Nucleic Acids Research, 2025, gkaf1154
https://doi.org/10.1093/nar/gkaf1154
Database issue

Therapeutic target database 2026: facilitating targeted therapies and precision medicine

Yintao Zhang [1,†], Ying Zhou [2,†], Hangwei Xu [1], Wanghao Jiang [1], Bo Li [3], Dianyu Lai [4], Cong Wan [5], Shanshan Wang [6], Mingxiao Zhao [6], Ying Tan [7], Songlin Lu [8], Tingting Fan [8], Xin Liu [3], Feng Zhu [1,2,*], Yuzong Chen [8,*]

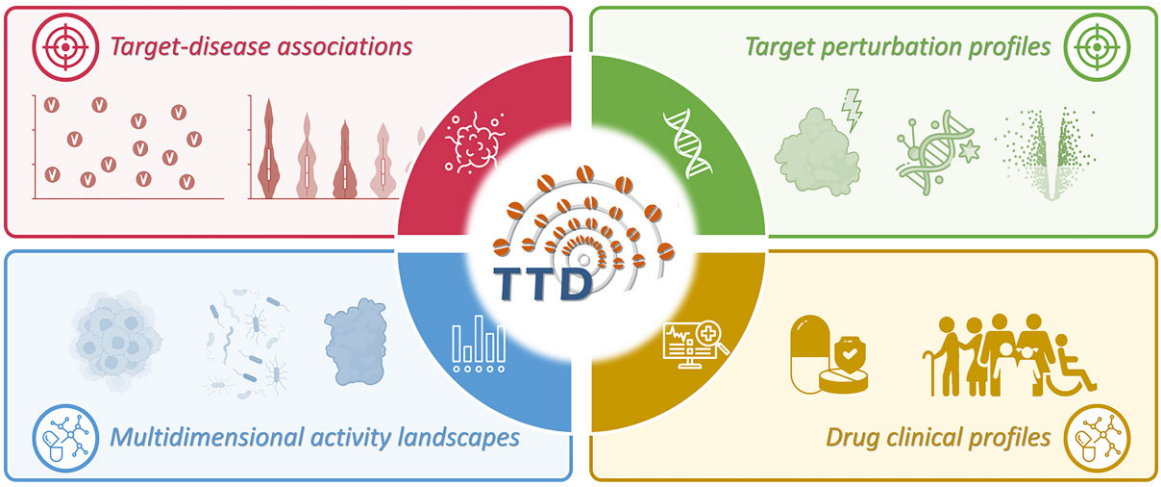
[1]College of Pharmaceutical Sciences, State Key Laboratory of Advanced Drug Delivery and Release Systems, Zhejiang University, Hangzhou 310058, China
[2]Department of Pharmacy, Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, China
[3]Wisdom Lake Academy of Pharmacy, Xi'an Jiaotong–Liverpool University, Suzhou 215123, China
[4]Chu Kochen Honors College, Zhejiang University, Hangzhou 310058, China
[5]School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, United Kingdom
[6]Qian Xuesen Collaborative Research Center of Astrochemistry and Space Life Sciences, Institute of Drug Discovery Technology, Ningbo University, Ningbo 315211, China
[7]State Key Laboratory of Chemical Oncogenomics, Key Laboratory of Chemical Biology, The Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China
[8]Institute of Biomedical Health Technology and Engineering, Shenzhen Bay Laboratory, Shenzhen 518000, China

*To whom correspondence should be addressed. Email: zhufeng@zju.edu.cn
Correspondence may also be addressed to Yuzong Chen. Email: chenyuzong@sz.tsinghua.edu.cn
†The first two author should be regarded as Joint First Authors.

Abstract

Development of targeted therapeutics begins with the discovery and validation of therapeutic targets, which builds the foundation for rational drug design. Extensive information on target–disease associations, target perturbation profiles, drug bioactivity landscapes, and clinical profiles is critical for identifying disease-relevant targets, elucidating their biological functions, and assessing the therapeutic potential. Therapeutic target database 2026 represents a major update that expands multiple layers of data essential for drug discovery. Key expansions include (i) 306 247 target–disease associations covering 2912 targets, (ii) 10 506 perturbation profiles triggered by genetic modification or chemical interference on 2368 targets, (iii) multidimensional activity landscapes of 17 806 drugs, encompassing cytotoxic, antimicrobial-, and molecular-level activities, and (iv) abundant clinical profiles for 2234 approved drugs. As a result, this update provides expanded and curated information on 3798 targets and 40 398 drugs. To enhance usability and scalability, the database framework has been rebuilt, which can be freely accessible without any login requirement at https://idrblab.org/ttd/.

Graphical abstract



Received: September 14, 2025. **Revised:** October 9, 2025. **Accepted:** October 11, 2025

Introduction

Development of targeted therapeutics starts from the discovery and validation of therapeutic targets, which builds the foundation for rational drug design. Identifying associations between macromolecules and diseases (target–disease associations) plays a critical role in discovering new therapeutic targets and guiding drug development efforts [1, 2]. Understanding transcriptomic changes when a particular target is perturbed (target perturbation profile) is valuable not only for uncovering protein function and disease mechanisms but also for drug screening and combination therapy design [3, 4]. Characterizing the multidimensional drug bioactivities (drug bioactivity landscape) is key to unlocking a drug's therapeutic potential, improving its safety profile and facilitating drug repurposing and personalized medicine [5, 6]. In other words, such data can not only elucidate molecular mechanisms underlying diseases and support the prioritization of targets [7] but also aid in the design of effective interventions [8] and the optimization of drug selectivity and safety [9].

So far, numerous databases are widely used in the field of target identification and drug discovery. Some databases, such as DrugBank [10], DrugCentral [11], and DrugMAP [12], are dedicated to offering the pharmacological information about drugs. Others, including therapeutic target database (TTD) [13] and Open Targets [7], primarily focus on providing extensive information on drug targets. Additionally, resources like PubChem [14], ChEMBL [15], and BindingDB [16] specialize in general molecule and bioactivity information. Although all these databases have garnered significant citations and made substantial contributions to the related fields, the need for fully integrating the diverse layers of information mentioned above remains unmet.

Here, we describe a major update of the TTD, featuring extensive data on target–disease associations, target perturbation profiles, and multidimensional drug activities. First, target–disease associations supported by genetic and transcriptomic evidence were systematically compiled. In particular, genomic mutations of targets associated with corresponding diseases were sourced from UniProt, while disease-specific expression patterns of targets were analyzed using gene expression omnibus (GEO) and The Cancer Genome Atlas (TCGA) data. This effort resulted in a total of 306 247 target–disease associations, covering 441 successful (targeted by at least one approved drug), 977 clinical trial (not targeted by any approved drug, but targeted by at least one clinical trial drug), 193 preclinical/patented (not targeted by any approved/clinical trial drug, but targeted by at least one preclinical/patented drug), and 1301 literature-reported (targeted by experimental drugs only) targets. Second, the expression profiles resulting from perturbations of 407 successful, 812 clinical trial, 166 patented/preclinical, and 983 literature-reported targets were provided based on the data derived from CMap/LINCS, covering 8047 genetic modifications and 2459 chemical interferences. Third, multidimensional drug activity landscapes were offered for 2443 approved, 2474 clinical/preclinical, and 12 889 investigational drugs, compiled from BindingDB, ChEMBL, PubChem, etc. Specifically, molecular-level activities of 16 983 drugs against 5253 proteins/targets were collected, while cytotoxic activities of 3221 drugs against 2367 disease cell lines and antimicrobial activities for 1238 drugs against 803 microbial organisms were explicitly documented.

More importantly, several core sections of TTD were updated in a timely manner based on recent resources (including Drugs@FDA, Clinicaltrials.gov, and company's official reports), leading to a total of 3798 targets and 40 398 drugs, together with the comprehensive FDA label information for 2234 approved drugs (showing detailed drug information of 'dosage and administration,' 'adverse reaction,' 'nonclinical toxicology,' 'drug interaction,' 'population-specific usage,' and 'pharmacodynamics and pharmacokinetics'). For facilitating more convenient and smoother access to the more expanded information contents, the server framework of TTD was reconstructed using Vue3 and Django to streamline the integrations of new features and to meet the rapidly growing user demand. TTD 2026 is accessible at https://idrblab.org/ttd/.

Updates to database content

TTD 2026 release encompasses four key aspects: target–disease associations, target perturbation profiles, multidimensional drug activity landscapes, and clinical profiles of approved drugs. Detailed descriptions of each aspect were described as follows.

Collection of target–disease associations from genetic and transcriptomic evidence

Investigating macromolecule–disease associations provides critical insights into the molecular mechanisms underlying diseases, facilitating the identification of new therapeutic targets and the development of innovative treatment strategies [1, 17]. To uncover these associations, a variety of approaches have been employed, including genome-wide association studies (GWAS), transcriptomic profiling, and cross-cancer correlation studies. For instance, GWAS have revealed thousands of genetic variants associated with disease susceptibility and relevant phenotypic traits [18]. Meanwhile, cross-cancer analysis of protein expression and inter-tumor correlations help to uncover conserved mechanisms and pan-cancer targets, such as CDK9 coexpression patterns [19], supporting drug repurposing and the design of shared immunotherapies [20].

Target–disease associations based on genetic mutation

In this update, the target–disease associations based on genetic mutation were collected through following steps. First, all amino acid mutation records for TTD protein targets were retrieved using the variant viewer application programming interface (API) of UniProt [21]. Second, only those variants supported by curated clinical or experimental evidence were retained to ensure higher confidence in the target–disease associations. Third, the detailed mutation information was gathered from its original resources including ClinVar [22], dbSNP [23], and ClinGen [24]. As illustrated in the left panel of Fig. 1A, all pathogenic variants of a specific protein target were displayed in a tabular format, summarizing key information such as position, amino acid change, consequence, and associated disease. Each entry can be expandable to present additional details, such as codon changes, mutation type, frequencies, clinical significance, genomic locations, and external cross-reference. Additionally, a panoramic visualization of all mutation–disease associations for a given target was also provided (as shown in the right panel of Fig. 1A), which allowed

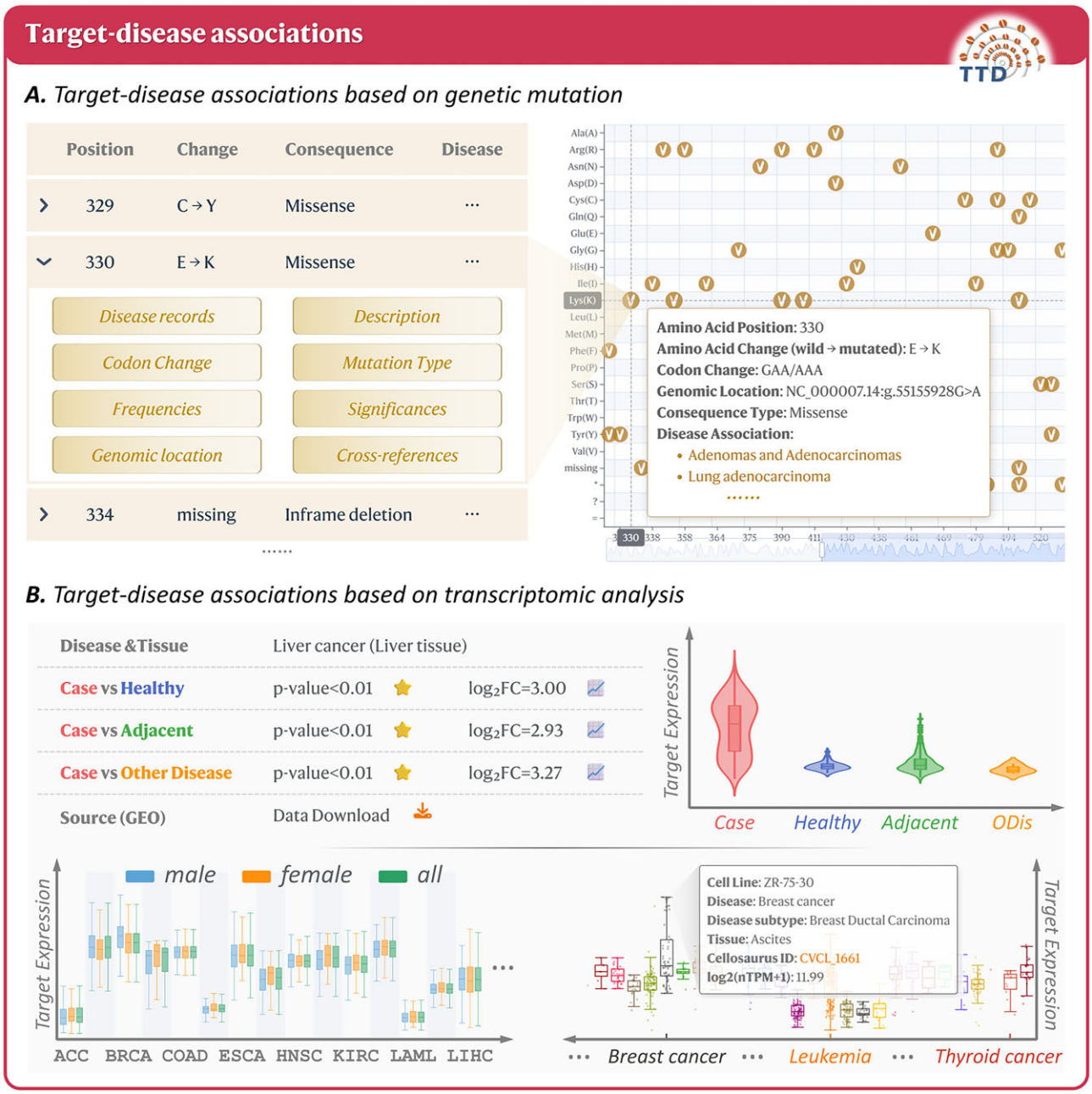


Figure 1. Enriched target–disease associations in TTD 2026. (**A**) Target–disease associations based on genetic mutation. The left panel lists target mutations and their associated disease, including key details such as mutation position, amino acid change, consequence, and disease. Each entry can be expanded to reveal detailed records, including codon change, mutation type, frequencies, significances, genomic location, and so on. The right panel provides a visualization of all mutation–disease associations, with interactive nodes displaying mutation details and corresponding diseases. (**B**) Target–disease associations based on transcriptomic analysis. The upper section shows differential expression of targets between disease and control tissues, including statistical metrics such as *P*-values and $\log_2$FC, accompanied by violin plots. The lower section illustrates expression patterns across cancers. Clinical sample data are shown as boxplots (left) stratified by cancer type and sex; and grouped cancer cell line data are presented as boxplots (right) with overlaid scatter points representing individual cell line values. ACC: adrenocortical carcinoma; BRCA: breast invasive carcinoma; COAD: colon adenocarcinoma; ESCA: esophageal carcinoma; HNSC: head and neck squamous cell carcinoma; KIRC: kidney renal clear cell carcinoma; LAML: acute myeloid leukemia; LIHC: liver hepatocellular carcinoma; ODis: other disease.

users to explore detailed information about mutations and associated diseases by hovering over individual nodes.

Target–disease associations based on transcriptomic analysis

In this update, the target–disease associations based on transcriptomic analyses include the following two parts: disease-specific differential expression and pan-cancer expression patterns.

For disease-specific differential expression analysis, the Affymetrix HG-U133 Plus 2.0 microarray datasets, specifically designed for human gene expression profiling with probes targeting ~22 000 human genes were retrieved from the GEO [25]. Each dataset was manually annotated with disease type and tissue origin, covering 98 disease classes and

59 tissues. Samples from the same disease within the same tissue were pooled, and raw intensity values were processed using the RMA function of the *affy* R package [26]. The median expression array was chosen as the reference baseline, followed by global normalization across all arrays. Differential expression was assessed by combining fold-change analysis with *t*-tests, comparing disease samples against healthy controls, adjacent tissues, or other disease samples (if available). A target was considered associated with a given disease if any comparison yielded a *P*-value ≤ 0.5 and an absolute $\log_2$ fold change ≥ 1. As shown in the upper part of Fig. 1B, a violin plot illustrated the expression differences of a given target for each disease, along with metadata on disease class, tissue, *P*-values, and fold changes.

For pan-cancer expression analysis, both clinical samples and cancer cell lines were examined within an integrated framework. Clinical RNA-seq data were obtained from TCGA [27] and Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project, which comprised 33 cancer types with metadata specifying sex and 9 pediatric cancer types, respectively. Expression values in TPM (transcripts per million) format were downloaded from the UCSC Xena platform [28]. Complementary RNA-seq data from cancer cell lines were obtained from the Human Protein Atlas [29], which provided normalized TPM expression profiles across a wide range of cancer lineages. In the latest version of TTD, expression patterns of targets from clinical samples were visualized as boxplots, reflecting variation across cancer types and between sexes (lower-left panel of Fig. 1B), whereas expression pattern from cell lines were presented as boxplots with overlaid scatter points, featuring interactive elements that allow inspection of cell line-specific information (lower-right panel of Fig. 1B).

As a result, a total of 306 247 target–disease associations derived from genetic and transcriptomic evidence were given in TTD, covering 441 successful, 977 clinical trial, 193 preclinical/patented, and 1301 literature-reported targets.

Curation of target perturbation data induced by different mechanisms

Analyzing the transcriptomic changes resulting from the perturbation of a specific target (described here as target perturbation profile) through different mechanisms, such as knockout, overexpression, or chemical interference, provides a powerful framework for characterizing the functional consequences of target modulation [30]. By capturing genome-wide transcriptional changes, these profiles can reveal the biological roles of genes, the pathways they regulate, their involvement in diseases, and the potential effects of their perturbation [3]. In drug discovery, such data not only enable the identification and validation of therapeutic targets but also help clarify mechanisms of action, uncover potential off-target effects, and support drug repurposing as well as the design of rational combination therapy [31]. Furthermore, comparing perturbation signatures across genes, compounds, and diseases allows researchers to identify functional connections and therapeutic opportunities. That is to say, target perturbation profile represents a valuable resource for advancing both fundamental biology and translational therapeutics [32].

Target perturbation profiles were systematically incorporated into TTD using data derived from CMap/LINCS, a large-scale repository that characterizes effects of genetic and pharmacologic perturbagens [33]. The detailed data procedure was as follows. First, the raw perturbation data (induced by gene knockdown, overexpression, and compound/antibody treatment) were retrieved from the CMap/LINCS Level 4 datasets, which provided gene-level *Z*-scores normalized across all samples on each plate. Samples sharing the same brew prefix (encoding cell line, time point, perturbation, and dose) were grouped together after excluding those that failed quality control. Within each group, perturbation samples were compared with matched controls to identify differential expression, using only the 978 landmark genes directly measured in the L1000 assay. Differential expression analysis was then carried out with the *limma* R package [34]. As shown in Fig. 2, in the case of *lapatinib*-mediated perturbation, recorded attributes include cell line, Cellosaurus ID, experimental conditions, and control group. Perturbation results were further annotated with differentially expressed genes (DEGs), including symbols, gene IDs, statistical metrics, and links to TTD Target (if available). To facilitate data exploration, interactive volcano plots and heatmaps were also provided, enabling users to visualize DEG distributions and expression changes associated with each perturbation experiment. As a result, the expression profiles resulting from perturbations of 407 successful, 812 clinical trial, 166 patented/preclinical, and 983 literature-reported targets were provided in TTD, covering 8047 genetic modifications and 2459 chemical interferences.

Integration of multidimensional activity landscape of drugs

Comprehensive drug activity landscapes, spanning cytotoxic, antimicrobial, and molecular-level data, give a multidimensional view of drugs' therapeutic potential, limitations, and mechanisms of action [35, 36]. Cytotoxicity profiles are useful to reveal resistance mechanisms, identify opportunities for indication expansion, and highlight predictive biomarkers that inform precision oncology [37]. Antimicrobial activity maps define susceptibility and resistance patterns, reveal evolutionary trajectories, and guide rational drug combinations or "cocktail" regimens [38]. At the molecular level, activity data clarify mechanisms of action, uncover secondary and off-target effects, and illuminate the structure–activity relationships, supporting optimization for improved efficacy and safety [39]. Collectively, these multidimensional activity landscapes offer critical insights for drug discovery, repurposing, and precision medicine.

Cytotoxic and antimicrobial activity landscape of drugs

In this update, raw cytotoxic and antimicrobial activity data for drugs in TTD were retrieved from GDSC [40], CCLE [41], and ChEMBL [15]. Drug entries from different sources were matched using synonym libraries and chemical identifiers to ensure unambiguous alignment. Only quantifiable activity metrics (e.g. IC_{50}, MIC, GI_{50}) were retained, and all values were normalized to nanomolar (nM) or micrograms per milliliter (μg/ml) units. Cell line information was cross-referenced with Cellosaurus [42] and CCLE [41] to incorporate additional annotations, including synonyms and disease indications. For microbial assays, taxonomic identifiers were mapped to the NCBI Taxonomy database [43], enabling hierarchical classification from kingdom down to species. For datasets obtained from the same platform, such

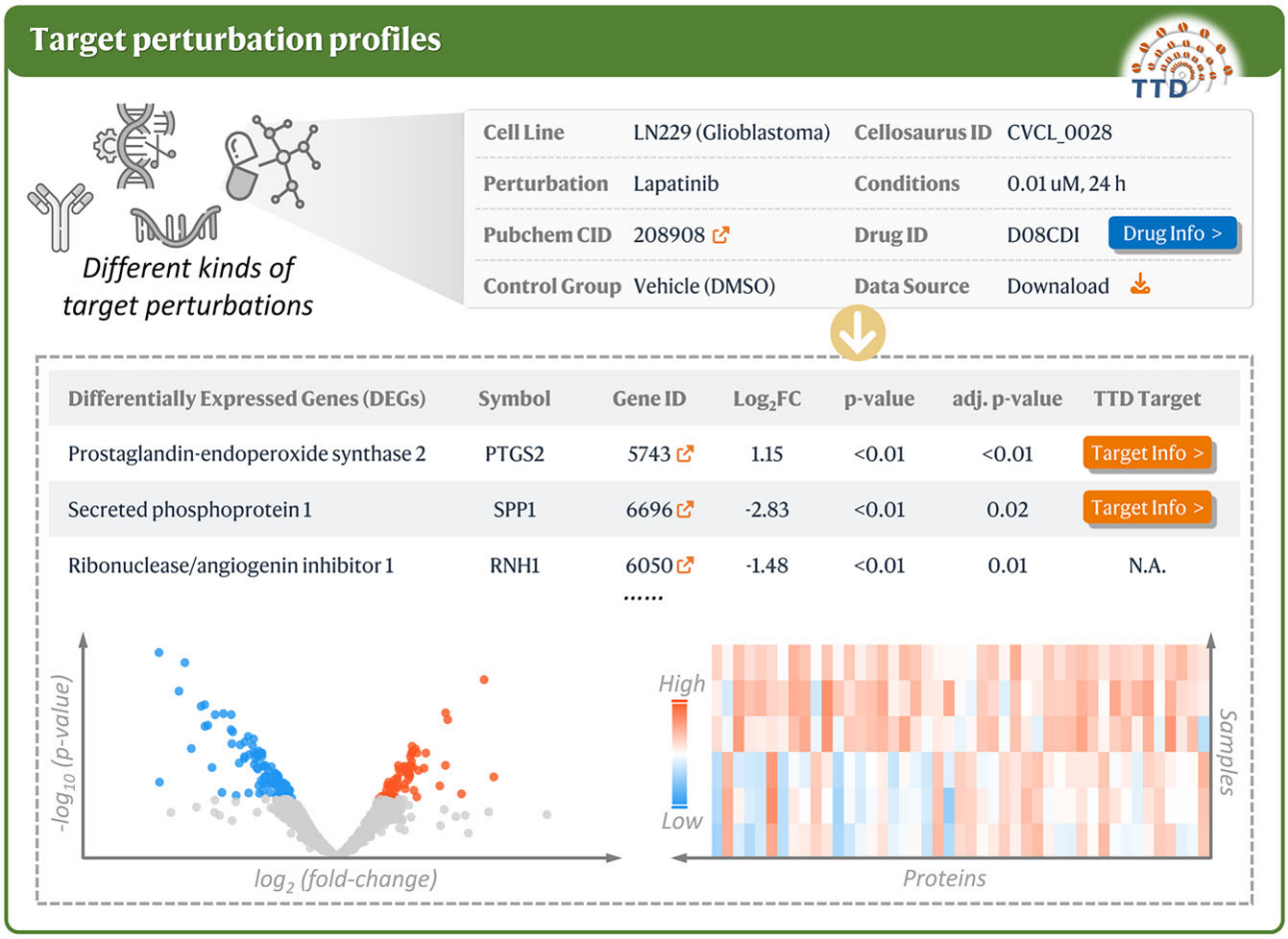


Figure 2. Target perturbation profiles provided in the TTD 2026. Target perturbation can be induced by different types of interventions, including genetic alterations (e.g. gene knockout, knockdown, and overexpression) and chemical perturbations. Taking the chemical perturbations as an example, recorded attributes include the cell line, Cellosaurus ID, perturbation name, experimental condition, and so on. The lower panel presents the outcomes of target perturbation, featuring DEGs along with their symbols, gene IDs, statistical values, and links to TTD Target (where available). Visualization is provided by volcano and heatmap plots, illustrating up- and down-regulated genes as well as their expression patterns across samples.

as GDSC1, GDSC2, and CCLE, cytotoxic activity was aggregated by disease type to facilitate systematic comparison of drug response profiles across cancer types and subtypes. As shown in Fig. 3A, cytotoxic activities were presented in a tabular format, listing cell lines, disease, activity values, and so on. Each entry can be expanded to reveal detailed records, such as synonyms and links to Cellosaurus, CCLE, and DepMap. Complementary visualizations include combined boxplot–scatterplots, depicting activity distributions across cell types. For antimicrobial activities, the data table summarizes the tested microbe, activity values, and contextual metadata, while the taxonomy hierarchy is interactively presented to enable rapid exploration of phylogenetic relationships.

Molecular-level activity landscape of drugs

Molecular-level activity data for drugs in TTD were collected from BindingDB [44], ChEMBL [15], and PubChem [14]. Considering drug activity data are highly sensitive to experimental conditions, the drug activity data provided in TTD are accompanied by metadata specifying the experimental conditions under which they were obtained. Data standardization followed the procedures described above, with all activity measures (IC_{50}, Kd, Ki, etc.) normalized to nM units. The compiled activity dataset comprised IC_{50} (37.37%), Kd (30.35%), Ki (26.27%), EC_{50} (5.22%), and others (0.79%). The six molecule classes with the highest number of activity records—counting proteins in every class they belong to—were enzymes, membrane receptors, ion channels, transporters, transcription factors, and epigenetic regulators. All target molecules were cross-referenced with UniProt to facilitate the retrieval of protein annotations. As illustrated in Fig. 3B, for each drug, the molecular-level activity data were categorized by activity type. Each record includes molecule name, UniProt ID, TTD Target (if available), and experimental activity values.

All in all, the activity data for a total of 2443 approved, 2474 clinical/preclinical drugs, and 12 889 experimental agents were provided in this study. Specifically, cytotoxic activity data were compiled for 3221 drugs across 2367 disease-related cell lines, antimicrobial activity data were assembled for 1238 drugs against 803 microbial organisms, and molecular bioactivities were curated for 16 983 drugs targeting 5253 proteins.

Compilation of the clinical profiles of FDA-approved drugs

Clinical profiles of drugs play a pivotal role in the development of targeted therapeutics by providing valuable insights

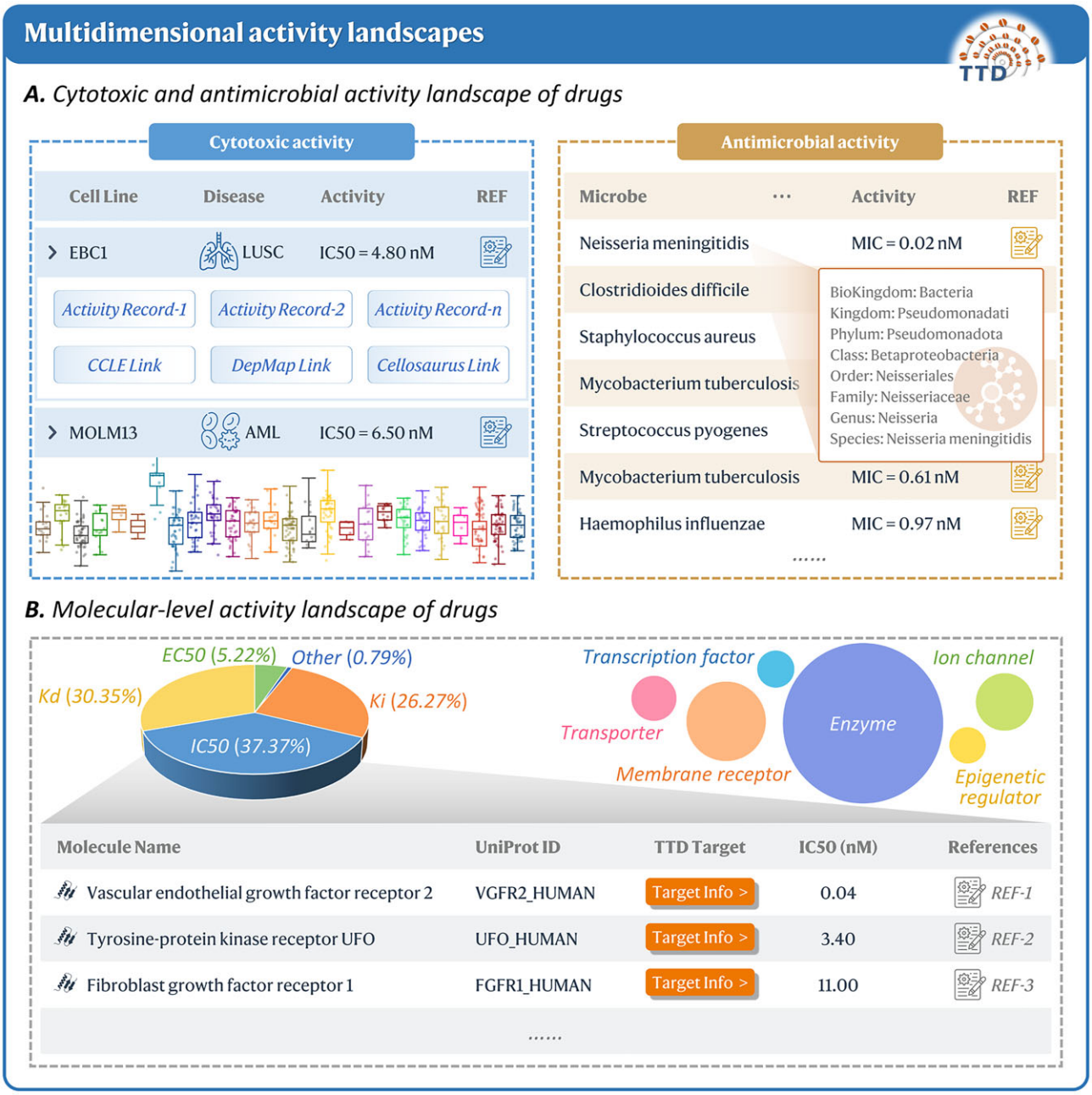


Figure 3. Multidimensional activity landscapes of drugs integrated in TTD 2026. (**A**) Cytotoxic and antimicrobial activity of drugs. The left panel summarizes cytotoxic activity data, such as cell line, disease, and activity values. Each entry can be expanded to access detailed information, including activity records and links to CCLE, DepMap, and Cellosaurus. A combined boxplot–scatterplot below the table illustrates the activity spectrum across cell types. The right panel displays the antimicrobial activity data, listing the microbe, activity values, and so on. (**B**) Molecular-level activity landscape of drugs. Bioactivity data are categorized by activity type, comprising IC_{50} (37.37%), Kd (30.35%), Ki (26.27%), EC_{50} (5.22%), and other (0.79%). Within each category, detailed records include molecule name, UniProt ID, TTD Target ID (if available), and activity values. The six molecular classes with the highest number of activity entries—counting proteins in every class they belong to—are enzyme, membrane receptor, ion channel, transporter, transcription factor, and epigenetic regulator.

that guide different stages of research and clinical trials [45]. Beyond knowledge of clinical endpoints, these profiles deliver up-to-date information on drug safety through pharmacovigilance [46], facilitate drug repurposing [47], and support advancements in precision medicine [48, 49]. An in-depth understanding of such data enables pharmaceutical researchers to optimize/deprioritize drug candidates [50], make informed decisions on indication expansion [51], define target patient populations [52], and address safety concerns [53].

In this update, a multistep workflow was implemented to systematically incorporate the clinical profiles of approved drugs into TTD. First, the complete and up-to-date collection of FDA-approved drug labels was retrieved in XML format from the FDA Online Label Repository. To ensure accurate and consistent integration, unique ingredient identifiers (UNIIs), along with their corresponding synonym dictionaries and chemical identifiers, were obtained from the FDA's Global Substance Registration System. For each drug label,

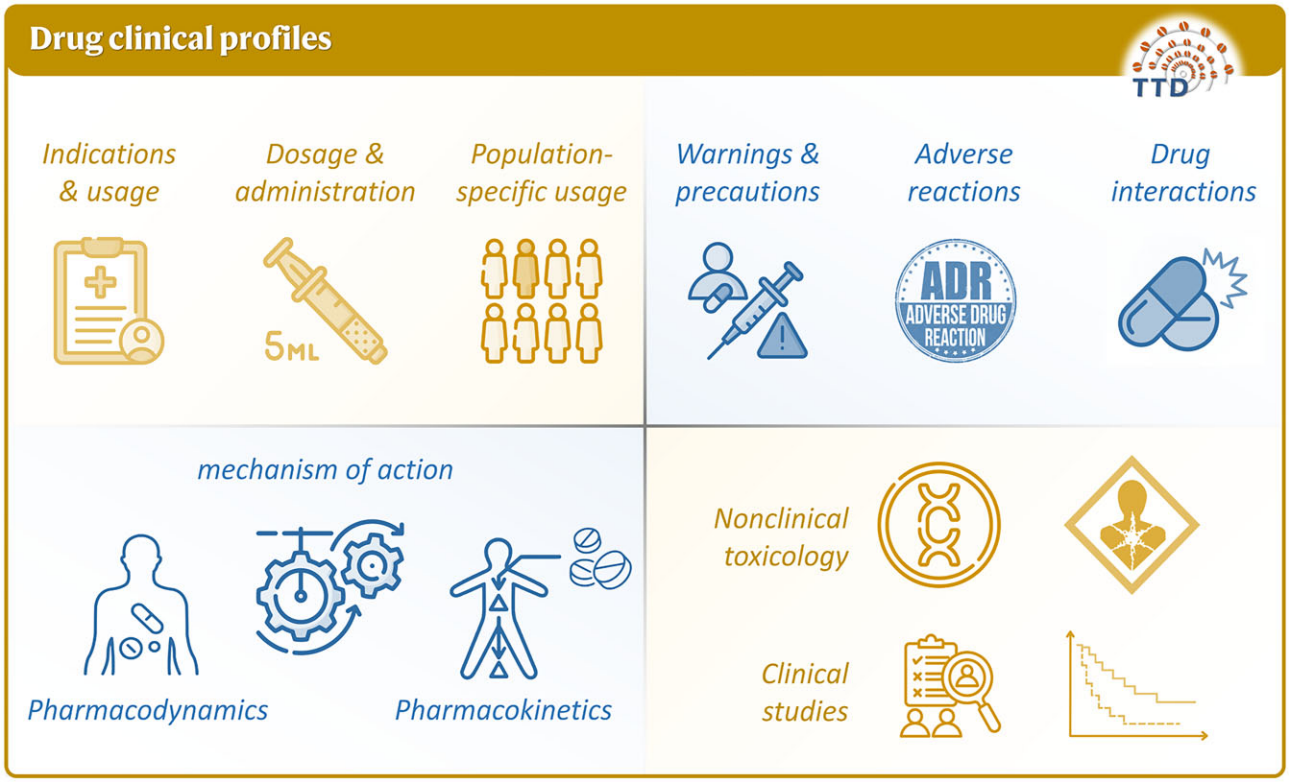


Figure 4. Overview of the newly incorporated clinical profiles for approved drugs in TTD 2026. For a specific approved drug, key clinical information includes indications and usage, dosage and administration, population-specific usage, warnings and precautions, drug interactions, adverse reaction, mechanism of action, pharmacodynamics, pharmacokinetics, nonclinical toxicology, and clinical studies (when available).

Table 1. Number of drugs and their corresponding therapeutic targets in different versions of TTD over the past decade

	Different versions of TTD published during the past decade					
	2026	**2024**	**2022**	**2020**	**2018**	**2016**
All targets	3798	3730	3578	3419	3101	2589
Successful targets	566	532	498	461	445	397
Clinical trial targets	1487	1442	1342	1191	1121	723
Preclinical/patented targets	232	239	185	155	0	0
Literature-reported targets	1513	1517	1553	1612	1535	1469
All drugs	40 398	39 862	38 760	37 102	34 019	31 614
Approved drugs	3019	2895	2797	2649	2544	2071
Clinical trial drugs	12 214	11 796	10 831	9465	8103	7291
Preclinical/patented agents	5040	5041	5009	4845	0	0
Experimental agents	20 125	20 130	20 123	20 143	18 923	17 803

the brand name, generic name, and active ingredient were extracted and mapped to the corresponding UNII. Using synonym matching and chemical identifier alignment, these entities were then linked to the TTD drug entries. As illustrated in Fig. 4, each profile integrates key clinical information, including indications and usage, dosage and administration, population-specific usage, warnings and precautions, adverse reactions, drug interactions, mechanism of action, pharmacodynamics, pharmacokinetics, nonclinical toxicology, and clinical study data. As a single drug can have multiple labels for different formulations (e.g. oral tablets, injectables) as well as for distinct brand names marketed by different manufacturers, 44 649 valid FDA drug labels were collected for 2234 approved drugs, comprising 2141 new drug applications, 331 biologics license applications, and 8889 abbreviated new drug applications.

Regular update and database reconstruction

The integration of newly emerged drugs and targets to TTD was also routinely conducted in this update. Approved drugs were systemically collected from Drugs@FDA, while new drugs in clinical trial were primarily sourced from ClinicalTrials.gov and company pipeline reports. Moreover, the trial status of drugs available in TTD were also updated based on the timely data provided in ClinicalTrials.gov. To facilitate cross-referencing with other biological resources, targets are standardized using both UniProt identifiers and HGNC symbols, small molecular drugs are mapped using PubChem CID, and diseases are annotated using the WHO ICD-11 classification system (if available). As summarized in Table 1, TTD 2026 includes a total of 3798 targets and 40 398 drugs, and the numbers of drugs and therapeutic targets in different versions of TTD over the past decade were also provided.

In this update, TTD database was reconstructed using a modernized architecture to enhance scalability, feature integration, and user accessibility. The backend is built on CentOS Linux, with Nginx as the web server and MySQL as the database management system, and is powered by Python using the Django framework. The frontend is developed with Vue3 and Element Plus, providing a responsive and user-friendly interface. Interactive visualization is enabled through ECharts, Plotly, and iCn3D plugins, offering advanced tools for data exploration. The platform is fully compatible with major browsers including Chrome, Edge, Firefox, and Safari.

Conclusions and future directions

Since its establishment in 2002, the TTD has undergone several major updates to broaden its content and enhance its functionality. The 2026 release marks a significant advancement, featuring extensive expansions in target–disease associations, target perturbation profiles, multidimensional drug activity landscapes, and clinical drug profiles. These enhancements are expected to provide a stronger foundation for mechanistic target discovery, targeted therapies development, and precision medicine. The complete reconstruction of the database architecture further improves usability, scalability, and data integration capacity, ensuring long-term accessibility and adaptability to the evolving needs of the biomedical research community.

Moving forward, there are some possible directions that our development effort can focus on. First, the implementation of AI and machine learning models can enable predictive analysis—for instance, forecasting drug–target interactions and inferring potential drug combinations. Second, the systematic inclusion of real-world clinical and pharmacogenomic data will enhance translational relevance and inform population-specific therapy optimization. Lastly, continued emphasis will be placed on user-centered design, refining the web interface and data visualization tools to accommodate increasingly diverse and complex user queries. With the rapid technological advances and collaborative efforts in drug discovery, the vast data accumulated in TTD and other databases will undoubtedly provide critical support for future target discovery and rational drug design.

Conflict of interest

None declared.

Funding

Funded by National Natural Science Funding of China (22220102001, 82373790, 62502424); National Key R&D Program of People's Republic of China (2024YFA1307503); Natural Science Foundations of Zhejiang (RG25H300001); Postdoctoral Fellowship Program of CPSF (GZC20252382); and Information Technology Center and State Key Lab of CAD&CG, Zhejiang University. Funding to pay the Open Access publication charges for this article was provided by National Natural Science Funding of China (22220102001).

Acknowledgement

Thanks for the supports of Information Technology Center and State Key Lab of CAD&CG, Zhejiang University.

Author contributions: Yintao Zhang (Data curation [equal], Methodology [equal], Software [lead], Visualization [lead], Writing—original draft [equal]), Ying Zhou (Data curation [equal], Methodology [equal], Writing—original draft [equal]), Hangwei Xu (Data curation [supporting]), Wanghao Jiang (Data curation [supporting]), Bo Li (Data curation [supporting]), Dianyu Lai (Data curation [supporting]), Cong Wan (Data curation [supporting]), Shanshan Wang (Data curation [supporting]), Mingxiao Zhao (Data curation [supporting]), Ying Tan (Data curation [supporting]), Songlin Lu (Data curation [supporting]), Tingting Fan (Data curation [supporting]), Xin Liu (Data curation [supporting]), Feng Zhu (Conceptualization [equal], Funding acquisition [equal], Supervision [equal], Writing—original draft [equal]), and Yu Zong Chen (Conceptualization [equal], Supervision [equal], Writing—original draft [equal]).

Data availability

TTD is freely accessible to all users without any login requirement at https://idrblab.org/ttd/.

References

1. Deng YT, You J, He Y *et al.* Atlas of the plasma proteome in health and disease in 53,026 adults. *Cell* 2025;188:253–71. https://doi.org/10.1016/j.cell.2024.10.045
2. Ferkingstad E, Sulem P, Atlason BA *et al.* Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet* 2021;53:1712–21. https://doi.org/10.1038/s41588-021-00978-w
3. Roohani Y, Huang K, Leskovec J. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nat Biotechnol* 2024;42:927–35. https://doi.org/10.1038/s41587-023-01905-6
4. Song B, Liu D, Dai W *et al.* Decoding heterogeneous single-cell perturbation responses. *Nat Cell Biol* 2025;27:493–504. https://doi.org/10.1038/s41556-025-01626-9
5. Fredin Haslum J, Lardeau CH, Karlsson J *et al.* Cell painting-based bioactivity prediction boosts high-throughput screening hit-rates and compound diversity. *Nat Commun* 2024;15:3470. https://doi.org/10.1038/s41467-024-47171-1
6. Gutierrez JG, Lau E, Dharmapalan S *et al.* Multi-output prediction of dose-response curves enables drug repositioning and biomarker discovery. *NPJ Precis Oncol* 2024;8:209. https://doi.org/10.1038/s41698-024-00691-x
7. Buniello A, Suveges D, Cruz-Castillo C *et al.* Open targets platform: facilitating therapeutic hypotheses building in drug discovery. *Nucleic Acids Res* 2025;53:D1467–75. https://doi.org/10.1093/nar/gkae1128
8. Elhussein A, Baymuradov U, Elhadad N *et al.* A framework for sharing of clinical and genetic data for precision medicine applications. *Nat Med* 2024;30:3578–89. https://doi.org/10.1038/s41591-024-03239-5
9. Racz A, Mihalovits LM, Beckers M *et al.* The changing landscape of medicinal chemistry optimization. *Nat Rev Drug Discov* 2025;24:870–7. https://doi.org/10.1038/s41573-41025-01225-41571
10. 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

11. Avram S, Bologa CG, Holmes J *et al.* DrugCentral 2021 supports drug discovery and repositioning. *Nucleic Acids Res* 2021;49:D1160–9. https://doi.org/10.1093/nar/gkaa997
12. Li F, Mou M, Li X *et al.* DrugMAP 2.0: molecular atlas and pharma-information of all drugs. *Nucleic Acids Res* 2025;53:D1372–82. https://doi.org/10.1093/nar/gkae791
13. Zhou Y, Zhang Y, Zhao D *et al.* TTD: therapeutic target database describing target druggability information. *Nucleic Acids Res* 2024;52:D1465–77. https://doi.org/10.1093/nar/gkad751
14. Kim S, Chen J, Cheng T *et al.* PubChem 2023 update. *Nucleic Acids Res* 2023;51:D1373–80. https://doi.org/10.1093/nar/gkac956
15. Zdrazil B, Felix E, Hunter F *et al.* The ChEMBL database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res* 2024;52:D1180–92. https://doi.org/10.1093/nar/gkad1004
16. Liu T, Hwang L, Burley SK *et al.* BindingDB in 2024: a FAIR knowledgebase of protein-small molecule binding data. *Nucleic Acids Res* 2025;53:D1633–44. https://doi.org/10.1093/nar/gkae1075
17. Namba S, Iwata M, Nureki SI *et al.* Therapeutic target prediction for orphan diseases integrating genome-wide and transcriptome-wide association studies. *Nat Commun* 2025;16:3355. https://doi.org/10.1038/s41467-025-58464-4
18. Morris JA, Caragine C, Daniloski Z *et al.* Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. *Science* 2023;380:eadh7699. https://doi.org/10.1126/science.adh7699
19. Savage SR, Yi X, Lei JT *et al.* Pan-cancer proteogenomics expands the landscape of therapeutic targets. *Cell* 2024;187:4389–407. https://doi.org/10.1016/j.cell.2024.05.039
20. Dagher OK, Schwab RD, Brookens SK *et al.* Advances in cancer immunotherapies. *Cell* 2023;186:1814. https://doi.org/10.1016/j.cell.2023.02.039e1811.
21. UniProt Consortium. UniProt: the universal protein knowledgebase in 2025. *Nucleic Acids Res* 2025;53:D609–17. https://doi.org/10.1093/nar/gkae1010
22. Landrum MJ, Chitipiralla S, Kaur K *et al.* ClinVar: updates to support classifications of both germline and somatic variants. *Nucleic Acids Res* 2025;53:D1313–21. https://doi.org/10.1093/nar/gkae1090
23. Phan L, Zhang H, Wang Q *et al.* The evolution of dbSNP: 25 years of impact in genomic research. *Nucleic Acids Res* 2025;53:D925–31. https://doi.org/10.1093/nar/gkae977
24. ClinGen Consortium. The clinical genome resource (ClinGen): advancing genomic knowledge through global curation. *Genet Med* 2025;27:101228. https://doi.org/10.1016/j.gim.2024.101228
25. Clough E, Barrett T, Wilhite SE *et al.* NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Res* 2024;52:D138–44. https://doi.org/10.1093/nar/gkad965
26. Gautier L, Cope L, Bolstad BM *et al.* affy—analysis of affymetrix GeneChip data at the probe level. *Bioinformatics* 2004;20:307–15. https://doi.org/10.1093/bioinformatics/btg405
27. Liu J, Lichtenberg T, Hoadley KA *et al.* An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 2018;173:400–16.e11. https://doi.org/10.1016/j.cell.2018.02.052
28. Goldman MJ, Craft B, Hastie M *et al.* Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* 2020;38:675–8. https://doi.org/10.1038/s41587-020-0546-8
29. Uhlen M, Zhang C, Lee S *et al.* A pathology atlas of the human cancer transcriptome. *Science* 2017;357:eaan2507. https://doi.org/10.1126/science.aan2507
30. Qian L, Sun R, Aebersold R *et al.* AI-empowered perturbation proteomics for complex biological systems. *Cell Genomics* 2024;4:100691. https://doi.org/10.1016/j.xgen.2024.100691
31. Wei Z, Si D, Duan B *et al.* PerturBase: a comprehensive database for single-cell perturbation data analysis and visualization. *Nucleic Acids Res* 2025;53:D1099–111. https://doi.org/10.1093/nar/gkae858
32. Qi X, Zhao L, Tian C *et al.* Predicting transcriptional responses to novel chemical perturbations using deep generative model for drug discovery. *Nat Commun* 2024;15:9256. https://doi.org/10.1038/s41467-024-53457-1
33. Subramanian A, Narayan R, Corsello SM *et al.* A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 2017;171:1437–52. https://doi.org/10.1016/j.cell.2017.10.049
34. Ritchie ME, Phipson B, Wu D *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47. https://doi.org/10.1093/nar/gkv007
35. Klaeger S, Heinzlmeir S, Wilhelm M *et al.* The target landscape of clinical kinase drugs. *Science* 2017;358:eaan4368. https://doi.org/10.1126/science.aan4368
36. Catalan P, Wood E, Blair JMA *et al.* Seeking patterns of antibiotic resistance in ATLAS, an open, raw MIC database with patient metadata. *Nat Commun* 2022;13:2917. https://doi.org/10.1038/s41467-022-30635-7
37. Zhao W, Li J, Chen MM *et al.* Large-scale characterization of drug responses of clinically relevant proteins in cancer cell lines. *Cancer Cell* 2020;38:829–43. https://doi.org/10.1016/j.ccell.2020.10.008
38. Kim MK, Chen Q, Echterhof A *et al.* A blueprint for broadly effective bacteriophage-antibiotic cocktails against bacterial infections. *Nat Commun* 2024;15:9987. https://doi.org/10.1038/s41467-024-53994-9
39. Chowdhury S, Zielinski DC, Dalldorf C *et al.* Empowering drug off-target discovery with metabolic and structural analysis. *Nat Commun* 2023;14:3390. https://doi.org/10.1038/s41467-023-38859-x
40. 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
41. Ghandi M, Huang FW, Jane-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
42. 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
43. Cox E, Tsuchiya MTN, Ciufo S *et al.* NCBI taxonomy: enhanced access via NCBI Datasets. *Nucleic Acids Res* 2025;53:D1711–5. https://doi.org/10.1093/nar/gkae967
44. Gilson MK, Liu T, Baitaluk M *et al.* BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res* 2016;44:D1045–53. https://doi.org/10.1093/nar/gkv1072
45. Fang H, Harris S, Liu Z *et al.* FDALabel for drug repurposing studies and beyond. *Nat Biotechnol* 2020;38:1378–9. https://doi.org/10.1038/s41587-020-00751-0
46. Suran M. FDA adds boxed warning to CAR T-cell therapies, but says benefits outweigh risks of secondary cancers. *JAMA* 2024;331:818–20. https://doi.org/10.1001/jama.2024.1011
47. Pushpakom S, Iorio F, Eyers PA *et al.* Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov* 2019;18:41–58. https://doi.org/10.1038/nrd.2018.168
48. Mehta D, Uber R, Ingle T *et al.* Study of pharmacogenomic information in FDA-approved drug labeling to facilitate application of precision medicine. *Drug Discov Today* 2020;25:813–20. https://doi.org/10.1016/j.drudis.2020.01.023
49. Bourgeois FT, Kesselheim AS. Promoting pediatric drug research and labeling—outcomes of legislation. *N Engl J Med* 2019;381:875–81. https://doi.org/10.1056/NEJMhle1901265
50. Wang S, Dong G, Sheng C. Structural simplification: an efficient strategy in lead optimization. *Acta Pharmaceutica Sinica B* 2019;9:880–901. https://doi.org/10.1016/j.apsb.2019.05.004

51. Miller KL, Lanthier M. Orphan drug label expansions: analysis of subsequent rare and common indication approvals. *Health Affairs* 2024;43:18–26. https://doi.org/10.1377/hlthaff.2023.00219
52. Tan WK, Ross-Innes CS, Somerset T *et al.* Biomarker risk stratification with capsule sponge in the surveillance of Barrett's oesophagus: prospective evaluation of UK real-world implementation. *Lancet* 2025;406:271–82. https://doi.org/10.1016/S0140-6736(25)01021-9
53. Beilmann M, Adkins K, Boonen HCM *et al.* Application of new approach methodologies for nonclinical safety assessment of drug candidates. *Nat Rev Drug Discov* 2025;24:705–25. https://doi.org/10.1038/s41573-025-01182-9

Received: September 14, 2025. **Revised:** October 9, 2025. **Accepted:** October 11, 2025