WISP2	20898	0.462	0.171	0.37	-1.833	0.124	0.919	0.381	Up
SPHK1	18418	0.462	0.067	0.146	-2.008	0.83	0.727	0.381	Up
EFEMP2	5364	0.461	0.073	0.159	-1.928	0.157	0.956	0.381	Up
PTGIS	15923	0.46	0.1	0.218	-1.716	1.675	0.231	0.381	Up
PTN	15939	0.46	0.155	0.337	-1.607	1.752	0.142	0.381	Up
LYSMD1	11921	0.46	0.063	0.136	1.893	-1.495	-0.398	0.381	Up
SCN3B	17192	0.46	0.082	0.178	-1.935	1.129	0.567	0.554	Up
TCEAL7	19092	0.459	0.063	0.137	-1.701	1.857	0.15	0.554	Up
SPIRE1	18445	0.459	0.051	0.112	0.952	-2.427	0.477	0.554	Up
RAB3A	16113	0.459	0.071	0.154	1.638	-1.89	-0.103	0.554	Up
NSUN7	13736	0.458	0.093	0.203	0.986	1.524	-1.137	0.554	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	contrast 1	contrast 2	contrast 3	adjusted *P* value (%)	Direction
SLC26A10	17801	0.458	0.063	0.138	-1.88	-0.082	1.027	0.554	Up
LOC100132891	10294	0.458	0.028	0.062	-2.262	0.461	1.01	0.554	Up
PLSCR4	15219	0.458	0.074	0.162	-1.763	1.67	0.258	0.554	Up
TNFRSF18	19731	0.457	0.09	0.197	1.589	-1.863	-0.088	0.554	Up
C6orf174	2545	0.457	0.06	0.131	-2.006	0.912	0.692	0.554	Up
SYNDIG1	18851	0.457	0.13	0.284	-1.833	0.125	0.919	0.554	Up
RECQL	16396	0.457	0.03	0.066	-2.213	0.379	1.017	0.554	Up
CDKN1C	3411	0.457	0.09	0.197	-1.81	1.47	0.364	0.554	Up
PANX2	14505	0.457	0.033	0.073	2.067	-1.614	-0.441	0.554	Up
ITGB3	8732	0.456	0.093	0.204	-1.918	0.441	0.836	0.554	Up
MYO1D	12974	0.456	0.053	0.115	-2.029	0.566	0.844	0.554	Up
C14orf21	1886	0.456	0.037	0.08	2.086	-1.389	-0.542	0.554	Up
CRISPLD2	4183	0.456	0.065	0.142	-1.985	0.885	0.692	0.554	Up
ADAMTS12	271	0.456	0.045	0.098	-1.858	-0.354	1.126	0.554	Up
OTUD7B	14397	0.455	0.055	0.12	1.922	-1.438	-0.436	0.554	Up
CDO1	3421	0.455	0.108	0.238	-1.553	1.845	0.076	0.554	Up
GOLGA7B	7287	0.454	0.064	0.141	-1.822	-0.206	1.047	0.554	Up
MITF	12471	0.454	0.053	0.116	-1.98	0.257	0.943	0.554	Up
AASS	27	0.454	0.057	0.126	-1.935	1.331	0.486	0.554	Up
AKR1C1	489	0.454	0.094	0.208	-1.291	2.1	-0.165	0.554	Up
CD248	3237	0.454	0.066	0.146	-1.969	0.542	0.822	0.554	Up
TP53I3	19830	0.454	0.067	0.147	-1.786	1.625	0.288	0.554	Up
RCAN2	16364	0.454	0.054	0.119	-1.922	1.422	0.442	0.554	Up
POTEF	15389	0.454	0.056	0.123	-1.98	1.149	0.583	0.554	Up
LRRN4CL	11831	0.454	0.072	0.158	-1.84	1.462	0.383	0.554	Up
WNK2	20903	0.453	0.063	0.138	1.97	-1.002	-0.637	0.554	Up
TMEM204	19560	0.453	0.072	0.158	-1.955	0.687	0.756	0.554	Up
LDB2	9574	0.453	0.051	0.113	-2.021	0.94	0.689	0.554	Up
GLIS2	7158	0.453	0.086	0.19	-1.872	0.189	0.914	0.554	Up
TSPAN18	20105	0.453	0.063	0.139	-1.977	0.821	0.714	0.554	Up
HHIPL1	7863	0.452	0.061	0.135	-1.923	0.173	0.947	0.554	Up
EDIL3	5333	0.452	0.118	0.26	-1.388	-0.93	1.109	0.554	Up
CPZ	4139	0.452	0.076	0.168	-1.913	1.108	0.564	0.554	Up
MAP4K5	12066	0.452	0.046	0.101	-2.052	0.801	0.762	0.554	Up
GNMT	7258	0.451	0.065	0.143	0.974	1.572	-1.15	0.554	Up
XG	20957	0.451	0.063	0.139	-1.949	0.372	0.88	0.554	Up
CASKIN1	2915	0.451	0.055	0.123	1.856	-1.544	-0.358	0.554	Up
KLK8	9270	0.451	0.219	0.486	-1.237	2.011	-0.158	0.554	Up
MIPOL1	12462	0.451	0.065	0.144	1.62	-1.873	-0.1	0.554	Up
TMEM109	19451	0.451	0.062	0.137	-1.307	2.171	-0.185	0.554	Up
PCSK7	14699	0.451	0.036	0.081	-2.068	1.339	0.553	0.554	Up
LOXL1	11682	0.451	0.104	0.23	-1.849	0.221	0.889	0.554	Up
COL8A1	4001	0.451	0.129	0.286	-1.749	-0.104	0.967	0.554	Up
PLAC9	15124	0.451	0.071	0.157	-1.908	1.171	0.536	0.554	Up
SNORA16B	18173	0.451	0.053	0.118	1.842	-1.598	-0.329	0.554	Up
C11orf93	1784	0.451	0.043	0.095	-1.68	1.994	0.083	0.554	Up
ZFAND5	21174	0.45	0.046	0.102	-2	0.282	0.944	0.554	Up
BAIAP3	1384	0.45	0.117	0.259	1.841	-1.129	-0.518	0.554	Up
MAP7D3	12072	0.45	0.041	0.091	-2.038	1.241	0.576	0.554	Up
SLC36A1	17880	0.45	0.062	0.139	-1.853	-0.056	1.002	0.554	Up
EMX2OS	5538	0.449	0.079	0.175	-1.915	0.943	0.632	0.554	Up
TBCE	19050	0.449	0.053	0.119	1.83	-1.6	-0.322	0.554	Up
JUN	8810	0.449	0.095	0.211	-1.756	1.479	0.331	0.554	Up
HOXA7	8093	0.449	0.126	0.28	-1.782	1.304	0.416	0.554	Up
PPP1R15A	15492	0.449	0.053	0.118	-1.748	1.754	0.216	0.554	Up
CAPN5	2876	0.449	0.078	0.174	-1.92	0.838	0.677	0.554	Up
ZFP36L1	21192	0.449	0.072	0.16	-1.854	1.333	0.442	0.554	Up
CYHR1	4466	0.449	0.051	0.114	1.947	-1.283	-0.511	0.554	Up
RCBTB2	16367	0.449	0.052	0.115	-1.85	1.566	0.346	0.554	Up
C12orf75	1834	0.449	0.086	0.191	-1.906	0.846	0.666	0.554	Up
CCDC82	3102	0.448	0.074	0.164	-1.905	1.064	0.578	0.554	Up
FKBP9	6419	0.448	0.033	0.073	-2.108	0.293	0.997	0.554	Up
PBX1	14585	0.448	0.059	0.131	1.97	-0.753	-0.738	0.554	Up
DECR2	4722	0.448	0.054	0.12	1.975	-1.013	-0.635	0.554	Up
PRELP	15614	0.447	0.025	0.056	-2.113	1.819	0.383	0.554	Up
GPC6	7333	0.447	0.076	0.17	-1.877	0.255	0.89	0.554	Up
IGFBP6	8408	0.446	0.083	0.186	-1.831	1.279	0.452	0.554	Up
SLIT2	18035	0.446	0.067	0.15	-1.936	0.741	0.725	0.554	Up
PGM5	14896	0.446	0.053	0.118	-1.756	1.713	0.237	0.554	Up
B4GALT3	1352	0.446	0.061	0.138	1.951	-0.744	-0.731	0.554	Up
KRT5	9347	0.446	0.197	0.443	-1.688	1.418	0.32	0.833	Up
MGC50722	12415	0.446	0.067	0.151	1.851	-1.328	-0.443	0.833	Up
SGMS2	17483	0.446	0.057	0.128	-1.931	0.317	0.893	0.833	Up
ALDH1A3	513	0.446	0.106	0.238	-1.403	1.93	-0.038	0.833	Up
CSRP2	4277	0.445	0.084	0.189	-1.865	1.097	0.543	0.833	Up
FHAD1	6379	0.445	0.03	0.068	-1.986	1.746	0.345	0.833	Up
CTHRC1	4325	0.445	0.117	0.263	-1.833	0.339	0.832	0.833	Up
AKR1B1	486	0.445	0.072	0.161	-1.585	1.825	0.101	0.833	Up
TMEM119	19459	0.445	0.099	0.222	-1.809	1.245	0.454	0.833	Up
SLC26A3	17804	0.445	0.085	0.191	-0.704	2.258	-0.539	0.833	Up
ALDH1L2	516	0.444	0.073	0.164	-1.814	-0.007	0.962	0.833	Up
CXCL2	4403	0.444	0.138	0.311	-1.695	1.452	0.31	0.833	Up
CYP2U1	4503	0.444	0.054	0.122	-1.943	1.103	0.582	0.833	Up
SETDB1	17425	0.444	0.038	0.084	1.952	-1.542	-0.41	0.833	Up
MEIS2	12286	0.444	0.065	0.145	-1.92	0.477	0.823	0.833	Up
FMOD	6645	0.444	0.111	0.249	-1.825	1.081	0.529	0.833	Up
ARPC2	1032	0.444	0.046	0.104	-2.006	0.815	0.732	0.833	Up
EXOC8	5753	0.443	0.039	0.088	1.967	-1.42	-0.467	0.833	Up
STX6	18750	0.443	0.048	0.108	1.957	-1.178	-0.559	0.833	Up
CNIH3	3895	0.443	0.079	0.177	-1.887	0.514	0.79	0.833	Up
PEX19	14840	0.443	0.057	0.128	1.952	-0.757	-0.727	0.833	Up
NVL	13837	0.443	0.069	0.156	1.896	-1.021	-0.59	0.833	Up
SOD3	18286	0.442	0.127	0.288	-1.163	2.05	-0.213	0.833	Up
C22orf13	2313	0.442	0.039	0.087	2.015	-0.313	-0.94	0.833	Up
CNTNAP3	3940	0.442	0.082	0.185	-1.775	1.382	0.381	0.833	Up
CLDN8	3777	0.441	0.091	0.207	-1.216	2.066	-0.191	0.833	Up
NR2F1	13676	0.441	0.137	0.312	-1.587	1.625	0.183	0.833	Up
FBXO28	6241	0.441	0.054	0.121	1.95	-0.932	-0.655	0.833	Up
ENPP2	5562	0.441	0.077	0.174	-1.544	1.817	0.083	0.833	Up
RPS6KA3	16898	0.441	0.058	0.133	-1.833	1.4	0.404	0.833	Up
PLD1	15152	0.441	0.048	0.109	-1.976	0.907	0.679	0.833	Up
HSPG2	8245	0.44	0.046	0.104	-1.943	0.23	0.935	0.833	Up
GJB2	7120	0.44	0.145	0.329	-1.722	-0.017	0.918	0.833	Up
L3MBTL3	9473	0.44	0.048	0.109	-1.689	1.802	0.165	0.833	Up
MRAS	12639	0.44	0.087	0.197	-1.737	1.43	0.341	0.833	Up
APOD	846	0.44	0.181	0.413	-0.431	2.145	-0.638	0.833	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	contrast 1	contrast 2	contrast 3	adjusted *P* value (%)	Direction
COG2	3953	0.44	0.052	0.118	1.933	-1.077	-0.587	0.833	Up
LHFPL2	9644	0.439	0.062	0.142	-1.913	0.864	0.663	0.833	Up
LY96	11899	0.439	0.072	0.163	-1.879	0.966	0.603	0.833	Up
OSGIN2	14363	0.439	0.047	0.107	1.98	-0.668	-0.777	0.833	Up
CDKL5	3408	0.439	0.065	0.147	-1.889	1.03	0.583	0.833	Up
ADPRH	346	0.439	0.058	0.132	-1.915	0.998	0.61	0.833	Up
PYGO2	16044	0.439	0.034	0.077	2.074	-1.06	-0.669	0.833	Up
sep.10	17341	0.439	0.065	0.148	-1.728	1.562	0.283	0.833	Up
EGFL6	5387	0.438	0.125	0.286	-1.821	0.771	0.652	0.833	Up
IQCJ-SCHIP1	8647	0.438	0.066	0.15	-1.768	1.46	0.345	0.833	Up
SPATS2L	18389	0.438	0.07	0.16	-1.831	0.155	0.905	0.833	Up
VEGFB	20656	0.438	0.047	0.108	-1.944	1.105	0.582	0.833	Up
P2RY1	14425	0.438	0.049	0.112	-1.938	1.069	0.593	0.833	Up
SV2B	18822	0.438	0.08	0.183	-1.558	1.754	0.116	0.833	Up
SLAMF9	17652	0.437	0.061	0.141	-1.708	-0.325	1.034	0.833	Up
COL14A1	3966	0.437	0.098	0.224	-1.731	1.363	0.365	0.833	Up
PLEKHH2	15184	0.437	0.051	0.116	-1.949	0.764	0.722	0.833	Up
KLHL12	9226	0.437	0.037	0.085	1.923	-1.508	-0.408	0.833	Up
NPC2	13617	0.437	0.061	0.14	-1.795	1.411	0.379	0.833	Up
COL4A2	3985	0.437	0.073	0.167	-1.874	0.556	0.766	0.833	Up
EBF1	5293	0.436	0.048	0.111	-1.899	1.248	0.501	0.833	Up
LGALS1	9618	0.436	0.075	0.172	-1.858	0.448	0.802	0.833	Up
TWIST1	20277	0.436	0.057	0.13	-1.866	1.211	0.498	0.833	Up
S1PR1	17046	0.436	0.071	0.162	-1.557	1.776	0.106	0.833	Up
DNA2	4986	0.436	0.046	0.105	1.965	-0.862	-0.691	0.833	Up
STK17B	18691	0.435	0.058	0.134	-1.864	1.181	0.509	0.833	Up
KLK4	9266	0.435	0.056	0.128	-1.917	0.611	0.767	0.833	Up
MMP7	12547	0.435	0.195	0.449	-1.766	0.914	0.564	0.833	Up
PKIA	15079	0.435	0.091	0.209	-1.341	1.934	-0.072	0.833	Up
SEMA4G	17316	0.435	0.033	0.077	2.063	-0.581	-0.856	0.833	Up
HYMAI	8289	0.435	0.071	0.162	-1.867	0.904	0.622	0.833	Up
FAM189B	5978	0.435	0.046	0.106	1.867	-1.396	-0.423	0.833	Up
FAM203A	6007	0.435	0.065	0.148	1.889	-0.765	-0.69	0.833	Up
CDH13	3350	0.435	0.067	0.154	-1.837	0.217	0.883	0.833	Up
SDHC	17243	0.435	0.047	0.108	1.931	-1.097	-0.578	0.833	Up
LOC728875	11559	0.435	0.082	0.189	-1.842	0.476	0.782	0.833	Up
GRID1	7502	0.435	0.023	0.053	-2.136	1.658	0.46	0.833	Up
FAM69A	6085	0.435	0.045	0.103	-1.921	0.236	0.921	0.833	Up
PDE1B	14728	0.434	0.042	0.097	-1.888	1.408	0.43	0.833	Up
PRR19	15746	0.434	0.056	0.129	1.816	-1.363	-0.41	0.833	Up
LIX1L	9715	0.434	0.04	0.092	-1.81	1.646	0.293	0.833	Up
PALMD	14494	0.434	0.073	0.167	-1.452	1.873	0.011	0.833	Up
DNM1P46	5067	0.434	0.053	0.123	-1.674	1.698	0.2	0.833	Up
LOC100131826	10220	0.433	0.048	0.11	-1.923	0.379	0.864	0.833	Up
C13orf33	1846	0.433	0.093	0.215	-1.768	1.197	0.451	0.833	Up
GABARAPL1	6854	0.433	0.055	0.126	-1.913	0.79	0.693	0.833	Up
ARFGAP2	897	0.433	0.035	0.081	2.026	-0.482	-0.877	0.833	Up
S1PR2	17047	0.433	0.054	0.125	-1.83	1.321	0.434	0.833	Up
ACTBL2	204	0.433	0.069	0.159	-1.801	1.24	0.452	0.833	Up
ANKH	647	0.432	0.069	0.16	-1.863	0.825	0.652	0.833	Up
EPB41L2	5587	0.432	0.058	0.133	-1.884	0.975	0.602	0.833	Up
A2M	5	0.432	0.084	0.195	-1.821	0.959	0.576	0.833	Up
TPST1	19874	0.432	0.067	0.155	-1.803	0.118	0.906	0.833	Up
SYPL2	18871	0.432	0.04	0.092	-1.744	1.747	0.217	0.833	Up
PCDH18	14608	0.432	0.052	0.12	-1.919	0.661	0.748	0.833	Up
MRPL21	12671	0.432	0.06	0.139	1.864	-0.363	-0.839	0.833	Up
LATS2	9526	0.432	0.043	0.1	-1.922	1.192	0.535	0.833	Up
FAM101B	5825	0.432	0.074	0.171	-1.836	0.417	0.803	0.833	Up
CSDC2	4233	0.432	0.042	0.097	-1.895	1.334	0.463	0.833	Up
HAX1	7724	0.431	0.045	0.105	1.949	-0.804	-0.706	0.833	Up
COL7A1	4000	0.431	0.093	0.216	-1.755	0.111	0.883	0.833	Up

Add. Table 4: Significantly different expressed genes between the metabolic clusters Mc1 and Mc2

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
GEM	7035	5.207	0.921	0.177	-0.455	0	Up
PLS3	15215	4.912	0.883	0.18	-0.713	0	Up
SKAP2	17634	4.875	0.616	0.126	-0.474	0	Up
TAGLN	18949	4.872	1.01	0.207	-0.586	0	Up
LMOD1	9739	4.865	0.467	0.096	-0.444	0	Up
PDLIM4	14775	4.745	0.559	0.118	-0.642	0	Up
EGR2	5397	4.735	0.822	0.174	-0.347	0	Up
CXorf57	4437	4.677	0.635	0.136	-1.253	0	Up
FERMT2	6321	4.645	0.588	0.127	-0.458	0	Up
ACTA2	202	4.631	1.035	0.224	-0.638	0	Up
GSN	7558	4.585	0.674	0.147	-0.368	0	Up
CYR61	4528	4.584	1.021	0.223	-0.451	0	Up
LIMS2	9684	4.544	0.604	0.133	-0.551	0	Up
CALD1	2825	4.528	0.783	0.173	-0.442	0	Up
BHLHE41	1507	4.501	0.851	0.189	-0.439	0	Up
EGR1	5396	4.461	1.151	0.258	-0.42	0	Up
CAV2	2944	4.439	0.446	0.1	-0.652	0	Up
FOXO1	6712	4.435	0.525	0.118	-0.591	0	Up
ISM1	8696	4.4	0.931	0.212	-0.515	0	Up
CNN1	3900	4.382	1.252	0.286	-0.764	0	Up
CTGF	4323	4.362	0.963	0.221	-0.484	0	Up
CD200	3228	4.351	0.417	0.096	-0.382	0	Up
FMO2	6639	4.303	1.205	0.28	-0.607	0	Up
CLDN11	3758	4.294	0.904	0.211	-0.611	0	Up
NEXN	13395	4.25	0.556	0.131	-0.384	0	Up
TGFB1I1	19244	4.226	0.601	0.142	-0.467	0	Up
PROS1	15712	4.222	0.433	0.103	-0.312	0	Up
AKAP12	465	4.22	0.697	0.165	-0.582	0	Up
MSRB3	12776	4.208	0.826	0.196	-0.453	0	Up
LTBP2	11867	4.195	0.533	0.127	-0.367	0	Up
C10orf10	1681	4.186	0.788	0.188	-0.481	0	Up
CCDC8	3099	4.185	0.74	0.177	-0.617	0	Up
GRASP	7484	4.165	0.589	0.141	-0.415	0	Up
FEZ1	6326	4.158	0.468	0.113	-0.7	0	Up
RHOJ	16542	4.151	0.584	0.141	-0.402	0	Up
MYLK	12958	4.146	0.881	0.212	-0.513	0	Up
CSRP1	4276	4.14	0.408	0.099	-0.813	0	Up
ARRDC3	1046	4.12	0.555	0.135	-0.683	0	Up
PTPN21	15959	4.113	0.644	0.156	-0.629	0	Up
CD36	3255	4.103	1.181	0.288	-0.591	0	Up
RGS2	16502	4.097	0.95	0.232	-0.935	0	Up
MRVI1	12739	4.097	0.683	0.167	-0.492	0	Up
ANXA1	754	4.094	0.828	0.202	-0.407	0	Up
NDN	13282	4.093	0.691	0.169	-0.499	0	Up
FXYD6	6823	4.069	0.55	0.135	-1.301	0	Up
C9orf125	2684	4.064	0.503	0.124	-0.409	0	Up
S100A10	17025	4.064	0.695	0.171	-0.475	0	Up
ETS2	5717	4.06	0.475	0.117	-0.95	0	Up
EMP1	5529	4.052	0.722	0.178	-0.643	0	Up
C5orf62	2509	4.049	0.721	0.178	-0.473	0	Up
FOSB	6676	4.047	1.253	0.31	-0.506	0	Up
PLK3	15203	4.044	0.499	0.123	-0.252	0	Up
PTRF	15990	4.038	0.421	0.104	-0.309	0	Up
MIR100HG	12463	4.034	0.739	0.183	-0.306	0	Up
DPYSL3	5150	4.029	0.88	0.219	-0.414	0	Up
TCF4	19112	4.025	0.606	0.15	-0.35	0	Up
CAV1	2943	4.02	0.757	0.188	-0.59	0	Up
CLMP	3830	4.02	0.701	0.174	-0.146	0	Up
DLC1	4936	4.019	0.353	0.088	-0.324	0	Up
RASD1	16251	4.019	1.088	0.271	-0.972	0	Up
DKK3	4932	4.015	0.571	0.142	-0.434	0	Up
CNN2	3901	3.995	0.809	0.202	-0.448	0	Up
FOS	6675	3.994	1.125	0.282	-0.618	0	Up
SVEP1	18824	3.99	0.418	0.105	-0.36	0	Up
GPX8	7474	3.987	0.697	0.175	-0.271	0	Up
NAP1L3	13066	3.984	0.496	0.124	-0.804	0	Up
DACT3	4555	3.969	0.78	0.197	-0.417	0	Up
RARRES2	16240	3.958	0.901	0.228	-0.376	0	Up
APCDD1	809	3.947	0.819	0.207	-0.647	0	Up
PDLIM3	14774	3.946	0.687	0.174	-0.718	0	Up
MB21D2	12148	3.937	0.473	0.12	-1.485	0	Up
PODN	15296	3.924	0.552	0.141	-0.395	0	Up
ADH1C	324	3.906	1.069	0.274	-0.719	0	Up
FNDC4	6658	3.904	0.485	0.124	-0.704	0	Up
MYL9	12956	3.901	0.759	0.195	-0.551	0	Up
SHOX2	17566	3.891	0.488	0.125	-0.334	0	Up
ARL4C	996	3.884	0.674	0.173	-0.562	0	Up
HSPB2	8233	3.869	0.815	0.211	-0.644	0	Up
GADD45B	6883	3.86	0.577	0.15	-0.418	0	Up
TNS1	19782	3.856	0.432	0.112	-0.133	0	Up
FBLN5	6200	3.845	0.737	0.192	-0.371	0	Up
FOXF2	6694	3.841	0.378	0.099	-0.809	0	Up
MPZL2	12634	3.839	0.726	0.189	-1.306	0	Up
DPYSL2	5149	3.837	0.577	0.15	-0.93	0	Up
NPC2	13617	3.818	0.408	0.107	-0.642	0	Up
H19	7663	3.817	0.844	0.221	-0.723	0	Up
PRKD1	15660	3.811	0.464	0.122	-0.551	0	Up
CELF2	3478	3.81	0.432	0.113	-0.288	0	Up
SGK1	17475	3.805	0.602	0.158	-0.437	0	Up
NNMT	13555	3.798	0.808	0.213	-0.253	0	Up
PGM5	14896	3.797	0.365	0.096	-0.6	0	Up
SERPING1	17402	3.796	0.478	0.126	-0.378	0	Up
FBLN2	6199	3.79	0.883	0.233	-0.392	0	Up
SEMA5A	17317	3.789	0.673	0.178	-0.431	0	Up
SNAI2	18133	3.781	0.508	0.134	-0.262	0	Up
SFRP4	17449	3.777	0.866	0.229	-0.933	0	Up
LIMS3L	9686	3.762	0.401	0.107	-0.429	0	Up
PRKCA	15649	3.758	0.67	0.178	-0.51	0	Up
BACH2	1366	3.756	0.29	0.077	-0.628	0	Up
PID1	14982	3.754	0.644	0.171	-0.732	0	Up
D4S234E	4545	3.751	0.64	0.171	-1.094	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
MMP19	12536	3.747	0.332	0.089	-0.074	0	Up
TNFRSF10D	19723	3.743	0.526	0.14	-0.781	0	Up
VEGFC	20657	3.738	0.499	0.133	-0.685	0	Up
IGFBP3	8405	3.737	0.608	0.163	-0.271	0	Up
PODNL1	15297	3.732	0.535	0.143	-0.585	0	Up
KRT17	9319	3.729	1.558	0.418	-0.797	0	Up
SV2B	18822	3.725	0.565	0.152	-0.662	0	Up
PTGIS	15923	3.723	0.697	0.187	-0.622	0	Up
CFI	3568	3.722	0.577	0.155	-0.55	0	Up
MOXD1	12603	3.715	0.723	0.195	-0.621	0	Up
VCL	20646	3.712	0.454	0.122	-1.301	0	Up
KRT14	9315	3.711	1.613	0.435	-0.791	0	Up
KRT16P2	9318	3.705	1.503	0.406	-0.783	0	Up
DOCK11	5084	3.703	0.5	0.135	-0.406	0	Up
FAM89A	6139	3.702	0.602	0.163	-0.506	0	Up
PDLIM7	14777	3.701	0.519	0.14	-0.566	0	Up
RND3	16628	3.696	0.81	0.219	-1.154	0	Up
PLSCR4	15219	3.682	0.512	0.139	-0.646	0	Up
LPCAT2	11695	3.679	0.331	0.09	-0.395	0	Up
DST	5189	3.677	0.611	0.166	-0.362	0	Up
CADM3	2813	3.673	0.401	0.109	-1.196	0	Up
RNF175	16670	3.666	0.308	0.084	-0.637	0	Up
LAYN	9528	3.665	0.545	0.149	-0.663	0	Up
SERPINF1	17400	3.663	0.722	0.197	-0.318	0	Up
FABP4	5804	3.662	0.88	0.24	-0.643	0	Up
KRT42P	9346	3.652	0.557	0.152	-0.924	0	Up
CXCL2	4403	3.651	0.939	0.257	-0.444	0	Up
CDO1	3421	3.65	0.766	0.21	-0.672	0	Up
ITPRIP	8761	3.649	0.509	0.14	-0.307	0	Up
KLF2	9204	3.649	0.579	0.159	-0.404	0	Up
SWAP70	18829	3.644	0.379	0.104	-1.081	0	Up
TGFBR2	19249	3.64	0.558	0.153	-0.634	0	Up
CCDC3	3053	3.638	0.713	0.196	-0.493	0	Up
SCN3B	17192	3.633	0.505	0.139	-0.372	0	Up
L3MBTL3	9473	3.633	0.337	0.093	-1.128	0	Up
SPARCL1	18363	3.632	0.831	0.229	-0.467	0	Up
FOXQ1	6719	3.629	0.865	0.238	-0.694	0	Up
AOC3	770	3.622	0.742	0.205	-0.418	0	Up
LHFP	9642	3.62	0.675	0.186	-0.291	0	Up
ARAP3	887	3.618	0.497	0.137	-0.462	0	Up
GAS1	6942	3.613	0.771	0.213	-0.526	0	Up
MALL	12005	3.61	0.923	0.256	-1.181	0	Up
AIM1	446	3.602	0.775	0.215	-2.535	0	Up
MFAP4	12347	3.602	1.096	0.304	-0.508	0	Up
ZFP36	21191	3.602	0.669	0.186	-0.367	0	Up
TSC22D3	20069	3.6	0.458	0.127	-0.652	0	Up
TSHZ2	20085	3.595	0.498	0.138	-0.42	0	Up
SOBP	18276	3.595	0.321	0.089	-0.327	0	Up
HOXA3	8089	3.591	0.708	0.197	-0.438	0	Up
MMP23B	12540	3.583	0.618	0.172	-0.756	0	Up
PRKCDBP	15652	3.581	0.632	0.176	-0.449	0	Up
SPRY2	18483	3.569	0.665	0.186	-0.431	0	Up
MYO1B	12972	3.565	0.385	0.108	-0.497	0	Up
SLIT3	18036	3.563	0.709	0.199	-0.265	0	Up
LRRN4CL	11831	3.562	0.479	0.135	-0.418	0	Up
C6orf145	2533	3.56	0.517	0.145	-0.868	0	Up
CCL14	3132	3.557	0.998	0.281	-0.76	0	Up
CMTM3	3874	3.556	0.532	0.15	-0.588	0	Up
C1S	2133	3.552	0.664	0.187	-0.309	0	Up
ANXA3	761	3.551	0.999	0.281	-1.096	0	Up
HOXA4	8090	3.546	0.361	0.102	-0.309	0	Up
PECAM1	14810	3.546	0.457	0.129	-0.399	0	Up
F2RL1	5787	3.542	0.534	0.151	-0.718	0	Up
FAS	6180	3.541	0.355	0.1	-0.259	0	Up
METTL7A	12335	3.532	0.482	0.137	-1.15	0	Up
PRKD3	15662	3.524	0.518	0.147	-0.76	0	Up
ADH1A	322	3.524	0.874	0.248	-0.594	0	Up
RBP4	16353	3.52	1.024	0.291	-0.5	0	Up
TUBB6	20251	3.517	0.665	0.189	-0.353	0	Up
EDN1	5334	3.513	0.765	0.218	-0.772	0	Up
VSNL1	20740	3.511	0.692	0.197	-0.697	0	Up
MT1X	12798	3.51	0.682	0.194	-0.254	0	Up
CLDN8	3777	3.51	0.638	0.182	-1.386	0	Up
HIC1	7872	3.509	0.463	0.132	-0.65	0	Up
GPR124	7367	3.499	0.554	0.158	-0.201	0	Up
CCDC80	3100	3.498	0.66	0.189	-0.35	0	Up
ADAMTS1	269	3.497	0.662	0.189	-0.566	0	Up
C1R	2131	3.496	0.648	0.185	-0.318	0	Up
FIBIN	6389	3.496	0.646	0.185	-0.383	0	Up
ARL4A	995	3.494	0.368	0.105	-1.084	0	Up
PRICKLE1	15627	3.488	0.79	0.226	-0.535	0	Up
TIMP2	19344	3.486	0.707	0.203	-0.281	0	Up
PRELP	15614	3.483	0.17	0.049	-0.483	0	Up
EHD2	5403	3.472	0.506	0.146	-0.385	0	Up
CPXM1	4137	3.47	0.546	0.157	-0.545	0	Up
SSPN	18583	3.467	0.596	0.172	-0.398	0	Up
GNG11	7241	3.465	0.597	0.172	-0.337	0	Up
CFH	3562	3.461	0.575	0.166	-0.27	0	Up
IQCJ-SCHIP1	8647	3.456	0.444	0.128	-0.699	0	Up
TPM2	19853	3.456	0.589	0.17	-0.403	0	Up
ADD3	321	3.45	0.68	0.197	-0.465	0	Up
KLF12	9198	3.442	0.436	0.127	-0.652	0.324	Up
ANKRD35	695	3.44	0.805	0.234	-0.463	0.324	Up
NID1	13456	3.438	0.645	0.188	-0.254	0.324	Up
KRT5	9347	3.432	1.336	0.389	-0.648	0.324	Up
TMEM200A	19554	3.431	0.449	0.131	-0.355	0.324	Up
FAM43A	6040	3.43	0.485	0.141	-0.508	0.324	Up
TWIST2	20278	3.43	0.655	0.191	-0.288	0.324	Up
C11orf96	1787	3.426	0.642	0.187	-0.414	0.324	Up
KIF26A	9151	3.425	0.382	0.112	-0.859	0.324	Up
HOXA7	8093	3.423	0.826	0.241	-0.47	0.324	Up
FSTL1	6776	3.422	0.745	0.218	-0.369	0.324	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
ENPP2	5562	3.418	0.541	0.158	-0.543	0.324	Up
GRP	7537	3.415	1.026	0.3	-0.79	0.324	Up
PDGFRA	14754	3.407	0.697	0.205	-0.232	0.324	Up
ZC3H12B	21103	3.407	0.253	0.074	-0.396	0.324	Up
WDR86	20856	3.406	0.817	0.24	-0.426	0.324	Up
PDE2A	14730	3.405	0.621	0.182	-0.516	0.324	Up
FLNA	6617	3.405	0.386	0.113	-0.554	0.324	Up
PKD2	15071	3.404	0.442	0.13	-0.421	0.324	Up
CLDN5	3774	3.401	0.749	0.22	-0.668	0.324	Up
SOD3	18286	3.399	0.885	0.26	-1.124	0.324	Up
SLC12A4	17666	3.397	0.384	0.113	-0.572	0.324	Up
AKR1C1	489	3.396	0.661	0.195	-0.87	0.324	Up
LRCH2	11717	3.394	0.275	0.081	-0.438	0.324	Up
PPP1R15A	15492	3.394	0.369	0.109	-0.957	0.324	Up
RUNX1T1	16999	3.393	0.608	0.179	-0.307	0.324	Up
GRID1	7502	3.391	0.155	0.046	-0.431	0.324	Up
SCARNA17	17147	3.388	0.658	0.194	-0.373	0.324	Up
HEG1	7809	3.38	0.511	0.151	-0.519	0.324	Up
WTIP	20938	3.378	0.553	0.164	-0.89	0.324	Up
TCEAL7	19092	3.377	0.441	0.131	-1.161	0.324	Up
HABP4	7685	3.376	0.448	0.133	-0.347	0.324	Up
MME	12524	3.376	0.588	0.174	-0.393	0.324	Up
COL14A1	3966	3.376	0.654	0.194	-0.613	0.324	Up
NOV	13598	3.375	0.734	0.218	-0.923	0.324	Up
KIAA0408	9027	3.373	0.346	0.103	-0.3	0.324	Up
ZNF300P1	21390	3.369	0.547	0.162	-0.396	0.324	Up
DUSP1	5217	3.369	0.669	0.198	-0.37	0.324	Up
SNX3	18263	3.359	0.283	0.084	-1.924	0.324	Up
ARHGAP28	926	3.359	0.379	0.113	-0.153	0.324	Up
AKR1B1	486	3.358	0.506	0.151	-0.939	0.324	Up
ODZ4	13894	3.354	0.255	0.076	-0.494	0.324	Up
SNCA	18152	3.349	0.344	0.103	-0.864	0.324	Up
CNTNAP3	3940	3.338	0.545	0.163	-0.438	0.324	Up
DNM1P46	5067	3.336	0.372	0.112	-0.542	0.324	Up
LRIG3	11730	3.336	0.578	0.173	-0.336	0.324	Up
DCN	4634	3.335	0.835	0.25	-0.323	0.324	Up
KLK8	9270	3.335	1.536	0.461	-1.787	0.324	Up
C11orf93	1784	3.332	0.304	0.091	-0.943	0.324	Up
JUN	8810	3.325	0.641	0.193	-0.424	0.324	Up
NHS	13452	3.322	0.296	0.089	-0.472	0.324	Up
VIM	20673	3.321	0.51	0.154	-0.175	0.324	Up
ALPK2	569	3.32	0.282	0.085	-1.512	0.324	Up
TSHZ3	20086	3.319	0.531	0.16	-0.084	0.324	Up
COL18A1	3970	3.318	0.593	0.179	-0.16	0.324	Up
JAM2	8783	3.316	0.443	0.133	-0.33	0.324	Up
ZEB1	21166	3.307	0.523	0.158	-0.159	0.324	Up
S1PR1	17046	3.306	0.497	0.15	-0.648	0.324	Up
SH3PXD2A	17528	3.302	0.46	0.139	-0.318	0.324	Up
CDKN1C	3411	3.301	0.604	0.183	-0.577	0.324	Up
SCG5	17165	3.299	0.531	0.161	-0.094	0.324	Up
SCHIP1	17177	3.293	0.611	0.185	-0.347	0.536	Up
KCNH8	8885	3.292	0.464	0.141	-0.775	0.536	Up
MMP2	12537	3.289	0.682	0.207	-0.446	0.536	Up
LOR	11679	3.287	0.225	0.068	-0.791	0.536	Up
CNRIP1	3925	3.285	0.274	0.083	-0.377	0.536	Up
MRC2	12641	3.284	0.657	0.2	-0.169	0.536	Up
PTN	15939	3.283	1.088	0.332	-0.77	0.536	Up
CCND2	3163	3.282	0.481	0.147	-0.669	0.536	Up
TRPC1	20038	3.282	0.469	0.143	-0.157	0.536	Up
ID3	8312	3.28	0.426	0.13	-0.581	0.536	Up
ETHE1	5713	3.28	0.378	0.115	-1.233	0.536	Up
LOC150622	10676	3.27	0.397	0.122	-2.746	0.536	Up
FHAD1	6379	3.259	0.208	0.064	-0.882	0.536	Up
ISLR	8694	3.259	0.686	0.211	-0.316	0.536	Up
ALOX5	564	3.258	0.308	0.095	-1.365	0.536	Up
C14orf139	1865	3.256	0.424	0.13	-0.538	0.536	Up
FGF1	6341	3.252	0.202	0.062	-0.266	0.536	Up
PCSK7	14699	3.248	0.231	0.071	-0.639	0.536	Up
PTGS2	15927	3.246	0.308	0.095	-0.776	0.536	Up
PKIA	15079	3.24	0.643	0.198	-1.226	0.536	Up
CYTH3	4535	3.24	0.565	0.174	-0.264	0.536	Up
EFHA2	5366	3.238	0.383	0.118	-0.044	0.536	Up
KCNMB4	8924	3.238	0.572	0.177	-0.352	0.536	Up
SRPX	18546	3.238	0.651	0.201	-0.119	0.536	Up
CFHR3	3565	3.237	0.507	0.157	-0.311	0.536	Up
SYPL2	18871	3.236	0.277	0.085	-1.091	0.536	Up
GPX3	7469	3.236	0.635	0.196	-0.642	0.536	Up
JAG1	8774	3.235	0.397	0.123	-0.372	0.536	Up
RPS6KA3	16898	3.233	0.387	0.12	-0.399	0.536	Up
CES1	3545	3.228	0.477	0.148	-0.157	0.536	Up
ITGB1	8727	3.227	0.364	0.113	-0.516	0.536	Up
PALMD	14494	3.226	0.514	0.159	-0.535	0.536	Up
FZD7	6836	3.226	0.639	0.198	-0.519	0.536	Up
NR2F1	13676	3.225	0.96	0.298	-0.664	0.536	Up
COL15A1	3967	3.225	0.489	0.152	-0.267	0.536	Up
FXYD1	6818	3.223	0.419	0.13	-0.626	0.536	Up
RCBTB2	16367	3.221	0.351	0.109	-0.447	0.536	Up
C17orf51	2000	3.221	0.427	0.133	-0.235	0.536	Up
JAZF1	8786	3.22	0.356	0.111	-0.412	0.536	Up
IGFBP6	8408	3.22	0.538	0.167	-0.447	0.536	Up
PELI2	14817	3.218	0.473	0.147	-0.323	0.536	Up
CKMT2	3734	3.212	0.299	0.093	-0.391	0.536	Up
VCAN	20645	3.21	0.727	0.226	-0.376	0.536	Up
MATN2	12138	3.209	0.611	0.19	-0.679	0.536	Up
AQP7P1	876	3.207	0.434	0.135	-0.448	0.536	Up
TMEM45B	19611	3.206	0.923	0.288	-1.778	0.536	Up
SLC2A14	17829	3.204	0.388	0.121	-0.086	0.536	Up
SFRP2	17448	3.203	0.882	0.275	-0.43	0.536	Up
KANK2	8818	3.203	0.418	0.131	-0.502	0.536	Up
ANGPTL2	633	3.202	0.608	0.19	-0.138	0.536	Up
ANTXR2	751	3.201	0.31	0.097	-0.379	0.536	Up
GLI1	7149	3.198	0.515	0.161	-0.698	0.536	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
LIX1L	9715	3.197	0.274	0.086	-0.63	0.536	Up
MAP7D3	12072	3.196	0.255	0.08	-0.325	0.536	Up
FAT4	6191	3.196	0.222	0.07	-0.309	0.536	Up
ALDH1A3	513	3.196	0.751	0.235	-0.743	0.536	Up
SLFN13	18029	3.195	0.512	0.16	-0.587	0.536	Up
FAT1	6188	3.194	0.653	0.204	-0.45	0.536	Up
PCSK5	14697	3.193	0.275	0.086	-0.211	0.536	Up
SMAD9	18062	3.192	0.238	0.075	-0.344	0.536	Up
C21orf34	2296	3.189	0.294	0.092	-0.148	0.536	Up
PDE1B	14728	3.186	0.277	0.087	-0.35	0.536	Up
TP53I3	19830	3.186	0.459	0.144	-0.808	0.536	Up
LSP1	11857	3.184	0.41	0.129	-0.218	0.536	Up
TMEM109	19451	3.184	0.433	0.136	-1.557	0.536	Up
F3	5790	3.182	0.76	0.239	-0.556	0.536	Up
FGF7	6360	3.18	0.368	0.116	-0.235	0.536	Up
LAMA4	9489	3.176	0.365	0.115	-0.156	0.536	Up
LCAT	9536	3.173	0.371	0.117	-0.192	0.536	Up
PLAC9	15124	3.162	0.443	0.14	-0.354	0.536	Up
COL16A1	3968	3.16	0.418	0.132	-0.321	0.536	Up
TRIM29	19958	3.155	0.943	0.299	-1.003	0.81	Up
IGF2	8396	3.153	0.415	0.132	-0.422	0.81	Up
DAB2	4549	3.149	0.458	0.145	-0.212	0.81	Up
MRGPRF	12648	3.148	0.293	0.093	-0.835	0.81	Up
BGN	1500	3.145	0.508	0.161	-0.47	0.81	Up
PDPN	14781	3.144	0.565	0.18	-0.451	0.81	Up
RNF144A	16653	3.144	0.39	0.124	-0.053	0.81	Up
LOC100190939	10363	3.143	0.477	0.152	-0.647	0.81	Up
CSDC2	4233	3.143	0.272	0.087	-0.932	0.81	Up
NMB	13533	3.142	0.599	0.191	-0.527	0.81	Up
TNNI2	19766	3.141	0.91	0.29	-0.59	0.81	Up
C3	2389	3.14	0.648	0.206	-0.823	0.81	Up
JAM3	8784	3.138	0.342	0.109	-0.323	0.81	Up
ACTBL2	204	3.138	0.444	0.141	-0.655	0.81	Up
RCAN2	16364	3.134	0.356	0.114	-0.456	0.81	Up
SLC38A2	17891	3.133	0.332	0.106	-0.555	0.81	Up
MRAS	12639	3.133	0.585	0.187	-0.421	0.81	Up
MAP1B	12033	3.132	0.401	0.128	-0.171	0.81	Up
SPRY1	18482	3.132	0.554	0.177	-0.453	0.81	Up
VPS36	20715	3.128	0.32	0.102	-0.983	0.81	Up
TMEM119	19459	3.128	0.64	0.205	-0.66	0.81	Up
CTSK	4353	3.128	0.704	0.225	-0.427	0.81	Up
SPNS2	18449	3.128	0.672	0.215	-0.806	0.81	Up
TMEM43	19608	3.126	0.266	0.085	-0.258	0.81	Up
NEFM	13353	3.125	0.282	0.09	-0.63	0.81	Up
RBMS1	16338	3.125	0.472	0.151	-0.189	0.81	Up
SDPR	17247	3.125	0.61	0.195	-0.316	0.81	Up
NEFH	13351	3.124	0.552	0.177	-0.22	0.81	Up
MAGEL2	11991	3.123	0.426	0.136	-0.503	0.81	Up
TCF7	19113	3.123	0.404	0.129	-1.431	0.81	Up
ABCA6	42	3.122	0.535	0.171	-0.095	0.81	Up
MT1M	12797	3.118	0.809	0.259	-0.129	0.81	Up
DNAJB5	5018	3.117	0.337	0.108	-0.299	0.81	Up
CH25H	3585	3.11	0.634	0.204	-0.333	0.81	Up
GGT5	7072	3.109	0.536	0.172	-0.333	0.81	Up
CXCL12	4398	3.106	0.427	0.138	-0.267	0.81	Up
ZFP36L2	21193	3.106	0.507	0.163	-0.556	0.81	Up
ARHGAP20	918	3.104	0.333	0.107	-0.296	0.81	Up
SRPX2	18547	3.103	0.631	0.203	-0.273	0.81	Up
sep.10	17341	3.103	0.445	0.143	-0.994	0.81	Up
ROBO3	16733	3.102	0.316	0.102	-0.583	0.81	Up
IL8	8535	3.086	0.244	0.079	-1.134	0.81	Up
MXRA5	12901	3.085	0.43	0.139	-0.277	0.81	Up
AGPAT4	404	3.084	0.331	0.107	-0.174	0.81	Up
RSPO3	16963	3.082	0.405	0.131	-0.396	0.81	Up
GLT8D2	7182	3.082	0.357	0.116	-0.141	0.81	Up
AKR1B15	488	3.08	0.778	0.253	-2.532	0.81	Up
SHF	17550	3.08	0.257	0.083	-1.88	0.81	Up
AKAP13	466	3.078	0.286	0.093	-0.338	0.81	Up
MEG3	12275	3.07	0.408	0.133	-0.125	0.81	Up
PRNP	15688	3.07	0.566	0.184	-0.285	0.81	Up
TPM4	19855	3.069	0.453	0.147	-0.2	0.81	Up
ADRB2	356	3.069	0.535	0.174	-0.868	0.81	Up
ANTXR1	750	3.065	0.54	0.176	-0.481	0.81	Up
UACA	20323	3.064	0.353	0.115	-0.106	0.81	Up
ZFHX4	21181	3.064	0.525	0.171	-0.325	0.81	Up
DACT1	4553	3.062	0.624	0.204	-0.403	0.81	Up
EFNB2	5378	3.062	0.462	0.151	-0.583	0.81	Up
ZCCHC24	21130	3.06	0.512	0.167	-0.83	0.81	Up
PALLD	14490	3.06	0.384	0.126	-0.333	0.81	Up
MYOM1	12994	3.055	0.303	0.099	-0.375	0.81	Up
ROR1	16743	3.055	0.298	0.098	-0.227	0.81	Up
AASS	27	3.053	0.371	0.121	-0.639	0.81	Up
KLF6	9208	3.053	0.318	0.104	-0.467	0.81	Up
BHLHE22	1504	3.05	0.508	0.167	-0.439	0.81	Up
NT5E	13747	3.045	0.354	0.116	-0.682	0.81	Up
IER2	8328	3.045	0.447	0.147	-0.825	0.81	Up
PDGFRB	14755	3.045	0.524	0.172	-0.129	0.81	Up
MRAP2	12638	3.041	0.38	0.125	-1.119	0.81	Up
EPGN	5599	3.031	0.196	0.065	-0.87	0.81	Up
MDFIC	12218	3.027	0.349	0.115	-0.016	0.81	Up
A2M	5	3.027	0.506	0.167	-0.289	0.81	Up
COL6A2	3995	3.023	0.615	0.204	-0.465	0.81	Up
ATF3	1147	3.023	0.484	0.16	-0.228	0.81	Up
BBOX1	1410	3.022	0.849	0.281	-0.984	0.81	Up
ANGPTL4	635	3.022	0.872	0.288	-0.38	0.81	Up
FHL3	6384	3.022	0.248	0.082	-1.279	0.81	Up
GPRASP1	7450	3.021	0.404	0.134	-0.738	0.81	Up
DDR2	4674	3.02	0.417	0.138	-0.334	0.81	Up
C1QTNF4	2124	3.019	0.34	0.113	-0.713	0.81	Up
CDKL5	3408	3.019	0.392	0.13	-0.656	0.81	Up
C20orf103	2234	3.018	0.697	0.231	-0.498	0.81	Up
CPZ	4139	3.017	0.471	0.156	-0.608	0.81	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
SPON1	18456	3.016	0.851	0.282	-0.345	0.81	Up
LOC283392	10795	3.012	0.419	0.139	-0.209	0.81	Up
MGC24103	12397	3.012	0.628	0.209	-0.123	0.81	Up
CLIP3	3821	3.011	0.343	0.114	-0.154	0.81	Up
NGFR	13437	3.01	0.538	0.179	-0.733	0.81	Up
SAA1	17051	3.008	1.287	0.428	-0.707	0.81	Up
IGFBP7	8409	3.008	0.496	0.165	-0.503	1.213	Up
C13orf33	1846	3.005	0.601	0.2	-0.549	1.213	Up
sep.11	17342	3.004	0.474	0.158	-0.33	1.213	Up
CYP2U1	4503	3.003	0.333	0.111	-0.205	1.213	Up
HLF	7999	3.002	0.651	0.217	-0.453	1.213	Up
RARB	16237	3.002	0.206	0.069	-1.064	1.213	Up
MET	12304	2.999	0.75	0.25	-0.475	1.213	Up
DSE	5180	2.999	0.399	0.133	-0.151	1.213	Up
ADAMTS5	283	2.996	0.596	0.199	-0.121	1.213	Up
FOXN3	6710	2.996	0.325	0.108	-0.652	1.213	Up
PCOLCE2	14688	2.996	0.703	0.235	-0.36	1.213	Up
RNF39	16702	2.995	0.362	0.121	-0.804	1.213	Up
CCDC50	3068	2.994	0.262	0.088	-0.986	1.213	Up
LOC572558	11293	2.993	0.221	0.074	-0.754	1.213	Up
KLK7	9269	2.993	1.023	0.342	-1.347	1.213	Up
EBF1	5293	2.992	0.31	0.104	-0.274	1.213	Up
LATS2	9526	2.991	0.27	0.09	-0.518	1.213	Up
KRT6C	9350	2.991	0.785	0.263	-1.191	1.213	Up
C3orf57	2423	2.988	0.378	0.127	-0.955	1.213	Up
SLC2A3	17831	2.987	0.198	0.066	0.016	1.213	Up
NPR2	13642	2.987	0.387	0.13	-0.252	1.213	Up
GFPT2	7054	2.986	0.503	0.169	-0.048	1.213	Up
KRT6B	9349	2.986	1.383	0.463	-0.787	1.213	Up
C9orf53	2729	2.984	0.634	0.212	-0.35	1.213	Up
ZEB2	21168	2.98	0.344	0.115	-0.097	1.213	Up
DZIP1	5276	2.98	0.411	0.138	-0.427	1.213	Up
MYH11	12932	2.98	0.967	0.324	-0.779	1.213	Up
CR2	4142	2.978	0.641	0.215	-2.027	1.213	Up
DZIP1L	5277	2.978	0.319	0.107	-0.14	1.213	Up
SYNPO	18864	2.977	0.34	0.114	-0.424	1.213	Up
EPHB6	5613	2.976	0.438	0.147	-0.589	1.213	Up
IGF1	8394	2.976	0.614	0.206	-0.406	1.213	Up
HIF1A	7874	2.976	0.406	0.136	-0.114	1.213	Up
CSRP2	4277	2.974	0.523	0.176	-0.475	1.213	Up
LAMA1	9486	2.972	0.312	0.105	-0.252	1.213	Up
ZFP36L1	21192	2.967	0.468	0.158	-0.547	1.213	Up
MMRN1	12550	2.965	0.461	0.155	-1.142	1.213	Up
GPIHBP1	7344	2.965	0.647	0.218	-0.259	1.213	Up
FILIP1L	6399	2.965	0.306	0.103	-1.125	1.213	Up
DEFB1	4736	2.964	0.808	0.273	-0.687	1.213	Up
SOCS5	18281	2.963	0.307	0.104	-0.91	1.213	Up
KCTD12	8949	2.959	0.45	0.152	-0.227	1.213	Up
SH3PXD2B	17529	2.958	0.439	0.148	-0.439	1.213	Up
SNCAIP	18153	2.955	0.167	0.056	-0.825	1.213	Up
EMX2OS	5538	2.955	0.465	0.157	-0.356	1.213	Up
P2RY1	14425	2.954	0.299	0.101	-0.379	1.213	Up
TBX15	19068	2.949	0.255	0.087	-0.131	1.213	Up
SOCS2	18278	2.948	0.81	0.275	-0.965	1.213	Up
FAP	6172	2.947	0.692	0.235	-0.285	1.213	Up
FAM107A	5835	2.946	0.604	0.205	-0.648	1.213	Up
HOXA11-AS1	8086	2.945	0.404	0.137	-0.112	1.213	Up
KLF7	9209	2.941	0.38	0.129	-0.427	1.213	Up
PARVA	14557	2.94	0.37	0.126	-0.558	1.213	Up
MAP4K4	12065	2.938	0.306	0.104	-0.316	1.213	Up
COL3A1	3983	2.935	0.77	0.262	-0.226	1.213	Up
IL33	8517	2.935	0.599	0.204	-1.046	1.213	Up
ACTN1	214	2.934	0.484	0.165	-0.533	1.213	Up
CRISPLD2	4183	2.932	0.371	0.127	-0.502	1.213	Up
CRYAB	4209	2.929	1.02	0.348	-0.43	1.213	Up
LINGO2	9696	2.926	0.31	0.106	-1.445	1.213	Up
C1QTNF5	2125	2.923	0.543	0.186	-0.608	1.213	Up
JDP2	8787	2.923	0.396	0.135	-0.323	1.213	Up
STARD8	18661	2.92	0.287	0.098	-0.238	1.213	Up
ARHGAP10	910	2.919	0.359	0.123	-0.647	1.213	Up
CYLD	4469	2.916	0.272	0.093	-0.194	1.213	Up
STK17B	18691	2.915	0.368	0.126	-0.302	1.213	Up
PPAPDC3	15424	2.914	0.271	0.093	-0.079	1.213	Up
GNAI1	7220	2.914	0.4	0.137	-0.689	1.213	Up
C1orf15-NBL1	2163	2.913	0.572	0.196	-0.314	1.213	Up
TEK	19172	2.912	0.444	0.153	-0.456	1.213	Up
C1QTNF1	2121	2.91	0.31	0.107	-0.619	1.213	Up
HOXA9	8094	2.909	0.409	0.141	-0.35	1.213	Up
STX7	18751	2.906	0.25	0.086	-0.443	1.213	Up
C20orf194	2258	2.904	0.283	0.097	-0.22	1.213	Up
SLFN11	18026	2.904	0.447	0.154	-0.133	1.213	Up
LY96	11899	2.903	0.429	0.148	-0.338	1.213	Up
RCAN1	16363	2.903	0.619	0.213	-0.799	1.213	Up
CLIP4	3822	2.902	0.412	0.142	-0.219	1.213	Up
RASGRP2	16260	2.897	0.442	0.153	-1.221	1.213	Up
TSPAN18	20105	2.895	0.355	0.123	-0.378	1.213	Up
SPON2	18457	2.894	0.347	0.12	-0.399	1.213	Up
ZMAT3	21243	2.893	0.264	0.091	-0.771	1.213	Up
NCOA7	13147	2.891	0.586	0.203	-0.751	1.213	Up
ADAMTS8	286	2.886	0.544	0.189	-0.639	1.213	Up
JUNB	8811	2.886	0.378	0.131	-0.745	1.213	Up
PPIC	15449	2.885	0.471	0.163	-0.286	1.213	Up
NUDT9P1	13803	2.882	0.275	0.095	-0.592	1.213	Up
LMO2	9735	2.881	0.406	0.141	-0.685	1.213	Up
RGMA	16483	2.881	0.623	0.216	-0.676	1.213	Up
CLEC1A	3790	2.881	0.347	0.12	-0.394	1.213	Up
CD93	3295	2.876	0.39	0.136	-0.878	1.213	Up
MXRA8	12903	2.876	0.579	0.201	-0.253	1.213	Up
ACVR2A	234	2.875	0.275	0.096	-0.285	1.213	Up
C6orf174	2545	2.873	0.345	0.12	-0.064	1.46	Up
RPL21P44	16792	2.873	0.45	0.157	-1.189	1.46	Up
SACS	17058	2.873	0.302	0.105	-0.258	1.46	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
MXRA7	12902	2.872	0.317	0.11	-0.43	1.46	Up
PRUNE2	15808	2.871	0.273	0.095	-0.51	1.46	Up
CLCA2	3742	2.871	0.762	0.266	-2.456	1.46	Up
MOBKL2B	12566	2.87	0.308	0.107	-0.338	1.46	Up
CCDC82	3102	2.87	0.449	0.157	-0.336	1.46	Up
EMILIN1	5520	2.869	0.494	0.172	-0.13	1.46	Up
EPAS1	5584	2.867	0.483	0.168	-0.55	1.46	Up
ITIH5	8744	2.866	0.128	0.045	-0.271	1.46	Up
PLD1	15152	2.865	0.278	0.097	-0.44	1.46	Up
ATL1	1174	2.864	0.403	0.141	-1.098	1.46	Up
HTRA1	8270	2.864	0.739	0.258	-0.247	1.46	Up
CAPN5	2876	2.862	0.444	0.155	-0.264	1.46	Up
LGALS3	9624	2.86	0.373	0.131	-0.722	1.46	Up
CIDEA	3706	2.859	0.581	0.203	-0.554	1.46	Up
EVC2	5727	2.858	0.215	0.075	-0.093	1.46	Up
ANKRD37	700	2.856	0.435	0.152	-0.822	1.46	Up
LIPE	9702	2.854	0.734	0.257	-0.44	1.46	Up
FNDC1	6655	2.853	0.605	0.212	-0.601	1.46	Up
LOC100505933	10461	2.852	0.224	0.078	-0.667	1.46	Up
SYDE1	18841	2.851	0.199	0.07	-0.52	1.46	Up
LDB2	9574	2.849	0.296	0.104	-0.159	1.46	Up
B3GAT1	1333	2.849	0.52	0.183	-0.545	1.46	Up
PDK4	14771	2.848	0.775	0.272	-0.437	1.46	Up
PDGFRL	14756	2.847	0.727	0.255	-0.237	1.46	Up
PRRX2	15775	2.847	0.519	0.182	-0.476	1.46	Up
SPARC	18362	2.847	0.669	0.235	-0.116	1.46	Up
MECOM	12236	2.844	0.317	0.112	-0.518	1.46	Up
NFIB	13413	2.844	0.571	0.201	-0.542	1.46	Up
PPAP2A	15418	2.843	0.39	0.137	-0.233	1.46	Up
P2RY14	14429	2.842	0.377	0.133	-0.7	1.46	Up
HSPA12B	8218	2.84	0.428	0.151	-0.377	1.46	Up
LOC100130776	10099	2.838	0.259	0.091	-0.658	1.46	Up
TPT1	19876	2.836	0.343	0.121	-0.854	1.46	Up
ETV1	5718	2.836	0.254	0.089	-0.219	1.46	Up
SFRP1	17447	2.833	0.972	0.343	-0.628	1.46	Up
CYS1	4529	2.833	0.547	0.193	-0.215	1.46	Up
PLXDC1	15224	2.832	0.301	0.106	-0.131	1.46	Up
PER1	14824	2.832	0.44	0.155	-0.366	1.46	Up
POTEF	15389	2.832	0.344	0.121	-0.541	1.46	Up
AEBP1	364	2.832	0.708	0.25	-0.371	1.46	Up
ANXA8L2	766	2.831	1.185	0.419	-0.64	1.46	Up
KIAA1462	9087	2.831	0.419	0.148	-0.271	1.46	Up
FREM1	6737	2.828	0.27	0.095	-0.397	1.46	Up
LOX	11680	2.828	0.576	0.204	-0.197	1.46	Up
RAB7L1	16138	2.826	0.378	0.134	-0.261	1.46	Up
KIRREL	9182	2.825	0.368	0.13	-0.151	1.46	Up
C14orf37	1892	2.825	0.361	0.128	-0.253	1.46	Up
DCLK2	4629	2.822	0.273	0.097	-0.405	1.46	Up
KLK5	9267	2.82	1.034	0.367	-1.642	1.46	Up
LYL1	11903	2.819	0.208	0.074	-1.314	1.46	Up
GLIPR1	7153	2.818	0.292	0.104	-0.506	1.46	Up
ACTN3	216	2.817	0.294	0.104	-0.686	1.46	Up
ITM2A	8749	2.815	0.639	0.227	-0.715	1.46	Up
S1PR2	17047	2.814	0.355	0.126	-0.456	1.46	Up
LOC339524	11001	2.813	0.241	0.086	-0.321	1.46	Up
TGM5	19259	2.811	0.144	0.051	-1.566	1.46	Up
C7orf10	2582	2.805	0.343	0.122	-0.218	1.46	Up
PLAGL1	15126	2.803	0.282	0.101	-0.331	1.46	Up
RECK	16395	2.802	0.32	0.114	-0.211	1.46	Up
MMP7	12547	2.801	1.168	0.417	-0.459	1.46	Up
CACNA2D1	2793	2.8	0.223	0.08	0.055	1.46	Up
BACH1	1365	2.799	0.208	0.074	-0.58	1.46	Up
ZFPM2	21206	2.796	0.405	0.145	-0.515	1.46	Up
ADPRH	346	2.796	0.346	0.124	-0.205	1.46	Up
PLEKHH2	15184	2.796	0.282	0.101	-0.023	1.46	Up
LRRC32	11774	2.795	0.504	0.18	-0.409	1.46	Up
SYN2	18847	2.794	0.33	0.118	-0.224	1.46	Up
NOTCH4	13595	2.794	0.38	0.136	-0.242	1.46	Up
LOC728392	11532	2.792	0.426	0.152	-0.564	1.46	Up
SLC7A10	17982	2.789	0.462	0.166	-0.515	1.46	Up
CDC14B	3306	2.786	0.196	0.07	-0.233	1.46	Up
PVRL3	16018	2.784	0.268	0.096	0.076	1.46	Up
GYPC	7653	2.781	0.221	0.079	-0.119	1.46	Up
VEGFB	20656	2.781	0.291	0.104	-0.436	1.46	Up
ADIPOQ	332	2.779	0.617	0.222	-0.404	1.46	Up
PAX1	14570	2.777	0.452	0.163	-0.59	1.46	Up
MEGF6	12279	2.775	0.46	0.166	-1.673	1.46	Up
RAB23	16092	2.774	0.389	0.14	-0.347	1.46	Up
SLC26A3	17804	2.77	0.527	0.19	-3.689	1.46	Up
DARC	4573	2.764	0.409	0.148	-1.126	1.46	Up
F2R	5786	2.764	0.309	0.112	-0.256	1.46	Up
TMEM204	19560	2.764	0.385	0.139	-0.342	1.46	Up
CYYR1	4539	2.762	0.308	0.111	-0.66	1.46	Up
EMX2	5537	2.761	0.269	0.097	-0.413	1.46	Up
SLC16A7	17695	2.761	0.201	0.073	-0.054	1.46	Up
DUSP6	5238	2.759	0.538	0.195	-0.344	1.46	Up
LOC100130876	10110	2.759	0.536	0.194	-0.184	1.46	Up
SCN2B	17190	2.758	0.264	0.096	-0.529	1.46	Up
TBC1D2	19021	2.757	0.367	0.133	-1.301	1.46	Up
ZDHHC3	21158	2.757	0.25	0.091	-1.215	1.46	Up
AKR1B10	487	2.756	0.729	0.264	-2.563	1.46	Up
MT1L	12796	2.756	0.545	0.198	-0.173	1.46	Up
HYMAI	8289	2.756	0.414	0.15	-0.188	1.46	Up
ARHGEF40	965	2.756	0.396	0.144	-0.165	1.46	Up
KLF10	9196	2.755	0.395	0.143	-0.453	1.46	Up
CCL23	3142	2.755	0.209	0.076	-1.068	1.46	Up
F10	5779	2.754	0.144	0.052	-0.724	1.46	Up
TMEM100	19441	2.752	0.345	0.126	-0.807	2.095	Up
COL17A1	3969	2.748	0.697	0.254	-0.338	2.095	Up
MAMDC2	12007	2.748	0.335	0.122	-0.147	2.095	Up
PEAK1	14806	2.747	0.312	0.114	-0.224	2.095	Up
TWIST1	20277	2.747	0.362	0.132	-0.532	2.095	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
EFNB1	5377	2.747	0.385	0.14	-0.267	2.095	Up
PLIN1	15195	2.746	0.718	0.261	-0.624	2.095	Up
AOX1	772	2.746	0.38	0.138	-0.545	2.095	Up
FAM20A	6009	2.746	0.401	0.146	-0.09	2.095	Up
C11orf75	1772	2.745	0.486	0.177	-0.72	2.095	Up
ITGAV	8725	2.744	0.436	0.159	-0.651	2.095	Up
CUX2	4383	2.743	0.297	0.108	-1.712	2.095	Up
GRB10	7486	2.74	0.39	0.142	-0.666	2.095	Up
RPL21	16791	2.739	0.377	0.138	-1.274	2.095	Up
CCL28	3147	2.738	0.844	0.308	-0.813	2.095	Up
FLRT2	6623	2.738	0.249	0.091	-0.702	2.095	Up
BTBD11	1637	2.738	0.384	0.14	-1.627	2.095	Up
FAM164A	5928	2.736	0.331	0.121	-0.749	2.095	Up
NXN	13843	2.736	0.297	0.109	-0.282	2.095	Up
GGTA1P	7076	2.735	0.401	0.146	-0.478	2.095	Up
SLIT2	18035	2.735	0.367	0.134	-0.287	2.095	Up
PHYHIP	14965	2.734	0.238	0.087	-1.041	2.095	Up
FLJ36031	6513	2.732	0.396	0.145	-0.444	2.095	Up
THSD7A	19305	2.731	0.345	0.126	-0.653	2.095	Up
CTNNB1	4332	2.731	0.324	0.119	-0.806	2.095	Up
LAMB1	9491	2.731	0.203	0.074	-0.185	2.095	Up
ETS1	5716	2.73	0.251	0.092	-0.971	2.095	Up
VGLL3	20667	2.724	0.234	0.086	-0.082	2.095	Up
PPP1R13L	15487	2.724	0.363	0.133	-1.193	2.095	Up
ABCA8	44	2.723	0.349	0.128	-0.564	2.095	Up
AMOTL2	606	2.721	0.375	0.138	-0.55	2.095	Up
KIAA0922	9051	2.72	0.352	0.129	-0.801	2.095	Up
CARD6	2898	2.719	0.399	0.147	-0.289	2.095	Up
FAM20C	6011	2.719	0.44	0.162	-0.472	2.095	Up
MAFB	11955	2.719	0.365	0.134	-0.488	2.095	Up
LOXL2	11683	2.718	0.542	0.2	-0.113	2.095	Up
LPHN2	11700	2.718	0.26	0.096	-0.242	2.095	Up
TBC1D2B	19033	2.718	0.271	0.1	-0.377	2.095	Up
LONRF3	11678	2.716	0.385	0.142	-0.582	2.095	Up
ID1	8310	2.714	0.597	0.22	-0.468	2.095	Up
FAM65B	6080	2.714	0.402	0.148	-1.606	2.095	Up
DDX43	4702	2.714	0.473	0.174	-0.263	2.095	Up
ICAM5	8304	2.713	0.286	0.105	-0.797	2.095	Up
LEPR	9606	2.713	0.183	0.067	-0.542	2.095	Up
FLJ42709	6570	2.712	0.338	0.125	-0.46	2.095	Up
EGR3	5398	2.708	0.692	0.255	-0.339	2.095	Up
DDX26B	4691	2.708	0.312	0.115	-0.766	2.095	Up
TFAP2B	19215	2.707	1.339	0.495	-1.887	2.095	Up
CPEB1	4099	2.707	0.447	0.165	-0.677	2.095	Up
F13A1	5783	2.705	0.422	0.156	-0.594	2.095	Up
LOC100127983	9769	2.702	0.172	0.064	-0.199	2.095	Up
PRDM8	15603	2.697	0.235	0.087	-0.322	2.095	Up
BMP2	1544	2.697	0.279	0.103	-0.273	2.095	Up
CMAHP	3864	2.692	0.523	0.194	-1.055	2.095	Up
FZD4	6833	2.689	0.42	0.156	-0.554	2.095	Up
KRT6A	9348	2.688	0.909	0.338	-1.97	2.095	Up
MARCKS	12111	2.687	0.386	0.144	-0.152	2.095	Up
FMOD	6645	2.686	0.688	0.256	-0.335	2.095	Up
SELP	17298	2.686	0.389	0.145	-1.578	2.095	Up
C12orf75	1834	2.684	0.49	0.182	-0.47	2.095	Up
SPHK1	18418	2.684	0.378	0.141	-0.469	2.095	Up
LRRC70	11811	2.683	0.293	0.109	-0.742	2.095	Up
BNC2	1562	2.681	0.323	0.12	-0.183	2.095	Up
CNTN1	3932	2.679	0.231	0.086	-0.329	2.095	Up
GSTP1	7578	2.675	0.716	0.267	-0.439	2.095	Up
KLF8	9210	2.675	0.297	0.111	-0.513	2.095	Up
C13orf15	1837	2.674	0.399	0.149	-0.376	2.095	Up
VENTX	20658	2.674	0.226	0.084	-1.079	2.095	Up
TGFBR1	19248	2.671	0.333	0.124	-0.811	2.095	Up
BACE1	1363	2.67	0.261	0.098	-0.574	2.095	Up
HSD17B1	8180	2.669	0.683	0.256	-3.07	2.095	Up
VDR	20654	2.668	0.35	0.131	-0.667	2.095	Up
HOXA6	8092	2.667	0.522	0.196	-0.223	2.095	Up
LOC401097	11164	2.667	0.231	0.087	-0.596	2.095	Up
MMD	12522	2.664	0.396	0.149	-0.182	2.095	Up
LAMA2	9487	2.662	0.513	0.193	-0.526	2.095	Up
MT1A	12788	2.661	0.444	0.167	-0.243	2.095	Up
LAMB3	9494	2.661	0.775	0.291	-0.334	2.095	Up
ZNF521	21525	2.66	0.447	0.168	-0.454	2.095	Up
SIRPA	17610	2.659	0.394	0.148	-0.391	2.095	Up
DIRAS3	4887	2.659	0.345	0.13	-0.267	2.095	Up
GLI2	7150	2.659	0.156	0.059	-0.512	2.095	Up
LRP1	11735	2.658	0.25	0.094	-0.107	2.095	Up
CLIC2	3813	2.656	0.369	0.139	-0.449	2.095	Up
CPNE5	4113	2.655	0.459	0.173	-0.895	2.095	Up
DPT	5137	2.653	0.673	0.254	-0.049	2.095	Up
MT1E	12791	2.653	0.582	0.219	-0.099	2.095	Up
ITGA7	8718	2.653	0.486	0.183	-0.364	2.095	Up
NOTCH3	13594	2.653	0.371	0.14	-0.53	2.095	Up
TNFSF11	19742	2.652	0.249	0.094	-0.47	2.095	Up
PDE1C	14729	2.651	0.385	0.145	-0.897	2.095	Up
IL18R1	8476	2.65	0.295	0.111	-0.67	2.095	Up
ADAMTS6	284	2.65	0.235	0.089	-0.205	2.095	Up
GLG1	7148	2.646	0.272	0.103	-0.678	2.095	Up
ABCB1	46	2.646	0.241	0.091	-0.81	2.095	Up
VWF	20765	2.643	0.385	0.146	-0.53	2.095	Up
LGALS7	9627	2.643	0.801	0.303	-0.758	2.095	Up
CDH11	3348	2.643	0.379	0.143	-0.195	2.905	Up
PPL	15463	2.642	0.428	0.162	-1.473	2.905	Up
MYO1E	12975	2.641	0.35	0.132	-0.667	2.905	Up
LRMP	11734	2.64	0.413	0.157	-0.877	2.905	Up
GULP1	7638	2.639	0.355	0.135	-0.234	2.905	Up
TFPI	19235	2.638	0.348	0.132	-0.159	2.905	Up
EBF4	5296	2.637	0.497	0.189	-1.532	2.905	Up
THBS2	19280	2.636	0.628	0.238	-0.346	2.905	Up
FUT4	6807	2.635	0.297	0.113	-0.562	2.905	Up
NAMPT	13056	2.634	0.406	0.154	-0.318	2.905	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
PCDH18	14608	2.631	0.276	0.105	-0.048	2.905	Up
FYN	6827	2.63	0.407	0.155	-0.414	2.905	Up
PCDHB6	14643	2.63	0.227	0.086	-0.649	2.905	Up
BAMBI	1386	2.63	0.978	0.372	-2.666	2.905	Up
BEND6	1483	2.629	0.188	0.071	-0.299	2.905	Up
LIMA1	9676	2.626	0.406	0.154	-0.767	2.905	Up
ZYX	21847	2.623	0.366	0.139	-0.801	2.905	Up
RBMS3	16340	2.623	0.209	0.08	-0.037	2.905	Up
ACTG2	207	2.622	0.764	0.292	-0.557	2.905	Up
PRMT2	15681	2.621	0.231	0.088	-2.371	2.905	Up
SCN9A	17198	2.621	0.164	0.063	-1.048	2.905	Up
GALNTL2	6919	2.62	0.499	0.191	-0.451	2.905	Up
NCRNA00241	13204	2.62	0.395	0.151	-0.06	2.905	Up
PNMT	15276	2.616	1.135	0.434	-4.565	2.905	Up
FLI1	6430	2.616	0.327	0.125	-0.453	2.905	Up
TNFAIP6	19715	2.616	0.548	0.209	-0.317	2.905	Up
ICAM2	8301	2.614	0.443	0.17	-0.368	2.905	Up
NRXN2	13716	2.612	0.277	0.106	-0.385	2.905	Up
CASP5	2924	2.612	0.344	0.132	-0.844	2.905	Up
PLCL2	15147	2.611	0.217	0.083	-0.36	2.905	Up
ITM2B	8750	2.609	0.31	0.119	-0.449	2.905	Up
ACSL4	189	2.609	0.254	0.097	-0.008	2.905	Up
RASIP1	16263	2.609	0.451	0.173	-0.964	2.905	Up
ADM	336	2.608	0.631	0.242	-0.369	2.905	Up
FAM129C	5879	2.606	0.249	0.096	-2.039	2.905	Up
POSTN	15383	2.605	0.756	0.29	-0.254	2.905	Up
ADAMTS3	281	2.604	0.215	0.083	-0.737	2.905	Up
ZNF699	21663	2.603	0.165	0.063	-0.38	2.905	Up
ID2	8311	2.603	0.321	0.123	-0.903	2.905	Up
ANKH	647	2.603	0.396	0.152	-0.256	2.905	Up
EDN2	5335	2.602	0.812	0.312	-1.532	2.905	Up
CACNA1C	2785	2.602	0.244	0.094	-0.461	2.905	Up
LUM	11879	2.602	0.704	0.27	-0.303	2.905	Up
BANK1	1389	2.598	0.478	0.184	-1.618	2.905	Up
G0S2	6840	2.598	0.768	0.296	-0.264	2.905	Up
PABPC4L	14448	2.597	0.288	0.111	-0.387	2.905	Up
EPB41L2	5587	2.596	0.345	0.133	-0.225	2.905	Up
PPFIBP1	15442	2.595	0.287	0.111	-0.097	2.905	Up
RNF145	16655	2.593	0.282	0.109	-0.285	2.905	Up
MMP14	12532	2.59	0.525	0.203	-0.222	2.905	Up
COL1A2	3973	2.59	0.721	0.278	-0.343	2.905	Up
CSDA	4232	2.59	0.514	0.198	-0.322	2.905	Up
ROBO4	16734	2.59	0.327	0.126	-0.445	2.905	Up
C5orf13	2477	2.589	0.32	0.123	-0.384	2.905	Up
NIPAL4	13470	2.588	0.304	0.117	-0.565	2.905	Up
GXYLT2	7646	2.587	0.65	0.251	-0.357	2.905	Up
TJP1	19355	2.586	0.233	0.09	-0.884	2.905	Up
KLK4	9266	2.584	0.291	0.113	-0.552	2.905	Up
ZBTB16	21064	2.583	0.606	0.235	-2.378	2.905	Up
TTC28	20163	2.582	0.177	0.068	-0.768	2.905	Up
CD59	3274	2.581	0.434	0.168	-0.604	2.905	Up
KLHL2	9232	2.58	0.363	0.141	-0.676	2.905	Up
CPNE8	4116	2.58	0.252	0.098	-0.382	2.905	Up
CLEC3B	3797	2.578	0.449	0.174	-0.2	2.905	Up
L1TD1	9469	2.577	0.329	0.128	-1.108	2.905	Up
CIDEC	3708	2.576	0.461	0.179	-0.285	2.905	Up
ANKRD57	716	2.576	0.485	0.188	-2.563	2.905	Up
APOD	846	2.575	1.03	0.4	-2.165	2.905	Up
EFCAB1	5351	2.573	0.282	0.109	-0.606	2.905	Up
GRRP1	7541	2.572	0.412	0.16	-0.883	2.905	Up
PCOLCE	14687	2.572	0.463	0.18	-0.176	2.905	Up
TNFAIP8	19716	2.572	0.358	0.139	-0.395	2.905	Up
ECSCR	5317	2.569	0.338	0.132	-1.181	2.905	Up
ERG	5657	2.568	0.152	0.059	-0.432	2.905	Up
OXER1	14410	2.566	0.448	0.175	-1.04	2.905	Up
LIMS1	9683	2.565	0.277	0.108	-0.205	2.905	Up
CHRDL1	3652	2.564	0.54	0.211	-0.424	2.905	Up
TAC1	18908	2.562	0.206	0.08	-0.39	2.905	Up
PRRX1	15774	2.562	0.523	0.204	-0.16	2.905	Up
TMEM45A	19610	2.56	0.665	0.26	-0.327	2.905	Up
LHFPL2	9644	2.56	0.36	0.14	-0.272	2.905	Up
SOX17	18311	2.558	0.415	0.162	-0.22	2.905	Up
MYO1D	12974	2.558	0.266	0.104	-0.395	2.905	Up
CXCL1	4395	2.558	0.676	0.264	-0.974	2.905	Up
NDRG1	13287	2.557	0.626	0.245	-2.834	2.905	Up
EMCN	5513	2.555	0.434	0.17	-0.922	2.905	Up
IFI16	8335	2.555	0.409	0.16	-0.455	2.905	Up
C1QTNF7	2127	2.555	0.256	0.1	-0.385	2.905	Up
SCN7A	17196	2.552	0.296	0.116	-0.531	2.905	Up
IL4R	8527	2.55	0.321	0.126	-0.405	2.905	Up
ANO6	739	2.549	0.177	0.069	-0.375	2.905	Up
ARHGEF4	964	2.549	0.248	0.097	-0.215	2.905	Up
ENTPD1	5570	2.548	0.258	0.101	-0.515	2.905	Up
CXorf36	4429	2.547	0.146	0.057	-0.793	2.905	Up
LOC100129940	9999	2.546	0.274	0.108	-0.065	2.905	Up
STK10	18686	2.545	0.292	0.115	-1.146	2.905	Up
FMNL3	6637	2.543	0.162	0.064	-0.44	2.905	Up
C2orf40	2349	2.543	0.883	0.347	-0.471	2.905	Up
SLC9A9	18012	2.539	0.318	0.125	-0.349	2.905	Up
MT1B	12789	2.538	0.498	0.196	-0.143	2.905	Up
TMEM132C	19474	2.538	0.287	0.113	-0.297	2.905	Up
COL5A1	3991	2.537	0.399	0.157	-0.195	2.905	Up
FAM3D	6035	2.537	0.684	0.27	-1.271	2.905	Up
GALNT12	6907	2.536	0.281	0.111	-0.652	3.574	Up
DPP4	5126	2.534	0.439	0.173	-0.484	3.574	Up
ECM2	5316	2.534	0.273	0.108	-0.193	3.574	Up
GLIPR2	7156	2.534	0.455	0.179	-0.202	3.574	Up
DPEP2	5112	2.533	0.267	0.105	-1.284	3.574	Up
ZNF295	21384	2.532	0.184	0.073	-1.279	3.574	Up
IFFO1	8333	2.53	0.188	0.074	-0.127	3.574	Up
MAP4K5	12066	2.529	0.252	0.1	-0.375	3.574	Up
ADH1B	323	2.528	0.317	0.125	-0.377	3.574	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
LBH	9529	2.528	0.202	0.08	-0.631	3.574	Up
RSU1	16968	2.528	0.158	0.063	-0.884	3.574	Up
LRRC3B	11784	2.527	0.171	0.068	-0.727	3.574	Up
RASA3	16246	2.523	0.354	0.14	-0.089	3.574	Up
DIO2	4879	2.522	0.348	0.138	-0.687	3.574	Up
ZSWIM4	21834	2.521	0.43	0.171	-0.573	3.574	Up
WBP5	20783	2.521	0.452	0.179	-0.461	3.574	Up
PTGIR	15922	2.518	0.306	0.121	-0.375	3.574	Up
BEND4	1481	2.517	0.14	0.056	-1.949	3.574	Up
SSC5D	18577	2.517	0.562	0.223	-0.252	3.574	Up
DNASE1L3	5057	2.517	0.448	0.178	-0.95	3.574	Up
ARSI	1056	2.516	0.29	0.115	-0.387	3.574	Up
PLEKHG1	15175	2.515	0.308	0.122	-0.352	3.574	Up
TLR4	19380	2.514	0.269	0.107	-0.269	3.574	Up
TLE4	19368	2.514	0.212	0.084	-0.365	3.574	Up
EGFL6	5387	2.513	0.708	0.282	-0.359	3.574	Up
COPZ2	4039	2.511	0.48	0.191	-0.192	3.574	Up
CD99	3298	2.511	0.253	0.101	-0.17	3.574	Up
EFEMP1	5363	2.511	0.627	0.25	-0.25	3.574	Up
COL4A2	3985	2.509	0.374	0.149	-0.229	3.574	Up
ARPC2	1032	2.509	0.258	0.103	-0.475	3.574	Up
PAPLN	14512	2.508	0.238	0.095	-1.26	3.574	Up
EFNB3	5379	2.507	0.32	0.127	-0.723	3.574	Up
KRT16	9317	2.507	0.77	0.307	-1.462	3.574	Up
TUSC5	20273	2.507	0.375	0.15	-0.263	3.574	Up
RCN1	16374	2.506	0.403	0.161	-0.445	3.574	Up
BNC1	1561	2.504	0.283	0.113	-1.074	3.574	Up
AFAP1L1	370	2.504	0.294	0.117	-0.394	3.574	Up
ADAM12	247	2.503	0.292	0.117	-0.121	3.574	Up
LOC441869	11264	2.502	0.432	0.173	-0.996	3.574	Up
C6	2512	2.5	0.26	0.104	-0.711	3.574	Up
MAPRE2	12097	2.5	0.348	0.139	-0.625	3.574	Up
ILK	8542	2.497	0.214	0.086	-0.769	3.574	Up
RUNX2	17000	2.494	0.262	0.105	0.173	3.574	Up
KLF5	9207	2.494	0.604	0.242	-1.477	3.574	Up
MFGE8	12350	2.494	0.572	0.229	-0.365	3.574	Up
PTGER4	15916	2.494	0.406	0.163	-0.402	3.574	Up
HVCN1	8280	2.492	0.289	0.116	-0.964	3.574	Up
TNNT3	19771	2.491	0.27	0.108	-0.201	3.574	Up
HAAO	7682	2.491	0.199	0.08	-1.387	3.574	Up
FAM49A	6054	2.49	0.314	0.126	-0.375	3.574	Up
MYL2	12949	2.489	0.16	0.064	-0.765	3.574	Up
ALDH2	517	2.489	0.475	0.191	-0.68	3.574	Up
HOXA5	8091	2.489	0.428	0.172	-0.503	3.574	Up
CCL19	3137	2.488	0.862	0.347	-1.616	3.574	Up
MEIS2	12286	2.488	0.318	0.128	-0.013	3.574	Up
MMP3	12546	2.487	0.677	0.272	-0.564	3.574	Up
BLK	1529	2.487	0.534	0.215	-2.064	3.574	Up
S100A2	17031	2.487	0.794	0.319	-0.532	3.574	Up
PRICKLE2	15628	2.485	0.425	0.171	-0.428	3.574	Up
GRAMD2	7478	2.483	0.309	0.124	-1.166	3.574	Up
STON1	18717	2.48	0.377	0.152	-0.447	3.574	Up
CLDN14	3760	2.479	0.233	0.094	-1.07	3.574	Up
NCRNA00256B	13215	2.477	0.346	0.14	-0.837	3.574	Up
GPNMB	7352	2.476	0.287	0.116	-0.325	3.574	Up
TMEM132E	19476	2.476	0.22	0.089	0.099	3.574	Up
THY1	19311	2.473	0.474	0.192	-0.172	3.574	Up
PGF	14883	2.471	0.361	0.146	-0.628	3.574	Up
C7	2581	2.47	0.421	0.171	-1.197	3.574	Up
MEF2C	12272	2.47	0.336	0.136	-0.569	3.574	Up
CCDC88A	3110	2.469	0.295	0.12	-0.207	3.574	Up
ADAM23	256	2.469	0.137	0.055	-1.022	3.574	Up
SESN1	17410	2.468	0.314	0.127	-0.911	3.574	Up
ADAMTS9	287	2.468	0.4	0.162	-0.329	3.574	Up
RNF166	16664	2.466	0.204	0.083	-1.033	3.574	Up
ACSL1	187	2.465	0.439	0.178	-0.952	3.574	Up
PHLDA3	14948	2.464	0.354	0.144	-0.289	3.574	Up
PMCHL1	15239	2.464	0.142	0.058	-1.91	3.574	Up
BCL6B	1454	2.463	0.327	0.133	-1.139	3.574	Up
CNIH3	3895	2.462	0.395	0.161	-0.254	3.574	Up
SLC25A27	17774	2.462	0.404	0.164	-0.229	3.574	Up
RIPK4	16583	2.461	0.317	0.129	-2.015	3.574	Up
MCAM	12178	2.461	0.392	0.159	-0.446	3.574	Up
BATF3	1402	2.458	0.29	0.118	-0.407	3.574	Up
GABARAPL1	6854	2.457	0.307	0.125	-0.094	3.574	Up
ARGLU1	908	2.455	0.298	0.121	-0.533	3.574	Up
ELTD1	5511	2.454	0.302	0.123	-0.435	3.574	Up
THBS1	19279	2.453	0.36	0.147	-0.41	3.574	Up
PCDHB18	14638	2.452	0.194	0.079	-0.811	3.574	Up
KDELC2	8968	2.451	0.272	0.111	-0.105	3.574	Up
FCER1A	6276	2.451	0.531	0.217	-1.746	3.574	Up
LOC732272	11649	2.451	1.034	0.422	-0.655	3.574	Up
TMBIM1	19408	2.45	0.372	0.152	-0.776	3.574	Up
CDKL1	3404	2.449	0.225	0.092	-0.209	3.574	Up
MYO9A	12987	2.446	0.225	0.092	-0.707	3.574	Up
ACAA2	116	2.446	0.395	0.161	-0.454	3.574	Up
CHN1	3641	2.444	0.318	0.13	-1.009	3.574	Up
CDK15	3382	2.442	0.118	0.048	-0.948	3.574	Up
KLHL29	9241	2.441	0.307	0.126	-0.293	3.574	Up
LOC554202	11289	2.439	0.25	0.103	-0.683	3.574	Up
ZNF469	21487	2.438	0.548	0.225	-0.354	3.574	Up
GREM1	7492	2.437	0.428	0.176	-0.321	3.574	Up
KRT15	9316	2.435	1.015	0.417	-0.714	4.317	Up
PPAP2B	15419	2.435	0.39	0.16	-0.199	4.317	Up
C4orf49	2469	2.435	0.234	0.096	-1.108	4.317	Up
FBN1	6202	2.433	0.409	0.168	-0.063	4.317	Up
ALDH1A1	511	2.433	0.352	0.145	-0.646	4.317	Up
LOC100131541	10184	2.431	0.443	0.182	-0.651	4.317	Up
EYA2	5772	2.431	0.545	0.224	-0.384	4.317	Up
GMFG	7206	2.431	0.407	0.167	-0.598	4.317	Up
TMEM130	19470	2.431	0.496	0.204	-0.496	4.317	Up
CASP4	2923	2.429	0.338	0.139	-0.514	4.317	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
STX12	18739	2.429	0.21	0.087	-0.974	4.317	Up
PLTP	15221	2.429	0.43	0.177	-0.425	4.317	Up
BACE2	1364	2.428	0.552	0.227	-0.234	4.317	Up
RECQL	16396	2.428	0.137	0.057	-0.122	4.317	Up
GAS6	6948	2.427	0.409	0.169	-0.361	4.317	Up
GOLIM4	7293	2.423	0.313	0.129	-0.347	4.317	Up
PLXNC1	15233	2.421	0.289	0.119	-0.253	4.317	Up
NTF4	13751	2.419	0.423	0.175	-0.744	4.317	Up
PNLIPRP3	15267	2.419	0.523	0.216	-3.339	4.317	Up
CNTNAP3B	3941	2.419	0.198	0.082	-0.49	4.317	Up
MSN	12772	2.418	0.35	0.145	-0.177	4.317	Up
DSG3	5184	2.418	0.674	0.279	-1.828	4.317	Up
DFNA5	4804	2.418	0.232	0.096	-1.107	4.317	Up
ITGB3	8732	2.416	0.45	0.186	-0.23	4.317	Up
TLR10	19377	2.416	0.275	0.114	-1.395	4.317	Up
ALPK3	570	2.416	0.304	0.126	-0.445	4.317	Up
LOC100192378	10366	2.415	0.176	0.073	-1.171	4.317	Up
CTIF	4326	2.414	0.319	0.132	-0.308	4.317	Up
AQP7	875	2.414	0.393	0.163	-0.397	4.317	Up
GLS	7175	2.412	0.294	0.122	-0.439	4.317	Up
CAMK2N1	2848	2.412	0.614	0.255	-0.805	4.317	Up
CX3CL1	4392	2.411	0.469	0.194	-0.637	4.317	Up
RAB7B	16137	2.41	0.44	0.183	-0.262	4.317	Up
ITGA5	8716	2.41	0.344	0.143	-0.25	4.317	Up
UBE2E2	20353	2.409	0.294	0.122	-0.89	4.317	Up
NPR1	13641	2.408	0.263	0.109	-0.252	4.317	Up
BMP6	1549	2.408	0.382	0.159	-0.734	4.317	Up
MFAP5	12348	2.405	0.518	0.215	-0.326	4.317	Up
C3orf58	2424	2.405	0.283	0.118	-0.917	4.317	Up
MITF	12471	2.404	0.232	0.096	-0.001	4.317	Up
OLFML2B	13917	2.403	0.386	0.161	-0.32	4.317	Up
RNF150	16659	2.403	0.71	0.295	-0.741	4.317	Up
PTGFR	15920	2.4	0.223	0.093	-0.247	4.317	Up
KCNJ12	8893	2.399	0.368	0.153	-0.668	4.317	Up
SLC16A2	17690	2.398	0.511	0.213	-1.236	4.317	Up
IFNGR1	8372	2.398	0.283	0.118	-0.338	4.317	Up
NCRNA00312	13250	2.398	0.341	0.142	-0.281	4.317	Up
CIB2	3702	2.392	0.46	0.192	-0.483	4.317	Up
SGCD	17470	2.392	0.334	0.14	-0.08	4.317	Up
ANKS6	729	2.391	0.508	0.212	-0.3	4.317	Up
NAALADL1	13032	2.39	0.137	0.057	-0.358	4.317	Up
CD24	3234	2.39	0.98	0.41	-1.203	4.317	Up
COMMD6	4016	2.389	0.233	0.097	-1.742	4.317	Up
CD79B	3285	2.386	0.309	0.129	-2.816	4.317	Up
DMD	4969	2.385	0.21	0.088	-0.359	4.317	Up
ODZ2	13892	2.383	0.513	0.215	-0.149	4.317	Up
THBS4	19282	2.383	0.696	0.292	-0.283	4.317	Up
FABP5	5805	2.38	0.517	0.217	-0.246	4.317	Up
PURG	16006	2.379	0.11	0.046	-0.201	4.317	Up
PALM2	14492	2.372	0.193	0.081	-0.982	4.317	Up
ST3GAL2	18610	2.372	0.23	0.097	-0.205	4.317	Up
ANO1	733	2.37	0.423	0.178	-0.684	4.317	Up
WLS	20901	2.37	0.393	0.166	-0.633	4.317	Up
PMEPA1	15241	2.369	0.472	0.199	-0.217	4.317	Up
SOCS3	18279	2.367	0.425	0.18	-0.201	4.317	Up
LOC100131826	10220	2.365	0.224	0.095	-0.16	4.317	Up
MSL3	12766	2.365	0.126	0.053	-0.93	4.317	Up
PAMR1	14497	2.365	0.448	0.189	-0.469	4.317	Up
TNC	19710	2.365	0.624	0.264	-0.416	4.317	Up
AKR1C3	490	2.364	0.62	0.262	-1.668	4.317	Up
CD248	3237	2.364	0.333	0.141	-0.169	4.317	Up
GJB5	7123	2.363	0.302	0.128	-1.834	4.317	Up
SLCO2A1	18018	2.362	0.371	0.157	-1.18	4.317	Up
LPXN	11713	2.362	0.38	0.161	-0.476	4.317	Up
COL8A2	4002	2.361	0.495	0.21	-0.161	4.317	Up
IL1R1	8482	2.361	0.376	0.159	-0.851	4.317	Up
RHOQ	16543	2.36	0.214	0.091	-0.023	4.317	Up
ATP11B	1188	2.36	0.327	0.139	-0.39	4.317	Up
LOC100505806	10454	2.36	0.449	0.19	-0.503	4.317	Up
CAST	2934	2.36	0.253	0.107	-0.207	4.317	Up
ACVR1C	233	2.359	0.211	0.09	-0.324	4.317	Up
LOC100132891	10294	2.359	0.134	0.057	-0.19	4.317	Up
MOCS1	12571	2.359	0.25	0.106	-1.091	4.317	Up
TNAP	19709	2.359	0.179	0.076	-0.093	4.317	Up
PHLDB1	14949	2.358	0.252	0.107	0.08	4.317	Up
MAGED4B	11986	2.358	0.541	0.23	-0.472	4.317	Up
HSD17B11	8182	2.358	0.344	0.146	-0.108	4.317	Up
SPSB1	18490	2.358	0.345	0.146	-0.347	4.317	Up
TSPAN11	20098	2.357	0.194	0.082	-0.174	4.317	Up
ADAMTS18	277	2.356	0.493	0.209	-0.371	4.317	Up
KPNA3	9293	2.355	0.175	0.074	-0.609	4.317	Up
DTD1	5193	2.355	0.378	0.161	-0.989	4.317	Up
COL1A1	3972	2.352	0.658	0.28	-0.381	4.317	Up
MAF	11952	2.352	0.272	0.116	-0.454	4.317	Up
CALCRL	2824	2.351	0.258	0.11	-0.346	4.317	Up
SYT11	18875	2.349	0.237	0.101	-1.362	4.317	Up
COL5A2	3992	2.349	0.623	0.265	-0.179	4.317	Up
DHRS12	4841	2.349	0.22	0.094	-1.407	4.317	Up
NPY2R	13657	2.349	0.305	0.13	-1.188	4.317	Up
MT2A	12799	2.348	0.476	0.203	-0.291	4.317	Up
GABRE	6871	2.348	0.79	0.336	-0.622	4.317	Up
SERPINE1	17397	2.348	0.246	0.105	-0.336	4.317	Up
SYNM	18863	2.347	0.829	0.353	-0.841	4.317	Up
TGFB3	19246	2.347	0.491	0.209	-0.628	4.317	Up
CHL1	3628	2.346	0.256	0.109	-1.326	4.317	Up
MAFF	11956	2.345	0.325	0.138	-0.442	4.317	Up
STYK1	18760	2.345	0.425	0.181	-1.426	4.317	Up
CORIN	4049	2.345	0.306	0.13	-0.973	4.317	Up
SEMA3D	17307	2.343	0.516	0.22	-0.798	4.317	Up
CST6	4285	2.341	0.712	0.304	-1.086	4.317	Up
ZNF662	21633	2.34	0.468	0.2	-1.157	4.317	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
UHMK1	20465	-4.276	-0.414	0.097	-0.619	1.46	Down
CHADL	3589	-4.249	-0.765	0.18	-0.625	1.46	Down
RPUSD1	16914	-4.174	-0.439	0.105	-0.57	1.46	Down
PSMD4	15868	-4.166	-0.329	0.079	-0.699	1.46	Down
RPRD2	16856	-4.04	-0.315	0.078	-0.782	2.095	Down
KRTCAP3	9459	-3.953	-0.641	0.162	-0.588	2.095	Down
PYCR2	16036	-3.937	-0.325	0.083	-0.762	2.095	Down
C1orf182	2177	-3.936	-0.432	0.11	-0.97	2.095	Down
PSMB4	15847	-3.883	-0.364	0.094	-0.926	2.905	Down
TMPRSS6	19690	-3.842	-0.555	0.145	-0.809	2.905	Down
SNRNP35	18218	-3.747	-0.324	0.087	-0.635	3.574	Down
LOC284939	10890	-3.657	-0.704	0.192	-1.424	3.574	Down
NIT1	13476	-3.609	-0.349	0.097	-0.638	3.574	Down
LYSMD1	11921	-3.605	-0.419	0.116	-0.704	3.574	Down
TARS2	18972	-3.572	-0.262	0.073	-0.931	3.574	Down
C1orf97	2231	-3.554	-0.484	0.136	-0.968	3.574	Down
SCAMP3	17128	-3.55	-0.334	0.094	-0.517	3.574	Down
TNFRSF18	19731	-3.547	-0.638	0.18	-1.194	3.574	Down
ZNF646	21624	-3.543	-0.174	0.049	-0.768	3.574	Down

Add. Table 5: Significantly different expressed genes between the metabolic clusters Mc1 and Mc3

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
CALD1	2825	5.825	0.746	0.128	-0.375	0	Up
MRC2	12641	5.811	0.837	0.144	-0.49	0	Up
NNMT	13555	5.757	0.921	0.16	-0.427	0	Up
RUNX2	17000	5.748	0.491	0.085	-0.552	0	Up
COL18A1	3970	5.674	0.727	0.128	-0.421	0	Up
FERMT2	6321	5.668	0.53	0.094	-0.315	0	Up
DPYSL3	5150	5.577	0.874	0.157	-0.404	0	Up
TAGLN	18949	5.537	0.875	0.158	-0.374	0	Up
CTGF	4323	5.527	0.896	0.162	-0.38	0	Up
GFPT2	7054	5.507	0.71	0.129	-0.478	0	Up
TSHZ3	20086	5.502	0.674	0.122	-0.374	0	Up
GPX8	7474	5.474	0.772	0.141	-0.408	0	Up
SRPX2	18547	5.458	0.812	0.149	-0.639	0	Up
CLMP	3830	5.457	0.758	0.139	-0.239	0	Up
FAP	6172	5.376	0.924	0.172	-0.716	0	Up
VCAN	20645	5.365	0.895	0.167	-0.694	0	Up
LOXL2	11683	5.345	0.794	0.149	-0.629	0	Up
NEXN	13395	5.336	0.525	0.098	-0.306	0	Up
TIMP2	19344	5.294	0.79	0.149	-0.432	0	Up
LOX	11680	5.232	0.811	0.155	-0.687	0	Up
COL3A1	3983	5.23	1.031	0.197	-0.641	0	Up
PTRF	15990	5.185	0.425	0.082	-0.321	0	Up
EMILIN1	5520	5.184	0.619	0.119	-0.417	0	Up
ANGPTL2	633	5.165	0.763	0.148	-0.428	0	Up
SLIT3	18036	5.16	0.799	0.155	-0.426	0	Up
PDGFRB	14755	5.141	0.658	0.128	-0.416	0	Up
MSRB3	12776	5.141	0.746	0.145	-0.313	0	Up
SPARC	18362	5.14	0.892	0.174	-0.488	0	Up
NID1	13456	5.132	0.725	0.141	-0.41	0	Up
DACT3	4555	5.087	0.75	0.147	-0.362	0	Up
FBLN2	6199	5.087	0.897	0.176	-0.414	0	Up
NOX4	13603	5.067	0.536	0.106	-0.789	0	Up
CNN2	3901	5.057	0.753	0.149	-0.348	0	Up
TGFB1I1	19244	5.029	0.555	0.11	-0.354	0	Up
TCF4	19112	5.018	0.586	0.117	-0.306	0	Up
HTRA3	8272	5.008	0.679	0.136	-0.816	0	Up
MGC24103	12397	5.003	0.81	0.162	-0.448	0	Up
PDLIM7	14777	4.991	0.499	0.1	-0.504	0	Up
TPM4	19855	4.988	0.575	0.115	-0.525	0	Up
TMEM200A	19554	4.971	0.479	0.096	-0.446	0	Up
MXRA5	12901	4.966	0.524	0.106	-0.559	0	Up
GLT8D2	7182	4.958	0.452	0.091	-0.444	0	Up
SHOX2	17566	4.956	0.511	0.103	-0.397	0	Up
COL5A1	3991	4.95	0.587	0.119	-0.762	0	Up
GEM	7035	4.95	0.723	0.146	-0.142	0	Up
PALLD	14490	4.946	0.443	0.09	-0.536	0	Up
INHBA	8573	4.937	0.847	0.172	-1.272	0	Up
SNAI2	18133	4.92	0.534	0.108	-0.327	0	Up
FSTL1	6776	4.909	0.799	0.163	-0.468	0	Up
PODNL1	15297	4.908	0.534	0.109	-0.58	0	Up
COL1A2	3973	4.882	0.976	0.2	-0.818	0	Up
COL5A2	3992	4.881	0.964	0.198	-0.824	0	Up
ZEB1	21166	4.873	0.608	0.125	-0.345	0	Up
RNF144A	16653	4.871	0.51	0.105	-0.375	0	Up
PODN	15296	4.868	0.538	0.111	-0.36	0	Up
DSE	5180	4.845	0.503	0.104	-0.452	0	Up
SCG5	17165	4.836	0.623	0.129	-0.282	0	Up
THY1	19311	4.828	0.696	0.144	-0.72	0	Up
DZIP1L	5277	4.828	0.407	0.084	-0.456	0	Up
TNFAIP6	19715	4.826	0.746	0.155	-0.793	0	Up
AEBP1	364	4.811	0.881	0.183	-0.707	0	Up
POSTN	15383	4.805	1.026	0.214	-0.703	0	Up
PLK3	15203	4.791	0.446	0.093	-0.119	0	Up
BGN	1500	4.784	0.582	0.122	-0.684	0	Up
HTRA1	8270	4.782	0.888	0.186	-0.499	0	Up
MIR100HG	12463	4.778	0.7	0.147	-0.237	0	Up
DACT1	4553	4.766	0.713	0.15	-0.603	0	Up
CDH11	3348	4.748	0.51	0.107	-0.61	0	Up
TEAD1	19159	4.746	0.41	0.086	-0.534	0	Up
SERPINF1	17400	4.745	0.699	0.147	-0.276	0	Up
ACTA2	202	4.738	0.802	0.169	-0.269	0	Up
FBN1	6202	4.734	0.621	0.131	-0.616	0	Up
TRPC1	20038	4.732	0.503	0.106	-0.24	0	Up
IGFBP3	8405	4.731	0.639	0.135	-0.336	0	Up
ARHGAP28	926	4.724	0.466	0.099	-0.419	0	Up
ITPRIP	8761	4.722	0.519	0.11	-0.332	0	Up
LMOD1	9739	4.719	0.381	0.081	-0.178	0	Up
S100A10	17025	4.712	0.61	0.129	-0.295	0	Up
GPR124	7367	4.711	0.571	0.121	-0.24	0	Up
VIM	20673	4.703	0.551	0.117	-0.269	0	Up
COL1A1	3972	4.695	0.949	0.202	-0.993	0	Up
THBS2	19280	4.69	0.835	0.178	-0.792	0	Up
ADAMTS6	284	4.669	0.316	0.068	-0.62	0	Up
CYR61	4528	4.664	0.823	0.176	-0.17	0	Up
DKK3	4932	4.657	0.54	0.116	-0.355	0	Up
PRKCDBP	15652	4.638	0.589	0.127	-0.352	0	Up
PRRX1	15774	4.633	0.695	0.15	-0.54	0	Up
PLS3	15215	4.624	0.651	0.141	-0.262	0	Up
LAMA4	9489	4.616	0.445	0.096	-0.412	0	Up
MARVELD1	12120	4.615	0.596	0.129	-0.561	0	Up
MMP14	12532	4.604	0.717	0.156	-0.669	0	Up
LIMS2	9684	4.603	0.449	0.097	-0.151	0	Up
FIBIN	6389	4.602	0.657	0.143	-0.406	0	Up
MRVI1	12739	4.597	0.595	0.129	-0.299	0	Up
C5orf62	2509	4.597	0.621	0.135	-0.269	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
DCN	4634	4.579	0.871	0.19	-0.381	0	Up
ISM1	8696	4.579	0.774	0.169	-0.26	0	Up
PCSK5	14697	4.573	0.302	0.066	-0.329	0	Up
RARRES2	16240	4.558	0.8	0.175	-0.221	0	Up
KCNMB4	8924	4.557	0.697	0.153	-0.647	0	Up
SRPX	18546	4.545	0.679	0.149	-0.167	0	Up
ANTXR1	750	4.544	0.609	0.134	-0.671	0	Up
ISLR	8694	4.542	0.739	0.163	-0.417	0	Up
MMP19	12536	4.525	0.343	0.076	-0.107	0	Up
TPM2	19853	4.518	0.599	0.132	-0.426	0	Up
CYS1	4529	4.518	0.671	0.149	-0.491	0	Up
PDLIM4	14775	4.507	0.418	0.093	-0.227	0	Up
SERPING1	17402	4.493	0.418	0.093	-0.204	0	Up
LTBP2	11867	4.485	0.497	0.111	-0.275	0	Up
PPAPDC3	15424	4.477	0.342	0.076	-0.362	0	Up
MYL9	12956	4.471	0.644	0.144	-0.316	0	Up
PPIC	15449	4.471	0.556	0.124	-0.518	0	Up
COL6A3	3996	4.468	0.594	0.133	-0.751	0	Up
MXRA8	12903	4.466	0.687	0.154	-0.487	0	Up
GSN	7558	4.46	0.499	0.112	-0.013	0	Up
CFH	3562	4.452	0.546	0.123	-0.206	0	Up
TUBB6	20251	4.446	0.666	0.15	-0.356	0	Up
MMP2	12537	4.444	0.689	0.155	-0.459	0	Up
COL6A1	3994	4.435	0.75	0.169	-0.645	0	Up
PRICKLE1	15627	4.425	0.727	0.164	-0.413	0	Up
SPON1	18456	4.418	0.921	0.209	-0.457	0	Up
MEG3	12275	4.412	0.47	0.107	-0.295	0	Up
C1S	2133	4.403	0.636	0.144	-0.254	0	Up
FGF1	6341	4.401	0.242	0.055	-0.513	0	Up
MDFIC	12218	4.386	0.397	0.091	-0.154	0	Up
RBMS1	16338	4.385	0.488	0.111	-0.228	0	Up
ADAM12	247	4.384	0.412	0.094	-0.577	0	Up
SLC2A3	17831	4.372	0.255	0.058	-0.27	0	Up
FN1	6649	4.369	0.89	0.204	-0.954	0	Up
GAS1	6942	4.364	0.685	0.157	-0.356	0	Up
MGC4294	12410	4.359	0.577	0.132	-0.835	0	Up
sep.11	17342	4.357	0.499	0.115	-0.403	0	Up
SGCD	17470	4.353	0.469	0.108	-0.515	0	Up
TNS1	19782	4.352	0.393	0.09	-0.033	0	Up
LIMS1	9683	4.343	0.34	0.078	-0.481	0	Up
LUM	11879	4.342	0.883	0.203	-0.635	0	Up
EVC2	5727	4.33	0.253	0.058	-0.285	0	Up
LCAT	9536	4.32	0.379	0.088	-0.218	0	Up
FEZ1	6326	4.316	0.374	0.087	-0.36	0	Up
TBX15	19068	4.314	0.291	0.067	-0.288	0	Up
EFHA2	5366	4.305	0.391	0.091	-0.065	0	Up
C1R	2131	4.303	0.612	0.142	-0.245	0	Up
SFRP2	17448	4.301	0.904	0.21	-0.466	0	Up
C1orf15-NBL1	2163	4.297	0.65	0.151	-0.493	0	Up
CLIP3	3821	4.29	0.375	0.087	-0.264	0	Up
ANXA1	754	4.289	0.63	0.147	-0.069	0	Up
NID2	13457	4.288	0.362	0.084	-0.653	0	Up
KIAA1462	9087	4.288	0.478	0.111	-0.45	0	Up
COPZ2	4039	4.282	0.62	0.145	-0.538	0	Up
PPAPDC1A	15421	4.281	0.857	0.2	-1.053	0	Up
GADD45B	6883	4.28	0.522	0.122	-0.283	0	Up
AGPAT4	404	4.279	0.36	0.084	-0.276	0	Up
PMEPA1	15241	4.273	0.588	0.138	-0.517	0	Up
RHOJ	16542	4.259	0.457	0.107	-0.097	0	Up
SERPINH1	17403	4.256	0.652	0.153	-0.704	0	Up
COL6A2	3995	4.254	0.657	0.154	-0.563	0	Up
RFX8	16471	4.249	0.375	0.088	-0.614	0	Up
PLXDC1	15224	4.249	0.38	0.089	-0.428	0	Up
ARSI	1056	4.231	0.375	0.089	-0.796	0	Up
VGLL3	20667	4.225	0.287	0.068	-0.326	0	Up
PCOLCE	14687	4.224	0.572	0.135	-0.452	0	Up
PDGFRA	14754	4.219	0.699	0.166	-0.235	0	Up
COL12A1	3964	4.218	0.813	0.193	-0.757	0	Up
DNAJB5	5018	4.216	0.359	0.085	-0.387	0	Up
MFAP5	12348	4.213	0.73	0.173	-0.869	0	Up
IKBIP	8438	4.213	0.288	0.068	-0.514	0	Up
NDN	13282	4.208	0.558	0.133	-0.21	0	Up
BNC2	1562	4.195	0.404	0.096	-0.481	0	Up
HIF1A	7874	4.195	0.465	0.111	-0.277	0	Up
PRRX2	15775	4.191	0.569	0.136	-0.619	0	Up
CYTH3	4535	4.186	0.522	0.125	-0.168	0	Up
PDGFRL	14756	4.185	0.815	0.195	-0.387	0	Up
RAB23	16092	4.181	0.491	0.117	-0.701	0	Up
PDPN	14781	4.18	0.594	0.142	-0.526	0	Up
C14orf37	1892	4.179	0.429	0.103	-0.489	0	Up
LRP1	11735	4.179	0.288	0.069	-0.275	0	Up
SEMA5A	17317	4.175	0.61	0.146	-0.297	0	Up
COL8A2	4002	4.17	0.695	0.167	-0.629	0	Up
SKAP2	17634	4.17	0.455	0.109	-0.089	0	Up
MAP4K4	12065	4.154	0.325	0.078	-0.398	0	Up
PXDN	16027	4.153	0.668	0.161	-0.709	0	Up
DAB2	4549	4.147	0.446	0.107	-0.179	0	Up
ASPN	1122	4.137	1.058	0.256	-0.92	0	Up
CCDC80	3100	4.136	0.607	0.147	-0.242	0	Up
SLC2A14	17829	4.136	0.454	0.11	-0.27	0	Up
MARCKS	12111	4.13	0.466	0.113	-0.391	0	Up
EHD2	5403	4.129	0.46	0.111	-0.257	0	Up
DLC1	4936	4.128	0.306	0.074	-0.148	0	Up
PVRL3	16018	4.127	0.353	0.086	-0.215	0	Up
NAP1L3	13066	4.125	0.339	0.082	-0.233	0	Up
GRASP	7484	4.122	0.466	0.113	-0.121	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
CLEC11A	3782	4.119	0.601	0.146	-0.58	0	Up
PHLDB1	14949	4.111	0.34	0.083	-0.244	0	Up
C7orf10	2582	4.107	0.442	0.108	-0.569	0	Up
EGR1	5396	4.105	0.835	0.203	-0.031	0	Up
SH3PXD2A	17528	4.102	0.464	0.113	-0.328	0	Up
PPFIBP1	15442	4.102	0.375	0.091	-0.433	0	Up
PLXDC2	15225	4.093	0.521	0.127	-0.621	0	Up
LHFP	9642	4.092	0.579	0.141	-0.107	0	Up
ETV1	5718	4.089	0.281	0.069	-0.347	0	Up
GXYLT2	7646	4.059	0.775	0.191	-0.617	0	Up
SLFN11	18026	4.058	0.467	0.115	-0.182	0	Up
MYLK	12958	4.056	0.678	0.167	-0.165	0	Up
KIRREL	9182	4.056	0.399	0.098	-0.247	0	Up
OLFML2B	13917	4.055	0.481	0.119	-0.644	0	Up
CTSK	4353	4.045	0.735	0.182	-0.489	0	Up
TSPO2	20120	4.041	0.335	0.083	-0.614	0	Up
ITGA11	8711	4.04	0.731	0.181	-0.955	0	Up
PRDM1	15591	4.037	0.487	0.121	-0.484	0	Up
FNDC1	6655	4.036	0.67	0.166	-0.771	0	Up
C1QTNF5	2125	4.034	0.54	0.134	-0.6	0	Up
COL10A1	3961	4.026	0.974	0.242	-1.47	0	Up
CACNA2D1	2793	4.025	0.292	0.073	-0.238	0	Up
ADAMTS14	273	4.024	0.712	0.177	-0.714	0	Up
SGIP1	17474	4.018	0.417	0.104	-0.632	0	Up
C9orf125	2684	4.008	0.412	0.103	-0.154	0	Up
UACA	20323	4.006	0.384	0.096	-0.204	0	Up
ECM2	5316	4.004	0.314	0.078	-0.373	0	Up
FBLN5	6200	4.002	0.612	0.153	-0.139	0	Up
HMCN1	8005	3.995	0.612	0.153	-0.555	0	Up
TMEM45A	19610	3.984	0.736	0.185	-0.468	0	Up
GREM1	7492	3.979	0.559	0.14	-0.723	0	Up
FOXF2	6694	3.97	0.285	0.072	-0.364	0	Up
SPON2	18457	3.97	0.362	0.091	-0.457	0	Up
C11orf96	1787	3.965	0.55	0.139	-0.213	0	Up
NEFH	13351	3.959	0.597	0.151	-0.319	0	Up
BEND6	1483	3.956	0.226	0.057	-0.565	0	Up
RECK	16395	3.952	0.364	0.092	-0.377	0	Up
SSC5D	18577	3.947	0.666	0.169	-0.485	0	Up
AKAP12	465	3.946	0.47	0.119	-0.068	0	Up
LEPRE1	9607	3.944	0.275	0.07	-0.666	0	Up
ADAMTS5	283	3.942	0.605	0.154	-0.139	0	Up
COL16A1	3968	3.942	0.415	0.105	-0.314	0	Up
KDELC2	8968	3.938	0.352	0.089	-0.429	0	Up
PRKD1	15660	3.935	0.355	0.09	-0.188	0	Up
ARL4C	996	3.927	0.518	0.132	-0.202	0	Up
ACTN1	214	3.925	0.47	0.12	-0.489	0	Up
DDR2	4674	3.919	0.415	0.106	-0.326	0	Up
HOXA11-AS1	8086	3.907	0.541	0.138	-0.488	0	Up
WNT2	20911	3.902	0.431	0.11	-0.645	0	Up
PTPRD	15972	3.902	0.592	0.152	-0.519	0	Up
PTGIR	15922	3.896	0.355	0.091	-0.595	0	Up
CFHR3	3565	3.89	0.454	0.117	-0.175	0	Up
HABP4	7685	3.888	0.368	0.095	-0.106	0	Up
TMEM2	19552	3.888	0.321	0.083	-0.568	0	Up
LAMB1	9491	3.888	0.223	0.057	-0.301	0	Up
NCRNA00241	13204	3.887	0.436	0.112	-0.17	0	Up
SSPN	18583	3.879	0.525	0.135	-0.232	0	Up
GRP	7537	3.876	0.94	0.243	-0.641	0	Up
PROS1	15712	3.872	0.362	0.093	-0.096	0	Up
ITGA1	8709	3.87	0.37	0.096	-0.444	0	Up
CD99	3298	3.869	0.303	0.078	-0.399	0	Up
CES1	3545	3.867	0.453	0.117	-0.101	0	Up
CCDC8	3099	3.866	0.538	0.139	-0.176	0	Up
MT1M	12797	3.865	0.768	0.199	-0.071	0	Up
PRDM6	15602	3.851	0.349	0.091	-0.738	0	Up
BHLHE41	1507	3.851	0.649	0.169	-0.097	0	Up
SPOCK1	18453	3.847	0.695	0.181	-0.802	0	Up
EGR2	5397	3.845	0.603	0.157	0.011	0	Up
CMTM3	3874	3.837	0.467	0.122	-0.393	0	Up
HEYL	7849	3.835	0.494	0.129	-0.727	0	Up
CNN1	3900	3.831	0.815	0.213	-0.149	0	Up
TWIST2	20278	3.827	0.603	0.157	-0.184	0	Up
LOC100129940	9999	3.82	0.388	0.101	-0.506	0	Up
LSP1	11857	3.807	0.368	0.097	-0.093	0	Up
ACSL4	189	3.806	0.302	0.079	-0.199	0	Up
COL5A3	3993	3.806	0.526	0.138	-0.895	0	Up
SH3PXD2B	17529	3.804	0.46	0.121	-0.506	0	Up
SALL4	17067	3.8	0.664	0.175	-0.938	0	Up
SOBP	18276	3.799	0.28	0.074	-0.158	0	Up
ITGA5	8716	3.799	0.43	0.113	-0.56	0	Up
CTSB	4346	3.795	0.437	0.115	-0.662	0	Up
WISP1	20897	3.793	0.314	0.083	-0.694	0	Up
RBMS3	16340	3.791	0.238	0.063	-0.179	0	Up
FLNA	6617	3.79	0.315	0.083	-0.269	0	Up
ZNF469	21487	3.788	0.66	0.174	-0.632	0	Up
KLF7	9209	3.786	0.388	0.102	-0.458	0	Up
LRRC32	11774	3.785	0.482	0.127	-0.348	0	Up
C9orf53	2729	3.785	0.669	0.177	-0.425	0	Up
LRCH2	11717	3.785	0.241	0.064	-0.257	0	Up
COMP	4020	3.773	0.878	0.233	-0.709	0	Up
FST	6775	3.766	0.776	0.206	-0.506	0	Up
WISP2	20898	3.766	0.982	0.261	-0.559	0	Up
ZFP36	21191	3.763	0.531	0.141	-0.085	0	Up
SVEP1	18824	3.762	0.326	0.087	-0.06	0	Up
JDP2	8787	3.761	0.431	0.115	-0.44	0	Up
FOXQ1	6719	3.76	0.765	0.204	-0.499	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
SCARNA17	17147	3.758	0.54	0.144	-0.128	0	Up
RAB7L1	16138	3.758	0.441	0.117	-0.471	0	Up
RASA3	16246	3.757	0.385	0.103	-0.185	0	Up
NTM	13753	3.755	0.329	0.088	-0.821	0	Up
LMCD1	9727	3.747	0.499	0.133	-0.456	0	Up
GLIPR2	7156	3.742	0.498	0.133	-0.316	0	Up
ZFHX4	21181	3.737	0.498	0.133	-0.258	0	Up
ST6GAL2	18617	3.736	0.477	0.128	-0.775	0	Up
LOC100132891	10294	3.725	0.161	0.043	-0.43	0	Up
DPT	5137	3.725	0.747	0.201	-0.164	0	Up
APCDD1	809	3.722	0.588	0.158	-0.183	0	Up
JAG1	8774	3.719	0.407	0.109	-0.404	0	Up
SULF1	18774	3.718	0.668	0.18	-0.945	0	Up
EIF5A2	5468	3.717	0.283	0.076	-0.3	0	Up
C12orf70	1830	3.708	0.415	0.112	-0.575	0	Up
LOXL1	11682	3.708	0.594	0.16	-0.63	0	Up
NXN	13843	3.707	0.344	0.093	-0.486	0	Up
ATF3	1147	3.692	0.471	0.128	-0.196	0	Up
XG	20957	3.69	0.356	0.097	-0.491	0	Up
NHS	13452	3.688	0.252	0.068	-0.254	0	Up
VEGFC	20657	3.687	0.383	0.104	-0.294	0	Up
ZEB2	21168	3.677	0.342	0.093	-0.09	0	Up
ITGBL1	8739	3.677	0.672	0.183	-0.522	0	Up
COL8A1	4001	3.674	0.742	0.202	-0.959	0	Up
ARHGEF40	965	3.673	0.408	0.111	-0.201	0	Up
C21orf34	2296	3.67	0.272	0.074	-0.062	0	Up
MT1X	12798	3.669	0.58	0.158	-0.067	0	Up
MYO1B	12972	3.664	0.312	0.085	-0.212	0	Up
EFEMP2	5364	3.659	0.422	0.115	-0.488	0	Up
PDLIM3	14774	3.658	0.467	0.128	-0.167	0	Up
ITGB3	8732	3.657	0.526	0.144	-0.437	0	Up
SLC12A4	17666	3.656	0.292	0.08	-0.193	0	Up
ACVR2A	234	3.655	0.264	0.072	-0.232	0	Up
LOC100130876	10110	3.647	0.578	0.159	-0.277	0	Up
MAMDC2	12007	3.647	0.342	0.094	-0.173	0	Up
FKBP9	6419	3.642	0.188	0.052	-0.455	0	Up
LPCAT2	11695	3.639	0.271	0.074	-0.142	0	Up
C5orf13	2477	3.639	0.366	0.101	-0.584	0	Up
LRIG3	11730	3.638	0.482	0.132	-0.113	0	Up
CD200	3228	3.635	0.334	0.092	-0.106	0	Up
SPSB1	18490	3.624	0.373	0.103	-0.46	0	Up
MXRA7	12902	3.62	0.314	0.087	-0.414	0	Up
PTPN21	15959	3.618	0.463	0.128	-0.172	0	Up
C17orf51	2000	3.612	0.428	0.118	-0.237	0	Up
KLF12	9198	3.611	0.34	0.094	-0.287	0	Up
CELF2	3478	3.608	0.346	0.096	-0.033	0	Up
FAM7A1	6115	3.607	0.261	0.072	-0.473	0	Up
LAYN	9528	3.6	0.399	0.111	-0.218	0	Up
PTGFRN	15921	3.6	0.398	0.11	-0.598	0	Up
SYNDIG1	18851	3.599	0.746	0.207	-0.833	0	Up
SCHIP1	17177	3.584	0.508	0.142	-0.121	0	Up
RECQL	16396	3.58	0.171	0.048	-0.398	0	Up
GGT5	7072	3.577	0.456	0.128	-0.134	0	Up
FGF7	6360	3.577	0.324	0.09	-0.085	0	Up
FOXO1	6712	3.57	0.332	0.093	-0.006	0	Up
FAM20A	6009	3.57	0.401	0.112	-0.089	0	Up
WDR86	20856	3.567	0.707	0.198	-0.233	0	Up
ZFPM2	21206	3.565	0.375	0.105	-0.402	0	Up
LOC541471	11281	3.564	0.452	0.127	-0.771	0	Up
MT1E	12791	3.56	0.584	0.164	-0.102	0	Up
KLF2	9204	3.559	0.471	0.132	-0.144	0	Up
LAMA1	9486	3.557	0.307	0.086	-0.23	0	Up
MAP4K5	12066	3.556	0.249	0.07	-0.356	0	Up
CERCAM	3536	3.554	0.347	0.098	-0.772	0	Up
DZIP1	5276	3.552	0.372	0.105	-0.29	0	Up
GLIS1	7157	3.55	0.429	0.121	-0.496	0	Up
JAM3	8784	3.55	0.322	0.091	-0.245	0	Up
SLC26A10	17801	3.547	0.365	0.103	-0.463	0	Up
ROR1	16743	3.545	0.305	0.086	-0.254	0	Up
CTHRC1	4325	3.542	0.667	0.188	-0.924	0	Up
PKD2	15071	3.542	0.354	0.1	-0.136	0	Up
SLC36A1	17880	3.536	0.359	0.102	-0.529	0	Up
GLIS2	7158	3.535	0.495	0.14	-0.534	0	Up
MITF	12471	3.532	0.303	0.086	-0.308	0	Up
NPR2	13642	3.526	0.388	0.11	-0.256	0	Up
GJB2	7120	3.526	0.836	0.237	-0.948	0	Up
ZFAND5	21174	3.522	0.263	0.075	-0.152	0	Up
GPC6	7333	3.519	0.435	0.124	-0.587	0	Up
CNRIP1	3925	3.519	0.227	0.065	-0.145	0	Up
CNIH3	3895	3.516	0.441	0.125	-0.399	0	Up
SGMS2	17483	3.514	0.326	0.093	-0.473	0	Up
STARD8	18661	3.514	0.273	0.078	-0.179	0	Up
P2RY1	14425	3.514	0.252	0.072	-0.161	0	Up
RUNX1T1	16999	3.514	0.509	0.145	-0.095	0	Up
JAZF1	8786	3.512	0.312	0.089	-0.235	0	Up
HSPB2	8233	3.507	0.564	0.161	-0.139	0	Up
NKD2	13486	3.506	0.386	0.11	-0.487	0	Up
COL15A1	3967	3.505	0.425	0.121	-0.1	0	Up
CD248	3237	3.504	0.371	0.106	-0.3	0	Up
HEG1	7809	3.502	0.404	0.115	-0.202	0	Up
MYO1D	12974	3.502	0.294	0.084	-0.544	0	Up
F2R	5786	3.499	0.312	0.089	-0.265	0	Up
PELI2	14817	3.497	0.425	0.121	-0.189	0	Up
ALDH1L2	516	3.493	0.421	0.121	-0.53	0	Up
LIMS3L	9686	3.489	0.351	0.101	-0.249	0	Up
HHIPL1	7863	3.475	0.35	0.101	-0.605	0	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
ZC3H12B	21103	3.475	0.196	0.056	-0.082	0	Up
ABCA6	42	3.461	0.479	0.138	0.019	0	Up
MT1L	12796	3.456	0.514	0.149	-0.106	0	Up
ARAP3	887	3.456	0.372	0.108	-0.094	0	Up
ADAMTS12	271	3.451	0.256	0.074	-0.91	0	Up
PRKCA	15649	3.451	0.446	0.129	-0.005	0	Up
FAS	6180	3.442	0.29	0.084	-0.03	0	Up
BMP1	1541	3.441	0.196	0.057	-0.798	0	Up
GOLGA7B	7287	3.44	0.367	0.107	-0.558	0	Up
CLIP4	3822	3.44	0.373	0.108	-0.104	0	Up
LGALS1	9618	3.427	0.423	0.124	-0.513	0	Up
C20orf103	2234	3.425	0.707	0.206	-0.519	0	Up
MEIS2	12286	3.417	0.364	0.107	-0.16	0	Up
FAM43A	6040	3.416	0.373	0.109	-0.16	0	Up
KANK2	8818	3.413	0.34	0.1	-0.221	0	Up
AKAP13	466	3.412	0.256	0.075	-0.198	0	Up
SLAMF9	17652	3.406	0.351	0.103	-0.873	0	Up
HSPG2	8245	3.402	0.263	0.077	-0.429	0	Up
TMEM43	19608	3.399	0.278	0.082	-0.315	0.17	Up
EGFL6	5387	3.396	0.675	0.199	-0.297	0.17	Up
C10orf10	1681	3.395	0.576	0.17	-0.082	0.17	Up
MAGEL2	11991	3.393	0.363	0.107	-0.281	0.17	Up
FAT4	6191	3.389	0.189	0.056	-0.115	0.17	Up
HIC1	7872	3.386	0.349	0.103	-0.244	0.17	Up
C20orf194	2258	3.385	0.301	0.089	-0.301	0.17	Up
ID3	8312	3.385	0.339	0.1	-0.259	0.17	Up
CXCL12	4398	3.384	0.38	0.112	-0.127	0.17	Up
SPATS2L	18389	3.382	0.404	0.119	-0.273	0.17	Up
MMP11	12529	3.381	0.816	0.241	-1.676	0.17	Up
TPST1	19874	3.378	0.385	0.114	-0.51	0.17	Up
MAP1B	12033	3.372	0.341	0.101	0.003	0.17	Up
LOC100131826	10220	3.372	0.271	0.08	-0.404	0.17	Up
ANTXR2	751	3.37	0.236	0.07	-0.047	0.17	Up
KLK4	9266	3.366	0.309	0.092	-0.648	0.17	Up
PEAK1	14806	3.359	0.28	0.083	-0.097	0.17	Up
ARHGAP20	918	3.351	0.278	0.083	-0.083	0.17	Up
DOCK11	5084	3.347	0.355	0.106	0	0.17	Up
SPECC1	18404	3.341	0.261	0.078	-0.242	0.17	Up
SPHK1	18418	3.34	0.364	0.109	-0.416	0.17	Up
NUAK1	13767	3.339	0.303	0.091	-0.658	0.17	Up
TMEM204	19560	3.333	0.395	0.119	-0.377	0.17	Up
COL7A1	4000	3.329	0.535	0.161	-0.523	0.17	Up
FZD7	6836	3.325	0.501	0.151	-0.193	0.17	Up
HOMER1	8072	3.321	0.243	0.073	-0.501	0.17	Up
DST	5189	3.321	0.452	0.136	-0.007	0.17	Up
ADAM19	251	3.32	0.272	0.082	-0.638	0.17	Up
GNG11	7241	3.318	0.451	0.136	-0.01	0.17	Up
SFRP4	17449	3.313	0.58	0.175	-0.294	0.17	Up
BACH2	1366	3.311	0.19	0.057	-0.063	0.17	Up
LOC286068	10952	3.309	0.159	0.048	-0.521	0.17	Up
TIMP3	19345	3.309	0.58	0.175	-0.666	0.17	Up
EFNB2	5378	3.306	0.412	0.125	-0.411	0.17	Up
ITGB1	8727	3.304	0.303	0.092	-0.264	0.17	Up
FAM101B	5825	3.303	0.418	0.127	-0.504	0.17	Up
P4HA3	14436	3.303	0.27	0.082	-0.928	0.17	Up
ARPC2	1032	3.297	0.25	0.076	-0.432	0.17	Up
SGK1	17475	3.297	0.422	0.128	-0.006	0.17	Up
ZNF300P1	21390	3.292	0.443	0.135	-0.131	0.17	Up
KCTD12	8949	3.291	0.376	0.114	-0.026	0.17	Up
FOSB	6676	3.291	0.753	0.229	0.095	0.17	Up
MMP13	12531	3.29	0.744	0.226	-0.813	0.17	Up
SLIT2	18035	3.289	0.365	0.111	-0.279	0.17	Up
EDNRA	5337	3.286	0.31	0.094	-0.592	0.17	Up
CDR1	3423	3.281	0.227	0.069	-0.256	0.17	Up
LZTS1	11934	3.278	0.221	0.068	-0.475	0.17	Up
CDH13	3350	3.277	0.383	0.117	-0.486	0.17	Up
FAM69A	6085	3.272	0.258	0.079	-0.247	0.17	Up
IGFBP7	8409	3.272	0.427	0.131	-0.294	0.17	Up
BMPR2	1556	3.271	0.205	0.063	-0.378	0.17	Up
PLXNC1	15233	3.271	0.36	0.11	-0.559	0.17	Up
PRNP	15688	3.27	0.498	0.152	-0.131	0.17	Up
SEC23A	17270	3.27	0.342	0.105	-0.571	0.17	Up
C12orf75	1834	3.266	0.458	0.14	-0.374	0.17	Up
STGC3	18680	3.263	0.231	0.071	-0.374	0.17	Up
CRISPLD2	4183	3.262	0.346	0.106	-0.401	0.17	Up
PLEKHG2	15176	3.257	0.258	0.079	-0.522	0.17	Up
ITGAV	8725	3.256	0.375	0.115	-0.422	0.17	Up
COL4A2	3985	3.25	0.407	0.125	-0.336	0.17	Up
TSPAN18	20105	3.249	0.341	0.105	-0.325	0.17	Up
LDB2	9574	3.248	0.271	0.083	-0.061	0.17	Up
AHNAK2	431	3.247	0.455	0.14	-0.832	0.17	Up
DUSP1	5217	3.246	0.5	0.154	-0.024	0.17	Up
CPXM1	4137	3.245	0.414	0.128	-0.172	0.17	Up
PLD1	15152	3.243	0.256	0.079	-0.326	0.17	Up
DDX43	4702	3.242	0.481	0.148	-0.283	0.17	Up
C6orf174	2545	3.241	0.319	0.098	0.016	0.17	Up
EMX2OS	5538	3.238	0.414	0.128	-0.208	0.17	Up
KLKP1	9273	3.236	0.356	0.11	-0.752	0.17	Up
EDIL3	5333	3.23	0.618	0.191	-1.289	0.17	Up
EMP1	5529	3.228	0.446	0.138	-0.015	0.17	Up
CCDC88A	3110	3.225	0.292	0.09	-0.192	0.17	Up
CDKN1A	3409	3.225	0.233	0.072	-0.35	0.17	Up
LHFPL2	9644	3.219	0.334	0.104	-0.18	0.17	Up
PCDH18	14608	3.217	0.285	0.089	-0.083	0.17	Up
PLEKHH2	15184	3.209	0.277	0.086	-0.007	0.17	Up
SPARCL1	18363	3.209	0.619	0.193	-0.093	0.17	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
CCRL1	3196	3.209	0.426	0.133	-0.728	0.17	Up
ANKH	647	3.207	0.37	0.116	-0.175	0.17	Up
PENK	14822	3.207	0.258	0.08	-0.535	0.17	Up
LOC728875	11559	3.2	0.463	0.145	-0.23	0.17	Up
PARVA	14557	3.199	0.306	0.096	-0.286	0.17	Up
CAPN5	2876	3.199	0.418	0.131	-0.191	0.17	Up
POTEF	15389	3.198	0.282	0.088	-0.263	0.17	Up
IGF2	8396	3.194	0.325	0.102	-0.114	0.17	Up
UCN2	20436	3.193	0.475	0.149	-0.596	0.17	Up
SYNPO	18864	3.193	0.272	0.085	-0.138	0.17	Up
FNDC4	6658	3.193	0.317	0.099	-0.114	0.17	Up
PLAT	15128	3.19	0.817	0.256	-0.349	0.17	Up
JAM2	8783	3.19	0.332	0.104	0.003	0.17	Up
IFFO1	8333	3.186	0.202	0.064	-0.212	0.17	Up
FAT1	6188	3.184	0.509	0.16	-0.13	0.17	Up
LOC100287221	10387	3.18	0.373	0.117	-0.907	0.17	Up
RNF175	16670	3.179	0.21	0.066	-0.116	0.17	Up
RFX2	16465	3.177	0.318	0.1	-0.31	0.17	Up
LRRC17	11759	3.176	0.574	0.181	-0.452	0.17	Up
VSNL1	20740	3.176	0.526	0.166	-0.291	0.17	Up
LY96	11899	3.167	0.374	0.118	-0.167	0.314	Up
MT1B	12789	3.165	0.472	0.149	-0.084	0.314	Up
PHACTR2	14906	3.163	0.262	0.083	-0.179	0.314	Up
SPRY2	18483	3.161	0.451	0.143	0.03	0.314	Up
KLF6	9208	3.155	0.247	0.078	-0.137	0.314	Up
EFNB1	5377	3.151	0.35	0.111	-0.149	0.314	Up
CPZ	4139	3.146	0.386	0.123	-0.318	0.314	Up
PYGO1	16043	3.145	0.266	0.085	-0.387	0.314	Up
THBS1	19279	3.135	0.392	0.125	-0.537	0.314	Up
SPRY1	18482	3.134	0.42	0.134	-0.101	0.314	Up
ADAMTS1	269	3.132	0.472	0.151	-0.116	0.314	Up
ANGPT2	630	3.131	0.399	0.127	-0.908	0.314	Up
RHOQ	16543	3.13	0.226	0.072	-0.081	0.314	Up
CAV2	2944	3.129	0.246	0.079	0.088	0.314	Up
PCSK7	14699	3.127	0.178	0.057	-0.261	0.314	Up
CCDC82	3102	3.121	0.376	0.12	-0.117	0.314	Up
PLAC9	15124	3.115	0.352	0.113	-0.075	0.314	Up
MSN	12772	3.106	0.358	0.115	-0.204	0.314	Up
CYP2U1	4503	3.103	0.276	0.089	0.001	0.314	Up
ANKRD35	695	3.102	0.582	0.188	-0.058	0.314	Up
METRNL	12309	3.096	0.324	0.105	-0.84	0.314	Up
MFGE8	12350	3.094	0.54	0.175	-0.29	0.314	Up
GAS6	6948	3.094	0.378	0.122	-0.256	0.314	Up
CFI	3568	3.09	0.381	0.123	-0.024	0.314	Up
DCLK2	4629	3.089	0.254	0.082	-0.31	0.314	Up
NEFM	13353	3.085	0.26	0.084	-0.501	0.314	Up
RGS2	16502	3.081	0.525	0.171	-0.07	0.314	Up
ROCK2	16736	3.081	0.244	0.079	-0.26	0.314	Up
HYMAI	8289	3.08	0.372	0.121	-0.067	0.314	Up
VMP1	20685	3.08	0.39	0.127	-0.762	0.314	Up
F13A1	5783	3.078	0.377	0.122	-0.422	0.314	Up
HOXA4	8090	3.077	0.276	0.09	0	0.314	Up
STK17B	18691	3.073	0.287	0.093	-0.015	0.314	Up
LOC401097	11164	3.063	0.204	0.067	-0.413	0.314	Up
ARRDC3	1046	3.061	0.364	0.119	-0.104	0.314	Up
LOC100128905	9881	3.061	0.117	0.038	-0.361	0.314	Up
S100A4	17033	3.06	0.435	0.142	-0.445	0.314	Up
TNFSF11	19742	3.06	0.236	0.077	-0.395	0.314	Up
HOXA3	8089	3.059	0.474	0.155	0.036	0.314	Up
RNF122	16639	3.054	0.312	0.102	-0.292	0.563	Up
H19	7663	3.053	0.554	0.182	-0.132	0.563	Up
NRP2	13711	3.053	0.206	0.068	-0.734	0.563	Up
FMOD	6645	3.052	0.557	0.183	-0.081	0.563	Up
GABARAPL1	6854	3.051	0.296	0.097	-0.055	0.563	Up
MMP7	12547	3.051	1.016	0.333	-0.268	0.563	Up
WTIP	20938	3.05	0.378	0.124	-0.291	0.563	Up
TSHZ2	20085	3.046	0.333	0.109	0.049	0.563	Up
ACTBL2	204	3.042	0.329	0.108	-0.226	0.563	Up
SCN3B	17192	3.041	0.412	0.136	-0.12	0.563	Up
SULF2	18775	3.041	0.538	0.177	-0.706	0.563	Up
DNAJB4	5017	3.041	0.28	0.092	-0.486	0.563	Up
BACE2	1364	3.04	0.55	0.181	-0.229	0.563	Up
FSCN1	6765	3.039	0.455	0.15	-0.468	0.563	Up
PRKD3	15662	3.038	0.344	0.113	-0.169	0.563	Up
GOLIM4	7293	3.037	0.313	0.103	-0.347	0.563	Up
PHLDA3	14948	3.037	0.321	0.106	-0.169	0.563	Up
TWSG1	20280	3.037	0.294	0.097	-0.87	0.563	Up
RASL11B	16267	3.036	0.459	0.151	-0.413	0.563	Up
F3	5790	3.032	0.576	0.19	-0.179	0.563	Up
COL4A1	3984	3.03	0.396	0.131	-0.48	0.563	Up
CAV1	2943	3.029	0.453	0.15	0.049	0.563	Up
VEGFB	20656	3.024	0.241	0.08	-0.19	0.563	Up
ARSE	1052	3.018	0.281	0.093	-0.845	0.563	Up
PLAU	15129	3.018	0.549	0.182	-0.879	0.563	Up
BCAT1	1431	3.014	0.326	0.108	-0.869	0.563	Up
QKI	16054	3.013	0.225	0.075	-0.178	0.563	Up
SYDE1	18841	3.006	0.157	0.052	-0.202	0.563	Up
FNDC3B	6657	3.004	0.271	0.09	-0.358	0.563	Up
MAP7D3	12072	3.002	0.203	0.068	-0.056	0.563	Up
CSRP2	4277	3.001	0.425	0.142	-0.199	0.563	Up
OLFML2A	13916	3	0.239	0.08	-0.733	0.563	Up
MOXD1	12603	3	0.46	0.153	-0.031	0.563	Up
CDKL5	3408	2.999	0.332	0.111	-0.403	0.563	Up
LATS2	9526	2.998	0.213	0.071	-0.198	0.563	Up
CCDC3	3053	2.995	0.465	0.155	0.026	0.563	Up
DAPK3	4570	2.995	0.254	0.085	-0.967	0.563	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
CDK6	3399	2.99	0.497	0.166	-0.178	0.563	Up
DOCK4	5087	2.99	0.299	0.1	-0.416	0.563	Up
ADPRH	346	2.988	0.3	0.1	-0.044	0.563	Up
MMD	12522	2.986	0.371	0.124	-0.108	0.563	Up
RFPL1-AS1	16453	2.982	0.235	0.079	-0.284	0.563	Up
TGFBR1	19248	2.981	0.275	0.092	-0.496	0.563	Up
LPHN2	11700	2.977	0.225	0.075	-0.073	0.563	Up
LOC401164	11168	2.976	0.233	0.078	-0.373	0.563	Up
CLDN11	3758	2.973	0.544	0.183	0.032	0.563	Up
LAMC1	9496	2.968	0.378	0.127	-0.419	0.563	Up
IGFBP6	8408	2.964	0.395	0.133	-0.062	0.563	Up
MME	12524	2.964	0.363	0.122	0.14	0.563	Up
LRRC15	11756	2.957	0.71	0.24	-1.313	0.563	Up
S100A2	17031	2.957	0.715	0.242	-0.381	0.563	Up
GULP1	7638	2.953	0.304	0.103	-0.055	0.563	Up
ZNF521	21525	2.953	0.403	0.136	-0.31	0.563	Up
TMEM119	19459	2.949	0.474	0.161	-0.23	0.563	Up
NAMPT	13056	2.948	0.385	0.131	-0.251	0.563	Up
KDELC1	8967	2.948	0.29	0.098	-0.581	0.563	Up
LOC100129397	9942	2.947	0.229	0.078	-0.564	0.563	Up
FOXL1	6705	2.945	0.335	0.114	-0.364	0.563	Up
GUCY1B3	7630	2.944	0.212	0.072	-0.481	0.563	Up
KIAA0408	9027	2.939	0.249	0.085	0.064	0.563	Up
CILP	3711	2.938	0.837	0.285	-0.46	0.563	Up
GLIPR1	7153	2.936	0.24	0.082	-0.235	0.563	Up
NRP1	13710	2.936	0.283	0.097	-0.147	0.563	Up
PDLIM2	14773	2.936	0.153	0.052	-0.527	0.563	Up
FLJ42709	6570	2.933	0.268	0.092	-0.16	0.758	Up
LEPREL1	9608	2.933	0.457	0.156	-0.136	0.758	Up
A2M	5	2.932	0.436	0.149	-0.111	0.758	Up
CDKL1	3404	2.929	0.221	0.075	-0.187	0.758	Up
ADM	336	2.928	0.599	0.204	-0.298	0.758	Up
FLJ41170	6550	2.928	0.177	0.06	-0.191	0.758	Up
EDN1	5334	2.926	0.473	0.161	-0.094	0.758	Up
AFAP1L1	370	2.925	0.271	0.092	-0.286	0.758	Up
TMEM217	19572	2.924	0.32	0.11	-0.785	0.758	Up
CARD6	2898	2.92	0.326	0.111	-0.051	0.758	Up
CAPZB	2891	2.92	0.136	0.047	-0.428	0.758	Up
ACTB	203	2.919	0.194	0.067	-0.658	0.758	Up
ADC	301	2.917	0.258	0.089	-0.516	0.758	Up
BACE1	1363	2.915	0.239	0.082	-0.442	0.758	Up
TWIST1	20277	2.915	0.278	0.095	-0.177	0.758	Up
EPB41L2	5587	2.914	0.3	0.103	-0.065	0.758	Up
PAX1	14570	2.912	0.458	0.157	-0.609	0.758	Up
FZD1	6829	2.909	0.386	0.133	-0.27	0.758	Up
GPNMB	7352	2.908	0.261	0.09	-0.201	0.758	Up
RAB3IL1	16119	2.907	0.236	0.081	-0.448	0.758	Up
ADD3	321	2.907	0.483	0.166	-0.042	0.758	Up
HSPA12B	8218	2.906	0.354	0.122	-0.137	0.758	Up
ICAM4	8303	2.904	0.21	0.072	-0.194	0.758	Up
F2RL1	5787	2.904	0.336	0.116	-0.081	0.758	Up
GLI2	7150	2.903	0.135	0.046	-0.306	0.758	Up
LOC100134240	10340	2.901	0.377	0.13	-1.111	0.758	Up
NINJ2	13462	2.9	0.38	0.131	-0.502	0.758	Up
ANXA2P3	760	2.9	0.347	0.12	-0.369	0.758	Up
PITX2	15059	2.898	0.324	0.112	-0.655	0.758	Up
P4HA2	14435	2.894	0.268	0.092	-0.761	0.758	Up
ODZ4	13894	2.894	0.256	0.089	-0.502	0.758	Up
ANKRD37	700	2.892	0.334	0.115	-0.398	0.758	Up
GPX7	7473	2.891	0.312	0.108	-0.39	0.758	Up
ROBO1	16731	2.888	0.219	0.076	-0.466	0.758	Up
CSDC2	4233	2.886	0.199	0.069	-0.411	0.758	Up
FOS	6675	2.882	0.632	0.219	0.091	0.758	Up
MFAP2	12344	2.878	0.488	0.169	-1.208	0.758	Up
EBF1	5293	2.877	0.236	0.082	0.028	0.758	Up
AASS	27	2.876	0.275	0.096	-0.215	0.758	Up
ST5	18615	2.872	0.264	0.092	-0.464	0.758	Up
FKBP7	6417	2.868	0.317	0.111	-0.451	0.758	Up
LOC645166	11387	2.867	0.353	0.123	-0.41	0.758	Up
PDE1B	14728	2.863	0.195	0.068	0.051	0.758	Up
MMP23B	12540	2.863	0.413	0.144	-0.174	0.758	Up
LOC100130776	10099	2.861	0.198	0.069	-0.272	0.758	Up
RAB7B	16137	2.858	0.446	0.156	-0.28	0.758	Up
CSRP1	4276	2.857	0.23	0.081	-0.023	0.758	Up
TGFBI	19247	2.856	0.44	0.154	-0.922	0.758	Up
C1QTNF6	2126	2.856	0.223	0.078	-0.949	0.758	Up
STON1	18717	2.856	0.327	0.115	-0.256	0.758	Up
FMNL3	6637	2.854	0.151	0.053	-0.347	0.758	Up
NOTCH3	13594	2.853	0.317	0.111	-0.306	0.758	Up
SACS	17058	2.85	0.261	0.092	-0.087	0.758	Up
AOC3	770	2.849	0.475	0.167	0.091	0.758	Up
ACTN3	216	2.847	0.249	0.087	-0.426	0.758	Up
DKFZP586K1520	4912	2.847	0.177	0.062	-0.366	0.758	Up
FBLN1	6198	2.842	0.447	0.157	-0.276	0.758	Up
DSEL	5181	2.841	0.345	0.121	-0.355	0.758	Up
SH3KBP1	17527	2.84	0.26	0.092	-0.122	0.758	Up
TRO	20034	2.837	0.373	0.131	-0.537	0.758	Up
PTH1R	15929	2.834	0.109	0.038	-0.089	0.758	Up
EGR3	5398	2.834	0.59	0.208	-0.142	0.758	Up
SERPINE1	17397	2.831	0.232	0.082	-0.262	0.758	Up
PLOD2	15209	2.83	0.439	0.155	-0.523	0.758	Up
TMEM158	19502	2.827	0.518	0.183	-0.637	0.758	Up
ANXA2	758	2.827	0.342	0.121	-0.183	0.758	Up
FMO2	6639	2.827	0.701	0.248	0.065	0.758	Up
LRCH1	11716	2.825	0.154	0.055	-0.391	0.758	Up
DPYSL2	5149	2.823	0.33	0.117	-0.103	0.758	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
MMP3	12546	2.822	0.637	0.226	-0.474	1.096	Up
HOXA7	8093	2.821	0.588	0.209	-0.047	1.096	Up
MT1A	12788	2.821	0.366	0.13	-0.025	1.096	Up
ROBO3	16733	2.819	0.203	0.072	-0.017	1.096	Up
NDEL1	13279	2.819	0.182	0.065	-0.376	1.096	Up
KRT17	9319	2.817	0.962	0.342	-0.11	1.096	Up
NCKAP5L	13137	2.816	0.121	0.043	-0.314	1.096	Up
RCAN2	16364	2.815	0.252	0.089	-0.029	1.096	Up
LRRN4CL	11831	2.811	0.323	0.115	0.045	1.096	Up
LOC283867	10834	2.809	0.42	0.15	-0.641	1.096	Up
DOCK6	5089	2.808	0.261	0.093	-0.311	1.096	Up
AKT3	502	2.806	0.356	0.127	-0.191	1.096	Up
TNFRSF12A	19726	2.806	0.458	0.163	-0.486	1.096	Up
C13orf33	1846	2.804	0.45	0.16	-0.16	1.096	Up
CAST	2934	2.797	0.235	0.084	-0.119	1.096	Up
PRICKLE2	15628	2.797	0.357	0.128	-0.2	1.096	Up
RNF145	16655	2.796	0.264	0.094	-0.199	1.096	Up
TMEM132E	19476	2.794	0.315	0.113	-0.288	1.096	Up
ZFP36L1	21192	2.792	0.337	0.121	-0.115	1.096	Up
MYO5A	12981	2.789	0.244	0.087	-0.7	1.096	Up
PTPRG	15975	2.786	0.146	0.052	-0.242	1.096	Up
CSDA	4232	2.785	0.476	0.171	-0.224	1.096	Up
TNFRSF1A	19733	2.782	0.212	0.076	-0.514	1.096	Up
CDC14B	3306	2.78	0.168	0.061	-0.058	1.096	Up
MOBKL2B	12566	2.779	0.246	0.089	-0.071	1.096	Up
ZNF532	21533	2.774	0.278	0.1	-0.32	1.096	Up
DPP4	5126	2.773	0.401	0.145	-0.356	1.096	Up
C8orf4	2639	2.773	0.74	0.267	-0.606	1.096	Up
TBX5	19077	2.77	0.158	0.057	-0.396	1.096	Up
HS3ST3A1	8166	2.77	0.168	0.061	-0.588	1.096	Up
ANO6	739	2.769	0.143	0.052	-0.114	1.096	Up
PECAM1	14810	2.767	0.277	0.1	0.152	1.096	Up
TGFB3	19246	2.767	0.451	0.163	-0.493	1.096	Up
CNTNAP3	3940	2.765	0.372	0.135	0.018	1.096	Up
C1QTNF1	2121	2.764	0.242	0.088	-0.264	1.096	Up
C1QTNF7	2127	2.762	0.206	0.075	-0.115	1.096	Up
DIO2	4879	2.76	0.267	0.097	-0.293	1.096	Up
EYA2	5772	2.759	0.487	0.177	-0.237	1.096	Up
MFAP4	12347	2.759	0.732	0.265	-0.008	1.096	Up
HAS2	7710	2.756	0.147	0.053	-0.344	1.096	Up
MYO1E	12975	2.755	0.283	0.103	-0.35	1.096	Up
CCDC36	3057	2.754	0.198	0.072	-0.341	1.096	Up
HOXA9	8094	2.75	0.338	0.123	-0.116	1.096	Up
UST	20599	2.748	0.256	0.093	-0.254	1.096	Up
FHAD1	6379	2.748	0.13	0.047	-0.176	1.096	Up
SCN2B	17190	2.746	0.209	0.076	-0.21	1.096	Up
NT5E	13747	2.741	0.263	0.096	-0.252	1.096	Up
SNAI1	18132	2.74	0.312	0.114	-0.499	1.096	Up
CORIN	4049	2.74	0.261	0.095	-0.687	1.096	Up
ATP11B	1188	2.74	0.309	0.113	-0.311	1.096	Up
ST3GAL2	18610	2.74	0.226	0.082	-0.184	1.096	Up
KRT5	9347	2.738	0.864	0.316	-0.065	1.096	Up
NKX3-2	13503	2.735	0.38	0.139	-1.129	1.096	Up
ACOT9	167	2.734	0.254	0.093	-0.21	1.096	Up
S1PR2	17047	2.733	0.255	0.093	-0.046	1.096	Up
DRD4	5161	2.731	0.396	0.145	-1.028	1.096	Up
BHMT2	1509	2.731	0.235	0.086	-0.062	1.096	Up
ETHE1	5713	2.729	0.244	0.09	-0.445	1.096	Up
PARP4	14551	2.729	0.164	0.06	-0.303	1.096	Up
GJC1	7126	2.729	0.379	0.139	-0.579	1.096	Up
PRELP	15614	2.727	0.108	0.04	0.059	1.096	Up
RALB	16195	2.726	0.245	0.09	-0.5	1.096	Up
MT1H	12794	2.724	0.307	0.113	-0.12	1.096	Up
CH25H	3585	2.724	0.468	0.172	0.017	1.096	Up
KRT14	9315	2.723	0.986	0.362	-0.095	1.096	Up
CHRNA7	3667	2.723	0.22	0.081	-0.365	1.096	Up
ANXA8L2	766	2.72	0.931	0.342	-0.288	1.096	Up
GOLGA8E	7289	2.719	0.252	0.093	-0.18	1.096	Up
SERPINE2	17398	2.718	0.401	0.148	-0.396	1.096	Up
KRT16P2	9318	2.716	0.87	0.32	-0.032	1.096	Up
C1orf54	2211	2.715	0.239	0.088	-0.136	1.096	Up
CCNO	3179	2.713	0.831	0.306	-0.487	1.096	Up
CDKN1C	3411	2.712	0.4	0.148	-0.045	1.096	Up
HSD17B11	8182	2.712	0.32	0.118	-0.031	1.096	Up
SLC16A7	17695	2.708	0.18	0.066	0.057	1.642	Up
RPS6KA3	16898	2.708	0.268	0.099	0.032	1.642	Up
COL14A1	3966	2.708	0.443	0.164	-0.092	1.642	Up
NEDD9	13350	2.708	0.294	0.109	-0.244	1.642	Up
ARHGEF4	964	2.705	0.234	0.086	-0.148	1.642	Up
KANK4	8820	2.701	0.595	0.22	-0.612	1.642	Up
FILIP1L	6399	2.699	0.208	0.077	-0.448	1.642	Up
GSTP1	7578	2.698	0.523	0.194	-0.051	1.642	Up
MEF2A	12269	2.696	0.263	0.098	-0.243	1.642	Up
LRP12	11738	2.692	0.316	0.118	-0.571	1.642	Up
PCDHGA8	14652	2.691	0.196	0.073	-0.136	1.642	Up
IQCJ-SCHIP1	8647	2.691	0.291	0.108	-0.112	1.642	Up
PID1	14982	2.691	0.371	0.138	0.003	1.642	Up
MECOM	12236	2.689	0.209	0.078	0	1.642	Up
CTIF	4326	2.687	0.292	0.109	-0.197	1.642	Up
ITGB6	8736	2.686	0.538	0.2	-0.264	1.642	Up
CYLD	4469	2.686	0.215	0.08	0.056	1.642	Up
KLF9	9211	2.683	0.361	0.134	-0.149	1.642	Up
PDE7B	14746	2.682	0.224	0.083	-0.149	1.642	Up
ZNF880	21783	2.681	0.214	0.08	-0.351	1.642	Up
TNC	19710	2.68	0.58	0.216	-0.317	1.642	Up
FOSL2	6678	2.676	0.403	0.151	-0.572	1.642	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
SGCB	17469	2.675	0.254	0.095	-0.398	1.642	Up
MAFB	11955	2.675	0.294	0.11	-0.196	1.642	Up
ZNF699	21663	2.674	0.137	0.051	-0.141	1.642	Up
LOC100505806	10454	2.673	0.425	0.159	-0.422	1.642	Up
KLF10	9196	2.672	0.316	0.118	-0.162	1.642	Up
GOLGA6L10	7281	2.671	0.306	0.114	-0.528	1.642	Up
GRID1	7502	2.669	0.106	0.04	0.021	1.642	Up
CDHR2	3370	2.668	0.224	0.084	-0.565	1.642	Up
SLC38A2	17891	2.667	0.235	0.088	-0.101	1.642	Up
EMX2	5537	2.667	0.217	0.081	-0.141	1.642	Up
NOTCH4	13595	2.666	0.308	0.116	-0.008	1.642	Up
RCBTB2	16367	2.666	0.225	0.085	0.07	1.642	Up
ODZ2	13892	2.664	0.525	0.197	-0.176	1.642	Up
LOC728061	11506	2.661	0.381	0.143	-0.493	1.642	Up
NRN1	13708	2.661	0.332	0.125	-0.239	1.642	Up
SPATA9	18385	2.661	0.132	0.05	-0.328	1.642	Up
MRAS	12639	2.659	0.384	0.144	0.067	1.642	Up
PTPN14	15954	2.659	0.362	0.136	-0.272	1.642	Up
JUN	8810	2.657	0.413	0.156	0.082	1.642	Up
RORA	16745	2.657	0.238	0.09	-0.203	1.642	Up
CILP2	3712	2.656	0.478	0.18	-0.848	1.642	Up
VDR	20654	2.653	0.293	0.11	-0.394	1.642	Up
TNFAIP8	19716	2.647	0.272	0.103	-0.062	1.642	Up
CDH2	3356	2.647	0.485	0.183	-1.018	1.642	Up
RASD1	16251	2.643	0.56	0.212	-0.016	1.642	Up
UBE2Q2P1	20373	2.636	0.239	0.091	-0.105	1.642	Up
SMAD9	18062	2.636	0.162	0.062	0.083	1.642	Up
XYLT1	20996	2.634	0.23	0.087	-0.308	1.642	Up
DNAJC18	5032	2.634	0.179	0.068	-0.068	1.642	Up
LOC400464	11131	2.63	0.194	0.074	-0.094	1.642	Up
PER1	14824	2.628	0.349	0.133	-0.083	1.642	Up
RUNX1	16998	2.626	0.294	0.112	-0.723	1.642	Up
FAM20C	6011	2.625	0.328	0.125	-0.096	1.642	Up
LARP6	9517	2.625	0.335	0.128	-0.188	1.642	Up
PURG	16006	2.623	0.128	0.049	-0.404	1.642	Up
FABP5	5805	2.621	0.431	0.164	-0.037	1.642	Up
C13orf15	1837	2.618	0.314	0.12	-0.081	1.642	Up
ADAMTSL2	289	2.617	0.271	0.104	-0.975	1.642	Up
EHD3	5404	2.617	0.197	0.075	-0.329	1.642	Up
SDC1	17226	2.615	0.448	0.171	-0.817	1.642	Up
TSC22D3	20069	2.615	0.323	0.123	-0.164	1.642	Up
PCNX	14684	2.615	0.23	0.088	-0.309	1.642	Up
FAM43B	6041	2.611	0.268	0.103	-0.301	1.642	Up
CHST12	3679	2.611	0.202	0.077	-0.136	1.642	Up
RCN1	16374	2.61	0.314	0.12	-0.123	1.642	Up
FAM89A	6139	2.609	0.334	0.128	0.164	1.642	Up
CXCL2	4403	2.607	0.598	0.229	0.08	1.642	Up
MIAT	12430	2.606	0.468	0.18	-0.516	1.642	Up
KCNJ15	8896	2.605	0.112	0.043	-0.733	1.642	Up
NPHP3	13626	2.603	0.203	0.078	-0.306	2.4	Up
LOC100190939	10363	2.601	0.347	0.133	-0.197	2.4	Up
IFI16	8335	2.599	0.324	0.124	-0.151	2.4	Up
EMILIN2	5521	2.598	0.335	0.129	-0.133	2.4	Up
CPE	4098	2.598	0.525	0.202	-0.193	2.4	Up
LOC100130000	10005	2.597	0.668	0.257	-0.297	2.4	Up
IL4R	8527	2.594	0.245	0.095	-0.073	2.4	Up
COX7A1	4079	2.593	0.459	0.177	-0.187	2.4	Up
ZCCHC11	21121	2.592	0.216	0.083	-0.133	2.4	Up
C6orf145	2533	2.59	0.301	0.116	-0.088	2.4	Up
PLAGL1	15126	2.589	0.212	0.082	-0.002	2.4	Up
AFAP1	368	2.589	0.186	0.072	-0.5	2.4	Up
CSPG4	4268	2.588	0.339	0.131	-0.52	2.4	Up
CD99P1	3300	2.588	0.212	0.082	-0.895	2.4	Up
ANKS6	729	2.587	0.417	0.161	-0.067	2.4	Up
UBTD2	20414	2.587	0.217	0.084	-1.039	2.4	Up
MAGEH1	11990	2.585	0.302	0.117	-0.225	2.4	Up
TP53I3	19830	2.585	0.279	0.108	-0.1	2.4	Up
CHST11	3678	2.585	0.264	0.102	-0.649	2.4	Up
ARHGAP10	910	2.583	0.274	0.106	-0.257	2.4	Up
GOPC	7300	2.58	0.189	0.073	-0.453	2.4	Up
C9orf3	2716	2.58	0.181	0.07	-0.525	2.4	Up
KATNAL1	8830	2.579	0.244	0.095	-0.407	2.4	Up
BACH1	1365	2.579	0.163	0.063	-0.239	2.4	Up
LBH	9529	2.579	0.169	0.065	-0.365	2.4	Up
IGDCC4	8393	2.579	0.295	0.114	-0.142	2.4	Up
TTC23	20157	2.579	0.239	0.093	-0.199	2.4	Up
HOOK3	8078	2.578	0.241	0.094	-0.239	2.4	Up
ENTPD1	5570	2.577	0.201	0.078	-0.179	2.4	Up
OLFML1	13915	2.577	0.176	0.068	-0.238	2.4	Up
LMNA	9731	2.576	0.216	0.084	-0.356	2.4	Up
MYH9	12944	2.576	0.208	0.081	-0.363	2.4	Up
HIVEP2	7964	2.576	0.283	0.11	-0.46	2.4	Up
PAPPA	14516	2.575	0.292	0.114	-0.806	2.4	Up
FHOD3	6387	2.571	0.404	0.157	-0.342	2.4	Up
CLEC3B	3797	2.568	0.378	0.147	-0.009	2.4	Up
ADCY7	314	2.567	0.273	0.106	-0.502	2.4	Up
TGFBR2	19249	2.563	0.296	0.115	0.134	2.4	Up
IL11	8451	2.56	0.23	0.09	-1.293	2.4	Up
AREG	890	2.56	0.815	0.318	-0.133	2.4	Up
WDR63	20837	2.556	0.172	0.067	-0.011	2.4	Up
BHLHE22	1504	2.555	0.35	0.137	0.008	2.4	Up
ASNS	1113	2.554	0.301	0.118	-0.498	2.4	Up
GALNTL2	6919	2.554	0.386	0.151	-0.12	2.4	Up
WIPF1	20892	2.549	0.207	0.081	-0.171	2.4	Up
STX2	18746	2.549	0.298	0.117	-0.361	2.4	Up
MT2A	12799	2.548	0.396	0.155	-0.074	2.4	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
DUSP6	5238	2.548	0.383	0.151	0.042	2.4	Up
ZNF438	21469	2.546	0.185	0.073	-0.237	2.4	Up
AFAP1L2	371	2.544	0.332	0.131	-0.214	2.4	Up
BNIP3L	1566	2.543	0.235	0.093	-0.62	2.4	Up
NPC2	13617	2.54	0.277	0.109	-0.113	2.4	Up
GABRB2	6868	2.54	0.156	0.062	-0.559	2.4	Up
PRR5L	15759	2.532	0.228	0.09	-0.815	2.4	Up
PCDHB14	14634	2.531	0.353	0.139	-0.217	2.4	Up
FOXN3	6710	2.529	0.222	0.088	-0.131	2.4	Up
LOC100507507	10534	2.528	0.237	0.094	-0.511	2.4	Up
LOC145694	10642	2.52	0.313	0.124	-0.698	2.4	Up
CLIC4	3815	2.518	0.256	0.102	-0.555	2.4	Up
UNC5B	20491	2.518	0.382	0.152	-1.797	2.4	Up
CLDN5	3774	2.517	0.411	0.163	0.084	2.4	Up
LOC653075	11484	2.51	0.312	0.124	-0.04	2.4	Up
CPNE8	4116	2.51	0.187	0.075	-0.029	2.4	Up
TFPI	19235	2.508	0.283	0.113	0.056	2.4	Up
ERMN	5670	2.504	0.16	0.064	-0.37	3.164	Up
VCL	20646	2.503	0.247	0.099	-0.251	3.164	Up
PCDHB16	14636	2.503	0.375	0.15	-0.069	3.164	Up
ARHGAP31	929	2.502	0.147	0.059	-0.345	3.164	Up
PPP3CC	15533	2.501	0.196	0.079	-0.376	3.164	Up
TNAP	19709	2.501	0.157	0.063	0.042	3.164	Up
KRT42P	9346	2.499	0.33	0.132	-0.14	3.164	Up
C3orf64	2427	2.499	0.143	0.057	-0.295	3.164	Up
TSPAN2	20107	2.497	0.179	0.072	-0.197	3.164	Up
GLG1	7148	2.496	0.2	0.08	-0.237	3.164	Up
PLEKHG1	15175	2.495	0.248	0.099	-0.09	3.164	Up
SYN2	18847	2.494	0.263	0.106	0.024	3.164	Up
TMEM117	19458	2.494	0.218	0.088	-0.408	3.164	Up
RAP2B	16225	2.493	0.171	0.069	-0.362	3.164	Up
SLC39A13	17903	2.49	0.179	0.072	-0.532	3.164	Up
C3orf57	2423	2.487	0.188	0.076	0.027	3.164	Up
ZSWIM4	21834	2.487	0.348	0.14	-0.271	3.164	Up
PLSCR4	15219	2.487	0.301	0.121	0.031	3.164	Up
RAPGEF2	16228	2.486	0.159	0.064	-0.042	3.164	Up
CYP27C1	4488	2.484	0.375	0.151	-0.495	3.164	Up
TBC1D2	19021	2.484	0.243	0.098	-0.526	3.164	Up
ARSB	1050	2.483	0.171	0.069	-0.682	3.164	Up
NIPAL4	13470	2.483	0.338	0.136	-0.739	3.164	Up
BTBD19	1641	2.482	0.114	0.046	-0.895	3.164	Up
ZBTB1	21060	2.481	0.201	0.081	-0.639	3.164	Up
FLRT2	6623	2.481	0.17	0.068	-0.158	3.164	Up
PDGFA	14750	2.481	0.22	0.089	-0.197	3.164	Up
sep.10	17341	2.478	0.272	0.11	-0.219	3.164	Up
PTGIS	15923	2.473	0.4	0.162	0.069	3.164	Up
PABPC4L	14448	2.472	0.228	0.092	-0.098	3.164	Up
RABAC1	16144	2.47	0.19	0.077	-0.901	3.164	Up
CTNNB1	4332	2.469	0.239	0.097	-0.329	3.164	Up
THSD7A	19305	2.467	0.246	0.1	-0.182	3.164	Up
VPS36	20715	2.465	0.208	0.084	-0.286	3.164	Up
PTTG1IP	15996	2.464	0.182	0.074	-0.535	3.164	Up
TCEAL7	19092	2.462	0.23	0.093	-0.125	3.164	Up
CHN1	3641	2.461	0.266	0.108	-0.68	3.164	Up
CXorf57	4437	2.461	0.262	0.106	0.071	3.164	Up
PPP1R13L	15487	2.458	0.241	0.098	-0.456	3.164	Up
FLJ36031	6513	2.458	0.275	0.112	-0.001	3.164	Up
LGALS7	9627	2.458	0.605	0.246	-0.329	3.164	Up
FYN	6827	2.456	0.313	0.127	-0.087	3.164	Up
ZCCHC9	21137	2.455	0.126	0.051	-0.231	3.164	Up
WDR41	20817	2.454	0.19	0.077	-0.585	3.164	Up
SPG20	18414	2.454	0.222	0.09	-0.177	3.164	Up
GLS	7175	2.453	0.239	0.098	-0.171	3.164	Up
STX7	18751	2.453	0.177	0.072	-0.022	3.164	Up
IL16	8462	2.45	0.154	0.063	-0.167	3.164	Up
PPAP2B	15419	2.449	0.327	0.134	-0.006	3.164	Up
STMN2	18710	2.447	0.471	0.193	-0.52	3.164	Up
SLC20A1	17720	2.446	0.199	0.081	-1.141	3.164	Up
ID1	8310	2.443	0.412	0.169	-0.014	3.164	Up
PPP1R15A	15492	2.442	0.207	0.085	-0.097	3.164	Up
MATN3	12139	2.441	0.428	0.175	-0.958	3.164	Up
ANGPTL4	635	2.44	0.587	0.24	0.071	3.164	Up
PGM5	14896	2.44	0.21	0.086	0.08	3.164	Up
CCDC102A	2982	2.439	0.257	0.105	-0.239	3.164	Up
RASGRF2	16258	2.439	0.229	0.094	-0.466	3.164	Up
HECW2	7808	2.436	0.294	0.121	-0.428	3.164	Up
PPAP2A	15418	2.436	0.273	0.112	0.139	3.164	Up
ACAA2	116	2.435	0.289	0.119	-0.063	3.164	Up
CLIP2	3820	2.435	0.163	0.067	-0.466	3.164	Up
MYL12B	12948	2.434	0.173	0.071	-0.703	3.164	Up
COG6	3957	2.434	0.121	0.05	-0.323	3.164	Up
LOC100127983	9769	2.432	0.194	0.08	-0.353	3.164	Up
LOC100128252	9801	2.432	0.378	0.156	-0.229	3.164	Up
IL1RAP	8484	2.432	0.227	0.094	-0.35	3.164	Up
LIX1L	9715	2.43	0.167	0.069	0.008	3.164	Up
CRYAB	4209	2.429	0.684	0.282	0.041	3.164	Up
LRRC8A	11813	2.429	0.191	0.079	-0.575	3.164	Up
ZYX	21847	2.427	0.253	0.104	-0.248	3.164	Up
FLNC	6619	2.423	0.146	0.06	-0.467	3.164	Up
KLHL28	9240	2.421	0.196	0.081	-0.759	3.164	Up
HOXA6	8092	2.418	0.432	0.178	-0.011	3.164	Up
ELK3	5487	2.417	0.145	0.06	-0.425	3.164	Up
COL11A1	3962	2.414	0.234	0.097	-1.074	3.164	Up
LOC375196	11053	2.413	0.312	0.129	-0.423	3.164	Up
PLTP	15221	2.412	0.304	0.126	-0.008	3.164	Up
SIRPA	17610	2.409	0.279	0.116	0.015	3.164	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
OXT	14416	2.409	0.178	0.074	-0.738	3.164	Up
LOC100129675	9969	2.408	0.207	0.086	-0.163	3.164	Up
ADAMTS9	287	2.408	0.309	0.128	-0.026	3.164	Up
PLA2G5	15113	2.405	0.201	0.084	-0.476	4.418	Up
ARL15	988	2.404	0.213	0.089	-0.225	4.418	Up
BAG2	1369	2.403	0.305	0.127	-0.884	4.418	Up
PCDH7	14611	2.403	0.299	0.124	-0.646	4.418	Up
GYPC	7653	2.401	0.163	0.068	0.174	4.418	Up
TTC7A	20183	2.401	0.219	0.091	-0.232	4.418	Up
PDE4DIP	14737	2.401	0.172	0.072	-0.28	4.418	Up
ANXA5	763	2.4	0.221	0.092	-0.144	4.418	Up
TNNI2	19766	2.398	0.519	0.216	0.094	4.418	Up
MALL	12005	2.398	0.403	0.168	0.047	4.418	Up
CCIN	3125	2.398	0.144	0.06	-0.553	4.418	Up
WBP5	20783	2.397	0.339	0.141	-0.096	4.418	Up
TTC28	20163	2.396	0.132	0.055	-0.321	4.418	Up
COL17A1	3969	2.396	0.544	0.227	-0.044	4.418	Up
CXCR7	4415	2.396	0.393	0.164	-0.845	4.418	Up
SOCS3	18279	2.394	0.364	0.152	-0.03	4.418	Up
SALL1	17064	2.393	0.097	0.041	-0.469	4.418	Up
ZNF575	21570	2.392	0.168	0.07	-0.606	4.418	Up
CALU	2839	2.391	0.266	0.111	-1.039	4.418	Up
SGTB	17494	2.391	0.18	0.075	-0.164	4.418	Up
CACNA1C	2785	2.391	0.184	0.077	-0.099	4.418	Up
NAALADL1	13032	2.391	0.111	0.046	-0.106	4.418	Up
CLEC2B	3793	2.39	0.272	0.114	-0.119	4.418	Up
SLC38A5	17894	2.388	0.247	0.103	-1.293	4.418	Up
PTPRB	15969	2.387	0.229	0.096	-0.029	4.418	Up
MYL12A	12947	2.387	0.236	0.099	-0.432	4.418	Up
CHSY3	3693	2.387	0.131	0.055	-0.621	4.418	Up
ZNF555	21549	2.386	0.184	0.077	-0.437	4.418	Up
PPM1K	15472	2.384	0.353	0.148	-0.725	4.418	Up
PIWIL4	15064	2.383	0.275	0.115	-0.089	4.418	Up
SYPL2	18871	2.382	0.155	0.065	-0.174	4.418	Up
ADAMTS16	275	2.378	0.326	0.137	-0.575	4.418	Up
C6orf204	2553	2.378	0.121	0.051	0.002	4.418	Up
MAN1A1	12014	2.378	0.308	0.129	-0.11	4.418	Up
C1orf123	2147	2.378	0.17	0.071	-0.287	4.418	Up
EDA2R	5322	2.375	0.175	0.074	-0.129	4.418	Up
KIAA0922	9051	2.375	0.256	0.108	-0.312	4.418	Up
IL8	8535	2.374	0.134	0.056	-0.169	4.418	Up
LEPREL2	9609	2.373	0.378	0.159	-0.634	4.418	Up
PHC2	14913	2.372	0.143	0.06	-0.319	4.418	Up
SLC35B4	17864	2.369	0.187	0.079	-0.443	4.418	Up
PLK2	15202	2.369	0.412	0.174	-0.398	4.418	Up
EVI2A	5728	2.368	0.282	0.119	-0.246	4.418	Up
ADAMTS10	270	2.366	0.339	0.143	-0.333	4.418	Up
GNPTAB	7264	2.366	0.188	0.079	-0.303	4.418	Up
RAB24	16093	2.363	0.221	0.093	-0.593	4.418	Up
STARD9	18662	2.362	0.134	0.057	-0.145	4.418	Up
CNTN1	3932	2.361	0.212	0.09	-0.217	4.418	Up
GAS7	6949	2.36	0.226	0.096	-0.144	4.418	Up
WLS	20901	2.36	0.313	0.132	-0.297	4.418	Up
CREB3L2	4157	2.359	0.23	0.098	-0.281	4.418	Up
KPNA3	9293	2.359	0.156	0.066	-0.435	4.418	Up
TBX18	19069	2.358	0.152	0.065	-0.065	4.418	Up
MMP10	12528	2.356	0.465	0.198	-0.691	4.418	Up
PRDM8	15603	2.355	0.17	0.072	0.045	4.418	Up
CYP1B1	4478	2.351	0.454	0.193	-0.169	4.418	Up
LOC100288615	10406	2.35	0.208	0.089	-0.087	4.418	Up
B3GALTL	1332	2.35	0.237	0.101	-0.898	4.418	Up
SLC39A14	17904	2.348	0.182	0.078	-0.752	4.418	Up
CORO1C	4052	2.347	0.223	0.095	-0.322	4.418	Up
LOC283392	10795	2.346	0.346	0.148	0.002	4.418	Up
DNM1P46	5067	2.346	0.207	0.088	0.143	4.418	Up
RAB31	16102	2.345	0.428	0.183	-0.644	4.418	Up
ARRDC4	1047	2.344	0.31	0.132	-0.018	4.418	Up
FLJ22536	6456	2.342	0.387	0.165	-0.358	4.418	Up
LAMB3	9494	2.341	0.558	0.238	0.039	4.418	Up
HEY1	7847	2.34	0.312	0.133	-0.335	4.418	Up
PTN	15939	2.339	0.567	0.242	0.078	4.418	Up
GNG2	7244	2.333	0.179	0.077	0.049	4.418	Up
CYB5R3	4453	2.332	0.155	0.067	-0.407	4.418	Up
DLX5	4962	2.332	0.329	0.141	-0.539	4.418	Up
SLC24A3	17751	2.331	0.501	0.215	-0.652	4.418	Up
RGL1	16479	2.331	0.272	0.117	0.101	4.418	Up
AVPR1A	1308	2.328	0.173	0.074	-0.076	4.418	Up
PDE2A	14730	2.328	0.317	0.136	0.227	4.418	Up
SOX17	18311	2.328	0.306	0.131	0.103	4.418	Up
MYH10	12931	2.328	0.194	0.084	-0.536	4.418	Up
LOC643650	11336	2.327	0.409	0.176	-0.11	4.418	Up
DEPDC7	4793	2.325	0.233	0.1	-1.047	4.418	Up
ZNF423	21457	2.325	0.213	0.092	-0.26	4.418	Up
SLC16A3	17691	2.324	0.394	0.169	-1.342	4.418	Up
CCND2	3163	2.323	0.28	0.121	0.028	4.418	Up
KIF26B	9152	2.322	0.155	0.067	-1.102	4.418	Up
FHL2	6383	2.321	0.324	0.139	-0.201	4.418	Up
TTC37	20175	2.32	0.169	0.073	-0.315	4.418	Up
LOR	11679	2.32	0.137	0.059	-0.092	4.418	Up
CUBN	4369	2.319	0.082	0.035	-0.034	4.418	Up
ATP10A	1184	2.317	0.189	0.082	-0.752	4.418	Up
COL13A1	3965	2.316	0.308	0.133	-0.555	4.418	Up
CAMK2N1	2848	2.316	0.461	0.199	-0.354	4.418	Up
HAPLN3	7702	2.315	0.53	0.229	-0.332	4.418	Up
MPZL3	12635	2.315	0.179	0.077	-0.337	4.418	Up
C5orf46	2497	2.313	0.32	0.138	-0.783	4.418	Up

Gene ID	Gene Name	Score (d)	Numerator (r)	Denominator (s+s0)	Fold Change	adjusted *P* value (%)	Direction
DOCK5	5088	2.312	0.21	0.091	-0.756	4.418	Up
LOC283143	10787	2.312	0.113	0.049	-0.289	4.418	Up
RAI14	16192	2.31	0.251	0.109	-1.22	4.418	Up
SDPR	17247	2.309	0.373	0.161	0.196	4.418	Up
GOLGA8A	7288	2.308	0.249	0.108	-0.04	4.418	Up
FAM49A	6054	2.308	0.233	0.101	-0.021	4.418	Up

Add. Table 6: Statistically over-represented annotation terms, according to DAVID, of differently expressed genes between metabolic cluster Mc1 and Mc2

Annotation Cluster 1		Enrichment Score: 12.55	Count	P_Value	Benjamini
	SP_PIR_KEYWORDS	signal	139	3.60E-19	1.40E-16
	UP_SEQ_FEATURE	signal peptide	139	6.30E-19	9.00E-16
	GOTERM_CC_FAT	extracellular region part	71	1.60E-18	4.00E-16
	SP_PIR_KEYWORDS	Secreted	83	6.40E-14	1.30E-11
	GOTERM_CC_FAT	extracellular region	99	6.70E-14	8.30E-12
	SP_PIR_KEYWORDS	disulfide bond	113	5.50E-12	7.30E-10
	UP_SEQ_FEATURE	disulfide bond	109	2.30E-11	1.70E-08
	GOTERM_CC_FAT	extracellular space	43	7.10E-09	3.50E-07
	SP_PIR_KEYWORDS	glycoprotein	137	1.30E-08	1.00E-06
	UP_SEQ_FEATURE	glycosylation site:N-linked (GlcNAc...)	129	1.70E-07	8.40E-05
Annotation Cluster 2		Enrichment Score: 10.65	Count	P_Value	Benjamini
	GOTERM_CC_FAT	extracellular region part	71	1.60E-18	4.00E-16
	GOTERM_CC_FAT	extracellular matrix	34	4.40E-12	3.70E-10
	SP_PIR_KEYWORDS	extracellular matrix	26	2.70E-11	2.70E-09
	GOTERM_CC_FAT	proteinaceous extracellular matrix	31	7.30E-11	4.50E-09
	GOTERM_CC_FAT	extracellular matrix part	11	3.90E-04	6.40E-03
Annotation Cluster 3		Enrichment Score: 8.78	Count	P_Value	Benjamini
	GOTERM_BP_FAT	cell adhesion	46	8.80E-11	9.00E-08
	GOTERM_BP_FAT	biological adhesion	46	9.10E-11	6.20E-08
	SP_PIR_KEYWORDS	cell adhesion	27	5.90E-07	3.90E-05
Annotation Cluster 4		Enrichment Score: 7.45	Count	P_Value	Benjamini
	GOTERM_BP_FAT	vasculature development	25	1.60E-09	6.70E-07
	GOTERM_BP_FAT	blood vessel development	23	2.50E-08	7.50E-06
	GOTERM_BP_FAT	angiogenesis	17	1.70E-07	2.90E-05
	GOTERM_BP_FAT	blood vessel morphogenesis	20	2.20E-07	3.50E-05
Annotation Cluster 5		Enrichment Score: 6.23	Count	P_Value	Benjamini
	GOTERM_BP_FAT	regulation of locomotion	22	1.60E-09	8.10E-07
	GOTERM_BP_FAT	regulation of cell migration	19	3.60E-08	9.30E-06
	GOTERM_BP_FAT	regulation of cell motion	19	2.80E-07	4.10E-05
	GOTERM_BP_FAT	positive regulation of locomotion	13	1.60E-06	2.00E-04
	GOTERM_BP_FAT	positive regulation of cell migration	11	2.60E-05	2.30E-03
	GOTERM_BP_FAT	positive regulation of cell motion	11	6.00E-05	4.30E-03
Annotation Cluster 6		Enrichment Score: 6.11	Count	P_Value	Benjamini
	GOTERM_BP_FAT	regulation of response to external stimulus	23	6.00E-12	1.20E-08
	GOTERM_BP_FAT	regulation of inflammatory response	13	9.60E-08	2.00E-05
	GOTERM_BP_FAT	negative regulation of defense response	9	9.20E-07	1.20E-04
	GOTERM_BP_FAT	negative regulation of inflammatory response	7	5.10E-05	3.90E-03
	GOTERM_BP_FAT	negative regulation of response to stimulus	11	7.20E-05	4.90E-03
	GOTERM_BP_FAT	negative regulation of response to external stimulus	8	1.10E-04	6.00E-03
Annotation Cluster 7		Enrichment Score: 5.56	Count	P_Value	Benjamini
	GOTERM_MF_FAT	polysaccharide binding	18	8.00E-08	4.00E-05
	GOTERM_MF_FAT	pattern binding	18	8.00E-08	4.00E-05
	GOTERM_MF_FAT	glycosaminoglycan binding	17	1.20E-07	3.00E-05
	GOTERM_MF_FAT	heparin binding	12	2.20E-05	1.60E-03
	SP_PIR_KEYWORDS	heparin-binding	9	5.00E-05	1.80E-03
	GOTERM_MF_FAT	carbohydrate binding	20	5.00E-04	2.30E-02
Annotation Cluster 8		Enrichment Score: 5.22	Count	P_Value	Benjamini
	GOTERM_CC_FAT	actin cytoskeleton	22	1.20E-06	5.10E-05
	GOTERM_MF_FAT	cytoskeletal protein binding	30	4.40E-06	7.40E-04
	SP_PIR_KEYWORDS	actin-binding	18	1.30E-05	6.50E-04
	GOTERM_MF_FAT	actin binding	22	1.90E-05	1.90E-03
Annotation Cluster 9		Enrichment Score: 4.96	Count	P_Value	Benjamini
	GOTERM_BP_FAT	cytoskeleton organization	27	4.40E-06	4.70E-04
	GOTERM_BP_FAT	actin cytoskeleton organization	18	1.10E-05	1.10E-03
	GOTERM_BP_FAT	actin filament-based process	18	2.60E-05	2.40E-03
Annotation Cluster 10		Enrichment Score: 4.2	Count	P_Value	Benjamini
	INTERPRO	EGF-like calcium-binding	14	1.40E-07	9.10E-05
	INTERPRO	EGF-like calcium-binding, conserved site	14	1.40E-07	9.10E-05
	INTERPRO	EGF-type aspartate/asparagine hydroxylation conserved site	14	1.60E-07	5.20E-05
	SMART	EGF_CA	14	1.10E-06	1.80E-04
	INTERPRO	EGF	14	3.60E-06	7.90E-04
	INTERPRO	EGF-like, type 3	17	4.90E-06	7.90E-04
	SP_PIR_KEYWORDS	egf-like domain	18	5.10E-06	2.90E-04
	INTERPRO	EGF-like region, conserved site	21	6.00E-06	7.80E-04
	INTERPRO	EGF-like	16	3.10E-05	3.40E-03
	UP_SEQ_FEATURE	domain:EGF-like 1	12	3.60E-05	1.00E-02
	UP_SEQ_FEATURE	domain:EGF-like 3; calcium-binding	7	9.60E-05	2.00E-02
	INTERPRO	EGF calcium-binding	9	1.50E-04	1.30E-02
	SMART	EGF	16	2.40E-04	1.90E-02
	UP_SEQ_FEATURE	domain:EGF-like 2	9	4.40E-04	6.20E-02
	UP_SEQ_FEATURE	domain:EGF-like 5; calcium-binding	6	8.90E-04	9.40E-02
	UP_SEQ_FEATURE	domain:EGF-like 4	7	1.20E-03	1.10E-01
	UP_SEQ_FEATURE	domain:EGF-like 6; calcium-binding	5	1.60E-03	1.20E-01
	UP_SEQ_FEATURE	domain:EGF-like 4; calcium-binding	5	4.40E-03	2.00E-01
	UP_SEQ_FEATURE	domain:EGF-like 2; calcium-binding	6	6.10E-03	2.30E-01
	UP_SEQ_FEATURE	domain:EGF-like 7; calcium-binding	4	1.70E-02	4.10E-01
	UP_SEQ_FEATURE	domain:EGF-like 10; calcium-binding	3	2.80E-02	5.00E-01

Add. Table 7: Statistically over-represented annotation terms, according to DAVID, of differently expressed genes between metabolic cluster Mc1 and Mc3

Annotation Cluster 1		Enrichment Score: 39.49	Count	P_Value	Benjamini
	SP_PIR_KEYWORDS	signal	270	8.60E-56	3.90E-53
	UP_SEQ_FEATURE	signal peptide	270	3.10E-55	6.70E-52
	SP_PIR_KEYWORDS	extracellular matrix	77	4.20E-54	9.60E-52
	GOTERM_CC_FAT	extracellular matrix	90	3.40E-50	1.10E-47
	GOTERM_CC_FAT	proteinaceous extracellular matrix	86	5.30E-49	8.50E-47
	GOTERM_CC_FAT	extracellular region part	132	5.90E-41	6.30E-39
	SP_PIR_KEYWORDS	Secreted	166	3.00E-40	4.60E-38
	GOTERM_CC_FAT	extracellular region	191	1.40E-37	1.10E-35
	SP_PIR_KEYWORDS	glycoprotein	272	5.90E-33	6.70E-31
	GOTERM_CC_FAT	extracellular matrix part	42	4.00E-29	2.60E-27
	UP_SEQ_FEATURE	glycosylation site:N-linked (GlcNAc...)	251	8.10E-27	8.70E-24
	SP_PIR_KEYWORDS	disulfide bond	197	1.20E-25	1.10E-23
	UP_SEQ_FEATURE	disulfide bond	192	3.90E-25	2.80E-22
Annotation Cluster 2		Enrichment Score: 29.11	Count	P_Value	Benjamini
	GOTERM_BP_FAT	cell adhesion	98	1.40E-33	3.20E-30
	GOTERM_BP_FAT	biological adhesion	98	1.60E-33	1.80E-30
	SP_PIR_KEYWORDS	cell adhesion	61	2.00E-22	1.50E-20
Annotation Cluster 3		Enrichment Score: 19.86	Count	P_Value	Benjamini
	GOTERM_CC_FAT	extracellular matrix part	42	4.00E-29	2.60E-27
	GOTERM_CC_FAT	basement membrane	27	2.90E-18	1.60E-16
	SP_PIR_KEYWORDS	basement membrane	17	2.20E-14	9.20E-13
Annotation Cluster 4		Enrichment Score: 17.3	Count	P_Value	Benjamini
	GOTERM_BP_FAT	extracellular matrix organization	34	2.50E-23	1.90E-20
	GOTERM_BP_FAT	extracellular structure organization	36	1.50E-18	4.70E-16
	GOTERM_BP_FAT	collagen fibril organization	14	3.40E-12	7.00E-10
Annotation Cluster 5		Enrichment Score: 15.2	Count	P_Value	Benjamini
	GOTERM_BP_FAT	vasculature development	47	4.80E-21	2.70E-18
	GOTERM_BP_FAT	blood vessel development	44	5.30E-19	2.40E-16
	GOTERM_BP_FAT	blood vessel morphogenesis	34	2.70E-13	6.60E-11
	GOTERM_BP_FAT	angiogenesis	25	2.40E-10	2.80E-08
Annotation Cluster 6		Enrichment Score: 11.76	Count	P_Value	Benjamini
	GOTERM_CC_FAT	extracellular matrix part	42	4.00E-29	2.60E-27
	GOTERM_MF_FAT	extracellular matrix structural constituent	28	1.00E-18	5.80E-16
	SP_PIR_KEYWORDS	hydroxylation	24	2.10E-17	1.30E-15
	GOTERM_CC_FAT	collagen	19	4.50E-17	2.10E-15
	SP_PIR_KEYWORDS	collagen	25	2.50E-15	1.30E-13
	INTERPRO	Collagen triple helix repeat	23	4.60E-14	5.40E-12
	SP_PIR_KEYWORDS	trimer	14	2.60E-13	9.90E-12
	SP_PIR_KEYWORDS	triple helix	14	4.50E-12	1.50E-10
	SP_PIR_KEYWORDS	hydroxylysine	14	4.50E-12	1.50E-10
	UP_SEQ_FEATURE	region of interest:Triple-helical region	12	3.80E-11	1.30E-08
	SP_PIR_KEYWORDS	hydroxyproline	14	6.50E-11	2.00E-09
	SP_PIR_KEYWORDS	pyroglutamic acid	9	1.00E-04	1.30E-03
	GOTERM_CC_FAT	anchoring collagen	5	2.20E-04	2.90E-03
	UP_SEQ_FEATURE	domain:VWFA 1	5	1.50E-03	5.70E-02
	UP_SEQ_FEATURE	domain:VWFA 2	5	1.90E-03	6.70E-02
Annotation Cluster 7		Enrichment Score: 11.42	Count	P_Value	Benjamini
	GOTERM_BP_FAT	skeletal system development	50	5.60E-19	2.10E-16
	GOTERM_BP_FAT	bone development	21	7.20E-09	7.10E-07
	GOTERM_BP_FAT	ossification	20	1.30E-08	1.20E-06
Annotation Cluster 8		Enrichment Score: 11.3	Count	P_Value	Benjamini
	GOTERM_BP_FAT	cell motion	55	7.40E-15	2.10E-12
	GOTERM_BP_FAT	cell migration	37	5.60E-12	9.60E-10
	GOTERM_BP_FAT	localization of cell	37	1.20E-10	1.70E-08
	GOTERM_BP_FAT	cell motility	37	1.20E-10	1.70E-08
Annotation Cluster 9		Enrichment Score: 10.46	Count	P_Value	Benjamini
	INTERPRO	EGF-like calcium-binding, conserved site	28	7.10E-18	5.90E-15
	INTERPRO	EGF-like region, conserved site	45	7.40E-17	4.60E-14
	INTERPRO	EGF-like calcium-binding	27	8.10E-17	3.10E-14
	SP_PIR_KEYWORDS	egf-like domain	38	6.20E-16	3.80E-14
	INTERPRO	EGF-type aspartate/asparagine hydroxylation conserved site	25	1.30E-14	2.60E-12
	INTERPRO	EGF calcium-binding	22	1.60E-14	2.70E-12
	INTERPRO	EGF-like, type 3	34	1.80E-14	2.50E-12
	SMART	EGF_CA	27	7.60E-14	1.50E-11
	INTERPRO	EGF	25	6.90E-12	7.20E-10
	INTERPRO	EGF-like	31	9.30E-12	8.60E-10
	UP_SEQ_FEATURE	domain:EGF-like 3; calcium-binding	14	7.00E-11	2.20E-08
	UP_SEQ_FEATURE	domain:EGF-like 2; calcium-binding	16	2.60E-10	6.90E-08
	UP_SEQ_FEATURE	domain:EGF-like 1	22	2.80E-10	6.80E-08
	UP_SEQ_FEATURE	domain:EGF-like 5; calcium-binding	13	1.10E-09	2.40E-07
	UP_SEQ_FEATURE	domain:EGF-like 4; calcium-binding	12	4.00E-09	7.80E-07
	SMART	EGF	31	8.40E-09	8.40E-07
	UP_SEQ_FEATURE	domain:EGF-like 6; calcium-binding	9	8.00E-07	1.30E-04
	UP_SEQ_FEATURE	domain:EGF-like 4	11	1.60E-05	1.80E-03
	UP_SEQ_FEATURE	domain:EGF-like 2	13	2.70E-05	2.70E-03
	UP_SEQ_FEATURE	domain:EGF-like 3	12	2.90E-05	2.70E-03
	UP_SEQ_FEATURE	domain:EGF-like 7; calcium-binding	7	1.90E-04	1.10E-02
Annotation Cluster 10		Enrichment Score: 9.06	Count	P_Value	Benjamini
	UP_SEQ_FEATURE	domain:VWFC	12	1.00E-12	4.40E-10
	INTERPRO	von Willebrand factor, type C	13	5.10E-09	3.90E-07
	SMART	VWC	13	1.30E-07	6.40E-06

Add. Table 8: Gene set enrichment analysis (GSEA) result for Gene Ontology (GO) gene sets. Metabolic cluster Mc1 was compared with Mc2 and Mc3.

NAME	SIZE	ES	NES	NOM p-val	**FDR q-val**	FWER p-val	RANK AT MAX	LEADING EDGE
COLLAGEN	23	0.7818	2.0623	0.0011	0.0020	0.0070	2064	tags=70%, list=9%, signal=77%
EXTRACELLULAR_MATRIX	96	0.6880	2.0895	0.0000	0.0020	0.0050	2225	tags=49%, list=10%, signal=54%
PROTEINACEOUS_EXTRACELLULAR_MATRIX	94	0.6915	2.1008	0.0000	0.0025	0.0040	2225	tags=49%, list=10%, signal=54%
EXTRACELLULAR_MATRIX_PART	55	0.7282	2.1274	0.0000	0.0040	0.0040	2129	tags=51%, list=10%, signal=56%
INTEGRIN_BINDING	30	0.7220	1.9686	0.0000	0.0080	0.0270	3129	tags=63%, list=14%, signal=74%
CELL_SUBSTRATE_ADHESION	37	0.6891	1.9336	0.0000	0.0109	0.0420	3380	tags=57%, list=15%, signal=67%
EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT	26	0.6938	1.9012	0.0000	0.0110	0.0550	1226	tags=46%, list=6%, signal=49%
CELL_MATRIX_ADHESION	36	0.6843	1.9120	0.0000	0.0112	0.0470	3380	tags=56%, list=15%, signal=66%
CELL_MATRIX_JUNCTION	16	0.7306	1.8324	0.0000	0.0257	0.1290	2484	tags=56%, list=11%, signal=63%
BASEMENT_MEMBRANE	35	0.6546	1.8215	0.0010	0.0268	0.1380	2684	tags=43%, list=12%, signal=49%
TRANSMEMBRANE_RECEPTOR_PROTEIN_KINASE_ACTIVITY	51	0.6166	1.8115	0.0000	0.0276	0.1540	3983	tags=49%, list=18%, signal=60%
AXON_GUIDANCE	22	0.6824	1.7839	0.0000	0.0323	0.1990	933	tags=27%, list=4%, signal=28%
BASAL_LAMINA	19	0.6938	1.7897	0.0020	0.0330	0.1900	2684	tags=42%, list=12%, signal=48%
SKELETAL_DEVELOPMENT	99	0.5846	1.7745	0.0000	0.0338	0.2130	2762	tags=37%, list=13%, signal=43%
REGULATION_OF_CELL_MIGRATION	27	0.6465	1.7558	0.0000	0.0393	0.2490	2952	tags=44%, list=14%, signal=51%
BASOLATERAL_PLASMA_MEMBRANE	31	0.6225	1.7124	0.0000	0.0539	0.3680	2798	tags=42%, list=13%, signal=48%
BLOOD_COAGULATION	42	0.6010	1.7209	0.0000	0.0543	0.3420	3903	tags=48%, list=18%, signal=58%
ADHERENS_JUNCTION	21	0.6590	1.7162	0.0000	0.0549	0.3610	2484	tags=48%, list=11%, signal=54%
HEMOSTASIS	47	0.5966	1.7230	0.0000	0.0556	0.3340	3903	tags=47%, list=18%, signal=57%
COAGULATION	43	0.5950	1.7030	0.0000	0.0570	0.4060	3903	tags=47%, list=18%, signal=57%
SULFURIC_ESTER_HYDROLASE_ACTIVITY	16	0.6727	1.6963	0.0010	0.0584	0.4200	3141	tags=50%, list=14%, signal=58%
LIPID_HOMEOSTASIS	16	0.6690	1.6824	0.0010	0.0611	0.4790	3348	tags=44%, list=15%, signal=52%
ANATOMICAL_STRUCTURE_FORMATION	52	0.5763	1.6844	0.0000	0.0618	0.4690	5237	tags=52%, list=24%, signal=68%
WOUND_HEALING	53	0.5772	1.6865	0.0000	0.0623	0.4610	4715	tags=51%, list=22%, signal=65%
BONE_REMODELING	28	0.6094	1.6632	0.0010	0.0668	0.5410	3983	tags=39%, list=18%, signal=48%
TISSUE_REMODELING	29	0.6081	1.6635	0.0000	0.0691	0.5410	3983	tags=38%, list=18%, signal=46%
EXTRACELLULAR_REGION_PART	322	0.5296	1.6665	0.0000	0.0696	0.5260	4482	tags=39%, list=21%, signal=49%
REGULATION_OF_BODY_FLUID_LEVELS	56	0.5625	1.6467	0.0000	0.0728	0.6060	4183	tags=45%, list=19%, signal=55%
TRANSMEMBRANE_RECEPTOR_PROTEIN_TYROSINE_KINASE_ACTIVITY	43	0.5705	1.6498	0.0000	0.0750	0.5920	3856	tags=42%, list=18%, signal=51%
POSITIVE_REGULATION_OF_CELL_DIFFERENTIATION	23	0.6161	1.6470	0.0000	0.0751	0.6060	2762	tags=43%, list=13%, signal=50%
GLYCOSAMINOGLYCAN_BINDING	30	0.5890	1.6322	0.0040	0.0806	0.6740	3108	tags=40%, list=14%, signal=47%
HOMEOSTASIS_OF_NUMBER_OF_CELLS	19	0.6164	1.6235	0.0071	0.0826	0.7060	1949	tags=37%, list=9%, signal=40%
POLYSACCHARIDE_BINDING	32	0.5863	1.6326	0.0020	0.0830	0.6730	3108	tags=38%, list=14%, signal=44%
CELL_MIGRATION	92	0.5447	1.6261	0.0000	0.0830	0.6990	2952	tags=35%, list=14%, signal=40%
MUSCLE_DEVELOPMENT	89	0.5369	1.6113	0.0000	0.0831	0.7460	5238	tags=46%, list=24%, signal=60%
RUFFLE	27	0.5885	1.6170	0.0000	0.0839	0.7260	3309	tags=37%, list=15%, signal=44%
TISSUE_DEVELOPMENT	135	0.5301	1.6127	0.0000	0.0840	0.7360	5011	tags=42%, list=23%, signal=54%
ATP_DEPENDENT_HELICASE_ACTIVITY	22	0.6085	1.6190	0.0050	0.0844	0.7200	6660	tags=68%, list=30%, signal=98%
REGULATION_OF_CYTOSKELETON_ORGANIZATION_AND_BIOGENESIS	28	0.5915	1.6140	0.0010	0.0849	0.7330	2952	tags=36%, list=14%, signal=41%
HEPARIN_BINDING	22	0.5979	1.6041	0.0020	0.0870	0.7710	3108	tags=41%, list=14%, signal=48%
ACTIN_BINDING	72	0.5291	1.5901	0.0000	0.0880	0.8240	4089	tags=38%, list=19%, signal=46%
CELL_CORTEX_PART	23	0.5909	1.5937	0.0030	0.0885	0.8110	3160	tags=35%, list=14%, signal=41%
REGULATION_OF_ORGANELLE_ORGANIZATION_AND_BIOGENESIS	37	0.5654	1.5914	0.0010	0.0886	0.8210	2952	tags=32%, list=14%, signal=37%
PROTEIN_COMPLEX_BINDING	54	0.5464	1.5850	0.0010	0.0887	0.8400	3232	tags=39%, list=15%, signal=46%
REGULATION_OF_RESPONSE_TO_EXTERNAL_STIMULUS	15	0.6496	1.5822	0.0092	0.0895	0.8460	3525	tags=53%, list=16%, signal=64%
ANGIOGENESIS	44	0.5509	1.5856	0.0010	0.0899	0.8380	5625	tags=52%, list=26%, signal=70%
EXTRACELLULAR_REGION	423	0.5050	1.5985	0.0000	0.0901	0.7930	4482	tags=37%, list=21%, signal=45%
MUSCLE_CELL_DIFFERENTIATION	20	0.6121	1.5939	0.0061	0.0904	0.8090	5130	tags=60%, list=23%, signal=78%
INTEGRIN_COMPLEX	19	0.6200	1.5960	0.0040	0.0908	0.8040	4500	tags=58%, list=21%, signal=73%
ACTIN_POLYMERIZATION_AND_OR_DEPOLYMERIZATION	23	0.5846	1.5773	0.0010	0.0924	0.8600	2848	tags=35%, list=13%, signal=40%
NEURON_PROJECTION	19	0.6090	1.5706	0.0060	0.0933	0.8780	5094	tags=58%, list=23%, signal=75%
CALMODULIN_BINDING	25	0.5821	1.5712	0.0030	0.0944	0.8760	7178	tags=64%, list=33%, signal=95%
SODIUM_CHANNEL_ACTIVITY	16	0.6191	1.5722	0.0061	0.0950	0.8740	1909	tags=38%, list=9%, signal=41%
VASCULATURE_DEVELOPMENT	51	0.5379	1.5645	0.0010	0.0975	0.9000	5188	tags=49%, list=24%, signal=64%
KINASE_INHIBITOR_ACTIVITY	25	0.5854	1.5627	0.0040	0.0978	0.9010	3734	tags=40%, list=17%, signal=48%
ECTODERM_DEVELOPMENT	79	0.5307	1.5594	0.0050	0.0990	0.9090	5011	tags=42%, list=23%, signal=54%
PROTEIN_KINASE_INHIBITOR_ACTIVITY	24	0.5865	1.5561	0.0050	0.1004	0.9140	3734	tags=42%, list=17%, signal=50%
ENERGY_RESERVE_METABOLIC_PROCESS	15	0.6309	1.5537	0.0082	0.1009	0.9160	5542	tags=60%, list=25%, signal=80%
ACTIN_FILAMENT_BINDING	23	0.5747	1.5501	0.0030	0.1009	0.9250	4089	tags=48%, list=19%, signal=59%
NEGATIVE_REGULATION_OF_DNA_BINDING	16	0.6063	1.5511	0.0112	0.1018	0.9240	4951	tags=56%, list=23%, signal=73%
EXTRACELLULAR_STRUCTURE_ORGANIZATION_AND_BIOGENESIS	29	0.5641	1.5443	0.0261	0.1046	0.9410	2999	tags=41%, list=14%, signal=48%
ACTIN_CYTOSKELETON	122	0.4982	1.5355	0.0000	0.1114	0.9560	3302	tags=30%, list=15%, signal=35%
CARBOHYDRATE_BIOSYNTHETIC_PROCESS	45	0.5400	1.5369	0.0030	0.1115	0.9540	5302	tags=44%, list=24%, signal=59%
ORGAN_MORPHOGENESIS	136	0.4979	1.5320	0.0000	0.1137	0.9580	6037	tags=48%, list=28%, signal=66%
METALLOENDOPEPTIDASE_ACTIVITY	27	0.5639	1.5297	0.0131	0.1144	0.9600	5134	tags=44%, list=23%, signal=58%
STRUCTURAL_CONSTITUENT_OF_CYTOSKELETON	55	0.5283	1.5242	0.0030	0.1187	0.9670	5240	tags=44%, list=24%, signal=57%
EPIDERMIS_DEVELOPMENT	70	0.5233	1.5215	0.0060	0.1205	0.9680	5011	tags=40%, list=23%, signal=52%
TRANSMEMBRANE_RECEPTOR_PROTEIN_SERINE_THREONINE_KINASE_SIGNALING_PATHWAY	47	0.5270	1.5184	0.0020	0.1226	0.9710	5843	tags=49%, list=27%, signal=67%
SERINE_HYDROLASE_ACTIVITY	44	0.5325	1.5073	0.0060	0.1352	0.9810	2788	tags=30%, list=13%, signal=34%
SULFUR_COMPOUND_BIOSYNTHETIC_PROCESS	16	0.5972	1.4970	0.0215	0.1419	0.9870	2772	tags=31%, list=13%, signal=36%
PHOSPHORIC_DIESTER_HYDROLASE_ACTIVITY	39	0.5283	1.4978	0.0020	0.1428	0.9870	4367	tags=41%, list=20%, signal=51%
CORTICAL_CYTOSKELETON	19	0.5744	1.4985	0.0071	0.1439	0.9850	3160	tags=32%, list=14%, signal=37%
NEGATIVE_REGULATION_OF_BINDING	17	0.5775	1.4941	0.0153	0.1441	0.9880	4951	tags=53%, list=23%, signal=68%
METALLOPEPTIDASE_ACTIVITY	46	0.5251	1.4996	0.0010	0.1444	0.9850	5789	tags=46%, list=26%, signal=62%
MYOBLAST_DIFFERENTIATION	17	0.5818	1.4862	0.0295	0.1479	0.9900	3890	tags=47%, list=18%, signal=57%
GROWTH_FACTOR_BINDING	32	0.5301	1.4874	0.0220	0.1480	0.9900	5197	tags=50%, list=24%, signal=66%
DENDRITE	16	0.5891	1.4885	0.0163	0.1486	0.9900	4156	tags=44%, list=19%, signal=54%
GENERATION_OF_A_SIGNAL_INVOLVED_IN_CELL_CELL_SIGNALING	27	0.5619	1.4894	0.0151	0.1490	0.9900	5004	tags=44%, list=23%, signal=58%
REGULATION_OF_INTRACELLULAR_TRANSPORT	22	0.5552	1.4797	0.0190	0.1556	0.9910	4825	tags=41%, list=22%, signal=52%
RAS_GTPASE_ACTIVATOR_ACTIVITY	26	0.5449	1.4743	0.0130	0.1628	0.9920	6320	tags=46%, list=29%, signal=65%
RECEPTOR_COMPLEX	56	0.5100	1.4702	0.0040	0.1630	0.9940	4969	tags=45%, list=23%, signal=58%
SERINE_TYPE_PEPTIDASE_ACTIVITY	43	0.5200	1.4686	0.0100	0.1635	0.9940	2788	tags=28%, list=13%, signal=32%
STRUCTURAL_CONSTITUENT_OF_MUSCLE	32	0.5360	1.4709	0.0090	0.1639	0.9930	5250	tags=44%, list=24%, signal=58%
PROTEIN_PROCESSING	45	0.5146	1.4720	0.0020	0.1640	0.9930	4954	tags=36%, list=23%, signal=46%
ENDOPEPTIDASE_ACTIVITY	111	0.4832	1.4645	0.0000	0.1681	0.9940	5238	tags=35%, list=24%, signal=46%
HELICASE_ACTIVITY	45	0.5127	1.4631	0.0070	0.1687	0.9950	8005	tags=62%, list=37%, signal=98%
PROTEASE_INHIBITOR_ACTIVITY	40	0.5187	1.4584	0.0050	0.1723	0.9960	681	tags=18%, list=3%, signal=18%
REGULATION_OF_MYELOID_CELL_DIFFERENTIATION	18	0.5734	1.4591	0.0234	0.1731	0.9950	1949	tags=33%, list=9%, signal=37%
PATTERN_BINDING	38	0.5131	1.4517	0.0110	0.1774	0.9970	4908	tags=42%, list=22%, signal=54%
PEPTIDASE_ACTIVITY	164	0.4698	1.4527	0.0000	0.1777	0.9970	5789	tags=37%, list=26%, signal=49%
REGULATION_OF_SECRETION	33	0.5276	1.4532	0.0120	0.1789	0.9960	5004	tags=42%, list=23%, signal=55%
AMINE_BIOSYNTHETIC_PROCESS	15	0.5867	1.4486	0.0346	0.1810	0.9970	8286	tags=80%, list=38%, signal=129%

NAME	SIZE	ES	NES	NOM p-val	**FDR q-val**	FWER p-val	RANK AT MAX	LEADING EDGE
OXIDOREDUCTASE_ACTIVITY_ACTING_ON_NADH_OR_NADPH	25	0.5327	1.4416	0.0221	0.1921	0.9990	6672	tags=64%, list=31%, signal=92%
NEGATIVE_REGULATION_OF_CELLULAR_COMPONENT_ORGANIZATION_AND_BIOGENESIS	26	0.5243	1.4385	0.0190	0.1958	0.9990	2745	tags=31%, list=13%, signal=35%
POSITIVE_REGULATION_OF_TRANSCRIPTION_FROM_RNA_POLYMERASE_II_PROMOTER	57	0.4909	1.4355	0.0040	0.1998	1.0000	4266	tags=35%, list=20%, signal=43%
TRANSPORT_VESICLE	29	0.5150	1.4181	0.0180	0.2030	1.0000	4937	tags=38%, list=23%, signal=49%
REGULATION_OF_NUCLEOCYTOPLASMIC_TRANSPORT	19	0.5413	1.4151	0.0322	0.2031	1.0000	4825	tags=37%, list=22%, signal=47%
DNA_HELICASE_ACTIVITY	22	0.5435	1.4325	0.0332	0.2036	1.0000	6302	tags=55%, list=29%, signal=77%
ELECTRON_CARRIER_ACTIVITY	75	0.4780	1.4185	0.0020	0.2040	1.0000	5315	tags=44%, list=24%, signal=58%
NEGATIVE_REGULATION_OF_TRANSCRIPTION_FROM_RNA_POLYMERASE_II_PROMOTER	78	0.4714	1.4165	0.0000	0.2042	1.0000	7409	tags=46%, list=34%, signal=70%
KERATINOCYTE_DIFFERENTIATION	15	0.5781	1.4152	0.0462	0.2048	1.0000	6886	tags=53%, list=32%, signal=78%
TUBE_DEVELOPMENT	18	0.5542	1.4104	0.0445	0.2050	1.0000	1962	tags=33%, list=9%, signal=37%
COPPER_ION_BINDING	15	0.5645	1.4189	0.0324	0.2052	1.0000	1565	tags=27%, list=7%, signal=29%
NUCLEOTIDE_BIOSYNTHETIC_PROCESS	19	0.5430	1.4090	0.0456	0.2056	1.0000	4828	tags=42%, list=22%, signal=54%
MEMBRANE_FUSION	27	0.5163	1.4128	0.0320	0.2058	1.0000	3927	tags=37%, list=18%, signal=45%
TRANSFORMING_GROWTH_FACTOR_BETA_RECEPTOR_SIGNALING_PATHWAY	36	0.5086	1.4288	0.0130	0.2063	1.0000	7420	tags=56%, list=34%, signal=84%
PROTEIN_BINDING_BRIDGING	59	0.4886	1.4106	0.0070	0.2065	1.0000	7526	tags=46%, list=34%, signal=70%
REGULATION_OF_CELL_DIFFERENTIATION	55	0.4901	1.4189	0.0040	0.2070	1.0000	2911	tags=29%, list=13%, signal=33%
PROTEIN_DNA_COMPLEX_ASSEMBLY	37	0.4978	1.4111	0.0171	0.2071	1.0000	8969	tags=68%, list=41%, signal=114%
CONTRACTILE_FIBER_PART	23	0.5362	1.4196	0.0332	0.2076	1.0000	1468	tags=26%, list=7%, signal=28%
SERINE_TYPE_ENDOPEPTIDASE_ACTIVITY	39	0.5094	1.4268	0.0190	0.2080	1.0000	2788	tags=28%, list=13%, signal=32%
RESPONSE_TO_HYPOXIA	27	0.5217	1.4289	0.0100	0.2083	1.0000	7014	tags=59%, list=32%, signal=87%
AXONOGENESIS	43	0.5047	1.4210	0.0070	0.2087	1.0000	3671	tags=30%, list=17%, signal=36%
ACTIN_FILAMENT	17	0.5527	1.4200	0.0515	0.2088	1.0000	4742	tags=35%, list=22%, signal=45%
RNA_DEPENDENT_ATPASE_ACTIVITY	15	0.5695	1.4246	0.0483	0.2099	1.0000	7365	tags=73%, list=34%, signal=111%
ENZYME_INHIBITOR_ACTIVITY	115	0.4685	1.4211	0.0000	0.2107	1.0000	4713	tags=34%, list=22%, signal=43%
REGULATION_OF_ANATOMICAL_STRUCTURE_MORPHOGENESIS	23	0.5341	1.4227	0.0272	0.2119	1.0000	3035	tags=35%, list=14%, signal=40%
CELL_CELL_ADHESION	81	0.4783	1.4215	0.0020	0.2120	1.0000	6050	tags=44%, list=28%, signal=61%
CELL_PROJECTION_PART	18	0.5419	1.3966	0.0406	0.2174	1.0000	6415	tags=56%, list=29%, signal=79%
POSITIVE_REGULATION_OF_RNA_METABOLIC_PROCESS	102	0.4592	1.3983	0.0010	0.2175	1.0000	4266	tags=32%, list=20%, signal=40%
ENZYME_LINKED_RECEPTOR_PROTEIN_SIGNALING_PATHWAY	135	0.4501	1.3957	0.0000	0.2175	1.0000	4274	tags=34%, list=20%, signal=42%
CELL_JUNCTION	76	0.4747	1.3998	0.0050	0.2183	1.0000	6471	tags=47%, list=30%, signal=67%
NEGATIVE_REGULATION_OF_TRANSCRIPTION	167	0.4474	1.4007	0.0000	0.2183	1.0000	7227	tags=43%, list=33%, signal=64%
BIOGENIC_AMINE_METABOLIC_PROCESS	16	0.5563	1.4015	0.0427	0.2184	1.0000	4828	tags=44%, list=22%, signal=56%
CELLULAR_MORPHOGENESIS_DURING_DIFFERENTIATION	49	0.4909	1.3987	0.0090	0.2186	1.0000	3035	tags=27%, list=14%, signal=31%
LEADING_EDGE	42	0.4842	1.3966	0.0190	0.2191	1.0000	3309	tags=29%, list=15%, signal=34%
PHOSPHOLIPID_BINDING	39	0.4973	1.3925	0.0240	0.2224	1.0000	4460	tags=38%, list=20%, signal=48%
POSITIVE_REGULATION_OF_TRANSCRIPTIONDNA_DEPENDENT	100	0.4568	1.3902	0.0010	0.2253	1.0000	4266	tags=32%, list=20%, signal=40%
ANATOMICAL_STRUCTURE_MORPHOGENESIS	356	0.4361	1.3832	0.0000	0.2258	1.0000	5963	tags=39%, list=27%, signal=53%
CELL_SURFACE	72	0.4643	1.3855	0.0040	0.2264	1.0000	7241	tags=49%, list=33%, signal=72%
RESPONSE_TO_WOUNDING	178	0.4479	1.3821	0.0080	0.2264	1.0000	6311	tags=41%, list=29%, signal=57%
RECEPTOR_SIGNALING_PROTEIN_SERINE_THREONINE_KINASE_ACTIVITY	34	0.4972	1.3887	0.0150	0.2267	1.0000	5214	tags=41%, list=24%, signal=54%
ACTIN_CYTOSKELETON_ORGANIZATION_AND_BIOGENESIS	101	0.4543	1.3869	0.0010	0.2268	1.0000	3500	tags=30%, list=16%, signal=35%
POSITIVE_REGULATION_OF_JNK_ACTIVITY	18	0.5382	1.3834	0.0455	0.2271	1.0000	4705	tags=44%, list=22%, signal=57%
POSITIVE_REGULATION_OF_TRANSCRIPTION	120	0.4520	1.3875	0.0010	0.2273	1.0000	6739	tags=43%, list=31%, signal=62%
TUBE_MORPHOGENESIS	15	0.5559	1.3856	0.0462	0.2278	1.0000	1962	tags=33%, list=9%, signal=37%
REGULATION_OF_TRANSCRIPTION_FACTOR_ACTIVITY	36	0.4904	1.3838	0.0200	0.2278	1.0000	4951	tags=39%, list=23%, signal=50%
CELL_PROJECTION	98	0.4512	1.3749	0.0040	0.2311	1.0000	2978	tags=27%, list=14%, signal=31%
GTPASE_ACTIVATOR_ACTIVITY	56	0.4723	1.3791	0.0100	0.2311	1.0000	6484	tags=45%, list=30%, signal=63%
DNA_DEPENDENT_ATPASE_ACTIVITY	19	0.5307	1.3755	0.0676	0.2316	1.0000	4923	tags=47%, list=23%, signal=61%
EXTRACELLULAR_SPACE	233	0.4436	1.3766	0.0000	0.2326	1.0000	3568	tags=28%, list=16%, signal=33%
NEGATIVE_REGULATION_OF_MULTICELLULAR_ORGANISMAL_PROCESS	30	0.5084	1.3773	0.0311	0.2329	1.0000	4878	tags=40%, list=22%, signal=51%
REGULATION_OF_ANGIOGENESIS	25	0.5071	1.3756	0.0302	0.2330	1.0000	5625	tags=44%, list=26%, signal=59%
NEGATIVE_REGULATION_OF_CELL_DIFFERENTIATION	27	0.5095	1.3679	0.0340	0.2434	1.0000	1501	tags=26%, list=7%, signal=28%
REGULATION_OF_NEUROTRANSMITTER_LEVELS	22	0.5257	1.3657	0.0493	0.2463	1.0000	4964	tags=45%, list=23%, signal=59%
TRANSCRIPTION_INITIATION_FROM_RNA_POLYMERASE_II_PROMOTER	18	0.5264	1.3646	0.0487	0.2470	1.0000	5685	tags=56%, list=26%, signal=75%

Add. Table 9: Gene set enrichment analysis (GSEA) result for Gene Ontology (GO) gene sets. Metabolic cluster Mc1 was compared with Mc2.

NAME	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE
INTEGRIN_BINDING	30	0.7361	2.0381	0.0000	0.0110	0.0110	2319	tags=60%, list=11%, signal=67%
CELL_MATRIX_JUNCTION	16	0.7521	1.9300	0.0000	0.0360	0.0460	1814	tags=56%, list=8%, signal=61%
ADHERENS_JUNCTION	21	0.6979	1.8518	0.0000	0.0620	0.1030	1814	tags=48%, list=8%, signal=52%
PROTEINACEOUS_EXTRACELLULAR_MATRIX	94	0.5694	1.7481	0.0050	0.0954	0.2720	5379	tags=55%, list=25%, signal=73%
EXTRACELLULAR_MATRIX	96	0.5696	1.7493	0.0050	0.1095	0.2680	3162	tags=45%, list=14%, signal=52%
COLLAGEN	23	0.6382	1.7129	0.0384	0.1243	0.3630	6011	tags=78%, list=28%, signal=108%
TRANSMEMBRANE_RECEPTOR_PROTEIN_KINASE_ACTIVITY	51	0.5924	1.7532	0.0010	0.1253	0.2550	3358	tags=45%, list=15%, signal=53%
EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT	26	0.6175	1.7028	0.0051	0.1254	0.3980	2994	tags=50%, list=14%, signal=58%
EXTRACELLULAR_MATRIX_PART	55	0.5969	1.7696	0.0130	0.1265	0.2220	5379	tags=58%, list=25%, signal=77%
GROWTH_FACTOR_BINDING	32	0.6009	1.6675	0.0010	0.1309	0.5230	5615	tags=66%, list=26%, signal=88%
CELL_SUBSTRATE_ADHESION	37	0.5911	1.6695	0.0000	0.1383	0.5160	3072	tags=46%, list=14%, signal=53%
LIPID_HOMEOSTASIS	16	0.6659	1.6505	0.0020	0.1391	0.5820	2598	tags=44%, list=12%, signal=50%
SERINE_TYPE_ENDOPEPTIDASE_ACTIVITY	39	0.5827	1.6553	0.0000	0.1393	0.5630	3056	tags=36%, list=14%, signal=42%
BASEMENT_MEMBRANE	35	0.5880	1.6724	0.0060	0.1458	0.4990	5379	tags=54%, list=25%, signal=72%
NEGATIVE_REGULATION_OF_DNA_BINDING	16	0.6452	1.6274	0.0031	0.1503	0.6830	3270	tags=44%, list=15%, signal=51%
BASOLATERAL_PLASMA_MEMBRANE	31	0.5887	1.6307	0.0000	0.1532	0.6660	1814	tags=35%, list=8%, signal=39%
SERINE_HYDROLASE_ACTIVITY	44	0.5650	1.6203	0.0000	0.1541	0.7100	3056	tags=34%, list=14%, signal=40%
PROTEIN_COMPLEX_BINDING	54	0.5525	1.6153	0.0010	0.1541	0.7320	3878	tags=43%, list=18%, signal=52%
SERINE_TYPE_PEPTIDASE_ACTIVITY	43	0.5853	1.6741	0.0000	0.1565	0.4930	3056	tags=35%, list=14%, signal=40%
HEMOSTASIS	47	0.5537	1.6083	0.0000	0.1582	0.7580	2123	tags=36%, list=10%, signal=40%
CELL_MATRIX_ADHESION	36	0.5787	1.6323	0.0000	0.1606	0.6600	3072	tags=44%, list=14%, signal=52%
EPIDERMIS_DEVELOPMENT	70	0.5423	1.5873	0.0010	0.1666	0.8300	3565	tags=37%, list=16%, signal=44%
NEURON_PROJECTION	19	0.6006	1.5887	0.0030	0.1711	0.8270	2527	tags=42%, list=12%, signal=48%
TRANSMEMBRANE_RECEPTOR_PROTEIN_TYROSINE_KINASE_ACTIVITY	43	0.5487	1.5957	0.0040	0.1735	0.8050	3358	tags=40%, list=15%, signal=47%
REGULATION_OF_MAPKKK_CASCADE	19	0.6113	1.5904	0.0091	0.1757	0.8230	3372	tags=42%, list=15%, signal=50%
NEGATIVE_REGULATION_OF_BINDING	17	0.6147	1.5664	0.0061	0.1798	0.8970	3270	tags=41%, list=15%, signal=48%
MUSCLE_DEVELOPMENT	89	0.5286	1.5750	0.0000	0.1826	0.8690	4652	tags=47%, list=21%, signal=60%
COAGULATION	43	0.5457	1.5666	0.0010	0.1857	0.8950	2123	tags=35%, list=10%, signal=39%
BLOOD_COAGULATION	42	0.5462	1.5666	0.0010	0.1924	0.8950	2123	tags=36%, list=10%, signal=39%
ECTODERM_DEVELOPMENT	79	0.5244	1.5467	0.0020	0.2059	0.9310	3173	tags=34%, list=15%, signal=40%
STRUCTURAL_CONSTITUENT_OF_CYTOSKELETON	55	0.5388	1.5494	0.0040	0.2068	0.9260	5410	tags=45%, list=25%, signal=60%
ACTIVATION_OF_PROTEIN_KINASE_ACTIVITY	24	0.5706	1.5224	0.0071	0.2084	0.9700	2429	tags=38%, list=11%, signal=42%
CELL_MIGRATION	92	0.5099	1.5270	0.0010	0.2102	0.9640	2625	tags=34%, list=12%, signal=38%
KINASE_INHIBITOR_ACTIVITY	25	0.5648	1.5227	0.0020	0.2134	0.9700	5182	tags=44%, list=24%, signal=58%
SKELETAL_DEVELOPMENT	99	0.4977	1.5271	0.0080	0.2156	0.9640	4851	tags=42%, list=22%, signal=54%
REGULATION_OF_BODY_FLUID_LEVELS	56	0.5194	1.5323	0.0010	0.2168	0.9560	2123	tags=34%, list=10%, signal=37%
PROTEIN_KINASE_INHIBITOR_ACTIVITY	24	0.5736	1.5351	0.0030	0.2175	0.9530	5182	tags=46%, list=24%, signal=60%
TRANSFORMING_GROWTH_FACTOR_BETA_RECEPTOR_SIGNALING_PATHWAY	36	0.5409	1.5151	0.0060	0.2186	0.9780	7287	tags=58%, list=33%, signal=87%
TRANSMEMBRANE_RECEPTOR_PROTEIN_SERINE_THREONINE_KINASE_SIGNALING_PATHWAY	47	0.5338	1.5276	0.0010	0.2205	0.9620	7287	tags=57%, list=33%, signal=86%
STRUCTURAL_CONSTITUENT_OF_MUSCLE	32	0.5553	1.5352	0.0040	0.2242	0.9530	2760	tags=41%, list=13%, signal=46%
EXTRACELLULAR_REGION_PART	322	0.4756	1.5086	0.0030	0.2281	0.9810	6443	tags=46%, list=29%, signal=64%
WOUND_HEALING	53	0.5178	1.5040	0.0020	0.2328	0.9810	2935	tags=38%, list=13%, signal=43%
EXTRACELLULAR_STRUCTURE_ORGANIZATION_AND_BIOGENESIS	29	0.5388	1.5001	0.0313	0.2359	0.9840	4817	tags=48%, list=22%, signal=62%
LIPID_TRANSPORT	28	0.5511	1.4774	0.0080	0.2488	0.9940	7203	tags=54%, list=33%, signal=80%

Add. Table 10: Gene set enrichment analysis (GSEA) result for Gene Ontology (GO) gene sets. Metabolic cluster Mc3 was compared with Mc1 and Mc2.

NAME	SIZE	ES	NES	NOM p-val	**FDR q-val**	FWER p-val	RANK AT MAX	LEADING EDGE
EXTRACELLULAR_MATRIX_PART	55	0.7515	2.1854	0.0000	0.0020	0.0040	1682	tags=47%, list=8%, signal=51%
COLLAGEN	23	0.8417	2.2165	0.0000	0.0040	0.0040	1682	tags=70%, list=8%, signal=75%
EXTRACELLULAR_MATRIX	96	0.6787	2.0428	0.0000	0.0055	0.0170	1682	tags=43%, list=8%, signal=46%
PROTEINACEOUS_EXTRACELLULAR_MATRIX	94	0.6836	2.0553	0.0000	0.0063	0.0160	1682	tags=43%, list=8%, signal=46%
INTEGRIN_BINDING	30	0.7168	1.9587	0.0000	0.0098	0.0330	2692	tags=50%, list=12%, signal=57%
EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT	26	0.7268	1.9400	0.0000	0.0113	0.0440	1997	tags=54%, list=9%, signal=59%
BASEMENT_MEMBRANE	35	0.6598	1.8425	0.0000	0.0345	0.1170	703	tags=31%, list=3%, signal=32%
BASAL_LAMINA	19	0.6669	1.7255	0.0000	0.1232	0.3500	703	tags=32%, list=3%, signal=33%
CELL_MATRIX_ADHESION	36	0.5977	1.6720	0.0000	0.1718	0.5340	2695	tags=39%, list=12%, signal=44%
CELL_SUBSTRATE_ADHESION	37	0.5963	1.6740	0.0010	0.1877	0.5230	2695	tags=38%, list=12%, signal=43%
PROTEASE_INHIBITOR_ACTIVITY	40	0.5840	1.6425	0.0000	0.2111	0.6390	1497	tags=20%, list=7%, signal=21%

Add. Table 11: Integrated pathway analysis result from comparison of the metabolic clusters Mc1 compared to Mc2

Pathway	Total	Expected	Hits	P.Value	Metabolite hits	Gene hits
Tyrosine metabolism	80	1.9872	8	0.00066125	L-Tyrosine	*AOC3, ALDH1A3, AOX1, ADH1C, MAOA, MAOB, PNMT*
Glycerolipid metabolism	72	1.7885	7	0.0017364		*AKR1B1, AKR1B10, ALDH2, AGPAT4, PPAP2A, PPAP2B, PNLIPRP3*
Fatty acid metabolism	83	2.0617	7	0.003944		*ACSL4, ACSL1, ACAA2, ADH1A, ADH1B, ADH1C, ALDH2*
Glycolysis / Gluconeogenesis	91	2.2604	7	0.0065727	Beta-D-glucose, Acetate	*ADH1A, ADH1B, ADH1C, ALDH2, ALDH1A3*
Ether lipid metabolism	51	1.2668	5	0.0078844		*LPCAT2, ENPP2, PLD1, PPAP2A, PPAP2B*
Glycerophospholipid metabolism	119	2.956	8	0.0082917	Glycerophosphocholine, Phosphocholine	*LCAT, LPCAT2, PPAP2A, PPAP2B, PLD1, AGPAT4*
Arachidonic acid metabolism	100	2.484	7	0.010899		*CYP2U1, ALOX5, GPX3, GGT5, PTGS2, PTGIS, AKR1C3*
Nicotinate and nicotinamide metabolism	39	0.96876	4	0.014869		*AOX1, NNMT, NAMPT, NT5E*
D-Glutamine and D-glutamate metabolism	9	0.22356	2	0.01955	Glutamate	*GLS*
Phenylalanine metabolism	29	0.72036	3	0.033755	L-Tyrosine	*AOC3, ALDH1A3*
Drug metabolism - cytochrome P450	127	3.1547	7	0.035916		*AOX1, ALDH1A3, ADH1A, ADH1B, ADH1C, GSTP1, FMO2*
Mucin type O-Glycan biosynthesis	32	0.79488	3	0.043479		*ST3GAL2, GALNTL2, GALNT12*

Add. Table 12: Integrated pathway analysis result from comparison of the metabolic clusters Mc1 compared to Mc3

Pathway	Total	Expected	Hits	P.Value	Metabolite hits	Gene hits
Glycerophospholipid metabolism	119	2.5977	9	0.0009	Glycerophosphocholine, Phosphocholine	*LCAT, PLA2G5, LPCAT2, PPAP2A, PPAP2B, PLD1, AGPAT4*
Ether lipid metabolism	51	1.1133	5	0.0045		*PPAP2A, PPAP2B, PLA2G5, PLD1, LPCAT2*
D-Glutamine and D-glutamate metabolism	9	0.19646	2	0.0153	Glutamate	*GLS*
Glycosaminoglycan biosynthesis - chondroitin sulfate	14	0.30561	2	0.0360		*XYLT1, CHSY3*

Paper III

Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Leslie R. Euceda[1], Tonje H. Haukaas[1,2], Guro F. Giskeødegård[1,3], Riyas Vettukattil[1], Jasper Engel[4], Laxmi Silwal-Pandit[2,5], Steinar Lundgren[6,7], Elin Borgen[2,8], Øystein Garred[2,8], Geert Postma[9], Lutgarde M.C. Buydens[9], Anne-Lise Børresen-Dale[2,5], Olav Engebraaten[2,10], Tone F. Bathen[1,2]

[1]*Department of Circulation and Medical Imaging, Faculty of Medicine, NTNU, Norwegian University of Science and Technology, Trondheim, Norway*

[2]*K.G. Jebsen Center for Breast Cancer Research, Institute of Clinical Medicine, University of Oslo, Oslo, Norway*

[3]*St. Olavs University Hospital, Trondheim, Norway*

[4]*NERC Biomolecular Analysis Facility Metabolomics Node (NBAF-B), School of Biosciences, University of Birmingham, Birmingham, UK*

[5]*Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, The Norwegian Radium Hospital, Oslo, Norway*

[6]*Department of Cancer Research and Molecular Medicine, NTNU, Trondheim, Norway*

[7]*Cancer Clinic, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway*

[8]*Department of Pathology, Division of Diagnostics and Intervention, Oslo University Hospital, Oslo, Norway*

[9]*Institute for Molecules and Materials, Radboud University Nijmegen, Nijmegen, The Netherlands*

[10]*Department of Oncology and Department of Tumor Biology, Oslo University Hospital, and Department of Clinical Medicine, University of Oslo, Norway*

Corresponding author: Leslie R. Euceda, Department of Circulation and Medical Imaging - MR Center, Faculty of Medicine, NTNU, Norwegian University of Science and Technology, Postboks 8905, Medisinsk Teknisk Forskningssenter, 7491 Trondheim, Norway; Phone: +47 73597449, Fax: +47 73598613, E.mail: leslie.e.wood@ntnu.no.

Acknowledgements: The authors would like to thank Øyvind Salvesen for useful discussions regarding linear mixed-effects models and Santosh Lamichhane for technical support during HR MAS MRS acquisition. The HR MAS MRS analysis was performed at the MR Core Facility, Norwegian University of Science and Technology (NTNU), which is funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority.

The study was funded in part by generous grants from:

(1) The Research Council of Norway, Imaging the breast cancer metabolome, Project no 221879
(2) The Pink Ribbon Movement and Norwegian Breast Cancer Society
(3) K. G. Jebsen Center for Breast Cancer Research
(4) Roche Norway
(5) Sanofi-Aventis Norway

Abstract

Introduction: Metabolomics investigates biochemical processes directly, potentially complementing transcriptomics and proteomics in providing insight into treatment outcome.

Objectives: This study aimed to use magnetic resonance (MR) spectroscopy on breast tumor tissue to explore the effect of neoadjuvant therapy on metabolic profiles, determine metabolic effects of the antiangiogenic drug bevacizumab, and investigate metabolic differences between responders and non-responders.

Methods: Breast tumors from 122 patients were profiled using high resolution magic angle spinning MR spectroscopy. All patients received neoadjuvant chemotherapy, and were randomized to receive bevacizumab or not. Tumors were biopsied prior, during, and after treatment.

Results: Principal component analysis showed clear metabolic changes indicating a decline in glucose consumption and a transition to normal breast adipose tissue as an effect of chemotherapy. Partial least squares-discriminant analysis revealed metabolic differences between pathological minimal residual disease patients and pathological non-responders after treatment (accuracy of 77%, $p<0.001$), but not before or during treatment. Lower glucose and higher lactate was observed in patients exhibiting a good response (≥90% tumor reduction) compared to those with no response (≤10% tumor reduction) before treatment, while the opposite was observed after treatment. Bevacizumab-receiving and chemotherapy-only patients could not be discriminated at any time point. Linear mixed-effects models revealed a significant interaction between *time* and *bevacizumab* for glutathione, indicating higher levels of this antioxidant in chemotherapy-only patients than in bevacizumab receivers after treatment.

Conclusion: MR spectroscopy showed potential in detecting metabolic response to treatment and complementing other molecular assays for the elucidation of underlying mechanisms affecting pathological response.

Keywords: bevacizumab, breast cancer, chemotherapy, HR MAS MRS, metabolomics, neoadjuvant

INTRODUCTION

Breast cancer prognosis has improved during the past decades (Ferlay et al. 2013), attributed to earlier detection through effective mammography screening and improved therapy (American Cancer Society 2011; Malmgren et al. 2014; Giordano et al. 2004). However, with the disease still being the leading cause of cancer death in women (Ferlay et al. 2013; American Cancer Society 2011), defining new tools to stratify patients to targeted therapy and to detect early response is warranted.

Neoadjuvant or preoperative therapy is used to treat locally advanced breast tumors where size is too large for surgical mastectomy with acceptable margins or to allow for breast conserving surgery (Miller et al. 2014). The effect of neoadjuvant treatment on progression-free survival (PFS) and overall survival (OS) is comparable to adjuvant or post-operative therapy (van der Hage et al. 2001), enabling the possibility for molecular analyses on tumor samples before, during, and after therapy, and a direct visualization of the treatment effect by pathological examination after removal of the tumor.

Angiogenesis, the formation of new blood vessels from existing vasculature, is a hallmark of cancer needed for the supply of nutrients and oxygen to the tumor (Hanahan and Weinberg 2000) and may be therapeutically targeted. Vascular endothelial growth factor (VEGF), a proangiogenic factor, can be inhibited by bevacizumab, a humanized monoclonal antibody (Ferrara et al. 2004). Addition of neoadjuvant bevacizumab to chemotherapy has resulted in better responses or increase in PFS, but has not increased OS (Miller et al. 2007). As for all targeted anticancer agents used in breast cancer, a reliable marker for the selection of patients who benefit from antiangiogenic therapy is needed.

Much of the variability in breast cancer treatment response has been attributed to tumor heterogeneity (Ng et al. 2012). Breast cancer has been divided into five molecular subtypes based on gene expression patterns found to have significant differences in clinical outcome (Perou et al. 2000; Sørlie et al. 2003). Metabolites are downstream of gene expression in the biochemical information flow known as the omics cascade, and altered metabolism is an emerging hallmark of cancer (Hanahan and Weinberg 2011). High resolution (HR) magic angle spinning (MAS) magnetic resonance spectroscopy (MRS) of intact tissue has been applied to study metabolites involved in important pathways in different cancers (Moestue et al. 2011; Bathen et al. 2010) . HR MAS MRS has been employed to relate metabolic changes after treatment to 5-year survival in breast cancer patients receiving neoadjuvant single agent chemotherapy followed by endocrine adjuvant therapy (Cao et al. 2012) and to reveal variability in metabolic profiles of breast tumors within the same gene-expression subtype (Borgan et al. 2010; Haukaas et al. in press). Thus, MR metabolomics can provide complimentary insight into breast cancer treatment outcome. This study aimed to use MRS to explore the metabolic changes in breast tumor tissue occurring as an effect of overall neoadjuvant therapy, to investigate metabolic differences between therapy responders and non-responders, and to determine the metabolic effects of antiangiogenic treatment with bevacizumab. Particular focus was given to

metabolites involved in the pathways of glycolysis, choline phospholipid metabolism, and glutaminolysis, due to their high relevance for cancer.

METHODS

Patient and tumor characteristics

Tumor tissue from 122 breast cancer patients included in the NeoAva study, a phase 2 randomized clinical trial, were examined. Women (age ≥ 18 years) with large (size ≥ 2.5 cm), non-metastatic, human epidermal growth factor receptor 2 (HER2) negative tumors were recruited in the period of November 2008 – July 2012. All patients received neoadjuvant chemotherapy in the form of FEC100 (5-fluorouracil 600 mg/m2, epirubicin 100 mg/m2, cyclophosphamide 600 mg/m2) followed by taxane-based therapy, while they were randomized to receive bevacizumab or not (See Online Resource 1). Ultrasound-guided needle biopsies were sampled prior to treatment (TP1) and 12 weeks into treatment (TP2), while biopsies at 24 weeks (TP3) were obtained from the surgically removed tumor. For details on sample handling, see Online Resource 1. The study was registered in the http://www.ClinicalTrials.gov/ database with the identifier NCT00773695. Patient and tumor characteristics are shown in Table 1, and the study design is illustrated in Fig. S1 (Online Resource 2).

Evaluation of treatment response

Analysis of only cases with pathological complete response (pCR) (20/122 patients) would result in a loss of information from the patients experiencing a good response, but not achieving pCR (44/122). Therefore, in addition to the pCR endpoint, we introduced a cut-off value of 1 cm for the tumor diameter measured pathologically at the time of surgery (TP3) to classify patients as having pathological minimal residual disease (pMRD) or as pathological non-responders (pNRs) (diameter <1cm and >1cm, respectively). Furthermore, a 'response ratio' inversely proportional to tumor reduction was calculated as: tumor diameter at TP3/tumor diameter at TP1. Response ratio >1 indicates an increase in tumor size with treatment. At TP1, the diameter was measured using magnetic resonance imaging (MRI). For patients with no available MRI measurement at TP1 (21/122 patients), the maximum value from ultrasound and/or mammography-measured diameters was used. Patients were categorized as having good, intermediate, or no response (GR, IR, NR) based on response ratio (≤0.10, >0.10 - <0.90, or ≥0.90, respectively).

Gene expression subtyping

Gene expression microarray data was acquired and processed as described in Online Resource 1. The PAM50 algorithm (Parker et al. 2009) was used to classify samples by gene expression subtype. Gene expression microarray data is available in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-4439.

HR MAS MRS Experiments and Data preprocessing

Tissue samples (4.1 ± 1.3 mg) were analyzed by HR MAS MRS in accordance with (Giskeødegård et al. 2015) as described in Online Resource 1.Sample availability at each time point varied for each patient (n=325 total analyzed samples). Samples with high lipid content (n=23) and giving spectra of poor quality, i.e. poor shimming and insufficient water suppression (n=17), were excluded. Additionally, fifteen samples collected from patients with tumors not evaluable for pathological measurement of diameter at TP3 were excluded, resulting in 270 samples for subsequent statistical analysis. Preprocessing before selection of spectral regions for further analysis was carried out in Matlab R2013b (The Mathworks, Inc., USA) as described in Online Resource 1. Lipid peaks at 4.37-4.27, 2.97-2.70, 2.31-2.23, 2.11-1.92, 1.64-1.49 ppm, and the contaminant peaks for ethanol and acetone at 3.71-3.63 and 2.23-2.21 ppm, respectively, were excluded. Regions containing peaks from lidocaine, a local anesthetic used for sampling biopsies at TP1 and TP2, at 3.38-3.30 and 2.21-2.17 ppm, were also excluded. The spectra were subsequently PQN (Dieterle et al. 2006) normalized. Normalization after lipid removal corrects for differences in sample size and tumor cell content, as it can be assumed that the majority of the lipid signals arise from normal adipose cells within the tumor specimen obtained.

Statistical analysis

Multivariate analyses

Multivariate analysis was carried out in Matlab R2013b (The Mathworks, Inc., USA) using PLS Toolbox 7.8.2 (Eigenvector Research Inc., U.S.A). Spectra were mean-centered prior to multivariate modelling. Principal component analysis (PCA) (Wold et al. 1987) was employed using residual explained variance plots to select the number of principal components (PCs). Partial least squares-discriminant analysis (PLS-DA) (Wold et al. 2001) models were built for each separate time point to discriminate between pCR+ and pCR-, pMRD and pNR, the two extreme response ratio groups of GR and NR, bevacizumab-treated and chemotherapy-only patients, and bevacizumab-treated pMRD and chemotherapy-only pMRD. PLS-DA models were validated using double-layered cross validation (CV) (Westerhuis et al. 2008) and permutation testing (Westerhuis et al. 2008) as described in Online Resource 1. For PLS-DA plots, y-variance was condensed into the first latent variable (LV) through orthogonal-PLS (OPLS) (Eigenvector Research 2013) when the number of optimal LVs >1.

Univariate analyses

Metabolites were assigned based on previous identification (Sitter et al. 2002) and relative levels were calculated by integrating fixed regions of preprocessed spectra (prior to normalization) corresponding to each metabolite in Matlab R2013b. The integrals were then subsequently PQN normalized to

correct for differences in sample weight. The integral of the metabolite lactate (Lac) was not calculable in 116 samples due to presence of an overlapping peak originating from the C_1H_2 of the glycerol backbone of phospholipids and triglycerides at 4.13 ppm. For these samples, the Lac levels were imputed in R 3.1.1 (R Core Team 2014) using the package for the method of multivariate imputation by chained equations (MICE) (Buuren and Groothuis-Oudshoorn 2011) set to predictive mean matching and 10 imputations (Rubin 1987). This procedure was validated as described in Online Resource 1. All metabolite integrals were log10 transformed before univariate tests were conducted to satisfy prerequisite assumptions of normality (Matlab R2013b). Student's t-tests were conducted to compare metabolite levels between different groups. Pearson correlations between metabolite relative levels and response ratio for each separate time point were calculated in SPSS Statistics 22.0, (IBM Corp., U.S.A.).

Linear mixed-effects models (LMM) (Pinheiro and Bates 2000) for individual metabolite multilevel analysis of data grouped according to different classification factors were built in R 3.1.1 using the function *lme* from the 'nlme' package (Pinheiro et al. 2014) employing the method of restricted maximum likelihood. LMM describe relationships between a response variable and these factors incorporating two types of effects: fixed which are controlled and random which depend on each individual. Here, the response variable was the metabolite level while the fixed effects were *time* (TP1, TP2, or TP3), *pMRD response* (pMRD or pNR), and *bevacizumab* (bevacizumab-treated or chemotherapy-only) and the random effect was the patient ID. Significance of interactions between fixed effects was evaluated as described in Online Resource 1. Visual inspection of residual q-q plots and histograms did not reveal obvious deviations from normality.

Multiple testing correction of the p-values resulting from Pearson correlations and LMM was performed by the Benjamini Hochberg method for false discovery rate (FDR) adjustment in R 3.1.1 using the 'stats' package (R Core Team 2014). Statistical significance was considered for adjusted p (q-value) ≤ 0.05.

RESULTS

Metabolic effects of neoadjuvant chemotherapy

The sample composition of the cohort according to gene expression subtype, time point, pMRD response group, and bevacizumab randomization group is summarized in Fig. S2 in Online Resource 2. PCA was employed on all spectra to investigate trends in the sample cohort as a whole (Fig. 1). Clear changes in metabolic profiles were observed with time (Fig. 1A). The scores plot labeled by gene-expression subtype (Fig. 1C) revealed a separation of the metabolic profiles of normal-like subtype samples from the rest. Comparing Fig. 1A and Fig. 1C, most TP3 samples were of the normal-

like gene expression subtype, also observed in Fig. S2F (Online Resource 2). The loadings plot (Fig. 1D) clearly indicated elevated phosphocholine (PCh), glycerophosphocholine (GPC), and taurine (Tau) at TP1, while glucose (Glc) and lipids were higher with increasing time of treatment and in normal-like samples.

Sixteen metabolites were identified as measurable, and their relative intensities were calculated by peak integration and log10 transformed (See Online Resource 3).

LMM revealed significant differences after multiple testing correction in 11/16 metabolites for the factor *time* (Table 2), including the metabolites detected in PCA with the exception of Tau. Four out of these 11 metabolites were significantly different for TP2 vs TP1, while all 11 of them were significant for TP3 vs TP2. Seven metabolites exhibited a significant interaction between *time* and *pMRD response*, while a significant interaction between *time* and *bevacizumab* was detected only for glutathione (GSH) (Supplementary Table 3 in Online Resource 4).

Metabolic differences between responders and non-responders

Results from PLS-DA discrimination of patients at each time point based on three different response criteria using metabolic profiles are summarized in Supplementary Table 4 in Online Resource 4. Metabolic differences between pCR+ and pCR- were detected at TP3 (Fig. 2A), but not TP1 or TP2, with an accuracy of 69% ($p=0.018$). Similarly, a significant discrimination between pMRD and pNRs was achieved only at TP3 (Fig. 2C), with PLS-DA model performance improving greatly (accuracy=77%, $p<0.001$). Compared to pCR- and pNR, both pCR+ and pMRD, respectively, showed elevated Glc and Lac and decreased creatine (Cr), glycine (Gly), Tau, GPC, PCh, choline (Cho), succinate (Succ), and alanine (Ala) levels (Fig. 2B, Fig. 2D).

PLS-DA significantly discriminated spectra from patients in the extreme response ratio groups of GR and NR at TP1 and TP3, but not TP2, with an accuracy of 76% ($p=0.001$) and 75% ($p=0.002$), respectively. The GR group showed lower levels of Glc and higher levels of ascorbate (Asc), Lac, Tyr, Cr, Gly, Tau, GPC, PCh, Cho, GSH, Succ, glutamate (Glu), and Ala compared to the NR group at TP1 (Fig. 3A-3B), while the opposite was observed at TP3 (Fig. 3C-3D). In addition, relative levels of two and eight metabolites were significantly correlated with the response ratio at TP1 and TP3, respectively (Supplementary Table 5 in Online Resource 4, Fig. S3 in Online Resource 2), but not at TP2. A switch in the direction of the correlation trend (positive or negative) was observed in the two metabolites with significant correlations at both time points (Glc and Tau).

Applying the multilevel approach of LMM, Lac was the only significant metabolite for the factor *pMRD response* (Table 2). However, the interaction between *time* and *pMRD response* was significant for seven metabolites (Supplementary Table 3 in Online Resource 4), indicating that the effect of the

pMRD response factor varies at different time points. *pMRD response* was significant at TP2 for Glc, Tyr, and Succ, while all seven metabolites were significant for this factor at TP3.

The most relevant results pertinent to treatment effect and comparison between responders and non-responders are summarized by pathway in Fig. 4, complemented by Supplementary Table 6 in Online Resource 4.

Glutathione metabolism identified as possibly affected by bevacizumab

PLS-DA could not discriminate bevacizumab-treated from chemotherapy-only patients at any time point. Both multivariate PLS-DA and univariate t-tests were employed to investigate differences between responding patients that were treated with additional bevacizumab and chemotherapy only. Significant metabolic differences between these groups could not be detected by either method at any time point (see Supplementary Table 7 and 8 in Online Resource 4).

For the multilevel LMM, *bevacizumab* was the only factor for which no metabolite was significant after multiple testing correction. However, a significant interaction between *time* and *bevacizumab* was detected for GSH (Table 2). GSH was significantly higher in controls than in bevacizumab-treated patients at TP3 ($p < 0.001$) (Supplementary Table 3 in Online Resource 4).

DISCUSSION

In this study we used MR based metabolic profiling of breast cancer tissue to investigate the effect of neoadjuvant chemotherapy in a large cohort of HER2 negative breast cancer patients randomized to receive additional bevacizumab treatment or not. The study design included three sampling time points over the treatment course, which allowed for a multilevel approach to investigate metabolic differences between patient groups.

Despite metabolic profiles not being able to predict pCR prior to treatment, a significant metabolic difference in pCR+ patients compared to pCR- was detected after neoadjuvant chemotherapy. This metabolic difference was even clearer when comparing pMRD patients and pNRs, with the former displaying similar metabolite trends as pCR+ patients. Moreover, a metabolic difference both before and after treatment was identified comparing patients with good response and no response. However, since the majority of cancer cells in pathological responders have responded to treatment, when comparing pMRD and pNRs at TP3, findings may also reflect metabolic differences between tumor and normal tissue. Previous studies using HR MAS MRS metabolic profiles of intact breast tumor tissue have successfully identified patients with more than 5-year survival, while association to treatment response was more subtle (Cao et al. 2012; Cao et al. 2011). Findings from the NASBP B-40 trial report a significant increase in OS of HER2 negative breast cancer patients receiving

chemotherapy with addition of neoadjuvant bevacizumab continued post-operatively (Bear et al. 2015). The significant discrimination of pathological response at surgery (TP3) based on metabolic profiles may be predictive of patients who would benefit from post-operative bevacizumab treatment. Metabolic signatures in patients with 5-year DFS, PFS and OS for this cohort will be investigated in due course.

Clear metabolic differences between time points in accordance with tumor reduction were observed. The increase in glucose with treatment progression points to a decline in glucose consumption, which is characteristically rapid in cancer cells, referred to as a glycolytic phenotype, to fulfill their high energy demands (Vander Heiden et al. 2009). Glucose was also found to be significantly higher in pMRD patients compared to pNRs after completion of neoadjuvant therapy, indicating an even greater decline in the consumption of glucose in the former group. The increase of lipids towards TP3 suggests a transition to normal breast adipose tissue. Lipid metabolism plays an important role in tumor cell apoptosis and necrosis resulting from treatment (Huang and Freter 2015). However, it can be assumed that most lipid signals arise from normal breast adipose cells. This is supported by samples at TP3 and of the normal-like subtype, which displays high expression of adipose cell genes (Perou et al. 2000), corresponding well in the PCA scores plot (Fig. 1A and 1C).

Metabolites that significantly decreased with treatment progression include glycerophosphocholine, phosphocholine, and choline, with the latter two also being significantly lower in pMRD patients compared to pNRs at TP3. These metabolites are involved in the metabolism of phosphatidylcholine, the most abundant phospholipid in eukaryote cell membranes (Gibellini and Smith 2010). Although the underlying mechanisms governing abnormal choline metabolism in cancer are still not fully understood (Glunde et al. 2011; Moestue et al. 2012), increase in the levels of choline containing metabolites are thought to be an indicator of malignant transformation and cancer aggressiveness (Aboagye and Bhujwalla 1999). Decreased phospholipid metabolism may be an indicator of a reduction in tumor malignancy and aggressiveness due to neoadjuvant chemotherapy. Further, studies showing that the behavior of choline metabolites varies between breast cancer subtypes (Giskeødegård et al. 2010; Grinde et al. 2014) indicate that the stratification of patients for treatment targeting phospholipid metabolism may be beneficial. This is further supported by our recent study (Haukaas et al. in press) in which we used metabolic profiles to derive naturally-occurring metabolic subgroups, revealing that one of these contained significantly higher phosphocholine and glycerophosphocholine levels and less active degradation of phosphatidylcholine determined by integrative pathway analysis.

LMM additionally revealed increased glutamine with treatment progression indicating a decrease in glutaminolysis. The normally non-essential amino acid glutamine becomes an important alternative source of carbon and nitrogen for glucose-deprived cancer cells, replenishing the TCA cycle and providing substrates for nucleotide, lipid, and protein biosynthesis (Hensley et al. 2013). Glutamine is

converted to glutamate, a precursor of non-essential amino acids and glutathione. Glutathione is a major cellular antioxidant which protects cancer cells against apoptosis caused by reactive oxygen species (ROS) (Franco and Cidlowski 2009). A significant interaction between *time* and *bevacizumab* was detected for glutathione, with chemotherapy-only patients exhibiting significantly higher levels than patients receiving bevacizumab at TP3. This is coherent with previously reported reduced levels of glutathione with bevacizumab treatment of glioblastoma (Fack et al. 2015). Bevacizumab may thus play a redox destabilizing role in cancer cells, promoting apoptosis.

A significant LMM interaction between *time* and *response* was detected for succinate, showing significantly lower levels in pMRD patients compared to pNRs at TP2 and TP3. Succinate is the substrate of succinate dehydrogenase (SDH), a tricarboxylic acid (TCA) cycle enzyme which has been identified as a tumor suppressor (Selak et al. 2005). Inactivation of SDH leads to accumulation of succinate, which mediates signaling pathways promoting resistance to apoptotic signals and inhibits the degradation of HIF alpha, a transcription factor that promotes the glycolytic phenotype characteristic of cancer cells (King et al. 2006). Decreased succinate in responders, observed at both TP2 and TP3, is therefore in accordance with a less malignant phenotype.

Unexpectedly, a significant increase in lactate with treatment progression and in pMRD patients compared to pNRs was detected, similarly as for glucose. Both increased lactate production and rapid glucose consumption are characteristic of the Warburg effect, which is the metabolic switch from anaerobic to aerobic glycolysis observed in most cancer cells (Vander Heiden et al. 2009). Lactate has been associated with poor prognosis in ER positive breast cancer (Giskeødegård et al. 2012) as well as in other cancers (Saraswathy et al. 2009; Walenta and Mueller-Klieser 2004; Walenta et al. 2000). Still, the glycolytic phenotype is not absolute in cancer cells, as a better-oxygenated subpopulation uses lactate as the main energy source fueling oxidative phosphorylation (Hanahan and Weinberg 2011; Semenza 2008; Feron 2009). A dual metabolic effect in cancer cells has been suggested (Xie et al. 2014), where lactic acidosis, although a consequence of the Warburg effect, can convert the usually dominant Warburg effect to a non-glycolytic phenotype, decreasing glucose consumption. Elevated lactate promoting this non-glycolytic phenotype is a possible explanation for glucose and lactate not behaving oppositely as expected. Another possibility is that chemotherapy may produce morphological changes in tissue that impede adequate lactate cellular excretion, leading to its accumulation. An issue to be kept in mind is that imputation was employed to calculate the relative levels of lactate for 116/270, which can be considered a limitation of this study. However, the imputation process was validated achieving a root mean square error (RMSE) of 0.006 and a coefficient of determination (R^2) of 0.9996, indicating high confidence for the predicted lactate values.

Regarding the multivariate comparison between good response and no response groups based on the response ratio, glucose and lactate behaved oppositely in both groups at both TP1 and TP3, contrary to

them exhibiting trends in the same direction in pMRD patients and pNRs at TP3. However, when including samples from all three response ratio groups, glucose and lactate were both significantly negatively correlated with tumor size increase at TP3. The contradictory lactate/glucose behavior using different response criteria may therefore be due to the exclusion of 69/122 patients when comparing the two extreme response ratio groups. Lower glucose and higher lactate in the good response group before treatment points to a higher Warburg effect compared to the no response group, which is subsequently reversed in these groups by the end of treatment. In the same way, compounds associated with a more malignant phenotype including the choline containing metabolites, glutathione, glutamate, and succinate were elevated in the good response group before treatment, while the opposite was observed at the end of treatment. Although models built with these extreme groups cannot be used for prediction of treatment response, since a large part of the population has been excluded, these findings suggest that patients with a more malignant metabolic profile are more prone to benefit from treatment in terms of tumor reduction.

Changes in glutathione metabolism were identified as a possible effect of bevacizumab based on a significant interaction between *time* and *bevacizumab* using LMM for this metabolite. A substantial metabolic effect due to chemotherapy in both bevacizumab-treated and chemotherapy-only patients was evidenced, suggesting the potent chemotherapeutic effect may mask many of the effects of bevacizumab. In addition, the increase in lactate observed with increasing time on therapy may counteract VEGF inhibition by bevacizumab, as it has been found that excreted lactate from tumor cells stimulates angiogenesis dependent of a different proangiogenic factor: interleukin 8 (IL-8) (Végran et al. 2011). IL-8 signaling has been found to compensate for inhibition of VEGF-dependent angiogenesis in bevacizumab-resistant cells of head and neck squamous cell carcinoma (Gyanchandani et al. 2013). Altered lactate metabolism as a mechanism of bevacizumab resistance should therefore be further investigated.

Conclusion

MR based metabolic profiles of non-metastatic breast tumor tissue reflected changes in all patient groups as an effect of chemotherapy. Metabolic profiles discriminated pNRs from pMRD patients after neoadjuvant chemotherapy, which may complement other molecular assays for the elucidation of the underlying mechanisms affecting pathological response. Although metabolic differences based on bevacizumab administration were not prominent, glutathione was identified to be possibly affected by the antiangiogenic drug.

Abbreviations

Ala: alanine; Asc: ascorbate; ATP: adenosine triphosphate; Cho: choline; Cr: creatine; CV: cross validation; ER: estrogen receptor; FDR: false discovery rate; FEC: 5-fluorouracil-epirubicin-cyclophosphamide; Glc: glucose; Glu: glutamate; Gly: glycine; GPC: glycerophosphocholine; GR: good response; GSH: glutathione; HER2: human epidermal growth factor receptor 2; HIF: hypoxia-inducible factor; HR MAS MRS: high resolution magic angle spinning magnetic resonance spectroscopy; IL-8: interleukin 8; IR: intermediate response; Lac: lactate; LMM: linear mixed-effects model; LV: latent variable; MICE: multiple imputation by chained equations; MR: magnetic resonance; MRI: magnetic resonance imaging; NR: no response; OPLS: orthogonal partial least squares; OS: overall survival; PAM50: prediction analysis of microarrays 50; PC: principal component; PCA: principal component analysis; PCh: phosphocholine; pCR: pathological complete response; PFS: progression-free survival; PLS-DA: partial least squares-discriminant analysis; pMRD: pathological minimal residual disease; pNR: pathological non-responder; PQN: probabilistic quotient normalization; q-q: quantile-quantile; RMSE: root mean square error; ROS: reactive oxygen species; R^2: coefficient of determination; SDH: succinate dehydrogenase; Succ: succinate; Tau: taurine; TCA: tricarboxylic acid; TP: time point; VEGF: vascular endothelial growth factor.

Authors' contributions

LRE, THH, GFG, RV, JE, LSP, GP, LMCB, ALBD, OE, and TFB participated in the design of the study. ALBD, OE, and TFB conceived the study. LRE, THH, GFG, RV, JE, GP, LMCB, ALBD, OE, and TFB interpreted the data. LRE performed the HR MAS MRS acquisition, statistical analysis, and drafted the manuscript. LSP, SL, EB, OG, ALBD, OE, and TFB participated in acquisition of the data. All authors have read and helped to revise the manuscript. The final manuscript is approved by all the authors.

Acknowledgements

The authors would like to thank Øyvind Salvesen for useful discussions regarding linear mixed-effects models and Santosh Lamichhane for technical support during HR MAS MRS acquisition. The HR MAS MRS analysis was performed at the MR Core Facility, Norwegian University of Science and Technology (NTNU), which is funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority.

(1) The Research Council of Norway, Imaging the breast cancer metabolome, Project no 221879
(2) The Pink Ribbon Movement and Norwegian Breast Cancer Society
(3) K. G. Jebsen Center for Breast Cancer Research

(4) Roche Norway
(5) Sanofi-Aventis Norway

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The NeoAva study was co-sponsored by Roche Norway and Sanofi-Aventis Norway. Oslo University Hospital is the main sponsor for the NeoAva study.

Compliance with ethical requirement

The study was approved for all centers involved by the Regional Ethics Committee (Approval number S-08354a) and the Norwegian Medical Agency. All procedures performed in studies involving human participants were in accordance with the Declaration of Helsinki, International Conference on Harmony/Good Clinical practice. Informed written consent was obtained from all patients.

References

Aboagye, E. O., & Bhujwalla, Z. M. (1999). Malignant Transformation Alters Membrane Choline Phospholipid Metabolism of Human Mammary Epithelial Cells. Cancer Research, 59(1), 80-84.

American Cancer Society (2011). Global Cancer Facts & Figures 2nd Edition. Atlanta: American Cancer Society.

Bathen, T. F., Sitter, B., Sjøbakk, T. E., Tessem, M.-B., & Gribbestad, I. S. (2010). Magnetic Resonance Metabolomics of Intact Tissue: A Biotechnological Tool in Cancer Diagnostics and Treatment Evaluation. *Cancer Research, 70*(17), 6692-6696, doi:10.1158/0008-5472.can-10-0437.

Bear, H. D., Tang, G., Rastogi, P., Geyer Jr, C. E., Liu, Q., Robidoux, A., et al. (2015). Neoadjuvant plus adjuvant bevacizumab in early breast cancer (NSABP B-40 [NRG Oncology]): secondary outcomes of a phase 3, randomised controlled trial. *The Lancet Oncology, 16*(9), 1037-1048, doi:http://dx.doi.org/10.1016/S1470-2045(15)00041-8.

Borgan, E., Sitter, B., Lingjærde, O. C., Johnsen, H., Lundgren, S., Bathen, T. F., et al. (2010). Merging trancriptomics and metaboloimics - advances in breast cancer profiling. *BMC Cancer, 10*, 628, doi:10.1186/1471-2407-10-628.

Buuren, S. v., & Groothuis-Oudshoorn, K. (2011). *Journal of Statistical Software, 45*(3), 1-67.

Cao, M. D., Giskeodegard, G. F., Bathen, T. F., Sitter, B., Bofin, A., Lonning, P. E., et al. (2012). Prognostic value of metabolic response in breast cancer patients receiving neoadjuvant chemotherapy. *BMC Cancer, 12*, doi:10.1186/1471-2407-12-39.

Cao, M. D., Sitter, B., Bathen, T. F., Bofin, A., Lonning, P. E., Lundgren, S., et al. (2011). Predicting long-term survival and treatment response in breast cancer patients receiving neoadjuvant chemotherapy by MR metabolic profiling. *NMR Biomed, 25*, doi:10.1002/nbm.1762.

Dieterle, F., Ross, A., Schlotterbeck, G., & Senn, H. (2006). Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. *Analytical Chemistry, 78*(13), 4281-4290, doi:10.1021/ac051632c.

Eigenvector Research (2013). Orthogonalizepls. http://wiki.eigenvector.com/index.php?title=Orthogonalizepls. Accessed August 2015.

Fack, F., Espedal, H., Keunen, O., Golebiewska, A., Obad, N., Harter, P., et al. (2015). Bevacizumab treatment induces metabolic adaptation toward anaerobic metabolism in glioblastomas. *Acta Neuropathologica, 129*(1), 115-131, doi:10.1007/s00401-014-1352-5.

Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., et al. (2013). GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. Accessed July 21 2015.

Feron, O. (2009). Pyruvate into lactate and back: From the Warburg effect to symbiotic energy fuel exchange in cancer cells. *Radiotherapy and Oncology, 92*(3), 329-333, doi:http://dx.doi.org/10.1016/j.radonc.2009.06.025.

Ferrara, N., Hillan, K. J., Gerber, H.-P., & Novotny, W. (2004). Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. [10.1038/nrd1381]. *Nat Rev Drug Discov, 3*(5), 391-400.

Franco, R., & Cidlowski, J. A. (2009). Apoptosis and glutathione: beyond an antioxidant. *Cell Death Differ, 16*(10), 1303-1314, doi:http://www.nature.com/cdd/journal/v16/n10/suppinfo/cdd2009107s1.html.

Gibellini, F., & Smith, T. K. (2010). The Kennedy pathway—De novo synthesis of phosphatidylethanolamine and phosphatidylcholine. *IUBMB Life, 62*(6), 414-428, doi:10.1002/iub.337.

Giordano, S. H., Buzdar, A. U., Smith, T. L., Kau, S.-W., Yang, Y., & Hortobagyi, G. N. (2004). Is breast cancer survival improving? *Cancer, 100*(1), 44-52, doi:10.1002/cncr.11859.

Giskeødegård, G., Cao, M., & Bathen, T. (2015). High-Resolution Magic-Angle-Spinning NMR Spectroscopy of Intact Tissue. In J. T. Bjerrum (Ed.), *Metabonomics* (Vol. 1277, pp. 37-50, Methods in Molecular Biology): Springer New York.

Giskeødegård, G. F., Grinde, M. T., Sitter, B., Axelson, D. E., Lundgren, S., Fjøsne, H. E., et al. (2010). Multivariate Modeling and Prediction of Breast Cancer Prognostic Factors Using MR Metabolomics. *Journal of Proteome Research, 9*(2), 972-979, doi:10.1021/pr9008783.

Giskeødegård, G. F., Lundgren, S., Sitter, B., Fjøsne, H. E., Postma, G., Buydens, L. M. C., et al. (2012). Lactate and glycine—potential MR biomarkers of prognosis in estrogen receptor-positive breast cancers. *NMR in Biomedicine, 25*(11), 1271-1279, doi:10.1002/nbm.2798.

Glunde, K., Bhujwalla, Z. M., & Ronen, S. M. (2011). Choline metabolism in malignant transformation. [10.1038/nrc3162]. *Nat Rev Cancer, 11*(12), 835-848.

Grinde, M. T., Skrbo, N., Moestue, S. A., Rødland, E. A., Borgan, E., Kristian, A., et al. (2014). Interplay of choline metabolites and genes in patient-derived breast cancer xenografts. *Breast Cancer Research, 16*, R5, doi:10.1186/bcr3597.

Gyanchandani, R., Sano, D., Ortega Alves, M. V., Klein, J. D., Knapick, B. A., Oh, S., et al. (2013). Interleukin-8 as a modulator of response to bevacizumab in preclinical models of head and neck squamous cell carcinoma. *Oral Oncology, 49*(8), 761-770, doi:http://dx.doi.org/10.1016/j.oraloncology.2013.03.452.

Hanahan, D., & Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell, 100*(1), 57-70, doi:http://dx.doi.org/10.1016/S0092-8674(00)81683-9.

Hanahan, D., & Weinberg, Robert A. (2011). Hallmarks of Cancer: The Next Generation. *Cell, 144*(5), 646-674, doi:http://dx.doi.org/10.1016/j.cell.2011.02.013.

Haukaas, T., Euceda, L., Giskeødegård, G., Lamichhane, S., Krohn, M., Jernström, S., et al. (in press). Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes *Cancer & Metabolism*.

Hensley, C. T., Wasti, A. T., & DeBerardinis, R. J. (2013). Glutamine and cancer: cell biology, physiology, and clinical opportunities. *The Journal of Clinical Investigation, 123*(9), 3678-3684, doi:10.1172/JCI69600.

Huang, C., & Freter, C. (2015). Lipid Metabolism, Apoptosis and Cancer Therapy. *International Journal of Molecular Sciences, 16*(1), 924.

King, A., Selak, M. A., & Gottlieb, E. (2006). Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. *Oncogene, 25*(34), 4675-4682.

Malmgren, J. A., Parikh, J., Atwood, M. K., & Kaplan, H. G. (2014). Improved Prognosis of Women Aged 75 and Older with Mammography-detected Breast Cancer. *Radiology, 273*(3), 686-694, doi:doi:10.1148/radiol.14140209.

Miller, E., Lee, H., Lulla, A., Hernandez, L., Gokare, P., & Lim, B. (2014). *Current treatment of early breast cancer: adjuvant and neoadjuvant therapy [v1; ref status: indexed, http://f1000r.es/3l2]* (Vol. 3, Vol. 198): F1000Research.

Miller, K., Wang, M., Gralow, J., Dickler, M., Cobleigh, M., Perez, E. A., et al. (2007). Paclitaxel plus Bevacizumab versus Paclitaxel Alone for Metastatic Breast Cancer. *New England Journal of Medicine, 357*(26), 2666-2676, doi:doi:10.1056/NEJMoa072113.

Moestue, S., Sitter, B., Bathen, T. F., Tessem, M.-B., & Gribbestad, I. S. (2011). HR MAS MR spectroscopy in metabolic characterization of cancer. *Current Topics in Medicinal Chemistry, 11*(1), 2-26.

Moestue, S. A., Giskeødegård, G. F., Cao, M. D., Bathen, T. F., & Gribbestad, I. S. (2012). Glycerophosphocholine (GPC) is a poorly understood biomarker in breast cancer. *Proceedings of the National Academy of Sciences, 109*(38), E2506, doi:10.1073/pnas.1208226109.

Ng, C. K., Pemberton, H. N., & Reis-Filho, J. S. (2012). Breast cancer intratumor genetic heterogeneity: causes and implications. *Expert Review of Anticancer Therapy, 12*(8), 1021-1032, doi:doi:10.1586/era.12.85.

Parker, J. S., Mullins, M., Cheang, M. C. U., Leung, S., Voduc, D., Vickery, T., et al. (2009). Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *Journal of Clinical Oncology, 27*(8), 1160-1167, doi:10.1200/jco.2008.18.1370.

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., et al. (2000). Molecular portraits of human breast tumours. [10.1038/35021093]. *Nature, 406*(6797), 747-752, doi:http://www.nature.com/nature/journal/v406/n6797/suppinfo/406747a0_S1.html.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2014) (2014). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-117, URL: http://CRAN.R-project.org/package=nlme.

Pinheiro, J. C., & Bates, D. M. (2000). Linear Mixed-Effects Models: Basic Concepts and Examples. In *Mixed-Effects Models in S and S-PLUS* (pp. 3-56, Statistics and Computing). New York, NY, USA: Springer New York.

R Core Team (2014). R: A language and environment for statistical computing. Viena, Austria. URL http://www.R-project.org/: R Foundation for Statistical Computing.

Rubin, D. B. (1987). *Mulitiple Imputation for Nonresponse in Surveys*. New York, USA: John Wiley & Sons.

Saraswathy, S., Crawford, F., Lamborn, K., Pirzkall, A., Chang, S., Cha, S., et al. (2009). Evaluation of MR markers that predict survival in patients with newly diagnosed GBM prior to adjuvant therapy. *Journal of Neuro-Oncology, 91*(1), 69-81, doi:10.1007/s11060-008-9685-3.

Selak, M. A., Armour, S. M., MacKenzie, E. D., Boulahbel, H., Watson, D. G., Mansfield, K. D., et al. (2005). Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase. *Cancer Cell, 7*(1), 77-85, doi:http://dx.doi.org/10.1016/j.ccr.2004.11.022.

Semenza, G. L. (2008). Tumor metabolism: cancer cells give and take lactate. *The Journal of Clinical Investigation, 118*(12), 3835-3837, doi:10.1172/JCI37373.

Sitter, B., Sonnewald, U., Spraul, M., Fjösne, H. E., & Gribbestad, I. S. (2002). High-resolution magic angle spinning MRS of breast cancer tissue. *NMR in Biomedicine, 15*(5), 327-337, doi:10.1002/nbm.775.

Sørlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., et al. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences, 100*(14), 8418-8423, doi:10.1073/pnas.0932692100.

van der Hage, J. A., van de Velde, C. J. H., Julien, J.-P., Tubiana-Hulin, M., Vandervelden, C., Duchateau, L., et al. (2001). Preoperative Chemotherapy in Primary Operable Breast Cancer: Results From the European Organization for Research and Treatment of Cancer Trial 10902. *Journal of Clinical Oncology, 19*(22), 4224-4237.

Vander Heiden, M. G., Cantley, L. C., & Thompson, C. B. (2009). Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science, 324*(5930), 1029-1033, doi:10.1126/science.1160809.

Végran, F., Boidot, R., Michiels, C., Sonveaux, P., & Feron, O. (2011). Lactate Influx through the Endothelial Cell Monocarboxylate Transporter MCT1 Supports an NF-κB/IL-8 Pathway that

Drives Tumor Angiogenesis. *Cancer Research, 71*(7), 2550-2560, doi:10.1158/0008-5472.can-10-2828.

Walenta, S., & Mueller-Klieser, W. F. (2004). Lactate: mirror and motor of tumor malignancy. *Seminars in Radiation Oncology, 14*(3), 267-274, doi:http://dx.doi.org/10.1016/j.semradonc.2004.04.004.

Walenta, S., Wetterling, M., Lehrke, M., Schwickert, G., Sundfør, K., Rofstad, E. K., et al. (2000). High Lactate Levels Predict Likelihood of Metastases, Tumor Recurrence, and Restricted Patient Survival in Human Cervical Cancers. *Cancer Research, 60*(4), 916-921.

Westerhuis, J., Hoefsloot, H. J., Smit, S., Vis, D., Smilde, A., van Velzen, E. J., et al. (2008). Assessment of PLSDA cross validation. *Metabolomics, 4*(1), 81-89, doi:10.1007/s11306-007-0099-6.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal Component Analysis. *Chemometr Intell Lab, 2*, 37-52.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109-130, doi:http://dx.doi.org/10.1016/S0169-7439(01)00155-1.

Xie, J., Wu, H., Dai, C., Pan, Q., Ding, Z., Hu, D., et al. (2014). Beyond Warburg effect - dual metabolic nature of cancer cells. [Article]. *Sci. Rep., 4*, doi:10.1038/srep04927
http://www.nature.com/srep/2014/140513/srep04927/abs/srep04927.html#supplementary-information.

Tables

Table 1. Patient and tumor characteristics.

		Total	Bevacizumab-treated	Chemotherapy Only
Number of patients		122	60	62
Mean age (range)	years	49.0 (25 – 70)	49.6 (27-70)	48.4 (25-68)
Therapy Response				
Pathological complete response (pCR)	pCR	20	13	7
	No pCR	102	47	55
Pathological minimal residual disease (pMRD)	pMRD	44	24	20
	Pathological non-responder (pNR)	78	36	42
Response ratio	Good response (GR)	31	18	13
	Intermediate response (IR)	69	32	37
	No response (NR)	22	10	12
Histology	Ductal	98	47	51
	Lobular	21	12	9
	Other	3	1	2
T status at TP1	T1	0	0	0
	T2	36	17	19
	T3	74	37	37
	T4	11	5	6
	NA	1	1	0
T status at TP3	T1	49	21	28
	T2	36	20	16
	T3	14	5	9
	T4	0	0	0
	NA	23	14	9
Node status at TP3	N0	45	23	22
	N1	30	14	16
	N2	14	5	9
	N3	4	0	4
	NA	29	18	11
ER status	+	101	48	53
	-	21	12	9
PgR status	+	70	32	38
	-	51	27	24
	NA	1	1	0

NA: not available

Table 2. LMM results for the fixed effects of *time*, *pMRD response*, and *bevacizumab*.

	TP2 vs TP1	TP3 vs TP2	pMRD vs pNR	Bev-treated vs Chemo-only	Significant Interactions[#]
Glc	0.130* (0.065)	0.309** (0.064)	-0.155 (0.078)	0.096 (0.055)	*Time* × *Response*
Asc	-0.036* (0.015)	-0.107** (0.014)	-0.016 (0.017)	0.005 (0.017)	None
Lac	0.173** (0.050)	0.398** (0.048)	0.157** (0.049)	0.053 (0.047)	None
Tyr	-0.047* (0.023)	-0.121** (0.022)	0.052* (0.026)	-0.023 (0.016)	*Time* × *Response*
mI	0.050** (0.016)	0.035** (0.015)	-0.014 (0.021)	-0.003 (0.020)	None
Gly	-0.029 (0.017)	0.004 (0.016)	0.011 (0.021)	0.018 (0.020)	None
Tau	<0.001 (0.021)	0.042* (0.021)	-0.059* (0.026)	0.014 (0.018)	*Time* × *Response*
GPC	-0.032 (0.018)	-0.057** (0.017)	-0.001 (0.022)	-0.022 (0.021)	None
PCh	-0.061** (0.023)	-0.101** (0.023)	0.057 (0.029)	-0.024 (0.021)	*Time* × *Response*
Cho	-0.055** (0.018)	-0.043** (0.018)	0.051* (0.022)	-0.034* (0.015)	*Time* × *Response*
Cr	-0.015 (0.017)	-0.048** (0.016)	-0.034 (0.018)	0.007 (0.017)	None
GSH	0.022 (0.030)	-0.045 (0.030)	0.044 (0.029)	0.057* (0.029)	*Time* × *Response;* *Time* × *Bevacizumab*
Gln	0.021 (0.022)	0.119** (0.021)	-0.011 (0.026)	-0.007 (0.025)	None
Succ	-0.035 (0.037)	0.035 (0.036)	0.048 (0.042)	-0.032 (0.027)	*Time* × *Response*
Glu	0.018 (0.021)	0.014 (0.020)	0.064* (0.025)	0.045 (0.024)	None
Ala	0.042 (0.029)	0.096** (0.028)	-0.016 (0.028)	0.039 (0.026)	None

Metabolite increase (positive estimate) or decrease (negative estimate) is shown for TP2, TP3, pMRD, and bevacizumab-treated patients in relation to TP1,TP2, pNR, and chemotherapy-only patients, respectively. Standard errors are shown in parenthesis. * and ** indicate significance ($p \leq 0.05$) before and after multiple testing correction, respectively.

[#]Final LMMs shown here were built including significant interaction terms, but interaction results are presented separately in Table 3.

pMRD: pathological minimal residual disease; pNR: pathological non-responder; Bev-treated: Bevacizumab-treated; Chemo-only: Chemotherapy-only; Glc: glucose; Asc: ascorbate; Lac: lactate; Tyr: tyrosine; mI: myo-inositol; Gly: glycine; Tau: taurine; GPC: glycerophosphocholine; PCh: phosphoscholine; Cho: choline; Cr: creatine; GSH: glutathione; Gln: glutamine; Succ: succinate; Glu: glutamate; Ala: alanine.

Figures

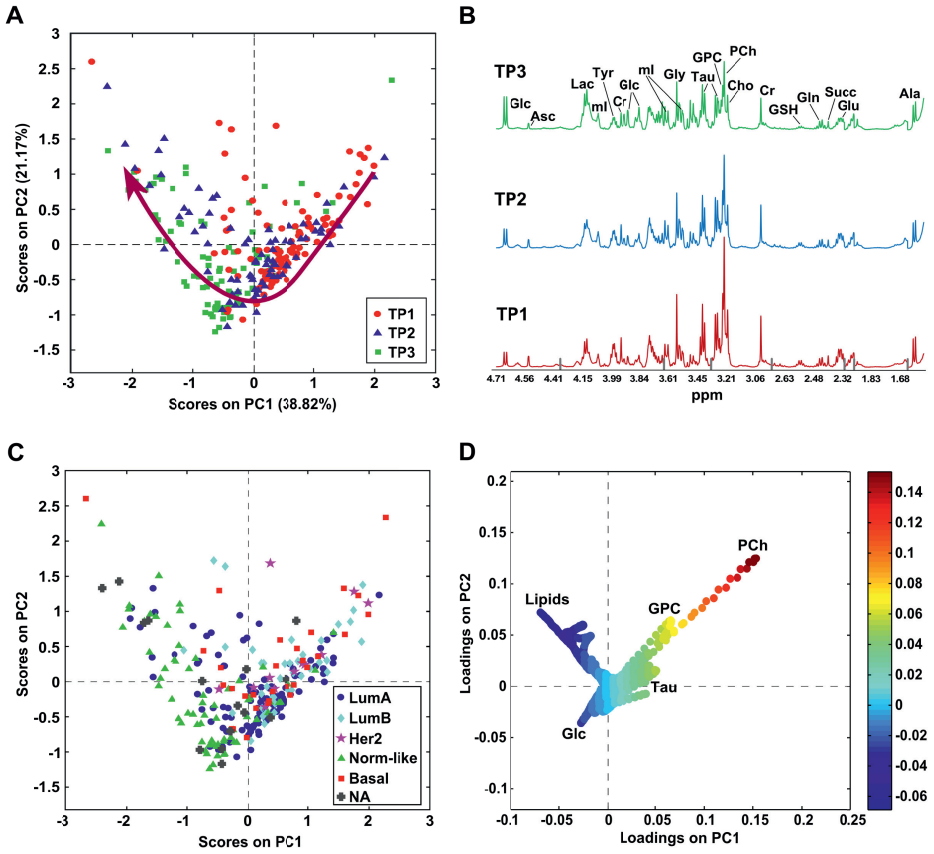


Fig. 1. PCA including all samples. (A) The scores plot shows a trend in the direction of the arrow with increasing time point. (B) PQN-normalized mean spectra at each time point. Gray bars indicate removed spectral regions. (C) The normal-like gene expression subtype is most clearly separated from the rest in the scores plot, showing a similar distribution as TP3 in A. (D) The loadings plot indicates higher phosphocholine, glycerophosphocholine, and taurine at TP1 and increasing glucose and lipids with increasing time of treatment and in normal-like samples. Loadings are colored according to LV1. LumA=luminal A; LumB=luminal B; Norm-like=normal-like; NA=not available; Glc=glucose; Asc=ascorbate; Lac=lactate; mI=myo-inositol; Tyr=tyrosine; Cr=creatine; Gly=glycine; Tau=taurine; GPC=glycerophosphocholine; PCh=phosphoscholine; Cho=choline; GSH=glutathione; Gln=glutamine; Succ=succinate; Glu=glutamate; Ala=alanine.

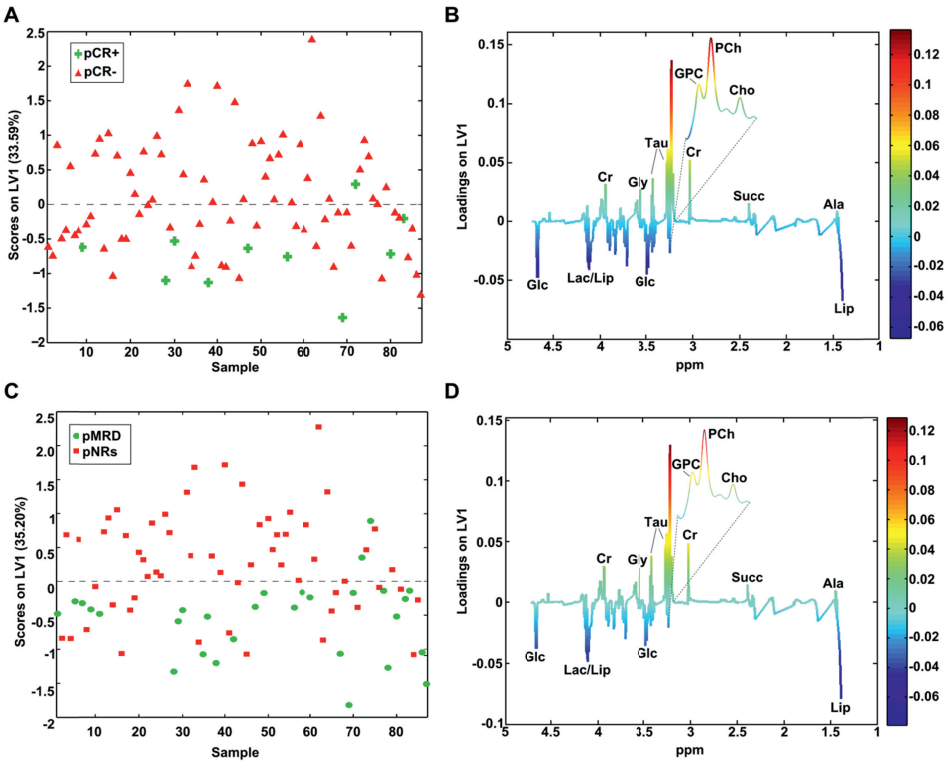


Fig. 2. PLS-DA scores (A) and loadings (B) plots of pathological complete responders (pCR+) vs patients with no pathological complete response (pCR-) at TP3, and PLS-DA scores (C) and loadings (D) plots of pathological minimal residual disease patients (pMRD) vs pathological non-responders (pNRs) at TP3. Both pCR+ (A) and pMRD (C) display lower scores along LV1 than pCR- and pNRs, respectively. Loadings show pCR+ (B) and pMRD (D) having higher glucose and lactate compared to pCR- and pNRs, respectively, and lower levels of the choline containing metabolites, creatine, glycine, taurine, succinate, and alanine. Loadings are colored according to LV1. Glc=glucose; Lac=lactate; Lip=lipids; Cr=creatine; Gly=glycine; Tau=taurine; GPC=glycerophosphocholine; PCh=phosphoscholine; Cho=choline; Succ=succinate; Ala=alanine.

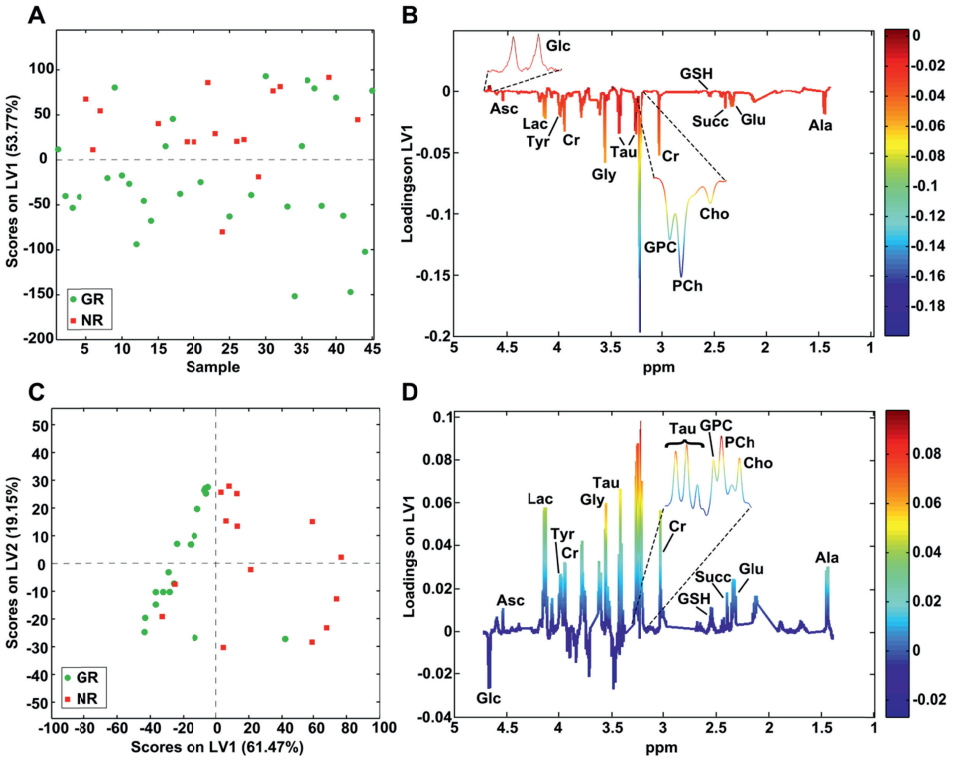


Fig. 3. PLS-DA scores and loadings plots of good response (GR) vs no response (NR) patient groups at TP1 (A-B) and TP3 (C-D). (A) At TP1, the GR group displays lower scores along LV1 than the NR group, with loadings showing the former group having lower glucose, and higher levels of lactate, taurine, the choline containing metabolites, glutathione, and succinate, among others (B). (C) At TP3, the GR group also displays lower scores along LV1, but the loading profile (D) is inverted when compared to TP1, indicating a metabolic switch related to tumor reduction with treatment time. Loadings are colored according to LV1. Glc=glucose; Asc=ascorbate; Lac=lactate; Tyr=tyrosine; Cr=creatine; Gly=glycine; Tau=taurine; GPC=glycerophosphocholine; PCh=phosphoscholine; Cho=choline; GSH=glutathione; Succ=succinate; Glu=glutamate; Ala=alanine.

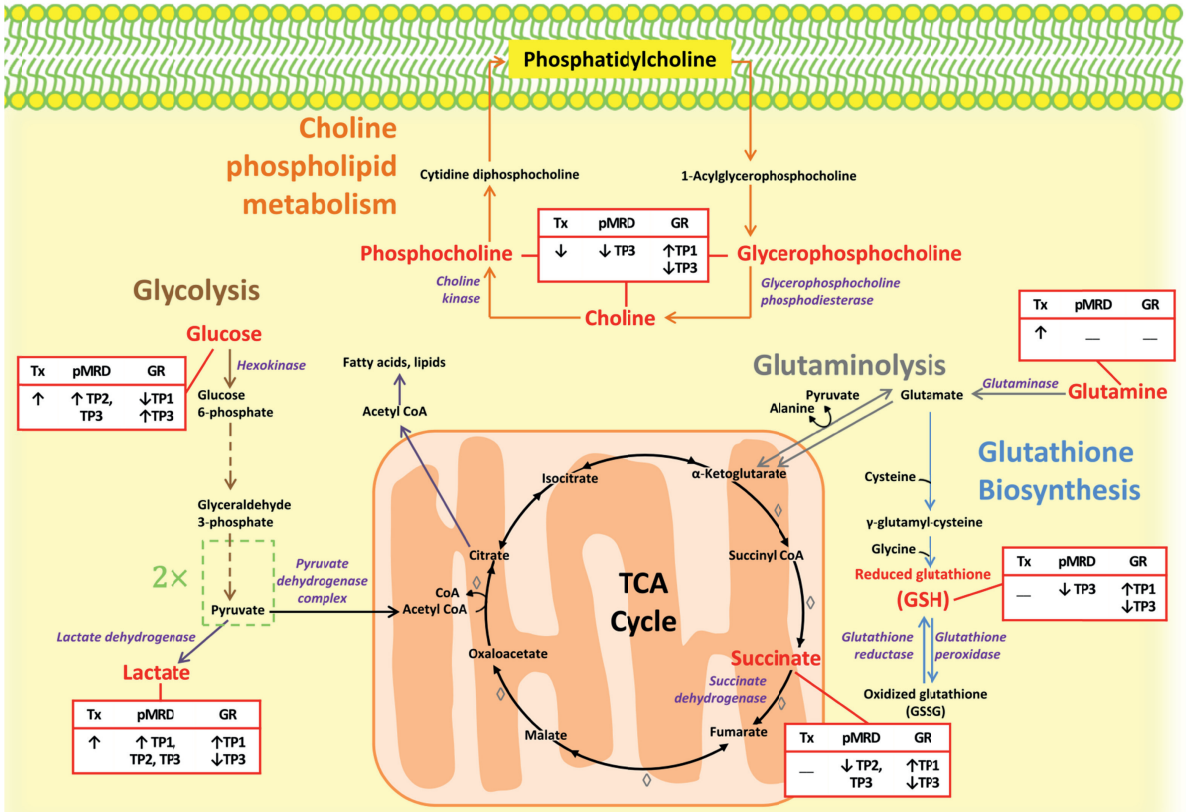


Fig. 4. Illustration of treatment effect and differences between responders and non-responders. Cancer-relevant pathways shown include glycolysis (brown), tricarboxylic acid (TCA) cycle (black), choline phospholipid metabolism (orange), glutaminolysis (gray), and glutathione biosynthesis (blue), with altered metabolites and related enzymes marked in red and purple, respectively . General trends in metabolite levels with increasing treatment time (Tx) and differences in pathological minimal residual disease patients compared to pathological non-responders (pMRD) and in the response ratio group of good response compared to the no response group (GR) at each time point (TP) indicated with arrows are shown in red boxes. TCA cycle reactions that are recruited by glutaminolysis are marked ◊.

Supplementary Information

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Supplementary File 1: Supplementary methods and references

Information on supplementary methods.

Supplementary File 2: Supplementary Figures

Supplementary Figures S1-S3.

Supplementary File 3: Supplementary Table 2

Supplementary Table 2.

Supplementary File 4: Supplementary Tables 3-8

Supplementary Tables 3-8.

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Corresponding author: Leslie R. Euceda, Department of Circulation and Medical Imaging, Faculty of Medicine, NTNU, Norwegian University of Science and Technology, Trondheim, Norway, E.mail: leslie.e.wood@ntnu.no.

Supplementary Methods

Patient treatment. All patients received neoadjuvant chemotherapy with four three-weekly cycles of anthracyclines in the form of FEC100 (5-fluorouracil 600 mg/m2, epirubicin 100 mg/m2, cyclophosphamide 600 mg/m2) followed by taxane-based therapy, (12 weekly infusion with paclitaxel 80 mg/m2 or four three-weekly cycles of docetaxel 100 mg/m2). Patients randomized to receive bevacizumab were all given 15 mg/kg every three weeks. After 12 weeks of FEC100, most patients were started on docetaxel, and received 15 mg/kg with this therapy. However, due to toxicity, a majority of the patients thereafter received paclitaxel with bevacizumab dosing 10 mg/kg every other week.

Sample handling. Biopsies were immediately snap-frozen after collection and stored at -80°C. At TP1 and TP2, three core needle biopsies were sampled. The middle section of the first biopsy was used for histopathological analysis, and approximately 4 mg were used for high resolution (HR) magic angle spinning (MAS) magnetic resonance (MR) analysis. The rest of the first needle biopsy was blended with the second needle biopsy, and divided equally for DNA and gene expression profiling. The third needle biopsy was used for protein expression analysis. At TP3, a piece between 80-100 mg was cut from the surgically removed tumor. One or two additional pieces were cut when the first piece had a mass lower than 80 mg. The two lateral sections of the first piece were used for histopathological analysis, and approximately 4 mg was used for HR MAS MR analysis. The rest of the first piece was blended, together with additional pieces, if any. 10-16 mg and 25-30 mg of the blended tumor material was used for protein expression analysis and DNA profiling, respectively, with the remaining material being used for gene expression profiling.

Acquisition and processing of gene expression microarray data. Total RNA was extracted using TRIzol® reagent (Carlsbad, CA, USA). Gene expression profiling was performed using one color Sureprint G3 Human GE 8x60k Microarrays (Agilent Technologies) according to the manufacturer's protocol (One-Color Microarray-Based Gene expression Analysis, Low Input Quick Amp Labeling, v.6.5, May 2010). The microarray data was log2 transformed, quantile normalized, and mean centered on both rows and columns.

High resolution (HR) magic angle spinning (MAS) magnetic resonance spectroscopy (MRS) experiments and spectral preprocessing. Samples were prepared within a short period (< 5min) on a dedicated work station designed to keep the samples frozen. Biopsies (4.1 ± 1.3 mg) were cut to fit into 30 μL disposable inserts containing 3.0 μL of 24.29 mM sodium formate (VWR BDH Prolabo, France) in D2O (Armar Chemicals, Switzerland) for shimming and locking purposes. Each insert was set tightly into a 4 mm o.d. MAS zirconium rotor. HR MAS MR spectra were acquired on a Bruker Avance DRX600 spectrometer (Bruker Biospin GmbH, Germany) equipped with a $^{1}H/^{13}C$ MAS probe. The run order of the samples was randomized and samples were blindly analyzed. Samples were spun at 5 kHz while kept at a temperature of 5°C to minimize tissue degradation. A spin-echo one dimensional experiment with presaturation (cpmgpr1d, Bruker BioSpin, Germany) was recorded for all samples, with effective echo time of 77 ms, a spectral width of 20 ppm (−5 to 15 ppm), and 256 scans. The free induction decays (FIDs) were multiplied by a 0.30 Hz exponential weighting function and Fourier transformed into 64K real points. Phase correction was performed automatically for each spectrum using TopSpin 3.1 (Bruker BioSpin GmbH). Further spectral preprocessing was carried out in Matlab R2013b (The Mathworks, Inc., USA). Chemical shifts were referenced to the creatine peak at 3.03 ppm. The spectral region between 1.40 – 4.70 ppm, containing the majority of low-molecular weight metabolites, was selected as the region of interest. The spectra were baseline corrected using asymmetric least squares (Eilers 2004) with parameters λ = 1e7 and p = 0.0001. Additionally, the mean value of the noise regions at 4.52-4.49, 3.10-3.06, 2.63-2.58, and 1.84-1.75 ppm were subtracted from each spectrum so that the baseline varied around zero. The spectra were peak aligned using icoshift (Savorani et al. 2010). Lipid peaks at 4.37-4.27, 2.97-2.70, 2.31-2.23, 2.11-1.92, 1.64-1.49 ppm, and the contaminant peaks for ethanol and acetone at 3.71-3.63 and 2.23-2.21 ppm, respectively, were excluded for further analysis. Regions containing peaks from lidocaine, a local anesthetic used for sampling biopsies at TP1 and TP2, at 3.38-3.30 and 2.21-2.17 ppm, were also excluded. The spectra were subsequently PQN (Dieterle et al. 2006) normalized. Normalization after lipid removal corrects for differences in sample size and tumor cell content, as it can be assumed that most of the lipid signals from breast biopsies do not originate from cancer cells but from adjacent adipose tissue.

Validation of partial least squares-discriminant analysis (PLS-DA) models. Double-layered cross validation (CV) (Westerhuis et al. 2008) was used to optimize the number of PLS LVs in the inner layer and to assess model performance in the outer layer, using 10% randomly selected samples as a test set and repeating the splitting until each sample had been used for testing once at each layer. The maximum number of LVs tested was 10. The final classification results and optimal number of LVs resulted from the average and mode, respectively, of 20 repetitions of the whole double CV procedure. An additional validation was carried out through permutation testing (Westerhuis et al. 2008) set to 1000 permutations. Models were considered significant if the final accuracy obtained from non-permuted double CV was higher than 95% of the permuted accuracy values ($p \leq 0.05$).

Lactate imputation validation. A resampling-type procedure was employed to validate the imputation of lactate (Lac) relative levels using the package for the method of multivariate imputation by chained equations (MICE) set to 10 imputations and predictive mean matching. From the original sample cohort of 325 breast tumor tissue samples analyzed with high resolution (HR) magic angle spinning (MAS) magnetic resonance (MR) spectroscopy, 40 were excluded due to high lipid content or poor spectral quality, resulting in 285 samples. Of these, 124 exhibited a peak originating from the C_1H_2 of the glycerol backbone of phospholipids and triglycerides at 4.13 ppm, which overlapped with the area integrated for Lac. The remaining 161 samples with Lac values successfully calculated were used as a training set. The imputation procedure was repeated 20 times, without sample normalization. Five per cent of the samples from the training set (n≈8) were randomly left out as a test set at each imputation repetition. This resulted in the unknown samples being predicted 20 times and each training set sample included in the test set only once. The mean of the 20 predicted values for the unknown samples was used as their final Lac value. Comparison between the predicted and real value for the training set yielded a range-normalized root mean squared error (RMSE) of 0.006 and a coefficient of determination (R^2) of 0.9996.

Evaluation of significance of interactions between LMM fixed effects. The significance of interactions between LMM fixed effects was evaluated by comparing a series of models, with the first, null model (M_0) including no interactions, and subsequent, more complex models incorporating an additional interacting factor. Models in the form $M_{A\times B}$ were the three possible single-interaction models: *time* interacting with *pMRD response* ($M_{T\times R}$) or *bevacizumab* ($M_{T\times B}$), and *pMRD response* interacting with *bevacizumab* ($M_{R\times B}$). The final, full model (M_F) incorporated all possible interactions ($M_{T\times R\times B}$). Likelihood ratio tests using the anova function from the 'stats' package (R Core Team 2014) in R 3.1.1 were employed to perform two comparisons for each single-interaction model: 1) M_0 versus $M_{A\times B}$ and 2) $M_{A\times B}$ versus M_F. The former model in both comparisons was the reduced model containing fewer interactions, while the latter was the more complex model. Differences between reduced and complex models were considered significant for multiple testing corrected p-value (q) ≤ 0.05, which indicated that the interaction excluded in the reduced model provided important information and should be included in the final LMM. If the first comparison (M_0 versus $M_{A\times B}$) did not yield a significant difference between the models, the second comparison ($M_{A\times B}$ versus M_F) was not considered significant, independently of the q-value. The comparison between M_0 and $M_{T\times R}$ was significant for seven metabolites, while M_0 versus $M_{T\times B}$ was significant for glutathione (GSH) only, and M_0 versus $M_{R\times B}$ for none (Supplementary Table 1). No significant q-values were detected for $M_{A\times B}$ versus M_F in those eight cases with significant M_0 versus $M_{A\times B}$. Therefore, no three-factor interactions were found to be significant.

Supplementary Table 1. q-values resulting from likelihood ratio tests to compare LMMs to determine interaction significance.

	Time point × pMRD response		Time point × bevacizumab		pMRD response × bevacizumab	
Metabolite	**M_0 vs $M_{T\times R}$**	**$M_{T\times R}$ vs M_F**	**M_0 vs $M_{T\times B}$**	**$M_{T\times B}$ vs M_F**	**M_0 vs $M_{R\times B}$**	**$M_{R\times B}$ vs M_F**
Glc	****9.28E-05**	7.86E-01	2.56E-01	*1.99E-03	7.82E-01	*8.05E-04
Asc	4.40E-01	7.86E-01	6.44E-01	4.97E-01	7.82E-01	4.24E-01
Lac	7.11E-02	7.86E-01	5.37E-01	1.95E-01	7.82E-01	1.41E-01
Tyr	****9.40E-06**	7.86E-01	6.74E-01	*1.68E-04	7.82E-01	*3.54E-04
mI	8.73E-02	7.86E-01	5.37E-01	3.17E-01	7.82E-01	2.38E-01
Gly	1.59E-01	7.86E-01	5.37E-01	3.67E-01	7.82E-01	3.61E-01
Tau	****3.49E-04**	7.86E-01	4.64E-01	*1.08E-02	7.82E-01	*3.79E-03
GPC	6.71E-01	7.86E-01	5.37E-01	7.96E-01	7.82E-01	7.59E-01
PCh	****4.33E-02**	7.86E-01	2.52E-01	2.72E-01	7.82E-01	7.63E-02
Cho	****1.91E-02**	7.86E-01	6.74E-01	9.77E-02	7.82E-01	1.50E-01
Cr	4.32E-01	9.51E-01	6.74E-01	7.96E-01	7.82E-01	7.77E-01
GSH	****1.52E-02**	2.45E-01	****1.80E-02**	1.89E-01	7.82E-01	*3.43E-03
Gln	2.60E-01	9.51E-01	8.36E-01	6.90E-01	7.82E-01	7.59E-01
Succ	****1.25E-02**	7.86E-01	5.37E-01	9.77E-02	7.82E-01	7.54E-02
Glu	7.22E-01	7.86E-01	5.37E-01	7.59E-01	7.82E-01	7.59E-01
Ala	6.86E-01	2.45E-01	7.04E-02	7.24E-01	7.82E-01	9.84E-02

Significant q-value (≤ 0.05) and interaction is indicated with ** and bold values. Significant q-value but not significant interaction is indicated with *. M_0: null model without interactions; M_F: full model with all interactions; ×: interaction; T: factor *time point*; R: factor *pMRD response*; B: factor *bevacizumab*; Glc:glucose; Asc:ascorbate; Lac:lactate; Tyr:tyrosine; mI:myo-inositol; Gly:glycine; Tau:taurine; GPC:glycerophosphocholine; PCh:phosphocholine; Cho:choline; Cr:creatine; GSH:glutathione; Gln:glutamine; Succ:succinate; Glu:glutamate; Ala:alanine.

Supplementary References

Dieterle, F., Ross, A., Schlotterbeck, G., & Senn, H. (2006). Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. *Analytical Chemistry, 78*(13), 4281-4290, doi:10.1021/ac051632c.

Eilers, P. H. C. (2004). Parametric Time Warping. *Analytical Chemistry, 76*(2), 404-411, doi:10.1021/ac034800e.

R Core Team (2014). R: A language and environment for statistical computing. Viena, Austria. URL http://www.R-project.org/: R Foundation for Statistical Computing.

Savorani, F., Tomasi, G., & Engelsen, S. B. (2010). icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *Journal of Magnetic Resonance, 202*(2), 190-202, doi:http://dx.doi.org/10.1016/j.jmr.2009.11.012.

Westerhuis, J., Hoefsloot, H. J., Smit, S., Vis, D., Smilde, A., van Velzen, E. J., et al. (2008). Assessment of PLSDA cross validation. *Metabolomics, 4*(1), 81-89, doi:10.1007/s11306-007-0099-6.

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Corresponding author: Leslie R. Euceda, Department of Circulation and Medical Imaging, Faculty of Medicine, NTNU, Norwegian University of Science and Technology, Trondheim, Norway, E.mail: leslie.e.wood@ntnu.no.

Supplementary Figures

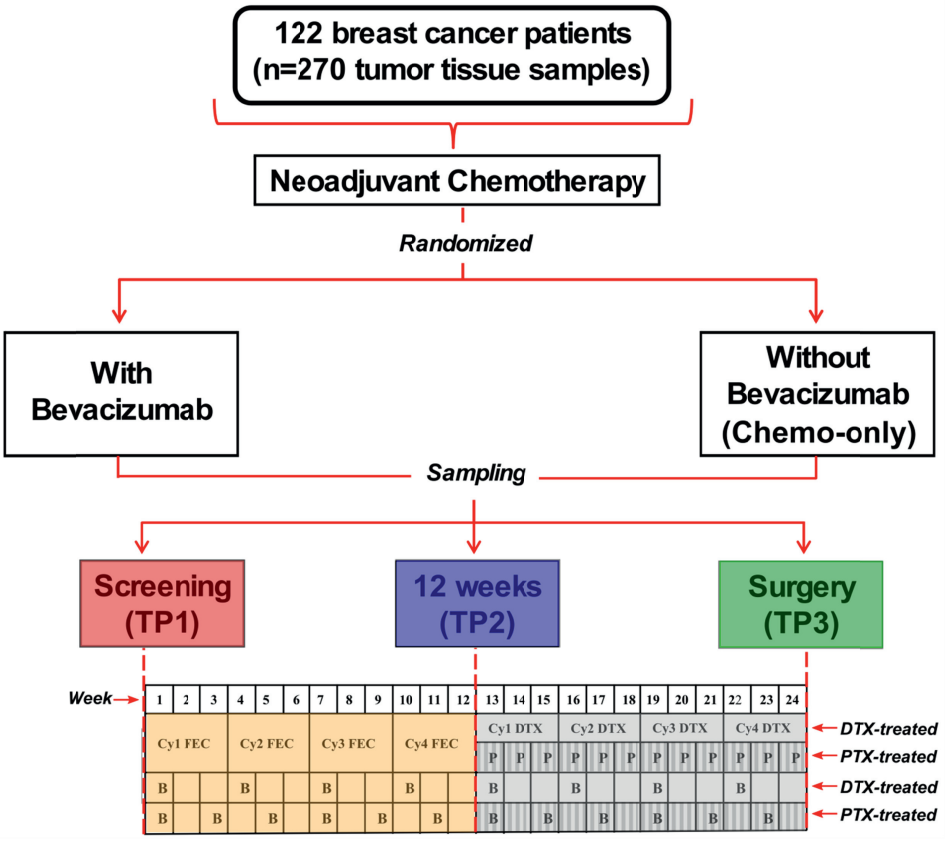


Fig. S1. Study Design. 122 patients receiving neoadjuvant chemotherapy (Chemo) were randomized to receive bevacizumab or not. Tissue biopsies were sampled at three time points

(TP): before (screening), during (after 12 weeks), and after treatment (during surgery). The lower box illustrates the treatment schedule: four three-weekly cycles (Cy) of FEC followed by 12 weeks of taxane-based treatment. For those patients receiving bevacizumab (B), this was administered once every three weeks and every other week for docetaxel (DTX) and paclitaxel (PTX) receivers, respectively. P:Paclitaxel administration.

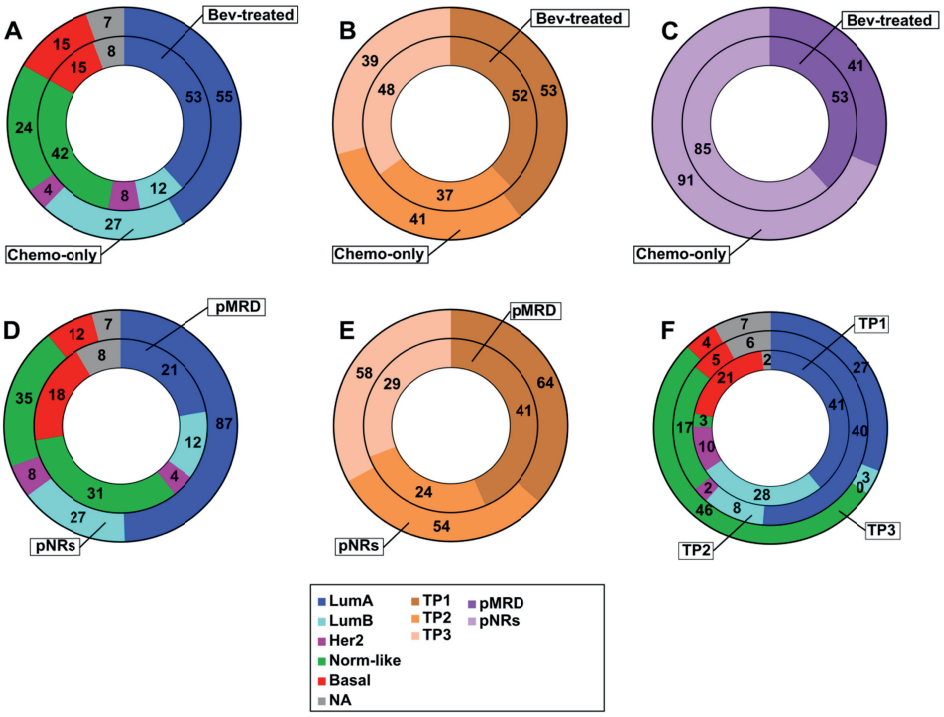


Fig. S2. Sample composition of the cohort (n=270). Each chart shows number of patients encompassing a pair of characteristics: bevacizumab randomization group combined with (A) gene expression subtype, (B) time point, and (C) pathological minimal residual disease (pMRD) response group; pMRD response group combined with (D) gene expression subtype and (E) time point; and (F) time point combined with gene expression subtype. Each ring in the charts represents an individual level of one characteristic with different shading indicating the numerical proportion of samples from the levels of the second characteristic within it. Chemo-only=chemotherapy only; Bev-treated=bevacizumab-treated; pNRs=pathological non-responders; LumA=luminal A; LumB=luminal B; Norm-like=normal-like; NA=not available.

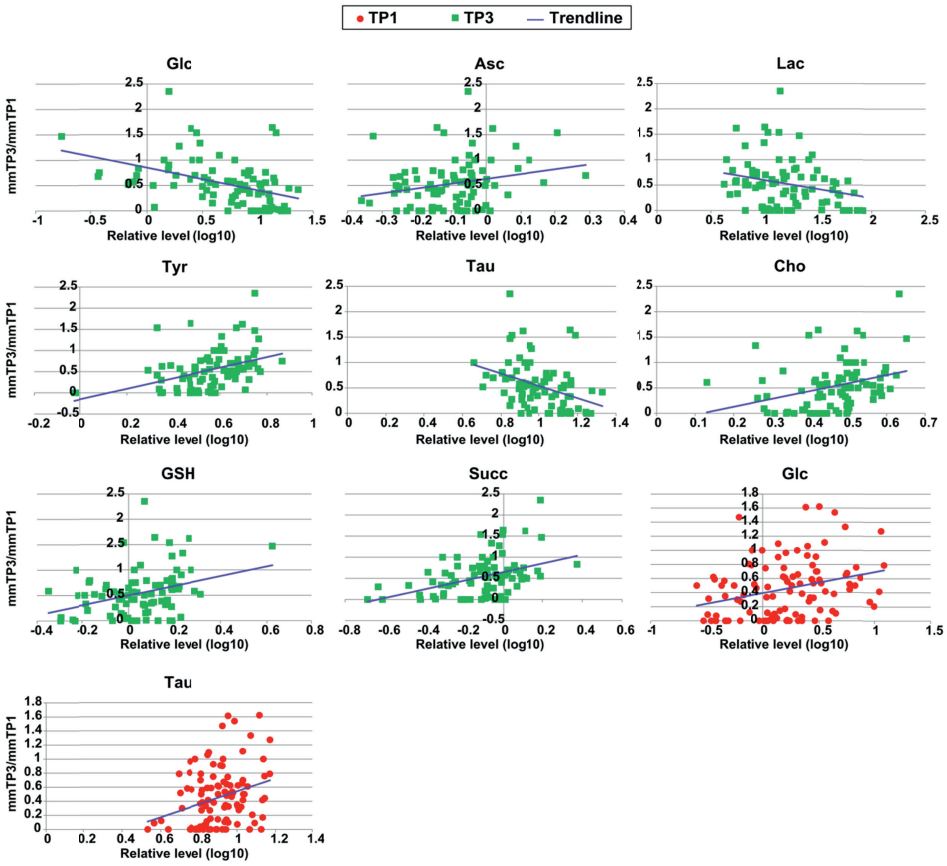


Fig. S3. Correlation trend between metabolite relative levels and response ratio. Only significantly correlated metabolites (q≤0.05) are shown. Glc=glucose; Asc=ascorbate; Lac=lactate; Tyr=tyrosine; Tau=taurine; Cho=choline; GSH=glutathione; Succ=succinate.

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Supplementary Table 2. Relative metabolite levels for each sample after log10 transformation. Samples are ordered according to time point (TP).

		Identification Level:	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		Human Metabolome Database ID	HMDB00122	HMDB00044	HMDB00190	HMDB00158	HMDB00211	HMDB00123	HMDB00251	HMDB00086	Not available	HMDB00097	HMDB00064	HMDB00125	HMDB00641	HMDB00254	HMDB03339	HMDB00161	
Patient Number	Time Point	Bevacizumab randomization group	Glucose	Ascorbate	Lactate	Tyrosine	Myo-inositol	Glycine	Taurine	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Alanine	
1	TP1	Chemotherapy only	0.0459	-0.0012	0.6017	0.7364	0.6295	0.5594	0.8960	0.5628	0.8136	0.6575	0.4178	0.1241	0.2511	-0.0906	0.7173	0.1974	
2	TP1	Bevacizumab treated	0.1823	0.1429	0.2516	0.7898	0.6135	0.3175	1.0048	0.7315	0.9565	0.6285	0.4864	0.0707	0.1435	-0.0860	0.6058	0.0745	
3	TP1	Bevacizumab treated	0.0845	-0.2176	1.4180	0.6092	0.6700	0.5888	0.8388	0.9588	1.1291	0.4961	0.3337	0.3085	0.1552	-0.3152	1.0070	0.2815	
4	TP1	Bevacizumab treated	-0.4841	0.4560	0.6901	0.7018	0.5204	0.7279	0.9354	0.4806	0.8812	0.5018	0.6106	0.4596	0.3679	0.0498	0.6535	-0.0028	
5	TP1	Chemotherapy only	0.1257	0.0322	0.7799	0.7195	0.5332	0.5057	0.8095	0.6322	0.9466	0.4714	0.5811	0.0489	0.3760	0.0181	0.6769	0.2974	
6	TP1	Bevacizumab treated	0.1373	-0.0063	0.6303	0.7283	0.5622	0.7306	0.8542	0.6125	1.0041	0.6335	0.4204	0.1943	0.2214	-0.0853	0.7395	0.2760	
7	TP1	Bevacizumab treated	-0.4238	-0.0754	1.0415	0.7475	0.7199	0.4614	0.6387	1.1811	0.9733	0.7166	0.4156	0.2403	-0.1144	0.0274	0.6626	0.0811	
8	TP1	Chemotherapy only	0.3974	-0.0359	0.4174	0.8236	0.5339	0.6952	0.8482	0.7840	1.0831	0.5826	0.7150	0.0087	0.1512	0.2147	0.3150	0.2738	
9	TP1	Chemotherapy only	0.8274	0.0034	0.3191	0.5544	0.4674	0.5291	1.1428	0.7649	0.9218	0.4723	0.4294	0.0114	0.2997	-0.1553	0.4558	0.1106	
10	TP1	Chemotherapy only	0.4095	0.0697	0.3745	0.7431	0.6322	0.5383	1.0094	0.5203	0.7713	0.5877	0.2120	0.2170	0.3099	-0.0293	0.6984	0.1680	
11	TP1	Bevacizumab treated	0.5479	-0.0252	0.7477	0.6633	0.6108	0.6656	1.0301	0.5359	0.7010	0.4597	0.4014	0.2102	0.2525	-0.3051	0.6791	0.2902	
12	TP1	Chemotherapy only	-0.3376	0.0984	0.6342	0.8239	0.4978	0.4384	1.0008	0.7414	0.8842	0.5500	0.4222	0.2873	0.2730	-0.2185	0.6761	-0.0268	
13	TP1	Bevacizumab treated	0.6377	-0.0335	0.6139	0.6867	0.6637	0.4805	0.9870	0.5654	0.9412	0.4045	0.5044	0.1376	0.3242	-0.1389	0.4937	0.2217	
14	TP1	Bevacizumab treated	1.0592	0.0096	1.0074	0.6279	0.5807	0.3396	1.1705	0.6186	0.7015	0.4122	0.3229	0.1644	0.3740	-0.2217	0.7731	0.4202	
15	TP1	Chemotherapy only	0.9962	-0.0431	1.2221	0.6226	0.5495	0.4814	1.0791	0.7440	0.9524	0.3962	0.3881	-0.1022	0.1703	-0.2045	0.6962	0.3352	
16	TP1	Bevacizumab treated	-0.1724	-0.0774	0.6368	0.8439	0.6038	0.9404	0.8099	0.6591	1.0981	0.6033	0.4622	-0.0364	0.3350	0.0629	0.9142	0.4982	
17	TP1	Bevacizumab treated	0.7304	-0.0409	0.5154	0.6883	0.6727	0.5837	1.0707	0.5498	0.8011	0.4186	0.3570	0.3951	0.2657	-0.0882	0.7316	0.2148	
18	TP1	Chemotherapy only	0.4730	-0.0731	0.4543	0.7502	0.7100	0.5016	0.8117	0.4824	1.0128	0.4817	0.4303	0.1290	0.6610	0.0057	0.6210	0.3524	
19	TP1	Bevacizumab treated	0.4207	0.0407	0.6876	0.7008	0.6590	0.4115	1.0037	0.5376	0.7706	0.6311	0.4302	0.1575	0.2082	-0.0050	0.7199	0.1628	
20	TP1	Bevacizumab treated	-0.0747	0.0967	1.3434	0.5796	0.5857	0.4365	0.7012	0.9217	0.7865	0.3341	0.5141	0.2661	0.4026	-0.0563	0.7252	0.1770	
21	TP1	Bevacizumab treated	0.1028	0.1739	0.8220	0.7710	0.4993	0.8243	0.9545	0.4347	0.7313	0.4984	0.7610	0.3943	0.0621	-0.0125	0.8040	0.2793	
22	TP1	Bevacizumab treated	0.4410	0.0559	0.6902	0.6939	0.6170	0.6672	0.8131	0.5526	0.8520	0.4973	0.5431	0.2349	0.3906	0.0998	0.5822	0.3125	
23	TP1	Chemotherapy only	0.3300	-0.0056	0.7301	0.7335	0.6960	0.3665	0.9401	0.7340	0.7462	0.5462	0.4021	0.2463	0.1590	-0.1712	0.5634	0.1138	
24	TP1	Chemotherapy only	-0.4315	-0.1280	0.9168	0.9181	0.4721	0.6128	0.8663	0.9784	0.7597	0.5939	0.5877	0.2837	0.2919	0.0647	0.5984	0.2603	
25	TP1	Bevacizumab treated	0.3950	-0.0663	0.2573	0.7271	0.7535	0.4279	0.8775	0.6349	0.7402	0.5907	0.5266	0.2120	0.3772	-0.0963	0.6590	0.2437	
26	TP1	Chemotherapy only	0.3816	0.0537	0.3477	0.7148	0.6743	0.4532	0.9539	0.7208	0.7954	0.5080	0.3798	0.1785	0.3014	-0.0223	0.5906	0.1170	
27	TP1	Bevacizumab treated	-0.0308	0.2201	0.8501	0.9339	0.7064	0.7328	0.8047	0.7340	1.0668	0.8831	0.3807	0.1664	-0.0906	0.0116	0.9219	0.2487	
28	TP1	Chemotherapy only	0.2575	0.0680	0.7590	0.7386	0.5937	0.5216	1.0385	0.7040	0.8405	0.3777	0.3586	-0.0789	0.3635	0.0028	0.7417	0.1850	
29	TP1	Chemotherapy only	-0.2137	-0.1852	1.1398	0.9087	0.3832	0.6391	0.9248	0.7073	0.6269	0.8270	0.3007	0.1832	0.1415	-0.0071	1.1490	0.4952	
30	TP1	Bevacizumab treated	0.5084	0.1485	1.3107	0.5886	0.7036	0.4591	0.8629	0.5812	0.8885	0.3305	0.3229	0.2810	0.2477	-0.2794	0.7322	0.3699	
31	TP1	Bevacizumab treated	-0.4609	-0.0341	0.8791	0.7753	0.5147	0.7776	0.5323	0.7728	0.9294	0.4444	0.5354	0.3518	0.1381	0.1181	0.8447	0.2714	
32	TP1	Chemotherapy only	-0.2578	-0.0253	0.8136	0.6660	0.6879	0.2475	0.8151	0.8019	0.6702	0.5674	0.4866	0.2231	0.2979	0.0357	0.7122	0.2390	
33	TP1	Chemotherapy only	0.6186	-0.0186	0.5136	0.6454	0.6864	0.5133	0.9455	0.5929	0.8114	0.4937	0.3443	0.1434	0.3219	-0.0235	0.6684	0.3877	
34	TP1	Bevacizumab treated	-0.1274	-0.1614	1.0292	0.8496	0.4275	0.7041	0.7517	0.6457	0.7322	0.7908	0.4586	0.1975	-0.2044	0.0681	0.9367	0.2619	
35	TP1	Bevacizumab treated	0.7367	-0.0750	1.6919	0.5040	0.6025	0.5777	0.9461	0.5144	0.5332	0.3749	0.2953	0.2705	0.2270	-0.5297	1.0236	0.4325	
36	TP1	Bevacizumab treated	0.2054	0.0655	1.3898	0.6030	0.5052	0.4531	0.7543	0.6586	0.9526	0.4365	0.5290	0.1970	0.0909	-0.0379	0.7956	0.2990	
37	TP1	Bevacizumab treated	0.6467	0.0337	1.1732	0.6214	0.6845	0.4674	0.9368	0.4432	0.6791	0.6617	0.3530	0.0262	0.3401	-0.2528	0.7274	0.3210	
38	TP1	Bevacizumab treated	0.1967	-0.0909	0.7411	0.6899	0.8837	0.3809	1.1333	0.3592	0.6586	0.5249	0.4105	0.2637	0.2938	0.0002	0.9019	0.1927	
39	TP1	Chemotherapy only	-0.0566	0.0338	0.7683	0.7027	0.7464	0.3256	0.9721	0.6079	0.8728	0.4729	0.4019	0.3136	0.2609	0.0072	0.6699	0.2771	
40	TP1	Chemotherapy only	-0.1983	-0.1428	0.7115	0.8306	0.6392	0.5919	0.8168	0.7011	1.1580	0.5747	0.4867	-0.3203	0.3484	0.0737	0.6137	0.0910	
41	TP1	Chemotherapy only	0.4343	-0.0272	0.6218	0.8106	0.6183	0.5949	0.8460	0.6255	0.8332	0.5477	0.3937	0.1405	0.3778	0.0531	0.5653	0.3264	
42	TP1	Bevacizumab treated	-0.2032	-0.0013	0.7223	0.7120	0.7591	0.4041	0.7836	0.9001	1.2244	0.5614	0.5928	0.2341	0.2808	-0.0929	0.7166	0.0832	
43	TP1	Chemotherapy only	-0.3283	0.1948	0.7767	0.8406	0.4740	0.6960	0.9392	0.4111	0.9411	0.6009	0.5655	0.3517	0.0671	0.0059	0.6951	0.1248	
44	TP1	Bevacizumab treated	0.2367	0.0689	0.3630	0.7115	0.6622	0.4304	0.9248	0.5072	0.8994	0.5340	0.4435	0.2116	0.3958	0.0354	0.6716	0.2640	
45	TP1	Bevacizumab treated	-0.3339	-0.0207	1.1654	0.8018	0.3364	0.3610	0.9434	0.3094	0.7936	0.6376	0.4675	0.3458	0.1883	-0.2528	0.6934	-0.5619	
46	TP1	Bevacizumab treated	0.0245	0.0618	0.3256	0.7455	0.4550	0.3113	0.9522	0.7816	0.9269	0.3919	0.6413	0.2434	0.3460	-0.0882	0.5667	0.0951	
47	TP1	Chemotherapy only	0.0312	-0.1088	1.4446	0.7490	0.6108	0.4639	0.8631	0.5978	0.9451	0.6773	0.3747	0.2855	0.0094	0.0591	0.8127	0.3449	
48	TP1	Bevacizumab treated	0.2407	-0.0840	0.6046	0.7849	0.2923	0.1980	1.0313	0.8371	0.7978	0.5321	0.4182	0.3487	0.2910	-0.0056	0.6807	0.1565	
49	TP1	Chemotherapy only	0.2919	-0.0924	0.6975	0.7320	0.3439	0.4752	1.0917	0.7328	0.6722	0.6036	0.3969	0.4290	0.3326	0.0570	0.9225	0.2342	
50	TP1	Chemotherapy only	0.4212	-0.0861	0.6379	0.6315	0.6270	0.4235	1.0217	0.5208	0.5752	0.5255	0.3932	0.2235	0.4218	0.0378	0.9118	0.3458	
51	TP1	Bevacizumab treated	0.4710	-0.0297	1.0460	0.5994	0.6749	0.4400	0.9148	0.5083	0.8796	0.4002	0.5981	0.1313	0.3716	-0.1288	0.6321	0.3918	
52	TP1	Chemotherapy only	0.4487	-0.0390	0.5086	0.7272	0.6134	0.5479	0.9373	0.4463	0.6970	0.5329	0.5498	0.1620	0.3432	-0.1198	0.6899	0.3591	
53	TP1	Chemotherapy only	0.7457	-0.0919	1.8729	0.5033	0.6983	0.3794	0.9317	0.6928	0.7207	0.5395	0.3474	0.1528	-0.5534	-1.9498	1.2228	0.4891	
54	TP1	Chemotherapy only	0.0344	0.2297	1.2731	0.7525	0.4666	0.4456	0.8379	0.6151	0.8355	0.5089	0.6237	0.1676	0.3504	-0.2489	0.6515	0.3474	

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Supplementary Table 2. Relative metabolite levels for each sample after log10 transformation. Samples are ordered according to time point (TP).

		Identification Level:	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		Human Metabolome Database ID	HMDB00122	HMDB00044	HMDB00190	HMDB00158	HMDB00211	HMDB00123	HMDB00251	HMDB00086	Not available	HMDB00097	HMDB00064	HMDB00125	HMDB00641	HMDB00254	HMDB03339	HMDB00161	
Patient Number	Time Point	Bevacizumab randomization group	Glucose	Ascorbate	Lactate	Tyrosine	Myo-inositol	Glycine	Taurine	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Alanine	
55	TP1	Chemotherapy only	-0.3686	-0.0880	0.8866	0.6689	0.7794	0.3927	0.9529	0.7570	1.2167	0.4890	0.6360	0.1852	0.2510	0.0175	0.7221	0.1622	
56	TP1	Bevacizumab treated	-0.3368	0.0138	0.7581	0.8889	0.4591	0.4508	0.8920	0.8824	0.7716	0.5480	0.5552	0.3390	0.1503	-0.1058	0.7898	0.3536	
57	TP1	Chemotherapy only	0.2127	-0.1089	1.4448	0.7271	0.5362	0.6332	0.8538	0.7890	0.7199	0.4975	0.4675	0.0542	0.5681	0.0396	0.9622	0.4775	
58	TP1	Bevacizumab treated	0.3466	0.1149	0.5659	0.7330	0.6153	0.5176	0.8293	0.6079	0.8508	0.5063	0.6123	0.1094	0.3264	0.0060	0.6287	0.2764	
59	TP1	Chemotherapy only	0.5014	0.1381	0.5389	0.7917	0.5423	0.4575	1.1162	0.5453	0.6901	0.5459	0.6625	0.3495	0.2857	-0.0828	0.8416	0.2521	
60	TP1	Chemotherapy only	0.3058	0.1448	1.5898	0.5697	0.6112	0.4676	0.7826	0.7718	1.0959	0.3107	0.5035	0.2803	0.1929	-0.4915	0.9225	0.4798	
61	TP1	Chemotherapy only	0.2186	0.1389	0.7145	0.7225	0.4502	0.6288	0.8188	0.8760	0.7789	0.6352	0.4512	0.1743	0.4539	0.0458	0.6681	0.1710	
62	TP1	Bevacizumab treated	1.0853	-0.0648	0.4803	0.6515	0.3426	0.4326	1.1697	0.7521	0.8356	0.4691	0.4002	0.2872	0.4987	-0.2819	0.4419	0.2865	
63	TP1	Chemotherapy only	0.4820	0.0266	0.4121	0.8255	0.5891	0.4152	1.0307	0.6680	0.8605	0.6399	0.4018	0.1613	0.1765	-0.1259	0.7417	0.2071	
64	TP1	Chemotherapy only	0.1371	-0.0076	0.4227	0.7109	0.6327	0.7969	0.9244	0.6499	0.6498	0.5265	0.4401	0.1790	0.1545	-0.0808	0.5926	0.0517	
65	TP1	Bevacizumab treated	-0.4192	-0.1147	0.7557	0.7841	0.7899	0.5388	0.8267	0.5416	1.2107	0.5248	0.3748	0.2267	0.5140	0.1318	0.8612	0.4101	
66	TP1	Bevacizumab treated	0.3485	-0.0207	0.3787	0.7758	0.7635	0.4822	0.8209	0.6767	0.8420	0.4854	0.3271	0.1384	0.2129	-0.2136	0.6046	0.3028	
67	TP1	Chemotherapy only	0.3477	-0.0084	0.2928	0.7157	0.7012	0.3686	0.8911	0.6589	0.8825	0.6534	0.5742	0.1023	0.2698	-0.1172	0.5203	0.1312	
68	TP1	Chemotherapy only	0.7688	-0.0123	0.4751	0.6983	0.6034	0.5552	0.9791	0.6266	0.6781	0.5170	0.2397	-0.1775	0.3216	-0.1753	0.6922	0.3173	
69	TP1	Bevacizumab treated	0.0739	-0.0838	0.5744	0.7570	0.4393	0.5419	0.7606	0.7382	0.8331	0.5849	0.4815	0.4019	0.0720	-0.0284	0.8504	0.0600	
70	TP1	Chemotherapy only	-0.1183	0.2716	0.5488	0.9994	0.2924	0.6936	0.6019	0.4704	0.7650	0.5708	0.4249	0.3583	0.5838	-0.0581	0.5669	0.3961	
71	TP1	Bevacizumab treated	-0.0045	0.4082	0.6343	0.9071	0.5836	0.2551	1.1368	0.5979	0.9785	0.5469	0.6190	0.1830	0.0156	0.0020	0.4189	0.1511	
72	TP1	Bevacizumab treated	0.1877	0.0703	0.9396	0.7905	0.9134	0.5760	0.8770	0.7494	1.3303	0.6639	0.2108	0.1368	0.2905	-0.1938	0.6813	0.0664	
73	TP1	Bevacizumab treated	0.5025	-0.0473	1.2902	0.6344	0.6079	0.3934	0.7371	0.7663	0.7622	0.3822	0.4005	0.3732	0.4341	-0.0102	0.8847	0.3910	
74	TP1	Bevacizumab treated	-0.5903	0.0899	0.8033	0.7213	0.5796	0.5108	0.9021	0.9738	0.8734	0.5696	0.8904	-0.2431	0.0341	0.1266	0.4829	0.1990	
75	TP1	Chemotherapy only	-0.1003	0.0382	1.0990	0.7929	0.5730	0.4188	0.6942	0.6104	0.8688	0.4948	0.4900	0.1779	0.4663	-0.0402	0.6746	0.2826	
76	TP1	Chemotherapy only	0.3113	-0.0420	0.7298	0.6953	0.5472	0.4555	1.0273	0.7878	0.9321	0.4760	0.6308	0.1805	0.4574	0.0881	0.6737	0.1221	
77	TP1	Bevacizumab treated	0.5176	0.1392	0.9453	0.7994	0.4775	0.3388	0.9754	0.6011	0.8520	0.4966	0.3683	-0.0376	0.3873	-0.0837	0.6192	0.1257	
78	TP1	Chemotherapy only	0.8176	-0.0519	1.1373	0.5214	0.2270	0.4011	1.0330	0.7395	0.7501	0.5872	0.3070	0.0734	0.3245	-0.0194	0.6957	0.2810	
79	TP1	Chemotherapy only	0.8675	-0.1286	0.7596	0.6512	0.3859	0.5212	1.0551	0.6401	0.7959	0.4787	0.3725	0.1939	0.5412	-0.2373	0.6351	0.1715	
80	TP1	Chemotherapy only	0.4623	0.1047	0.6842	0.9425	0.3984	0.5717	1.0657	0.5827	0.9472	0.6872	0.2290	0.3801	0.2887	-0.0912	0.8783	0.2570	
81	TP1	Chemotherapy only	1.0427	-0.1985	0.9128	0.5183	0.3226	0.5270	1.1320	0.6191	0.7252	0.5047	0.3967	0.1386	0.5042	-0.1261	0.5587	0.4804	
82	TP1	Bevacizumab treated	0.7697	-0.0345	0.2877	0.8007	0.6020	0.5165	1.1434	0.5700	0.8371	0.3649	0.3946	0.1463	0.4226	-0.1308	0.6302	0.1711	
83	TP1	Chemotherapy only	-0.4925	-0.1503	0.8670	0.7830	0.6674	0.4963	0.5653	1.1767	1.0908	0.4952	0.5097	0.0837	0.3671	0.1542	0.6954	0.1893	
84	TP1	Chemotherapy only	0.2672	-0.0139	0.8122	0.6725	0.6089	0.3582	0.9646	0.6047	1.0506	0.4282	0.4030	0.2601	0.3932	-0.0658	0.6783	0.1466	
85	TP1	Chemotherapy only	-0.5299	-0.1031	1.0377	0.9942	0.7478	0.2899	0.7891	0.9631	1.0673	0.8022	0.4516	0.2080	-0.3250	0.0471	0.5906	0.0567	
86	TP1	Bevacizumab treated	0.0120	-0.1353	1.0542	0.8401	0.4538	0.5948	0.7514	0.4994	0.5964	0.5689	0.6341	0.3297	0.5232	0.0198	0.9387	0.4495	
87	TP1	Bevacizumab treated	0.0446	-0.0052	0.6984	0.7237	0.5496	0.6596	0.9465	0.4453	0.7679	0.5158	0.4481	0.2301	0.3660	0.1503	0.6903	0.2333	
88	TP1	Chemotherapy only	-0.4170	-0.1272	0.8213	0.7946	0.5629	0.3668	1.0017	0.3497	0.9088	0.5818	0.4888	0.2501	0.2751	0.0904	0.8480	0.2189	
89	TP1	Chemotherapy only	-0.4419	-0.1851	0.8577	0.9207	0.3646	0.8627	0.9007	0.7378	0.7009	0.7746	0.5593	0.1944	0.4654	0.1903	0.7460	0.2654	
90	TP1	Chemotherapy only	-0.0970	0.0776	0.7890	0.7854	0.6089	0.4905	0.9282	0.5081	0.7706	0.5122	0.4582	0.3582	0.3138	-0.0091	0.7795	0.2452	
91	TP1	Chemotherapy only	0.9578	-0.0567	1.1648	1.1582	0.7107	0.4376	0.8598	0.6551	0.8352	0.4602	0.3489	0.2584	0.1738	-0.2378	0.6944	0.2021	
92	TP1	Bevacizumab treated	0.0887	0.1633	0.5714	0.9387	0.3500	0.1813	0.9415	0.8245	0.7176	0.5804	0.4055	0.2439	0.3385	0.0560	0.4900	0.0620	
93	TP1	Bevacizumab treated	0.1979	0.0324	0.5686	0.7628	0.4681	0.4397	0.9054	0.7617	0.7161	0.5089	0.4368	0.2008	0.4787	0.0867	0.6904	0.2450	
94	TP1	Bevacizumab treated	0.3137	0.0419	0.4746	0.9094	0.7624	0.4394	1.1259	0.6056	1.1071	0.6276	0.2844	0.1762	0.1584	-0.1423	0.5264	0.0599	
95	TP1	Chemotherapy only	-0.3449	-0.0768	1.0947	0.8125	0.7464	0.4510	0.7624	0.8456	0.9713	0.7675	0.2910	-0.0079	-0.2025	0.0389	0.7711	0.2480	
96	TP1	Bevacizumab treated	-0.3094	-0.0366	0.8265	0.7416	0.5801	0.5016	0.7597	0.6322	1.3177	0.8315	0.3978	0.2457	-0.0084	0.0362	0.8834	0.1121	
97	TP1	Bevacizumab treated	0.5543	-0.0133	0.6902	0.7477	0.5142	0.4583	0.8507	0.7799	0.8132	0.4846	0.3541	0.3103	0.3023	-0.5220	0.7067	0.3213	
98	TP1	Bevacizumab treated	0.1517	0.0355	0.6251	0.7256	0.5684	0.6561	0.9012	0.5545	0.9418	0.4444	0.5134	0.1924	0.5096	-0.0038	0.6791	0.3349	
99	TP1	Chemotherapy only	0.5825	0.0271	1.3650	0.5563	0.6749	0.3982	0.7867	0.5354	0.8347	0.3796	0.4275	0.2429	0.2997	-0.3459	0.6301	-0.2888	
100	TP1	Bevacizumab treated	-0.2295	-0.0406	0.8464	0.8366	0.4968	0.9732	0.7096	0.4424	0.8630	0.6449	0.4292	0.3486	0.2149	0.1678	0.9081	0.3094	
101	TP1	Bevacizumab treated	0.4498	-0.0257	0.3196	0.8201	0.5733	0.6309	0.9562	0.5234	0.8014	0.4742	0.3846	0.1851	0.5495	-0.1199	0.7938	0.3094	
102	TP1	Chemotherapy only	0.0416	0.1741	0.3700	0.8062	0.5148	0.8291	0.8442	0.5288	0.9339	0.5692	0.4177	0.1838	0.4017	-0.1945	0.8064	0.3726	
103	TP1	Bevacizumab treated	0.6313	-0.0599	1.0963	0.7095	0.6279	0.4434	0.8797	0.5596	0.7131	0.4013	0.3982	0.4503	0.3468	-0.3576	0.4586	0.3979	
104	TP1	Bevacizumab treated	0.2178	-0.0427	0.3387	0.8692	0.6827	0.5541	0.7893	0.7826	1.0294	0.5652	0.2712	0.2402	0.3416	-0.1366	0.7869	0.0683	
105	TP1	Chemotherapy only	0.0947	0.0179	0.4614	0.7445	0.8194	0.4651	0.7594	0.7741	1.1718	0.5206	0.4015	0.0983	0.1232	-0.0559	0.6869	0.0893	
106	TP2	Chemotherapy only	0.6647	0.2389	1.3825	0.6248	0.6053	0.4094	0.9534	0.7193	0.6640	0.4885	0.3848	0.0765	0.4582	-0.1559	0.4701	0.4268	
1	TP2	Chemotherapy only	0.6044	-0.1288	1.4248	0.5316	0.6729	0.4879	0.7897	0.4662	0.6085	0.5491	0.4085	0.1378	0.1981	-0.1885	0.9127	0.3822	
2	TP2	Bevacizumab treated	0.3375	0.0728	0.8884	0.7180	0.7703	0.4085	0.9322	0.6093	0.8565	0.4340	0.4767	0.1541	0.1382	-0.1464	0.6591	0.3204	

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Supplementary Table 2. Relative metabolite levels for each sample after log10 transformation. Samples are ordered according to time point (TP).

		Identification Level:	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		Human Metabolome Database ID	HMDB00122	HMDB00044	HMDB00190	HMDB00158	HMDB00211	HMDB00123	HMDB00251	HMDB00086	Not available	HMDB00097	HMDB00064	HMDB00125	HMDB00641	HMDB00254	HMDB03339	HMDB00161	
Patient Number	Time Point	Bevacizumab randomization group	Glucose	Ascorbate	Lactate	Tyrosine	Myo-inositol	Glycine	Taurine	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Alanine	
107	TP2	Bevacizumab treated	0.7165	-0.1625	1.7528	0.5977	0.7449	0.5616	0.9111	0.5421	0.5800	0.4111	0.3178	0.2246	-0.0617	-0.3373	1.0804	0.3615	
108	TP2	Chemotherapy only	-0.3313	-0.0221	0.8969	0.6977	0.5990	0.4904	0.5567	0.6000	0.8874	0.5335	0.5142	0.0392	0.1736	0.0797	0.8219	0.3042	
7	TP2	Bevacizumab treated	0.2560	0.0696	0.8099	0.7046	0.7339	0.5710	0.8420	0.6756	0.6143	0.6045	0.3060	0.3399	0.2334	-0.0383	0.7418	0.0941	
8	TP2	Chemotherapy only	0.8672	-0.1490	1.5328	0.5601	0.6362	0.5404	0.9735	0.5748	0.6809	0.3989	0.3410	0.1472	0.2560	-0.6734	0.8983	0.3934	
9	TP2	Chemotherapy only	0.9715	-0.1003	0.4638	0.5934	0.5251	0.5649	1.0837	0.6676	0.7884	0.6013	0.2656	-0.2950	0.4661	-0.1583	0.4789	0.3440	
10	TP2	Chemotherapy only	0.4235	0.0289	1.2264	0.7237	0.7071	0.4806	0.8313	0.6669	0.8595	0.4958	0.3654	0.2317	0.1992	-0.0968	0.6944	0.2859	
109	TP2	Bevacizumab treated	0.4675	-0.0774	0.5146	0.6465	0.6874	0.3605	0.9337	0.6641	0.8121	0.4969	0.4413	0.1252	0.2924	-0.0590	0.6883	0.3216	
110	TP2	Chemotherapy only	-0.3490	-0.0386	0.8170	0.7748	0.6148	0.4907	0.8414	0.9522	0.9734	0.4789	0.7042	0.2975	0.2569	0.1184	0.6899	0.2105	
12	TP2	Chemotherapy only	0.1138	-0.0329	0.9568	0.8224	0.5630	0.4168	1.0115	0.8426	0.8105	0.4830	0.4808	0.3332	0.2108	-0.2968	0.6202	0.1471	
111	TP2	Bevacizumab treated	0.6911	-0.0165	1.0975	0.6582	0.6523	0.4483	0.9450	0.5990	0.6225	0.5615	0.3939	0.0280	0.1985	-0.1254	0.8272	0.2151	
15	TP2	Chemotherapy only	0.6257	-0.0402	1.4649	0.5726	0.7014	0.4695	0.8815	0.4264	0.5364	0.5456	0.3402	0.2397	0.1527	-0.2015	0.8524	0.3056	
112	TP2	Chemotherapy only	0.9902	-0.0334	0.9042	0.5680	0.6061	0.5524	1.0181	0.5022	0.6801	0.2841	0.3090	0.2872	0.4825	-0.3446	0.5210	0.3415	
17	TP2	Bevacizumab treated	0.4072	-0.0230	1.0358	0.6335	0.6714	0.9764	1.1344	0.4838	0.7476	0.4140	0.3756	0.4666	0.3488	-0.0913	0.7699	0.1988	
19	TP2	Bevacizumab treated	0.9766	-0.0467	0.8254	0.6276	0.7266	0.5244	1.0203	0.6013	0.7085	0.4965	0.3248	0.1627	0.3825	-0.2818	0.7351	0.2660	
21	TP2	Bevacizumab treated	0.0615	0.0833	0.7745	0.8228	0.6277	0.4755	0.9514	0.5921	0.7069	0.5894	0.3889	0.2384	-0.0380	-0.0116	0.7258	0.1398	
23	TP2	Chemotherapy only	0.3597	-0.0423	0.9990	0.6924	0.7019	0.4277	0.9319	0.6919	0.7555	0.4930	0.4264	0.2782	0.4114	-0.0810	0.6445	0.2354	
24	TP2	Chemotherapy only	-0.1160	-0.1176	0.6889	0.8355	0.3544	0.6382	0.9526	0.8177	0.7622	0.6061	0.5207	0.2986	0.3151	-0.0366	0.6132	0.0530	
28	TP2	Chemotherapy only	0.6489	0.0237	0.5077	0.7182	0.6249	0.5201	1.0209	0.7197	0.7698	0.4775	0.3717	0.1433	0.3085	-0.1748	0.6297	0.2212	
29	TP2	Chemotherapy only	0.0693	-0.1701	0.9848	0.7749	0.4265	0.3865	0.9389	0.8428	0.8614	0.7483	0.1532	0.6076	0.2233	-0.0791	1.1283	0.3453	
113	TP2	Bevacizumab treated	0.5187	-0.0591	0.8857	0.6773	0.6912	0.6316	0.9120	0.5720	0.8878	0.4654	0.4488	0.1673	0.4483	-0.0964	0.7265	0.3521	
33	TP2	Chemotherapy only	0.5847	-0.0821	1.4747	0.5141	0.7436	0.8167	0.7920	0.6697	0.6642	0.4207	0.3555	0.3462	0.2312	-0.2266	0.6877	0.2881	
114	TP2	Chemotherapy only	0.2971	-0.0478	0.1743	0.7523	0.8230	0.2171	1.0495	0.6096	0.9474	0.5043	0.3774	0.0473	0.0681	-0.0350	0.4681	0.1205	
36	TP2	Bevacizumab treated	0.7827	-0.0349	0.6743	0.7075	0.6020	0.3745	1.0906	0.5017	0.7217	0.5001	0.3607	0.1561	0.4992	0.0110	0.6250	0.2445	
115	TP2	Bevacizumab treated	0.8165	-0.1914	1.9086	0.3437	0.5914	0.6439	0.8457	0.6167	0.7272	0.3075	0.4998	0.4020	-0.0675	-0.3143	1.2645	0.4241	
37	TP2	Bevacizumab treated	0.7442	-0.1783	1.8506	0.4767	0.6556	0.2730	0.9500	0.5941	0.6914	0.3487	0.4149	0.5531	0.2450	-0.9759	1.0888	-0.0308	
39	TP2	Chemotherapy only	0.0706	0.2658	0.5513	0.7524	0.7149	0.3502	0.8641	0.6527	0.9438	0.3939	0.5033	0.3811	0.2068	0.0541	0.5573	0.2220	
40	TP2	Chemotherapy only	-0.5611	-0.0913	0.7300	0.8232	0.6657	0.5358	0.6634	0.7171	1.0601	0.5765	0.4825	-0.3695	0.3974	0.1129	0.5134	0.1796	
41	TP2	Chemotherapy only	0.2842	-0.0349	0.4579	0.7407	0.8734	0.4714	0.9100	0.4331	0.8564	0.4488	0.4307	0.1800	0.2990	0.1230	0.4746	0.0646	
42	TP2	Bevacizumab treated	0.5062	-0.0410	0.8614	0.6922	0.7380	0.3676	0.8606	0.6517	0.7988	0.4890	0.2932	0.1919	0.1420	-0.3190	0.7251	0.2096	
46	TP2	Bevacizumab treated	0.5274	-0.0598	0.7780	0.5737	0.8248	0.6001	0.9722	0.4326	0.7160	0.4012	0.4003	0.2288	0.5012	-0.1014	0.7822	0.2865	
48	TP2	Bevacizumab treated	0.6771	-0.0608	0.7860	0.8122	0.4214	0.4681	1.1219	0.7937	0.7334	0.3647	0.3944	0.2997	0.3996	-0.0696	0.5831	0.2144	
50	TP2	Chemotherapy only	0.2670	-0.1024	0.3340	0.6633	0.7404	0.4640	0.9021	0.4981	0.6812	0.5038	0.5526	0.2266	0.4610	0.0177	0.7080	0.2467	
51	TP2	Bevacizumab treated	0.4207	-0.1353	1.7874	0.4489	0.8159	0.3653	0.7214	0.7910	0.7529	0.4298	0.7029	0.3591	0.2387	-0.1444	0.9778	0.3473	
52	TP2	Chemotherapy only	0.4866	-0.1810	1.4298	0.5964	0.6392	0.4701	0.7839	0.4139	0.7685	0.3581	0.5517	0.2154	0.3783	-0.3188	0.6970	0.4906	
54	TP2	Chemotherapy only	0.4881	0.0978	1.3261	0.6246	0.5786	0.3182	0.9314	0.5946	0.7745	0.2835	0.4363	0.1781	0.3087	-0.2982	0.7897	0.3402	
55	TP2	Chemotherapy only	-0.0238	-0.0744	0.4741	0.7039	0.8252	0.3553	0.8425	0.7709	1.1689	0.5908	0.4868	0.2748	0.1377	-0.1030	0.6325	0.0376	
116	TP2	Bevacizumab treated	0.5050	0.0193	1.4225	0.5621	0.6104	0.1909	0.9248	0.7496	0.5959	0.2837	0.5618	0.2516	0.1137	-0.2616	0.6910	0.2662	
58	TP2	Bevacizumab treated	0.7779	-0.3087	1.1843	0.8580	0.5366	0.2634	0.9349	0.5846	0.9056	0.2825	1.2177	0.2781	0.6588	-0.0551	0.7016	0.2025	
59	TP2	Chemotherapy only	0.6934	-0.0249	0.3774	0.7116	0.6018	0.6244	0.9408	0.4514	0.6591	0.4548	0.4605	0.1274	0.3106	-0.1437	0.7105	0.3420	
60	TP2	Chemotherapy only	-0.2713	0.0974	0.6631	0.8281	0.7093	0.4689	0.8107	0.8781	1.1959	0.4990	0.5368	0.2262	0.2907	0.0360	0.5943	0.2523	
62	TP2	Bevacizumab treated	0.9059	-0.0754	0.5284	0.5035	0.5937	0.4815	1.0539	0.6421	0.7246	0.5609	0.3571	-0.0361	0.4805	-0.2685	0.5222	0.3942	
63	TP2	Chemotherapy only	0.0029	-0.1074	0.8959	0.7125	0.7462	0.4640	0.8808	0.3664	0.7677	0.5009	0.4102	0.1862	0.3089	-0.0512	0.8253	0.3768	
65	TP2	Bevacizumab treated	1.0002	0.1629	0.7416	0.2754	0.5822	0.4299	1.1260	0.5464	0.6342	0.7050	0.4273	-0.2692	0.1898	-0.3492	0.8490	0.4345	
66	TP2	Bevacizumab treated	0.0928	0.0811	0.5122	0.6877	0.8031	0.6592	0.6621	0.5794	0.9183	0.3932	0.5857	0.2415	0.3178	0.0065	0.5291	0.2249	
67	TP2	Chemotherapy only	0.2165	0.1286	0.4412	0.7873	0.7680	0.5241	0.8656	0.5885	1.0458	0.4652	0.6823	0.1814	0.2938	-0.0598	0.5030	0.2672	
68	TP2	Chemotherapy only	0.5159	-0.0188	0.9087	0.6111	0.6500	0.5514	0.8692	0.5952	0.6539	0.3449	0.4007	0.2013	0.2203	-0.3724	0.7870	0.2888	
70	TP2	Chemotherapy only	0.6661	-0.0912	1.5556	0.4561	0.4050	0.4253	0.7722	0.5799	0.8124	0.2979	0.4980	0.1978	0.0993	-0.4336	0.8983	0.5596	
71	TP2	Bevacizumab treated	-0.1767	0.3530	0.5919	0.8648	0.6683	0.2763	1.0921	0.6019	0.9383	0.5555	0.5979	0.1795	0.1639	-0.1222	0.4630	0.1921	
72	TP2	Bevacizumab treated	0.7377	-0.0596	0.9115	0.5977	0.5979	0.5337	0.9406	0.6270	0.8639	0.4172	0.2756	0.0108	0.4524	-0.2423	0.6148	0.2383	
73	TP2	Bevacizumab treated	0.1818	-0.0700	0.9061	0.7338	0.5847	0.4699	0.6326	0.6953	0.7316	0.5247	0.5640	0.2086	0.3625	-0.0170	0.7096	0.3213	
74	TP2	Bevacizumab treated	0.4740	0.0261	0.5668	0.6976	0.4447	0.4544	1.1201	0.7467	0.8030	0.5089	0.6799	0.0085	0.2465	-0.1771	0.6928	0.2192	
75	TP2	Chemotherapy only	0.3760	0.0221	1.4533	0.5596	0.6205	0.3510	0.7367	0.6371	0.7976	0.3804	0.6104	0.3127	0.3222	-0.2146	0.6369	0.3849	
117	TP2	Chemotherapy only	1.0686	-0.0160	0.5248	0.5204	0.7707	0.7681	1.1421	0.5636	0.6149	0.3591	0.3198	-0.0304	0.4395	-0.1099	0.4123	0.3200	
76	TP2	Chemotherapy only	0.6367	-0.2321	0.4294	0.6554	0.6232	0.2183	1.1691	0.5433	0.6668	0.4029	0.4155	0.1916	0.2218	-0.3786	0.8189	0.2976	

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Supplementary Table 2. Relative metabolite levels for each sample after log10 transformation. Samples are ordered according to time point (TP).

		Identification Level:	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		Human Metabolome Database ID	HMDB00122	HMDB00044	HMDB00190	HMDB00158	HMDB00211	HMDB00123	HMDB00251	HMDB00086	Not available	HMDB00097	HMDB00064	HMDB00125	HMDB00641	HMDB00254	HMDB03339	HMDB00161	
Patient Number	Time Point	Bevacizumab randomization group	Glucose	Ascorbate	Lactate	Tyrosine	Myo-inositol	Glycine	Taurine	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Alanine	
78	TP2	Chemotherapy only	-0.0564	0.1017	0.8472	0.8959	0.4712	0.6078	0.8956	0.9667	0.9798	0.7612	0.3323	0.0719	0.4487	0.0808	0.6760	0.2982	
81	TP2	Chemotherapy only	1.2016	-0.0544	1.6588	0.5564	0.5812	0.4780	1.2006	0.6791	0.7668	0.4319	0.3658	-0.0466	0.4040	-0.7767	0.6264	0.3564	
82	TP2	Bevacizumab treated	0.9168	-0.1839	0.9239	0.5557	0.3143	0.5170	1.0306	0.5411	0.7613	0.0294	0.4724	0.1891	0.4866	-0.3444	0.6178	0.2562	
83	TP2	Chemotherapy only	0.0332	-0.0005	0.7995	0.7691	0.5498	0.4070	0.8269	0.8859	0.9965	0.6723	0.4293	0.2881	0.2501	-0.0471	0.7256	0.1695	
84	TP2	Chemotherapy only	0.3429	-0.0513	0.6423	0.6917	0.6449	0.2441	0.9985	0.7895	1.0626	0.6636	0.4731	0.0817	0.2950	0.1914	0.7285	0.2425	
86	TP2	Bevacizumab treated	0.6574	0.0189	0.9386	0.6380	0.4050	0.4672	0.9545	0.4641	0.5234	0.5007	0.3448	0.2490	0.4513	-0.2617	0.7738	0.3869	
87	TP2	Bevacizumab treated	-0.3239	0.0088	0.8149	0.6999	0.6937	0.6888	0.9374	0.4120	0.8688	0.4313	0.5529	0.2586	0.5579	0.2000	0.6246	0.2271	
89	TP2	Chemotherapy only	0.2682	0.0012	0.5092	0.9622	0.0063	0.8103	0.9177	0.6729	0.6836	0.6029	0.6756	0.0365	0.3838	0.0720	0.6712	0.2458	
90	TP2	Chemotherapy only	0.4667	-0.0324	0.5763	0.6782	0.7393	0.5073	0.8767	0.5679	0.6702	0.5702	0.3680	0.1759	0.3485	-0.1290	0.7630	0.2853	
91	TP2	Chemotherapy only	0.1170	0.0477	0.4619	0.8469	0.6741	0.3508	1.0233	0.5835	0.8706	0.6122	0.1615	0.2541	0.0767	-0.0899	0.8117	0.0898	
92	TP2	Bevacizumab treated	0.6179	0.0320	0.5046	0.8507	0.6680	0.2266	0.8951	0.8614	0.7430	0.4853	0.3281	0.2511	0.3076	-0.1213	0.5693	0.1868	
93	TP2	Bevacizumab treated	0.5526	0.0782	0.9751	0.6574	0.5250	0.3821	1.0291	0.5394	0.5447	0.3540	0.4196	0.2019	0.4457	-0.0829	0.6975	0.2563	
94	TP2	Bevacizumab treated	0.5084	-0.2084	1.8007	0.2277	0.6542	0.2087	0.7188	0.6652	0.6019	0.2731	0.4722	0.3677	0.0954	-0.7215	1.2043	0.1628	
97	TP2	Bevacizumab treated	0.2508	0.0034	1.3455	0.7327	0.6055	0.3350	0.7967	0.5874	0.6849	0.6420	0.3833	0.1833	0.0446	-0.3243	0.9466	0.3372	
98	TP2	Bevacizumab treated	0.7169	-0.1132	1.2778	0.6313	0.6334	0.4308	1.0228	0.4475	0.6549	0.4333	0.3660	0.2736	0.3090	-0.2962	0.8373	0.3754	
99	TP2	Chemotherapy only	0.6881	-0.1090	1.4204	0.6109	1.0357	0.4737	0.8698	0.6210	0.7085	0.3390	0.3227	0.2602	0.3998	-0.2664	0.5273	0.3426	
100	TP2	Bevacizumab treated	0.0490	0.0060	0.6521	0.8222	0.7467	0.7311	0.9077	0.4626	0.7738	0.5681	0.3287	0.2205	0.3046	0.0828	0.9549	0.1828	
101	TP2	Bevacizumab treated	0.8598	-0.1262	1.1383	0.5551	0.5633	0.5426	0.9628	0.6114	0.6463	0.2706	0.3151	0.0399	0.3826	-0.3702	0.6852	0.3721	
102	TP2	Chemotherapy only	0.4642	0.1192	1.0645	0.6567	0.6133	0.6882	0.9214	0.6311	0.8410	0.3863	0.4276	0.3620	0.4337	-0.2270	0.5549	0.3423	
103	TP2	Bevacizumab treated	0.7895	-0.1532	1.7681	0.4991	0.5601	0.4551	0.9176	0.6531	0.7903	0.2368	0.4225	0.4306	0.2684	-0.4875	0.8021	0.5177	
104	TP2	Bevacizumab treated	0.1374	-0.0387	0.6714	0.9481	0.7418	0.5782	0.7923	0.6672	0.9653	0.5617	0.2545	0.3146	0.4379	-0.1483	0.8120	0.2755	
106	TP3	Chemotherapy only	1.0986	-0.3596	1.0626	0.4711	0.6455	0.5200	1.1277	0.5626	0.6842	0.5038	0.2822	-0.1762	0.5573	-0.0929	0.4852	0.3910	
2	TP3	Bevacizumab treated	0.9500	-0.1991	1.5544	0.3484	0.5528	0.5508	1.0259	0.5665	0.6418	0.5349	0.3706	-0.0978	0.3664	-0.2939	0.9356	0.4149	
118	TP3	Chemotherapy only	0.1493	0.0138	1.2561	0.5746	0.7205	0.1898	0.6579	0.8143	0.9562	0.5399	0.5174	0.2699	0.0007	-0.2332	0.6170	0.2181	
108	TP3	Chemotherapy only	0.5378	-0.1271	1.6303	0.4412	0.7092	0.4164	0.8148	0.4259	0.3737	0.5670	0.4183	0.1671	0.2674	-0.0113	0.8863	0.4057	
3	TP3	Bevacizumab treated	1.1120	-0.2191	1.1146	0.4774	0.3496	0.5855	1.1201	0.7620	0.7060	0.3852	0.3990	-0.2968	0.4965	-0.3526	0.7037	0.3438	
4	TP3	Bevacizumab treated	0.5965	0.0621	0.6764	0.7237	0.7069	0.4633	0.9059	0.4891	0.7547	0.4196	0.4087	0.1581	0.4272	0.0666	0.6942	0.2684	
5	TP3	Chemotherapy only	1.1331	-0.1551	1.0275	0.4688	0.5735	0.5163	1.1567	0.7020	0.7463	0.4967	0.3600	-0.0182	0.6151	-0.0886	0.6357	0.4165	
6	TP3	Bevacizumab treated	0.6500	-0.0489	1.4406	0.6052	0.7052	0.5921	0.8948	0.5125	0.7381	0.4769	0.3759	0.0509	0.3439	-0.0075	0.8404	0.8708	
7	TP3	Bevacizumab treated	1.0839	-0.0553	1.2115	0.5800	0.6085	0.5350	1.1549	0.5129	0.7013	0.4939	0.3481	-0.0692	0.5638	-0.2032	0.6702	0.4289	
9	TP3	Chemotherapy only	1.0593	-0.2344	0.9121	0.5172	0.5984	0.5078	1.1447	0.5649	0.7200	0.5078	0.3567	-0.1698	0.5579	-0.0475	0.5357	0.3586	
10	TP3	Chemotherapy only	0.5000	-0.0428	1.3808	0.5773	0.7330	0.5210	0.8535	0.4588	0.6783	0.4965	0.3422	0.1840	0.2433	-0.0822	0.7313	0.6837	
109	TP3	Bevacizumab treated	0.2541	-0.1285	0.8883	0.7191	0.7135	0.6083	0.7576	0.5555	0.6599	0.5072	0.4203	-0.0954	0.3611	0.0192	0.6897	0.4117	
110	TP3	Chemotherapy only	0.1929	-0.0518	1.1453	0.7485	0.4982	0.5158	0.8439	0.6968	0.8941	0.6327	0.5809	0.0680	0.3312	0.1901	0.6795	0.2861	
13	TP3	Bevacizumab treated	1.1530	-0.1213	1.1342	0.3277	0.3935	0.6983	1.1892	0.6667	0.6626	0.3958	0.2817	-0.0212	0.6420	-0.1152	0.9109	0.3046	
14	TP3	Bevacizumab treated	0.2831	0.0852	0.8176	0.7648	0.6824	0.3014	0.9613	0.5004	0.7810	0.4882	0.4335	0.1863	0.3096	-0.0222	0.6589	0.2421	
111	TP3	Bevacizumab treated	1.1413	-0.2341	1.6357	0.4246	0.4522	0.4621	1.1668	0.5966	0.6789	0.4294	0.4104	0.0194	0.5236	-0.2630	0.8394	0.5252	
119	TP3	Chemotherapy only	0.4479	0.2070	1.0328	0.6699	0.7117	0.5326	0.8545	0.5608	0.8385	0.5374	0.3755	0.1752	0.3607	-0.0056	0.6975	0.2172	
16	TP3	Bevacizumab treated	1.0256	-0.0363	1.3735	0.5679	0.6278	0.5616	1.0873	0.5371	0.6186	0.4824	0.3829	0.1077	0.5308	-0.2557	0.6866	0.3546	
112	TP3	Chemotherapy only	1.1185	-0.1409	1.0035	0.4723	0.5404	0.4939	1.1617	0.6007	0.7337	0.4201	0.3220	0.1120	0.6753	-0.0030	0.4018	0.3201	
17	TP3	Bevacizumab treated	0.4783	-0.0397	0.9976	0.6073	0.8208	0.4223	0.9501	0.3193	0.5568	0.2579	0.3744	0.2321	0.3906	-0.0611	0.6569	0.2580	
18	TP3	Chemotherapy only	0.9035	-0.0408	0.8962	0.6444	0.7944	0.5976	1.0493	0.3926	0.6722	0.4742	0.3977	0.1363	0.5409	-0.0824	0.6343	0.2886	
19	TP3	Bevacizumab treated	1.0877	-0.1405	0.7460	0.6652	0.7531	0.4351	1.1626	0.5946	0.7582	0.4901	0.4103	-0.1836	0.5561	-0.0822	0.6964	0.3846	
20	TP3	Bevacizumab treated	0.6169	-0.0521	0.6232	0.6363	0.7672	0.5869	0.9030	0.5050	0.8257	0.5015	0.3867	-0.1461	0.3545	-0.0818	0.5942	0.2176	
22	TP3	Bevacizumab treated	0.9115	-0.0106	0.8857	0.6762	0.7387	0.4733	1.0757	0.5621	0.6583	0.4946	0.3738	-0.1667	0.3279	-0.0955	0.7401	0.4460	
23	TP3	Chemotherapy only	0.8309	-0.1154	1.0766	0.6170	0.7586	0.4795	0.9222	0.6569	0.6086	0.4883	0.3898	0.0970	0.5000	-0.0242	0.6210	0.2639	
28	TP3	Chemotherapy only	0.2350	0.0241	0.9535	0.7718	0.6103	0.5598	0.9826	0.6924	0.8830	0.4921	0.3694	-0.1882	0.2867	0.0423	0.6503	0.1807	
29	TP3	Chemotherapy only	-0.7706	-0.3258	1.3207	0.7481	0.3421	0.4110	0.8448	0.8792	0.7816	0.6519	0.2163	0.6228	0.4089	0.1955	1.1742	0.4507	
113	TP3	Bevacizumab treated	0.9316	-0.1537	1.8558	0.3427	0.6617	0.6138	0.9464	0.5976	0.6071	0.4094	0.4593	0.0132	0.2831	-0.6117	1.2095	0.5386	
30	TP3	Bevacizumab treated	0.5889	-0.3365	1.5334	0.5606	0.5799	0.4001	0.8444	0.7618	0.8657	0.4938	0.4856	0.0727	0.0483	-0.0849	0.8834	0.4727	
31	TP3	Bevacizumab treated	1.0137	-0.1844	1.2514	0.4141	0.5915	0.5841	1.1131	0.4841	0.6411	0.4278	0.3424	-0.2984	0.4393	0.0076	0.5915	0.3797	
120	TP3	Bevacizumab treated	-0.1645	-0.1145	1.1054	0.6831	0.6175	0.4784	0.9351	0.6161	1.0176	0.5646	0.6373	0.1446	0.3034	0.1902	0.6501	0.2386	
34	TP3	Bevacizumab treated	0.1875	-0.0663	1.2772	0.7192	0.4421	0.7096	0.8037	0.5977	0.6817	0.5194	0.4888	0.1831	0.3781	0.0532	0.9197	0.4643	
114	TP3	Chemotherapy only	-0.1213	-0.0737	0.8210	0.7613	0.7637	0.3868	0.9544	0.4739	1.0090	0.5471	0.4873	-0.0636	0.1480	0.0041	0.4449	0.1042	

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Supplementary Table 2. Relative metabolite levels for each sample after log10 transformation. Samples are ordered according to time point (TP).

		Identification Level:	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		Human Metabolome Database ID	HMDB00122	HMDB00044	HMDB00190	HMDB00158	HMDB00211	HMDB00123	HMDB00251	HMDB00086	Not available	HMDB00097	HMDB00064	HMDB00125	HMDB00641	HMDB00254	HMDB03339	HMDB00161	
Patient Number	Time Point	Bevacizumab randomization group	Glucose	Ascorbate	Lactate	Tyrosine	Myo-inositol	Glycine	Taurine	Glycerophosphocholine	Phosphocholine	Choline	Creatine	Glutathione	Glutamine	Succinate	Glutamate	Alanine	
35	TP3	Bevacizumab treated	1.1316	-0.0915	1.4523	0.3852	0.3425	0.5116	1.1659	0.7487	0.5905	0.2756	0.3627	-0.0004	0.4500	-0.3794	0.7376	0.4133	
37	TP3	Bevacizumab treated	0.8119	-0.0008	1.6785	0.4114	0.6878	0.3920	0.9378	0.5435	0.7641	0.4995	0.4744	-0.0210	0.2330	-0.4217	1.1077	0.2466	
38	TP3	Bevacizumab treated	0.7571	-0.0759	1.5751	0.5714	0.7648	0.4417	0.9202	0.5134	0.5959	0.3909	0.4241	0.0832	0.1966	0.0207	1.0596	0.4551	
41	TP3	Chemotherapy only	0.7227	-0.0382	0.7354	0.6900	0.8308	0.5782	0.9305	0.5929	0.6750	0.4430	0.3787	-0.3504	0.3900	-0.0408	0.7078	0.0813	
44	TP3	Bevacizumab treated	0.9435	-0.1927	1.7276	0.3469	0.6286	0.5016	1.0478	0.6773	0.6927	0.4731	0.3640	0.0306	0.2865	-0.1679	0.9128	0.4029	
46	TP3	Bevacizumab treated	0.8723	-0.0438	0.9289	0.5564	0.6237	0.4742	1.0389	0.5800	0.7414	0.4380	0.3252	0.0141	0.4562	-0.1565	0.5712	0.4871	
121	TP3	Chemotherapy only	-0.0929	0.2886	0.9274	0.7454	0.6863	0.3464	1.0537	0.7750	1.0997	0.5837	0.3327	0.1153	0.2274	0.0620	0.6128	0.1351	
48	TP3	Bevacizumab treated	1.3487	-0.2291	1.4400	0.3278	0.3939	0.6736	1.3284	0.5363	0.9049	0.4540	0.3699	-0.0404	0.4313	-0.3554	0.5555	0.6858	
49	TP3	Chemotherapy only	1.2265	-0.0855	1.3091	0.3367	0.3405	0.6745	1.2046	0.8258	0.6279	0.2803	0.4039	-0.0785	0.6423	-0.3023	0.5210	0.4677	
122	TP3	Chemotherapy only	1.0605	0.1681	0.7327	0.5561	0.6149	0.5317	1.1309	0.5528	0.7834	0.4229	0.3684	0.0362	0.6094	-0.0783	0.6022	0.3243	
53	TP3	Chemotherapy only	-0.4333	-0.2484	1.2034	0.6790	0.6857	0.4069	0.8214	0.8685	0.9810	0.4831	0.8965	0.2173	0.1827	0.1760	0.7586	0.2933	
54	TP3	Chemotherapy only	1.0891	-0.1497	1.6796	0.4357	0.5702	0.4426	1.1511	0.6305	0.6709	0.2875	0.3454	-0.0102	0.5536	-0.4791	0.9874	0.4238	
55	TP3	Chemotherapy only	0.8759	-0.2599	0.8209	0.5333	0.5892	0.4041	1.1152	0.6101	0.7647	0.5038	0.4163	0.0673	0.5539	0.1206	0.6587	0.3598	
56	TP3	Bevacizumab treated	1.2029	-0.0630	1.0568	0.5101	0.5544	0.5938	1.2529	0.6179	0.7799	0.4279	0.3772	0.0121	0.7129	-0.1982	0.5857	0.5560	
116	TP3	Bevacizumab treated	-0.0774	0.0912	1.2933	0.5481	0.7587	0.2788	0.8784	0.6961	0.6760	0.3292	0.6034	0.1880	0.1666	0.3738	0.5232	0.3988	
58	TP3	Bevacizumab treated	1.0568	-0.1356	0.9767	0.5061	0.6206	0.5163	1.1125	0.5166	0.6928	0.4659	0.2331	-0.0029	0.5343	-0.0751	0.6420	0.2903	
59	TP3	Chemotherapy only	0.3926	0.0185	0.7344	0.6956	0.6731	0.5238	0.9123	0.4391	0.7871	0.5203	0.5657	0.2615	0.2882	0.1109	0.6552	0.2492	
60	TP3	Chemotherapy only	0.7868	-0.0891	0.6448	0.6177	0.7588	0.5242	0.9214	0.5662	0.8053	0.5038	0.3932	-0.2293	0.4232	-0.0536	0.6702	0.3314	
61	TP3	Chemotherapy only	0.5725	-0.0521	0.7274	0.6992	0.6601	0.4687	0.9070	0.6687	0.8736	0.4301	0.4404	0.0160	0.3932	-0.0097	0.6193	0.2177	
62	TP3	Bevacizumab treated	0.9220	-0.1521	0.7009	0.5957	0.6578	0.4649	1.0393	0.6384	0.7648	0.5942	0.4563	0.0765	0.5335	-0.2313	0.5703	0.2640	
63	TP3	Chemotherapy only	0.4332	-0.0935	1.0101	0.6562	0.7858	0.4593	0.8542	0.3713	0.8512	0.5145	0.3781	0.1905	0.3576	0.0254	0.7542	0.3393	
64	TP3	Chemotherapy only	0.1912	-0.0493	0.8438	0.7443	0.7348	0.5868	0.9364	0.6417	0.6143	0.5978	0.4059	0.2258	0.2780	-0.0816	0.5370	0.1319	
65	TP3	Bevacizumab treated	1.2009	-0.0304	0.9975	0.4874	0.5761	0.5858	1.2469	0.5777	0.7560	0.4222	0.3349	0.0353	0.6163	-0.2144	0.5928	0.3042	
66	TP3	Bevacizumab treated	0.8105	-0.1310	1.1971	0.6328	0.7087	0.5349	0.9088	0.6007	0.7744	0.4243	0.3454	0.1417	0.4188	-0.1646	0.6817	0.3598	
67	TP3	Chemotherapy only	1.0822	-0.0715	1.0029	0.4683	0.6137	0.4651	1.1642	0.6275	0.7556	0.4420	0.3726	0.0248	0.5405	-0.2002	0.5787	0.3250	
68	TP3	Chemotherapy only	-0.1675	-0.1597	0.9819	0.6574	0.5549	0.5624	0.7044	0.5690	0.7919	0.5837	0.6033	0.3136	0.4021	0.1058	0.8470	0.2917	
69	TP3	Bevacizumab treated	1.1773	-0.0122	1.0401	0.5510	0.6059	0.6203	1.2437	0.8496	0.7244	0.3408	0.2961	-0.2459	0.6704	-0.0957	0.6942	0.2244	
71	TP3	Bevacizumab treated	0.4731	0.1245	1.1352	0.7493	0.6093	0.4979	0.9541	0.5322	0.6799	0.4930	0.5266	0.0244	0.1949	-0.3064	0.4898	0.2877	
72	TP3	Bevacizumab treated	-0.4229	-0.1287	1.1714	0.8645	0.8992	0.5266	0.7211	1.1130	1.3523	0.6253	0.5717	0.0290	0.2760	0.1001	0.5427	0.2114	
73	TP3	Bevacizumab treated	0.7741	-0.0556	1.7699	0.4953	0.7899	0.4481	0.8929	0.6623	0.7212	0.4704	0.3784	0.2876	0.1778	-0.1474	1.0829	0.5394	
74	TP3	Bevacizumab treated	0.0533	-0.0030	0.9797	0.6707	0.3428	0.5065	0.9748	0.7750	0.6831	0.5229	0.7008	-0.2216	0.2317	0.1141	0.3615	0.2862	
75	TP3	Chemotherapy only	0.5887	-0.2045	1.4868	0.5442	0.7610	0.4244	0.8056	0.5066	0.5657	0.5969	0.4211	0.1643	0.2550	-0.1226	0.8496	0.4352	
76	TP3	Chemotherapy only	0.5165	-0.2680	1.3373	0.5871	0.5981	0.3630	0.9811	0.5858	0.7472	0.4133	0.4881	0.1831	0.3064	-0.1174	1.0112	0.1740	
77	TP3	Bevacizumab treated	1.0073	-0.0409	1.6610	0.2892	0.6048	0.7303	1.0129	0.5085	0.6474	0.4302	0.3043	-0.0353	0.3857	-0.0933	1.0279	0.4732	
78	TP3	Chemotherapy only	1.0176	-0.0486	0.8505	0.4939	0.7205	0.4765	1.1262	0.5785	0.6950	0.4913	0.2754	-0.0032	0.5997	-0.0297	0.5511	0.2869	
80	TP3	Chemotherapy only	1.0632	-0.2587	1.9183	-0.0216	0.4251	0.4446	1.1169	0.6569	0.6861	0.3994	0.2637	0.1777	0.3866	-0.6866	1.0388	0.6703	
81	TP3	Chemotherapy only	0.8976	-0.2532	1.2752	0.6513	0.6412	0.5125	1.0309	0.5159	0.7126	0.4777	0.3434	0.0233	0.3926	-0.2420	0.7129	0.3159	
82	TP3	Bevacizumab treated	1.2658	-0.1656	0.9173	0.4926	0.4799	0.6803	1.2753	0.6682	0.7592	0.3770	0.3661	-0.0295	0.6807	-0.1713	0.5797	0.3203	
86	TP3	Bevacizumab treated	0.7262	0.0114	0.8197	0.4687	0.5774	0.4609	0.9018	0.6840	0.6285	0.4053	0.2809	0.0183	0.5322	-0.0214	0.7064	0.3425	
87	TP3	Bevacizumab treated	0.5003	-0.2509	0.9786	0.6655	0.6804	0.5542	0.8914	0.5196	0.7593	0.6094	0.2311	-0.1145	0.4193	0.0906	0.7995	0.2652	
88	TP3	Chemotherapy only	0.0638	-0.1543	0.9803	0.7078	0.6901	0.4497	0.9775	0.4716	0.8221	0.3797	0.6009	0.2242	0.4165	-0.0143	0.6631	0.2378	
89	TP3	Chemotherapy only	0.3847	0.0015	1.1803	0.7233	0.3185	0.5792	1.0309	0.5729	0.5617	0.4107	0.4977	0.2133	0.4534	0.0819	0.6445	0.2045	
90	TP3	Chemotherapy only	0.4031	-0.0326	1.4085	0.5932	0.8179	0.4414	0.8362	0.4384	0.6362	0.5164	0.4367	0.2064	0.2028	-0.1074	0.6961	0.2840	
91	TP3	Chemotherapy only	0.7658	-0.0859	1.2665	0.5157	0.6114	0.5754	0.9063	0.4783	0.5963	0.5650	0.4196	-0.0565	0.2440	-0.0523	0.7416	-2.1078	
92	TP3	Bevacizumab treated	0.9169	-0.2649	1.8283	0.4681	0.5924	0.5437	0.9929	0.5784	0.5978	0.5873	0.4183	0.0982	0.2849	-0.3126	1.0223	0.4838	
93	TP3	Bevacizumab treated	0.5460	-0.0907	1.0699	0.4929	0.6498	0.5837	0.9089	0.4429	0.4946	0.1307	0.4363	0.1862	0.3916	0.0189	0.6306	0.2745	
96	TP3	Bevacizumab treated	1.2543	-0.2295	1.1201	0.3453	0.4495	0.5253	1.2693	0.8083	0.7045	0.3576	0.3145	-0.2368	0.6215	-0.0439	0.5773	0.5762	
97	TP3	Bevacizumab treated	0.8403	-0.1469	1.2714	0.6375	0.5577	0.5178	0.9698	0.6009	0.6331	0.5841	0.3440	0.0125	0.3745	-0.2454	0.6914	0.4357	
98	TP3	Bevacizumab treated	0.8594	-0.0529	1.2891	0.6072	0.6207	0.6774	1.0653	0.5038	0.7133	0.4512	0.2653	-0.0321	0.1425	-0.3007	0.7940	0.4070	
99	TP3	Chemotherapy only	0.9589	-0.0666	1.1429	0.5635	0.6055	0.4731	1.0559	0.7091	0.7575	0.3378	0.3764	0.1262	0.4914	-0.1122	0.5815	0.3030	
100	TP3	Bevacizumab treated	0.7928	-0.1963	1.6957	0.3452	0.3946	0.4978	0.9555	0.6324	0.7091	0.2634	0.3923	0.0697	-0.2079	-0.6329	1.1179	0.4942	
101	TP3	Bevacizumab treated	0.9945	-0.0677	1.1753	0.5257	0.6344	0.6606	1.0674	0.6367	0.6246	0.4317	0.3769	0.0627	0.4958	-0.2778	0.6246	0.3711	
104	TP3	Bevacizumab treated	1.0962	-0.0815	1.5287	0.3512	0.4678	0.5932	1.1224	0.6427	0.6778	0.2884	0.3217	0.0755	0.5063	-0.3420	0.7746	0.5633	
105	TP3	Chemotherapy only	1.0793	-0.1228	1.7828	-0.0300	0.4241	0.4912	1.1178	0.5945	0.5885	0.3073	0.1924	0.1858	0.4194	-0.4182	1.1175	0.4794	

Article Title: Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy

Journal Name: Metabolomics

Authors: Leslie R. Euceda, Tonje H. Haukaas, Guro F. Giskeødegård, Riyas Vettukattil, Jasper Engel, Laxmi Silwal-Pandit, Steinar Lundgren, Elin Borgen, Øystein Garred, Geert Postma, Lutgarde M.C. Buydens, Anne-Lise Børresen-Dale, Olav Engebraaten, Tone F. Bathen

Corresponding author: Leslie R. Euceda, Department of Circulation and Medical Imaging - MR Center, Faculty of Medicine, NTNU, Norwegian University of Science and Technology, Postboks 8905, Medisinsk Teknisk Forskningssenter, 7491 Trondheim, Norway; Phone: +47 73597449, Fax: +47 73598613, E.mail: leslie.e.wood@ntnu.no.

Supplementary Table 3. LMM results for significant cross-level interactions between *time* and *pMRD response*, and *time* and *bevacizumab*.

	pMRD vs pNRs (TP1)	pMRD vs pNRs (TP2)	pMRD vs pNRs (TP3)	Bev-treated vs Chemo-only (TP1)	Bev-treated vs Chemo-only (TP2)	Bev-treated vs Chemo-only (TP3)
Glc	-0.155 (0.078)	0.286** (0.112)	0.508** (0.106)	0.096 (0.055)	Not App.	Not App.
Tyr	0.052* (0.026)	-0.097** (0.039)	-0.206** (0.038)	-0.023 (0.016)	Not App.	Not App.
Tau	-0.059* (0.026)	0.040 (0.037)	0.152** (0.035)	0.014 (0.018)	Not App.	Not App.
PCh	0.057 (0.029)	-0.073 (0.040)	-0.102** (0.038)	-0.024 (0.021)	Not App.	Not App.
Cho	0.051* (0.022)	-0.047 (0.031)	-0.093** (0.030)	-0.034* (0.015)	Not App.	Not App.
GSH	0.044 (0.029)	0.008 (0.041)	-0.106** (0.039)	0.057* (0.029)	-0.034 (0.039)	-0.129*[#] (0.038)
Succ	0.048 (0.042)	-0.171** (0.063)	-0.180** (0.060)	-0.032 (0.027)	Not App.	Not App.

Metabolite increase (positive estimate) or decrease (negative estimate) is shown at each time point for pMRD and bevacizumab-treated patients in relation to pNR and chemotherapy-only patients, respectively. Standard errors are shown in parenthesis.* and ** indicate significance ($p \leq 0.05$) before and after multiple testing correction, respectively.

[#]Multiple testing does not apply since GSH was the only metabolite for which the *Time* × *Bevacizumab* interaction term was included in the LMM.

Not App.: not applicable; pMRD: pathological minimal residual disease; pNR: pathological non-responder; Bev-treated: Bevacizumab-treated; Chemo-only: Chemotherapy only; Glc: glucose; Asc: ascorbate; Lac: lactate; Tyr: tyrosine; mI: myo-inositol; Gly: glycine; Tau: taurine; GPC: glycerophosphocholine; PCh: phosphoscholine; Cho: choline; Cr: creatine; GSH: glutathione; Gln: glutamine; Succ: succinate; Glu: glutamate; Ala: alanine.

Supplementary Table 4. Classification results from PLS-DA.

Samples included in the model	Discriminated Classes	n	No. of LVs	Class. Accuracy (%)	Sensitivity/ Specifcity (%)	Permutation p-value
TP1	pCR + vs pCR-	105	1	59	42/77	0.105
TP2	pCR + vs pCR-	78	1	54	31/77	0.418
TP3	pCR + vs pCR-	87	1	69	71/67	0.018*
TP1	pMRD vs pNRs	105	1	57	45/70	0.112
TP2	pMRD vs pNRs	78	1	59	38/81	0.096
TP3	pMRD vs pNRs	87	1	77	88/67	<0.001*
TP1	GR vs NR	45	1	76	65/88	0.001*
TP2	GR vs NR	27	1	49	51/76	0.573
TP3	GR vs NR	33	2	75	94/57	0.002*
TP1	Bev-treated vs Chemo-only	105	1	55	36/54	0.792
TP2	Bev-treated vs Chemo-only	78	1	46	66/41	0.340
TP3	Bev-treated vs Chemo-only	87	1	44	62/48	0.262

n: number of patients; pCR: pathological complete responders; pMRD: pathological minimal residual disease; pNRs: pathological non-responders; GR: good response; NR: no response; Bev: Bevacizumab; Chemo: Chemotherapy; Class.: classification; Sensitivity/Specificity are reported for pMRD/GR/Bev-treated.
* indicates significant p-values (≤0.05).

Supplementary Table 5. Pearson correlation between metabolite relative levels (log10) and response ratio.

	TP1	TP2	TP3
Glc	0.250**	0.038	-0.425**
Asc	0.017	-0.026	0.252**
Lac	0.008	-0.16	-0.276**
Tyr	-0.171*	0.033	0.448**
mI	-0.069	-0.064	0.165
Gly	-0.04	0.148	-0.125
Tau	0.291**	0.207*	-0.363**
GPC	-0.094	0.166	-0.011
PCh	-0.196*	0.045	0.159
Cho	-0.167*	-0.002	0.364**
Cr	0.063	-0.037	0.201*
GSH	-0.098	0.04	0.360**
Gln	0.021	0.044	-0.187*
Succ	-0.104	0.169	0.423**
Glu	-0.048	-0.102	-0.111
Ala	0.13	0.002	-0.16

* and ** indicate significance ($p \leq 0.05$) before and after multiple testing correction, respectively. Glc: glucose; Asc: ascorbate; Lac: lactate; Tyr: tyrosine; mI: myo-inositol; Gly: glycine; Tau: taurine; GPC: glycerophosphocholine; PCh: phosphoscholine; Cho: choline; Cr: creatine; GSH: glutathione; Gln: glutamine; Succ: succinate; Glu: glutamate; Ala: alanine.

Supplementary Table 6. Summary of results regarding treatment effect and differences between responders and non-responders by pathway.

Pathway	Metabolite	Effect with treatment progression	pMRD compared to pNRs	GR group compared to NR group
Glycolysis	**Glc**	General increasing trend, significant from TP2 to TP3.	Significantly higher at TP2 and TP3.	-Lower at TP1 -Higher at TP3
	Lac	Significant increase from TP1 to TP2 and from TP2 to TP3.	Significantly higher at all time points.	-Higher at TP1 -Lower at TP3
Choline phospholipid metabolism	**Cho**	Significant decrease from TP1 to TP2 and from TP2 to TP3.	Significantly lower at TP3.	-Higher at TP1 -Lower at TP3
	PCh	General decreasing trend, significant from TP1 to TP2 and from TP2 to TP3.	Significantly lower at TP3.	-Higher at TP1 -Lower at TP3
	GPC	General decreasing trend, significant from TP2 to TP3.	Lower at TP3.	-Higher at TP1 -Lower at TP3
Glutamine metabolism	**Gln**	Significant increase from TP2 to TP3.	None.	None.
	GSH	None.	Significantly lower at TP3.	-Higher at TP1 -Lower at TP3
TCA cycle	**Succ**	None.	Significantly lower at TP2 and TP3	-Higher at TP1 -Lower at TP3
Others	**Tau**	General decreasing trend.	Lower at TP3 (multivariate), significantly higher at TP3 (univariate)[#].	-Higher at TP1 -Lower at TP3
	Tyr	Significant decrease from TP2 to TP3.	Significantly lower at TP2 and TP3.	-Higher at TP1 -Lower at TP3
	Cr	Significant decrease from TP2 to TP3.	Significantly lower at TP3.	-Higher at TP1 -Lower at TP3

[#]Tau was the only metabolite for which multivariate and univariate results were not in accordance. pMRD: pathological minimal residual disease; pNRs: pathological non-responders; GR: good response; NR: no response; TCA: tricarboxylic acid; Glc: glucose; Lac: lactate; Cho: choline; PCh: phosphoscholine; GPC: glycerophosphocholine; Gln: glutamine; GSH: glutathione; Succ: succinate; Tau: taurine; Tyr: tyrosine; Cr: creatine.

Supplementary Table 7. PLS-DA classification results for bevacizumab-treated responders vs chemotherapy-only responders.

Samples included in the model	n Bev-treated/ n Chemo-only	No. of LVs	Class. Accuracy (%)	Sensitivity/ Specifcity (%)
TP1	22/19	1	34	27/42
TP2	13.nov	2	55	63/47
TP3	18.nov	2	43	7/80

n: number of samples; Bev: Bevacizumab; Chemo:Chemotherapy; Class.:Classification. Sensitivity/Specificity are reported for Bev-treated.

Supplementary Table 8. Student t-test p-values when comparing metabolite relative levels (log10 transformed) of bevacizumab (Bev)-treated responders vs chemotherapy (Chemo)-only responders (no multiple testing correction).

	TP1	TP2	TP3
n Bev-treated/ n Chemo-only	22/19	13/11	18/11
Glc	0.279	0.698	0.372
Asc	0.632	0.205	0.4
Lac	0.85	0.67	0.619
Tyr	0.725	0.833	0.618
mI	0.473	0.527	0.755
Gly	0.855	0.216	0.247
Tau	0.661	0.783	0.871
GPC	0.786	0.436	0.678
PCh	0.516	0.181	0.95
Cho	0.954	0.94	0.88
Cr	0.436	0.602	0.956
GSH	0.364	0.453	0.057
Gln	0.552	0.847	0.655
Succ	0.796	0.62	0.849
Glu	0.293	0.032	0.213
Ala	0.517	0.204	0.248

n:number of samples; Glc: glucose; Asc: ascorbate; Lac: lactate; Tyr: tyrosine; mI: myo-inositol; Gly: glycine; Tau: taurine; GPC: glycerophosphocholine; PCh: phosphocholine; Cho: choline; Cr: creatine; GSH: glutathione; Gln: glutamine; Succ: succinate; Glu: glutamate; Ala: alanine.

Paper IV

Metabolic response to everolimus in patient-derived xenografts of triple negative breast cancer

Leslie R. Euceda[1], Deborah K. Hill[1,2], Endre Stokke[1], Rana Hatem[3,4], Rania El Botty[5], Ivan Bièche[3,6], Elisabetta Marangoni[5], Tone F. Bathen[1], Siver A. Moestue[1,7]

[1]*Department of Circulation and Medical Imaging, NTNU, The Norwegian University of Science and Technology, Trondheim, Norway*

[2]*Department of Medical Imaging, St Olavs University Hospital, Trondheim, Norway*

[3]*Institut Curie, PSL Research University, Genetics Department, Paris, France*

[4]*Faculty of Pharmacy, Aleppo University, Aleppo, Syria*

[5]*Institut Curie, PSL Research University, Translational Research Department, Paris, France*

[6]*EA7331, University of Paris Descartes, Paris, France*

[7]*Department of Laboratory Medicine, Children's and Women's Health, NTNU, The Norwegian University of Science and Technology, Trondheim, Norway*

ABSTRACT

Triple negative breast cancer (TNBC) patients are unresponsive to both endocrine and anti-HER2 pharmacotherapy, limiting their therapeutic options to chemotherapy. TNBC is frequently associated with abnormalities in the PI3K signaling pathway, and drugs targeting this pathway are currently being evaluated in patients with TNBC. However, response is variable, partly due to heterogeneity within TNBC, and there is a need to identify biomarkers predicting response and resistance to targeted therapies. In this study, we used a metabolomics approach to assess response to the mTOR inhibitor everolimus in a panel of TNBC patient-derived xenografts (PDX) (n=103 animals). Metabolic profiles were acquired using high-resolution magic angle spinning magnetic resonance spectroscopy. Partial least squares-discriminant analysis (PLS-DA) on relative metabolite concentrations discriminated treated animals from untreated controls with an accuracy of 67% (p=0.003). Multilevel linear mixed-effects models (LMM) revealed similar treatment group trends observed by PLS-DA, strongly indicating reduced glycolytic lactate production and glutaminolysis after treatment, consistent with PI3K/AKT pathway inhibition. Although inherent metabolic heterogeneity between different PDX models seemed to hinder prediction of treatment response, the metabolic effects following treatment were more pronounced in responding xenografts compared to non-responders. Additionally, the metabolic information predicted p53 mutation status, which may provide complimentary insight into the interplay between PI3K signaling and other drivers of disease progression.

BACKGROUND

Pharmacological treatment of breast cancer has progressed notably in the past decades, accounting for much of the improvement in patient survival [1, 2]. Currently available systemic treatments include chemotherapy, endocrine therapy, and novel targeted therapies. Chemotherapy has been shown to be the most potent of these treatments [3], but severe adverse effects may limit its use. Endocrine treatment was the first targeted therapy for breast cancer; it is considered to be safe and effective, but is only useful for the treatment of hormone (estrogen and progesterone) receptor positive patients [3]. Novel targeted therapies include the human epidermal growth factor receptor 2 (HER2) inhibitor trastuzumab, which can provide additional benefit to patients overexpressing the cell growth-promoting HER2 (ERBB2) protein [3].

Triple negative breast cancer (TNBC) refers to any breast cancer lacking expression of estrogen receptors (ER), progesterone receptors (PgR), and HER2 protein. These cancers are consequently unresponsive to both endocrine and anti-HER2 therapy, limiting the therapeutic option for these patients to chemotherapy alone. Breast tumor heterogeneity is not only found at the protein level; Perou and Sørlie *et al.* divided the disease into five naturally-occurring, clinically relevant subtypes based on gene expression [4, 5]. TNBC displays a high degree of overlap with the basal-like gene expression subtype, which is associated with the worst prognosis and a high frequency of abnormalities in the phosphoinositide 3-kinase (PI3K)/Akt signaling pathway [6, 7] (Figure 1). This pathway is involved in cell survival, proliferation, and growth, and has been shown to be deregulated in various human cancers [8].

PI3K inhibitors have been developed for use as cancer therapies, many of which act upon the mammalian target of rapamycin (mTOR), one of the major effectors downstream of Akt. One such inhibitor is everolimus, a rapamycin analog that binds to the FKBP-12 protein of which mTOR complex 1 (mTORC1) is a direct target [9]. This blocks FKBP-12 binding to mTORC1, inhibiting further downstream signaling of the PI3K/Akt pathway. Everolimus is approved for treatment of hormone receptor positive [10, 11] and HER2 positive [12] breast cancer in combination with hormonal or anti-HER2 therapy, respectively, with the addition of the mTOR inhibitor significantly prolonging progression-free survival (PFS). However, with abnormal PI3K signaling occurring very frequently in TNBC compared to the other breast cancer subtypes, everolimus has been considered a potential candidate for targeted therapy of this breast cancer subtype.

TNBC, however, is a heterogeneous disease in terms of clinical, histological, and molecular aspects [13]. Tumor heterogeneity contributes greatly to the variability in breast cancer treatment response. Therefore, it is important to define reliable markers that can predict the outcome of treatment with targeted drugs. Previous studies have demonstrated variable responses to everolimus in TNBC, and robust molecular biomarkers for prediction of response/resistance to treatment are not yet identified [14]. Metabolites provide information at a level closer to the phenotype compared to genes and proteins, and altered metabolism is one of the more recently acknowledged hallmarks of cancer [15]. Consequently, it has been suggested that metabolic biomarkers or global metabolic profiles can guide patient selection for targeted drug treatment, and detect response/resistance to therapy [16]. High resolution (HR) magic angle spinning (MAS) magnetic resonance spectroscopy (MRS) allows for non-destructive, ex-vivo analysis of biological tissue samples and has been applied to study cancer-related metabolic pathways [17, 18]. Cao *et al.* [19] applied this technique to metabolically characterize TNBC and demonstrated that it was metabolically different from ER+/Pgr+/HER2+ breast cancer. In addition, tumors from the same gene expression subtype have been shown to have a variety of different metabolic profiles, as assessed by HR MAS MRS [20, 21]. The potential of HR MAS MRS in assessing response to therapy on a metabolic level has also been studied, with metabolic profiles discriminating responders from non-responders and relating metabolic changes after treatment in 5-year survival [22, 23] [submitted manuscript Euceda *et al*.]. Moestue *et al*. have demonstrated how response to treatment with a dual PI3K/mTOR inhibitor, BEZ235, was associated with altered levels of choline-containing metabolites as well as altered glucose and lactate levels in basal-like breast cancer xenografts [7].

In this study, we have assessed the metabolic responses to treatment with everolimus in a heterogeneous panel of TNBC patient-derived xenografts (PDX) using HR MAS MRS, aiming to identify metabolic biomarkers for response/resistance. We additionally explored whether the expression, phosphorylation, or mutation status of molecules regulating PI3K/Akt signaling could be determined based on metabolite information.

METHODS

Ethics

All procedures and experiments involving animals were carried out according to the institutional guidelines of the French Ethical Committee and the European Convention for the Protection of Vertebrates used for Scientific Purposes.

Animal Models

Thirteen TNBC PDX were previously established in female Swiss nude mice (Charles River, Les Arbresles, France) from primary surgical specimens with patient written informed consent, as described in [24]. Based on gene expression profiles, the models have been further subclassified (Table 1). Treated animals ($n \approx 4$ per PDX model, $n_{treated}=50$) received everolimus (Novartis, Basel, Switzerland) at an oral dose of 2.5 mg/kg three times a week for 4-5 weeks; fifty three animals ($n \approx 4$ per PDX model) were untreated controls, resulting in tumor tissue samples from a total of 103 animals (Figure 2). The tumor tissue was harvested on the last day of treatment, snap-frozen immediately, and stored at -80°C. The tumor material was divided for mutation screening, immunohistochemistry analysis, western blotting, and metabolic profiling. TNBC xenograft molecular and histological traits, previously characterized, are summarized in Table 1; for further information on the effect of everolimus treatment on protein and gene expression for this cohort, refer to [25].

Evaluation of Treatment Response

Response was classified based on relative tumor volume (RTV) measured using external calipers. For each PDX model, a Student's t test was performed comparing the RTV of treated tumors with that of the untreated tumors at the end of the treatment period. PDX models exhibiting a significant difference ($p \leq 0.05$) in RTVs of treated and untreated animals were classified as responders, while those not meeting this criterion were classified as non-responders.

HR MAS MRS analysis

Tumor tissue samples (12.29 ± 2.97 mg) were prepared and analyzed as described in the HR MAS MRS protocol by Giskeødegård *et al*. [26]. In short, samples were cut on a dedicated, cooled work station to fit into 30 μL disposable inserts containing 3.0 μL of 24.29 mM sodium formate (VWR BDH Prolabo, France) in D2O (Armar Chemicals, Switzerland) for shimming and locking purposes. Each insert was set tightly into a 4 mm MAS zirconium rotor. Each sample was prepared within five minutes, and HR MAS MR spectra were subsequently acquired on a Bruker Avance DRX600 spectrometer (Bruker Biospin GmbH, Germany) equipped with a $^{1}H/^{13}C$ MAS probe. Samples were spun at 5 kHz while maintaining the probe temperature at 5°C to minimize tissue degradation. Proton spectra were acquired using a spin-echo Carr-Purcell Meiboom-Gill (CPMG) sequence (cpmgpr1d, Bruker

BioSpin, Germany) as previously described in [21], with effective echo time of 77 ms, a spectral width of 20 ppm (−5 to 15 ppm), and 256 scans.

Immunohistochemical staining and mutation screening

Formalin-fixed, paraffin-embedded samples were analyzed by immunohistochemistry for the expression of phosphatase and tensin homolog (PTEN), which negatively regulates mTOR, and for the expression of the PI3K downstream member pAKT, as described in [25]. Furthermore, mutations in the tumor suppressor gene p53 were screened for as described in [25]. This gene has been found to negatively regulate mTOR and its downstream targets [27, 28].

Data preprocessing

From the original sample cohort of 108 samples, four were excluded due to high levels of an unknown contaminant at 3.70 ppm and one for poor spectral quality, resulting in 103 samples for subsequent statistical analysis. Following acquisition, the free induction decays (FIDs) were Fourier transformed into 64k real points after 0.30 Hz line broadening. Phase correction was performed automatically for each spectrum using TopSpin 3.1 (Bruker BioSpin GmbH, Germany).

The following spectral preprocessing steps were performed in Matlab R2013b (The Mathworks, Inc., USA).The spectral region between 1.40 – 4.70 ppm, containing the majority of low-molecular weight metabolites, was selected for further processing. Chemical shifts were referenced to the creatine peak at 3.03 ppm. The spectra were baseline corrected using asymmetric least squares [29] with parameters $\lambda = 1e7$ and $p = 0.0001$, after setting the lowest point in each spectrum to zero. Peak alignment was carried out using icoshift [30]. Lipid peaks at 4.33–4.28, 4.18–4.13, 2.88–2.72, 2.30–2.21, 2.11–2.00, 1.65–1.55 ppm, and the contaminant peaks for the previously mentioned unknown compound and ethanol at 3.73–3.69 and 3.67–3.62 ppm, respectively, were excluded from further analysis. The resulting spectra were normalized to total area to correct for differences in sample size and tumor cell content.

Statistical Analysis

Metabolite peak assignment was performed based on previous identification [31]. Relative levels of 17 metabolites were calculated by integrating fixed regions of preprocessed spectra

corresponding to the metabolite of interest in Matlab R2013b. The metabolite ratios of lactate/glucose, taurine/creatine, and glycerophosphocholine/phosphocholine were determined, and were combined with the relative levels of individual metabolites to make a single dataset which will be referred to as 'metabolite integrals'. All metabolite integrals were log10 transformed prior to analysis (Supplementary Table 1).

Multivariate analysis

Multivariate analysis was carried out for metabolite integrals in Matlab R2013b (The Mathworks, Inc., USA) using PLS Toolbox 7.8.2 (Eigenvector Research Inc., U.S.A). Integrals were autoscaled prior to multivariate model building. Principal component analysis (PCA) [32] was performed on metabolite data from untreated controls and treated animals separately, to explore naturally-occurring trends. The optimal number of principle components (PCs) was selected based on visual inspection of residual explained variance plots. Partial least squares-discriminant analysis (PLS-DA) [33] was used to build a classification model of all samples to discriminate treated animals from untreated controls. Additional classification models were also built for treated animals and untreated controls separately, to discriminate between responding and non-responding PDX models. Finally, PLS-DA was employed to build models of untreated controls discriminating tumors expressing (+) or not expressing (–) PTEN and pAKT, as well as tumors with mutant or wild type p53. PCA and PLS-DA loadings plots were employed to relate variables or metabolites to samples or animals in the scores plots.

The PLS-DA classification performance parameters of accuracy, sensitivity, and specificity were obtained by employing double-layered cross validation (CV) [34] to avoid model overfitting. This method optimizes the number of PLS latent variables (LVs) in the inner CV layer and assesses model performance in the outer layer. The procedure consisted of splitting the samples into a training and test set in each layer; this was performed by randomly selecting 20% of the samples to be excluded from model building, or training, for subsequent testing of the model built. The splitting was repeated until each sample was used for testing once and only once, keeping all animals of the same PDX model type together in either the training or test set. A maximum of ten LVs were considered for optimization. The whole double CV procedure was repeated 20 times, with the final classification results and optimal number of LVs resulting from the mean and mode, respectively. Permutation testing [34] was carried out as an additional model validation. For this, the sample class labels were randomly

shuffled (permuted) before PLS-DA model building. Single-layered CV leaving out 20% of samples as a test set at each data split, using all spectra of the same PDX model type for either training or testing, was used to assess permuted models. The number of LVs used was the optimal number determined from double CV. A thousand permutations were performed for each non-permuted model being validated, obtaining a permuted accuracy distribution. Models were considered significant if the final accuracy obtained from non-permuted double CV was higher than 95% of the permuted accuracy values ($p \leq 0.05$).

Univariate analysis

To compare treated animals versus untreated controls and responding versus non-responding PDX models simultaneously, linear mixed-effects models (LMM) [35] were employed as a multilevel approach. LMMs incorporate two types of effects to describe relationships between a response variable and different categorical factors. Fixed effects are controlled and systematic, originating from differences between factor levels, while random effects originate from the between-PDX model variation, each being derived from a different patient. Here, a LMM was built in R 3.1.1 [36] using the function *lme* from the 'nlme' package [37] and the method of restricted maximum likelihood. The response variable was the metabolite level, the fixed effects were *treatment group* (treated or untreated) and *response group* (responding or non-responding PDX), and the random effect was the PDX model.

LMM was additionally employed to correct for repeated PDX model measurements while determining individual metabolite differences between treated animals and untreated controls in responding and non-responding PDX models separately. These two LMM were built as described above, except that only the *treatment group* fixed effect was included, as samples were divided to perform the analysis on each individual *response group*. All LMM were built without interaction terms after determining that interactions were not significant, as described in [submitted manuscript Euceda *et al.*]. Obvious deviations from normality were not observed from LMM residual q-q plots and histograms.

LMM p-values were corrected for multiple testing by the Benjamini Hochberg method for false discovery rate (FDR) adjustment in R 3.1.1 using the 'stats' package [36]. Adjusted p (q-value) ≤ 0.05 was considered to be statistically significant.

RESULTS

Metabolic effects of everolimus treatment

PLS-DA of 20 metabolite integrals from all samples was performed to discriminate between treated animals and untreated controls (Figure 3A), achieving an accuracy of 67% (p=0.003) (Table 2). Treated animals exhibited higher glucose (Glc), glutamine (Gln), and alanine (Ala) levels, and lower phosphocholine (PCh), glycerophosphocholine (GPC), and lactate/glucose (Lac/Glc) (Figure 3B). When employing LMM as a multilevel approach, the same metabolites, with the exception of GPC/PCh instead of GPC, were found to be significantly different for the fixed effect of *treatment group* after multiple testing correction (Table 3), with estimates agreeing with PLS-DA loading trends; this indicates that both techniques revealed the same metabolic differences between treated animals and untreated controls.

Metabolic differences between responding and non-responding PDX models

Responding and non-responding PDX models could not be discriminated by multivariate PLS-DA, neither using metabolic information from untreated controls nor treated animals (Table 2). Similarly, for the multilevel LMM, no metabolites were significantly different for the *response group* fixed effect. In addition, PCA of untreated controls (Figure 4A) and treated animals (Figure 5A) did not reveal any separation between responding and non-responding PDX models. Separate PCA models of untreated controls including only responding (n=37) (Figure 4B) or only non-responding models (n=16) (Figure 4C) showed clear groupings in the scores plot by PDX model, reflecting the metabolic heterogeneity among the PDX models. Initial metabolic differences within the two groups may thus hinder prediction of response to everolimus treatment. PDX model heterogeneity persisted after everolimus treatment, as evidenced by the groupings in separate PCA models of treated animals including only responding (n=36) (Figure 5B) or only non-responding PDX models (n=14) (Figure 5C).

Metabolite information from responding PDX models (n=73) and non-responding PDX models (n=30) were analyzed separately using LMM to correct for repeated PDX model measurements, individual metabolite differences between treated animals and controls were determined within these two groups. For responding models, treated animals exhibited significantly higher Glc, Gln, and Ala, and significantly lower PCh, GPC/PCh, and Lac/Glc compared to untreated controls (Figure 6), which was similar to the findings from LMM of the whole cohort simultaneously (Table 3). For non-responding models, however, only PCh

and Gln were significantly lower and higher, respectively, with treatment (Figure 6). Collectively, these results indicate that the metabolic response to everolimus treatment is more pronounced in responding PDX models, with increased levels of metabolites representing the central carbon metabolism and decreased levels of phosphocholine.

Metabolic discrimination of PI3K pathway protein expression and p53 mutation status

Protein expression determined from immunohistochemistry and p53 mutation status for the PDX models is given in Table 1. PLS-DA results to classify untreated controls (n=53) according to expression or no expression of proteins PTEN and pAKT and to p53 mutation status are summarized in Table 2. p53 mutation status was successfully discriminated with an accuracy of 71% (p=0.008) (Figure 7A), with the mutant gene associated with increased lactate (Lac), glycine (Gly), scyllo-inositol (sI) , and creatine (Cr), and decreased taurine (Tau), GPC, choline (Cho), and glutathione (GSH) (Figure 7B). Metabolite integrals were unable to predict pAKT expression, while PTEN expression discrimination approached significance (Accuracy=63%, p=0.069) (Figure 7C). Wild type p53 exhibited similar metabolic trends as PTEN expression (Figure 7B, Figure 7D), with increased PCh and decreased Ala additionally being associated with the latter. Altogether, this indicates that the expression of tumor suppressor genes regulating signal transduction in the PI3K pathway is reflected in the metabolic profile of the tumor.

DISCUSSION

Triple negative breast cancer is the most clinically challenging histopathological category of the disease owing to its highly aggressive nature, poor prognosis, and lack of available targeted therapies [38]. Activation of the PI3K signaling pathway is frequently seen in TNBC, and it has therefore been proposed that drugs targeting the PI3K/Akt/mTOR axis may be of particular benefit in these patients. In this study, we used MR spectroscopy to obtain metabolic profiles from patient-derived triple negative breast tumor xenografts following everolimus treatment, demonstrating more pronounced metabolic effects in responding xenografts compared to non-responders.

Since this drug blocks PI3K signaling via mTOR inhibition, we additionally explored whether metabolic data reflects the expression of key proteins and mutation status for the tumor suppressor gene p53, which regulate signaling activity in this pathway.

Successful discrimination of treated animals and untreated controls was achieved by applying a multivariate approach where 17 metabolites and three metabolite ratios, collectively referred

to as metabolite integrals, were analyzed simultaneously. Five out of six metabolite integrals that contributed the most to the discrimination were also found to be significantly different between the treatment groups when the relative metabolite levels were analyzed individually using LMM; treatment group trends observed by each method were in accordance with each other. These agreeing results, summarized by pathway in Figure 8, strongly indicate metabolic alterations in TNBC patient-derived tumor xenografts following everolimus treatment.

Glucose levels can be affected by many factors, including cell number and degree of perfusion. The elevated glucose observed as an effect of treatment points to the normalization of cellular energy metabolism. Via glycolysis, glucose is metabolized to pyruvate, most of which is fed into the mitochondrial tricarboxylic acid cycle (TCA) under aerobic conditions, and ultimately, into oxidative phosphorylation to produce energy in the form of adenosine triphosphate (ATP). On the other hand, anaerobic glycolysis leads to the majority of pyruvate being converted to lactate, producing much less ATP (Figure 8). In cancer and normal proliferating cells, the latter occurs independently of oxygen availability, a phenomenon known as the Warburg effect [39] or glycolytic phenotype. PI3K activation has been shown to promote this effect [40], and its inefficiency in energy production leads to an increase in glucose consumption to fulfill the energy requirements of the cell. Higher glucose as well as lower lactate has been repeatedly observed in normal breast tissue compared to tumor tissue [41-43], and was found by Moestue *et al.* in basal-like xenografts following PI3K inhibition [7]. Although we could not detect any significant change in lactate with treatment, the lactate-to-glucose ratio was significantly lower in treated animals, further suggesting a decrease in glucose consumption in this group.

Another finding indicative of a positive effect of everolimus treatment was the observed decrease in phosphocholine. This metabolite is an intermediate in the synthesis of the cell membrane phospholipid phosphatidylcholine (Figure 8), and has been associated with malignant transformation [44]. Phosphocholine can be observed in vivo together with choline and glycerophosphocholine as one peak referred to as total choline (tCho), whose decrease has been suggested as a marker of treatment response [45, 46]. Although it is widely accepted that choline metabolism is abnormal in cancer cells, the mechanisms driving this are still poorly understood [47]. The relationship between phosphocholine and glycerophosphocholine is currently under discussion, with some studies suggesting increased PCh/GPC as a marker for malignancy [44, 48] while others have found the inverse GPC/PCh to be elevated in breast cancer subtypes with worse prognosis [49, 50]. Here, GPC/PCh was observed to be

significantly higher in the treated animals compared to untreated controls, while no significant differences were observed for either glycerophosphocholine nor choline. This suggests that phosphocholine is the tCho-constituent most affected by everolimus, and its decrease may reflect a reduction in tumor malignancy and aggressiveness due to everolimus treatment.

Significant differences in glutamine were also observed between treated animals and controls. Although glutamine is a non-essential amino acid in non-cancerous cells, cancer cells depend on it to replace glucose, feeding the TCA cycle via production of alpha-ketoglutarate (Figure 8). In addition, it plays an important role in supplying carbon and nitrogen for macromolecular synthesis needed to sustain cell proliferation [51]. Furthermore, glutamine provides protection against oxidative stress, as it is a precursor of glutamate-derived glutathione (GSH), an important cellular antioxidant [51, 52]. The significantly higher levels of glutamine observed in the treated animals suggests its lower consumption in this group, and may reflect a decrease in glutamine addiction with everolimus treatment.

A key objective of the study was to compare metabolic responses in responding and non-responding PDX models. Using multivariate PLS-DA, significant differences between responding and non-responding PDX models could not be detected in either treated animals or untreated controls. Further multivariate PCA was performed to investigate metabolic heterogeneity both independent of (untreated control group) and after everolimus treatment (treated group). Although all patient-derived tumor xenografts included in this study were triple negative invasive ductal carcinomas (IDC), distinct metabolic characteristics for each PDX model were identified through the groupings on the PCA plots for both responding and non-responding models. This metabolic heterogeneity was expected, as the PDX differed with regards to gene expression traits. The inter-tumor heterogeneity therefore contributed to the unsuccessful prediction of response to treatment based on metabolite integrals. Building the multilevel LMM including the random effect of 'PDX model' should account for the between-PDX model heterogeneity, but differences between the response groups could still not be detected using this technique. The observed heterogeneity suggests that treatment management of TNBC may benefit from further stratification of this histopathological subtype. Moestue *et al.* [7] found that a basal-like, but not a luminal-like, breast cancer PDX responded to PI3K inhibitors, and basal-like cancers have been found to be more homogenous than TNBC [53]. In our cohort, however, both responding and non-responding PDX models included basal-like, luminal-like, and HER2-enriched tumors (Table 1), suggesting that other factors contributed to the heterogeneous treatment efficacy in terms of tumor growth

inhibition within the gene expression subtypes. Additional sources of this heterogeneity, perhaps with an initial focus on basal-like breast cancers, may be worth investigating.

Although comparing each response group with each other proved to be unsuccessful in detecting metabolic differences between them, separating these groups to investigate differences between treated animals and untreated controls revealed a more substantial metabolic effect of everolimus treatment in responding PDX models. With only phosphocholine and glutamine significantly decreasing and increasing, respectively, with treatment in non-responding models, significant changes in metabolites involved in glycolysis were only observed in the responding models. This is consistent with findings from Foster *et al.* [54], who observed glucose-dependent growth in cells with mutations in PIK3CA, the gene encoding the alpha catalytic subunit of PI3K. In contrast, they found that both wild-type and PIK3CA-mutated cells depended on glutamine to grow. This supports our observed significant increase in glutamine with treatment, independent of whether tumor growth was significantly inhibited (responding PDX) or not (non-responding PDX). Previous work studying the same animal cohort examined here, however, found that PIK3CA mutations were rare and could not predict response to everolimus on their own [25].

Despite findings related to the response groups not being as straightforward as the metabolic differences observed in the treatment groups, it is important to note that the response criteria used here was based on tumor size, since this is the traditional approach. Imaging techniques are widely used in the clinic to assess response to treatment based on tumor reduction [55, 56]. There has been increasing interest in the development of imaging modalities that can detect biological and physiological changes in tissue rather than just imaging morphology [57, 58]. Functional imaging approaches provide possibilities for treatment assessment monitoring based on biological changes in the tumor, which may occur long before any reduction in size [59]. In line with this, we could detect metabolic differences between treated animals and untreated controls, pointing to biological changes occurring as an effect of treatment, in all PDX models, whether they exhibited a significant reduction in tumor volume or not. Nevertheless, PDX models responding to everolimus therapy exhibited a more significant metabolic response than resistant models.

The study by Hatem *et al.* on the effect of everolimus treatment on protein and gene expression for the animal cohort included in this study suggested post treatment activation of AKT to be a potential biomarker for early drug response monitoring [25]. We could not see

differences in AKT activation being translated to the metabolic level, however, as the metabolite information could not discriminate post treatment samples expressing or not expressing pAKT (results not shown). Although discrimination of untreated tumors expressing or not expressing proteins involved in PI3K/Akt signaling based on metabolic information could not be achieved, p53 mutation status could be predicted from the metabolite levels. The potent tumor suppressor p53 is involved in cell cycle control via transcriptional regulation of its target genes and has been found to decrease glycolytic activity. Up-regulation of the genes of several glycolytic enzymes has been observed in mutated p53 breast tumors [60]. Although differences in glucose levels were not observed between p53 wild type and mutant PDX models, increased lactate was associated with the latter. This is in accordance with wild type p53 negatively regulating pyruvate dehydrogenase kinase-2, which inactivates the pyruvate dehydrogenase complex (PDC) [61]. Since PDC converts pyruvate into Acetyl CoA, p53 promotes a non-glycolytic phenotype by directing pyruvate into the TCA cycle instead of towards lactate production. Other metabolites that were found to be associated with p53 mutation were creatine and glycine. p53 increases guanidinoacetate methyltranferase (GAMT), which catalyzes the production of creatine from guanidinoacetate (GAA) [62], and glycine is a precursor of GAA. Differences in creatine and glycine levels between wild type and mutant p53 PDX models may therefore reflect a p53-associated dysregulation of creatine synthesis. This may result as a compensation mechanism for impaired glycolytic energy production in mutants, since creatine metabolism is tightly connected with ATP homeostasis via the reversible phosphorylation of creatine by creatine kinase with ATP/ADP [63]. Protein expression discrimination based on metabolic information approached significance for PTEN (Table 2), which is another tumor suppressor. Similar metabolic trends were observed with lack of PTEN expression and mutant p53. The metabolome may therefore reflect the connection between PI3K/AKT pathway and other signaling mechanisms.

In conclusion, clear metabolic differences between everolimus-treated animals and untreated controls were detected, indicating reduced glycolytic lactate production and glutaminolysis after treatment, consistent with PI3K/AKT signaling pathway inhibition. Although inherent metabolic heterogeneity between different PDX models seemed to hinder prediction of treatment response, significant changes in glucose, alanine, lactate/glucose, and glycerophosphocholine/phosphocholine with treatment were detected in responding, but not in non-responding, PDX models. p53 mutation status could be predicted using MR based

metabolite levels, which may provide complimentary insight into the interplay between PI3K signaling and other drivers of disease progression.

References

1. Giordano SH, Buzdar AU, Smith TL, Kau S-W, Yang Y, Hortobagyi GN. Is breast cancer survival improving? Cancer. 2004;100(1):44-52. doi:10.1002/cncr.11859.
2. American Cancer Society. Global Cancer Facts & Figures 2nd Edition. Atlanta, GA, USA: American Cancer Society; 2011.
3. Yiannakopoulou E. Breast Cancer Therapy–Classical Therapy, Drug Targets, and Targeted Therapy. In: Barh D, editor. Omics Approaches in Breast Cancer. Springer India; 2014. p. 483-98.
4. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52. doi:http://www.nature.com/nature/journal/v406/n6797/suppinfo/406747a0_S1.html.
5. Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences. 2003;100(14):8418-23. doi:10.1073/pnas.0932692100.
6. López-Knowles E, O'Toole SA, McNeil CM, Millar EKA, Qiu MR, Crea P et al. PI3K pathway activation in breast cancer is associated with the basal-like phenotype and cancer-specific mortality. International Journal of Cancer. 2010;126(5):1121-31. doi:10.1002/ijc.24831.
7. Moestue SA, Dam CG, Gorad SS, Kristian A, Bofin A, Mælandsmo GM et al. Metabolic biomarkers for response to PI3K inhibition in basal-like breast cancer. Breast Cancer Res. 2013;15:R16.
8. Vara JÁF, Casado E, de Castro J, Cejas P, Belda-Iniesta C, González-Barón M. PI3K/Akt signalling pathway and cancer. Cancer Treatment Reviews. 2004;30(2):193-204. doi:http://dx.doi.org/10.1016/j.ctrv.2003.07.007.
9. Houghton PJ. Everolimus. Clinical cancer research : an official journal of the American Association for Cancer Research. 2010;16(5):1368-72. doi:10.1158/1078-0432.CCR-09-1314.
10. Baselga J, Campone M, Piccart M, Burris HA, Rugo HS, Sahmoud T et al. Everolimus in Postmenopausal Hormone-Receptor–Positive Advanced Breast Cancer. New England Journal of Medicine. 2012;366(6):520-9. doi:doi:10.1056/NEJMoa1109653.
11. Ejlertsen B, Heinrich G, Jerusalem M, Hurvitz SA, De Boer RH, Taran T et al. BOLERO-6: Phase II study of everolimus plus exemestane versus everolimus or capecitabine monotherapy in HR+, HER2- advanced breast cancer. J Clin Oncol; 2013.
12. André F, O'Regan R, Ozguroglu M, Toi M, Xu B, Jerusalem G et al. Everolimus for women with trastuzumab-resistant, HER2-positive, advanced breast cancer (BOLERO-3): a randomised, double-blind, placebo-controlled phase 3 trial. The Lancet Oncology. 2014;15(6):580-91. doi:http://dx.doi.org/10.1016/S1470-2045(14)70138-X.
13. Metzger-Filho O, Tutt A, de Azambuja E, Saini KS, Viale G, Loi S et al. Dissecting the Heterogeneity of Triple-Negative Breast Cancer. Journal of Clinical Oncology. 2012. doi:10.1200/jco.2011.38.2010.
14. Hatem R, Botty RE, Chateau-Joubert S, Servely J-L, Labiod D, Plater Ld et al. Targeting mTOR pathway inhibits tumor growth in different molecular subtypes of triple-negative breast cancers. 2016. 2016.
15. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144. doi:10.1016/j.cell.2011.02.013.
16. Palmnas M, Vogel H. The Future of NMR Metabolomics in Cancer Therapy: Towards Personalizing Treatment and Developing Targeted Drugs? Metabolites. 2013;3(2):373.
17. Bathen TF, Sitter B, Sjøbakk TE, Tessem M-B, Gribbestad IS. Magnetic Resonance Metabolomics of Intact Tissue: A Biotechnological Tool in Cancer Diagnostics and Treatment Evaluation. Cancer Research. 2010;70(17):6692-6. doi:10.1158/0008-5472.can-10-0437.
18. Moestue SA, Sitter B, Bathen TF, Tessem M-B, Gribbestad I. HR MAS MR Spectroscopy in Metabolic Characterization of Cancer. Current Topics in Medicinal Chemistry. 2011;11(1):2-26. doi:http://dx.doi.org/10.2174/156802611793611869.

19. Cao MD, Lamichhane S, Lundgren S, Bofin A, Fjøsne H, Giskeødegård GF et al. Metabolic characterization of triple negative breast cancer. BMC Cancer. 2014;14(1):1-12. doi:10.1186/1471-2407-14-941.
20. Borgan E, Sitter B, Lingjærde OC, Johnsen H, Lundgren S, Bathen TF et al. Merging transcriptomics and metabolomics - advances in breast cancer profiling. BMC Cancer. 2010;10(1):1-14. doi:10.1186/1471-2407-10-628.
21. Haukaas TH, Euceda LR, Giskeødegård GF, Lamichhane S, Krohn M, Jernström S et al. Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. Cancer & Metabolism. 2016;4(1):1-14. doi:10.1186/s40170-016-0152-x.
22. Cao MD, Sitter B, Bathen TF, Bofin A, Lønning PE, Lundgren S et al. Predicting long-term survival and treatment response in breast cancer patients receiving neoadjuvant chemotherapy by MR metabolic profiling. NMR in Biomedicine. 2012;25(2):369-78. doi:10.1002/nbm.1762.
23. Cao MD, Giskeødegård GF, Bathen TF, Sitter B, Bofin A, Lønning PE et al. Prognostic value of metabolic response in breast cancer patients receiving neoadjuvant chemotherapy. BMC Cancer. 2012;12(1):1-11. doi:10.1186/1471-2407-12-39.
24. Marangoni E, Vincent-Salomon A, Auger N, Degeorges A, Assayag F, de Cremoux P et al. A New Model of Patient Tumor-Derived Breast Cancer Xenografts for Preclinical Assays. Clinical Cancer Research. 2007;13(13):3989-98. doi:10.1158/1078-0432.ccr-07-0078.
25. Hatem R, Botty RE, Chateau-Joubert S, Servely J-L, Labiod D, Plater Ld et al. Targeting mTOR pathway inhibits tumor growth in different molecular subtypes of triple-negative breast cancers. Oncotarget. 2016. doi:10.18632/oncotarget.10195.
26. Giskeødegård G, Cao M, Bathen T. High-Resolution Magic-Angle-Spinning NMR Spectroscopy of Intact Tissue. In: Bjerrum JT, editor. Metabonomics. Methods in Molecular Biology: Springer New York; 2015. p. 37-50.
27. Feng Z, Zhang H, Levine AJ, Jin S. The coordinate regulation of the p53 and mTOR pathways in cells. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(23):8204-9. doi:10.1073/pnas.0502857102.
28. Akeno N, Miller AL, Ma X, Wikenheiser-Brokamp KA. p53 suppresses carcinoma progression by inhibiting mTOR pathway activation. Oncogene. 2015;34(5):589-99. doi:10.1038/onc.2013.589.
29. Eilers PHC. Parametric Time Warping. Anal Chem. 2004;76(2):404-11. doi:10.1021/ac034800e.
30. Savorani F, Tomasi G, Engelsen SB. icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. J Magn Reson. 2010;202(2):190-202. doi:http://dx.doi.org/10.1016/j.jmr.2009.11.012.
31. Sitter B, Sonnewald U, Spraul M, Fjösne HE, Gribbestad IS. High-resolution magic angle spinning MRS of breast cancer tissue. NMR in Biomedicine. 2002;15(5):327-37. doi:10.1002/nbm.775.
32. Wold S, Esbensen K, Geladi P. Principal Component Analysis. Chemometr Intell Lab. 1987;2:37-52.
33. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems. 2001;58(2):109-30. doi:http://dx.doi.org/10.1016/S0169-7439(01)00155-1.
34. Westerhuis J, Hoefsloot HJ, Smit S, Vis D, Smilde A, van Velzen EJ et al. Assessment of PLSDA cross validation. Metabolomics. 2008;4(1):81-9. doi:10.1007/s11306-007-0099-6.
35. Pinheiro JC, Bates DM. Linear Mixed-Effects Models: Basic Concepts and Examples. Mixed-Effects Models in S and S-PLUS. Statistics and Computing. New York, NY, USA: Springer New York; 2000. p. 3-56.
36. R Core Team (2014) R: A language and environment for statistical computing. http://www.R-project.org/.Viena, Austria. URL http://www.R-project.org/: R Foundation for Statistical Computing.
37. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2014). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-117, URL: http://CRAN.R-project.org/package=nlme. 2014.
38. Kaplan HG, Malmgren JA. Impact of Triple Negative Phenotype on Breast Cancer Prognosis. The Breast Journal. 2008;14(5):456-63. doi:10.1111/j.1524-4741.2008.00622.x.
39. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Sci Signal. 2009;324.
40. Plas DR, Thompson CB. Akt-dependent transformation: there is more to growth than just surviving. Oncogene. 0000;24(50):7435-42.

41. Gribbestad IS, Petersen SB, Fjosne HE, Kvinnsland S, Krane J. 1H NMR spectroscopic characterization of perchloric acid extracts from breast carcinomas and non-involved breast tissue. NMR Biomed. 1994;7(4):181-94.
42. Tessem M-B, Selnæs KM, Sjursen W, Tranø G, Giskeødegård GF, Bathen TF et al. Discrimination of Patients with Microsatellite Instability Colon Cancer using 1H HR MAS MR Spectroscopy and Chemometric Analysis. Journal of Proteome Research. 2010;9(7):3664-70. doi:10.1021/pr100176g.
43. Bathen TF, Geurts B, Sitter B, Fjøsne HE, Lundgren S, Buydens LM et al. Feasibility of MR Metabolomics for Immediate Analysis of Resection Margins during Breast Cancer Surgery. PLoS ONE. 2013;8(4):e61578. doi:10.1371/journal.pone.0061578.
44. Aboagye EO, Bhujwalla ZM. Malignant transformation alters membrane choline phospholipid metabolism of human mammary epithelial cells. Cancer Res. 1999;59(1):80-4.
45. Jagannathan NR, Kumar M, Seenu V, Coshic O, Dwivedi SN, Julka PK et al. Evaluation of total choline from in-vivo volume localized proton MR spectroscopy and its response to neoadjuvant chemotherapy in locally advanced breast cancer. British Journal of Cancer. 2001;84(8):1016-22. doi:10.1054/bjoc.2000.1711.
46. Baek H-M, Chen J-H, Nalcioglu O, Su M-Y. Proton MR spectroscopy for monitoring early treatment response of breast cancer to neo-adjuvant chemotherapy. Annals of Oncology. 2008;19(5):1022-4. doi:10.1093/annonc/mdn121.
47. Moestue SA, Giskeodegard GF, Cao MD, Bathen TF, Gribbestad IS. Glycerophosphocholine (GPC) is a poorly understood biomarker in breast cancer. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(38):E2506; author reply E7. doi:10.1073/pnas.1208226109.
48. Iorio E, Ricci A, Bagnoli M, Pisanu ME, Castellano G, Di Vito M et al. Activation of phosphatidylcholine-cycle enzymes in human epithelial ovarian cancer cells. Cancer research. 2010;70(5):2126-35. doi:10.1158/0008-5472.CAN-09-3833.
49. Moestue S, Borgan E, Huuse E, Lindholm E, Sitter B, Borresen-Dale AL et al. Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models. BMC Cancer. 2010;10. doi:10.1186/1471-2407-10-433.
50. Giskeødegård GF, Grinde MT, Sitter B, Axelson DE, Lundgren S, Fjosne HE et al. Multivariate modeling and prediction of breast cancer prognostic factors using MR metabolomics. J Proteome Res. 2010;9(2):972-9. doi:10.1021/pr9008783.
51. Hensley CT, Wasti AT, DeBerardinis RJ. Glutamine and cancer: cell biology, physiology, and clinical opportunities. The Journal of clinical investigation. 2013;123(9):3678-84. doi:10.1172/jci69600.
52. Franco R, Cidlowski JA. Apoptosis and glutathione: beyond an antioxidant. Cell Death Differ. 2009;16(10):1303-14. doi:http://www.nature.com/cdd/journal/v16/n10/suppinfo/cdd2009107s1.html.
53. Bertucci F, Finetti P, Cervera N, Esterni B, Hermitte F, Viens P et al. How basal are triple-negative breast cancers? International Journal of Cancer. 2008;123(1):236-40. doi:10.1002/ijc.23518.
54. Foster R, Griffin S, Grooby S, Feltell R, Christopherson C, Chang M et al. Multiple Metabolic Alterations Exist in Mutant PI3K Cancers, but Only Glucose Is Essential as a Nutrient Source. PLoS ONE. 2012;7(9):e45061. doi:10.1371/journal.pone.0045061.
55. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L et al. New Guidelines to Evaluate the Response to Treatment in Solid Tumors. Journal of the National Cancer Institute. 2000;92(3):205-16. doi:10.1093/jnci/92.3.205.
56. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer. 2009;45(2):228-47. doi:http://dx.doi.org/10.1016/j.ejca.2008.10.026.
57. Juweid ME, Cheson BD. Positron-Emission Tomography and Assessment of Cancer Therapy. New England Journal of Medicine. 2006;354(5):496-507. doi:doi:10.1056/NEJMra050276.
58. Gutte H, Hansen AE, Johannesen HH, Clemmensen AE, Ardenkjær-Larsen JH, Nielsen CH et al. The use of dynamic nuclear polarization (13)C-pyruvate MRS in cancer. American Journal of Nuclear Medicine and Molecular Imaging. 2015;5(5):548-60.
59. Brindle K. New approaches for imaging tumour responses to treatment. Nat Rev Cancer. 2008;8(2):94-107.

60. Harami-Papp H, Pongor LS, Munkacsy G, Horvath G, Nagy AM, Ambrus A et al. TP53 mutation hits energy metabolism and increases glycolysis in breast cancer. Oncotarget. 2016. doi:10.18632/oncotarget.11594.
61. Contractor T, Harris CR. p53 Negatively Regulates Transcription of the Pyruvate Dehydrogenase Kinase Pdk2. Cancer Research. 2012;72(2):560-7. doi:10.1158/0008-5472.can-11-1215.
62. da Silva RP, Clow K, Brosnan JT, Brosnan ME. Synthesis of guanidinoacetate and creatine from amino acids by rat pancreas. The British journal of nutrition. 2014;111(4):571-7. doi:10.1017/s0007114513003012.
63. Ide T, Chu K, Aaronson SA, Lee SW. GAMT joins the p53 network: branching into metabolism. Cell cycle (Georgetown, Tex). 2010;9(9):1706-10. doi:10.4161/cc.9.9.11473.

Tables

Table 1. Patient-derived tumor xenograft characteristics.

PDX model	nTreated/ nControls	TNBC subtype	PTEN expression	pAKT expression	53 mutations	Response Group
HBCx2	4/4	Luminal-like (AR+FOXA+)	+	+	p.A276D	Non-Resp.
HBCx12A	2/4	HER2 enriched	-	+	WT	Non-Resp.
HBCx16	4/4	Basal (KRT5+KRT17+)	+	+	WT	Non-Resp.
HBCx30	4/4	Basal (KRT5+KRT17+)	-	-	p.F134L	Non-Resp.
HBCx39	4/4	Basal (KRT5+KRT17+) / HER2 enriched	+	-	p.Y220C	Responder
HBCx31	4/4	Luminal-like (AR+FOXA1+)	+	+	p.R175H	Responder
HBCx66	4/4	Basal (KRT17+)	-	+	p.R273C	Responder
HBCx10	4/4	HER2 enriched	-	+	p.Q144fs	Responder
HBCx51	3/2	HER2 enriched	+	-	WT	Responder
HBCx4B	7/8	Basal (KRT5+ KRT17+)	-	+	p.S149fs	Responder
HBCx52	3/4	Luminal-like AR+ FOXA1+	+	NA	WT	Responder
HBCx24	4/3	Basal (KRT5+)	-	+	p.K292fs	Responder
HBCx63	3/4	unclassified	-	+	p.S149fs	Responder

All PDX models were of the invasive ductal carcinoma (IDC) histological type.

PDX: patient-derived xenograft; TNBC: triple negative breast cancer; Non-resp.: non-responder; AR: androgen receptor; FOXA: forkhead box protein A; KRT: keratin; NA: not available, WT:wild type.

Table 2. Classification results from PLS-DA.

Samples included in the model	Discriminated Classes	n	No. of LVs	Classification Accuracy (%)	Sensitivity/ Specificity (%)	Permutation p-value
All	Treated vs Untreated	106	1	67	66/68	0.003*
Untreated	Resp. PDX vs Non-resp. PDX	53	1	47	63/32	0.613
Treated	Resp. PDX vs Non-resp. PDX	50	1	57	68/46	0.237
Untreated	PTEN+ vs PTEN-	53	1	63	68/57	0.069
Untreated	pAKT+ vs pAKT-	49[#]	1	57	21/94	0.244
Untreated	p53 mutant vs Wild type	53	1	71	81/61	0.008*

Sensitivity/Specificity are reported for Treated/Resp. PDX /INPP4B+/PTEN+/pAKT+.
* indicates significant p-values (≤0.05).
[#]4/53 samples were of a PDX model with undetermined pAKT expression, and were therefore excluded.
n: number of samples; No. of LVs: number of latent variables; Resp. PDX: Responding patient-derived xenograft models; Non-resp PDX: non-responding patient-derived xenograft models; Class: classification.

Table 3. LMM results for the fixed effects of treatment group and response group.

Metabolite	Treated animals vs untreated controls			Responding vs non-responding PDX		
	Estimate	Std. Error	q-value	Estimate	Std. Error	q-value
Glc	0.130	0.038	4.45E-03*	-0.049	0.116	9.25E-01
Asc	0.016	0.019	6.63E-01	0.186	0.063	2.51E-01
Lac	0.015	0.018	6.63E-01	-0.057	0.055	8.39E-01
Tyr	-0.012	0.021	8.82E-01	0.045	0.065	9.09E-01
Gly	0.015	0.017	6.63E-01	0.008	0.077	9.76E-01
mI	0.007	0.018	9.24E-01	-0.054	0.126	9.25E-01
Tau	-0.004	0.021	9.24E-01	0.002	0.055	9.76E-01
sI	0.038	0.040	6.63E-01	0.028	0.118	9.59E-01
GPC	-0.035	0.025	4.80E-01	0.093	0.165	9.25E-01
PCh	-0.129	0.028	1.26E-04**	0.032	0.093	9.25E-01
Cho	-0.006	0.022	9.24E-01	0.054	0.051	8.39E-01
Cr	-0.003	0.019	9.28E-01	0.103	0.112	8.39E-01
GSH	0.007	0.027	9.24E-01	-0.054	0.057	8.39E-01
Gln	0.144	0.021	2.59E-08**	0.115	0.098	8.39E-01
Succ	0.019	0.018	6.63E-01	0.040	0.041	8.39E-01
Glu	0.007	0.019	9.24E-01	-0.081	0.061	8.39E-01
Ala	0.060	0.015	7.02E-04**	-0.071	0.075	8.39E-01
Lac/Glc	-0.115	0.041	2.05E-02*	-0.008	0.132	9.76E-01
Tau/Cr	0.000	0.015	9.87E-01	-0.103	0.128	8.83E-01
GPC/PCh	0.095	0.025	1.63E-03*	0.059	0.172	9.25E-01

Metabolite increase (positive estimate) or decrease (negative estimate) is shown for treated animals with respect to untreated controls and responding PDX models with respect to non-responding PDX models, respectively. * and ** indicate significance (q≤0.05 and q≤0.001, respectively).

Std.: standard; Glc: glucose; Asc: ascorbate; Lac: lactate; Tyr: tyrosine; Gly: glycine; mI: myo-inositol; Tau: taurine; sI: scyllo-inositol; GPC: glycerophosphocholine; PCh: phosphoscholine; Cho: choline; Cr: creatine; GSH: glutathione; Gln: glutamine; Succ: succinate; Glu: glutamate; Ala: alanine.

Figures

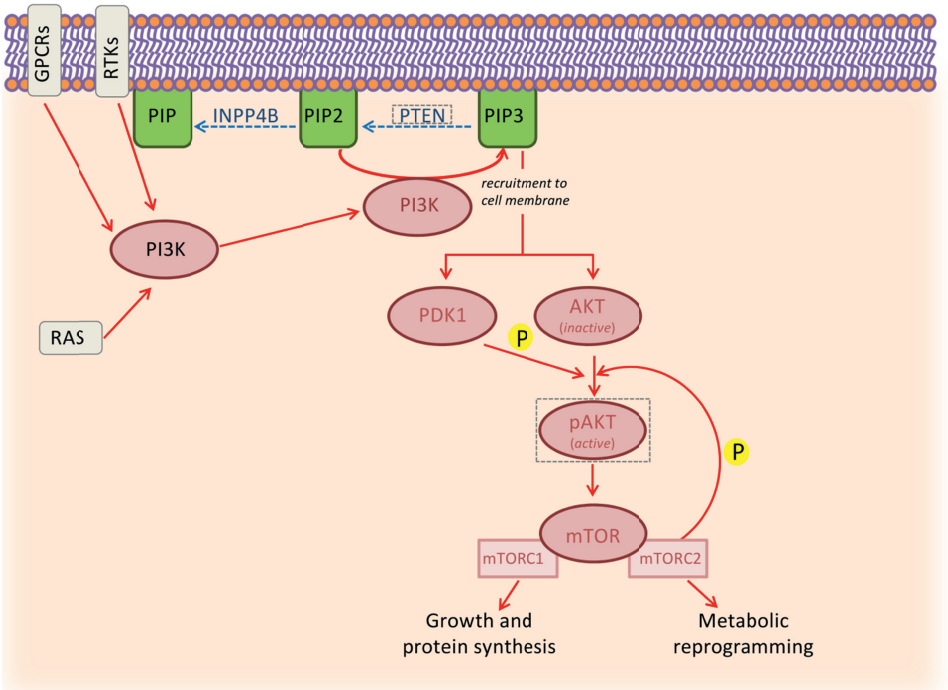


Figure 1. PI3K/Akt signaling pathway in cancer. PI3K can be activated by receptor tyrosine kinases (RTKs), G protein-coupled receptors (GPCRs), and RAS. Activated PI3K phosphorylates phosphatidylinositol 4,5-bisphosphate (PIP2), a plasma membrane phospholipid, and converts it into phosphatidylinositol 3,4,5-trisphosphate (PIP3). PIP3 binds to 3-phosphoinositide dependent protein kinase-1 (PDK1) and Akt, recruiting them to the plasma membrane. This process is negatively regulated by PTEN and INPP4B, which convert PIP3 back to PIP2, and PIP2 back to phosphatidylinositol phosphate (PIP), respectively. Once at the plasma membrane, Akt is activated via phosphorylation by PDK1. Activated Akt then phosphorylates mTOR (mammalian target of rapamycin), which acts as a catalytic subunit in the protein complexes of mTORC1 and mTORC2. mTORC1 is involved in growth and protein synthesis, while mTORC2 regulates metabolic reprogramming and activates Akt, creating a positive feedback loop. In the present study, prediction of the expression of the proteins PTEN and pAKT (in gray, dashed boxes) using metabolic data was attempted.

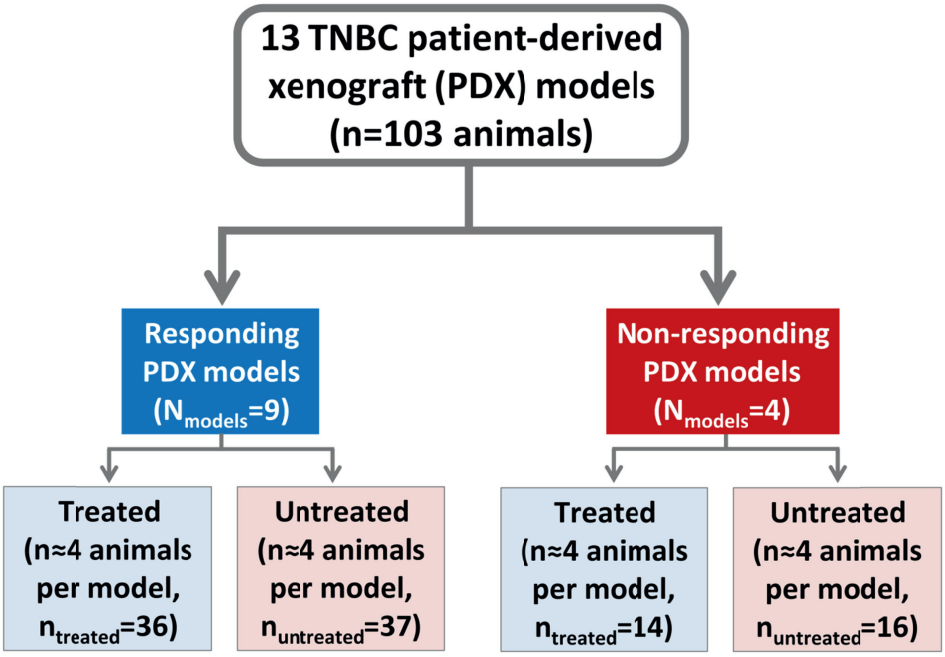


Figure 2. Study Design. Thirteen triple negative breast cancer (TNBC) patient-derived xenograft (PDX) models were included in the study. Nine PDX models were found to respond to everolimus based on significant differences in relative tumor volume between treated animals and untreated controls. For each PDX model, approximately 4 animals were treated with everolimus and 4 were untreated controls. The total number of samples obtained were $n_{treated}$=50, and $n_{untreated}$=53.

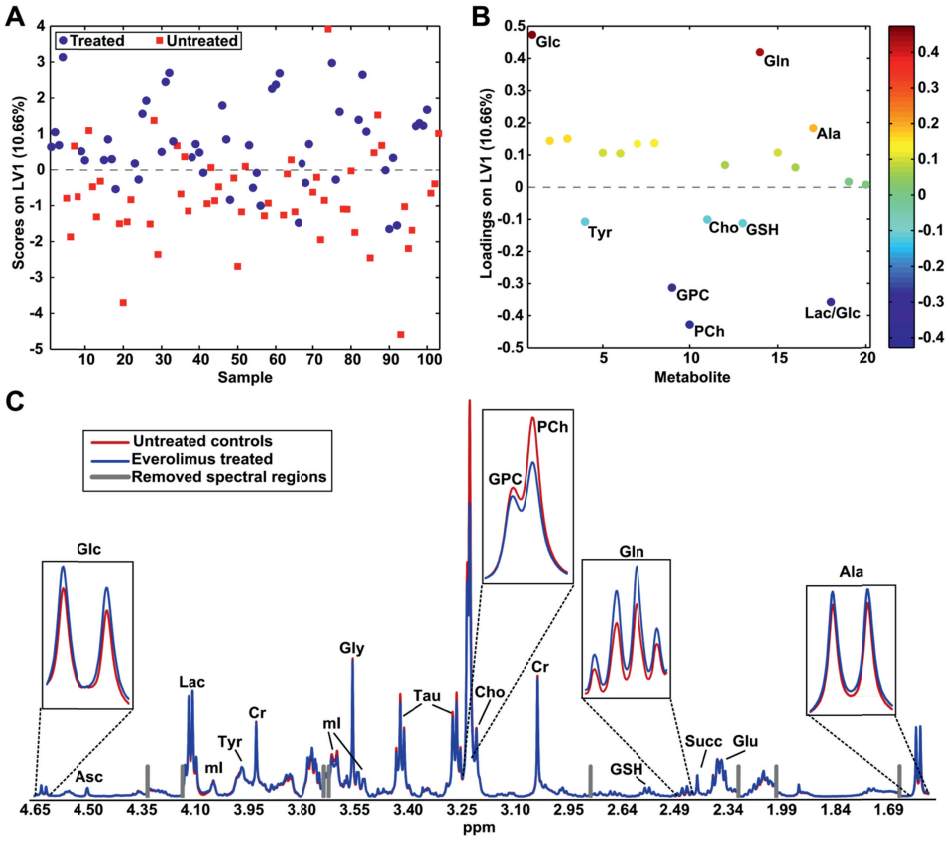


Figure 3. PLS-DA scores (A) and loadings plots (B) of treated animals vs untreated controls (n=103), and mean spectra (C) of treated animals and untreated controls. Loadings are colored according to latent variable (LV) 1. Ala: alanine; Asc: ascorbate; Cho: choline; Cr: creatine; Glc: glucose; Gln: glutamine; Glu: glutamate; Gly: glycine; GPC: glycerophosphocholine; GSH: glutathione; Lac: lactate; mI: myo-inositol; PCh: phosphocholine; Succ: succinate; Tau: taurine; Tyr: tyrosine.

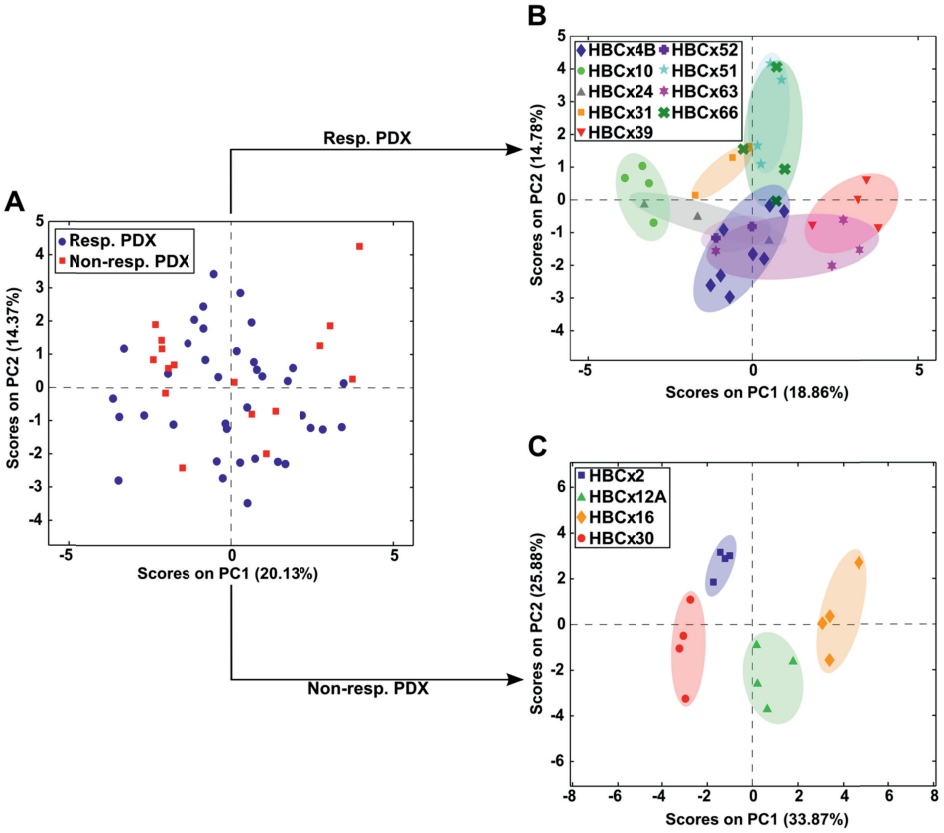


Figure 4. PCA of all untreated controls (A), untreated responding (Resp.) patient-derived xenografts (PDX) only (B), and untreated non-responding (Non-resp.) PDX only (C). Samples in B and C are colored according to PDX model. Ovals were drawn manually to illustrate clusters.

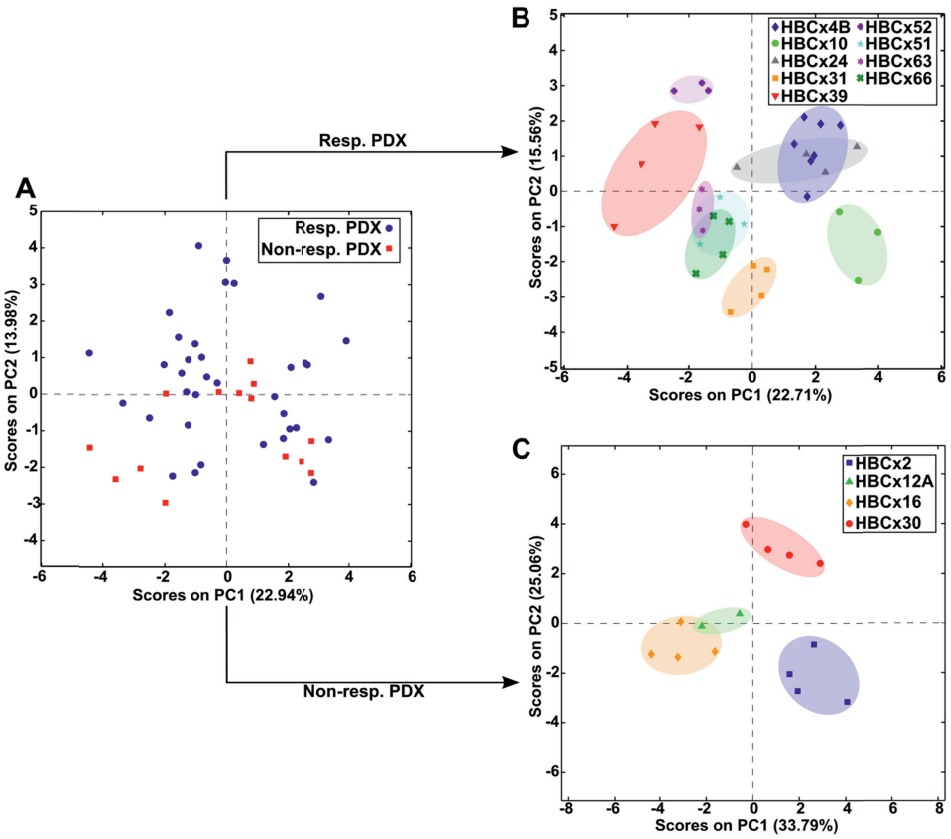


Figure 5. PCA of all treated animals (A), treated responding (Resp.) patient-derived xenografts (PDX) only (B), and treated non-responding (Non-resp.) PDX only. Samples in B and C are colored according to PDX model. Ovals were drawn manually to illustrate clusters.

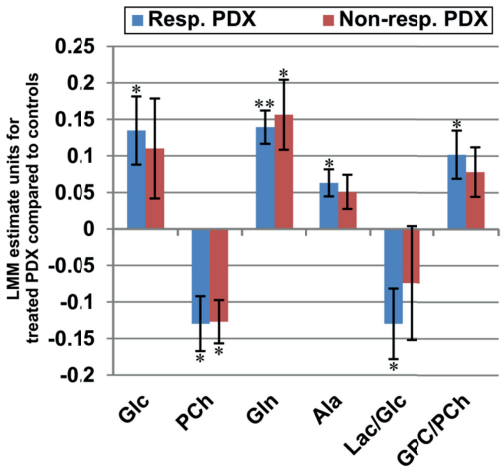


Figure6. LMM estimates for treated animals compared to untreated controls carried out separately for responding models (Resp. PDX) and non-responding models (Non-resp. PDX). * and ** indicate significance (q≤0.05 and q≤0.001, respectively). Ala: alanine; Glc: glucose; Gln: glutamine; GPC: glycerophosphocholine; Lac: lactate; PCh: phosphocholine.

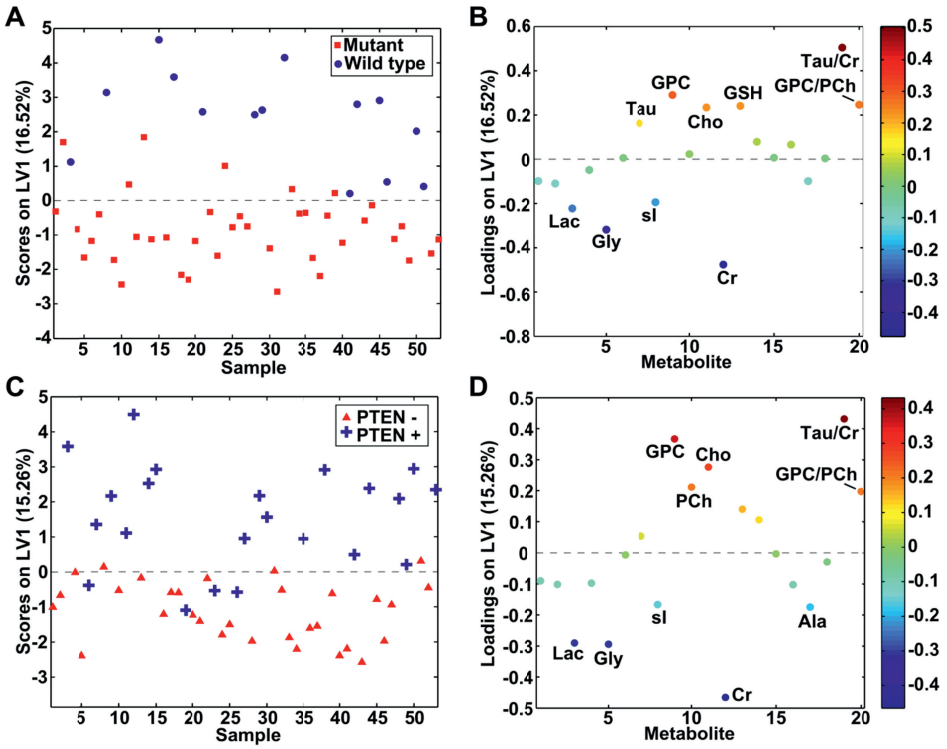


Figure 7. PLS-DA scores (A) and loadings plots (B) of p53 mutant vs p53 wild type untreated controls, and PLS-DA scores (C) and loadings plots (D) of PTEN- vs PTEN+ untreated controls. Loadings are colored according to latent variable (LV) 1. Ala: alanine; Cho:Choline; Cr: creatine; Gly: glycine; GPC: glycerophosphocholine; GSH:glutathione; Lac: lactate; PCh: phosphocholine; sI: scyllo-inositol; Tau: taurine.

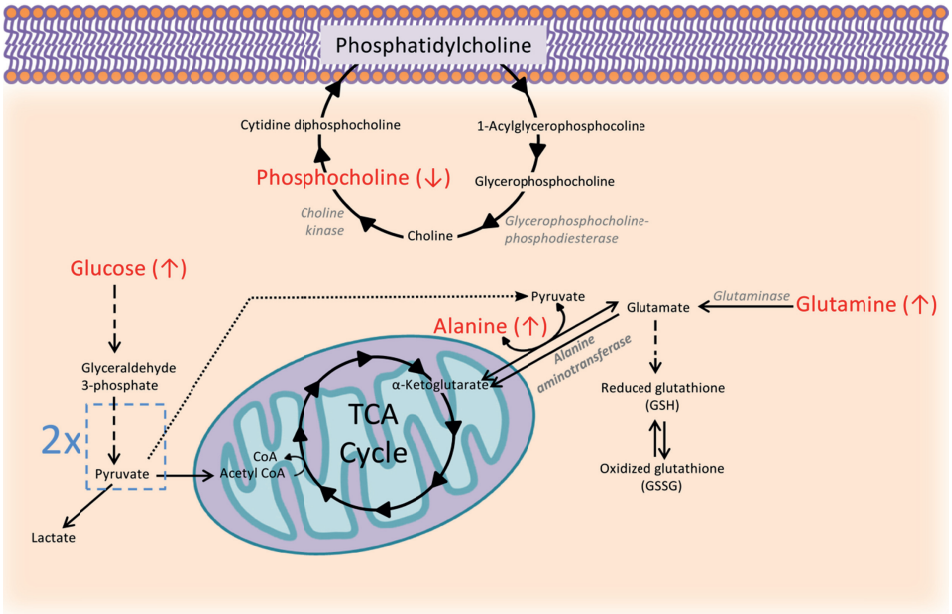


Figure 8. Overview of pathways involving metabolites found to be significantly different between treated animals and untreated controls. Significantly different metabolites are marked in red, with arrows indicating the trend in treated animals with respect to untreated controls. Related enzymes are marked in gray.

Supplementary Information

Article Title: Metabolic response to everolimus in patient-derived xenografts of triple negative breast cancer

Supplementary File 1: Supplementary Table 1.

Supplementary Table 1.

Supplementary Table 1. Metabolite relative levels (log10 transformed) per PDX model.

PDX model	Treatment Group	n	Glc 4.635 - 4.612 ppm		Asc 4.505 - 4.497 ppm		Lac 4.129 - 4.086 ppm		Tyr 3.986 -3.944 ppm		Gly 3.557 - 3.546 ppm		mI 3.539 - 3.513 ppm		Tau 3.432 - 3.402 ppm		sI 3.355 - 3.347 ppm		GPC 3.232 - 3.228 ppm		PCh 3.224 - 3.217 ppm		Cho 3.207 - 3.198 ppm	
			Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HBCx2	Treated	4	1.408	0.193	1.116	0.080	2.840	0.101	2.251	0.080	2.319	0.049	1.848	0.079	2.654	0.092	1.405	0.315	2.118	0.094	2.817	0.104	2.065	0.076
	Untreated	4	1.429	0.112	1.091	0.090	2.861	0.082	2.299	0.030	2.363	0.031	1.921	0.096	2.631	0.047	1.290	0.187	2.126	0.082	2.912	0.091	2.124	0.099
HBCx12A	Treated	2	1.649	0.055	1.194	0.072	2.697	0.052	2.461	0.015	2.415	0.098	2.136	0.001	2.569	0.041	0.894	0.101	2.460	0.033	2.682	0.002	2.089	0.074
	Untreated	4	1.655	0.148	1.167	0.131	2.713	0.070	2.518	0.036	2.456	0.052	2.032	0.021	2.611	0.042	0.919	0.152	2.381	0.036	2.712	0.032	2.140	0.046
HBCx16	Treated	4	1.640	0.114	1.043	0.021	2.770	0.124	2.345	0.044	2.095	0.025	2.092	0.056	2.626	0.060	1.048	0.087	2.434	0.173	2.633	0.136	2.254	0.260
	Untreated	4	1.381	0.304	0.954	0.149	2.711	0.060	2.363	0.029	2.081	0.035	2.037	0.099	2.638	0.062	0.967	0.092	2.609	0.070	2.806	0.042	2.329	0.181
HBCx30	Treated	4	1.645	0.164	1.125	0.087	2.904	0.055	2.430	0.113	2.273	0.028	2.406	0.081	2.489	0.027	1.131	0.237	1.808	0.106	2.509	0.040	1.984	0.027
	Untreated	4	1.438	0.207	1.092	0.058	2.792	0.115	2.307	0.041	2.343	0.060	2.540	0.090	2.506	0.037	1.052	0.076	1.857	0.079	2.689	0.059	1.961	0.127
HBCx39	Treated	4	1.541	0.246	1.156	0.049	2.640	0.116	2.578	0.093	2.226	0.058	2.373	0.041	2.550	0.128	0.845	0.056	2.573	0.074	2.619	0.071	2.205	0.035
	Untreated	4	1.395	0.176	1.081	0.058	2.616	0.080	2.597	0.023	2.205	0.085	2.343	0.051	2.552	0.081	0.895	0.173	2.619	0.078	2.707	0.109	2.146	0.076
HBCx31	Treated	4	1.858	0.076	1.331	0.067	2.695	0.045	2.353	0.021	2.449	0.051	1.980	0.018	2.496	0.069	1.029	0.097	2.309	0.073	3.025	0.032	2.243	0.041
	Untreated	4	1.727	0.176	1.282	0.028	2.689	0.036	2.372	0.024	2.412	0.050	1.927	0.042	2.511	0.075	0.949	0.086	2.408	0.049	3.091	0.051	2.224	0.094
HBCx66	Treated	4	1.763	0.058	1.320	0.044	2.649	0.035	2.474	0.027	2.320	0.035	2.201	0.070	2.600	0.025	1.515	0.427	2.390	0.038	2.559	0.037	2.126	0.096
	Untreated	4	1.838	0.355	1.369	0.101	2.695	0.061	2.446	0.116	2.291	0.069	2.252	0.120	2.648	0.054	1.502	0.328	2.331	0.147	2.674	0.158	2.167	0.078
HBCx10	Treated	4	1.540	0.126	1.343	0.159	2.760	0.056	2.269	0.363	2.429	0.313	1.849	0.066	2.630	0.335	1.310	0.320	1.892	0.211	2.653	0.425	1.976	0.262
	Untreated	4	1.475	0.236	1.352	0.049	2.776	0.147	2.506	0.053	2.561	0.067	1.828	0.037	2.443	0.103	1.180	0.311	1.874	0.089	2.904	0.137	2.132	0.101
HBCx51	Treated	3	1.101	0.033	1.190	0.024	2.615	0.039	2.164	0.072	2.284	0.055	2.207	0.091	2.500	0.091	0.968	0.186	2.673	0.062	2.872	0.091	2.258	0.074
	Untreated	2	1.053	0.128	1.153	0.065	2.642	0.090	2.173	0.002	2.327	0.009	2.152	0.042	2.362	0.021	1.032	0.101	2.667	0.009	3.057	0.020	2.204	0.061
HBCx4B	Treated	7	1.502	0.191	1.217	0.089	2.904	0.076	2.520	0.115	2.359	0.044	1.936	0.057	2.600	0.065	0.952	0.087	2.102	0.173	2.466	0.165	2.200	0.077
	Untreated	8	1.313	0.180	1.188	0.111	2.850	0.043	2.497	0.098	2.277	0.061	1.916	0.111	2.724	0.116	0.940	0.077	2.212	0.110	2.581	0.207	2.181	0.097
HBCx52	Treated	3	1.672	0.137	1.278	0.053	2.694	0.045	2.453	0.101	2.272	0.070	1.970	0.140	2.751	0.045	1.299	0.159	2.009	0.068	2.609	0.018	2.213	0.062
	Untreated	4	1.307	0.318	1.505	0.091	2.715	0.083	2.389	0.050	2.127	0.034	1.928	0.083	2.824	0.057	0.977	0.054	2.081	0.057	2.834	0.033	2.200	0.065
HBCx24	Treated	4	1.558	0.121	1.457	0.124	2.912	0.134	2.379	0.053	2.384	0.027	2.268	0.118	2.639	0.078	1.064	0.160	2.207	0.238	2.671	0.163	2.227	0.041
	Untreated	3	1.395	0.301	1.433	0.154	2.846	0.232	2.362	0.167	2.332	0.079	2.310	0.134	2.566	0.125	1.345	0.149	2.208	0.306	2.836	0.246	2.172	0.116
HBCx62	Treated	3	1.431	0.111	1.251	0.035	2.679	0.036	2.474	0.077	2.160	0.042	1.926	0.114	2.619	0.039	1.258	0.229	2.573	0.073	2.680	0.141	2.166	0.081
	Untreated	4	1.280	0.273	1.159	0.128	2.731	0.151	2.480	0.068	2.078	0.097	1.876	0.172	2.591	0.102	1.229	0.360	2.499	0.229	2.690	0.134	2.104	0.151

The chemical shift interval of fixed regions integrated to calculate the relative level for each metabolite is additionally reported. For metabolite ratios, the mean of the chemical shift integral is reported for the numerator over that of the denominator. PDX: patient-derived xenograft;

Std. Dev.: standard deviation; Glc: glucose; Asc: ascorbate; Lac: lactate; Tyr: tyrosine; Gly: glycine; mI: myo-inositol; Tau: taurine; sI: scyllo-inositol; GPC: glycerophosphocholine; PCh: phosphocholine; Cho: choline; Cr: creatine; GSH: glutathione; Succ: succinate; Glu: glutamate; Ala: alanine

Supplementary Table 1. Metabolite relative levels (log10 transformed) per PDX model.

PDX model	Treatment Group	n	Cr 3.042 - 3.021 ppm		GSH 2.582 - 2.538 ppm		Gln 2.478 - 2.429 ppm		Succ 2.422 - 2.414 ppm		Glu 2.386 - 2.334 ppm		Ala 1.509 - 1.487 ppm		Lac/Glc 4.107/4.624 ppm		Tau/Cr 3.417/3.031 ppm		GPC/PCh 3.230/3.221 ppm	
			Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HBCx2	Treated	4	2.429	0.067	1.820	0.244	1.688	0.058	1.506	0.155	2.483	0.031	2.485	0.043	1.432	0.242	0.225	0.061	-0.699	0.114
	Untreated	4	2.394	0.042	1.671	0.165	1.561	0.030	1.516	0.077	2.501	0.036	2.494	0.050	1.431	0.042	0.237	0.063	-0.786	0.101
HBCx12A	Treated	2	2.183	0.059	1.919	0.019	1.654	0.109	1.510	0.062	2.572	0.039	2.343	0.042	1.047	0.106	0.385	0.100	-0.222	0.031
	Untreated	4	2.242	0.037	1.892	0.089	1.616	0.145	1.484	0.097	2.590	0.019	2.258	0.078	1.058	0.206	0.369	0.051	-0.332	0.031
HBCx16	Treated	4	1.946	0.110	1.903	0.138	1.733	0.030	1.344	0.022	2.563	0.008	2.339	0.061	1.131	0.180	0.679	0.071	-0.200	0.075
	Untreated	4	1.916	0.073	1.927	0.029	1.505	0.148	1.463	0.047	2.547	0.048	2.271	0.093	1.330	0.258	0.722	0.039	-0.197	0.061
HBCx30	Treated	4	2.328	0.032	1.818	0.061	1.856	0.253	1.500	0.039	2.668	0.047	2.412	0.054	1.259	0.217	0.161	0.042	-0.701	0.113
	Untreated	4	2.403	0.021	1.910	0.073	1.664	0.126	1.480	0.060	2.719	0.068	2.344	0.063	1.354	0.321	0.103	0.045	-0.832	0.130
HBCx39	Treated	4	2.182	0.032	1.875	0.034	1.946	0.081	1.492	0.195	2.419	0.042	2.194	0.078	1.099	0.316	0.368	0.152	-0.046	0.050
	Untreated	4	2.203	0.090	1.910	0.067	1.754	0.097	1.465	0.030	2.471	0.036	2.107	0.079	1.221	0.192	0.349	0.071	-0.088	0.094
HBCx31	Treated	4	2.417	0.030	1.802	0.060	1.941	0.052	1.440	0.073	2.387	0.043	2.186	0.040	0.837	0.094	0.078	0.061	-0.717	0.058
	Untreated	4	2.464	0.033	1.778	0.039	1.884	0.044	1.466	0.109	2.301	0.025	2.152	0.053	0.962	0.202	0.048	0.102	-0.683	0.101
HBCx66	Treated	4	2.343	0.043	1.899	0.054	1.685	0.034	1.517	0.086	2.617	0.032	2.268	0.048	0.886	0.081	0.256	0.061	-0.169	0.053
	Untreated	4	2.342	0.086	1.801	0.129	1.683	0.156	1.534	0.092	2.561	0.102	2.196	0.040	0.857	0.365	0.306	0.131	-0.343	0.032
HBCx10	Treated	4	2.557	0.331	1.661	0.196	1.958	0.122	1.538	0.102	2.547	0.359	2.366	0.204	1.220	0.101	0.073	0.069	-0.761	0.244
	Untreated	4	2.408	0.059	1.650	0.112	1.755	0.018	1.427	0.115	2.717	0.038	2.329	0.010	1.301	0.196	0.036	0.154	-1.030	0.085
HBCx51	Treated	3	2.058	0.055	1.764	0.029	1.832	0.075	1.533	0.029	2.534	0.078	2.426	0.065	1.513	0.008	0.442	0.046	-0.199	0.029
	Untreated	2	1.990	0.020	1.685	0.163	1.662	0.005	1.728	0.060	2.579	0.053	2.524	0.006	1.589	0.038	0.372	0.042	-0.390	0.029
HBCx4B	Treated	7	2.419	0.064	1.698	0.218	1.783	0.114	1.592	0.057	2.564	0.034	2.515	0.067	1.403	0.192	0.181	0.052	-0.364	0.069
	Untreated	8	2.507	0.125	1.872	0.196	1.624	0.121	1.565	0.087	2.504	0.061	2.423	0.095	1.537	0.195	0.217	0.092	-0.369	0.179
HBCx52	Treated	3	2.117	0.048	1.918	0.079	2.263	0.025	1.579	0.058	2.545	0.054	2.327	0.040	1.022	0.171	0.635	0.057	-0.600	0.084
	Untreated	4	2.127	0.016	1.915	0.027	2.047	0.079	1.461	0.073	2.596	0.038	2.237	0.048	1.409	0.401	0.697	0.056	-0.754	0.033
HBCx24	Treated	4	2.460	0.097	1.711	0.172	1.620	0.138	1.664	0.083	2.382	0.088	2.478	0.079	1.354	0.078	0.179	0.080	-0.464	0.317
	Untreated	3	2.449	0.117	1.603	0.111	1.488	0.125	1.549	0.106	2.258	0.170	2.379	0.025	1.451	0.084	0.116	0.012	-0.628	0.161
HBCx62	Treated	3	2.519	0.017	2.021	0.015	1.622	0.079	1.439	0.021	2.568	0.024	2.164	0.043	1.248	0.137	0.100	0.024	-0.107	0.181
	Untreated	4	2.456	0.099	1.830	0.181	1.489	0.075	1.340	0.089	2.460	0.072	2.094	0.086	1.451	0.292	0.135	0.023	-0.192	0.179

The chemical shift interval of fixed regions integrated to calculate the relative level for each metabolite is additionally reported. For metabolite ratios, the mean of the chemical shift integral is reported for the numerator over that of the denominator. PDX: patient-derived xenograft;

Std. Dev.: standard deviation; Glc: glucose; Asc: ascorbate; Lac: lactate; Tyr: tyrosine; Gly: glycine; mI: myo-inositol; Tau: taurine; sI: scyllo-inositol; GPC: glycerophosphocholine; PCh: phosphocholine; Cho: choline; Cr: creatine; GSH: glutathione; Succ: succinate; Glu: glutamate; Ala: alanine