Identification of novel genetic variants associated with cardiorespiratory fitness☆

Anja Bye a,b,⁎, Marie Klevjer a,b, Einar Ryeng c, Gustavo Jose Justo da Silva d, Jose Bianco Nascimento Moreira a,b, Dorte Stensvold a, Ulrik Wissloff a,e

a Cardiac Exercise Research Group (CERG) at Dept. of Circulation and Medical Imaging, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
b Department of Cardiology, St. Olavs Hospital, Trondheim, Norway
c Genomics Core Facility (GCF), Norwegian University of Science and Technology (NTNU), Trondheim, Norway
d Institute for Experimental Medical Research, Oslo University Hospital and University of Oslo, Oslo, Norway
e School of Human Movement & Nutrition Sciences, University of Queensland, Brisbane, Australia

ARTICLE INFO

Article history:
3 February 2020
3 February 2020

Keywords:
Maximal oxygen uptake
Genetics
Cardiovascular disease risk

ABSTRACT

Introduction: Low maximal oxygen uptake (VO2max) is a strong and independent risk factor for all-cause and cardiovascular disease (CVD) mortality. For other CVD risk factors, numerous genetic association studies have been performed, revealing promising risk markers and new therapeutic targets. However, large genomic association studies on VO2max are still lacking, despite the fact that VO2max has a large genetic component.

Methods: We performed a genetic association study on 123,545 single-nucleotide polymorphisms (SNPs) and directly measured VO2max in 3470 individuals (exploration cohort). Candidate SNPs from the exploration cohort were analyzed in a validation cohort of 718 individuals, in addition to 7 wild-card SNPs. Sub-analyses were performed for each gender. Validated SNPs were used to create a genetic score for VO2max. In silico analyses and genotype-phenotype databases were used to predict physiological function of the SNPs.

Results: In the exploration cohort, 41 SNPs were associated with VO2max (p < 5.0 × 10⁻⁴). Six of the candidate SNPs were associated with VO2max also in the validation cohort, in addition to three wild-card SNPs (p < 0.05, in men, women or both). The cumulative number of high-VO2max-SNPs correlated negatively with CVD risk factors, e.g. waist-circumference, visceral fat, fat %, cholesterol levels and BMI. In silico analysis indicated that several of the VO2max-SNPs influence gene expression in adipose tissue, skeletal muscle and heart.

Conclusion: We discovered and validated new SNPs associated with VO2max and proposed possible links between VO2max and CVD. Studies combining several large cohorts with directly measured VO2max are needed to identify more SNPs associated with this phenotype.

© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

Introduction ................................................................. 342
Material and methods ....................................................... 342
Study participants ............................................................ 342
Clinical measurements ...................................................... 342
Testing maximal oxygen uptake (VO2max) .................................... 342
Questionnaire-based information ........................................... 342
Blood analysis ............................................................... 343

Abbreviations and acronyms: ACTN3, alpha-actinin-3; ADRB3, beta-3 adrenergic receptor; APOA1, apolipoprotein A1; APOER2, apolipoprotein E receptor 2; BAHDI, Bromo adjacent homology domain containing 1; BMI, body mass index; CRP, C-reactive protein; CVD, cardiovascular disease; DNA, deoxyribonucleic acid; EDN1, endothelin 1; EROs, estrogen receptor alpha; GWAS, genome-wide association studies; HDL, high-density lipoprotein; HUNT, Nord-Trøndelag Health Study; IGF2, insulin-like growth factor 2; KCNQ1, potassium voltage-gated channel subfamily Q member 1; LDL, low-density lipoprotein; LDLR, low density lipoprotein receptor; LXR, liver X receptor; MAF, minor allele frequency; MET, metabolic equivalent; MYLIP, myosin regulatory light chain interacting protein; MYOC, myocardin; PROX1, Prospero homeobox protein 1; SNP, single-nucleotide polymorphism; VIP, vasoactive intestinal peptide; VIPR2, vasoactive intestinal peptide receptor 2; VLDLR, very low density lipoprotein receptor; VO2max, maximal oxygen uptake.

⁎ Statement of Conflict of Interest: see page 348.

⁎ Address reprint requests to Anja Bye, Department of Circulation and Medical Imaging, PO Box 8905, Medical Technical Research Center, 7491 Trondheim, Norway.

E-mail address: Anja.Bye@ntnu.no (A. Bye).

https://doi.org/10.1016/j.pcad.2020.02.001
0033-0620/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Genotyping of exploration cohort ........................................................... 343
Genotyping of validation cohort ........................................................... 343
In silico analysis of transcription starting sites ........................................ 343
Genotype-Tissue Expression (GTEx) database ........................................ 343
BXD mouse database ........................................................................... 343
Statistical analyses ............................................................................. 343
Results ................................................................................................. 344
Discussion ........................................................................................... 346
Limitations ............................................................................................ 348
Conclusion ............................................................................................. 348
Statement of conflict of interest ............................................................ 348
Acknowledgements ............................................................................... 348
References ........................................................................................... 348

Introduction

Low aerobic fitness, quantified as maximal oxygen uptake (VO2max), is a strong and independent predictor of all-cause and cardiovascular mortality in healthy individuals and in patients with cardiovascular disease (CVD).1–4 VO2max is determined by a combination of genetic and environmental factors, and the genetic contribution is suggested to be ~50%.5,6 Identification of genes and genomic variations associated with VO2max would lead to a better understanding of this complex trait, and provide possible links between VO2max and CVD. Previously, a few genes and genomic loci have been associated with VO2max.7–9 However, most studies are limited in size and employ the conventional hypothesis-driven approach of searching for pre-specified genomic associations, which limits the discovery of new genetic loci. Hence, the scientific community call for a large-scale systematic screening of genetic variants associated with directly measured VO2max in a large well-characterized population.10

By taking advantage of one of the world’s largest databases of directly measured VO2max, we report the first large-scale systematic screening for genetic variants associated with VO2max. Furthermore, we explore potential associations between VO2max-related single-nucleotide polymorphisms (SNPs) and CVD risk factors, and their potential biological implications by using in silico tools and genotype-phenotype databases.

Material and methods

Study participants

The Nord-Trøndelag Health Study (HUNT) is one of the largest health studies ever performed. It includes a unique database of questionnaire data, clinical measurements and biological samples. During the past 35 years, 120,000 individuals have contributed throughout four waves of the HUNT study (HUNT1 in 1984–86, HUNT2 in 1995–97, HUNT3 in 2006–08 and HUNT4 in 2017–19) in Norway. Participants in the present study attended a sub project during the third wave of HUNT (HUNT3 Fitness Study) designed to directly measure maximal oxygen uptake (VO2max) in a healthy adult population.11 Exclusion criteria for the HUNT3 Fitness Study were present or previous heart- or lung disease, cancer, and medical contraindication or orthopedic limitation to exercise. This study was approved by the Regional committee for medical research ethics (4.2008.2792), the Nord-Trøndelag Health Study, the Norwegian Data Inspectorate, and by the National Directorate of Health. The study was in conformity with Norwegian laws and the Helsinki declaration, and a signed informed consent was obtained from all participants.

Clinical measurements

Weight and height were measured on a combined scale (Model DS-102, Arctic Heating AS, Netterøy, Norway), and body mass index (BMI) was calculated as weight divided by height squared (kg/m²). Fat, muscle percentage and visceral fat were obtained using the InBody 720 scale (Biospace, Seoul, Korea). Testing maximal oxygen uptake (VO2max)

An individualized protocol was applied to measure VO2max.12,14 Each test-subject was familiarized with treadmill walking during the warm-up of 8–10 min, also to ensure safety and avoid handrail grasp when this was not absolutely necessary. Oxygen uptake kinetics were measured directly by a portable mixing chamber gas-analyzer (Cortex MetaMax II, Cortex, Leipzig, Germany) with the participants wearing a tight face mask (Hans Rudolph, Germany) connected to the MetaMax II device. The system has previously been found reliable and valid in our laboratory. Heart rate was measured by radio telemetry (Polar S610i, Polar Electro Oy, Kempele, Finland). From the warm-up pace, the load was regularly increased. When the participants reached an oxygen consumption that was stable for more than 30 s, treadmill inclination (1–2%) or velocity (0.5–1 km/h) were increased stepwise until the participants were exhausted. A maximal test was achieved when the respiratory quotient reached >1.05 or when the oxygen uptake did not increase >2 ml/kg/min despite increased workload. VO2max was measured as liters of oxygen per minute (l/min), and subsequently calculated as VO2max relative to body mass (ml/kg/min) and VO2max scaled (ml/kg0.75/min).

Questionnaire-based information

Physical activity is likely to be the most important behavioral factor influencing VO2max, and has to be adjusted for to isolate the genetic contribution to the phenotype. Physical activity was registered based on the responses to a self-administered questionnaire.15 The questionnaire included three questions and each participant’s response to the questions (i.e. numbers in brackets) were multiplied to calculate a physical activity index score (Kurtze score): Question 1: “How frequently do you exercise?”, with the response options: “Never” (0), “Less than once a
Genotyping raw data was subjected to systematic quality control using as a continuous variable, as this provides the best statistical power. The

week” (0), “Once a week” (1), “2–3 times per week” (2.5) and “Almost every day” (5). Question 2: “If you exercise as frequently as once or more
times a week: How hard do you push yourself?” with the response options: “I take it easy without breaking a sweat or losing my breath” (1),
“I push myself so hard that I lose my breath and break into sweat” (2) and “I push myself to near exhaustion” (3). Question 3: “How long
does each session last?”, with the response options: “Less than 15 minutes” (0.1), “16–30 minutes” (0.38), “30 minutes to 1 hour” (0.75)
and “More than 1 hour” (1.0). As the second and third question only ad-
dressed people who exercised at least once a week, both “Never” and
“Less than once a week” yielded an index score of zero. Participants
with a zero score were categorized as inactive, 0.05–1.5 as low activity,
1.51–3.75 as medium activity, and 3.76–15.0 as high activity.

Blood analysis

Standard biochemical analyses were performed on fresh venous non-fasting blood samples at Levanger Hospital, Norway. Non-fasting
glucose was analyzed by hexokinase/G-G-PDH methodology reagent kit 3L82-20/3L82-40 Glucose (Abbott Diagnostics, Illinois, US), high-
density lipoprotein (HDL) cholesterol by the Accelerator selective deter-
gent methodology reagent kit 3K33-20 Ultra HDL (Abbott Diagnostics),
total cholesterol by enzymatic cholesterol esterase methodology re-
genkit 7D62-20 Cholesterol (Abbott Diagnostics), triglycerides by
Glycerol Phosphate Oxidase methodology reagent kit 7D74 Triglyceride
(Abbott Diagnostics) and C-reactive protein (CRP) by the Areoset CRP
Vario kit (Abbott Diagnostics). Triglycerides and CRP were measured
in only 80% of the HUNT population. Low-density lipoprotein (LDL)
cholesterol was calculated based on information on total cholesterol, HDL-
cholesterol and triglycerides.

Genotyping of exploration cohort

Deoxyribonucleic acid (DNA) was extracted from blood samples
stored in the HUNT biobank as described elsewhere. DNA samples
were analyzed by the custom-made Cardio-MetaBiochip including ap-
proximately 210.000 SNPs (Illumina, CA, US). The annotation on the
basis is based on Genome build 36.3. The Cardio-Metabochip was de-
signed by representatives of the following genome-wide association
studies (GWAS) meta-analysis consortia: CARDioGRAM (coronary ar-
tery disease), DIAGRAM (type 2 diabetes), GIANT (height and weight),
MAGIC (glycemic traits), Lipids (lipids), ICBP-GWAS (blood pressure),
and QT-ICC (QT interval). The candidate SNPs were selected accord-
ing to five sets of criteria: (I) individual SNPs displaying evidence for associ-
ation in GWAS meta-analysis to diseases and traits relevant to metabolic
and atherosclerotic-cardiovascular endpoints, (II) detailed fine map-
ning of loci validated at genome-wide significance from these meta-
analyses, (III) all SNPs associated at genome-wide significance with any human trait, (IV) “wildcards” selected by each consortium for
consortium-specific purposes, and (V) other useful content, including
SNPs that tag common copy number polymorphisms, SNPs in the
human leukocyte antigen region, SNPs marking the X and Y chromo-
somes and mitochondrial DNA, and for sample fingerprinting.

The study was designed as a quantitative trait approach with VO2max
as a continuous variable, as this provides the best statistical power. The
genotyping raw data was subjected to systematic quality control using the
statistical software PLINK. Individuals with low genotype call rate (~90%) were excluded. SNPs with a genotype call rate <95% or a
minor allele frequency <1% were also excluded. Furthermore, SNPs
that clearly deviate from the expected Hardy-Weinberg Equilibrium
were excluded (p < 10−7). Individuals who showed gender discrepan-
cies based on the heterozygosity rate from chromosome X were also ex-
duced. After pre-processing the raw data, 123.545 SNPs were left for
association analyses.

Genotyping of validation cohort

Candidate SNPs from the exploration cohort, as well as a 7 wild-card
SNPs not included on the Cardio-MetaBiochip, were genotyped using the
Agena Biosciences MassARRAY® platform (Agena Bioscience, San
Diego, CA, US). SNP multiplexes were designed using Assay Design
Suite v1.0 software (Agena Bioscience). Genotyping was performed ac-
cording to the manufacturer’s protocol using iPLEX Gold assay (Agena
Bioscience) and analyzed using the MassARRAY Analyzer 4 platform.
Mass signals for the different alleles were captured with high accuracy
by matrix-assisted laser desorption/ionization time-of-flight mass spec-
trometry (MALDI-TOF MS). Genotype clustering and individual sample
 genotype calls were generated using Sequenom TyperAnalyzer v.4.0
software (Agena Bioscience).

In silico analysis of transcription starting sites

To determine if validated SNPs were located in transcription factors
binding sites, we performed in silico analysis of predicted transcription
factor binding sites using the software PROMO.17

Genotype-Tissue Expression (GTEx) database

To explore the relationship between validated SNPs and gene ex-
pression in different human tissues/organisms, we used the GTEx database.
The database includes ~900 post-mortem donors and opens the possi-
bility for studying the effects of genetic variation in multiple human ref-
rence tissues.18

BXD mouse database

The BXD database is an open-access web service for systems ge-
netics (www.genenetwork.org) to explore the genetic control of
multiple phenotypes.19 The database includes ~2000 phenotypes
across a large panel of isogenic but diverse strains of mice (BXD type).
The database contains phenotypes such as heart rate, oxygen
consumption and blood parameters, such as hematocrit and iron
levels, highly relevant for exploring the functional importance of
VO2max-related genes. We tested potential correlations between expres-
sion levels of genes under the influence of VO2max—SNPs and rel-
levant phenotypes in mouse on both chow diet and high fat diet,
independently.

Statistical analyses

The association between the final 123.545 SNPs and VO2max were
analyzed by linear regression using PLINK. The main covariates for the
VO2max phenotype were gender, age (years) and physical activity level
(Kurtze score). The cut-off for significance were set to (p < 5.0 × 10−4)
in the exploration cohort, as findings reaching the traditional genome-
wide significance were considered unlikely due to the low number of
available cases and that VO2max is a complex trait. To overcome the
issue of using a moderately stringent p-value, validation of the find-
ings in a separate cohort was necessary. In the validation cohort, associations
between VO2max and candidate SNPs were tested using the same statisti-
cal analyses as in the exploration cohort. Nominal p-value was consid-
ered significant (p < 0.05). A genetic score was created using a
combination of 9 SNPs associated with VO2max. Each participant was
scored according to the sum of high VO2max genotypes carried. The dif-
fferences in VO2max between participants with increasing numbers of fa-
orable genotypes were calculated by one-way ANOVA using the LSD
post hoc test.
After filtration of genotyping data, 123,545 SNPs were tested for their association with VO2max. 41 SNPs were significantly associated with VO2max in the exploration cohort after adjusting for age, gender and physical activity level ($p < 5.0 \times 10^{-4}$). Locus zoom plots can be found in Supplementary Fig. 1. The candidate SNPs were subsequently genotyped in a validation cohort, in addition to 7 wild-card SNPs not included on the chip used for the exploration cohort. The association between VO2max and six novel SNPs were replicated in the validation cohort ($p < 0.05$, Table 2). The SNP in the promoter region of the Myosin Regulatory Light Chain Interacting Protein (MYLIP) (rs3757354) did not pass the significance threshold in the validation cohort, however sub analyses for each gender showed a highly significant association in women, and the SNP was therefore included in Table 2. Three of the 7 wild-card SNPs in the genes beta-3 adrenergic receptor (ADRB3), alpha-actinin-3 (ACTN3) and endothelin 1 (EDN1) were associated with VO2max in men, women or both genders ($p < 0.05$, Table 2). Candidate SNPs that failed to be replicated in the validation cohort can be found in Supplementary Table 1.

Considering that VO2max is a complex trait influenced by multiple genetic factors, we assessed whether a cumulative effect existed between the number of favorable genotypes and VO2max. By using a combination of the 9 SNPs from Table 2, and scoring the high VO2max-associated genotypes 1 and low VO2max genotypes 0, we calculated a genetic score for each participant estimating inborn VO2max. In the validation cohort, the variations in VO2max ranged from 63 ml/kg$^{0.75}$/min to 98 ml/kg$^{0.75}$/min, for participants scoring 1 or 7, respectively (Fig. 1A). This corresponded to unscaled VO2max-values ranging from 22.3 ml/ kg/min to 32.7 ml/kg/min, for participants scoring 1 or 7, respectively. To illustrate that the power of this allele combination was independent of physical activity levels, we split the participants into two subgroups, participants below (inactive) and above (active) the median physical activity level. Interestingly, the proposed score appears to be robust even with the reduced sample power of this sub-analysis (Fig. 1B).

Using the same SNPs as basis, we also found a cumulative effect of the number of favorable SNPs and the decline in several CVD risk factors, e.g. waist circumference, visceral fat, % body fat, cholesterol and BMI (Fig. 2). In addition, among the participants with 1–4 favorable SNPs, 36% were on treatment for hypertension, compared to 23% in those with >4 favorable SNPs ($p < 0.05$). Among participants reporting little or no physical activity (Kurtze score $>3.75, n = 235$) those with 1–4 favorable SNPs had higher fat percentage (+2%), visceral fat (+9%), total cholesterol (+5%) and LDL-cholesterol (+6%) compared to those with no favorable SNPs. 

### Results

Characteristics of the participants in the exploration cohort (HUNT3 Fitness Study) and the validation cohort (Generation 100 Study) are shown in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Participant characteristics</th>
<th>HUNT3 Fitness</th>
<th>Generation 100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exploration cohort</td>
<td>Valuation cohort</td>
</tr>
<tr>
<td>N</td>
<td>3470</td>
<td>718</td>
</tr>
<tr>
<td>Females, No. (%)</td>
<td>1563 (45%)</td>
<td>390 (54%)</td>
</tr>
<tr>
<td>Males, No. (%)</td>
<td>1907 (55%)</td>
<td>328 (46%)</td>
</tr>
<tr>
<td>Age, years</td>
<td>47 (19–84)</td>
<td>73 (70–77)</td>
</tr>
<tr>
<td>Height, cm</td>
<td>173 (148–200)</td>
<td>171 (147–195)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>78 (39–135)</td>
<td>75 (42–132)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>26 (17–44)</td>
<td>25 (18–42)</td>
</tr>
<tr>
<td>Waist circumference, cm</td>
<td>90 (57–134)</td>
<td>93 (68–129)</td>
</tr>
<tr>
<td>Total body fat (%)</td>
<td>–</td>
<td>29 (11–50)</td>
</tr>
<tr>
<td>Visceral fat (cm²)</td>
<td>–</td>
<td>111 (26–270)</td>
</tr>
<tr>
<td>Systolic blood pressure, mmHg</td>
<td>126 (79–190)</td>
<td>135 (90–203)</td>
</tr>
<tr>
<td>Diastolic blood pressure, mmHg</td>
<td>72 (36–117)</td>
<td>74 (47–104)</td>
</tr>
<tr>
<td>Diabetes, No. (%)</td>
<td>48 (1.4%)</td>
<td>17 (2.4%)</td>
</tr>
<tr>
<td>Smoking status, No. (%)</td>
<td>1800 (51.9%)</td>
<td>359 (505)</td>
</tr>
<tr>
<td>Never</td>
<td>458 (13.2%)</td>
<td>51 (7.1%)</td>
</tr>
<tr>
<td>Current</td>
<td>1173 (33.8%)</td>
<td>269 (37.5%)</td>
</tr>
<tr>
<td>Former</td>
<td>39 (11.3%)</td>
<td>39 (54%)</td>
</tr>
<tr>
<td>Biochemical data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol, mmol/l</td>
<td>5.4 (2.3–10.0)</td>
<td>5.6 (2.6–9.9)</td>
</tr>
<tr>
<td>LDL cholesterol, mmol/l</td>
<td>–</td>
<td>3.4 (0.95–5.98)</td>
</tr>
<tr>
<td>HDL cholesterol, mmol/l</td>
<td>1.0 (0.5–3.4)</td>
<td>1.7 (0.6–4.5)</td>
</tr>
<tr>
<td>Triglycerides, mmol/l</td>
<td>1.8 (0.2–7.5)</td>
<td>1.1 (0.3–4.9)</td>
</tr>
<tr>
<td>Glucose, mmol/l</td>
<td>5.4 (2.6–12.4)</td>
<td>5.6 (3.1–10.5)</td>
</tr>
<tr>
<td>High-sensitivity CRP, mg/l</td>
<td>–</td>
<td>2.0 (0.1–36.8)</td>
</tr>
<tr>
<td>Treadmill data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VO2max, ml/kg$^{0.75}$/min</td>
<td>123 (56–222)</td>
<td>94 (42–161)</td>
</tr>
<tr>
<td>Resting heart rate, beats/min</td>
<td>59 (34–120)</td>
<td>64 (40–101)</td>
</tr>
<tr>
<td>Peak heart rate, beats/min</td>
<td>181 (107–231)</td>
<td>159 (96–203)</td>
</tr>
<tr>
<td>Physical activity level, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inactive</td>
<td>421 (12.1%)</td>
<td>52 (7.2%)</td>
</tr>
<tr>
<td>Low</td>
<td>761 (21.9%)</td>
<td>110 (15.3%)</td>
</tr>
<tr>
<td>Medium</td>
<td>1274 (36.7%)</td>
<td>350 (48.8%)</td>
</tr>
<tr>
<td>High</td>
<td>991 (28.6%)</td>
<td>193 (26.9%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>23 (0.7%)</td>
<td>13 (1.8%)</td>
</tr>
</tbody>
</table>

Data is shown as mean (min-max) or as number (percentage of the participants). LDL: Low-density lipoprotein, HDL: High-density lipoprotein, CRP: C-reactive protein, VO2max: Maximal oxygen uptake.

### Table 2

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr.</th>
<th>Proximal gene</th>
<th>SNP location</th>
<th>High VO2max genotype</th>
<th>Minor allele</th>
<th>Exploration cohort</th>
<th>Validation cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MAF</td>
<td>BETA</td>
</tr>
<tr>
<td>rs10494973</td>
<td>1</td>
<td>PROX1</td>
<td>Intron</td>
<td>GG</td>
<td>C</td>
<td>0.03</td>
<td>−1.01</td>
</tr>
<tr>
<td>rs5370</td>
<td>6</td>
<td>EDN1</td>
<td>Intron</td>
<td>TT/TG</td>
<td>T</td>
<td>0.23</td>
<td>−0.74</td>
</tr>
<tr>
<td>rs3757354</td>
<td>6</td>
<td>MYLIP</td>
<td>Upstream</td>
<td>GG</td>
<td>A</td>
<td>0.03</td>
<td>−4.68</td>
</tr>
<tr>
<td>rs6950857</td>
<td>7</td>
<td>VIPK2</td>
<td>Intron</td>
<td>GG</td>
<td>A</td>
<td>0.03</td>
<td>−1.53</td>
</tr>
<tr>
<td>rs4994</td>
<td>8</td>
<td>ADRB3</td>
<td>Intron</td>
<td>CC/CT</td>
<td>C</td>
<td>0.07</td>
<td>1.75</td>
</tr>
<tr>
<td>rs1815739</td>
<td>11</td>
<td>ACTN3</td>
<td>Intron</td>
<td>TT/TG</td>
<td>C</td>
<td>0.45</td>
<td>−0.73</td>
</tr>
<tr>
<td>rs2074238</td>
<td>11</td>
<td>KCNJ1</td>
<td>Intron</td>
<td>TT</td>
<td>T</td>
<td>0.09</td>
<td>−1.55</td>
</tr>
</tbody>
</table>

SNPs: Single-nucleotide polymorphisms, VO2max: Maximal oxygen uptake, Chr.: Chromosome, MAF: Minor allele frequency, PROX1: Prospero homeobox protein 1, EDN1: Endothelin 1, MYLIP: Myosin regulatory light chain interacting protein, VIPK2: Vasointestinal peptide receptor 2, ADRB3: Beta-3 adrenergic receptor, ACTN3: Alpha-actinin-3, KCNJ1: Potassium voltage-gated channel subfamily Q member 1, BAHD1: Bromo adjacent homology domain containing 1, MYOCD: Myocardin, NS: Not significant NC: Non-coding.
To explore and predict physiological consequences of the VO2max-SNPs, we used in silico tools and genotype-phenotype databases. The non-synonymous SNP rs3803357, located in the first exon of the Bromo adjacent homology domain containing 1 (BAHD1) gene, cause a shift from the amino acid glycine to lysine. The group of participants homozygote for the rs3803357 minor allele (TT) (24%) had a 3 ml/kg0.75/min lower VO2max than the group carrying the heterozygote allele (GT) (50%) or the common allele homozygotes (GG) 26% (Fig. 3A). In the validation cohort, the group of participant's homozygote for the rs3803357 minor allele (TT) (24%) had 4% and 7% lower VO2max compared to those harboring the (GT) and (GG) variants, respectively (Fig. 3B). SNPs located outside the promoter region or within introns and exons may influence transcription of proximal genes. Using the Genotype-Tissue Expression (GTEx) database, rs3803357 was found to be associated with differential expression BAHD1 in the left ventricle (p = 9.0e−9) (Fig. 3C). By using the BXD mice population, we found significant negative correlations between cardiac expression of Bahd1 and basal VO2 (in an untrained state), as well as with myocardial mass (Fig. 3D).

Another SNP that was found to be associated with VO2max in women, rs3757354, is located within the 2 KB upstream region of MYLIP. Women homozygote for the rs3757354 common allele (GG) (56%) had a 3 ml/kg0.75/min higher VO2max than the group carrying the heterozygote genotype (AG) (37%) or the minor allele homozygotes (AA) (7%) (Fig. 4A). To determine if rs3757354 could interfere with transcription factor binding, we performed in silico analysis to discover possible transcription factor binding sites. The analysis predicted that having the A allele at rs3757354 creates a perfect binding site for the estrogen receptor alpha (ER-α) targeting the sequence TGACC, whereas having the G allele at rs3757354 is likely to disable the binding of ER-α, potentially reducing estrogen-induced expression of MYLIP (Fig. 4B). Using the GTEx database, we found that rs3757354 was associated with

---

**Fig. 1.** The distribution of SNPs favoring high VO2max in participants from the validation cohort. A; Mean VO2max among participants in the validation cohort grouped according to number of favorable VO2max-SNPs. B; Stratified analysis of inactive and active participants in the exploration cohort (based on a Kurtze-score, lower or higher than the median of the cohort). Group differences are calculated using one-way ANOVA with post hoc tests. Data are shown as mean and SEM. The number of individuals in each category is displayed inside the columns. SNPs: Single nucleotide polymorphisms, VO2max: Maximal oxygen uptake, SEM: Standard error of mean.

**Fig. 2.** The number of favorable SNPs in each participant associated with cardiovascular risk factors in the validation cohort. Data are shown as mean and SEM. SNPs: Single nucleotide polymorphisms, SEM: Standard error of mean, BMI: Body mass index, LDL: Low-density lipoprotein.
differential expression of MYLIP in the adipose tissue, skeletal muscle and the heart \((p < 0.05)\). By using the BXD mice population, we found significant negative correlations between cardiac expression of Mylip and heart mass (Fig. 4C). Participants harboring the high-VO2max genotype (GG) had significantly lower waist, BMI, visceral fat, fat percentage, CRP-levels, as well as significantly higher HDL-cholesterol as compared to the low-VO2max genotypes (AA) and (AG) (Fig. 4D). Furthermore, among those with the low-VO2max genotypes (AA) and (AG) significantly more of the participants were on treatment for hypercholesterolemia (10%) compared to those with the high-VO2max genotype (GG) (3%) (Fig. 4E).

**Discussion**

Here we report the first large-scale screening for genetic variants associated with VO2max. So far, the lack of large studies directly measuring VO2max has limited the possibilities for large genetic association studies for this phenotype. In this present study, we validated 6 new genetic variants associated with VO2max, and replicate associations with 3 SNPs previously associated with fitness-related traits.\(^{10,21,22}\) Based on these nine SNPs we propose a genetic score reflecting inborn VO2max. The mean difference in VO2max between those with 1 favorable SNP compared to those with 7 favorable SNPs was 10.4 ml/kg/min, which is equal to a difference in 3 metabolic equivalents (METs) \((\approx 3.5 \text{ ml/kg/min})\). In other prospective studies, it has been suggested that a decrease of 1 MET is associated with increased risk of diabetes, hypertension and the metabolic syndrome.\(^{23–25}\) Interestingly, the number of favorable VO2max-SNPs carried correlated negatively with several CVD risk factors, like waist circumference, BMI, visceral fat, fat percentage, total cholesterol and LDL-cholesterol. Furthermore, among the participants with 1–4 favorable SNPs, significantly more participants were on treatment for hypertension, compared to those with 5 or more SNPs \((p < 0.05)\). Furthermore, sedentary participants with 1–4 favorable SNPs had higher fat percentage, more visceral fat, higher total cholesterol and higher LDL-cholesterol compared with those with 5 or more favorable SNPs. This indicated that inborn high VO2max is associated with decreased CVD risk.

Since VO2max is a strong predictor of cardiovascular health,\(^{2,3,24,27}\) SNPs associated with VO2max may provide physiological explanation for the link between VO2max and CVD. In this present study, a significant association was found between VO2max and a missense mutation in the exon of BAHD1 \((rs3803357)\), which involves transcription of different amino acids depending on genotype. According to data from the GTEx database, rs3803357 is associated with differences in MYLIP expression by binding to its promoter and recruiting heterochromatin proteins.\(^{28}\) Interestingly, we have previously shown that Igf2 is one of the most significantly upregulated genes in the left ventricle of rats with inherited high VO2max.\(^{29}\) In addition to the links to cardiac phenotype, we also found trends toward lower fat percentage and total cholesterol levels in participants with high VO2max genotypes of BAHD1. Further studies are needed to explore the links between these genomic loci and VO2max.

Another interesting SNP found to be associated with VO2max was located in an intron of the vasoactive intestinal peptide receptor 2 \((VIPR2)\). VIPR2 encodes a neuropeptide receptor that is expressed in the heart and the coronary arteries.\(^{30}\) In the heart, VIPR2 regulates cardiomyocyte contractility in response to binding of vasoactive intestinal peptide \((VIP)\).\(^{31}\) The release of VIP also increases coronary artery...
vasodilatation. Interestingly, several studies have shown that physical activity induces the release of VIP, hence VIPR2 is likely to be important for cardiovascular adaptations during exercise. Furthermore, in rats and humans with cardiomyopathy, the levels of VIPR2 are reduced both in heart and serum, suggesting also a link between VIPR2 and cardiovascular disease.

One of the validated SNPs (rs3757354) was located in the promoter region of MYLIP, potentially interfering with transcription factor binding sites. SNPs in promoter regions may cause loss of transcription factor binding sites or formation of a novel binding sites, which may influence how the gene is transcribed upon different stimuli. The SNP in the promoter region of MYLIP was significantly associated with VO2max in women in the validation cohort, but not in men, indicating that this genetic variation may affect VO2max differently in men and women. Furthermore, a previous transcriptome characterization of estrogen-treated human myocardium identified MYLIP as a sex-specific element influencing contractile function, more specifically showing a negative correlation between cardiac expression of Mylip and contractile function. In line with our data, several other studies have reported gender-specific associations with MYLIP genotypes. For instance, Yan et al. report that G allele-carrying women from the Bai Ku Yao population had higher levels of HDL-cholesterol than the non-carriers. Furthermore, G allele-carrying women from the Han population had decreased levels of total cholesterol and apolipoprotein A1 (ApoA1) compared to non-carriers. None of these associations were seen in men. In our study, rs3757354 was also found to be significantly associated with HDL-cholesterol, and several other CVD risk factors like waist circumference, BMI, visceral fat, fat percentage and high-sensitivity CRP levels. Furthermore, G-allele homozygotes were less prone to hypercholesterolemia. Studies in mice show that increased levels of total cholesterol and apolipoprotein A1 (ApoA1) in male mice significantly reduced the level of LDLR and increased the level of LXR agonists, MYLIP degrades the LDLR, apolipoprotein E receptor 2 (ApoER2) and the very low-density lipoprotein receptor (VLDLR) thereby raising circulating LDL-cholesterol. Furthermore, cells lacking Mylip exhibit markedly elevated levels of LDLR and increased rates of

Fig. 4. MYLIP upstream variant at rs3757354. A: Distribution of VO2max among women according to rs3803357 genotype in the Generation 100 cohort (n = 390) displayed as mean and SE, without adjustments for age and physical activity level (p < 0.002). B: DNA sequence showing the A allele at rs3757354, which creates a perfect binding site for ER-α. C: Correlations between cardiac expression of Mylip in BXD mice strains and heart mass. D: Waist circumference (cm), BMI, visceral fat (cm²), fat percentage, HDL-cholesterol (mmol/l) and high-sensitivity CRP (mg/l) among women according to rs3757354 in the Generation 100 cohort (n = 390) displayed as mean and SE. E: Percentage of women on cholesterol-lowering drugs according to rs3757354 in the Generation 100 cohort (n = 390) displayed as mean and SE. VO2max: Maximal oxygen uptake, DNA: Deoxyribonucleic acid, BMI: Body mass index, HDL: High-density lipoprotein, CRP: C-reactive protein, ER-α: Estrogen receptor alpha., SE: Standard error.
LDL-uptake. Overall, the literature provides compelling evidence suggesting important physiological consequences of MYLIP genetic variation. Based on the gender-specific associations of rs3757354, and the previous reported associations with longevity and CVD, these findings may shed new light on the gender-differences in CVD and the influence of sex-specific hormones. Furthermore, the location of rs3757354 in a potential transcription factor binding site that is under the control of estrogen encourages this hypothesis.

The potassium voltage-gated channel subfamily Q, member 1 (KCNQ1) is a well-characterized gene involved in potassium handling in cardiomyocytes. The rs20744238 located in an intron of KCNQ1 was associated with VO2max both in the exploration and validation cohort. In previous meta-analysis, the minor allele T of this particular SNP has been associated with a shortening of the QT interval, a measure of myocardial repolarization time. Prolongation of the QT interval duration, is a risk factor for drug-induced arrhythmias and sudden cardiac death. Other studies have also reported that this particular SNP influence QT interval in healthy Europeans. In our study, participants harboring the genotype previously associated with prolonged QT interval, had a significantly higher VO2max, compared to the other genotypes. This may shed new information on the U-shaped association between risk of arrhythmias and VO2max. However, only mechanistic studies will be able to identify the true functional consequences of rs20744238.

Due to large differences in human physiology between men and women, and that gender is a major determinant of VO2max, it is likely that some genetic variants have stronger effects in one gender compared to the other. In our study, the lack of similar dependency among men and women for some of the reported SNPs indicates that they may influence VO2max, in a gender-specific manner. A previous study suggests that androgenic hormones are likely to make a significant contribution to VO2max in men, hence, the relative effect of the VO2max-related SNPs may be lower in men than in women. Since our approach only covers a part of the genome, we do not have sufficient evidence to fully evaluate the genetic contribution to VO2max in men compared to women. As DNA-sequencing technology becomes more accessible, future studies will hopefully be able to explain gender differences with greater confidence.

Limitations

There are some limitations related to this study not discussed previously. First, the age distribution of the validation cohort is different from the exploration cohort, hence, we may fail to validate some of the SNPs for the exploration cohort due to their importance in different stages of life. Next, as this study only includes individuals with Caucasian decent, the results are not necessarily valid for other ethnicities, and would have to be validated in other cohorts. Furthermore, estimation of physical activity level is an important source of bias, as this parameter is included as a covariate in the genetic association analyses. Nevertheless, as regular physical activity has large influence on VO2max, this was considered a necessary covariate despite the use of self-reported data.

Conclusion

This is the first large genetic association study on directly measured VO2max. We discovered and validated new genetic loci associated with VO2max, and explored their physiological importance using genotype-phenotype databases and in silico tools. We proposed a genetic signature of inborn VO2max, consisting of 9 SNPs that could distinguish high vs. low fitness individuals based on simultaneous carriage of multiple favorable alleles. Interestingly, the number of favorable SNPs correlated negatively with the presence of several CVD risk factors. Future studies combining several large cohorts with directly measured VO2max are needed to identify more SNPs associated with this complex phenotype. Supplementary data to this article can be found online at https://doi.org/10.1016/j.pcad.2020.02.001.

Statement of conflict of interest

None of the authors have any conflicts of interests with regard to this publication.

Acknowledgements

This study was founded by grants from the Kristian Gerhard Jebsen Foundation, the Norwegian Health Association, the Liaison Committee between the Central Norwegian Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU), St. Olavs Hospital and the Faculty of Medicine and Health Sciences at NTNU and Foundation for Cardiovascular Research at St. Olav’s Hospital and NTNU. The chip genotyping was provided by the Genomics Core Facility (GCF), NTNU. GCF is funded by the Faculty of Medicine and Health Sciences at NTNU and Central Norwegian RHA. Genotyping of candidate SNPs was provided by the Centre for Integrative Genetics (CIGENE) at the Norwegian University of Life Sciences. The Nord-Trøndelag Health Study (HUNT) is a collaboration between HUNT Research Center, Faculty of Medicine and Health Sciences, NTNU, Nord-Trøndelag County Council, Central Norway RHA, and the Norwegian Institute of Public Health. Professor Johan Auwerx and researcher Hao Li at the École Polytechnique Fédérale Lausanne, Switzerland contributed with important data and analyses from the BXD database. Testing of VO2max in Generation 100 was provided by NeXt Move core facility at NTNU. NeXt Move is funded by the Faculty of Medicine and Health Sciences, NTNU and Central Norwegian RHA. The results of the study are presented clearly, honestly, and without fabrication, falsification or inappropriate data manipulation.

References

33. Van TT, Yin RX, Li Q, et al. Association of MYLIP rs3757354 SNP and several environmental factors with serum lipid levels in the Guangxi Bai Ku Yao and Han populations. Lipids Health Dis 2012;11:141.