Nucleic Acids Research, 2026, **54**, D470–D480
https://doi.org/10.1093/nar/gkaf1265
Advance access publication date: 17 November 2025
Database issue

Mapping drug mechanisms with ProteomicsDB: unified omics and cell sensitivity data at scale

Mario Picciani[1,†], Armin Soleymaniniya[1,†], Julian Müller[2,†], Amirhossein Sakhteman[2], Konstantinos Tzanakis[1], Judith Bernett[3], Elias Kahl[1], Firas Hamood[2], Johannes Kersting[3], Mohsen Pourjam[3,4], Luis Augusto Eijy Nagai[3,5], Markus List[3,6], Bernhard Kuster[2,6], Matthew The[2], Mathias Wilhelm[1,6,*]

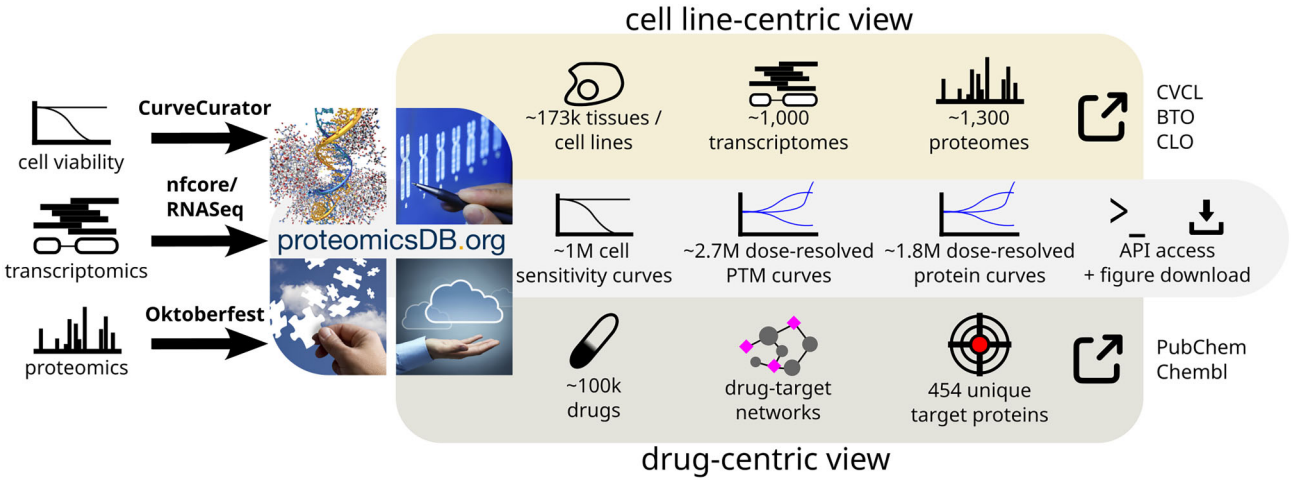
[1]Computational Mass Spectrometry, Technical University of Munich, 85354 Freising, Bavaria, Germany
[2]Proteomics and Bioanalytics, Technical University of Munich, 85354 Freising, Bavaria, Germany
[3]Data Science in Systems Biology, Technical University of Munich, 85354 Freising, Bavaria, Germany
[4]Institute of Artificial Intelligence in Medicine and Healthcare, Technical University of Munich, 81675 Munich, Bavaria, Germany
[5]Laboratory of Computational Genomics, Institute for Quantitative Biosciences, The University of Tokyo, 113-0032 Bunkyo, Japan
[6]Munich Data Science Institute, Technical University of Munich, 85748 Garching, Bavaria, Germany

*To whom correspondence should be addressed. Email: mathias.wilhelm@tum.de
†The first three authors should be regarded as Joint First Authors.

Abstract

Proteomic and phenotypic cell sensitivity datasets are increasingly important for understanding chemoproteomics and the underlying drug mechanisms of action. Yet, integrating such heterogeneous datasets remains challenging due to inconsistent annotations, incompatible IDs, and variable data processing methods. Here, a major update to ProteomicsDB (https://www.proteomicsdb.org) is presented that combines over 1300 proteomic and 1000 transcriptomic profiles with phenotypic cell sensitivity data across >1500 human cancer cell lines and 1470 drugs. Harmonizing cell line and drug names and applying a standardized normalization and refitting pipeline for dose–response curves enables consistent, statistically robust analysis across studies. Three new graphical user interfaces support interactive exploration of cell sensitivity data, exploring the protein targets and dose-resolved changes in protein expression in the presence of a drug, and comparing the expression profiles of cell lines. With this update, ProteomicsDB is strengthening its future role as a central hub for proteomics and multi-omics, providing researchers with a unified framework to explore phenotypic cell sensitivity in combination with dose-resolved expression proteomics at the molecular level, supporting biomarker discovery, drug repurposing, and precision medicine applications.

Graphical abstract



Received: September 15, 2025. **Revised:** October 23, 2025. **Accepted:** October 24, 2025

Introduction

ProteomicsDB (https://www.proteomicsdb.org) is a public multi-organism database primarily focused on proteomics experiments. Built on an in-memory database powered by SAP HANA [1], it enables real-time exploration, analysis, and visualization of large-scale omics datasets without requiring specialized computational resources or expertise. Since its launch in 2014, ProteomicsDB has evolved from a protein-centric database [2] into an integrative multi-omics platform, now covering transcriptomics, protein turnover, phenomics, and cell sensitivity data using a FAIR-oriented approach [3, 4].

As the field of proteomics increasingly shifts toward integrative and clinically oriented analyses, there is a growing demand to link molecular expression profiles with functional phenotypic outcomes. Several large public resources provide access to cell sensitivity datasets, such as the Cancer Therapeutics Response Portal (CTRP) [5–7] or the Cancer Cell Line Encyclopedia (CCLE) [8], typically providing their data for download, with limited online exploration and visualization options. The Genomics of Drug Sensitivity in Cancer (GDSC) [9–11] project offers visualizations of its data on the project website (https://www.cancerrxgene.org) but does not combine this with data from other programs and lacks comprehensive analytic tools, customization, and download options. Integrative programs, such as DepMap [12], attempt to collect cell sensitivity data from various programs and combine it with omics profiles but rely on inconsistent response measures. For instance, typical cell sensitivity metrics, including the area under the dose–response curve (AUC) and the concentration at half-maximal effect (IC50), can differ substantially across studies due to variations in the measured concentration range or extrapolation errors. DepMap provides customizable visualization options through its DataMiner interface (https://dataminer.depmap.sanger.ac.uk/), providing multiple layers of comparisons across omics layers, tissues, or cancer types. However, the sensitivity visualizations are limited to the IC50 and AUC, while the associated dose–response curves from which these measures were derived cannot be displayed. Visualizations are also still limited to the GDSC project, and harmonization for cell line and drug identifiers or links to external resources is not provided, hampering systematic cross-study comparisons. Finally, the CellMiner Cross-Database (CellMinerCDB) [13] provides an interface (https://discover.nci.nih.gov/cellminercdb/) with customizable visualization tools across multiple public cell line panels, including comparison between cell sensitivity and associated genomic profiles; however, it does not cover proteomics data.

This creates barriers for comparative analyses and exploring phenotypic responses in the context of protein-level changes. In particular, dose-resolved expression proteomics, examining changes in protein intensities [14] and post-translational modifications [15] upon treatment with a drug, captures global cellular responses beyond known drug–target interactions. Such studies provide deeper insights into mechanisms of action, resistance, and adaptive cellular responses to treatment. Beyond this, revealing the proteomic target space of a drug provides valuable insight into shared mechanisms of action, which could complement network medicine approaches such as *Drugst.One* [16], a tool focused on drug repurposing.

ProteomicsDB addresses these challenges by online integration of phenotypic cell sensitivity data with dose-resolved expression proteomics, experimentally validated drug targets, and data harmonization using established public ontology and annotation databases. To the best of our knowledge, it is currently the only resource that allows real-time cross-study examination of perturbation data at both the molecular and phenotypic levels, without requiring local data handling. This makes ProteomicsDB a unique hub for mapping expression profiles with functional outcomes.

In this update, major extensions of ProteomicsDB since the last report in 2022 are presented. These include expanded data content for baseline proteomics and transcriptomics data for human cancer cell lines, integration of dose-resolved expression proteomics with phenotypic cell sensitivity data and drug targets, integration of the *Drugst.One* network viewer for tissue- and drug-centric exploration, and incorporation of external metadata to facilitate cross-study comparisons. Together, these developments extend ProteomicsDB into a platform for biomarker discovery, drug repurposing, and translational research, thereby laying the foundation for future applications in precision medicine.

Results

Newly integrated data and resources

With this update, ProteomicsDB extends its scope toward research centering on human cancer cell lines and drug interactions by incorporating rich metadata for tissues and drugs, along with phenotypic cell sensitivity data and dose-resolved expression proteomics. To ensure comparability, metadata from three public ontologies and annotation databases, describing tissues and cell lines, were imported, namely the Brenda Tissue Ontology (BTO) [17], which provides a hierarchical representation of tissues and cell line descriptions; Cellosaurus [18], which offers a comprehensive resource on cell lines across species covering disease classifications, cancer types, synonyms, and cross-references; and Cell Line Ontology (CLO) [19], all of which were incorporated into ProteomicsDB. In total, this integration provides structured annotations for ~173 000 tissues, cell lines, and fluids across multiple species, complementing existing ProteomicsDB resources and enabling consistent mapping of datasets that make use of different ontologies for sample identification.

On top of these annotations, large-scale baseline expression profiles for over 1300 cell lines were added since the last update (Supplementary Tables S1 and S2). ProteomicsDB makes use of standardized importer scripts, including the established nfcore/rnaseq [20] pipeline for transcriptomics data, as well as Oktoberfest [21] for proteomics data, to ensure consistency in data preprocessing. The newly added profiles include reprocessed RNASeq data for ~1000 human cancer cell lines as part of the CCLE [22] and proteomics data for ~1200 human cancer cell lines from 375 TMT-based [23] and 949 DIA [24] samples (Supplementary Methods). For the latter, the post-translational modification (PSM) viewer [25] was extended to support DIA data (Supplementary Fig. S1). With these imports, the number of unique tissues for which experimental data exists has increased to 2657.

The database was further expanded with ~100 000 drug entries from PubChem [26] and ChEMBL [27], covering small molecules, natural compounds, and antibodies. From those resources, the main drug IDs and common names, as well as SMILES and InChI notations, were imported where available. In addition, metadata, including molecular and biomed-

ical properties, are dynamically loaded through the APIs of ChEMBL and PubChem when querying ProteomicsDB for a specific drug. This approach allows ProteomicsDB to map and integrate drug-related information, enabling cross-examination of public datasets containing experiments for the same compounds. The total number of drug entries in ProteomicsDB is now at 107 281, and for 4014 of those, the platform contains experimental data.

Building on this foundation, ProteomicsDB now provides large-scale data for dose-resolved expression proteomics, including measurements for ∼1.8 million drug–protein [14, 28] and ∼2.7 million drug–PTM combinations [15, 28, 29] for 162 and 187 compounds, respectively. These data are complemented by experimentally validated target proteins spanning inhibitors of ATR [29], KDACs [30, 31], and natural compounds [32], resulting in ∼14 600 drug–protein pairs for 1232 unique drugs across 454 unique target proteins in a dose-resolved fashion (Supplementary Table S3), allowing ProteomicsDB to build networks between the drug and its targets, which can be explored and expanded to other drugs and associated diseases using the *Drugst.One* network viewer. Finally, ProteomicsDB was extended with data from six phenotypic cell sensitivity experiments [5, 6, 8, 9, 11, 33], comprising more than one million dose–response curves for ∼800 unique drugs across ∼2500 cell lines (Supplementary Table S4). To view and analyze these data, the existing cell sensitivity analytics tool was redesigned and expanded in functionality. For phenotypic cell sensitivity and dose-resolved expression proteomics data, the dose–response fitting tool CurveCurator [34] is used, facilitating calculation of statistical measures for quality assessment and filtering data for significantly regulated curves (Supplementary Methods). A full description of imported datasets can be found in Supplementary Tables S1--S4.

Intuitive exploration of phenotypic cell sensitivity data

The cell sensitivity analytics tool (https://www.proteomicsdb.org/analytics/cellSensitivity) has undergone a comprehensive overhaul, introducing an intuitive, interactive visualization that allows users to efficiently explore and analyze phenotypic cell sensitivity data (Fig. 1). The process begins with the selection of a dataset of interest, which dynamically loads the list of available cell lines and drugs. Users can then filter the data down to specific cell lines and drugs, ensuring that only relevant combinations are loaded for exploration. In the following, the application is showcased for all cell lines of the CTRPv2 dataset, examined for sensitivity against the drug afatinib, an inhibitor of the epidermal growth factor receptor (EGFR), a receptor tyrosine kinase and key oncogene in many cancers (Fig. 1A). The user interface makes use of a visual approach that presents the data along various coordinates, each corresponding to a parameter derived from fitting a dose–response curve (Supplementary Methods). These coordinates are fully customizable, allowing users to select or deselect parameters based on their analytical needs. To refine the selection further, users can limit parameter ranges and, for instance, filter out downregulated dose–response curves within a specific pEC50 (−log10 EC50) range when analyzing a particular drug target of interest. In this case, the selection is limited to a pEC50 range of 4–7 (10^{-7} to 10^{-4} M) that is representative of afatinib-sensitive cell lines (Fig. 1B). Based on the statistical measures obtained through CurveCurator, the plot also contains a regulation coordinate, which can be used to filter, e.g., only significantly downregulated curves. In addition, a selection for plotting the filtered data over selected coordinates can be made (Fig. 1B). This allows investigating the relation of two parameters and allows users to select individual cell line–drug combinations. This allows the selection of, e.g., significantly downregulated curves of cell lines that are highly sensitive in a specific concentration range and exhibit a certain effect size (Fig. 1C). Once selected, the corresponding dose–response curves are visualized in a dedicated plot (Fig. 1D), providing an immediate view of the sensitivity profile across the measurement range, including the fitted dose–response curve.

To facilitate fine-grained control over drug, cell line, and curve fit parameters, the data are also provided as a table that allows filtering and selection (Fig. 1E). In line with ProteomicsDB's commitment to providing ready-to-use visualizations, all plots can be immediately downloaded as high-resolution PNG images or as editable SVG vector graphics, ensuring publication-quality output (Fig. 1). In addition, the tabular data are available as either a CSV or an Excel-formatted file. To facilitate programmatic data retrieval, a new OData-compliant API endpoint was created, allowing the retrieval of fitted parameters for a selected dataset and cell lines.

As a result, this newly designed frontend serves as a highly flexible, visually driven exploration tool, allowing researchers to interact with and analyze cell sensitivity data without requiring external data processing or offline visualization tools, making it the perfect tool for quick sensitivity checks and drawing dose–response curves, while connecting directly with drug and cell line information allows further investigation.

Drug-centric view

Currently, ProteomicsDB contains IDs for a total of 107 281 unique drugs, ranging from small molecules, over natural products, to antibodies, 4014 of which are linked to experimental data. To reflect the increasing importance of chemoproteomics data in ProteomicsDB, a drug-centric view was developed that is structured in an analogous way to the protein-centric view. In order to retrieve information about a drug, users have three possibilities to reach the drug-centric view: first, by using the drug name search and selection from one of the matching results; second, through URL access in case the ProteomicsDB drug ID is known; and third, by using direct interlinks ProteomicsDB provides in various views, such as the analytics tool for cell sensitivity in data lists and tables for immediate access.

The summary page is dynamically generated and makes use of the PubChem and ChEMBL APIs to provide additional information about molecular and biochemical features of the drug, including on-the-fly visualization of the compound's structure using SmilesDrawer [35] if a SMILES string is available (Supplementary Fig. S2). A list of tabs connects the drug-centric view to various experimental data that are collected in ProteomicsDB, which are outlined in the following for the example of afatinib.

A compound's protein targets can be experimentally determined by pulldown assays, quantifying each drug-target relationship in a dose-resolved fashion. These data can be accessed in the "Targets" tab (https://www.proteomicsdb.org/drug/46/targets). The target proteins are sorted by the drug's target binding affinity. This highlights a high affinity of afatinib toward EGFR in a concentration range of ∼3 nM (Fig. 2A). In this case, the drug has eight other tar-

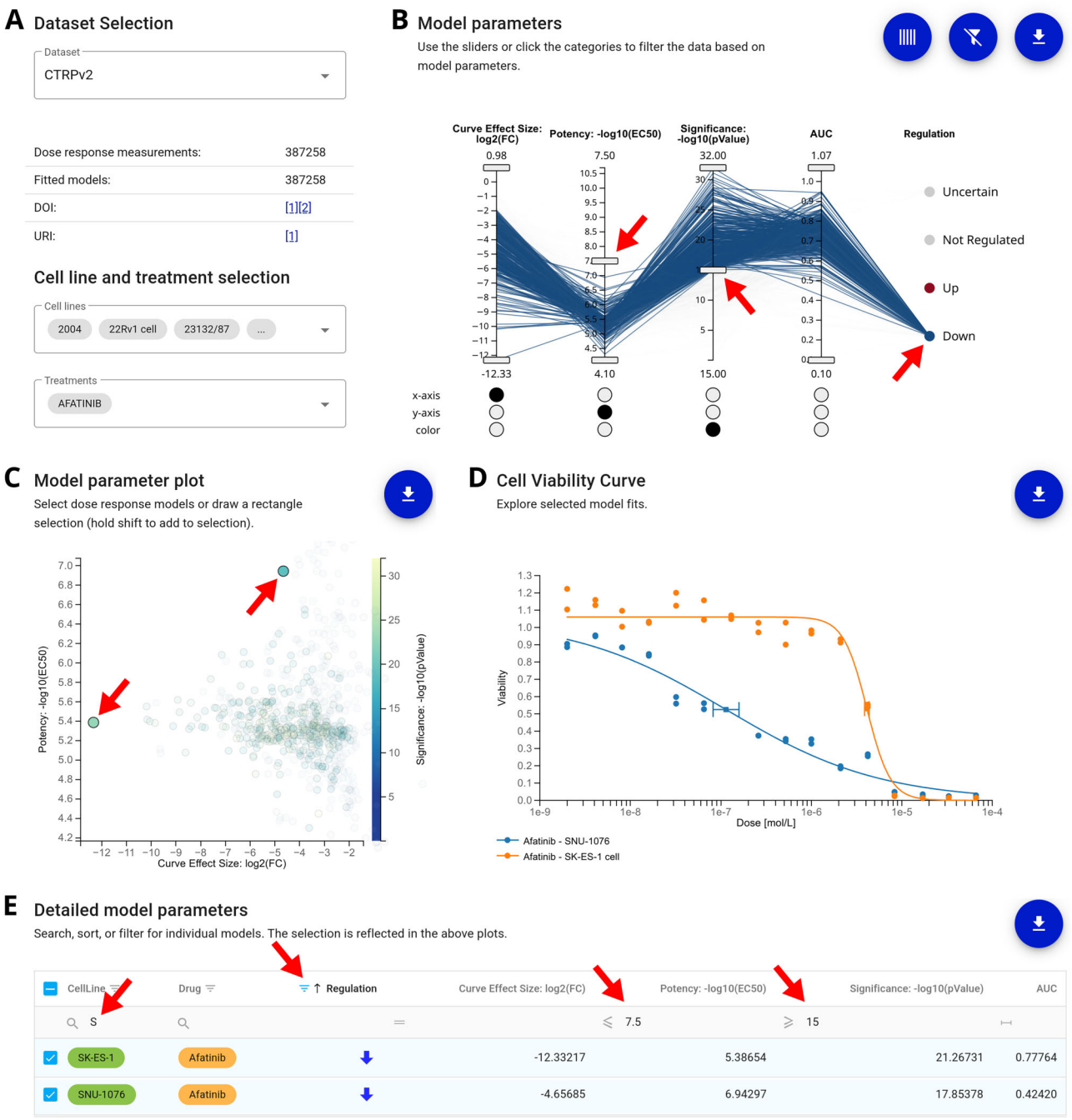


Figure 1. The cell sensitivity view for visual exploration and selection of cell line–drug combinations, showcasing the selection of afatinib-sensitive cell lines. (**A**) From a variety of available datasets and measured cell line–drug combinations, afatinib is selected over all cell lines of the CTRPv2 dataset. (**B**) The parallel coordinate plot is used to filter for highly significant, downregulated curves, with a drug potency that falls into the range of EGFR–target engagement using the draggable sliders of each coordinate (arrows). The numbers at the end of each coordinate indicate the current upper and lower boundaries of the selection. The curve fold change (effect size) and pEC50 are used to separate individual cell line–drug combinations into a two-dimensional space, colored by significance using the buttons below the respective coordinates. Three buttons allow selection of individual coordinates to be visualized, resetting the filters, and downloading the plot. (**C**) Two highly significant, downregulated curves within the expected pEC50 range are selected for plotting the fitted dose–response curves (arrows highlight the selection). The button downloads the plot. (**D**) The curves of the selected cell line–drug combinations are plotted in individual colors, including the normalized response measures for each replicate and the fitted dose–response curve. The button downloads the plot. (**E**) The selected dataset is shown in a table, representing each selected coordinate as a column. Filters on top allow precise data selection and connect with the sliders in the parallel coordinate plot. Selection with tick boxes provides an alternative to control which cell line–drug combinations are selected for curve plotting. Upon change, all plots are updated automatically (arrows highlight active filters). The cell line and drug names act as links to the cell line-centric and drug-centric views of the displayed cell line and drug, respectively. The button downloads the data as a CSV- or Excel-formatted file.

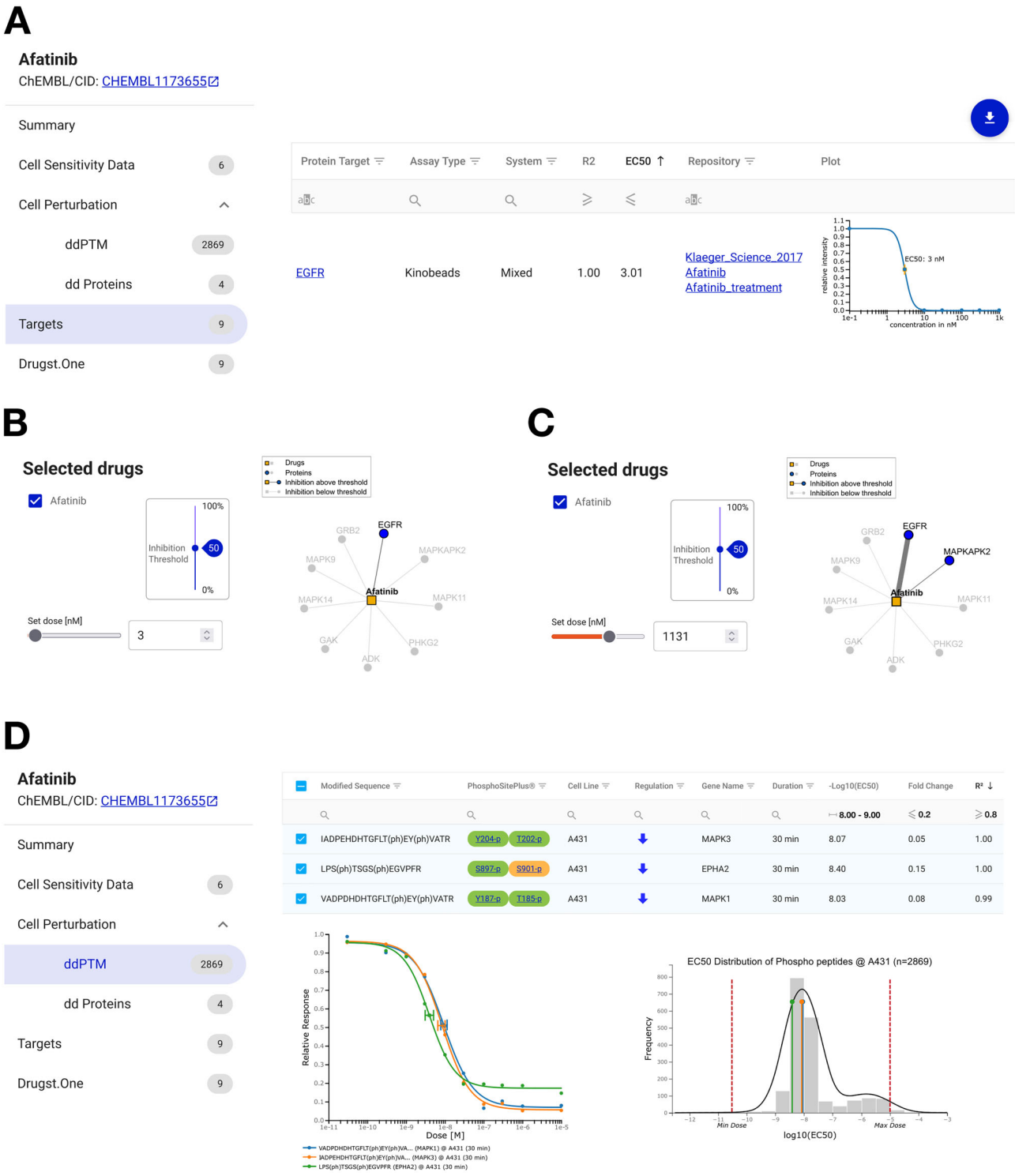


Figure 2. Drug-centric analysis in ProteomicsDB.(**A**) "Targets" tab of afatinib showing the data for its most potently targeted protein, EGFR. (**B**) The protein–drug interaction analytics tool is used to visualize target engagement using the horizontal (dose) and vertical (threshold) sliders: The graph on the right shows only those interactions in color that fulfill the current slider settings. At a dose of 3 nM afatinib, the only protein showing 50% effect is EGFR. (**C**) At a higher dose (1131 nM), a second protein, MAPKAPK2, shows 50% response. The dose can directly be set to the EC50 of a protein–drug interaction by clicking on the respective edge in the graph. (**D**) The ddPTM tab of afatinib's drug-centric page provides a table of PTMs. It is filtered for curves with a strong negative fold change (≤0.2) and an EC50 within the range of EGFR engagement (1–10 nM) and sorted by decreasing quality of curve fit (R^2), showcasing the response of three phosphorylated peptides over the measured dose range (left plot) and in the context of the overall distribution of EC50s, which shows an overall high frequency of responses within the EGFR engagement range (right plot).

gets (as indicated by the number next to the tab), which are hit at higher concentrations. ProteomicsDB also contains an analytics tool at https://www.proteomicsdb.org/analytics/proteinDrugInteraction that allows dynamically setting the concentration cutoff for target identification using a slider to see which targets are hit at which concentration. This can be used to determine a drug's affinity toward each of its targets. Afatinib is highly potent for EGFR, as it is the only target that is hit at a concentration of 3 nM (Fig. 2B), while the next target, MAPKAPK2, requires >300 times higher concentration to show the same response (Fig. 2C).

In order to study the cellular mode of action, the "Cell Perturbation" tabs for the protein and PTM level (https://www.proteomicsdb.org/drug/46/drugddptm) can be used. Since EGFR is a kinase, which exerts its effect by means of phosphorylation, the drug effect is primarily visible on the phosphoproteome (Fig. 2D). To study the downstream effects of EGFR inhibition, the fitted dose–response curves can be filtered for a strong negative fold change and an EC50 value similar to the EC50 of target engagement between afatinib and EGFR from the "Targets" tab. The data table in this tab can be manipulated with custom sorting and filters for each column. When sorting by the quality of fit (R^2), the top curves include well-studied indirect effector sites of EGFR (MAPK1-T185&Y187, MAPK3-T202&Y204, and EPHA2-S897&S901) [36, 37]. Each site and each protein is linked to the PhosphoSitePlus [38] and Uniprot [39] databases for further investigations.

The "Cell Sensitivity" tab (https://www.proteomicsdb.org/drug/46/cellSensitivity) allows combining the changes on the proteomic level with phenotypic cell sensitivity data. This tab is very similar to the standalone cell sensitivity analytics tool (Fig. 1), with the exception that the drug is preselected and datasets are filtered to include only the ones for which the drug was investigated, as indicated by the number next to the tab.

Lastly, the drug-centric view provides the "*Drugst.One*" tab (https://www.proteomicsdb.org/drug/46/drugstOne). *Drugst.One* generates interaction networks between proteins, drugs, and diseases. For example, ProteomicsDB provides protein targets as UniProt IDs, which *Drugst.One* uses to build a network linking the drug to its known targets. The tool enriches this network with protein–protein interactions from STRING [40], protein-disease and drug-disease associations from NeDRex [41], protein annotations from UniProt, and drug annotations with target information from DrugBank [42]. Loaded data include direct links to the corresponding external resources and can be accessed by clicking on nodes or edges in the network. Functionally related proteins are connected, and the network can be expanded with additional drugs or diseases (Supplementary Fig. S3).

Thus, the new drug-centric view serves as a central entry point for investigating drug actions within ProteomicsDB. It integrates a diverse panel of experiments, thereby enabling researchers to explore the selectivity and mechanisms of action of drugs as well as cell sensitivity. The addition of a network-based approach holds potential for drug repurposing studies.

Cell line-centric view

ProteomicsDB has evolved from a protein-centric resource to an integrated platform linking phenotypic cell sensitivity with dose-resolved expression proteomics and targets, enabling a sample-first perspective. Therefore, a new tissue/cell-line–centric view was introduced that unifies cell line/tissue baseline proteome/transcriptome profiles with dose-resolved proteomic responses and cell-sensitivity readouts. This organization is crucial for translational use, as cohorts are typically defined by tissue of origin [43], and it provides a natural locus for biomarker discovery and cohort stratification. To harmonize identifiers, tissues and cell lines are mapped to BTO, Cellosaurus, and CLO, yielding ∼173 000 standardized entries (tissues, cell lines, fluids) that are searchable and cross-resolvable by external IDs. This backbone enables reliable cross-study comparisons and seamless pivots between baseline expression, perturbation effects, and drug–target information within the same sample context.

To access the cell line-centric view, users can make use of the new cell line search to find cell lines by name that are matched against all the information that was imported from the external resources, which almost always guarantees to provide a match in case of various spellings. Alternatively, ProteomicsDB provides direct URL access using alternative accessions at https://proteomicsdb.org/cell-line/≤resource≥_≤ID≥, where < resource > is one of PRDBTISSUE, BTO, CVCL, or CLO, and < ID > is the resource-specific accession ID. This provides a flexible entry point that can be used for cross-referencing ProteomicsDB without having to know the native observable PRDBTISSUE accession. Finally, cell line references in other views, e.g. the cell sensitivity analytics tool, interlink to the cell line-centric view for quick access.

Following the idea of the protein-centric and drug-centric views, a summary page (Fig. 3A) is first presented when entering the cell line-centric view, which consolidates general information about the cell line (or tissue or fluid), mapped metadata with outgoing links to the source databases, and two tables listing available baseline expression and cell sensitivity experiments, respectively. These provide a quick overview of the available data, including direct links to the project-centric views. In case hierarchical information through the BTO is available, the cell line is also shown in a tree structure, providing context information about its tissue of origin in the organism. Specific experimental data for the cell line can be accessed through a list of tabs, which are presented in the following (Fig. 3A).

The "Expression" tab summarizes baseline proteome or transcriptome measurements for a selected tissue or cell line by aggregating available samples with ProteomicsDB's standardized expression calculation pipeline. It provides a holistic view of the expression landscape and readily places any gene product's abundance in the context of the entire cell line or tissue (Fig. 3B). The "Comparative analysis" tab enables pairwise comparisons between two samples, offering a succinct assessment of overall similarity while highlighting proteins with divergent abundance (Fig. 4). Together, these views support evaluating, e.g., how representative a cell line is of its tissue of origin, prioritizing candidate biomarkers, and detecting potential batch or other systematic effects. Users can further drill down to the protein level for immediate follow-up and export the underlying data for downstream analyses.

In addition, the cell line-centric view includes tabs for the previously mentioned experimental data. The "Cell Perturbation" tab (data available for currently 22 tissues) provides access to dose-resolved information on the protein and PTM level, similar to the drug-centric view (Fig. 2D), with the

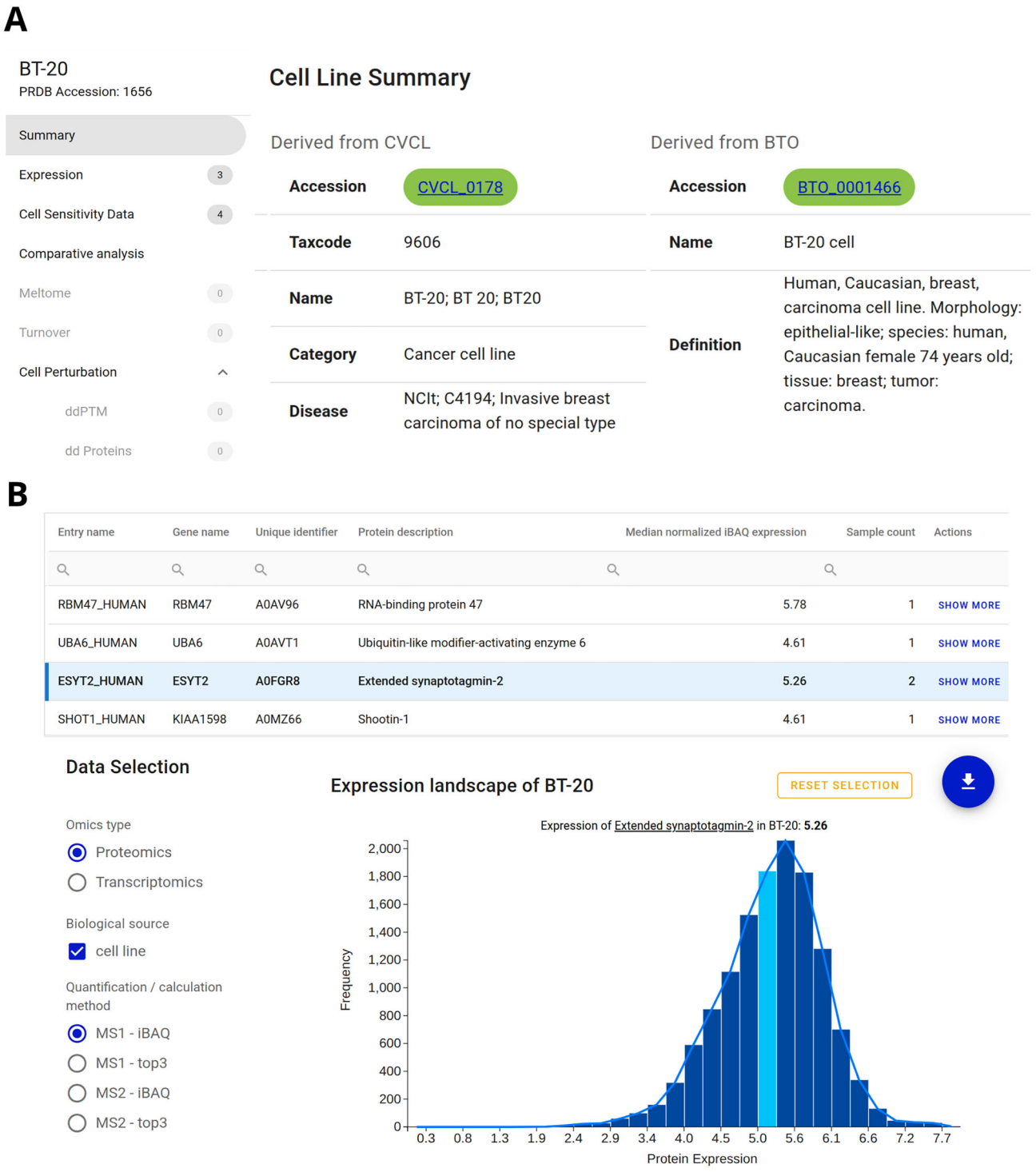


Figure 3. The cell line-centric view's summary and expression profile tab for the cell line BT-20. (**A**) The cell line summary provides lists with key-value annotations for ProteomicsDB's internal entry and all external databases, in this case, Cellosaurus (left) and the Brenda Tissue Ontology (right) figure for the cell line-centric view. The accession names act as links to the landing page of the cell line in the external resource. The left navigation lists tabs for available experimental data; numbers indicate how much data are available for each type. (**B**) The expression profile is visualized as a table and histogram for the selected omics type and quantification/calculation method. The expression data are provided as a table, which can be searched and filtered for individual entries. Selected entries, in this case the protein ESYT2, are highlighted in the histogram in the bin corresponding to the expression value. Table values are aggregated across all available samples for each protein/transcript using the median of normalized expression values. "Show more" opens the list of samples for an entry; clicking any row reveals the observed peptides for the corresponding protein, supporting the identification of the selected protein (proteomics). Selecting a peptide jumps to the protein-centric PSM view for full spectral context. The download button allows retrieving the figure as a high-resolution PNG image or an editable SVG vector graphic.

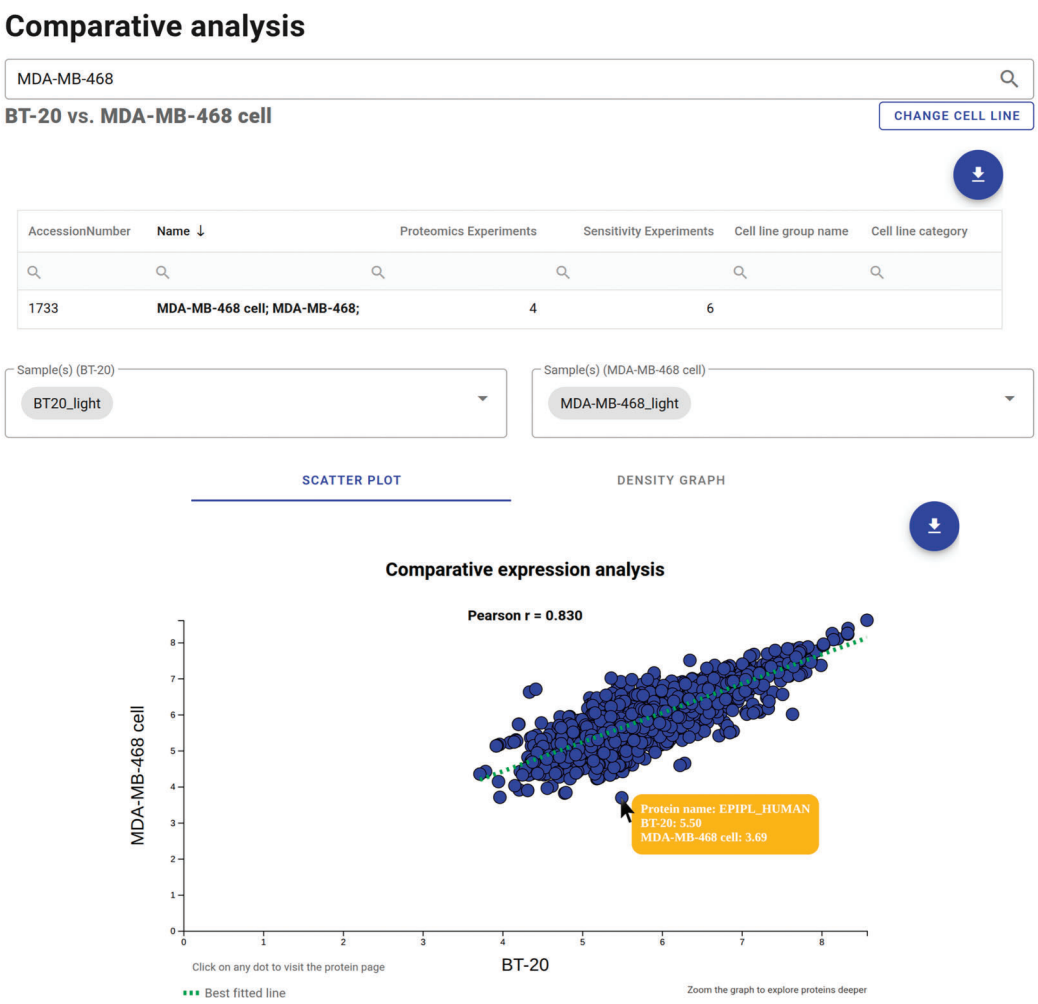


Figure 4. Expression profile comparison to other cell lines/tissues. This feature enables a bivariate expression analysis by allowing users to select a second cell line or tissue. Each protein is represented as an individual dot, with expression values aggregated (median) for the chosen samples of each cell line/tissue (current cell line on the *x*-axis and the selected cell line on the *y*-axis). A best-fit line and Pearson's correlation summarize concordance. For instance, the breast cancer cell lines BT-20 and MDA-MB-468 demonstrate a high correlation (Pearson's $r = 0.83$), as anticipated. Outliers, such as EPIPL in this example, can be identified by hovering over a dot, which displays protein information and exact expression values for both cell lines (rectangle); clicking on a dot navigates to the protein-centric view. Additionally, users can visualize the density of points instead of individual data points by selecting the option above the plot. Both plots are available for download as high-resolution PNGs or editable SVG vector graphics.

difference that experiments are collected for all drugs that were examined for the cell line. Similarly, the "Meltome" and "Turnover" tabs, which were introduced previously as part of the protein-centric view [3], are now available in a cell line-centric context. Currently, a total of ~13 000 protein melting points from thermal proteome profiling is available for 19 cell lines, and synthesis/degradation (turnover) curves for >6000 proteins are available for one (HeLa) cell line. Finally, the "Cell Sensitivity" tab embeds a pre-filtered cell sensitivity analytics tool (Fig. 1), containing data specific to the cell line for the experiments listed in the summary page.

Hence, the new cell line-centric view represents the centerpiece of ProteomicsDB's perturbation update, providing tools to analyze cell lines individually or compare their proteomics profiles across different (cancer) tissues, enabling drill-down to proteins and peptides connecting to the protein- and peptide-centric views, while linking to cell line ontologies and resources allows further investigations with external resources.

Conclusion and outlook

This update highlights ProteomicsDB's effort to interconnect heterogeneous experimental data, extending beyond its prior protein-, peptide-, and project-centric views. Specifically, the addition of the new drug- and cell line-centric views allows real-time data exploration from various angles. This makes ProteomicsDB a central hub for mapping baseline expression profiles at the transcriptomic and proteomic level to phenotypic outcomes, including dose-resolved expression on the protein and PTM level, phenotypic cell sensitivity data, and experimentally validated target proteins. Connecting ProteomicsDB to established external resources allows metadata retrieval for tissue and drugs to map and integrate public datasets, while the addition of the *Drugst.One* network viewer allows for exploration of similarities in drug target space, complementing available protein target data. While *Drugst.One* does not currently make use of expression information, ProteomicsDB fills this gap by coupling drugs with experimentally validated protein targets from dose-resolved

expression proteomics experiments. Further integration and exchange between these resources is planned, supporting ProteomicsDB as an additional source for experimentally validated drug–target protein relationships, allowing users of other platforms that make use of the *Drugst.One* viewer to access this information.

The platform's organism-agnostic design makes it well-suited for incorporating dose-resolved expression proteomics studies across diverse species, particularly in the context of preclinical drug development. An example of this is the *NHPig* project (https://www.nhpig.eu), where ProteomicsDB is serving as the project's data backbone for integrating mini- and micropig proteomics and metadata to support non-clinical safety assessment. Another application in the clinical direction is the integration of bacterial proteomes, which could enable systematic exploration of antibiotic responses and resistance mechanisms on the proteomic level.

Beyond this, plant proteomics represents an important research area for the future in light of challenges related to environmental conditions and a changing climate. With the addition of cell sensitivity support, ProteomicsDB is now, in principle, capable of handling similar dose–response type data, e.g., pesticide exposure, drought stress, or radiation, providing insights into resilience and adaptation to a changing environment at the molecular level. An ongoing example is *The Proteomes that Feed the World* (https://www.elitenetzwerk.bayern.de/en/home/funding-programs/doctoral-excellence/overview-of-international-doctorate-programs/proteomes-that-feed-the-world), a project mapping the proteomes of 100 crop species critical for human nutrition.

The cell line-centric approach is also laying the groundwork for single-cell and spatial proteomics, which promises deep insights into, e.g., cancer heterogeneity on a cellular level. This is a comparatively young discipline that is expected to produce large amounts of data that may prove of great value in understanding the phenotypic behavior of cancer subtypes or rare diseases [44]. ProteomicsDB's in-memory database and real-time visualization capabilities make it the perfect platform for such an application. ProteomicsDB's APIs already allow a connection to analysis pipelines, e.g., in the context of cell sensitivity prediction for future treatment recommendations, which requires rich and comprehensive training and metadata. One such tool is DrEval [45], a framework for fair comparison and evaluation of model performance.

Together, these advancements highlight how ProteomicsDB is evolving from a former protein-centric database into an integrative, multi-species, multi-omics resource linking molecular profiles to phenotypic effects. Its real-time analytics, extensible architecture, capacity for integration with complementary tools, and AI approaches make it a uniquely positioned platform to foster future applications in precision medicine and biomarker discovery in humans, animals, plants, and beyond.

Acknowledgements

The authors thank the SAP University Competence Center at the Technical University of Munich for hosting the SAP HANA infrastructure; Lisa Spindler for delivering support for the *Drugst.One* integration; and all members of the wilhelmlab, daisybio, and kusterlab for their input and fruitful discussions. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Author contributions: Mario Picciani (Conceptualization [equal], Data curation [equal], Software [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Armin Soleymaniniya (Conceptualization [equal], Data curation [equal], Software [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Julian Mueller(Conceptualization [equal], Data curation [equal], Software [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Amirhossein Sakhteman (Data curation [supporting], Software [equal], Writing—original draft [supporting], Writing—review & editing [supporting]), Judith Bernett (Data curation [supporting], Writing—review & editing [supporting]), Elias Davide Kahl (Software [supporting]), Konstantinos Tzanakis (Software [supporting]), Johannes Kersting (Software [supporting], Writing—review & editing [supporting]), Mohsen Pourjam (Software [supporting], Writing—review & editing [supporting]), Luis Augusto Eijy Nagai (Software [supporting], Writing—review & editing [supporting]), Firas Hamood (Writing—review & editing [supporting]), Markus List (Funding acquisition [supporting], Resources [supporting], Supervision [supporting], Writing—review & editing [supporting]), Bernhard Kuster (Funding acquisition [supporting], Resources [supporting], Supervision [supporting]), Matthew The (Funding acquisition [supporting], Resources [supporting], Supervision [supporting]), and Mathias Wilhelm (Conceptualization [lead], Funding acquisition [lead], Resources [lead], Supervision [lead], Writing—review & editing [equal])

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

M.W. and B.K. are founders and shareholders of MSAID GmbH and members of the scientific advisory board of Momentum Biotechnologies, with no operational role in either company. M.L. consults for mbiomics GmbH. Neither company affiliation had any influence on the results presented in this study.

Funding

This work was partially funded by the European Union under the Horizon Europe grant (REPO4EU, 101057619) and the ERC grants (ORIGIN, 101077037) and (TOPAS, 833710). This work was also partially supported by the German Federal Ministry of Research, Technology and Space (BMFTR) within the framework of the CompLS funding concept (DROP2AI, 031L0305A) and the MSCoreSys funding concept (CLINSPECT-M-2, 16LW0243K). It was further partially supported by the German Research Foundation (DFG) (PhoSAIC, 537476536), the Elite Network of Bavaria (F-6-M5613.6.K-NW-2021-411/1/1), and the Swiss State Secretariat for Education, Research, and Innovation (SERI) under contract No. 22.00115. Open Access funding enabled and organized by Projekt DEAL.

Data availability

ProteomicsDB is available at https://www.proteomicsdb.org. No previously unpublished data were created as part of this manuscript. A full description of imported datasets since the last update is available in Supplementary Tables S1--S4. Documentation for API endpoints is available at https://www.proteomicsdb.org/api. The website is optimized for usage with Google Chrome (≥ version 127) and Mozilla Firefox (≥ version 136).

References

1. Farber F, May N, Lehner W *et al.* The SAP HANA database—an architecture overview. *IEEE Comput Soc Data Eng Bull* 2012;35:28–33.
2. Schmidt T, Samaras P, Frejno M *et al.* ProteomicsDB. *Nucleic Acids Res* 2018;46:D1271–81. https://doi.org/10.1093/nar/gkx1029
3. Samaras P, Schmidt T, Frejno M *et al.* ProteomicsDB: a multi-omics and multi-organism resource for life science research. *Nucleic Acids Res* 2020;48:D1153–63.
4. Lautenbacher L, Samaras P, Muller J *et al.* ProteomicsDB: toward a FAIR open-source resource for life-science research. *Nucleic Acids Res* 2022;50:D1541–52. https://doi.org/10.1093/nar/gkab1026
5. Basu A, Bodycombe NE, Cheah JH *et al.* An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* 2013;154:1151–61. https://doi.org/10.1016/j.cell.2013.08.003
6. Seashore-Ludlow B, Rees MG, Cheah JH *et al.* Harnessing connectivity in a large-scale small-molecule sensitivity dataset. *Cancer Discov* 2015;5:1210–23. https://doi.org/10.1158/2159-8290.CD-15-0235
7. Rees MG, Seashore-Ludlow B, Cheah JH *et al.* Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat Chem Biol* 2016;12:109–16. https://doi.org/10.1038/nchembio.1986
8. Barretina J, Caponigro G, Stransky N *et al.* The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;483:603–7. https://doi.org/10.1038/nature11003
9. Garnett MJ, Edelman EJ, Heidorn SJ *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* 2012;483:570–5. https://doi.org/10.1038/nature11005
10. Yang W, Soares J, Greninger P *et al.* Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* 2013;41:D955–61. https://doi.org/10.1093/nar/gks1111
11. Iorio F, Knijnenburg TA, Vis DJ *et al.* A landscape of pharmacogenomic interactions in cancer. *Cell* 2016;166:740–54. https://doi.org/10.1016/j.cell.2016.06.017
12. Arafeh R, Shibue T, Dempster JM *et al.* The present and future of the cancer dependency map. *Nat Rev Cancer* 2025;25:59–73. https://doi.org/10.1038/s41568-024-00763-x
13. Luna A, Elloumi F, Varma S *et al.* CellMiner cross-database (CellMinerCDB) version 1.2: exploration of patient-derived cancer cell line pharmacogenomics. *Nucleic Acids Res* 2021;49:D1083–93. https://doi.org/10.1093/nar/gkaa968
14. Eckert S, Berner N, Kramer K *et al.* Decrypting the molecular basis of cellular drug phenotypes by dose-resolved expression proteomics. *Nat Biotechnol* 2025;43:406–15. https://doi.org/10.1038/s41587-024-02218-y
15. Zecha J, Bayer FP, Wiechmann S *et al.* Decrypting drug actions and protein modifications by dose- and time-resolved proteomics. *Science* 2023;380:93–101. https://doi.org/10.1126/science.ade3925
16. Maier A, Hartung M, Abovsky M *et al.* Drugst.One—a plug-and-play solution for online systems medicine and network-based drug repurposing. *Nucleic Acids Res* 2024;52:W481–8. https://doi.org/10.1093/nar/gkae388
17. Gremse M, Chang A, Schomburg I *et al.* The BRENDA Tissue Ontology (BTO): the first all-integrating ontology of all organisms for enzyme sources. *Nucleic Acids Res* 2011;39:D507–13. https://doi.org/10.1093/nar/gkq968
18. Bairoch A. The cellosaurus, a cell-line knowledge resource. *J Biomol Tech* 2018;29:25–38. https://doi.org/10.7171/jbt.18-2902-002
19. Sarntivijai S, Lin Y, Xiang Z *et al.* CLO: the cell line ontology. *J Biomed Semantics* 2014;5:37. https://doi.org/10.1186/2041-1480-5-37
20. Patel H, Ewels P, Peltzer A *et al.* nf-core/rnaseq: nf-core/rnaseq v3.10.1 - Plastered Rhodium Rudolph. Version 3.10.1. Zenodo. https://zenodo.org/records/7505987, 5 January 2023.
21. Picciani M, Gabriel W, Giurcoiu VG *et al.* Oktoberfest: open-source spectral library generation and rescoring pipeline based on Prosit. *Proteomics* 2024;24:2300112. https://doi.org/10.1002/pmic.202300112
22. Ghandi M, Huang FW, Jané-Valbuena J *et al.* Next-generation characterization of the cancer cell line encyclopedia. *Nature* 2019;569:503–8. https://doi.org/10.1038/s41586-019-1186-3
23. Nusinow DP, Szpyt J, Ghandi M *et al.* Quantitative proteomics of the cancer cell line Encyclopedia. *Cell* 2020;180:387–402. https://doi.org/10.1016/j.cell.2019.12.023
24. Gonçalves E, Poulos RC, Cai Z *et al.* Pan-cancer proteomic map of 949 human cell lines. *Cancer Cell* 2022;40:835–49. https://doi.org/10.1016/j.ccell.2022.06.010
25. Schmidt T, Samaras P, Dorfer V *et al.* Universal spectrum explorer: a standalone (web-)application for cross-resource spectrum comparison. *J Proteome Res* 2021;20:3388–94. https://doi.org/10.1021/acs.jproteome.1c00096
26. Kim S, Chen J, Cheng T *et al.* PubChem 2025 update. *Nucleic Acids Res* 2025;53:D1516–25. https://doi.org/10.1093/nar/gkae1059
27. Mendez D, Gaulton A, Bento AP *et al.* ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res* 2019;47:D930–40. https://doi.org/10.1093/nar/gky1075
28. Kabella N, Bayer FP, Stamatiou K *et al.* Proteomic analyses identify targets, pathways, and cellular consequences of oncogenic KRAS signaling. *Sci Signal* 2025;18:eadt6552. https://doi.org/10.1126/scisignal.adt6552
29. Höfer S, Frasch L, Brajkovic S *et al.* Gemcitabine and ATR inhibitors synergize to kill PDAC cells by blocking DNA damage response. *Mol Syst Biol* 2025;21:231–53. https://doi.org/10.1038/s44320-025-00085-6
30. Chang YC, Gnann C, Steimbach RR *et al.* Decrypting lysine deacetylase inhibitor action and protein modifications by dose-resolved proteomics. *Cell Rep* 2024;43:114272. https://doi.org/10.1016/j.celrep.2024.114272
31. Lechner S, Malgapo MIP, Grätz C *et al.* Target deconvolution of HDAC pharmacopoeia reveals MBLAC2 as common off-target. *Nat Chem Biol* 2022;18:812–20. https://doi.org/10.1038/s41589-022-01015-5
32. Prokofeva P, Höfer S, Hornisch M *et al.* Merits of diazirine photo-immobilization for target profiling of natural products and cofactors. *ACS Chem Biol* 2022;17:3100–9. https://doi.org/10.1021/acschembio.2c00500
33. Lee CY, The M, Meng C *et al.* Illuminating phenotypic drug responses of sarcoma cells to kinase inhibitors by phosphoproteomics. *Mol Syst Biol* 2024;20:28–55. https://doi.org/10.1038/s44320-023-00004-7
34. Bayer FP, Gander M, Kuster B *et al.* CurveCurator: a recalibrated F-statistic to assess, classify, and explore significance of dose–response curves. *Nat Commun* 2023;14:7902. https://doi.org/10.1038/s41467-023-43696-z

35. Probst D, Reymond JL. SmilesDrawer: parsing and drawing SMILES-encoded molecular structures using client-side JavaScript. *J Chem Inf Model* 2018;58:1–7. https://doi.org/10.1021/acs.jcim.7b00425
36. Cioce M, Fazio VM. EphA2 and EGFR: friends in life, partners in crime. Can EphA2 be a predictive biomarker of response to anti-EGFR agents? *Cancers* 2021;13:700. https://doi.org/10.3390/cancers13040700
37. Shi X, Wang B. Caught in the "Akt": cross-talk between EphA2 and EGFR through the Akt-PIKfyve axis maintains cellular sensitivity to EGF. *Sci Signal* 2018;11:eaau1207. https://doi.org/10.1126/scisignal.aau1207
38. Hornbeck PV, Zhang B, Murray B *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* 2015;43:D512–20. https://doi.org/10.1093/nar/gku1267
39. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2025. *Nucleic Acids Res* 2025;53:D609–17. https://doi.org/10.1093/nar/gkae1010
40. Szklarczyk D, Nastou K, Koutrouli M *et al.* The STRING database in 2025: protein networks with directionality of regulation. *Nucleic Acids Res* 2025;53:D730–7. https://doi.org/10.1093/nar/gkae1113
41. Sadegh S, Skelton J, Anastasi E *et al.* Network medicine for disease module identification and drug repurposing with the NeDRex platform. *Nat Commun* 2021;12:6848. https://doi.org/10.1038/s41467-021-27138-2
42. Knox C, Wilson M, Klinger CM *et al.* DrugBank 6.0: the DrugBank knowledgebase for 2024. *Nucleic Acids Res* 2024;52:D1265–75. https://doi.org/10.1093/nar/gkad976
43. Hoadley KA, Yau C, Hinoue T *et al.* Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* 2018;173:291–304. https://doi.org/10.1016/j.cell.2018.03.022
44. Nordmann TM, Anderton H, Hasegawa A *et al.* Spatial proteomics identifies JAKi as treatment for a lethal skin disease. *Nature* 2024;635:1001–9. https://doi.org/10.1038/s41586-024-08061-0
45. Bernett J, Iversen P, Picciani M *et al.* From hype to health check: critical evaluation of drug response prediction models with DrEval. bioRxiv, https://doi.org/10.1101/2025.05.26.655288, 29 May 2025, preprint: not peer reviewed.

Received: September 15, 2025. **Revised:** October 23, 2025. **Accepted:** October 24, 2025