

Breeze 2.0: an interactive web-tool for visual analysis and comparison of drug response data

Swapnil Potdar^{1,*}, Philipp Ianevski^{1,†}, Aleksandr Ianevski^{1,†}, Ziaurrehman Tanoli¹, Krister Wennerberg², Brinton Seashore-Ludlow^{3,4}, Olli Kallioniemi⁴, Päivi Östling⁴, Tero Aittokallio^{1,5,6} and Jani Saarela^{1,*}

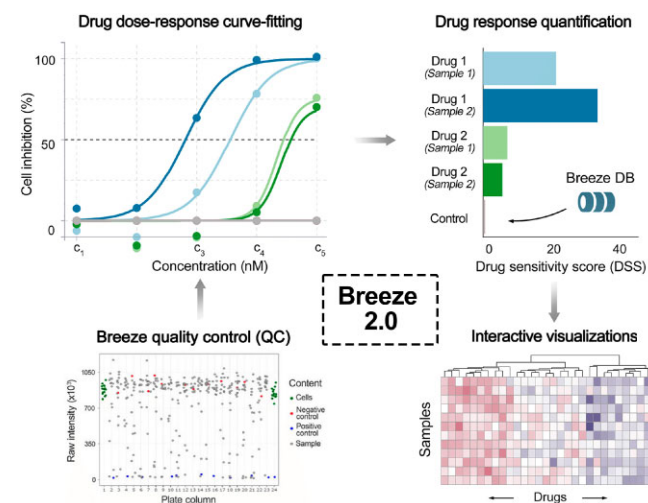
¹Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Finland, ²Biotech Research & Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark, ³Department of Medical Biochemistry and Biophysics, Chemical Biology Consortium Sweden (CBCS), Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden, ⁴Department of Oncology-Pathology, Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden, ⁵Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, Norway and ⁶Centre for Biostatistics and Epidemiology (OCBE), Faculty of Medicine, University of Oslo, Norway

Received March 26, 2023; Revised April 28, 2023; Editorial Decision April 30, 2023; Accepted May 11, 2023

ABSTRACT

Functional precision medicine (fPM) offers an exciting, simplified approach to finding the right applications for existing molecules and enhancing therapeutic potential. Integrative and robust tools ensuring high accuracy and reliability of the results are critical. In response to this need, we previously developed Breeze, a drug screening data analysis pipeline, designed to facilitate quality control, dose-response curve fitting, and data visualization in a user-friendly manner. Here, we describe the latest version of Breeze (release 2.0), which implements an array of advanced data exploration capabilities, providing users with comprehensive post-analysis and interactive visualization options that are essential for minimizing false positive/negative outcomes and ensuring accurate interpretation of drug sensitivity and resistance data. The Breeze 2.0 web-tool also enables integrative analysis and cross-comparison of user-uploaded data with publicly available drug response datasets. The updated version incorporates new drug quantification metrics, supports analysis of both multi-dose and single-dose drug screening data and introduces a redesigned, intuitive user interface. With these enhancements, Breeze 2.0 is anticipated to substantially broaden its potential applications in diverse domains of fPM.

GRAPHICAL ABSTRACT



INTRODUCTION

Development of new drugs and repurposing of existing ones for new indications is a critical and ongoing process with significant potential for enhancing future disease management strategies (1–9). With the escalating prevalence of various diseases and the growing demand for innovative drug development methods, high-throughput screening (HTS) has emerged as a systematic approach for identifying potential hits for drug discovery by profiling thousands of chemical compounds. (10). However, interpreting and analyzing the vast amounts of drug response data generated from the HTS experiments is a complex task,

*To whom correspondence should be addressed. Tel +35 8503175485; Email: jani.saarela@helsinki.fi

Correspondence may also be addressed Swapnil Potdar. Email: swapnil.potdar@helsinki.fi

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

requiring specialized expertise in statistical analysis and programming. Furthermore, the attainment of accurate, interpretable, and reproducible results is essential to identify robust and reliable drug candidates (11,12). One of the key applications of high throughput drug screening lies in the realm of drug repurposing/repositioning. Drug repurposing entails the discovery of novel therapeutic uses for existing drugs, providing a highly effective strategy for the development of drug molecules with innovative therapeutic indications. This process involves the examination of a panel of drugs against specific targets, followed by a systematic comparison with diverse datasets. In response to this challenge, we developed Breeze, a web application for interactive quality control, analysis and visualization of drug dose-response data (13).

Breeze streamlines the analysis and visualization of drug responses generated from cell-based drug screening experiments by implementing comprehensive quality control (QC) procedures, robust dose-response curve-fitting, diverse response quantification metrics, and interactive visualizations. Breeze's QC process plays a crucial role in identifying and quantifying potential errors in data generated from HTS assays, which are prone to common technical issues such as spatial plate variability, striping, and edge effects. Breeze provides a comprehensive set of QC metrics and visualizations, enabling researchers to monitor and identify technical problems, ensuring the accuracy and reproducibility of the screening results. The next critical step involves dose-response curve fitting, which utilizes mathematical modeling to describe the relationship between drug concentrations and the observed responses, such as cell viability or toxicity. Finally, the fitted drug responses are quantified and summarized into single metrics such as half-maximal inhibitory concentration (IC50), half-maximal effective concentration (EC50), area under the curve (AUC), or drug sensitivity score (DSS) to enable comparison across different compounds and concentrations, identifying clinically relevant dose ranges and the most potent and efficient compounds for a given target, patient or disease.

However, Breeze 1.0 did not allow researchers to integrate and compare analyzed datasets with publicly available drug response data, which is crucial for establishing reliable reference baselines for response comparison, validating results, and gaining insights into the broader implications of findings (14). Furthermore, the absence of user-friendly features and automated procedures in Breeze 1.0 posed some challenges for researchers with limited computational expertise. To address these limitations, we have implemented the Breeze 2.0 web-application, which introduces a curated database for easy data integration and comparison, novel interactive visualizations, new drug response metrics, and a redesigned, intuitive user interface. Breeze 2.0 supports analysis of both multi-dose and single-dose drug screening experiments and it utilizes machine learning to flag poor-quality dose-response curves. We believe that the updated web platform will become an even more useful tool, allowing comprehensible and interpretable analysis of drug response data, thus expediting the identification of novel treatment options for various diseases.

Table 1. Comparison of key features between Breeze 1.0 and Breeze 2.0

Feature	Release 1.0	Release 2.0
Comparison with publicly available datasets	No	Yes
Quality control	Plate-based statistics	Plate-based statistics, curve-fitting outlier detection
Minimum number of measured drug doses	4 doses	1 dose
Visualization options	Interactive QC plots, summary barplots and heatmaps	Interactive QC plots, summary barplots, heatmaps, volcano plots, interactive curve fits
Antiviral data analysis	No	Yes

MATERIALS AND METHODS

Overview of the workflow

Breeze 2.0 introduces a number of novel features and improvements for interactive analysis and visualization of drug response data; these include: (i) a curated database of published drug screening data that facilitates easy integration and cross-comparison of user provided data with the publicly available datasets, including standardized comparison with healthy controls or other reference datasets; (ii) novel interactive visualization options for integrative analysis of user-provided and published data; (iii) implementation of new response metrics for antiviral data analysis; (iv) implementation of a machine learning-based approach for automated identification of poor-quality dose-response curves and (v) analysis of both multi-dose and single-dose drug screening data. Additionally, Breeze 2.0 introduces a re-designed user interface that is more intuitive and user-friendly. Table 1 provides a detailed comparison of the features between Breeze releases 1.0 and 2.0. The users of the Breeze web-application provided valuable input, beta testing and suggestions for improvements which were implemented into Breeze 2.0.

Data processing pipeline

The Breeze 2.0 pipeline starts with processing of the raw data to generate a comprehensive QC report, featuring plate-specific heatmaps, scatterplots, control barplots, and an in-depth summary of QC statistics emphasizing control well performance. For each drug-dose data point, percent inhibition/viability is determined with reference to the plate controls (Figure 1A). Subsequently, dose-response curve fitting is carried out using four-parameter logistic modeling of percent inhibition values as a function of drug concentration (Figure 1B, left panel).

The curve fitting quality can be visually evaluated using the dose-response curve fitting plots and by analyzing the fitting errors. To ensure improved accuracy, we also employed a machine learning-based model (see the Implementation section) that automatically detects curve fitting errors and flags them in the summary curve fit table. Finally, based on the curve fitting parameters, various drug

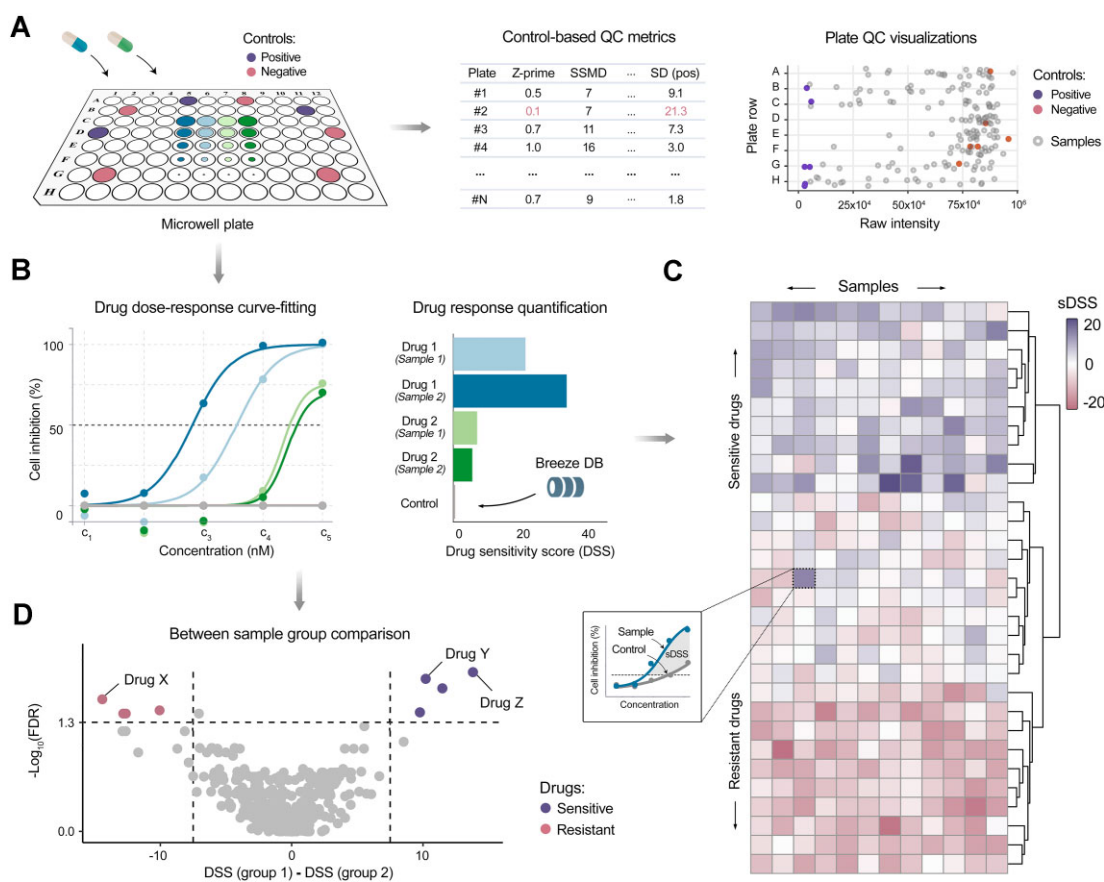


Figure 1. General workflow of Breeze 2.0 web-application. (A) Breeze data analysis starts with a quality control (QC) procedure that includes multiple plate control-based QC metrics, such as Z-prime and SSMD (middle panel). Additionally, Breeze features a range of plate-specific visualizations, e.g. scatterplots (right panel), facilitating the detection of anomalies and experimental errors not identified through numerical analysis alone. In this example, Plate #2 demonstrates a poor quality, as evidenced by its low Z-prime score and high standard deviations of positive controls (red highlights in the middle panel). (B) Drug dose-response curve fitting is the first step in quantifying drug responses into single metrics (left panel). Subsequently, Breeze 2.0 supports the calculation of various drug performance metrics, including IC₅₀, EC₅₀, AUC and DSS, allowing for direct comparisons between compounds and relative metrics, such as sDSS and SI index that allow for comparison between samples and controls. The barplot illustrates an example where the DSS score was used as the quantification metric (12). Additionally, Breeze offers the possibility to cross-compare user-provided data with previously reported drug responses incorporated into the Breeze database, serving as reference controls for comparison (right panel). (C) Next, an interactive heatmap is generated to compare drug responses across different samples (e.g. cell lines or experimental conditions), with sDSS scores shown as an example to highlight the selective efficacy of the drugs, while other metrics can also be used in the heatmap. (D) As an alternative to heatmap, statistically significant differences in drug responses between two groups of samples can be identified using a volcano plot.

quantification metrics, such as IC₅₀, EC₅₀, DSS, AUC, are calculated and reported in the summary table (Figure 1B, right panel). In addition, Breeze 2.0 allows calculation of relative metrics, such as selective DSS (sDSS), selective AUC (sAUC) and selectivity index (SI), that enable comparison between samples and controls, and help in the joint assessment of drug efficacy and toxicity. For example, in antiviral drug screening, SI is calculated by dividing a drug's cytotoxicity (its ability to kill cells) by its antiviral activity (its ability to inhibit viral replication), resulting in a ratio that reflects the drug's selectivity for viral targets over host cells. These metrics facilitate the identification of clinically relevant dose ranges and the most potent, effective and selective compounds for specific targets, patients or diseases.

Breeze 2.0 database

Breeze 2.0 introduces a curated database that allows researchers to integrate and compare their own datasets

with publicly available data, which is essential for establishing solid comparison baselines, validating results, and for exploring the wider implications of their findings. The drug responses from the database can serve as a reference control for comparison of user-provided drug responses in the same condition/disease, control groups, or cell lines, identifying disease-specific drug responses and uncovering potent targeted therapies. The database used in the Breeze 2.0 includes data from Malani *et al.* (15), which includes drug sensitivity data from 186 AML patient samples and 17 healthy controls. In addition, we incorporated the PharmacDB, the most comprehensive database that consolidates pharmacogenomics cell line data from multiple sources such as CCLE, GDSC, NCI-60, CTRP and others (14). In the future, we aim to expand the Breeze database by incorporating additional curated and published datasets, to improve coverage of drug response patterns across a wider range of tissues and drug classes.

Visualizations

The results of Breeze 2.0 pipeline are visualized in the form of multiple interactive plots such as heatmaps (Figure 1C), volcano plots (Figure 1D), barplots, scatterplots, and circular trees, allowing easy investigation of the results. The details on how to obtain and interpret each visualization plot are explained in the Breeze technical documentation: https://breeze.fimm.fi/DSRT_documentation/docs.html. Within the Breeze interface, users can access the 'Curve Fitting' tab and select one or more dose-response curves from a dropdown menu. The software also allows users to incorporate reference data from the database, including information on healthy controls, enabling integrative analysis of drug response data (see e.g. Figure 1B). The resulting plots can be exported as PDF, PNG and HTML files, while a summary table displays drug quantification metrics for selected drugs and screens, which can be downloaded as a spreadsheet.

Implementation

The Breeze 2.0 web-server is powered by PHP and MySQL for database support. The data processing pipeline utilizes the R programming language and a variety of R packages, while the interactive visualizations are created using GG-Plot, Plotly and D3.JS in JavaScript. To ensure the accuracy of curve fitting, an Adaboost machine learning classifier has been trained using Breeze's extensive in-house data set of over 10 000 expert-curated and classified 5-point dose-response curves to flag poor-quality curve-fits in the user data. The final model employs conformal prediction with a 0.8 confidence threshold, which allows users to exclude low-confidence model predictions, thereby flagging only the most confident low-quality curve fits. Currently, the model is capable to flag dose-response curves with a minimum of four doses, up to five drug doses.

RESULTS AND DISCUSSION

Breeze 2.0 represents a major upgrade to the existing drug screening data analysis pipeline presented in the first version of the tool. Designed to be accessible by researchers with no programming skills, Breeze 2.0 requires only the raw screening data as an input (either raw responses or percentage inhibition/viability), and it automatically analyzes, quantifies and visualizes the drug response data, thereby significantly reducing the manual time required for the analysis of large-scale drug screening experiments.

By incorporating novel data exploration capabilities, users can cross-compare their data in the context of published drug response datasets, enabling the identification of sample-specific, selective drug responses, and avoiding the prioritization of false positive hits, such as generally toxic drugs, when comparing against responses observed in healthy controls. Additionally, Breeze 2.0 enables users to better understand the dose-response relationship with other drugs targeting the same target pathways.

The Breeze 2.0 pipeline's flexible input format and data upload functionality makes it suitable for a wide range of readouts, including numeric data obtained from microscopic images, RNAi experiments, and other sources. This

makes Breeze 2.0 a versatile tool that can be used across a variety of different research applications, extending its utility beyond traditional drug development workflows. To the best of our knowledge, there are no similar and equally comprehensive web-platforms for drug response data analysis.

Moreover, Breeze 2.0 features a completely redesigned user interface that is more intuitive and user-friendly. It also includes novel interactive visualization options, which were requested by the users, and essential for avoiding false positive/negative findings. Moving forward, we plan to further expand Breeze by integrating existing and emerging large-scale drug screening resources, with an emphasis on healthy controls from diverse tissues.

DATA AVAILABILITY

Breeze 2.0 is freely accessible at <https://breeze.fimm.fi/v2/> or <https://breezetestool.app> without any login requirements. Extensive documentation of all the features is available at https://breeze.fimm.fi/DSRT_documentation/docs.html.

The Breeze 2.0 database includes published datasets, with links to the original sources provided below. The PharmacDB datasets were integrated using the PharmacGx 2.6.0 R/Bioconductor package (<https://bioconductor.org/packages/release/bioc/html/PharmacGx.html>), and were downloaded using the downloadPSet function. The Malani *et al.* dataset can be accessed at <https://zenodo.org/record/7274740>.

ACKNOWLEDGEMENTS

We would like to thank all the users of Breeze web-application for valuable input, beta testing and suggestions for improvements.

FUNDING

FIMM HTB unit is supported by the University of Helsinki and Biocenter Finland; T.A., B.S.L. and P.Ö. were supported by the European Union's Horizon Europe Research & Innovation Programme [REMEDI4ALL project, 101057442]; T.A. by European Union's Horizon 2020 Research and Innovation Programme (ERA PerMed JAKSTAT-TARGET and CLL-CLUE projects), and Academy of Finland [326238, 340141, 344698, 345803]; P.Ö., B.S.L. were supported by Knut and Alice Wallenberg Foundation [KAW 2020.0241, V-2020-0699, KAW 2020-0182]; O.K. was supported by (2015.0291). The Finnish Unix User Group Foundation (FUUG.fi) (to S.P., F.I., A.I., J.S.). Funding for open access charge: Finnish Unix User Group (FUUG.fi).

Conflict of interest statement. None declared.

REFERENCES

1. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C. *et al.* (2019) Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Discov.*, **18**, 41–58.
2. Tanoli, Z., Alam, Z., Ianevski, A., Wennerberg, K., Vaha-Koskela, M. and Aittokallio, T. (2020) Interactive visual analysis of drug-target interaction networks using Drug Target Profiler, with applications to

- precision medicine and drug repurposing. *Briefings Bioinf.*, **21**, 211–220.
3. Pemovska, T., Kontro, M., Yadav, B., Edgren, H., Eldfors, S., Szwajda, A., Almusa, H., Bespalov, M.M., Ellonen, P., Elonen, E. *et al.* (2013) Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer Discov.*, **3**, 1416–1429.
 4. Pemovska, T., Johnson, E., Kontro, M., Repasky, G.A., Chen, J., Wells, P., Cronin, C.N., McTigue, M., Kallioniemi, O., Porkka, K. *et al.* (2015) Axitinib effectively inhibits BCR-ABL1(T315I) with a distinct binding conformation. *Nature*, **519**, 102–105.
 5. Malani, D., Yadav, B., Kumar, A., Potdar, S., Kontro, M., Kankainen, M., Javarappa, K.K., Porkka, K., Wolf, M., Aittokallio, T. *et al.* (2020) KIT pathway upregulation predicts dasatinib efficacy in acute myeloid leukemia. *Leukemia*, **34**, 2780–2784.
 6. Talwelkar, S.S., Mäyränpää, M.I., Soraas, L., Potdar, S., Bao, J., Hemmes, A., Linnavirta, N., Lømo, J., Räsänen, J., Knuutila, A. *et al.* (2021) Functional diagnostics using fresh uncultured lung tumor cells to guide personalized treatments. *Cell Rep. Med.*, **2**, 100373.
 7. Tuomainen, K., Hyytiäinen, A., Al-Samadi, A., Ianevski, P., Ianevski, A., Potdar, S., Turunen, L., Saarela, J., Kuznetsov, S., Wahbi, W. *et al.* (2021) High-throughput compound screening identifies navitoclax combined with irradiation as a candidate therapy for HPV-negative head and neck squamous cell carcinoma. *Sci. Rep.*, **11**, 14755.
 8. Skaga, E., Kuleskiy, E., Brynjulvsen, M., Sandberg, C.J., Potdar, S., Langmoen, I.A., Laakso, A., Gaál-Paavola, E., Perola, M., Wennerberg, K. *et al.* (2019) Feasibility study of using high-throughput drug sensitivity testing to target recurrent glioblastoma stem cells for individualized treatment. *Clin. Transl. Med.*, **8**, 33.
 9. Brodin, B.A., Wennerberg, K., Lidbrink, E., Brosjö, O., Potdar, S., Wilson, J.N., Ma, L., Moens, L.N., Hesla, A., Porovic, E. *et al.* (2019) Drug sensitivity testing on patient-derived sarcoma cells predicts patient response to treatment and identifies c-Sarc inhibitors as active drugs for translocation sarcomas. *Br. J. Cancer*, **120**, 435–443.
 10. Moffat, J.G., Rudolph, J. and Bailey, D. (2014) Phenotypic screening in cancer drug discovery - past, present and future. *Nat. Rev. Drug Discov.*, **13**, 588–602.
 11. Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G. and Pangalos, M.N. (2014) Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nat. Rev. Drug Discov.*, **13**, 419–431.
 12. Yadav, B., Pemovska, T., Szwajda, A., Kuleskiy, E., Kontro, M., Karjalainen, R., Majumder, M.M., Malani, D., Murumägi, A., Knowles, J. *et al.* (2014) Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci. Rep.*, **4**, 5193.
 13. Potdar, S., Ianevski, A., Mpindi, J.P., Bychkov, D., Fiere, C., Ianevski, P., Yadav, B., Wennerberg, K., Aittokallio, T., Kallioniemi, O. *et al.* (2020) Breeze: an integrated quality control and data analysis application for high-throughput drug screening. *Bioinformatics*, **36**, 3602–3604.
 14. Smirnov, P., Kofia, V., Maru, A., Freeman, M., Ho, C., El-Hachem, N., Adam, G.A., Ba-Alawi, W., Safikhani, Z. and Haibe-Kains, B. (2018) PharmacDB: an integrative database for mining in vitro anticancer drug screening studies. *Nucleic Acids Res.*, **46**, D994–D1002.
 15. Malani, D., Kumar, A., Brück, O., Kontro, M., Yadav, B., Hellesøy, M., Kuusanmäki, H., Dufva, O., Kankainen, M., Eldfors, S. *et al.* (2022) Implementing a functional precision medicine tumor board for acute myeloid leukemia. *Cancer Discov.*, **12**, 388–401.