British Journal of Pharmacology

British Journal of Pharmacology (2018) **175** 168–180 168

Themed Section: Inventing New Therapies Without Reinventing the Wheel: The Power of Drug Repurposing

REVIEW ARTICLE

Drug repurposing from the perspective of pharmaceutical companies

Correspondence Daphna Laifenfeld, Global Research and Development, Teva Pharmaceutical Industries, Netanya, Israel. E-mail: daphna.laifenfeld@teva.co.il

Received 7 February 2017; **Revised** 6 March 2017; **Accepted** 8 March 2017

Y Cha[1], T Erez[2], I J Reynolds[3], D Kumar[1], J Ross[1], G Koytiger[1], R Kusko[1], B Zeskind[1], S Risso[3], E Kagan[2], S Papapetropoulos[4], I Grossman[2] and D Laifenfeld[2]

[1]Immuneering Corporation, Cambridge, MA, USA, [2]Global Research and Development, Teva Pharmaceutical Industries, Netanya, Israel, [3]Global Research and Development, Teva Pharmaceutical Industries, West Chester PA, USA, and [4]Global Research and Development, Teva Pharmaceutical Industries, Frazer PA, USA

Drug repurposing holds the potential to bring medications with known safety profiles to new patient populations. Numerous examples exist for the identification of new indications for existing molecules, most stemming from serendipitous findings or focused recent efforts specifically limited to the mode of action of a specific drug. In recent years, the need for new approaches to drug research and development, combined with the advent of big data repositories and associated analytical methods, has generated interest in developing systematic approaches to drug repurposing. A variety of innovative computational methods to enable systematic repurposing screens, experimental as well as through *in silico* approaches, have emerged. An efficient drug repurposing pipeline requires the combination of access to molecular data, appropriate analytical expertise to enable robust insights, expertise and experimental set-up for validation and clinical development know-how. In this review, we describe some of the main approaches to systematic repurposing and discuss the various players in this field and the need for strategic collaborations to increase the likelihood of success in bringing existing molecules to new indications, as well as the current advantages, considerations and challenges in repurposing as a drug development strategy pursued by pharmaceutical companies.

LINKED ARTICLES

This article is part of a themed section on Inventing New Therapies Without Reinventing the Wheel: The Power of Drug Repurposing. To view the other articles in this section visit http://onlinelibrary.wiley.com/doi/10.1111/bph.v175.2/issuetoc

Abbreviations

ALS, amyotrophic lateral sclerosis; AMD, age-related macular degeneration; EMRs, electronic medical records; EOSC, European Open Science Cloud; HTS, high-throughput screening; IP, intellectual property; LCM, life cycle management; MoA, mechanism of action; NCE, new chemical entity; OPCs, oligodendrocyte precursor cells; RWD, real-world data

DOI:10.1111/bph.13798

Introduction

Identification of novel indications for existing compounds through drug repurposing holds the potential to complement traditional drug discovery by mitigating the high monetary- and time- related costs and risks associated with the latter. With a failure rate of ~45% associated with safety or toxicity issues (Ashburn and Thor, 2004), mitigating the safety risk, in addition to the saving of up to 5–7 years in average drug development time (Ashburn and Thor, 2004), confers an attractive prospect for drug developers and patients alike. The latter stand to gain significant advantage in quicker access to drugs that not only would otherwise not be available but also have well-documented safety profiles.

For the purpose of this review, existing compounds refer to those that have a proven safety and tolerability profile based on successful Phase I or Phase II clinical trials. Thus, a candidate repurposing drug will have a well-established safety and toxicity profile, with data already accumulated toward gaining regulatory approval. A multitude of examples exist for repurposed drugs, which can primarily be categorized as stemming either from serendipitous findings - with OnabotulinumtoxinA (BOTOX®; Allergan) with eight different approved indications constituting the prototypical example - or from focused mechanism of action (MoA)-based research efforts pursued by biotech and pharmaceutical companies. Such research into new indications for existing drugs are most often pursued in the context of life cycle management (LCM) activities by which biotech and pharmaceutical companies attempt to extend the patent life of products, through application to adjacent diseases, beyond the primary sought for or obtained indication, as in the example of **infliximab** treatment for ulcerative colitis and rheumatoid arthritis. In some cases, however, in-depth understanding of drug and disease MoA have led to successful development of drugs to entirely new therapeutic arenas – such was the development of **azidothymidine**, which failed as a chemotherapy drug but became a treatment for human immunodeficiency virus (Volberding *et al.*, 1990), and **mycophenolate mofetil** for lupus nephritis (Yong and D'Cruz, 2008). Examples of successfully repurposed drugs, together with the ever-expanding high costs and failures of traditional drug discovery and the advent of new data and technologies, have led to the emergence of a new field of drug repurposing. The historically unintentional, serendipitous or constrained research effort is now being replaced by systematic, high-throughput and rational pursuit of new therapeutic uses for marketed drugs and drugs in development or as a drug-salvaging strategy.

In this review on the approach of pharmaceutical companies to repurposing, we will describe the methodologies, tools and data types that support the new field of systematic, high-throughput, drug repurposing. Covered methodologies span high-throughput screening (HTS) platforms and *in silico* approaches including omics-based repurposing and the use of real-world data (RWD) to support utility of approved drugs in new indications. This review does not cover structure-based repurposing approaches which have recently been reviewed by Martorana *et al.* (2016). We will make the case for combined use of methodologies for generating robust repurposing candidates with a high likelihood of success and discuss current considerations and challenges in repurposing as a drug development strategy.

Major players in the repurposing field

Three key players can be identified in the field of repurposing, including academia and research institutes, pharmaceutical companies and repurposing technology companies. While all share elements of a common approach to drug development through identifying new indications for existing drugs, the emphasis, field of expertise or capabilities, as well as business models and incentives, differ greatly. Academia can be viewed as divided between the approach of developing a molecule or indication and that of focusing on methodologies such as *in silico* or HTS. It is less constrained by the need of an economical or commercial success but is dependent on scientific breakthroughs that attract talent, government funding and partnerships with for-profit institutions. Repurposing technology companies, on the other hand, are bound by the choice of their business model, which may vary depending on the company's vision and capability. These models include consulting services, service provider such as compound evaluation, database and/or platform for drug libraries, screening platform and drug pipeline. Often, a mixed model is used, which is likely to increase both the development outcome of repurposing and the business success. Repurposing technology companies often utilize advanced technologies, expertise, focused research and agility but may lack access to the resources needed to rapidly and successfully pursue preclinical and clinical development of a drug in a regulatory-oriented environment. On the other hand, larger pharmaceutical companies may be more focused on LCM activities of a specific product or molecule, which is often done as part of the late-stage development of the product or post-marketing. For these reasons, partnerships between smaller repurposing technology companies and larger pharmaceutical companies, when effectively executed, represent an attractive means to combine advanced repurposing capabilities with deep expertise in drug development. Such an example is the case of Adamas Pharmaceuticals, a small biotech company that developed a new drug combination for the innovative marketed **memantine**, which was combined with **donepezil** for the treatment of Alzheimer's disease (Howard *et al.*, 2012). Forrest Labs, the original proprietor of memantine, bought the new therapeutic entity for exclusivity rights.

Repurposing approaches

Systematic repurposing approaches can be largely divided into those based on experimental screening approaches, and *in silico* approaches that employ existing data to identify potential new drug–disease associations.

Experimental screening approaches

Experimental screening approaches are used as a source of hits for both drug discovery and drug repurposing, with notable differences in their application and outcomes. Searches in drug discovery programmes are typically done for *de novo* candidate hits, fuelled by an HTS campaign, which requires highly specialized screening facilities and compound libraries containing several million compounds. Repurposing programmes focus on advanced known molecules either approved or failed with some knowledge of their safety or MoA available, led by in-depth screening, and with smaller size compound libraries. Typical approved compound libraries containing 500–2000 compounds and a similar number of existing but unapproved compounds are thought to be available. Table 1 summarizes the different compound collections available to date, with most being approved, or known from clinical trial reports. Some may contain annotation and/or data on safety and MoA. Compound libraries represent a scale that is manageable in many academic drug discovery laboratories, opening up opportunities for additional players and approaches to

Table 1

Compound library

Library	Number of Small Molecules	Reference	Comments
The Drug Repurposing Hub	~5000 compounds from preclinical to launched phase	https://clue.io/repurposing	The Drug Repurposing Hub is a close collaboration between the Broad Institute Cancer Program, Center for the Development of Therapeutics and the Connectivity Map group
FDA-approved anticancer drugs	The current set (AODVII) consists of 129 most current FDA-approved anticancer drugs	https://dtp.cancer.gov/organization/dscb/obtaining/available_plates.htm	The collection contains the most current FDA-approved anticancer drugs. The current set (AODVII) consists of 129 agents and is intended to enable cancer research, drug discovery and combination drug studies
NIH Small Molecule Repository	Specialty sets are included in this library comprising bioactive compounds such as known drugs and toxins	https://mlsmr.evotec.com/	NIH Molecular Libraries Small Molecule Repository collects samples for high-throughput biological screening
The Microsource Spectrum Collection: US Drug Collection	1360 compounds that have reached clinical trials in the USA	http://www.msdiscovery.com/usdrug.html	Compounds marketed in the USA
The Microsource Spectrum Collection: International Drug Collection	400 compounds that are or have been marketed in Europe and/or Asia	http://www.msdiscovery.com/intdrug.html	Compounds marketed in Europe and/or Asia
John S. Dunn Gulf Coast Consortium	~125 compounds	http://www.gulfcoastconsortia.org/	Fifty-seven per cent of the compounds are currently used in the clinic for the treatment of various forms of cancer, and 37% of the compounds are in clinical trials
LOPAC®1280	1280 compounds	http://www.sigmaaldrich.com/life-science/cell-biology/bioactive-small-molecules/lopac1280-navigator.html	Collection of inhibitors, receptor ligands, pharma-developed tools and approved drugs
SCREEN-WELL® FDA-approved drug library V2	Over 770 compounds	http://www.enzolifesciences.com/BML-2843/screen-well-fda-approved-drug-library-v2	FDA-approved compounds library
FDA-approved drug library	1447 FDA-approved drugs	http://www.selleckchem.com/screening/fda-approved-drug-library.html	FDA-approved compounds library
Teva Screening Set	640 FDA-and foreign-approved drugs	Contact the authors	Compounds approved in the USA and Europe, most of which are marketed by Teva

identification of candidate hits for a clinical development programme.

Another key difference between drug discovery and repurposing is the fate of the hits. For a repurposing screen, a compelling hit is a drug molecule that is a candidate to be advanced into development. This is quite distinct from an HTS hit in drug discovery that becomes a starting point for a newly found medicinal chemistry programme which is then evolved iteratively. Lastly, for drug discovery screens, an efficient assay must be simple and fast in order to manage the number of potential compounds for review, whereas for repurposing screens, the limited scale allows a broader range of complexity of assay types.

Two types of screens, commonly used and adaptable for HTS, are cell-free or cell-based, target-focused screens and phenotypical screens. The first is based on a specific compound activity in relation to a specific mechanism. The second is based on cell behaviour (e.g. growth and death) with or without the compound addition. Phenotypical screens have historically been more successful than target screening in generating new drugs (Swinney and Anthony, 2011; Eder *et al.*, 2014). A potentially powerful variant of phenotypical assays is the use of screens that report signals relevant to disease biology. However, the use of *in vitro* assays is challenged by its validity in presenting the human disease, and an ongoing debate around the use of more representative cell-based models such as human-induced pluripotent stem cell-derived tissues for the case of degenerative diseases is likely to continue (Horvath *et al.*, 2016). Screening with multicellular model organisms [e.g. worms (*Caenorhabditis elegans*), fruit flies (*Drosophila melanogaster*) and larval zebrafish (*Danio rerio*)] introduced with human genes and/or disease-causing mutations often provides a more compelling approach for phenotypical screening (Pandey and Nichols, 2011). While there are obviously limitations to the extent to which the human disease phenotype is fully represented in these models, they clearly have the potential to contribute to the identification and prioritization of active molecules, early in a repurposing effort.

A successful case-study of repurposing potential through screening approaches is provided by the convergent results that have emerged from several laboratories, working independently in search of compounds that promote myelin repair. Myelin repair involves both enhancing the differentiation of oligodendrocyte precursor cells (OPCs) into oligodendrocytes and the promotion of axon remyelination. A phenotypical screen using zebrafish model reported OPC migration and increased production of myelin basic protein, a key marker of differentiation (Buckley *et al.*, 2010). A subsequent screen using primary rat optic nerve OPCs identified the muscarinic antagonist **benztropine**, a synthetic compound used in the treatment of Parkinson's disease (PD), as a potential hit for development through repurposing (Deshmukh *et al.*, 2013). Importantly, benztropine penetrated the blood-brain barrier, a key requirement for myelin repair. Mei *et al.* (2014) used a novel micropillar pseudoaxonal substrate to track OPC differentiation into cells that form myelin and identified a cluster of muscarinic antagonists including benztropine and **clemastine**, another brain-penetrant compound with muscarinic and histamine antagonist activity. More recently, Najm *et al.* (2015) used stem cell-derived OPCs for a remyelination screen and once again confirmed benztropine as a hit and two additional unreported positives, **miconazole** and **clobetasol**. **Quetiapine** is another muscarinic antagonist compound found independently of repurposing screens (Xiao *et al.*, 2008) but was confirmed in relation to remyelination in a primary rat OPC differentiation assay (Lariosa-Willingham *et al.*, 2016). Both clemastine and quetiapine have rapidly advanced into clinical trials in multiple sclerosis (clinicaltrials.gov study numbers: NCT02040298 and NCT02087631, respectively) to evaluate their efficacy in myelin repair. The ability of multiple independent screens to identify similar classes of compound and the potential to advance repurposing hits rapidly into clinical development effectively highlight the potential of repurposing screens.

In silico repurposing approaches

In silico repurposing approaches apply sophisticated analytical methods to existing data identifying new potential associations between drug and disease. Approaches can be broadly divided into two categories: (i) molecular approaches, which are based on understanding of drug activity and disease pathophysiology and are often powered by large-scale molecular data (i.e. 'omic data'), such as genomic, transcriptomic or proteomic data, as well as data on drug targets and chemical structure, and (ii) RWD approaches, focusing on identification of unknown, and at times unexpected, relationships between drugs and diseases or their symptoms, based on RWD – data on individuals' health, habits and behaviour that are captured without intervention from the environment or biases introduced through data collection methodologies.

Molecular approaches

Understanding the MoA of a drug and matching it with a disease indication different from that for which it was originally approved or developed are at the heart of molecular approaches to drug repurposing. Access to increasing amounts of data largely attributed to the rapid advance of molecular biology methodologies and omic data enables the application of systematic analyses to drug and disease MoA (Sanseau *et al.*, 2012; Rastegar-Mojarad *et al.*, 2015).

Of the various types of omic data available, transcriptomics and genomics are the two data types most widely used to support drug repurposing, due to the combination of availability of datasets on drugs and diseases and the robustness and reproducibility of the data (Iorio *et al.*, 2013). Transcriptomics involves measuring the expression levels of thousands of genes, often by quantifying RNA using RNA-Seq or gene expression microarrays. One approach of applying transcriptomics to drug repurposing is based on the concept that reversal of gene expression signatures may translate into clinical benefit (Lamb *et al.*, 2006; Sirota *et al.*, 2011). In an early example of this approach, published along with the report of one of the first large-scale datasets of drug-induced gene expression profiles (the Broad Institute's Connectivity Map), **sirolimus**, an **mTOR** inhibitor, was shown to reverse the signature of

dexamethasone resistance in acute lymphoblastic leukaemia (Lamb *et al.*, 2006). The *in silico* findings were subsequently confirmed *in vitro* through a dramatic reduction in viability of cells exposed to sirolimus in combination with dexamethasone, compared with dexamethasone alone. Furthermore, the sirolimus–dexamethasone combination provided durable remission in xenografted acute lymphoblastic leukaemia (Teachey *et al.*, 2008), supporting the transcriptomic-based repurposing approach. Clinical validation for the concept of using sirolimus or related drugs to overcome resistance to glucocorticoids is, however, still pending (Fransecky *et al.*, 2015).

The utility of gene expression data to support disease indications has been further demonstrated in a recent publication (Geva *et al.*, 2016) which found that pridopidine up-regulates expression of genes associated with various neuroprotective pathways down-regulated in Huntington's disease. In addition to transcriptomic approaches, genomics (mutations and genetic variants associated with disease) is also commonly utilized as means to identify repurposing candidate, for example, drugs targeting genes associated with genome-wide association study hits in a given disease (Sanseau *et al.*, 2012). Consistently, it was demonstrated that drugs with genetic support have substantially higher likelihood of success in clinical trials (Nelson *et al.*, 2015). Another recent publication demonstrated methods by which the combination of transcriptomic and genomic data can be used to identify potential repurposing candidates for Alzheimer's disease (Fowler *et al.*, 2015).

In recent years, a rich set of resources has become available for repurposing activities, falling into three broad categories. Drug-focused databases (as shown in Table 2) provide information on the gene expression profiles, known targets, clinical status and other aspects of various drugs. Disease-focused databases (shown in Table 3) provide detailed information on the gene expression profiles and proteomic, genetic and epigenetic characteristics of diseases. Finally, various tools to link drugs with diseases have been made available (Table 4). Each of these databases and tools has strengths and limitations: listing here does not imply endorsement, and appropriate caution and quality control are advised.

RWD approaches to drug repurposing

RWD, non-interventional data on individuals' activities and health, are characterized by large, complex, intricately structured datasets, often containing several years of data on millions of patients. Data sources for RWD can stem either

Table 2

Drug database

Name of the database	Description	Link
PubChem (Bolton *et al.*, 2008)	Biological activity of >60 million unique compounds	http://pubchem.ncbi.nlm.nih.gov/
ChEMBL (Gaulton *et al.*, 2012)	Curated database with compound activity against target genes	https://www.ebi.ac.uk/chembl
LINCS (Vidović *et al.*, 2014)	Follow-up project to CMap, L1000-based expression profiles of drug-treated cancer cell lines	http://lincscloud.org/
Project Achilles (Cowley *et al.*, 2014)	Cancer cell line RNAi screen to ID genes relevant to cell survival	https://portals.broadinstitute.org/achilles
CMap (Lamb *et al.*, 2006)	Expression profiles of drug-treated cancer cell lines	http://www.broadinstitute.org/cmap
CTRP (Basu *et al.*, 2013; Seashore-Ludlow *et al.*, 2015; Rees *et al.*, 2016)	Screen of 860 cancer cell lines for sensitivity to 480 drugs and probes. Portal allows comparison to mutation, expression, CNV data	http://www.broadinstitute.org/ctrp.v2.2
ImmPort (Bhattacharya *et al.*, 2014)	Portal containing 222 studies with 37k subjects includes ELISA, ELISPOT and flow cytometry data	https://immport.niaid.nih.gov/
PharmGKB (Hewett, 2002)	Expert curated gene–drug genotype–phenotype connections, dosing guidelines and drug labels	https://www.pharmgkb.org/
e-Drug3D (Pihan *et al.*, 2012)	Mirrors US pharmacopoeia of small drugs, 1822 molecular structures	http://chemoinfo.ipmc.cnrs.fr/MOLDB/index.html
DailyMED	Catalogue of drug listings/drug label information	https://dailymed.nlm.nih.gov/dailymed/
Comparative Toxicogenomics Database (Mattingly *et al.*, 2006; Pihan *et al.*, 2012)	~1.5 M chemical–gene, ~2 M chemical–disease and ~20 M gene–disease interactions	http://ctdbase.org/

CNV, copy number variation.

Table 3

Disease database

Type	Name	Description	Link
Multi-omic level	TCGA (The Cancer Genome Atlas Research Network *et al.*, 2013)	Clinical and multi-omic data of 33 different tumour types	http://tcga-data.nci.nih.gov
	ICGC (International Cancer Genome Consortium *et al.*, 2010)	Genomic, transcriptomic and epigenomic data from ±25 000 tumours	http://icgc.org/
	CCLE (Barretina *et al.*, 2012)	Cancer cell line database of gene expression arrays, CNVs, mutations, and IC50s	http://software.broadinstitute.org/software/cprg/
Genomic	dbGAP (Mailman *et al.*, 2007)	Database of genotype and sequence data phenotype	http://www.ncbi.nlm.nih.gov/gap
	DisGeNET (Piñero *et al.*, 2015)	Curated databases together with text-mining derived associations to generate human gene–disease associations	http://www.disgenet.org/web/DisGeNET/menu
	dbSNP (Sherry, 2001)	Database for SNPs, DIPs and STRs for human and model organisms	http://www.ncbi.nlm.nih.gov/snp
	dbVar (Lappalainen *et al.*, 2013)	Database of human genome structural variations including CNVs	http://www.ncbi.nlm.nih.gov/dbvar
	1000 Genome Project (1000 Genomes Project Consortium *et al.*, 2015)	Large resource of human variation and genotype (2504 samples)	http://www.1000genomes.org/
	COSMIC (Forbes *et al.*, 2014)	Cancer focused, contains expert curation based on literature and systematic screen data from other databases	http://cancer.sanger.ac.uk/cosmic
Transcriptomic	GTex (Lonsdale *et al.*, 2013)	Connects genotype with tissue-specific expression level, for comprehensive human eQTLs	http://www.gtexportal.org/
	GEO (Barrett *et al.*, 2013)	Raw and processed transcriptomic data from microarrays, RNA-Seq and other platforms	http://www.ncbi.nlm.nih.gov/geo
	ArrayExpress (Kolesnikov *et al.*, 2015)	Raw and processed transcriptomic data	https://www.ebi.ac.uk/arrayexpress
	Allen Brain Atlas (Allen Institute for Brain Science, 2017)	Expression data from human and mouse brain compartments	http://www.brain-map.org/
Proteomic	Human Proteome Map (Kim *et al.*, 2014)	LC–MS/MS proteomics of multiple organs/tissues	http://www.humanproteomemap.org/
	StringDB (von Mering *et al.*, 2005)	Predicted and known protein–protein interaction database including >9 million proteins	http://string-db.org/
Epigenetic	ENCODE (The ENCODE Project Consortium, 2004)	Annotation of human genome function with TF ChIP-Seq and RNA-Seq	https://genome.ucsc.edu/ENCODE/
	Roadmap (Bernstein *et al.*, 2010)	Epigenetic data including ChIP-Seq and DNA methylation from many cell types and tissues	http://www.roadmapepigenomics.org/
	PsychEncode (PsychENCODE Consortium *et al.*, 2015)	Epigenetic data from psychiatric disease and healthy brains, focused on non-coding genome	https://www.synapse.org//#!Synapse:syn4921369/wiki/235539

CNV, copy number variation; eQTLs, expression quantitative trait loci.

Table 4

Omics analytical tools

Tool Name	Description	Link
ksRepo (Brown *et al.*, 2016)	Integrates gene expression and drug datasets from different platforms	https://github.com/adam-sam-brown/ksRepo
GoPredict (Louhimo *et al.*, 2016)	Gene ontology-based drug prioritization for breast and ovarian cancer	http://csblcanges.fimm.fi/GOPredict/
PREDICT (Gottlieb *et al.*, 2011)	Drug–drug and disease–disease relatedness	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3159979/
RE:fine drugs (Moosavinasab *et al.*, 2016)	Interactive portal linking 916 drugs, 567 genes to 1770 diseases using GWAS and PheWAS	https://drug-repurposing.nationwidechildrens.org/search
RANKS (Valentini *et al.*, 2016)	Graph-based node label ranking, can be applied to repurposing	https://cran.r-project.org/web/packages/RANKS/index.html
COGENA (Jia *et al.*, 2016)	Gene co-expression analysis with repositioning-oriented enrichment analysis	https://github.com/zhilongjia/cogena
DR.PRODIS (Zhou *et al.*, 2015)	Tests user input molecule against pre-computed target library	http://cssb.biology.gatech.edu/repurpose
GIFT (Zu *et al.*, 2015)	Infers chemogenomic features from drug chemical substructures	http://bioinfo.au.tsinghua.edu.cn/software/GIFT/
NFFinder (Setoain *et al.*, 2015)	Search engine of gene expression profiles to identify conditions similar to or opposite query	http://nffinder.cnb.csic.es/
PROMISCUOUS (von Eichborn *et al.*, 2011)	Searchable database of protein–protein and drug–protein connections	http://bioinformatics.charite.de/promiscuous/
MANTRA (Iorio *et al.*, 2010)	Utilizes network theory and non-parametric statistics of expression data to identify alternative mechanisms of action	http://mantra.tigem.it/
DSigDB (Yoo *et al.*, 2015)	GSEA-formatted gene sets of drug binding data, CMap data, kinase signatures and computational drug–gene interaction predictions	http://tanlab.ucdenver.edu/DSigDB/DSigDBv1.0/

GWAS, genome-wide association study; PheWAS, Phenome-wide association study; GSEA, gene set enrichment analysis.

from observational, simple trials (i.e. pragmatic trials) or from registries, administrative data, health surveys, electronic medical records (EMRs), medical chart reviews or adverse event reporting and even social media. This type of data can address a wide range of challenges across drug development and has been mainly used to support health economics research. However, RWD constitute a fertile, and largely untapped, ground for generating and validating drug repurposing candidates. For example, Xu *et al.* (2014) used EMR data to demonstrate that **metformin**, a common oral medication for type 2 diabetes, was associated with decreased mortality after a cancer diagnosis compared with diabetic cancer patients on metformin, and compared with non-diabetic cancer patients not on metformin. In another example, Brilliant *et al.* (2016) combined the use of EMR and insurance claim-based data to test the protective potential of **L-DOPA** (Levodopa) against age-related macular degeneration (AMD). The researchers were able to demonstrate that AMD occurred significantly later in patients with L-DOPA prescription (applied for the treatment of other indications in these patients, primarily PD) compared with those without and found that the odds ratio of developing AMD significantly negatively correlated with L-DOPA use. While neither of the above studies has been validated in a prospective clinical study, they demonstrate the potential in using RWD to identify patterns of unexpected efficacy for existing compounds. The ability to utilize RWD, often stemming from disparate sources each with its own annotation system and limited accompanying clinical phenotypical data, is highly dependent on the advent of sophisticated analytical methods such as artificial intelligence and deep learning, and their application to healthcare. The appropriate demonstration of utility in combining and integrating multiple data types with varying levels of ascertainment and metadata is critical for the growing need to enable the extraction of additional insights from RWD and other big data modalities.

Systematically combining drug repurposing approaches

Each of the methodologies described above addresses the question of pairing diseases to potentially beneficial drugs from a different perspective. HTS approaches aim to affect a specific phenotype or target; molecular/omic approaches are

based on the mechanistic principle that efficacy relies on the ability of the drug to alter disease-related pathways, and RWD assess unexpected occurrences and biases that may reflect novel utility for a given drug. Ultimately, robust repurposing candidates with high likelihood of success will stem from a systematic and rigorous combination of various data types, approaches and analytics. One such approach is based on the notion that shared drug side effects may imply similar molecular mechanisms. The former may then be utilized to identify drugs similar to those already known to treat a disease by combination of RWD [e.g. the Food and Drug Administration (FDA) adverse events reporting system] and databases on drug side effects from randomized-controlled clinical trials (Yang and Agarwal, 2011; Bisgin *et al.*, 2014; Ye *et al.*, 2014; Kuhn, 2016). Additional methodologies combining multiple approaches have been demonstrated, for example, through the use of phenotypical, chemical and genomic similarity among drugs (Gottlieb *et al.*, 2011; Yu *et al.*, 2016). Structural similarity can be used to infer new targets for existing therapies and combined with both phenotypical and gene expression-based signatures to infer drugs with similar functions to existing therapies for a disease (Tan *et al.*, 2014).

A repurposing approach involving combination of different methodologies and datasets was recently reported by Paik *et al.* (2015), who generated disease and drug pair similarity scores independently in genomics data and in EMR-extracted lab test data, based on the underlying assumption that similar drugs can be used to treat similar diseases. The result was identification of **terbutaline sulfate**, a **β_2-adrenoceptor** agonist widely used for the treatment of asthma, as a candidate for treatment of amyotrophic lateral sclerosis (ALS), based on similarity between terbutaline sulfate and **ursodeoxycholic acid**, on the one hand, and Kawasaki syndrome and ALS, on the other. The potential therapeutic benefit of terbutaline sulfate for ALS was then demonstrated and validated via prevention of defects in axons and neuromuscular junction degeneration in a zebrafish model of ALS, but to date, no clinical trial has demonstrated a similar benefit in humans.

Pharmaceutical companies and drug repurposing

Drug repurposing activities involve the identification of robust candidates with promising new indications and the subsequent clinical development of these drugs in their new indications, as far as regulatory submission and approval. The academic contribution to the development of analytical methods and approaches such as those described above to enable systematic analysis of data to generate drug repurposing insights is critical, as is the concentrated focus of repurposing-focused biotech companies. The core expertise of pharmaceutical companies, including that of clinical development, together with their overall remit to bring effective drugs to patients, positions them well to pursue a systematic approach to drug repurposing, while relying on partnerships and collaborations to increase likelihood of success in candidate identification. Within this overall scheme, pharmaceutical companies have several advantages, in particular with regard to data access and clinical development, as well as some specific constraints, inherent in their need for an underlying business model.

Drug data access

Drugs constituting potential candidates for repurposing can be divided into three categories: (i) generic drugs, that is, drugs approved that are no longer protected by patents; (ii) failed drugs – those that have been through some stage of clinical development but are currently not on the market; and (iii) patented drugs, either approved or in the late stage of clinical development.

Generic drugs constitute, by and large, a common resource for all major players in drug repurposing, that is, academia, biotech and pharmaceutical companies. These drugs have been commercially available for a significant amount of time, have well-known and largely documented safety profiles and are readily accessible for studying preclinically as well as for clinical development. However, with regard to drugs falling into the two other categories, that is, failed drugs or those still under patent, the pharmaceutical companies hold an advantage in access to drug-related data, for their respective compounds. Thus, in many cases, access to data is biased toward the pharmaceutical companies, rendering them a key partner in an effort to systematically identify new indications for existing molecules.

Clinical development principles for drug repurposing

Clinical development applies findings from basic science to enhance human health and well-being. In a medical research context, it aims to 'translate' findings in fundamental research into medical practice and meaningful health outcomes (Woolf, 2008). These efforts often continue past the approval of the drug and extend our knowledge of the compound and its use. Clinical development principles for drug repurposing should be adopted from those applied for new chemical entity (NCE) drug development, in which the pharmaceutical companies are well versed. There may be elements in the process that can be shortened because of existing knowledge. However, the trajectory of clinical development is generally similar between drug discovery and repurposing efforts. Two major apparent categories in drug development can be linked to failure in drug approval, namely, safety and efficacy (Schuster *et al.*, 2005). The first relates to the toxicological and pharmacokinetic profiles of the NCE, which are mainly addressed in Phases I and IIa clinical trials, following many preclinical evaluations. The second refers to target engagement and biological activities that are subject to therapeutic interference and is linked to the clinical efficacy of the NCE under investigation.

The clinical testing of a repurposed drug necessitates the integrated understanding of the fundamental pharmacokinetic/pharmacodynamic principles of exposure at the site of action, target binding and expression of functional pharmacological activity (termed together as the 'three pillars of survival'; Morgan *et al.*, 2012). Early stages of development (pre-proof of concept) determine the likelihood of either newly or repurposed candidate compound to continue in Phase II trials and improve the

chance of progression to Phase III and ultimate approval success. For drug repurposing programmes, some elements may have been elucidated during the development or approval of another indication providing some measure of early risk reduction in terms of both cost and time. On one end of the drug repurposing spectrum is an approved drug being examined at similar or lower doses than the maximum dose already approved by regulatory agencies to target the same molecular pathway/mechanism in a different patient population (Oprea *et al.*, 2011). The large body of clinical data and the experience accumulated in Phase III (efficacy) and Phase IV (post-marketing) trials for the drug in question offer a good understanding of its profile in terms of adverse events, long-term and chronic toxicity and on- and off-label effects limiting the need for Phase I and Phase IIa/IIb clinical trials, resulting in shorter development timelines.

On the other end of the spectrum are approved drugs being evaluated for novel therapeutic indications through the discovery of biologically and clinically relevant affinities for new targets, which play a determinant role in those indications (Oprea and Mestres, 2012). In this case, identifying the 'right compound' during repurposing efforts is just the beginning and caution should be applied against over-relying on prior clinical experience; all elements – far beyond the 'three pillars' – should be thoroughly re-explored with the following objective: the right dose, reaching the right compartment and to engage the right target, for the right time, in the right patient population, measured for the right duration, with the right tools/methodology. Consequently, under these conditions, a full clinical development programme similar to that of an NCE might ensue without necessarily saving time or cost.

Current challenges and prospects for repurposing activities in pharmaceutical companies

Drug repurposing by pharmaceutical companies faces many and varied challenges. In addition to the scientific challenges of identifying favourable and robust candidate compounds, there is a need to establish business models to support bringing existing molecules as therapies for new indications. Despite the potentially abbreviated clinical development path for repurposed drugs, there remains a significant commitment in the need to demonstrate the efficacy of the molecules in new indications, and pharmaceutical companies are faced with a challenge in trying to recoup the investment needed to bring a repurposed product to market. Existing incentives to support investment in the development of repurposing drugs include 3 years of market exclusivity in the USA and 10 years in Europe. In order to take advantage of this exclusivity, products rely on the generation of new intellectual property (IP). Such IP might arise from the repurposing concept itself, a new formulation or any of a number of other avenues (e.g. dose level, drug combination and route of administration) that might improve the drug in tailoring it specifically to a new patient population. Nevertheless, and in particular for compounds for which substantial data exist, new IP may be challenging and may constrain the portfolio of compounds that can be pursued by pharmaceutical companies. Additionally, the inventions underlying the new IP often create an additional tier of risk associated with the product's development process. As opposed to the abbreviated regulatory pathway that repurposed drugs are expected to follow, such as the FDA's 505(b)(2), for new formulation or technology, an almost complete clinical programme may be required in order to characterize the new technology/formulation, beyond the efficacy for the new indication. This can lead to lengthy development timelines and high costs and create another barrier to bringing the repurposed medicine to the clinic. Giovannoni *et al.* (2015) proposes several policy-oriented approaches, beyond extended exclusivity, for incentivizing drug repurposing in the pharmaceutical companies, including the development of more flexible regulatory paradigms to enable earlier approvals followed by post-marketing studies. The FDA's Qualified Infectious Disease Product designation, which provides expedited approval and additional years of marketing exclusivity for the approval of new antibiotics (Woodcock, 2014), may serve as a template for such new policies for drug repurposing. Importantly, the business model limitations described above often do not apply to academics and non-profit patient advocacy organizations, and novel ways of engaging in multi-partner collaborations may help support drug repurposing activities by combining the needs, expertise, and demands of the various entities. For instance, a variety of non-profit organizations provide funding for repurposing activities, including Cures Within Reach, the Michael J. Fox Foundation and the Alzheimer's Drug Discovery Foundation (Shineman *et al.*, 2014).

Summary and future directions

Despite a significant increase in spending by the pharmaceutical companies, the rate of new drugs being approved remains steady, mostly due to high attrition rates. This, combined with a large unmet need in treating a wide array of diseases and syndromes, results in an acute demand for innovation in bringing effective therapies to market. In recent years, drug repurposing has emerged as a viable strategy to increase the overall productivity of drug discovery. The surge of biomedical data, including genomic data as well as availability of big data through electronic capture (e.g. EMRs, claims data, social media and sensor data), and advances in accompanying analytical methods provide a critical substrate enabling the systematic assessment of repurposing candidates. Indeed, the massive accumulation of orthogonal data types supports the holistic understanding of the drugs and diseases for effective and data-driven repurposing.

The ability to realize the full promise of data and analytical methods for successful repurposing programmes depends on two key aspects. First, the scientific aspect which involves ensuring not only high-quantity but also high-quality data generation, with the appropriate validation metrics in place, as well as the ability to develop approaches across the fields of computer science, informatics and biotechnology to support the integration of multiple data

types. In the pharmaceutical companies, cross-disciplinary teams, involving biostats, translational science and clinical aspects, will be essential to systematically evaluate and improve upon the algorithms and methods underpinning drug repurposing. Second, the business aspect, including incentives put in place to provide an enabling ecosystem and to support the value chain in developing such drugs. Both large pharmaceutical companies and smaller biotech companies bring important strengths and expertise to the repurposing field. The pharmaceutical companies contribute extensive drug discovery and development capabilities along with a deep understanding of the regulatory, IP and commercial aspects, and the biotech companies provide innovative and quickly evolving technologies for identifying optimal repurposing candidates, along with the agility to pursue the most promising opportunities rapidly. For these reasons, partnerships between large pharmaceutical companies and smaller biotech groups present a particularly compelling means to identify and develop repurposing concepts as quickly as possible.

Three examples illustrate the types of future development and resources that will enable additional repurposing activities. First, following the European Medicines Agency policy 70, which grants open access to clinical trial study data for all clinical drug applications submitted after 1 January 2015, clinical trial results in the EU will be made publicly available through the European Clinical Trials Database (EudraCT 2016). Second, in order to facilitate a standardized research infrastructure, the EU is developing the European Open Science Cloud (EOSC), a comprehensive resource for sharing and analysing biomedical and clinical data. This approach will require the complicated task of establishing standards for data sharing as well as addressing geopolitical rules for sharing and protecting this information. However, once completed (estimated by 2020), the EOSC will comprise harmonized biomedical and clinical data to enable researchers to systematically identify unexpected connections and drug-driven phenotypes in patients that could be used for repurposing. Third, with the advent of wearable devices and the appropriate regulatory path to approved associated algorithms, continuous 24 h monitoring of vital physiological metrics such as heart rate, blood pressure and motor activity has become possible (Pantelopoulos and Bourbakis, 2010). In addition to wearable data, other types of data are increasingly becoming available outside the context of clinical trials, including genomic data obtained through direct to consumer companies, as well as consumer data and data from social media resources. The art of weaving together these types of big data is still in its infancy but holds great potential for efficient repurposing efforts.

Nomenclature of targets and ligands

Key protein targets and ligands in this article are hyperlinked to corresponding entries in http://www.guidetopharmacology.org, the common portal for data from the IUPHAR/BPS Guide to PHARMACOLOGY (Southan *et al.*, 2016), and are permanently archived in the Concise Guide to PHARMACOLOGY 2015/16 (Alexander *et al.*, 2015a,b).

Acknowledgements

Authors appreciate the assistance of Pippa Loupe, PhD (Global Research and Development, Teva Pharmaceutical Industries, Overland Park, KS, USA), in manuscript preparation.

Conflict of interest

Authors B.Z., R.K., G.K., Y.C., J.R. and D.K. are employees of Immuneering Corporation (Cambridge, MA, USA; company is 51% owned by Teva Pharmaceutical Industries). Authors T.E., D.L., I.R., S.R., E.K., S.P. and I.G. are employees in the Research and Development Division of Teva Pharmaceutical Industries (Netanya, Israel).

References

1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM *et al.* (2015). A global reference for human genetic variation. Nature 526: 68–74.

Alexander SPH, Fabbro D, Kelly E, Marrion N, Peters JA, Benson HE *et al.* (2015a). The Concise Guide to PHARMACOLOGY 2015/16: Enzymes. Br J Pharmacol 172: 6024–6109.

Alexander SPH, Davenport AP, Kelly E, Marrion N, Peters JA, Benson HE *et al.* (2015b). The Concise Guide to PHARMACOLOGY 2015/16: G protein-coupled receptors. Br J Pharmacol 172: 5744–5869.

Allen Institute for Brain Science. Allen Brain Atlas (Online). Available at: http://brain-map.org/ (Accessed on 27 January 2017).

Ashburn TT, Thor KB (2004). Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov 3: 673–683.

Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S *et al.* (2012). The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483: 603–607.

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M *et al.* (2013). NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res 41: D991–D995.

Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI *et al.* (2013). An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. Cell 154: 1151–1161.

Bernstein BE, Stamatoyannopoulos JA, Costello JF, Bing R, Aleksandar M, Alexander M *et al.* (2010). The NIH roadmap epigenomics mapping consortium. Nat Biotechnol 28: 1045–1048.

Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J *et al.* (2014). ImmPort: disseminating data to the public for the future of immunology. Immunol Res 58: 234–239.

Bisgin H, Liu Z, Hong F, Reagan K, Xiaowei X, Tong W (2014). A phenome-guided drug repositioning through a latent variable model. BMC Bioinformatics 15: 267–279.

Bolton EE, Wang Y, Thiessen PA, Bryant SH (2008). PubChem: integrated platform of small molecules and biological activities. In: Wheeler RA, Spellmeyer DC (eds). Annual Reports in Computational Chemistry, Vol. 4. Elsevier: Amsterdam, pp. 217–241.

Brilliant MH, Vaziri K, Connor TB Jr, Schwartz SG, Carroll JJ, McCarty CA *et al.* (2016). Mining retrospective data for virtual prospective drug repurposing: L-DOPA and age-related macular degeneration. Am J Med 129: 292–298.

Brown AS, Kong SW, Kohane IS, Patel CJ (2016). ksRepo: a generalized platform for computational drug repositioning. BMC Bioinformatics 17: 78–83.

Buckley CE, Marguerie A, Roach AG, Goldsmith P, Fleming A, Aldreton WK *et al.* (2010). Drug reprofiling using zebrafish identifies novel compounds with potential pro-myelination effects. Neuropharmacology 59: 149–159.

Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Mills Shaw KR, Ozenberger BA *et al.* (2013). The Cancer Genome Atlas Pan-Cancer Analysis project. Nat Genet 45: 1113–1120.

Cowley GS, Weir BA, Hahn WC (2014). Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. Sci Data 1: 14035. https://doi.org/10.1038/sdata.2014.35.

Deshmukh VA, Tardif V, Lyssiotis CA, Green CC, Kerman B, Kim HJ (2013). A regenerative approach to the treatment of multiple sclerosis. Nature 502: 327–350.

Eder J, Sedrani R, Wiesmann C (2014). The discovery of first-in class drugs: origins and evolution. Nat Rev Drug Discov 13: 577–587.

Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H *et al.* (2014). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res 43: D805–D811.

Fowler KD, Funt JM, Artyomov MN, Zeskind B, Kolitz SE, Towfic F (2015). Leveraging existing data sets to generate new insights into Alzheimer's disease biology in specific patient subsets. Sci Rep 5: 14324–14332.

Fransecky L, Mochmann LH, Baldus CD (2015). Outlook on PI3K/AKT/mTOR inhibition in acute leukemia. Mol Cell Ther 3: 2–19.

Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A *et al.* (2012). ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 40: D1100–D1107.

Geva M, Kusko R, Soares H, Fowler KD, Birnberg T, Barash S *et al.* (2016). Pridopidine activates neuroprotective pathways impaired in Huntington disease. Hum Mol Genet . https://doi.org/10.1093/hmg/ddw238.

Giovannoni G, Baker D, Schmierer K (2015). The problem with repurposing: is there really an alternative to Big Pharma for developing new drugs for multiple sclerosis? Mult Scler Relat Disord 4: 3–5.

Gottlieb A, Stein GY, Ruppin ER, Altman RB, Sharan R (2011). PREDICT: a method for inferring novel drug indications with application to personalized medicine. Mol Syst Biol 7: 496–505. https://doi.org/10.1186/1741-7015-11-194.

Hewett M (2002). PharmGKB: the Pharmacogenetics Knowledge Base. Nucleic Acids Res 30: 163–165.

Horvath P, Aulner N, Bickle M, Davies AM, Del Nery E, Ebner D *et al.* (2016). Screening out irrelevant cell-based models of disease. Nat Rev Drug Discov 15: 715–769.

Howard R, McShane R, Lindesay J, Ritchie C, Baldwin A, Barber *et al.* (2012). Donepezil and memantine for moderate-to-severe Alzheimer's disease. NEJM 366: 893–903.

International Cancer Genome Consortium, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C *et al.* (2010). International network of cancer genome projects. Nature 464: 993–998.

Iorio F, Rittman T, Ge H, Menden M, Saez-Rodriguez J (2013). Transcriptional data: a new gateway to drug repositioning? Drug Discov Today 18: 350–357.

Iorio F, Bosotti R, Scacheri E, Belcastro V, Mithbaokar P, Ferriero R *et al.* (2010). Discovery of drug mode of action and drug repositioning from transcriptional responses. Proc Natl Acad Sci U S A 107: 14621–14626.

Jia Z, Liu Y, Guan N, Bo X, Luo Z, Barnes MR (2016). Cogena, a novel tool for co-expressed gene-set enrichment analysis, applied to drug repositioning and drug mode of action discovery. BMC Genomics 17: 414.

Kim MS, Pinto SM, Derese G, Nirujogi RS, Manda SS, Raghothama C *et al.* (2014). A draft map of the human proteome. Nature 509: 575–581.

Kolesnikov N, Hastings E, Keays M, Melnichuk O, Tang YA, Williams E *et al.* (2015). ArrayExpress update–simplifying data submissions. Nucleic Acids Res 43: D1113–D1116.

Kuhn M (2016). The SIDER database of drugs and side effects. Nucleic Acids Res 44 (Database issue): D1075–D1079.

Lamb J, Crawford ED, Peck D, Modell JW, Blatt IC, Wrobel MJ *et al.* (2006). The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science 313: 1929–1935.

Lappalainen I, Lopez J, Skipper L, Hefferon T, Spalding JD, Garner J *et al.* (2013). DbVar and DGVa: public archives for genomic structural variation. Nucleic Acids Res 41: D936–D941.

Lariosa-Willingham KD, Rosler ES, Tung JS, Dugas JC, Collins TL, Leonoudakis D (2016). A high throughput drug screening assay to identify compounds that promote oligodendrocyte differentiation using acutely dissociated and purified oligodendrocyte precursor cells. BMC Res Notes 9: 419–433.

Lonsdale J, Jeffrey T, Mike S, Rebecca P, Edmund L, Saboor S *et al.* (2013). The Genotype-Tissue Expression (GTEx) project. Nat Genet 45: 580–585.

Louhimo R, Laakso M, Belitskin D, Klefström J, Lehtonen R, Hautaniemi S (2016). Data integration to prioritize drugs using genomics and curated data. BioData Min 9: 21.

Mailman MD, Michael F, Yumi J, Masato K, Kimberly T, Rinat B *et al.* (2007). The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 39. https://doi.org/10.1038/ng1007-1181.

Martorana A, Perricone U, Lauria A (2016). The repurposing of old drugs or unsuccessful lead compounds by in silico approaches: new advances and perspectives. Curr Top Med Chem 16: 2088–2106.

Mattingly CJ, Rosenstein MC, Colby GT, Forrest JN Jr, Boyer JL (2006). The comparative toxicogenomics database (CTD): a resource for comparative toxicological studies. J Exp Zool A Comp Exp Biol 305: 689–692.

Mei F, Fancy SPJ, Shen Y-A, Niu J, Zhao C, Presley B *et al.* (2014). Micropillar arrays as a high-throughput screening platform for therapeutics in multiple sclerosis. Nat Med 20: 954–960.

Moosavinasab S, Patterson J, Strouse R, Rastegar-Mojarad M, Regan K, Payne PRO *et al.* (2016). RE:fine drugs: an interactive dashboard to access drug repurposing opportunities. Database (Oxford) . https://doi.org/10.1093/database/baw083.

Morgan P, Van Der Graaf PH, Arrowsmith J, Feltner DE, Drummond KS, Wegner CD *et al.* (2012). Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving phase II survival. Drug Discov Today 17: 419–424.

Najm FJ, Madhavan M, Zaremba A, Shick E, Karl RT, Factor DC (2015). Drug-based modulation of endogenous stem cells promotes functional remyelination in vivo. Nature 522: 216–235.

Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y *et al.* (2015). The support of human genetic evidence for approved drug indications. Nat Genet 47: 856–860.

Oprea TI, Bauman JE, Bologa CG, Buranda T, Chigaev A, Edwards BS *et al.* (2011). Drug repurposing from an academic perspective. Drug Discov Today Ther Strateg 8: 61–69.

Oprea TI, Mestres J (2012). Drug repurposing: far beyond new targets for old drugs. AAPS J 14: 759–763.

Paik H, Chung AY, Park HC, Park RW, Suk K, Kim J *et al.* (2015). Repurpose terbutaline sulfate for amyotrophic lateral sclerosis using electronic medical records. Sci Rep 5: 8580–8588.

Pandey UB, Nichols CD (2011). Human disease models in *Drosophila melanogaster* and the role of the fly in therapeutic drug discovery. Pharmacol Rev 63: 411–436.

Pantelopoulos A, Bourbakis NG (2010). A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans Syst Man Cybern C Appl Rev 40: 1–12.

Pihan E, Colliandre L, Guichou JF, Douguet D (2012). E-Drug3D: 3D structure collections dedicated to drug repurposing and fragment-based drug design. Bioinformatics 28: 1540–1541.

Piñero J, Queralt-Rosinach N, Bravo A, Deu-Pons J, Bauer-Mehren A, Baron M *et al.* (2015). DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. Database (Oxford) 2015 bav028. doi: https://doi.org/10.1093/database/bav028.

PsychENCODE Consortium, Akbarian S, Liu C, Knowles JA, Vaccarino FM, Farnham PJ *et al.* (2015). The PsychENCODE project. Nat Neurosci 18: 1707–1712.

Rastegar-Mojarad M, Zhan Y, Kolesar JM, Hebbring SJ, Lin SM (2015). Opportunities for drug repositioning from phenome-wide association studies. Nat Biotechnol 33: 342–345.

Rees MG, Seashore-Ludlow B, Cheah JH, Adams DJ, Price EV, Gill S *et al.* (2016). Correlating chemical sensitivity and basal gene expression reveals mechanism of action. Nat Chem Biol 12: 109–116.

Sanseau P, Agarwal P, Barnes MR, Pastinen T, Richards JB, Cardon LR *et al.* (2012). Use of genome-wide association studies. Nat Biotechnol 30: 317–320.

Schuster D, Laggner C, Langer T (2005). Why drugs fail – a study on side effects in new chemical entities. Curr Pharm Des 11: 3545–3559.

Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME *et al.* (2015). Harnessing connectivity in a large-scale small-molecule sensitivity dataset. Cancer Discov 5: 1210–1223.

Setoain J, Franch M, Martínez M, Tabas-Madrid D, Sorzano COS, Bakker A *et al.* (2015). NFFinder: an online bioinformatics tool for searching similar transcriptomics experiments in the context of drug repositioning. Nucleic Acids Res 43: W193–W199.

Sherry ST (2001). dbSNP: The NCBI Database of Genetic Variation. Nucleic Acids Res 29: 308–311.

Shineman DW, Alam J, Anderson M, Black SE, Carman AJ, Cummings JL *et al.* (2014). Overcoming obstacles to repurposing for neurodegenerative disease. Ann Clin Transl Neurol 1: 512–518. https://doi.org/10.1002/acn3.76.

Sirota M, Dudley JT, Kim J, Chiang AP, Morgan AA, Sweet-Cordero A *et al.* (2011). Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med 3 96ra77. https://doi.org/10.1126/scitranslmed.3001318.

Southan C, Sharman JL, Benson HE, Faccenda E, Pawson AJ, Alexander SP *et al.* (2016). The IUPHAR/BPS Guide to PHARMACOLOGY in 2016: towards curated quantitative interactions between 1300 protein targets and 6000 ligands. Nucleic Acids Res 44: D1054–D1068.

Swinney DC, Anthony J (2011). How were new medicines discovered? Nat Rev Drug Discov 10: 507–519.

Tan F, Yang R, Xu X, Chen X, Wang Y, Ma H *et al.* (2014). Drug repositioning by applying 'expression profiles' generated by integrating chemical structure similarity and gene semantic similarity. Mol Biosyst 10: 1126–1138.

Teachey DT, Sheen C, Hall J, Ryan T, Brown VI, Fish J *et al.* (2008). mTOR inhibitors are synergistic with methotrexate: an effective combination to treat acute lymphoblastic leukemia. Blood 112: 2020–2023.

The ENCODE Project Consortium (2004). The ENCODE (ENCyclopedia of DNA Elements) project. Science 306: 636–640.

Valentini G, Armano G, Frasca M, Lin J, Mesiti M, Re M (2016). RANKS: a flexible tool for node label ranking and classification in biological networks. Bioinformatics 32: 2872–2874. https://doi.org/10.1093/bioinformatics/btw235.

Vidović D, Koleti A, Schürer SC (2014). Large-scale integration of small molecule-induced genome-wide transcriptional responses, kinome-wide binding affinities and cell-growth inhibition profiles reveal global trends characterizing systems-level drug action. Front Genet 5: 342.

Volberding PA, Lagakos SW, Koch MA, Pettinelli C, Myers MW, Booth DK *et al.* (1990). Zidovudine in asymptomatic human immunodeficiency virus infection. A controlled trial in persons with fewer than 500 CD4-positive cells per cubic millimeter. The AIDS Clinical Trials Group of the National Institute of Allergy and Infectious Diseases. N Engl J Med 322: 941–949.

von Eichborn J, Murgueitio MS, Dunkel M, Koerner S, Bourne PE, Preissner R (2011). PROMISCUOUS: A database for network-based drug-repositioning. Nucleic Acids Res 39: D1060–D1066.

von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M *et al.* (2005). STRING: known and predicted protein–protein associations, integrated and transferred across organisms. Nucleic Acids Res 33: D433–D437.

Woodcock J (2014). Three encouraging steps toward new antibiotics. Available at: https://blogs.fda.gov/fdavoice/index.php/2014/09/three-encouraging-steps-towards-new-antibiotics (Accessed on March 6 2017).

Woolf SH (2008). The meaning of translational research and why it matters. JAMA 299: 211–213.

Xiao L, Xu H, Zhang Y, Wei Z, He J, Jiang W *et al.* (2008). Quetiapine facilitates oligodendrocyte development and prevents mice from myelin breakdown and behavior changes. Mol Psychiatry 13: 697–708.

Xu H, Aldrich MC, Chen Q, Liu H, Peterson NB, Dai Q *et al.* (2014). Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. J Am Med Inform Assoc 22: 179–191.

Yang L, Agarwal P (2011). Systematic drug repositioning based on clinical side-effects. PLoS One 6: e28025.

Ye H, Liu Q, Wei J (2014). Construction of drug network based on side effects and its application for drug repositioning. PLoS One 9: e87864.

Yong PFK, D'Cruz DP (2008). Mycophenolate mofetil in the treatment of lupus nephritis. Biologics 2: 297–310.

Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S *et al.* (2015). DSigDB: drug signatures database for gene set analysis. Bioinformatics 31: 3069–3071.

Yu L, Ma X, Zhang L, Zhang J, Gao L (2016). Prediction of new drug indications based on clinical data and network modularity. Sci Rep 6: 32530. https://doi.org/10.1038/srep32530.

Zhou H, Gao M, Skolnick J (2015). Comprehensive prediction of drug–protein interactions and side effects for the human proteome. Sci Rep 5: 11090.

Zu S, Chen T, Li A (2015). Global optimization-based inference of chemogenomic features from drug–target interactions. Bioinformatics 31: 2523–2529.