cancers

MDPI

Review

Applications of Genome-Wide Screening and Systems Biology Approaches in Drug Repositioning

Elyas Mohammadi [1,2], Rui Benfeitas [3], Hasan Turkez [4], Jan Boren [5], Jens Nielsen [6,7], Mathias Uhlen [1] and Adil Mardinoglu [1,8,*]

[1] Science for Life Laboratory, KTH–Royal Institute of Technology, SE-17121 Stockholm, Sweden; e.mohammadi@mail.um.ac.ir (E.M.); mathias.uhlen@scilifelab.se (M.U.)
[2] Department of Animal Science, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran
[3] National Bioinformatics Infrastructure Sweden (NBIS), Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, SE-10691 Stockholm, Sweden; rui.benfeitas@nbis.se
[4] Department of Medical Biology, Faculty of Medicine, Atatürk University, Erzurum 25240, Turkey; hturkez@atauni.edu.tr
[5] Department of Molecular and Clinical Medicine, University of Gothenburg, The Wallenberg Laboratory, Sahlgrenska University Hospital, SE-41345 Gothenburg, Sweden; Jan.Boren@wlab.gu.se
[6] Department of Biology and Biological Engineering, Chalmers University of Technology, SE-41296 Gothenburg, Sweden; nielsenj@chalmers.se
[7] BioInnovation Institute, DK-2200 Copenhagen N, Denmark
[8] Centre for Host-Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London, London SE1 9RT, UK
* Correspondence: adilm@scilifelab.se

Received: 30 July 2020; Accepted: 16 September 2020; Published: 21 September 2020

Simple Summary: Drug repurposing is an accelerated route for drug development and a promising approach for finding medications for orphan and common diseases. Here, we compiled databases that comprise both computationally- or experimentally-derived data, and categorized them based on quiddity and origin of data, further focusing on those that present high throughput omic data or drug screens. These databases were then contextualized with genome-wide screening methods such as CRISPR/Cas9 and RNA interference, as well as state of art systems biology approaches that enable systematic characterizations of multi-omic data to find new indications for approved drugs or those that reached the latest phases of clinical trials.

Abstract: Modern drug discovery through de novo drug discovery entails high financial costs, low success rates, and lengthy trial periods. Drug repositioning presents a suitable approach for overcoming these issues by re-evaluating biological targets and modes of action of approved drugs. Coupling high-throughput technologies with genome-wide essentiality screens, network analysis, genome-scale metabolic modeling, and machine learning techniques enables the proposal of new drug–target signatures and uncovers unanticipated modes of action for available drugs. Here, we discuss the current issues associated with drug repositioning in light of curated high-throughput multi-omic databases, genome-wide screening technologies, and their application in systems biology/medicine approaches.

Keywords: drug repositioning; genomic screens; machine learning; systems pharmacology; systems medicine

1. Introduction

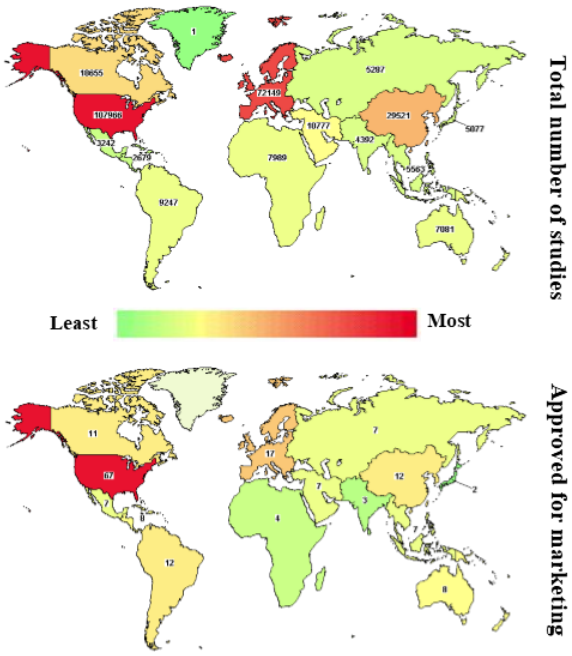
One of the first intentional and remarkable efforts for drug repositioning was carried out in 1987, when human immunodeficiency virus (HIV) became a full-blown pandemic with no known

treatments. Prior to that time, azidothymidine (AZT) had been under development as an anticancer drug and had failed during clinical trials due to the absence of efficacy. Surprisingly, it took just over two years to move from the initial demonstration of AZT's anti-HIV property to its FDA approval [1]. Empirical investigations of repositionable drugs have led to the rescue of the utility of many FDA-approved chemicals [2,3]. Empirical drug repositioning represents the process of finding an unknown target for a known drug in vitro. It has been proposed that empirical drug repositioning could be a starting point for the process of drug repurposing by the in vitro screening of known drugs or drug-like molecules, in order to identify and validate candidates for repositioning which result in the following advantages: The knowledge on a potential new disease setting will be increased; a serendipitous, or hypothesis-free, assessment of compounds will be achieved by testing multiple compounds with different modes of action; and data-driven choices will be produced for further investigations in more complex phenotypic or in vivo tests. As an example, an FDA-approved drug library of 640 compounds was screened and 10 of them, including tamoxifen and raloxifene, were proposed as compounds that are able to protect hair cell loss in response to known inner ear mechanosensory hair cell toxins, such as neomycin, gentamicin, kanamycin, and cisplatin, which cause hearing impairment and balance disorders [2]. In rational drug repurposing, the target is known from the beginning and subsequently, the goal is to find a previously FDA-approved compound which can interact with the target of interest [4]. In empirical drug repositioning, a large number of combinatorial compounds are prepared and tested. Rational drug repurposing reduces the number of candidates in the early stages of an investigation and identifies entities with a high probability of success [5]. Drug repurposing, also known as drug rescuing, repositioning, redirecting, tasking, recycling, or reprofiling, is a suitable approach for reusing existing drugs.

There are several reasons for the global need to repurpose currently approved drugs, instead of developing them de novo, including low success rates. At the beginning of 2020, there were 261,163 ongoing clinical trial studies (February 2020, www.clinicaltrials.gov) distributed all over the world (Figure 1a). In 2019, there were 318,901 ongoing clinical trials; while it is unclear how many of these clinical trials featured de novo drug design, only 48 novel drugs were approved for usage by the FDA (www.fda.gov). Taking this >95% failure rate into account, drug development is an extremely time- and resource-consuming process [3] insofar as the rate of success is almost 1 out of 10,000 (Figure 1b) [4].

a

Region Name	Total number of studies	Approved for marketing
World	261,163	131
Africa	7,989	4
Central America	2,679	8
East Asia	29,521	12
Japan	5,077	2
Europe	72,149	17
Middle East	10,777	7
North America	120,012	73
Canada	18,655	11
Mexico	3,242	7
United States	107,966	67
North Asia	5,287	7
Pacifica	7,081	8
South America	9,247	12
South Asia	4,392	3
Southeast Asia	5,563	7



b

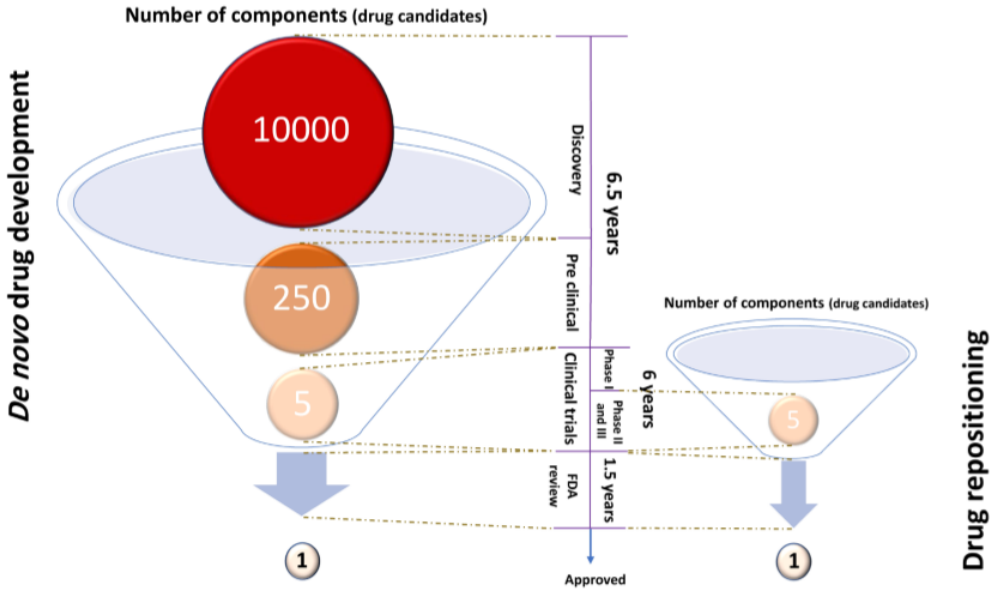


Figure 1. (**a**) Drug repositioning significantly cuts drug development costs and lengths of time. The distribution of the total number of conventional drug development studies and approved drugs for marketing is shown in the table and map. The USA, north North America, and Europe have the highest number of ongoing clinical trials [5]. (**b**) Two different methods for drug development are compared. The number of components (biological and/or chemical drug candidates), as input into de novo drug development, is 2000 times higher than drug repositioning, while drug repositioning may start in more advanced clinical trial phases and save several years when trying to find new treatments for emerging diseases.

On the other hand, whilst swift development for emerging diseases such as Covid-19 caused by SARS-Cov-2 is crucial, conventional procedures for drug development last more than 10 years [6,7].

FDA-approved drugs, or even those which have passed at least a phase I clinical trial (i.e., with proven human safety), could be considered as alternatives for rapid drug development. Considering that these drugs have passed at least a phase I clinical trial for one particular affliction, they could directly enter phase II or III of clinical trials for a second and unrelated indication. This path would take about two years, which is considerably shorter than the >14 years required for de novo drug development (Figure 1b), with an associated substantial cost reduction [3,8].

A single gene may be involved in two or more biological pathways or diseases and by targeting this gene, two unrelated diseases may be remedied. In addition, several mutations may disrupt the same biological pathway. In this case, all drugs which have similar effects on this pathway could be a candidate for repurposing studies [9]. For instance, trametinib is a mitogen-activated protein kinase kinase (MEK) pathway inhibitor, and is used for the treatment of leukemia and pancreatic cancer [9]. On the other hand, drug off-targeting could be another procedure employed for drug repositioning when the biological activity of a drug is different from its intended biological target. Despite being known as a common contributor to drug side effects, in some cases, off-target activity can be advantageous for therapeutic purposes. For example, the repurposing of the antimineralocorticoid and diuretic spironolactone, which is known to produce feminization and gynecomastia due to antiandrogen activity as side effects, could be used as an antiandrogen in the treatment of conditions like acne and hirsutism [10]. Additionally, sildenafil (Viagra®, Revatio®) [11] and minoxidil (Rogaine®) were originally developed for hypertension [12], but their off-target effects during trials were further explored in erectile dysfunction and promoted hair growth, respectively. Recent studies have aimed to make use of high-throughput datasets and computational approaches for the prediction of off-targets, in order to propose suitable repurposable drugs [13,14]. For instance, Rao et al. [13] developed the computational Off-Target Safety Assessment (OTSA) using more than 1 million compounds to identify potential off-target interactions that could be linked to predictable safety issues. In addition, Huang et al. [14] combined in vitro and in silico approaches to develop an integrative framework for the systematic identification of on-/off-target pathways and clarification of the underlying regulatory mechanisms.

Drug repositioning can be done based on prior knowledge or serendipity. This can be aided with the application of high-throughput screening, including CRISPR-Cas9 screening, RNAi Screening, and the application of systems medicine approaches for purposeful drug repositioning. One of the goals of genome-wide screening experiments is to generate and screen a population of genes to identify a particular phenotype or pathway and accordingly propose drug targets. The outcomes from screening studies could prepare the stage for systems medicine to integrate multi-omic data for accurate predictions of a drug's mode of action, targets, and disease associations.

2. High-Throughput Resources for the Identification of Drug Targets and Repositioning of Drugs

Precision medicine aims to develop efficient therapeutic strategies with fewer side effects matching the unique disease signatures in specific cohorts, thus making use of precision pharmacology [15]. However, the accurate identification of suitable targets and repurposable drugs relies on databases (DBs) of chemical and biological compounds and high-throughput omic datasets. Here, we classified DBs indicated as suitable resources for the identification of drug targets and revealing the chemical–phenotype association according to the quiddity and origin of data, as well as whether they rely on computational- or wet lab-based approaches.

One may distinguish five groups of DBs [16] and their content (Figure 2) as follows: i) Raw DBs, whose data are indirectly involved in drug repositioning and include literature curation, manual data uploading, data curation, integration, and clinical data; ii) Target-based DBs containing drug targets and related information, including information on pathways, enzymes, side effects, genes, and protein targets; iii) Specific DBs, which are those associated with specific software, disease data, and geographical information; iv) Drug Design DBs, which typically include small molecules, 3D molecular structures, and molecular replacement information; and v) Tool-based DBs, which embrace drug repositioning-related tools and include net-based tools (data mining through network

theory) and simple tools (other approaches for data mining). These DBs rely on many different types of biological and chemical data and present different content availabilities and assays (Table S1). Several DBs present particularly useful resources for data-driven analyses given their biological content, which often comprises high-throughput datasets (Table 1).

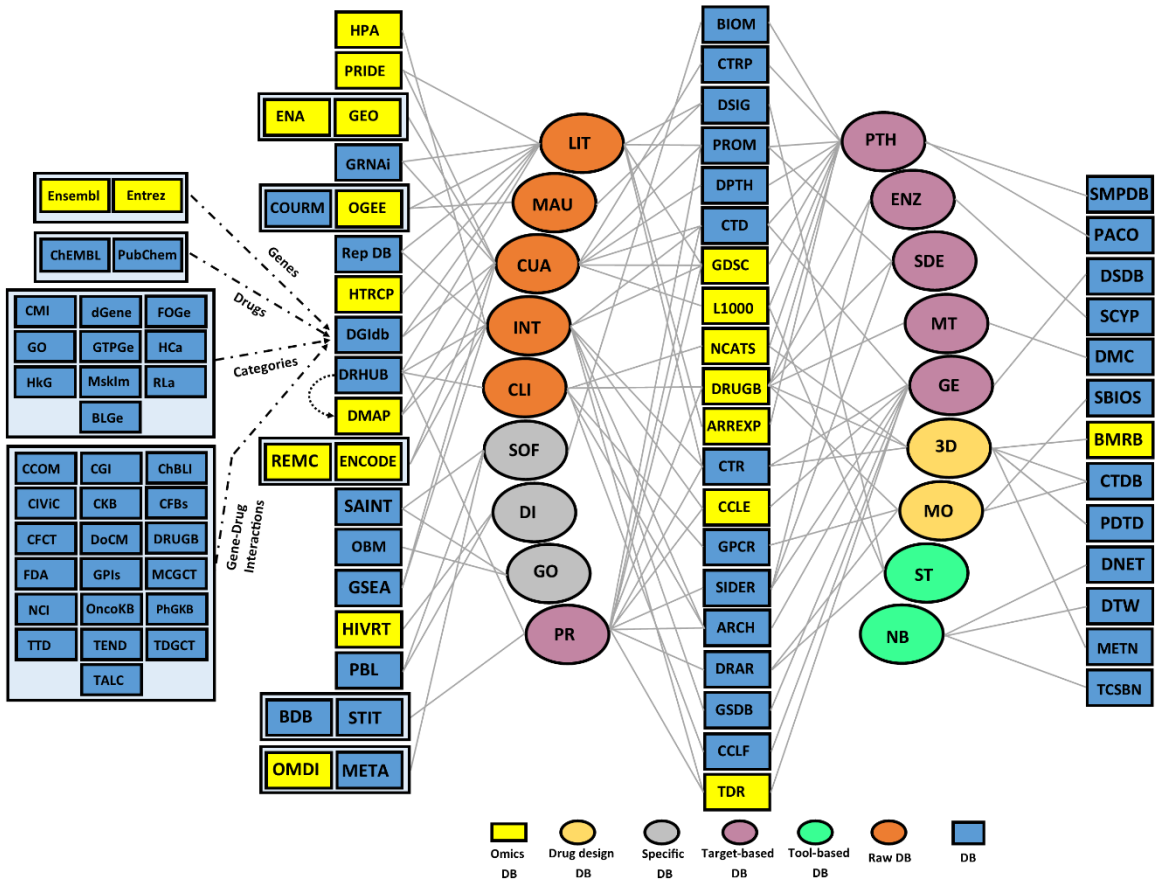


Figure 2. Key databases as primary resources for drug repositioning. Databases (DBs) can be divided into five classes, including Drug Design DB, Specific DB, Target-based DB, Tool-based DB, and Raw DB. The content of each DB is shown with an edge to the subcategories. Refer to Table S1 for a full description. DBs with omics and high-throughput datasets are colored in yellow. Subcategories: 3D: three-dimensional information; CLI: clinical; CUA: curation; DI: disease; ENZ: enzyme; GE: gene target; GO: geographic; INT: integration; Lit: literature; MAU: manual; MO: molecular information; MT: multidrug; NB: net-base; PR: protein target; PTH: pathway; SDE: side effect; ST: simple tool; TR: software. Databases: ACHI: Achilles; ARREXP: ArrayExpress; BLGe: BaderLabGenes; BIOM: Biomodels; BMRB: Biological Magnetic Resonance Data Bank; CCOM: Cancer Common; CCLF: cancer cell line factory; CCLE: Cancer Cell Line Encyclopedia; CFBs: Clearity Foundation Biomarkers; CFCT: Clearity Foundation Clinical Trial; CGI: Cancer Genome Interpreter; ChBLI: The ChEMBL Bioactivity Database; CIViC: Clinical Interpretations of Variants in Cancer; CKB: The Jackson Laboratory Clinical Knowledgebase; CTDB: Cheminformatic Tools and Databases; CMI: Caris Molecular Intelligence; CORUM: comprehensive resource of mammalian protein complexes; CTR: Cancer Therapeutics Response Portal; CTD: The Comparative Toxicogenomics Database; CTRP: Clinical Trials Reporting Program; dGene: The Druggable Gene List; DGIdb: drug–gene interaction database; DMAP: Dependency Map; DMC: Drug Map Central; DNET: Drug-Disease Network database; DoCM: Database of Curated Mutations; DPTH: Drug Pathway database; DRUGB: DrugBank; DRHUB: drug repurposing hub; DRAR: Drug Repurposing Adverse Reaction; DSDB: Drug Signatures Database; DSIG: DrugSig; DTW: Drug Target Web; ENCODE: Encyclopedia of DNA Elements, Entrez, Ensembl; FDA: FDA Pharmacogenomic Biomarkers; FOGe: Foundation One Gene; GEO: Gene Expression Omnibus; GPIs: Guide To Pharmacology Interactions; GPCR: G-protein-coupled receptors assay bank; GDSC: Genomics of Drug Sensitivity in Cancer; GO: The Gene

Ontology; GRNAi: GenomeRNAi; GSEA(MSigDB): Gene Set Enrichment Analysis; GSDB: Gene Set Database; GTPGe: Guide To Pharmacology Genes; HCa: Hingorani Casas; HIVRT: HIV Drug Resistance Database; HkG: Hopkins Groom; HTRCP: host transcriptional response connectivity map; L1000: Library of Integrated Network-based Cellular Signatures (LINCS) L1000; MCGCT: My Cancer Genome Clinical Trial; METN: Metlin; META: Metabolic Atlas; MskIm: Memorial Sloan Kettering IMPACT; NCI: Cancer Gene Index; NCATS: National Center for Advancing Translational Sciences; NNFIN: network-based similarity finder; ODB: Ontario database; OGEE: Online GEne Essentiality; OMDI: Omics Discovery Index (omicsdi); OncoKB: A Precision Oncology Knowledge Base; PACO: pathway commons; HPA: Protein Atlas; PBL: ProBiS-ligands; PDTD: Potential Drug-Target Database; PhGKB: The Pharmacogenomics Knowledgebase; PRIDE: PRoteomicsIDEntifications (PRIDE) database; PROM: Promiscuous; RepDB: Repurpose DB; RLa: RussLampel; SAINT: Significance Analysis of INTeractome; SBIOS: Swiss BIOisostere; SCYP: Super Cytochrome P450; SIDER: Side Effect Resource; SMPDB: Small Molecule Pathway Database; TALC: Targeted Agents in Lung Cancer; TCSBN: tissue- and cancer-specific biological networks; TDR: Tropical Diseases Research; TDGCT: The Druggable Genome (TDG) Clinical Trial; TEND: trends in the exploitation of novel drug targets; TTD: Therapeutic Target Database.

Table 1. DBs and their associated omic data as useful resources for drug repositioning.

Database	Type of Omics data	Link	Ref.
ARREXP	Functional genomics data	https://www.ebi.ac.uk/arrayexpress/	[17]
BMRB	Proteomics, metabolomics	http://www.bmrb.wisc.edu/	[18]
CCLE	Genomic data	https://portals.broadinstitute.org/ccle	[19]
DRUGB	Transcriptomics	https://www.drugbank.ca/	[20]
DMAP	Transcriptomics, proteomics	https://depmap.org/portal/	[21]
ENA	Genomics, transcriptomics	https://www.ebi.ac.uk/ena	[22]
ENCODE	Genomics, transcriptomics	https://www.encodeproject.org/	[23]
ENTREZ	Genomics, transcriptomics, proteomics, taxonomics	https://www.ncbi.nlm.nih.gov/Class/MLACourse/Original8Hour/Entrez/	[24]
ENSEMBL	Genomics, transcriptomics	https://www.ensembl.org/index.html	[25]
GDSC	Transcriptomics	https://www.cancerrxgene.org/	[26]
GEO	Functional genomics, transcriptomics, proteomics, phenomics	https://www.ncbi.nlm.nih.gov/geo/	[27]
HIVRT	Genomics, phenomics	https://hivdb.stanford.edu/	[28]
HTRCP	Transcriptomics	http://biotech.bmi.ac.cn/papers/2015/luhan.html	[29]
L1000	Transcriptomics	http://www.lincsproject.org/LINCS/tools/workflows/find-the-best-place-to-obtain-the-lincs-l1000-data	[30]
META	Metabolomics	https://metabolicatlas.org/	[31]
OMDI	Genomics, proteomics, transcriptomics and metabolomics	https://www.omicsdi.org/	[32]
HPA	Proteomics	https://www.proteinatlas.org/	[33]
PRIDE	Proteomics	https://www.ebi.ac.uk/pride/archive/	[34]
NCATS	Chemical genomics	https://ncats.nih.gov/	[35]
OGEE	Genomics, transcriptomics	http://ogee.medgenius.info/browse/	[36]
REMC	Epigenomics	http://www.roadmapepigenomics.org/	[37]
TDR	Genomics	https://tdrtargets.org/	[38]

3. Computational- and Wet Lab-Based Assay DBs

Among the assay-based DBs, one may consider two general groups of databases based on the main approaches they use to obtain data, including computational- and wet lab-based assays. Computational-based databases are those which only benefit from in silico approaches, whether

using primary assays such as data curation or advanced approaches, including network analysis. Wet lab-based datasets acquire data through wet lab assays.

3.1. Computational-Based Assay DBs

Various computational methods may be employed to acquire or collect biological data. For instance, text mining is a powerful technology for swiftly sifting key information through the vast quantities of biomedical literature. Text-mining techniques discover and extract new and unknown information from different resources, including databases, websites, books, reviews, and articles [39]. Manual data curation filters research results as they are generated, which means it unifies data for long-term sustainability and usage. In addition, this method monitors the reliability, reusability, and accessibility of obtained data [40]. Through integration, data from different sources may be processed and combined into a single, unified view [41,42].

Many datasets have followed one or a combination of these procedures, with or without experimental results, to design a comprehensive biological DB. For instance, GenomeRNAi (GRNAi) [43] is a manual extraction of phenotypes from RNAi screening in *Drosophila* and *Homo sapiens* literature, in addition to information about resources of RNAi reagents and their predicted quality. A dependency map (DMAP) [19] uses predictive modeling to collect and curate data from subdivisions of the DepMAP project, in order to propose gene–drug–disease associations. Gene Set Enrichment Analysis (GSEA)-MSigDB benefits from the gene set enrichment analysis method [42] and the Guide To Pharmacology Interactions (GPIs) [44], and Gene Set Database (GSDB) [45] benefit from data integration to present results.

In the case of methods used to retrieve drug–target associations, the Drug-Disease Network database (DNET) [46] uses differential co-expression analysis to obtain 1326 disease relationships among 108 diseases. Tissue- and cancer-specific biological networks (TCSBN) apply co-expression and integrated network analysis for 17 cancers and three tissues [33,47]. Additionally, Promiscuous (PROM) [48] organizes and curates data based on network concepts to predict new usages for existing drugs, considering the side effects. The Comparative Toxicogenomics Database (CTD) [49] provides manually curated information about chemical–gene/protein–disease associations, and the drug repurposing hub (DRHUB) [50] drug–gene interaction database (DGIdb) [51], DrugBank (DRUGB) [20], Small Molecule Pathway Database (SMPDB) [52], and Repurpose DB (RepDB) are comprehensive resources for gene targets, drugs and their categories, gene–drug interactions (DGIdb), and small molecule pathways (SMPDB). RepDB [53] is a compendium of drug targets, repurposed drugs, and their associated primary and secondary diseases. This database combines information on 253 drugs and 1125 diseases and using enrichment analysis, it determines key biological pathways, functional mechanisms, physicochemical features, and side effects associated with successfully repositioned drugs. These resources can help other researchers to design a better investigation for finding new targets for approved or even repositioned drugs.

Several repositories allow for manual and ununified dataset depositing, and include ArrayExpress [17] and Gene Expression Omnibus (GEO) [27], which contain functional genomics data from high-throughput functional genomics experiments. The Biological Magnetic Resonance Data Bank (BMRB) [18] contains experimental and derived data gathered from nuclear magnetic resonance (NMR) spectroscopic studies of biological molecules. BioModels [54] is a repository of mathematical models of biological and biomedical systems. In order to systematically analyze these repositories, some computational tools have aimed to integrate information for purposes that include proposing therapeutic targets. For instance, the Omics Discovery Index (OMDI) [32], as an open-source platform that enables access to and discovery and dissemination of omics data sets, may be used to integrate proteomics, genomics, metabolomics, and transcriptomics data sets from dozens of databases and repositories which have agreed on a common metadata structure framework and exchange format, and have contributed to OMDI. Similarly, the Metabolic Atlas [31] integrates open source genome-scale metabolic models. In addition, CTD utilizes various tools to discover the drug–target association through data provided by many datasets and repositories, including The

Pharmacogenomics Knowledgebase (PhGKB) [55], DRUGB [20], geographic (GO) [56], and NCBI [57].

3.2. Wet Lab-Based Assay DBs

DBs containing information which is created through experimental approaches can provide the information required for many other projects and computational assay-based DBs. In this case, three different purposes might be considered. In the first group, projects may encompass large amounts of ununified raw data (i.e., Big Data), but this does not enable their direct application in drug repositioning. The second group includes projects which benefit from a specific approach and directly create relevant information about drugs, gene targets, protein targets, or any other data which clearly pave the way for drug repositioning. Finally, the third group is comprised of projects that aim to develop experimental methods in order to escalate the efficiency of experimental approaches.

The value of data provided by the first group is undeniable and large amounts of time and many efforts have been devoted to developing them. However, whilst most of the provided data are not directly aimed at drug repositioning, they may nevertheless be mined by advanced computational approaches. Many consortiums have achieved strikingly crucial advances in coordinating these efforts. For instance, the goal of the ENCODE consortium [23] is to build a comprehensive list of functional elements in the human genome, including elements that act at the protein and RNA levels, and regulatory elements that control cells and circumstances in which a gene is active. Systematically mapped regions of transcription, transcription factor associations, the chromatin structure, and histone modification in ENCODE have enabled researchers to assign biochemical functions for 80% of previously unknown protein-coding genomic regions. REMC [37], which is a similar project to ENCODE, aimed to identify DNA methylation, histone modifications, chromatin accessibility, and small RNA transcripts in primary human tissues. Another interesting example in this group is the LINCS L1000 [58,59] project, which developed its own method to quantify transcriptomic data from treated and untreated cancer cell lines. L1000 is a cost-effective high-throughput method that can be used to estimate genome-wide mRNA expression on a large scale based on the direct measurement of a reduced representation of the transcriptome (978 landmark genes), which captures most of the information contained in the entire transcriptome [58] and computationally infers the remaining transcriptome [27].

Other DBs enable user queries to identify drug targets and repositionable drugs. For instance, the Achilles (ACHI) project, which is one of the subdivisions of DMAP and provides its essential data, systematically identifies and catalogs gene essentiality in genomically characterized cancer cell lines through genome-scale RNAi and CRISPR-Cas9 genetic perturbation reagents, in order to identify those genes that affect cell survival [21]. We discuss RNAi and CRISPR-Cas9 screens below. Identifying and targeting cancer dependencies with small molecules is the aim of the Cancer Therapeutics Response Portal (CTR) [60], which has accelerated the discovery of patient-matched cancer therapeutics. CTR has generated a set of 481 small-molecule probes and drugs that selectively target distinct nodes in cell circuitry and that collectively modulate a broad array of cell processes. In addition, the sensitivity of 860 deeply characterized cancer-cell lines toward Informer set compounds is quantitatively measured and connected to cancer features, including mutations, gene expression, copy-number variation, and lineages, and relevant results are freely available on the CTR database [61]. The Cancer Cell Line Encyclopedia (CCLE) provides public access to genomic data, analysis, and visualization for cancer cell lines [19]. It is one of the many large panels of comprehensively characterized human cancer models and helps to study genetic variants, candidate targets, and small-molecule and biological therapeutics. In addition, it may be used to identify new marker-driven cancer dependencies by examining genetic, RNA splicing, DNA methylation, histone H3 modification, microRNA expression, and reverse-phase protein array data for 1457 cell lines [19]. The Genomics of Drug Sensitivity in Cancer (GDSC) DB characterized many human cancer cell lines and screened them with hundreds of compounds, and accordingly, made drug response data and genomic markers of sensitivity available [26].

Among the projects that aim to escalate the experimental efficiency is the National Center for Advancing Translational Sciences (NCATS) consortium at NIH. The aim of this consortium is to develop, demonstrate, and disseminate tools and solutions to be used by all translational investigators. NCATS arms researchers with translational science information by making new methods, data, and information publicly available. For instance, the Assay Guidance Manual (AGM), as a procedure for data transparency and release in NCATS, resulted in the successful development of robust, early-stage drug discovery assays [62]. In addition, in the case of the availability of data, NCATS serves the pharmacological industry through PubChem [63] (small organic molecules and their activities against biological assays), the NCATS pharmaceutical collection [64] of active pharmaceutical ingredients and small molecules, and small molecule probes [65].

Another example of efforts aiming to escalate the experimental efficiency is the design and implementation of high-throughput G-protein-coupled receptor (GPCR) assays in GPCR Assay Bank that allow the cost-effective screening of large compound libraries, in order to identify novel drug candidates [66]. GPCRs are some of the most successful therapeutic target assays for a broad spectrum of diseases and mediate many important physiological functions. The GPCR Assay Bank is the largest readily accessible collection of approved drugs and has several subdivisions (https://web.duke.edu/gpcr-assay/my_screening.html), including Johns Hopkins Clinical Compound Library (JHCCL) (1514 compounds), KinaseGold Library (3893 compounds), PresTwick Library (1120 compounds), Sigma Kinase Library (97 compounds), Steroid Drug Library (1659 compounds), Tripos Library (50K compounds), TimTec NPL (Natural Product Library, 720 compounds), ActiProbe-5K (5K compounds), and Chemdiv Library (https://www.chemdiv.com/) (299,408 compounds).

4. Application of Genome-Wide Screening in Drug Repositioning

Functional genetic screening connects genes or genetic elements to phenotypes-of-interest. In this case, several methods have been gradually developed [67]. RNA interference (RNAi) oligonucleotides for loss-of-function studies [68] and cDNA overexpression libraries for gain-of-function studies [69] were performed prior to the clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas9 endonuclease system [67] and systems biology approaches [70]. All of these methods, especially CRISPR-Cas9, revealed a wealth of mechanistic insights, from drug resistance in cancer to neuronal toxicity in amyotrophic lateral sclerosis, and also made prerequisite information for drug repositioning by introducing drug targets available.

4.1. Drug Target Identification through RNA Interference Screening

RNA interference (RNAi) is a post-transcriptional gene regulating system which benefits from small noncoding RNAs and is utilized for the sequence-specific knockdown of a gene's function [4,71]. Small interfering RNAs (siRNAs), individually or as a pool of synthetic RNAs, are designed to specifically target unique messenger RNAs (mRNA) for a period of time or permanently [72]. High-throughput RNAi screening leads to the discovery of gene functions and its applications have been proven in many fields, such as infection, cancer, obesity, and aging.

RNAi can be introduced to the cell in various ways (Figure 3a), but all of them will end up with the siRNA duplex(es) entering the RNA-induced Silencing Complex (RISC) pathway. RISC can unwind and cleave the RNAi and realize the guide strand to pair with the target mRNA via perfect sequence complementarity. Afterwards, the RISC which is activated and specified by guide strand cleaves the target mRNA and accordingly, the protein production is abolished [73].

In Figure 3b, we provide a general workflow of siRNA screening towards drug repositioning. After transfection of the cell culture system with an siRNA library, a particular disease model (e.g., specific cancer) is induced in the cultured cells. Pre-determined endpoint assays are conducted at desired time points to yield hits (the genes which have an influence on disease) and lists of gene hits are compiled. Subsequently, these genes are validated in various conditions, such as different cell types, viruses, inductions, and different end-point assays. After the detection of a target gene, validation can be conducted by single siRNAs instead of siRNA pools. Eventually, these genes can be potentially targeted as a therapeutic disease intervention strategy [72].

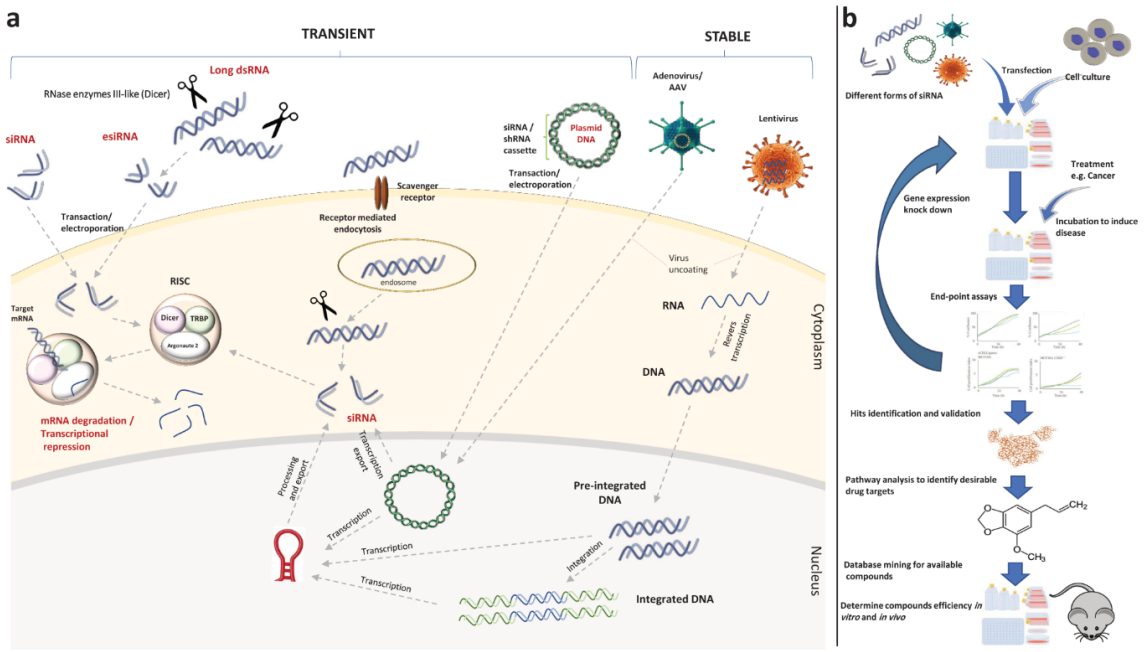


Figure 3. Various RNA interference (RNAi) constructs used for temporary or permanent gene silencing and a general workflow of small interfering RNA (siRNA) screening towards drug repositioning (adapted from [72]). (**a**) siRNA, endoribonuclease-prepared siRNA (esiRNA), and long double-strand RNA (dsRNA), as naked RNAs, can be introduced to the cell through many pathways, such as lipid-based transfection reagents or electroporation, or active uptake by target cells through receptor-mediated endocytosis. Inside the cytoplasm, dsRNA is cleaved by endonucleases. Then, the obtained siRNA, along with the directly introduced siRNAs and siRNAs coming from the nucleus of hard-to-transfect cells which were achieved from viral vectors carrying siRNA or short hairpin RNA (shRNA) expression cassettes, are loaded into the RNA-induced Silencing Complex (RISC) to facilitate gene expression knock down. (**b**) Drug targets and available compounds are determined by in silico pathway analyses and database mining after two-step in vitro screening. Prior to clinical analysis, the efficiency of these drug candidates can be validated by in vitro and in vivo approaches.

Among the successful experiments involving genome-wide RNAi screening is the application of a library of 74,905 retroviral shRNAs targeting 32,293 unique human transcripts to determine synthetic lethal interactions with the Kirsten rat sarcoma viral oncogene homolog (KRAS) oncogene, which is the isoform most commonly mutated in human cancers [74]. More recently, Takai et al. carried out a genome-wide RNAi through the GeneNet h50K shRNA library, which is composed of ~200,000 lentiviral shRNAs and more than 47,000 transcript targets, in order to identify genes with a synthetic lethal interaction with epithelial cellular adhesion molecule (EpCAM) as a potential therapeutic target for the EpCAM + AFP + Hepatocellular carcinoma subtype [75].

RNAi high-throughput screening, in spite of allowing for the swift discovery of the molecular basis of many diseases, and the identification of potential pathways for developing safe and effective treatments [72], presents issues associated with off-targeting [71].

4.2. Drug Target Identification through Genome-Wide CRISPR-Cas9 Screening

CRISPR-Cas9 genome-wide essentiality screens present suitable technologies for improving target prediction while minimizing off-targeting, thus improving on RNAi screens. Repurposing CRISPR as an RNA-guided platform for the sequence-specific control of gene expression was developed in 2013 [76]. Since then, many efforts have been made to conduct gene screening through this approach [9,77,78]. CRISPR-Cas9-based technology allows the discovery of efficient and precise new or repurposed drug candidates for genomic or proteomic targets. Systematic screenings of genes associated with a disorder can be conducted by combining the Cas9 endonuclease and a pooled guide RNA (gRNAs) library [75,79,80]. The gRNAs direct the Cas9 enzyme for DNA double strand

cleavage, while different kinds of mutations might be introduced by the error-prone Non-Homologous End Joining (NHEJ) DNA repair mechanism [81,82]. Therefore, it is possible to perturb many target genes simultaneously by using predesigned pooled gRNAs.

CRISPR-based gene activation (CRISPRa) and inhibition (CRISPRi), representing more powerful tools than RNA interference (RNAi) libraries, are highly useful in screening for gain and loss of function studies, respectively. Positive and negative selection using CRISPR libraries can clarify drug resistance mechanisms and survival-essential genes as drug targets [83]. CRISPR-Cas9 screening is currently employed to systematically investigate genes associated with lethal phenotypes and subsequently, propose drug targets at the gene and/or protein levels and accordingly identify likely repositionable drug candidates [84–87] (Figure 4).

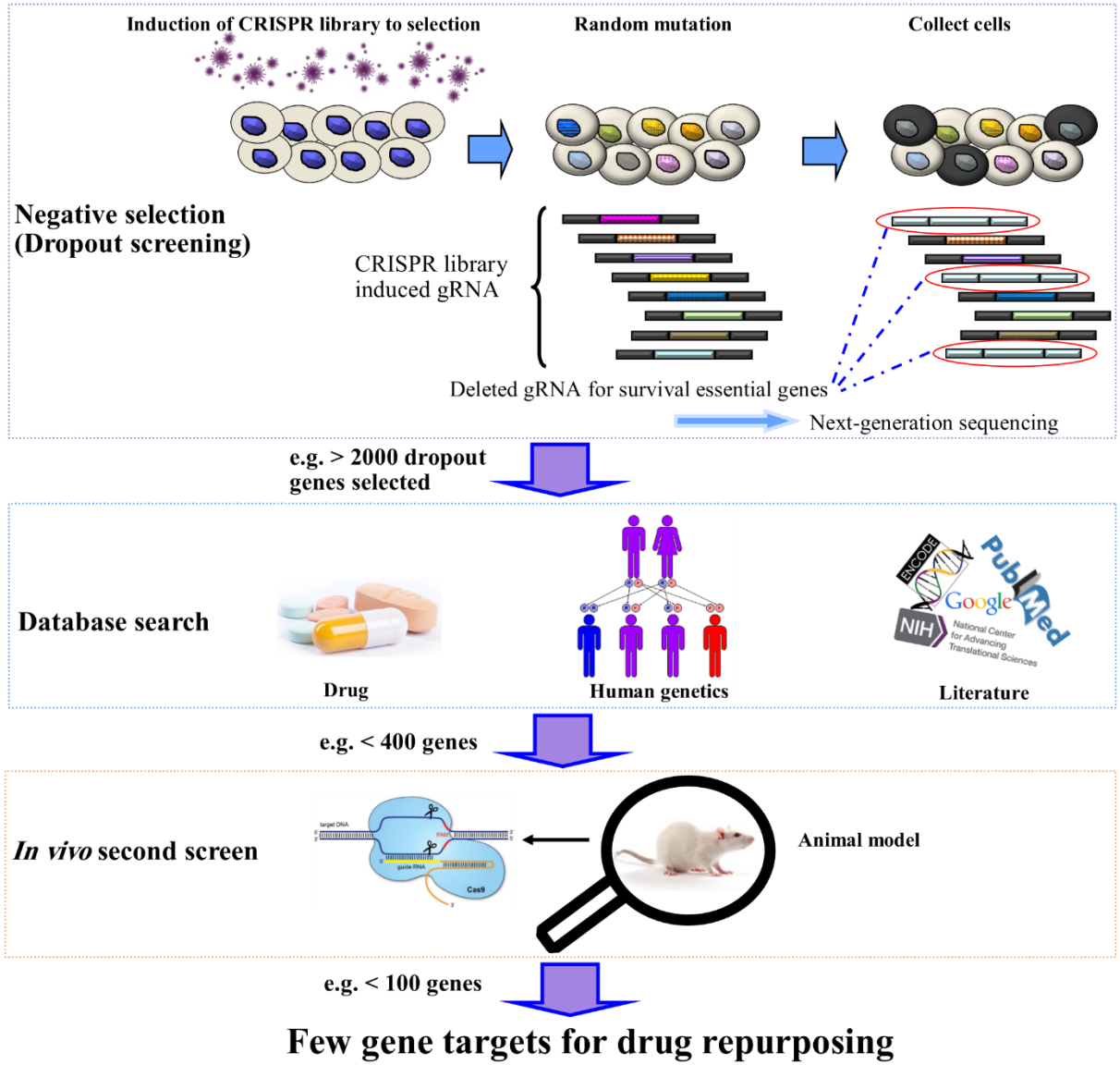


Figure 4. Pipeline of drug repositioning for a specific cancer or other disease. First, clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 dropout screening is applied for a negative selection of survival essential genes. Second, the number of dropped out selected genes is diminished by collecting information from databases and literature reviews. Third, in vivo CRISPR-Cas9 screening of an animal model of cancer or another disease will result in a few gene targets. Finally, FDA-approved drugs or those which have passed primitive phases of clinical trials could be tested as inhibitors for discovering essential survival genes.

Fused genes are two independent genes formed through structural rearrangements, the transcription of adjacent genes, or the splicing stage of pre-messenger RNA, which may lead to the dysfunctionality of modified genes and deregulation of associated genes, and afterward, the probable overexpression of oncogenes and/or decreased expression of tumor suppressor genes [88].

One of the first large-scale systematic analyses of thousands of gene fusions in human cancers analyzed 1034 samples from 1011 unique cancer cell lines to primarily define gene fusions [9]. Subsequently, whole-genome CRISPR-Cas9 drop-out screening was applied to identify gene fusions required for cancer cell fitness [84–87,89]. Similar techniques may be applied to identify repositionable anti-cancer drugs and propose new treatments by determining several fused genes as potential targets.

5. Systems Biology Application in Drug Repositioning

Wet lab high-throughput approaches examine all possible targets to find a repurposable drug [90]. In this regard, computer-assisted systematic drug repurposing methods may overcome the limitations associated with wet lab high-throughput techniques using focused resources based on knowledge of the gene/protein target and literature [91].

Personalized drug repurposing may now be employed systematically [90] by employing a systems medicine approach to handle the massive amounts of medical information which is accumulated in DBs [92]. Five major categories may be employed to classify in silico drug repurposing methods, namely network-based, machine learning, chemoinformatic-, bioinformatic-, and signature-based approaches [90].

5.1. Network-Based Methods in Drug Repositioning

Network-based methods integrate various high-throughput data to address issues in drug repositioning and drug combination. Specifically, they rely on known biological relationships of drug–gene–metabolite–protein–disease to generate networks, of which DisGeNET [93], the Therapeutic Target Database (TTD) [94], BindingDB [95], STRING [96], STITCH [97], and TCSBN [47] employ network-based approaches. Key studies in the field have highlighted how network-based methods may be employed for drug repositioning and drug combination. For instance, by quantifying the network proximity of disease genes and drug targets [98] and also drug targets and disease proteins [99] in the human (protein–protein) interactome, many novel drug–disease relationships for over 900 FDA-approved drugs and clinically efficacious drug combinations for specific diseases were identified. It should be noted that network-based methods are not alternatives, but complementary, to experimental approaches, capable of systematically capturing and characterizing coordinated cellular responses that would otherwise be technically challenging experimentally, further focusing clinical-phase attention on precise objectives [100]. To accomplish these tasks, network-based methods make use of various data, including phenotypic, drug-related, and molecular data, to construct various network types, such as Drug–Drug Interaction Networks, Drug–Target Interaction Networks, Drug–Disease Interaction Networks, and Multilayer Networks. By providing different biological characterizations, network-based methods are thus excellent frameworks in drug repositioning [100]. The idea behind network-based cluster approaches which can reveal the drug–disease or drug–target associations is that biological entities (e.g., diseases, drugs, and proteins) in the same module (also known as subnetworks or groups) share similar features in a biological network. Network-based approaches apply concepts from graph theory, where genes, drugs, diseases, or metabolites share biologically-demonstrated or putative associations. A module can consist of drug–disease, drug–drug, or drug–target relationships and this information can be extracted using the topological structure of a network. Density-based spatial clustering of applications with noise (DBSCAN) [101], Ordering Points To Identify the Clustering Structure (OPTICS), and the Statistical Information Grid (STING) [102] are some examples of the algorithms employed in network analyses. Importantly, in many cases, pleiotropic biological functions may occur; for instance, the proteins that are related to different cellular functions and consequently contribute to pleiotropic phenotypes when mutated [103]. Employing a *k*-means-based network

cluster algorithm, chemical–protein interactions were recently used to uncover repositionable drugs in small-cell lung cancer [104].

Network-based propagation methods may be divided into local (i.e., consider information from a subset of the network's nodes) [105] and global (i.e., derive information from the entire network) methods [106]. These approaches apply prerequisite information from the source node(s) to other network nodes [90,107]. The efficiency of these methods were proven for discovering disease–target, disease–gene, and disease–drug connections [108], and have been shown to provide a higher efficacy compared to other methods [106]. In addition, a comprehensive procedure based on formulating constraints on a score function related to the smoothness of the disease–gene network may lead to the discovery of disease–gene and disease–protein relationships [109]. This method could successfully predict gene targets for type 2 diabetes, Alzheimer's disease, and prostate cancer [90]. Additionally, methotrexate, gabapentin, cisplatin, donepezil, and risperidone could be repurposed for a second indication through a disease–gene–drug network propagation approach [110]. These observations indicate that drug repurposing conducted through network-based approaches presents an efficient and systematic approach for proposing repurposable drugs.

Computational methods may also derive information from genome-scale metabolic networks through the utilization of genome-scale metabolic models (GEMs). These are models that depict gene–protein–reaction relationships for entire metabolic genes in an organism, and can be simulated to predict metabolic fluxes for various systems-level metabolic studies [111]. For instance, a prostate cancer (PRAD)-specific GEM was reconstructed to investigate prostate cancer metabolism by the integration of global gene expression profiling of cell lines treated with more than 1000 different drugs [112]. As a result, among sulfamethoxypyridazine, azlocillin, hydroflumethiazide, and ifenprodil, as repositionable drugs proposed through an in silico cell viability assay, ifenprodil could be validated using an in vitro cell assay. In another study [113], GEMs used for drug design were assessed. In this regard, a list of human metabolites with KEGG [114] identifiers and their structures were retrieved using an updated human GEM [115]. Drug structures were compiled from DRUGB [20] and drug-metabolite pairs were analyzed. Finally, the effects of lipoamide analogs on MCF7 (a breast cancer cell line) and airway smooth muscle (ASM) cells were proposed and subsequently successfully demonstrated experimentally.

5.2. The Application of Machine Learning in Drug Repositioning

Machine-learning techniques have recently been employed to propose new drug candidates [116,117], where supervised approaches predict potential associations between approved drugs and health disorders [118]. For instance, the application of classification algorithms, including logistic regression, random forest, and support vector machine algorithms, has been conducted to construct prediction models and predict previously unknown pharmacological effects of different compounds [119]. In this case, in order to define the associations between drugs and diseases, four classes of drug-drug similarity (chemical structure, side-effects, gene ontology, and targets) and three categories of disease-disease similarity (phenotypes, human phenotype ontology, and gene ontology) were used. Finally, random forest methods exhibited the highest performance and could predict novel indications for 20 existing drugs and 31 compounds, which were subsequently validated using clinical trial data [119].

A swift escalation in the volume of biomedical literature has made it impossible to manually discover all meaningful connections between biomedical concepts, especially drug repositioning, which requires the integration of various aspects of biomedical knowledge. Hence, the automated text mining of literature applies natural language processing to transform unstructured text from various sources (databases, websites, books, reviews, and articles) into normalized, structured, and quantifiable data and their connections [120]. Co-occurrence methods reveal the association between two non-co-occurrence terms through discovering a third linking term that appears directly with each of them [121]. Recently, a comprehensive study was conducted to evaluate different drug repurposing strategies for Parkinson's disease by text mining the scientific literature through comparing various methods, including the extraction of biomedical entities and their relationships,

construction of a knowledge graph for Parkinson's disease, knowledge representation learning, and machine learning-based prediction [122]. As a result, unstructured biomedical literature data were effectively transformed to structured data that could be directly used by modern computational methods, such as machine learning [122]. Text-mining methods were recently applied, together with network analyses, to create disease-specific drug–protein connectivity maps [123]. These approaches extract disease–protein relationships from molecular interaction networks through network mining and investigate PubMed abstracts to predict novel identify drug–protein associations. In this regards, diltiazem and quinidine, as hypertension and arrhythmia drugs, respectively, can be repurposed for Alzheimer's disease, which has been confirmed by clinical evidence.

The application of deep learning as a promising approach for capturing complex and highly non-linear and heterogenous network structures has been actively explored. For instance [124], 10 drug–disease, one drug–side-effect, one drug–target, and seven drug–drug networks were integrated by a network-based deep-learning approach for in silico drug repurposing. Eventually, in addition to detecting an approved drug–disease association from the ClinicalTrials.gov database, repositionable drugs and novel drug targets were proposed for Alzheimer's disease (e.g., risperidone and aripiprazole) and Parkinson's disease (e.g., methylphenidate and pergolide).

5.3. Chemoinformatic-Based Methods in Drug Repositioning

Chemoinformatic approaches are one of the primary methods employed for screening repositionable drugs and create chemical libraries submitted to an in silico screen. These approaches are the basic components of many datasets, including DRUGB [20]. Using a similarity-based approach that correlates the anatomical therapeutic chemical classes of drug indications and their similarity, related therapeutic indications are proposed based on chemically similar drugs [125].

Ligand-based and structure-based approaches can be embedded in chemoinformatic-based drug repurposing. The story behind these two methods is that analogous compounds are likely to have similar biological properties [126] and proteins with almost the same structures tend to have similar functions and bind identical compounds [127]. However, the drugs identified through these two methods may share the same pathway with the template drug and accordingly, for a specific disease, if the template drug has already been used as a treatment, the role of these techniques becomes limited [128]. In addition, the lack of specificity is a serious drawback in this strategy. For instance, thalidomide has two chiral forms (the same chemical composition, but having mirrored structures). One can treat morning sickness, while the other form can have teratogen effects [129]. In this regard, molecular docking has been introduced as the most commonly used tool for structure-based virtual screening, in order to repurpose available drugs [130,131]. In this regard, Pinzi et al. repurposed cannabigerol, which is a non-psychoactive cannabinoid, as a low micromolar inhibitor of the enoyl acyl carrier protein reductase enzyme using an integrated ligand-based and structure-based study.

5.4. Bioinformatics-Based Methods in Drug Repositioning

From sequence alignment to domain similarity identification tools, a disease-centric approach is known as one of the principles of bioinformatical drug repositioning that tends to find diseases with the same dysfunctional proteins, which is often not clearly identified in experimental assays [90].

Local and global structural similarities are two different concepts; one can only be considered for similar binding sites of non-evolutionary proteins with variable folding and functions, and the other one is related to the whole structure of evolutionary proteins with considerable homology, respectively [132]. Ehrt et al. [127] claimed that the algorithms developed so far for binding site similarity assessments are not convincing since adequate benchmarking studies have not been developed. Accordingly, they suggested using a combination of in silico and experimental methods in a workflow to enhance the accuracy.

5.5. Signature-Based Drug Repositioning

The idea behind signature-based drug repositioning is that if the gene expression signature of a particular drug is opposite to the gene expression signature of a disease, that drug may have a potential therapeutic effect on the disease [133]. Drug signatures can be retrieved by comparing the gene expression profile of a cell line before and after its exposure to a specific drug. However, it should be considered that gene signatures are often used to address the off-targets of drugs in network pharmacology and these methods may not explain the relation between gene-expression alterations [134]. Some studies have tried to overcome this drawback through network-based analysis [133]. However, in complex disorders such as cancer [135], Alzheimer's disease [136], and inflammatory bowel disease, promising results were achieved through signature-based methods [137]. The Connectivity Map (cMap) [138] project was later joined with the LINCS L1000 project [30], which are two comprehensive and pioneer databases in the field of signature-based drug repurposing providing the gene expression profiles of dozens of cell lines treated with thousands of small molecules. It is obvious that dealing with such a large dataset would be challenging for computational systems biologists who aim to analyze and visualize Big Data. The cMap project presents a valuable resource for conducting large-scale assays of many small molecules. However, it is not without limitations. First, while comprehensive, it has tested few cell lines. Second, it presents limited drug perturbation data and does not provide a comprehensive view of drug dosages and temporal dynamics upon drug exposure. Third, because of this, its application in other cell lines and biological samples is very limited and has been demonstrated to not be very robust [139]. However, extensively examining a high number of small molecules, with different dosages and time points, is an extremely expensive and time-consuming process. Seeking to overcome these issues, others sought to devise systematic approaches to test a large number of samples and compounds, while aiming to also provide a systematic view of intracellular behavior (L1000). However, doing so is not without its perils, as previously mentioned with regards to the computationally heavy process [139]. Accordingly, a systematic compound signature discovery pipeline, called Enrichment of Gene Effect to a Molecule (EGEM), was developed by Liu et al. [140], which was derived from a rank-based gene set enrichment analysis and showed a better performance in comparison to the original. To date, several drug candidates have been proposed thanks to signature-based studies of cMap and L1000 data, including KM-00927 and BRD-K75081836 as novel histone deacetylase inhibitors and mitomycin C as a topoisomerase IIB inhibitor [141].

6. Conclusions and Future Direction

The integration of large assays is crucial for accurate predictions of potential therapeutic targets from high-throughput assays. DBs store both computationally- and laboratory-based created data and, subsequently, systems medicine makes use of these resources to decipher potential disease mechanisms, targetable genes and pathways, and prognostic or theranostic features of biomarkers (Figure 5). Omic data have been coupled with essentiality screens, including RNAi screening or the more accurate CRISPR-Cas9 screens.

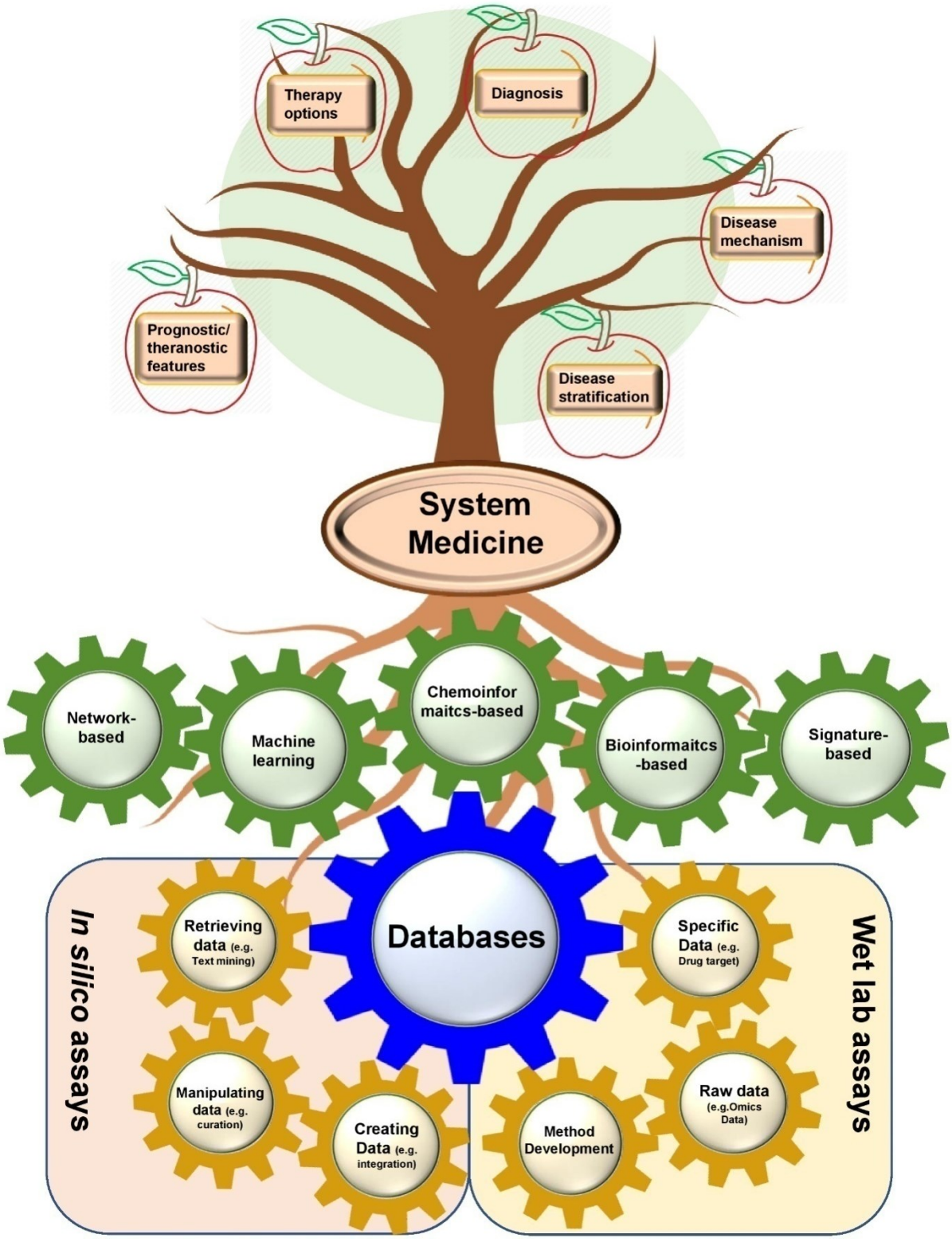


Figure 5. From bottom to the top, this is the whole story of current efforts to find repositionable drug candidates for emerging disease. DBs, as the foundation of drug repositioning, can be created from various computational and experimental approaches. The second way of categorizing DBs (assay-based classification) is depicted here as three in silico and three laboratory-based approaches in yellow gears to construct a DB. Afterwards, DBs can pave the way for various categorizations of systems medicine (shown in green gears) to decipher many probabilities in the case of disease mechanisms; disease stratification; therapy options; and diagnostic, prognostic, or theranostic features of biomarkers.

Several criteria are critical to facilitating DB usage in pharmaceutical science. First, one of the most important features of a database is that it has to be publicly available. Second, the availability of data after publication is another important criterion. Third, being updated in both data and interface aspects is crucial for drug discovery. Fourth, having an application programming interface (API) can help expert users to conveniently retrieve their required information, and many currently devise APIs to enable computational scientists to easily retrieve and analyze their data (e.g., LINCS L1000 and CheMBL). Fifth, data and metadata should be uniformly deposited and easily retrievable. Sixth, fully downloadable data are necessary for several experts to develop their own methods. LINCS L1000, ENCODE, TTD, and DSDB are instances of databases with data that are available to download. Seventh, internal tools such as the drug repositioning package of LINCS L1000 and SCYP can facilitate the use of databases. Eighth, advanced searches and queries increase the capabilities of databases, for instance, DrugDB and LINCS L1000 (Slicer, LINCS Canvas Browser, and L1000 Viewer). Researchers can make use of the feature list and characterization provided here (Tables 1 and S1, and Figure 2) as a starting point to identify suitable datasets and other information for their research purposes.

Previously FDA-approved chemicals as drugs may be repurposed for new targets and their mode of action and side effects can be investigated. Systematic drug repurposing requires a comprehensive collection of all these small-molecule drugs. There are many comprehensive databases encompassing various data types, and integrated perspectives such as those given by the DrugBank may facilitate drug repositioning. Systems biology and systems medicine may assist in this effort by applying machine learning and integrative approaches to predict repurposable drugs based on data from different sources. In this case, method and algorithm development seems necessary to escalate the efficiency of in silico screenings.

Supplementary Materials: The following are available online at www.mdpi.com/2072-6694/12/9/2694/s1: Table S1: An overview of all DBs discussed in this review in terms of the type of assay(s), data, and availability of content.

Funding: This research was funded by the Knut and Alice Wallenberg Foundation.

Conflicts of Interest: The authors declare no commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Fischl, M.A.; Richman, D.D.; Grieco, M.H.; Gottlieb, M.S.; Volberding, P.A.; Laskin, O.L.; Leedom, J.M.; Groopman, J.E.; Mildvan, D.; Schooley, R.T. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. *N. Engl. J. Med.* **1987**, *317*, 185–191.
2. Vlasits, A.L.; Simon, J.A.; Raible, D.W.; Rubel, E.W.; Owens, K.N. Screen of FDA-approved drug library reveals compounds that protect hair cells from aminoglycosides and cisplatin. *Hear. Res.* **2012**, *294*, 153–165.
3. DiMasi, J.A.; Hansen, R.W.; Grabowski, H.G. The price of innovation: New estimates of drug development costs. *J. Health Econ.* **2003**, *22*, 151–185.
4. Perwitasari, O.; Tripp, R.A. RNAi Screening to Facilitate Drug Repurposing. *Front. Rnai* **2014**, *19*, 247–265.
5. Clinicaltrials.gov. Available online: https://clinicaltrials.gov. (accessed on 20 February 2020).
6. Chong, C.R.; Sullivan, D.J. New uses for old drugs. *Nature* **2007**, *448*, 645–646.
7. Altay, O.; Mohammadi, E.; Lam, S.; Turkez, H.; Boren, J.; Nielsen, J.; Uhlen, M.; Mardinoglu, A. Current status of COVID-19 therapies and drug repositioning applications. *Iscience* **2020**, *23*, 101303.
8. Ekins, S.; Williams, A.J.; Krasowski, M.D.; Freundlich, J.S. In silico repositioning of approved drugs for rare and neglected diseases. *Drug Discov. Today* **2011**, *16*, 298–310.
9. Picco, G.; Chen, E.D.; Alonso, L.G.; Behan, F.M.; Gonçalves, E.; Bignell, G.; Matchan, A.; Fu, B.; Banerjee, R.; Anderson, E. Functional linkage of gene fusions to cancer cell fitness assessed by pharmacological and CRISPR-Cas9 screening. *Nat. Commun.* **2019**, *10*, 1–12.
10. Friedman, A.J. Spironolactone for adult female acne. *Cutis* **2015**, *96*, 216–217.

11. Terrett, N.K.; Bell, A.S.; Brown, D.; Ellis, P. Sildenafil (Viagra), a potent and selective inhibitor of type 5 cGMP phosphodiesterase with utility for the treatment of male erectile dysfunction. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1819–1824.
12. Goldenberg, M.M. Safety and efficacy of sildenafil citrate in the treatment of male erectile dysfunction. *Clin. Ther.* **1998**, *20*, 1033–1048.
13. Rao, M.; Gupta, R.; Liguori, M.; Hu, M.; Huang, X.; Mantena, S.; Mittelstadt, S.; Blomme, E.; Van Vleet, T. Novel computational approach to predict off-target interactions for small molecules. *Front. Big Data* **2019**, *2*, 25.
14. Huang, Y.; Furuno, M.; Arakawa, T.; Takizawa, S.; de Hoon, M.; Suzuki, H.; Arner, E. A framework for identification of on-and off-target transcriptional responses to drug treatment. *Sci. Rep.* **2019**, *9*, 1–9.
15. Lyman, G.H.; Moses, H.L. Biomarker Tests for Molecularly Targeted Therapies—The Key to Unlocking Precision Medicine. *N. Engl. J. Med.* **2016**, *375*, 4.
16. Masoudi-Sobhanzadeh, Y.; Omidi, Y.; Amanlou, M.; Masoudi-Nejad, A. Drug databases and their contributions to drug repurposing. *Genomics* **2019**, *112*, 1087–1095, doi:10.1016/j.ygeno.2019.06.021.
17. Kolesnikov, N.; Hastings, E.; Keays, M.; Melnichuk, O.; Tang, Y.A.; Williams, E.; Dylag, M.; Kurbatova, N.; Brandizi, M.; Burdett, T. ArrayExpress update—Simplifying data submissions. *Nucleic Acids Res.* **2015**, *43*, D1113–CD1116.
18. Ulrich, E.L.; Akutsu, H.; Doreleijers, J.F.; Harano, Y.; Ioannidis, Y.E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z. BioMagResBank. *Nucleic Acids Res.* **2007**, *36*, D402–CD408.
19. Ghandi, M.; Huang, F.W.; Jané-Valbuena, J.; Kryukov, G.V.; Lo, C.C.; McDonald, E.R.; Barretina, J.; Gelfand, E.T.; Bielski, C.M.; Li, H. Next-generation characterization of the cancer cell line encyclopedia. *Nature* **2019**, *569*, 503–508.
20. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082.
21. Tsherniak, A.; Vazquez, F.; Montgomery, P.G.; Weir, B.A.; Kryukov, G.; Cowley, G.S.; Gill, S.; Harrington, W.F.; Pantel, S.; Krill-Burger, J.M. Defining a cancer dependency map. *Cell* **2017**, *170*, 564–576.e516.
22. Harrison, P.W.; Alako, B.; Amid, C.; Cerdeño-Tárraga, A.; Cleland, I.; Holt, S.; Hussein, A.; Jayathilaka, S.; Kay, S.; Keane, T. The European nucleotide archive in 2018. *Nucleic Acids Res.* **2019**, *47*, D84–D88.
23. Dunham, I.; Birney, E.; Lajoie, B.R.; Sanyal, A.; Dong, X.; Greven, M.; Lin, X.; Wang, J.; Whitfield, T.W.; Zhuang, J. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.
24. Maglott, D.; Ostell, J.; Pruitt, K.D.; Tatusova, T. Entrez Gene: Gene-centered information at NCBI. *Nucleic Acids Res.* **2010**, *39*, D52–D57.
25. Zerbino, D.R.; Achuthan, P.; Akanni, W.; Amode, M.R.; Barrell, D.; Bhai, J.; Billis, K.; Cummins, C.; Gall, A.; Girón, C.G. Ensembl 2018. *Nucleic Acids Res.* **2018**, *46*, D754–D761.
26. Yang, W.; Soares, J.; Greninger, P.; Edelman, E.J.; Lightfoot, H.; Forbes, S.; Bindal, N.; Beare, D.; Smith, J.A.; Thompson, I.R. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **2012**, *41*, D955–D961.
27. Barrett, T.; Troup, D.B.; Wilhite, S.E.; Ledoux, P.; Rudnev, D.; Evangelista, C.; Kim, I.F.; Soboleva, A.; Tomashevsky, M.; Edgar, R. NCBI GEO: Mining tens of millions of expression profiles—Database and tools update. *Nucleic Acids Res.* **2007**, *35*, D760–D765.
28. Tzou, P.L.; Descamps, D.; Rhee, S.-Y.; Raugi, D.N.; Charpentier, C.; Taveira, N.; Smith, R.A.; Soriano, V.; de Mendoza, C.; Holmes, S.P. Expanded Spectrum of Antiretroviral-Selected Mutations in Human Immunodeficiency Virus Type 2. *J. Infect. Dis.* **2020**, *221*, 1962–1972.
29. Han, L.; He, H.; Li, F.; Cui, X.; Xie, D.; Liu, Y.; Zheng, X.; Bai, H.; Wang, S.; Bo, X. Inferring infection patterns based on a connectivity map of host transcriptional responses. *Sci. Rep.* **2015**, *5*, 15820.
30. Subramanian, A.; Narayan, R.; Corsello, S.M.; Peck, D.D.; Natoli, T.E.; Lu, X.; Gould, J.; Davis, J.F.; Tubelli, A.A.; Asiedu, J.K. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **2017**, *171*, 1437–1452.e1417.
31. Robinson, J.L.; Kocabaş, P.; Wang, H.; Cholley, P.-E.; Cook, D.; Nilsson, A.; Anton, M.; Ferreira, R.; Domenzain, I.; Billa, V., et al. An atlas of human metabolism. *Sci. Signal.* **2020**, *13*, eaaz1482, doi:10.1126/scisignal.aaz1482.

32. Perez-Riverol, Y.; Bai, M.; da Veiga Leprevost, F.; Squizzato, S.; Park, Y.M.; Haug, K.; Carroll, A.J.; Spalding, D.; Paschall, J.; Wang, M. Discovering and linking public omics data sets using the Omics Discovery Index. *Nat. Biotechnol.* **2017**, *35*, 406–409.
33. Uhlen, M.; Zhang, C.; Lee, S.; Sjöstedt, E.; Fagerberg, L.; Bidkhori, G.; Benfeitas, R.; Arif, M.; Liu, Z.; Edfors, F. A pathology atlas of the human cancer transcriptome. *Science* **2017**, *357*, eaan2507.
34. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D.J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **2019**, *47*, D442–D450.
35. Huang, R.; Grishagin, I.; Wang, Y.; Zhao, T.; Greene, J.; Obenauer, J.C.; Ngan, D.; Nguyen, D.-T.; Guha, R.; Jadhav, A. The NCATS BioPlanet–an integrated platform for exploring the universe of cellular signaling pathways for toxicology, systems biology, and chemical genomics. *Front. Pharmacol.* **2019**, *10*, 445.
36. Chen, W.-H.; Lu, G.; Chen, X.; Zhao, X.-M.; Bork, P. OGEE v2: An update of the online gene essentiality database with special focus on differentially essential genes in human cancer cell lines. *Nucleic Acids Res.* **2016**, gkw1013, doi:10.1093/nar/gkw1013.
37. Polak, P.; Karlić, R.; Koren, A.; Thurman, R.; Sandstrom, R.; Lawrence, M.S.; Reynolds, A.; Rynes, E.; Vlahovičeк, K.; Stamatoyannopoulos, J.A. Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **2015**, *518*, 360–364.
38. Urán Landaburu, L.; Berenstein, A.J.; Videla, S.; Maru, P.; Shanmugam, D.; Chernomoretz, A.; Agüero, F. TDR Targets 6: Driving drug discovery for human pathogens through intensive chemogenomic data integration. *Nucleic Acids Res.* **2020**, *48*, D992–D1005.
39. Allahyari, M.; Pouriyeh, S.; Assefi, M.; Safaei, S.; Trippe, E.D.; Gutierrez, J.B.; Kochut, K. A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv* **2017**, arXiv:1707.02919.
40. Odell, S.G.; Lazo, G.R.; Woodhouse, M.R.; Hane, D.L.; Sen, T.Z. The art of curation at a biological database: Principles and application. *Curr. Plant Biol.* **2017**, *11*, 2–11.
41. Gligorijević, V.; Pržulj, N. Methods for biological data integration: Perspectives and challenges. *J. R. Soc. Interface* **2015**, *12*, 20150571.
42. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550.
43. Schmidt, E.E.; Pelz, O.; Buhlmann, S.; Kerr, G.; Horn, T.; Boutros, M. GenomeRNAi: A database for cell-based and in vivo RNAi phenotypes, 2013 update. *Nucleic Acids Res.* **2013**, *41*, D1021–D1026.
44. Pawson, A.J.; Sharman, J.L.; Benson, H.E.; Faccenda, E.; Alexander, S.P.; Buneman, O.P.; Davenport, A.P.; McGrath, J.C.; Peters, J.A.; Southan, C. The IUPHAR/BPS Guide to PHARMACOLOGY: An expert-driven knowledgebase of drug targets and their ligands. *Nucleic Acids Res.* **2014**, *42*, D1098–D1106.
45. Araki, H.; Knapp, C.; Tsai, P.; Print, C. GeneSetDB: A comprehensive meta-database, statistical and visualisation framework for gene set analysis. *Febs Open Bio* **2012**, *2*, 76–82.
46. Yang, J.; Wu, S.-J.; Yang, S.-Y.; Peng, J.-W.; Wang, S.-N.; Wang, F.-Y.; Song, Y.-X.; Qi, T.; Li, Y.-X.; Li, Y.-Y. DNetDB: The human disease network database based on dysfunctional regulation mechanism. *Bmc Syst. Biol.* **2016**, *10*, 36.
47. Lee, S.; Zhang, C.; Arif, M.; Liu, Z.; Benfeitas, R.; Bidkhori, G.; Deshmukh, S.; Al Shobky, M.; Lovric, A.; Boren, J. TCSBN: A database of tissue and cancer specific biological networks. *Nucleic Acids Res.* **2018**, *46*, D595–D600.
48. Von Eichborn, J.; Murgueitio, M.S.; Dunkel, M.; Koerner, S.; Bourne, P.E.; Preissner, R. PROMISCUOUS: A database for network-based drug-repositioning. *Nucleic Acids Res.* **2010**, *39*, D1060–D1066.
49. Davis, A.P.; Murphy, C.G.; Johnson, R.; Lay, J.M.; Lennon-Hopkins, K.; Saraceni-Richards, C.; Sciaky, D.; King, B.L.; Rosenstein, M.C.; Wiegers, T.C. The comparative toxicogenomics database: Update 2013. *Nucleic Acids Res.* **2013**, *41*, D1104–D1114.
50. Corsello, S.M.; Bittker, J.A.; Liu, Z.; Gould, J.; McCarren, P.; Hirschman, J.E.; Johnston, S.E.; Vrcic, A.; Wong, B.; Khan, M. The Drug Repurposing Hub: A next-generation drug library and information resource. *Nat. Med.* **2017**, *23*, 405–408.
51. Cotto, K.C.; Wagner, A.H.; Feng, Y.-Y.; Kiwala, S.; Coffman, A.C.; Spies, G.; Wollam, A.; Spies, N.C.; Griffith, O.L.; Griffith, M. DGIdb 3.0: A redesign and expansion of the drug–gene interaction database. *Nucleic Acids Res.* **2018**, *46*, D1068–D1073.

52. Frolkis, A.; Knox, C.; Lim, E.; Jewison, T.; Law, V.; Hau, D.D.; Liu, P.; Gautam, B.; Ly, S.; Guo, A.C. SMPDB: The small molecule pathway database. *Nucleic Acids Res.* **2010**, *38*, D480–D487.
53. Shameer, K.; Glicksberg, B.S.; Hodos, R.; Johnson, K.W.; Badgeley, M.A.; Readhead, B.; Tomlinson, M.S.; O'Connor, T.; Miotto, R.; Kidd, B.A. Systematic analyses of drugs and disease indications in RepurposeDB reveal pharmacological, biological and epidemiological factors influencing drug repositioning. *Brief. Bioinform.* **2018**, *19*, 656–678.
54. Chelliah, V.; Juty, N.; Ajmera, I.; Ali, R.; Dumousseau, M.; Glont, M.; Hucka, M.; Jalowicki, G.; Keating, S.; Knight-Schrijver, V., et al. BioModels: Ten-year anniversary. *Nucleic Acids Res.* **2014**, *43*, D542–D548, doi:10.1093/nar/gku1181.
55. McDonagh, E.M.; Whirl-Carrillo, M.; Garten, Y.; Altman, R.B.; Klein, T.E. From pharmacogenomic knowledge acquisition to clinical applications: The PharmGKB as a clinical pharmacogenomic biomarker resource. *Biomark. Med.* **2011**, *5*, 795–806.
56. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29.
57. Coordinators, N.R. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **2017**, *45*, D12.
58. Duan, Q.; Flynn, C.; Niepel, M.; Hafner, M.; Muhlich, J.L.; Fernandez, N.F.; Rouillard, A.D.; Tan, C.M.; Chen, E.Y.; Golub, T.R. LINCS Canvas Browser: Interactive web app to query, browse and interrogate LINCS L1000 gene expression signatures. *Nucleic Acids Res.* **2014**, *42*, W449–W460.
59. Han, H.-W.; Hahn, S.; Jeong, H.Y.; Jee, J.-H.; Nam, M.-O.; Kim, H.K.; Lee, D.H.; Lee, S.-Y.; Choi, D.K.; Yu, J.H. LINCS L1000 dataset-based repositioning of CGP-60474 as a highly potent anti-endotoxemic agent. *Sci. Rep.* **2018**, *8*, 1–9.
60. Aksoy, B.A.; Dančík, V.; Smith, K.; Mazerik, J.N.; Ji, Z.; Gross, B.; Nikolova, O.; Jaber, N.; Califano, A.; Schreiber, S.L. CTD2 Dashboard: A searchable web interface to connect validated results from the Cancer Target Discovery and Development Network. *Database* **2017**, *2017*, bax054.
61. Cancer Therapeutics Response Portal: A CTD² Network Resource for Mining Candidate Cancer Dependencies. Available online: https://ocg.cancer.gov/e-newsletter-issue/issue-11/cancer-therapeutics-response-portal-ctd%C2%B2-network (Accessed on 21 September 2020).
62. Coussens, N.P.; Sittampalam, G.S.; Guha, R.; Brimacombe, K.; Grossman, A.; Chung, T.D.; Weidner, J.R.; Riss, T.; Trask, O.J.; Auld, D. Assay guidance manual: Quantitative biology and pharmacology in preclinical drug discovery. *Clin. Transl. Sci.* **2018**, *11*, 461–470.
63. Kim, S.; Thiessen, P.A.; Bolton, E.E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B.A. PubChem substance and compound databases. *Nucleic Acids Res.* **2016**, *44*, D1202–D1213.
64. Huang, R.; Zhu, H.; Shinn, P.; Ngan, D.; Ye, L.; Thakur, A.; Grewal, G.; Zhao, T.; Southall, N.; Hall, M.D. The NCATS Pharmaceutical Collection: A 10-year update. *Drug Discov. Today* **2019**, *24*, 2341–2349.
65. Probe Reports from the NIH Molecular Libraries Program, National Center for Biotechnology Information (US) Available oline: https://www.ncbi.nlm.nih.gov/books/NBK47352/ (accessed on 21 September 2020).
66. Zhang, R.; Xie, X. Tools for GPCR drug discovery. *Acta Pharmacol. Sin.* **2012**, *33*, 372–384.
67. So, R.W.; Chung, S.W.; Lau, H.H.; Watts, J.J.; Gaudette, E.; Al-Azzawi, Z.A.; Bishay, J.; Lin, L.T.-W.; Joung, J.; Wang, X. Application of CRISPR genetic screens to investigate neurological diseases. *Mol. Neurodegener.* **2019**, *14*, 41.
68. Johannessen, C.M.; Johnson, L.A.; Piccioni, F.; Townes, A.; Frederick, D.T.; Donahue, M.K.; Narayan, R.; Flaherty, K.T.; Wargo, J.A.; Root, D.E. A melanocyte lineage program confers resistance to MAP kinase pathway inhibition. *Nature* **2013**, *504*, 138–142.
69. Berns, K.; Hijmans, E.M.; Mullenders, J.; Brummelkamp, T.R.; Velds, A.; Heimerikx, M.; Kerkhoven, R.M.; Madiredjo, M.; Nijkamp, W.; Weigelt, B. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **2004**, *428*, 431–437.
70. Turanli, B.; Altay, O.; Borén, J.; Turkez, H.; Nielsen, J.; Uhlen, M.; Arga, K.Y.; Mardinoglu, A. Systems biology based drug repositioning for development of cancer therapy. *Semin. Cancer Biol.* **2019**, doi:10.1016/j.semcancer.2019.09.020.
71. Mohr, S.E.; Perrimon, N. RNAi screening: New approaches, understandings, and organisms. *Wiley Interdiscip. Rev. Rna* **2012**, *3*, 145–158.
72. Perwitasari, O.; Bakre, A.; Tompkins, S.M.; Tripp, R.A. siRNA genome screening approaches to therapeutic drug repositioning. *Pharmaceuticals* **2013**, *6*, 124–160.

73. Fareh, M.; Yeom, K.-H.; Haagsma, A.C.; Chauhan, S.; Heo, I.; Joo, C. TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments. *Nat. Commun.* **2016**, *7*, 1–11.
74. Luo, J.; Emanuele, M.J.; Li, D.; Creighton, C.J.; Schlabach, M.R.; Westbrook, T.F.; Wong, K.-K.; Elledge, S.J. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* **2009**, *137*, 835–848.
75. Takai, A.; Dang, H.; Oishi, N.; Khatib, S.; Martin, S.P.; Dominguez, D.A.; Luo, J.; Bagni, R.; Wu, X.; Powell, K. Genome-wide RNAi Screen identifies PMPCB as a therapeutic vulnerability in EpCAM+ hepatocellular carcinoma. *Cancer Res.* **2019**, *79*, 2379–2391.
76. Qi, L.S.; Larson, M.H.; Gilbert, L.A.; Doudna, J.A.; Weissman, J.S.; Arkin, A.P.; Lim, W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **2013**, *152*, 1173–1183.
77. Joung, J.; Konermann, S.; Gootenberg, J.S.; Abudayyeh, O.O.; Platt, R.J.; Brigham, M.D.; Sanjana, N.E.; Zhang, F. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **2017**, *12*, 828.
78. Shi, J.; Wang, E.; Milazzo, J.P.; Wang, Z.; Kinney, J.B.; Vakoc, C.R. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **2015**, *33*, 661.
79. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819–823.
80. Wang, H.; Yang, H.; Shivalila, C.S.; Dawlaty, M.M.; Cheng, A.W.; Zhang, F.; Jaenisch, R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **2013**, *153*, 910–918.
81. Doench, J.G.; Hartenian, E.; Graham, D.B.; Tothova, Z.; Hegde, M.; Smith, I.; Sullender, M.; Ebert, B.L.; Xavier, R.J.; Root, D.E. Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nat. Biotechnol.* **2014**, *32*, 1262.
82. Rouet, P.; Smih, F.; Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.* **1994**, *14*, 8096–8106.
83. Kurata, M.; Yamamoto, K.; Moriarity, B.S.; Kitagawa, M.; Largaespada, D.A. CRISPR/Cas9 library screening for drug target discovery. *J. Hum. Genet.* **2018**, *63*, 179–186.
84. Wang, T.; Wei, J.J.; Sabatini, D.M.; Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **2014**, *343*, 80–84.
85. Shalem, O.; Sanjana, N.E.; Hartenian, E.; Shi, X.; Scott, D.A.; Mikkelsen, T.S.; Heckl, D.; Ebert, B.L.; Root, D.E.; Doench, J.G. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **2014**, *343*, 84–87.
86. Shalem, O.; Sanjana, N.E.; Zhang, F. High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet.* **2015**, *16*, 299–311.
87. Koike-Yusa, H.; Li, Y.; Tan, E.-P.; Velasco-Herrera, M.D.C.; Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **2014**, *32*, 267.
88. Kumar-Sinha, C.; Kalyana-Sundaram, S.; Chinnaiyan, A.M. Landscape of gene fusions in epithelial cancers: Seq and ye shall find. *Genome Med.* **2015**, *7*, 129.
89. Wang, T.; Birsoy, K.; Hughes, N.W.; Krupczak, K.M.; Post, Y.; Wei, J.J.; Lander, E.S.; Sabatini, D.M. Identification and characterization of essential genes in the human genome. *Science* **2015**, *350*, 1096–1101.
90. Talevi, A. Drug repositioning: Current approaches and their implications in the precision medicine era. *Expert Rev. Precis. Med. Drug Dev.* **2018**, *3*, 49–61.
91. John Harris, C.; D Hill, R.; W Sheppard, D.; J Slater, M.; FW Stouten, P. The design and application of target-focused compound libraries. *Comb. Chem. High Throughput Screen.* **2011**, *14*, 521–531.
92. Goh, K.-I.; Cusick, M.E.; Valle, D.; Childs, B.; Vidal, M.; Barabási, A.-L. The human disease network. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8685–8690.
93. Piñero, J.; Bravo, À.; Queralt-Rosinach, N.; Gutiérrez-Sacristán, A.; Deu-Pons, J.; Centeno, E.; García-García, J.; Sanz, F.; Furlong, L.I. DisGeNET: A comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.* **2016**, *45*, gkw943.
94. Zhu, F.; Han, B.; Kumar, P.; Liu, X.; Ma, X.; Wei, X.; Huang, L.; Guo, Y.; Han, L.; Zheng, C. Update of TTD: Therapeutic target database. *Nucleic Acids Res.* **2010**, *38*, D787–D791.
95. Gilson, M.K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res.* **2016**, *44*, D1045–D1053.

96. Szklarczyk, D.; Santos, A.; von Mering, C.; Jensen, L.J.; Bork, P.; Kuhn, M. STITCH 5: Augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* **2016**, *44*, D380–D384.
97. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K.P. STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **2015**, *43*, D447–D452.
98. Cheng, F.; Desai, R.J.; Handy, D.E.; Wang, R.; Schneeweiss, S.; Barabási, A.-L.; Loscalzo, J. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **2018**, *9*, 1–12.
99. Cheng, F.; Kovács, I.A.; Barabási, A.-L. Network-based prediction of drug combinations. *Nat. Commun.* **2019**, *10*, 1–11.
100. Conte, F.; Fiscon, G.; Licursi, V.; Bizzarri, D.; D'Antò, T.; Farina, L.; Paci, P. A paradigm shift in medicine: A comprehensive review of network-based approaches. *Biochim. Et Biophys. Acta (Bba)-Gene Regul. Mech.* **2020**, *1863*, 194416.
101. Sander, J.; Ester, M.; Kriegel, H.-P.; Xu, X. Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data Min. Knowl. Discov.* **1998**, *2*, 169–194.
102. Wang, W.; Yang, J.; Muntz, R. STING: A statistical information grid approach to spatial data mining. In Proceedings of the VLDB, Athens, Greece, 25–29 August 1997; pp. 186–195.
103. Hodgkin, J. Seven types of pleiotropy. *Int. J. Dev. Biol.* **2002**, *42*, 501–505.
104. Lu, J.; Chen, L.; Yin, J.; Huang, T.; Bi, Y.; Kong, X.; Zheng, M.; Cai, Y.-D. Identification of new candidate drugs for lung cancer using chemical–chemical interactions, chemical–protein interactions and a K-means clustering algorithm. *J. Biomol. Struct. Dyn.* **2016**, *34*, 906–917.
105. Mei, J.-P.; Kwoh, C.-K.; Yang, P.; Li, X.-L.; Zheng, J. Drug–target interaction prediction by learning from local information and neighbors. *Bioinformatics* **2013**, *29*, 238–245.
106. Köhler, S.; Bauer, S.; Horn, D.; Robinson, P.N. Walking the interactome for prioritization of candidate disease genes. *Am. J. Hum. Genet.* **2008**, *82*, 949–958.
107. Aittokallio, T.; Schwikowski, B. Graph-based methods for analysing networks in cell biology. *Brief. Bioinform.* **2006**, *7*, 243–255.
108. Emig, D.; Ivliev, A.; Pustovalova, O.; Lancashire, L.; Bureeva, S.; Nikolsky, Y.; Bessarabova, M. Drug target prediction and repositioning using an integrated network-based approach. *PLoS ONE* **2013**, *8*, e60618.
109. Vanunu, O.; Magger, O.; Ruppin, E.; Shlomi, T.; Sharan, R. Associating genes and protein complexes with disease via network propagation. *PLoS Comput. Biol.* **2010**, *6*, e1000641.
110. Martinez, V.; Navarro, C.; Cano, C.; Fajardo, W.; Blanco, A. DrugNet: Network-based drug–disease prioritization by integrating heterogeneous data. *Artif. Intell. Med.* **2015**, *63*, 41–49.
111. Gu, C.; Kim, G.B.; Kim, W.J.; Kim, H.U.; Lee, S.Y. Current status and applications of genome-scale metabolic models. *Genome Biol.* **2019**, *20*, 121.
112. Turanli, B.; Zhang, C.; Kim, W.; Benfeitas, R.; Uhlen, M.; Arga, K.Y.; Mardinoglu, A. Discovery of therapeutic agents for prostate cancer using genome-scale metabolic modeling and drug repositioning. *EBioMedicine* **2019**, *42*, 386–396.
113. Raškevičius, V.; Mikalayeva, V.; Antanavičiūtė, I.; Ceslevičienė, I.; Skeberdis, V.A.; Kairys, V.; Bordel, S. Genome scale metabolic models as tools for drug design and personalized medicine. *PLoS ONE* **2018**, *13*, e0190636.
114. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462.
115. Mardinoglu, A.; Agren, R.; Kampf, C.; Asplund, A.; Nookaew, I.; Jacobson, P.; Walley, A.J.; Froguel, P.; Carlsson, L.M.; Uhlen, M. Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Mol. Syst. Biol.* **2013**, *9*, 649.
116. Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **2019**, *18*, 463–477.
117. Réda, C.; Kaufmann, E.; Delahaye-Duriez, A. Machine learning applications in drug development. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 241–252.
118. Gottlieb, A.; Stein, G.Y.; Ruppin, E.; Sharan, R. PREDICT: A method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.* **2011**, *7*, 496.

119. Kim, E.; Choi, A.-s.; Nam, H. Drug repositioning of herbal compounds via a machine-learning approach. *BMC Bioinform.* **2019**, *20*, 247.
120. Bijalwan, V.; Kumar, V.; Kumari, P.; Pascual, J. KNN based machine learning approach for text and document mining. *Int. J. Database Theory Appl.* **2014**, *7*, 61–70.
121. Cohen, T.; Widdows, D.; Schvaneveldt, R.W.; Davies, P.; Rindflesch, T.C. Discovering discovery patterns with predication-based semantic indexing. *J. Biomed. Inform.* **2012**, *45*, 1049–1065.
122. Zhu, Y.; Jung, W.; Wang, F.; Che, C. Drug repurposing against Parkinson's disease by text mining the scientific literature. *Libr. Hi Tech* **2020**, 10.1108/LHT-08-2019-0170, doi:10.1108/LHT-08-2019-0170.
123. Li, J.; Zhu, X.; Chen, J.Y. Building disease-specific drug-protein connectivity maps from molecular interaction networks and PubMed abstracts. *PLoS Comput. Biol.* **2009**, *5*, e1000450.
124. Zeng, X.; Zhu, S.; Liu, X.; Zhou, Y.; Nussinov, R.; Cheng, F. deepDR: A network-based deep learning approach to in silico drug repositioning. *Bioinformatics* **2019**, *35*, 5191–5198.
125. Wu, L.; Ai, N.; Liu, Y.; Wang, Y.; Fan, X. Relating anatomical therapeutic indications by the ensemble similarity of drug sets. *J. Chem. Inf. Modeling* **2013**, *53*, 2154–2160.
126. Sawada, R.; Iwata, H.; Mizutani, S.; Yamanishi, Y. Target-based drug repositioning using large-scale chemical–protein interactome data. *J. Chem. Inf. Modeling* **2015**, *55*, 2717–2730.
127. Ehrt, C.; Brinkjost, T.; Koch, O. Impact of binding site comparisons on medicinal chemistry and rational molecular design. *J. Med. Chem.* **2016**, *59*, 4121–4151.
128. Yang, G.; Ma, A.; Qin, Z.S. An integrated system biology approach yields drug repositioning candidates for the treatment of heart failure. *Front. Genet.* **2019**, *10*, 916.
129. Vargesson, N. Thalidomide-induced teratogenesis: History and mechanisms. *Birth Defects Res. Part C Embryo Today: Rev.* **2015**, *105*, 140–156.
130. Tanhaian, A.; Mohammadi, E.; Vakili-Ghartavol, R.; Saberi, M.R.; Mirzayi, M.; Jaafari, M.R. In silico and In vitro investigation of a likely pathway for anti-cancerous effect of Thrombocidin-1 as a novel anticancer peptide. *Protein Pept. Lett.* **2020**, doi:10.2174/0929866527666200219115129.
131. Kumar, S.; Kumar, S. Molecular Docking: A Structure-Based Approach for Drug Repurposing. In *In Silico Drug Design*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 161–189.
132. Haupt, V.J.; Daminelli, S.; Schroeder, M. Drug promiscuity in PDB: Protein binding site similarity is key. *PLoS ONE* **2013**, *8*, e65894.
133. Peyvandipour, A.; Saberian, N.; Shafi, A.; Donato, M.; Draghici, S. A novel computational approach for drug repurposing using systems biology. *Bioinformatics* **2018**, *34*, 2817–2825.
134. Akhoon, B.A.; Tiwari, H.; Nargotra, A. In Silico Drug Design Methods for Drug Repurposing. In *In Silico Drug Design*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 47–84.
135. Shigemizu, D.; Hu, Z.; Hung, J.-H.; Huang, C.-L.; Wang, Y.; DeLisi, C. Using functional signatures to identify repositioned drugs for breast, myelogenous leukemia and prostate cancer. *PLoS Comput. Biol.* **2012**, *8*, e1002347.
136. Siavelis, J.C.; Bourdakou, M.M.; Athanasiadis, E.I.; Spyrou, G.M.; Nikita, K.S. Bioinformatics methods in drug repurposing for Alzheimer's disease. *Brief. Bioinform.* **2016**, *17*, 322–335.
137. Dudley, J.T.; Sirota, M.; Shenoy, M.; Pai, R.K.; Roedder, S.; Chiang, A.P.; Morgan, A.A.; Sarwal, M.M.; Pasricha, P.J.; Butte, A.J. Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci. Transl. Med.* **2011**, *3*, 96ra76.
138. Lamb, J.; Crawford, E.D.; Peck, D.; Modell, J.W.; Blat, I.C.; Wrobel, M.J.; Lerner, J.; Brunet, J.-P.; Subramanian, A.; Ross, K.N. The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **2006**, *313*, 1929–1935.
139. Musa, A.; Ghoraie, L.S.; Zhang, S.-D.; Glazko, G.; Yli-Harja, O.; Dehmer, M.; Haibe-Kains, B.; Emmert-Streib, F. A review of connectivity map and computational approaches in pharmacogenomics. *Brief. Bioinform.* **2018**, *19*, 506–523.

140. Liu, C.; Su, J.; Yang, F.; Wei, K.; Ma, J.; Zhou, X. Compound signature detection on LINCS L1000 big data. *Mol. Biosyst.* **2015**, *11*, 714–722.
141. Liu, T.-P.; Hsieh, Y.-Y.; Chou, C.-J.; Yang, P.-M. Systematic polypharmacology and drug repurposing via an integrated L1000-based Connectivity Map database mining. *R. Soc. Open Sci.* **2018**, *5*, 181321.