Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Pharmacology

Next-generation drug repurposing using human genetics and network biology

Serguei Nabirotchkin[1,5], Alex E Peluffo[2,5], Philippe Rinaudo[2], Jinchao Yu[2], Rodolphe Hajj[3] and Daniel Cohen[4]

Drug repurposing has attracted increased attention, especially in the context of drug discovery rates that remain too low despite a recent wave of approvals for biological therapeutics (e.g. gene therapy). These new biological entities-based treatments have high costs that are difficult to justify for small markets that include rare diseases. Drug repurposing, involving the identification of single or combinations of existing drugs based on human genetics data and network biology approaches represents a next-generation approach that has the potential to increase the speed of drug discovery at a lower cost. This *Pharmacological Perspective* reviews progress and perspectives in combining human genetics, especially genome-wide association studies, with network biology to drive drug repurposing for rare and common diseases with monogenic or polygenic etiologies. Also, highlighted here are important features of this next generation approach to drug repurposing, which can be combined with machine learning methods to meet the challenges of personalized medicine.

Addresses
[1] Network Biology & Drug Discovery Department, Pharnext, 11 rue René Jacques, 92130 Issy-les-Moulineaux, France
[2] Data Science Department, Pharnext, 11 rue René Jacques, 92130 Issy-les-Moulineaux, France
[3] Preclinical Research and Pharmacology Department, Pharnext, 11 rue René Jacques, 92130 Issy-les-Moulineaux, France
[4] Chief Executive Officer, Pharnext, 11 rue René Jacques, 92130 Issy-les-Moulineaux, France

Corresponding author: Peluffo, Alex E (apeluffo@pharnext.com)
[5] These authors contributed equally.

Current Opinion in Pharmacology 2020, **51**:78–92

This review comes from a themed issue on **Pharmacological Perspective**

Edited by **Michael Williams** and **David Trist**

For a complete overview see the Issue and the Editorial

Available online 22nd January 2020

https://doi.org/10.1016/j.coph.2019.12.004

Introduction

While U.S. Food and Drug Administration (FDA) drug approvals have increased approximately twofold in the past two years from a historical average of 26.8 per year [1,2], to 46 and 59 drugs, small molecules and biologics, being approved by the FDA in 2017[6] and 2018[7] respectively, the success rate of clinical trials continues to be low. Based on data from 185 000 trials involving 21 000 compounds, Wong *et al.* [3] estimated that between 2000 and 2015, the overall probability of success of candidates from Phase I to approval was 3.4% for oncology, 15% for Central Nervous System, 20% for Metabolic/Endocrinology/Cardiovascular and 26% for infectious diseases. The overall probability of success for orphan drugs to treat rare diseases was 6.2%, less than half the average of 14% for all candidates. Biologics represent a consistent element of the increased approvals, from an average of 9.4 entities per year over the five years to 2017 to 17 entities in both 2017 and 2018.

Nonetheless, the overall rate at which drugs are coming to the market remains slow [4]. Additionally, while biologics, antibodies, gene and cell therapies [5,6], often show remarkable clinical results, they require immense investments in R&D which makes them among the most expensive drugs on the market [7–9] often contributing to financial toxicity for patients [10]. In this context, for all types of diseases, but especially for diseases with low prevalence and consequently financially unappealing markets, there is a growing need for efficient, competitive, and complementary drug development strategies.

Drug repurposing is a solution that comes with reduced safety concerns and R&D costs as the drugs being repurposed have already been approved with often thorough post-market surveillance data or well-characterized during clinical development. Repurposing success stories include the phosphodiesterase 5 inhibitor, sildenafil — repurposed from angina, for which it was never approved, to erectile dysfunction and later to pulmonary hypertension [11], the cyclooxygenase 2 inhibitor, celecoxib — repurposed from inflammation to treat familial adenomatous polyps, the lymphocyte migration

[6] https://www.fda.gov/drugs/new-drugs-fda-cders-new-molecular-entities-and-new-therapeutic-biological-products/novel-drug-approvals-2017.
[7] https://www.fda.gov/drugs/new-drugs-fda-cders-new-molecular-entities-and-new-therapeutic-biological-products/novel-drug-approvals-2018.

modulator, fingolimod — repurposed from transplant rejection to the treatment of multiple sclerosis and the immunomodulator, thalidomide which was successfully repurposed from treating morning sickness and leprosy to the treatment of multiple myeloma [12–14]. More recently, the 'mast cell stabilizer', cromolyn sodium, approved for the treatment of mastocytosis, and asthma, has been reported to have potential in the treatment of amyotrophic lateral sclerosis [15].

Drug repurposing programs have benefited from substantial investments in the past [13,16], and although the repurposing of an existing marketed drug may be a challenge when it comes to securing market exclusivity [16], there are currently numerous financial and development program incentives from governments to encourage repurposing, especially for orphan diseases [16,17].

Identifying new uses for existing drugs can be facilitated by evidence that can link them to known or yet undiscovered drug targets and human disease states to develop new therapeutic indications.

Genome-Wide Association Studies (GWASs) [18] that involve the use of human genetic data to link genes to specific diseases have already resulted in candidate targets for drug discovery and repurposing. However, a major theme in pharmaceutical R & D over the past three decades has been a focus on developing therapeutics based on a single drug acting at a single target [19]. Conceptually this offers a means to improves efficacy while reducing side effects, an approach that is considered to be inefficient [19,20]. In this context, GWAS studies usually identify more than one associated gene per disease [18]. Advances in genotyping technology, the use of larger sample sizes, and the advent of network biology together with fundamental discoveries regarding the nature of biological networks, and their underlying pathways have led to evidence that, in many instances, human diseases arise from a series of perturbations which result in several pathways becoming unbalanced, rather than by the disruption of a single element within a single pathway. It is thus becoming clear that combining multiple drugs in one treatment can yield improved therapeutics, especially when each drug in a combination targets a different signaling cascade. These combinations may act additively or in synergy, in which case side effects can be minimized using lower doses.

This *Pharmacological Perspective* reviews the use of GWAS and network biology approaches to associate well established and novel targets with disease phenotypes to repurpose previously approved, or yet to be approved, drugs for new indications. It also discusses how these approaches can be combined with machine learning/artificial intelligence methods to yield a new generation of repurposed drugs.

A short introduction to genome-wide association studies for drug development

GWAS studies rely on genome-wide genotyping of hundreds to millions of individuals to statistically link genetic variation to diseases or diseases-related-phenotypes [18]. A 'big data' extension of linkage studies between genetic loci and human diseases, the first GWAS became possible in the mid-00s following the early efforts in human genome mapping and sequencing [21–23][8] . The main output of a GWAS is a list of all genome-wide-genotyped genes and the p-value of their association to the disease or phenotype of interest. Usually, the inputs are, for each individual, 1) a series (100 000 to greater than 1 000 000) of single nucleotide polymorphisms (SNPs) which serve as markers that are assumed to be physically or genetically linked (i.e. linkage disequilibrium) with the disease-casual genes, and 2) phenotypic information, for example, 'diseased' or 'healthy' state for case-control studies or quantitative information such as blood pressure, body weight, disease biomarker levels, drug response, and so on, linked to diseases. The output is usually obtained through several data processing steps to select genetic variation which is expected to correlate with the disease rather than with other processes which could bias the results and create spurious associations, such as genetic variation due to population structure, differences in ethnicity or natural selection (i.e. not at an Hardy-Weinberg equilibrium). Data processing is followed by statistical association calculation (p-values) between each SNP and the phenotype using simple machine-learning approaches (e.g. logistic regression, linear regression).

It is also possible to use the opposite approach. Instead of correlating numerous SNPs to one specific phenotype, the effect of one or a few SNPs, from one or a few genes, on hundreds of phenotypes can be studied. This approach is termed PheWAS (Phenome-Wide Association Study) and was recently used by Denny *et al.* [24] to verify independent associations between five SNPs and several complex diseases (e.g. coronary artery disease, Crohn's disease, multiple sclerosis) using over 200 000 diagnostic codes from electronic medical records. In addition to confirming the link between the targeted SNPs and the diseases, they discovered nineteen previously unknown statistical associations between the few SNPs and the medical records, thereby demonstrating the power of the inverse approach.[9]

Another critical step in GWASs is the mapping of each SNP to the gene it affects. This is not always as

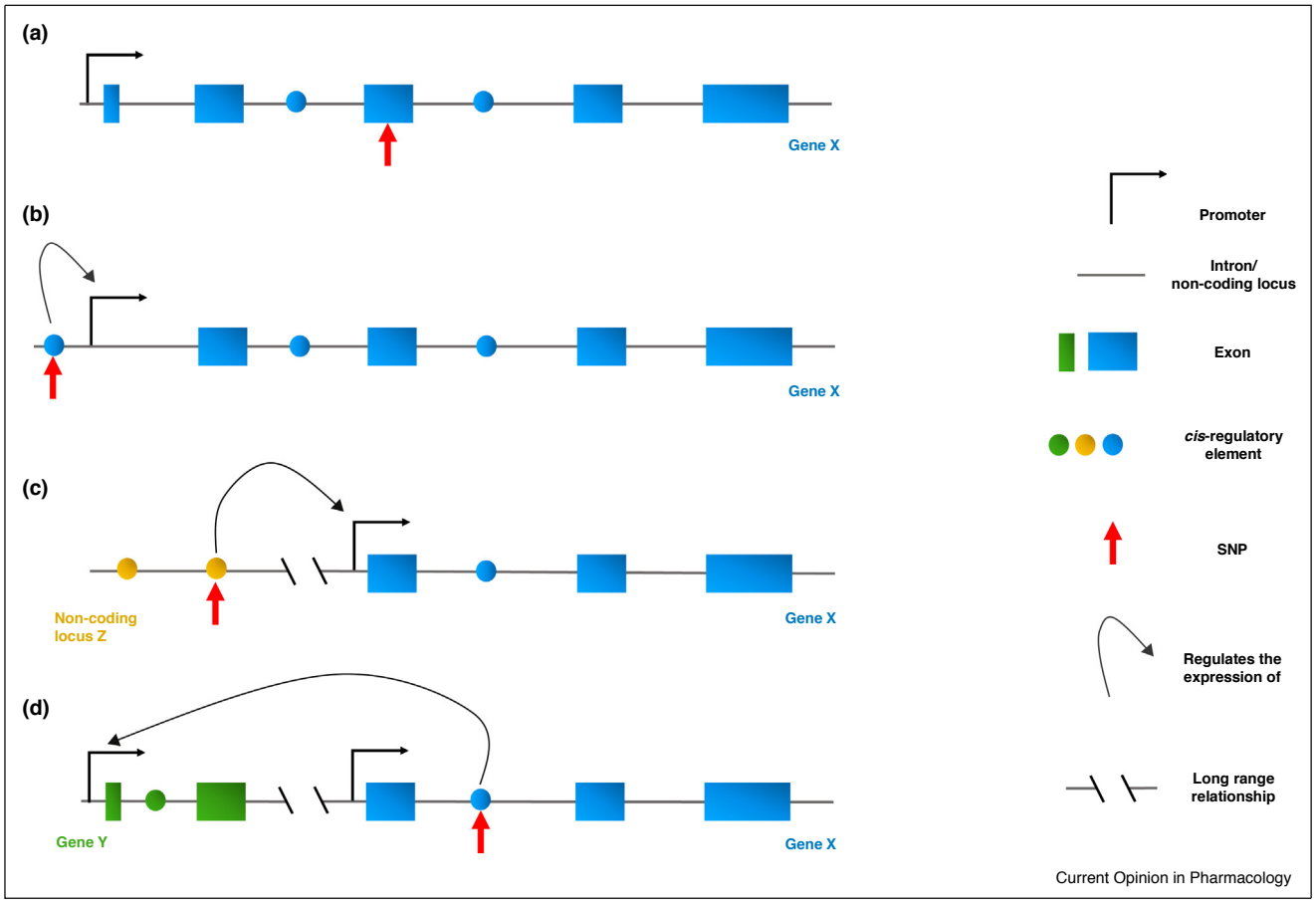
[8] For a first introduction of the GWAS concept to human genomic data see also U.S. Patent No. US 6,537,751 B1, Cohen et al. "Biallelic markers for use in constructing a high-density disequilibrium map of the human genome", filed Oct. 20th, 1999.

[9] For an early approach to PheWAS, see also U.S. Patent No. US 2002/0081584A, Blumenfeld *et al.* '"Genes, protein and biallelic markers related to central nervous system disease", filed Oct. 12th, 1999.

straightforward as checking whether the SNP is within the genomic locus of a gene, mainly because the candidate SNP may be located in an intergenic region or within a distant gene regulatory element or even within the intron of another gene [25–27] (Figure 1). When multiple SNPs map to one gene, several methods are available to aggregate the results into a single association value at the level of the gene, some of which account for the gene length, the cis-regulatory context, and so on. The corresponding *P*-values are compared to a multiple-comparisons test, usually Bonferroni, adjusted threshold (see Refs. [28–30]). Ultimately, the leverageable output of a GWAS is a list of candidate genes, the statistical support for each of these, and the calculated effect of each underlying SNPs in the sample population, on the studied phenotypic trait/disease.

In the early days of GWAS, reproducibility was a significant challenge because of limited sample sizes: a small number of individuals (100–1000) and a low number of SNPs per genome (0.1–0.5 million). Now, new government and private healthcare sponsored biobanks, where sample sizes are often higher than 100 000 individuals, have opened a new era in the genetics of human complex diseases [31]. The number of SNPs has also increased to several million using missing genotype imputation methods that use machine learning and whole-genome sequence references panels, for example, the 1000 Genomes Project [32] or the 65 000 haplotypes from the Haplotype Reference Consortium [33], to infer the genotype of individuals at ungenotyped SNPs based on their similarity to the genomes of the reference panel. However, a major limitation is the transferability of GWAS results from one ethnic background to another. Most GWASs, reference panels and methods have been established and developed using European ancestry populations [34]. Based on worldwide whole-genome studies, it is estimated that a quarter of a population's

Figure 1



The difficulty of mapping single nucleotide polymorphisms (SNPs) to their respective genes. It can be straightforward to link a SNP to the gene it affects when it occurs in the exon of that gene **(a)**. It becomes increasingly difficult when the SNP occurs in a *cis*-regulatory element. When the cis-regulatory element is located next to the gene it regulates, linking the SNP to its gene by physical proximity is still valid **(b)**. However, if the SNP is located several megabases upstream of the gene it regulates, in a non-coding locus **(c)** or within the intron of another gene **(d)**, mapping the SNP to its genes requires other approaches than methods based on physical proximity, such as combining genome-wide genotyping with RNA sequencing to link each SNP to gene-expression profiles.

genetic variation is unique to that population or to populations from the same continent [35]. That is why some genes are linked to diseases in specific populations only. For instance, variants in *APOL1* (Apolipoprotein L1) are associated with kidney disease in African descent only, which is thought to be the result of linked selection at nearby variants in *MYH9* (Myosin Heavy Chain 9) which protect against malaria [36]. Other examples include *MAPT* (microtubule-associated protein tau) variants linked to Parkinson's disease and *TREM2* (triggering receptor expressed on myeloid cells 2) variants involved in Alzheimer's disease in populations of Caucasian descent only [37,38].

Although a limitation when not carefully considered, genetic diversity between populations must be seen as an advantage in genetically-driven pharmacological research. One way of leveraging this diversity is by combining data obtained from different populations, in a PheWAS study framework, to test the clinical effect of a drug without going through a clinical trial. For instance, several GWASs reported an association between *PLA2G7* and coronary heart disease in populations with European ancestry, a biologically relevant association since the protein product of *PLA2G7* is a lipoprotein-associated phospholipase A_2 (Lp-PLA_2) which produces pro-inflammatory products implicated in endothelial dysfunction and which modulates the effects of platelet-activating factor [39]. Following the failure of inhibitors of Lp-PLA_2 in the treatment of coronary heart disease in U.S. and European Phase III clinical trials, it was found that the 10% of the individuals in the China Kadoorie Biobank (100 000 individuals in total at the time) carry a functional loss-of-function variant in the *PLA2G7* gene, which mimics the effect of Lp-PLA_2 inhibition, and that this variant was not associated with variations in major vascular events — vascular death, myocardial infarction, stroke, blood pressure, adiposity, blood glucose or lung function [40]. Thus, the failure of Lp-PLA_2 inhibitor in the clinic could have been predicted since 10% of Chinese individuals, who exhibit a loss-of-function variant in the *PLA2G7* gene and do not produce this protein, have the same risk of coronary heart disease. In this case, genetic diversity, the finding of variants that only exist in specific populations, allows for a thorough understanding of the functional role of a particular gene and can, therefore, be useful for prioritizing candidate drug targets.

GWAS-associated genes as drug targets in drug discovery and repurposing for complex and monogenic diseases

By combining most of the available data: 7000 genes and 2500 diseases forming 16 000 disease-gene associations, as well as 19 000 drug target-disease associations, Nelson *et al.* [41] found that selecting genetically supported targets can double the success rates of clinical development programs. Their conclusions were based on the finding that target genes for drugs approved in the US or the EU are enriched (more than expected by chance) among genes found in GWASs or other types of genetic studies. They found that 61% of approved drugs/target genes pairs are supported by at least one genetic association and 40% by more than five genetic associations with human diseases demonstrating the value and power of genetic data to successfully identify targets involved in human diseases which are likely to yield efficacious therapeutics. King *et al.* [42] performed the same analyses with additional data confirming that drugs with genetically supported targets were more likely to be successful in Phases II and III.

The propensity of genetic data to increase the success of drug development programs does not imply that other omics approaches should be ignored but rather that human genetic data should be seen as a particularly relevant starting point in drug development, followed (rather than preceded) by functional preclinical animal studies specifically designed to validate the genetically suggested candidate drug targets.

For novel drug discovery, multiple GWASs have either confirmed already suspected drug targets, such as *SLC30A8* for type 2 diabetes [43], PCSK9 for abdominal aortic aneurysm [44] and PDE3B for blood lipids and coronary diseases [45,46] or contributed to the identification of novel targets such as IL-23 for Crohn's disease [47,48] and psoriasis [49].

For drug repurposing, GWASs can link potential targets, one at the time, to several diseases and thereby suggest new indications for approved drugs. For instance, raloxifene, a selective estrogen receptor modulators (SERM), was initially developed for osteoporosis and then successfully repositioned to breast cancer, two indications genetically associated with estrogen receptors [12,50,51]. The most famous repurposed drug is the PDE5A inhibitor, sildenafil, initially developed for angina and successfully repurposed to erectile dysfunction [12]. In addition to confirming the genetic link between *PDE5A*, Coronary Artery Disease [52,53] and erectile dysfunction [54][10] in patient cohorts, recent GWASs have suggested that *PDE5A* is also associated to body-fat distribution, assessed by waist-to-hip ratio adjusted for body mass index [55] myeloid white blood cell count [56] and bone mineral density [57]. The EMBL-EBI ChEMBL database references 516 past or present clinical trials for sildenafil and other PDE5A inhibitors, for example, tadalafil including several Phase IV trials to assess efficacy in reducing obesity.[11] Another notable example is the tumor necrosis factor, *TNFSF11*, which has been genetically associated with monogenic familial osteopetrosis [58]

[10] At nominal value.
[11] https://www.targetvalidation.org/target/ENSG00000138735.

and bone mineral density [51,59] and against which an inhibitor, the monoclonal antibody, denosumab, has been successfully developed and approved in the treatment of osteoporosis. GWASs also reported an association between *TNFSF11* and inflammatory bowel disease suggesting denosumab as a potential therapeutic for Crohn's disease and ulcerative colitis [60,61].

Most GWASs have been performed on complex traits and diseases, that is, known or suspected to be under the control of many different genes (e.g. height, body mass index, blood lipids, blood pressure, Alzheimer's disease, coronary artery disease, type II diabetes, etc.); often, these complex diseases are also common diseases which affect an important part of the population. The GWAS framework can also be used for monogenic diseases, usually rare, by identifying multiple gene modifiers of those diseases. For example, by genotyping 7000 children with such disorders, plus an independent sample of 728 trios (child plus both parents), Niemi *et al.* [62] showed that common genetic variants contribute to the risk of rare severe neurodevelopmental disorders, thought to be monogenic, meaning that almost 10% of the variance in the risk of having the disease is attributable to inherited common genetic variation. The most common form of the Charcot-Marie-Tooth neuropathies, the type 1A (CMT1A), is thought to be monogenic and caused by a duplication of the peripheral myelin protein 22 (*PMP22*). Tao *et al.* [63,64] recently identified several genes involved in the severity of the CMT1A symptoms, including *SIPA1L2* for its effect on foot dorsiflexion. These studies suggest the possibility of leveraging GWASs results to develop drugs in the context of monogenic diseases.

Whether studying complex or monogenic diseases, in most cases, GWASs report SNP-associations with small effects (i.e. the mean phenotypic difference, or prevalence of the disease for case-control studies, between individuals carrying the mutation and those not carrying it, is small). It now appears clear that the effect of each SNP, that is, mutation, on a disease, is inversely proportional to its frequency in the population [65,66]. In other words, common mutations found in human populations are expected to have small effects on their bearers, while rare mutations appear to have stronger effects. It was elegantly demonstrated recently that this inverse relationship between frequency and effect is the result of negative natural selection on human populations that purge large effect mutations [67]. However, it is crucial to keep in mind that small phenotypic effects of genes and their underlying SNPs do not imply that drugs targeting the products of those genes will not have a strong effect on the disease [18,68]. Indeed, in most instances, a drug does not have the same effect on a gene-product than a single or even multiple SNPs. Thus, a disease-associated SNP found in patient cohorts may weakly impact the function of the gene product by lowering its expression or its enzymatic activity while a drug, like an inhibitor, may completely abolish its activity. It is essential to understand the association between SNPs, genes and the disease as an indication that the identified gene is a potential therapeutic target and not a quantitative predictor of the effect of a drug developed against that gene.

Targeting disease-associated protein-networks: using pleiotropy and molecular pathways to go beyond the 'one disease, one gene, one drug' paradigm

The mantra of 'one disease, one drug' or even that of 'one disease, one gene, one drug' has defined the drug discovery approach over the last three decades, especially in the context of the impact of GWASs and other genetic methods [69]. However, the fundamental genetic concept of pleiotropy, which, in current usage, refers to the action of one gene over several phenotypes [70], implies that a single gene and its variants can be involved in several diseases with different pathological phenotypes. The power of GWASs in drug repurposing is that they can identify links between genes and diseases, with no *a priori* assumption, and therefore independent, of *a posteriori* functional evidence. Recently, Finan *et al.* [71] found via a meta-analysis of numerous GWAS studies, that an important number of positions in the genome are individually associated to multiple diseases demonstrating that pleiotropy can be found at all levels of genetic studies, from large genomic loci to coding genes and single nucleotide changes. These findings are of especial interest in the context of drug repurposing as they show that the pleiotropic loci genetically associated with several different diseases are enriched with druggable genes (i.e. targets or homologs of targets of drugs in development or already approved). In recent years, GWASs have revealed what Visscher and Yang [72] termed a 'plethora of pleiotropy'. For instance, by compiling the results of GWAS studies over 42 human traits, Pickrell *et al.* [73] found that the *SLC39A8* gene that encodes for the Zinc transporter ZIP11 can influence seven traits including the risk of schizophrenia, height, and Parkinson disease. They also found that the lipid transport protein *APOE* is associated with waist-hip ratio, several lipid traits, especially cholesterol, and confirmed its well-established association with Alzheimer's disease [74,75]). Recently Watanabe *et al.* [76] analyzed more than 4000 publicly available GWASs and found that more than 90% of human genomic loci, 63% of human gene loci and 31% of GWAS-associated SNPs are pleiotropic, that is, they are associated to a trait/disease in at least two GWASs. Certainly, such a significant percentage of pleiotropic loci can be partly explained by the fact that those 'pleiotropic' genomic loci and genes contain *cis*-regulatory elements within their non-coding regions and that two SNPs physically located in the same gene may affect two

different traits via entirely different regulatory effects on different genes. A notable example is the gene *FTO* which has long been associated with obesity through genetic associations [77,78] and GWAS [79], hence the name 'Fat mass and obesity-associated protein'. Smemo *et al.* [27] found that the non-coding SNPs in an intron of *FTO*, which are associated with obesity, do not regulate the expression of *FTO* but rather that of the *IRX3* gene located several kilobases upstream of *FTO*, and which is known to play a role in neural development. Such results are a reminder that that functional characterization and validation of disease-associated SNPs remains a high priority in assessing GWASs results in the context of drug development and repurposing [80]. Nevertheless, 31% of pleiotropy at the SNP level implies that a single SNP, within a single gene, may affect several phenotypes and that, therefore, an important number of SNPs identified by GWASs may offer repurposing opportunities.

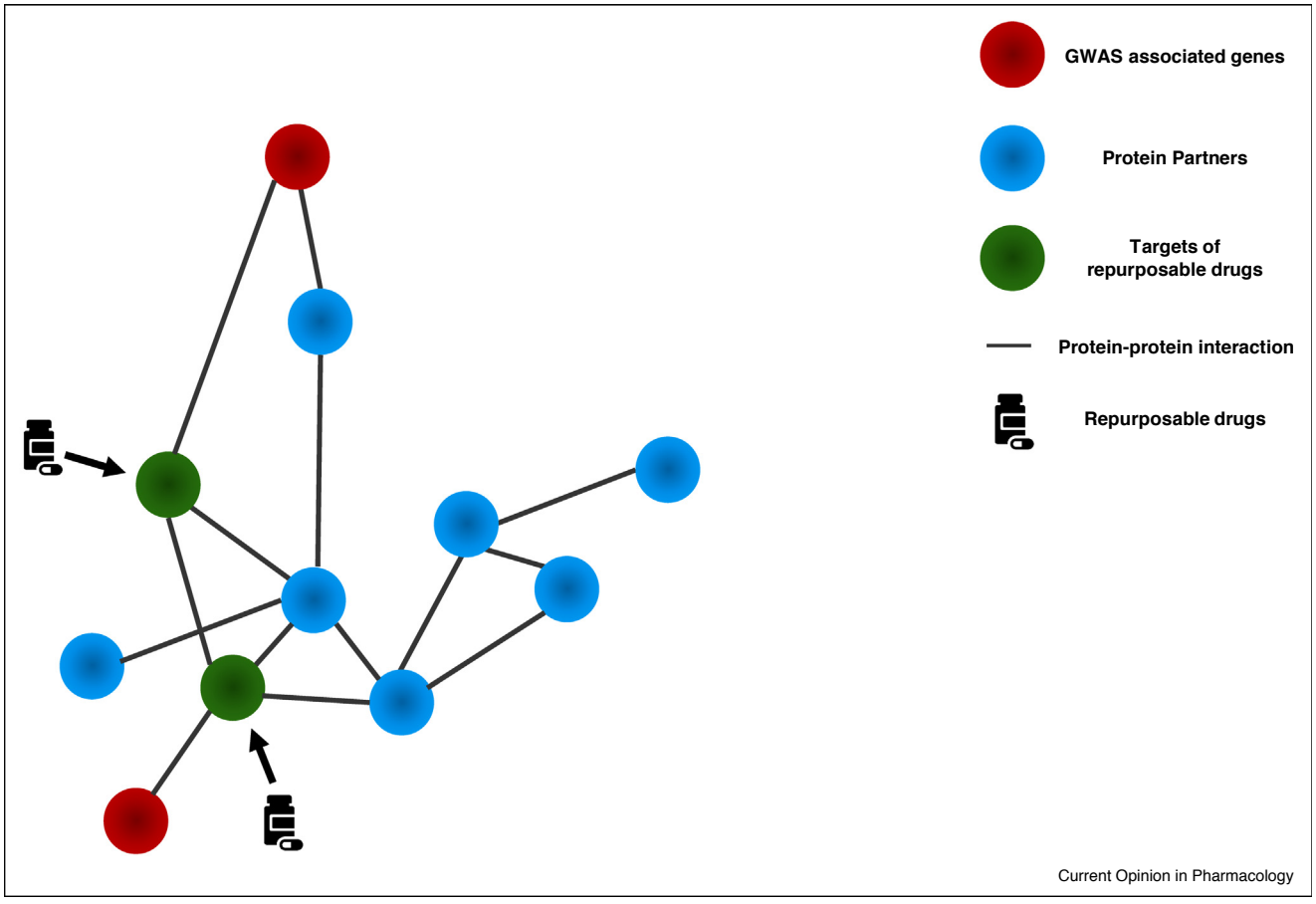
The pleiotropic nature of genes comes from the fact that gene-products are connected to each other by numerous mechanisms: DNA-protein, RNA–protein, metabolite–protein and protein–protein interactions integrated into so-called biological networks [81]. One gene may affect several diseases when its corresponding pathway participates in different types of physiological processes. This is exemplified by the multiple roles played by VEGF (vascular endothelial growth factor) receptors in several different diseases and several different tissues via diverse processes such as angiogenesis [82], macrophage migration [83], and neurogenesis [84] and by their association, in GWASs, to an important number of different phenotypes such as, for VEGFR-2, red and white blood cell count, systemic lupus erythematosus, height and body mass index, bone mineral density, cognitive performance, baldness [56,59,85–88]. The association of one receptor to numerous different diseases could be explained by the fact that perturbations, for example, mutations, in VEGFR-2, combine with different perturbations in other signaling pathways for each specific disease.

If findings regarding pleiotropy make it clear that one gene can often be associated with numerous diseases, another important series of findings make it clear that one disease is often associated with multiple genes and to the disruption of their respective, often inter-connected, pathways. Recent massive-scale genomics projects involving GWASs for complex diseases for hundreds of thousands of individuals have found that the genetic architecture of those diseases sometimes involves hundreds of genes [89,90]. In addition, as mentioned previously, even diseases thought to be monogenic are impacted by mutations at several genes, meaning that at the genetic level, for all types of diseases, polygenicity appears to be the rule. Moreover, fundamental biology research suggests that biological networks are extremely robust to perturbations which affect one node of a biological network, and therefore, that most diseases are unlikely to result from single mutation or environmental perturbations because there exist compensatory pathways or functional feed-back loops which can counterbalance or mitigate the advent of that perturbation [91–93]. In human populations, it was recently discovered that an individual carries an average of a hundred loss-of-function mutations and at least one recessive lethal mutation [94,95]. An important number of these loss-of-function mutations occur in genes with redundant functions, which suggests that molecular networks overcome single-node perturbations via some level of redundancy, whether via duplication of essential genes or compensatory pathways [96]. These considerations have triggered the advent of a series of approaches linked to network biology, the so-called 'Network Pharmacology' [69,97] and 'Network Medicine' [98]. These approaches combine human genetic data from studies such as GWASs with knowledge of functional interactions (e.g. protein–protein interactions) between associated and non-associated genes to identify and extend the repertoire of molecular pathways and new or repurposable drug targets. Considering an associated gene and its molecular partners, rather than only the gene itself, substantially extends the repertoire of potential drug-targets, and therefore, potential drugs (Figure 2). Following this approach, Okada *et al.* [99] performed a GWAS meta-analysis over 100 000 individuals and identified 42 novel rheumatoid arthritis (RA) risk loci in addition to the 59 already known. By mapping these 101 loci to their genes using a combination of genomic, transcriptomic and epigenomic data, they identified 98 RA risk genes and then used protein-protein interaction (PPI) databases to identify their direct PPI partners. Twenty height proteins encoded by genes associated to RA were found to interact directly with ten proteins that were targets of already approved RA drugs and with two proteins, *CDK6* and *CDK4* that are targetable by drugs with other disease indications, for instance, the CDK4/CDK6 cyclin-dependent kinase inhibitor, paldociclib (breast cancer and mantle cell lymphoma) and the CDK4 inhibitor/ tumor suppressor gene p16 modulator, capridine-beta (advanced solid tumors and psoriatic disorders), two promising drug repurposing candidates for RA.

Building human diseases networks with human genetic data and network biology

The construction of human disease networks is an obvious essential step of network pharmacology. Briefly, the biological network of a disease represents interactions between gene-products and cellular metabolites involved in the etiology and symptoms of that disease (see Refs. [98,100] for a definition). In general, the construction of the biological network of disease starts with the creation of a global molecular network derived from the results of large-scale experiments that have characterized the human interactome. Examples of key resources are the Human Protein

Figure 2



Combining GWAS data with Protein–Protein Networks to Identify Targets of Repurposable Drugs. If the disease-associated genes identified by GWASs are not targets of known drugs, it is possible to combine the GWASs results with molecular interaction data (especially protein–protein interactions) to integrate those genes in a biological network and identify direct or indirect partners of those genes, in the same pathways, which are targets of known repurposable drugs.

Reference Database (HPRD, http://www.hprd.org/), the Biological General Repository for Interaction Datasets (BioGRID, http://www.thebiogrid.org) and STRING (https://string-db.org/cgi/input.pl) for protein-protein interactions, the Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.genome.jp/kegg/) and OmniPath (http://omnipathdb.org/) for metabolic pathways; the Genotype-Tissue Expression (GTEx, https://www.gtexportal.org/home/) database is also a key resource to identify gene-regulatory interactions as well as the privately-owned Qiagen Ingenuity Pathway Analysis tool (https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/). The molecular disease network is then refined by reducing the global interactome to a subpart that contains: i) the genes associated to the disease and, ii) the key cellular partners of those genes [98]. The genes associated with the disease can be identified through functional studies and family linkage-studies, an important part of which is compiled by the Online Mendelian Inheritance in Man database (OMIM, https://www.omim.org/). The genes can also be identified using GWASs. The NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog, https://www.ebi.ac.uk/gwas/) combines all the statistically significant SNP-gene-disease associations (at $P < 10^{-5}$) for more than 150 000 studies. Similar results, combined with several of the resources mentioned above and other information such as drug data are also available via the public-private online initiative OpenTargets (https://www.opentargets.org/) that is sponsored by GSK, Sanofi, Biogen, Takeda, Celgene, the Sanger Institute and the EMBL-European Bioinformatics Institute.

The use of network biology data and methods also offers a way to rescue false-negative associations in GWASs, which may be the result of the stringent multiple testing

Bonferroni correction or of factors such as population heterogeneity, which impact the statistical power required to detect true associations. A GWAS is essentially an aggregation of thousands to millions of statistical tests (one for each SNP) and, therefore, the *P*-values of those tests need to be corrected for multiple testing. The Bonferroni correction simply multiplies the *P*-value of the test by the number of tests performed, and it yields a high rate of false-negative tests [100]. These false negatives are so many missed opportunities to identify potential drug targets. As a solution, it is possible to combine all nominally significant genes (without correction for multiple testing) rather than only the genes which are statistically associated at the Bonferroni corrected level, and to search for statistically significative enrichment of specific pathways containing those genes which could be involved in the etiology of the disease. Following this approach, Zipp *et al.* [101] performed a network-based pathway analysis for multiple sclerosis (MS), a complex inflammatory neurological disease, where they collected, from several GWASs, all the non-synonymous or potentially functionally deleterious SNPs associated to multiple sclerosis at nominal value and aggregated the results at the gene level. This dataset was then used together with a carefully curated collection of PPI databases to build a biological network connecting those genes and to search for subparts of that network significantly enriched with nominally significant genes with respect to the MS GWASs. They found that approximately 80 genes are arranged in a small protein-protein subnetwork, of which two-thirds had nominally significant *p*-values in independent meta-analyses of multiple MS GWASs. By performing a gene set enrichment analysis over those genes, which consists in annotating a list of genes with molecular/cellular/physiological/features (GeneOntology Terms, Functional Group, Tissue expression, etc.) and finding features which are statistically overrepresented [102,103], Zipp *et al.* [102] found that a third of the 80 genes that were highly expressed in immune-related cell types and half being highly expressed in the CNS, some of which, such as *S1PR1*, the gene that encodes for the Sphingosine-1-Phosphate Receptor 1 are known targets for MS. Greene *et al.* [104] developed a similar approach, termed 'NetWAS', which combines nominally significant genes from GWAS and tissue-specific networks built from protein-protein interaction databases, gene expression and transcription factor data. Using GWASs of hypertension, their method yielded tissue-specific networks (i.e. a list of GWASs nominally associated genes and their PPI partners in a specific tissue) that were highly enriched for the GeneOntology term 'regulation of blood pressure', at a higher frequency than the GWAS-alone top genes, and which contained an overrepresented proportion of targets of antihypertensive drugs from DrugBank (https://www.drugbank.ca/) and three other drug databases. NetWAS and several recent upgrades are now available as an online tool 'GIANT2' (http://giant-v2.princeton.edu/). In addition to NetWAS, there now exist several methods [105–107] to identify feature enrichment among gene lists, whether from GWAS or other omics experiments (e.g. RNAseq) and to 'prioritize' relevant genes for a phenotype of interest in the perspective of drug development (see Fine *et al.* [108] for a benchmarking of those methods).

Combinational repurposing via genetics-informed network pharmacology

The inescapable conclusion of the success of network biology methods is that diseases, like most complex biological phenomena, are better treated by targeting several nodes of their underlying biological networks. The fundamental biology findings regarding the robustness of biological networks, especially the fact that there exist, in most organisms, compensatory mechanisms against the perturbation of most canonical signaling pathways must be considered in assessing drug discovery and development programs. If a disease is unlikely to be the result of a single perturbation at the level of a single gene, any treatment against that disease is more likely to succeed when targeting several factors involved in that same disease.

For infectious diseases, there exist a few examples of combinational therapies, including HIV therapies [109] and the treatment of malaria with artemisinin-based combination therapies [110], which have demonstrated the benefits of combining small molecule therapeutics with Traditional Chinese Medicine remedies [111]. Between 1943 and 2018, the FDA approved 419 drug combinations from a total of 328 small molecules mostly for the nervous system (n = 78) and infectious (n = 74) disease indications [112] indicating that combinational treatments are a potentially fruitful approach but which remains overshadowed by monotherapy-based approaches. This is especially true in the context of drug repurposing. Recently, Sun *et al.*, in assessing FDA drug approvals between 2006 and 2015, found that combinations of repurposed drugs are more likely to succeed than single repurposed drugs [113]. Following this approach, Hung et al. [114] reported the successful results of a Phase IIb/III clinical trial combining repurposed clarithromycin, naproxen, and oseltamivir in the treatment of Influenza A (H3N2) infections. They nevertheless caution that such combinations, at high doses, given chronically, may induce severe potential side effects. Side effects are a serious safety concern for combinational treatments applied with high doses of each drug: when the curative effects of the drugs combine, so can their side effects.

Synergistic combinations

To overcome this issue, one strategy is to take advantage of the synergistic effect of the drug combo, that is, a greater effect than the sum of the effects of each

compound taken independently. In most cases, the synergy is thought to result from the targeting of different signaling pathways that cooperate in provoking the disease phenotypes, in which case therapeutic selectivity is increased, thus allowing the use of lower doses to restrict potential side effects [113,115,116]. Several synergistic repurposed combinations have been recently identified as promising against multidrug-resistant pathogens [117], viral infectious diseases [118], and cutaneous melanoma [119]. Importantly, the therapies mentioned above were identified by simple drug combination screening but were not explicitly driven by a knowledge of the disease biological networks. In addition, most of these initiatives have focused on infectious diseases. When the biological network of the disease can be built, it is possible to identify repurposable drugs that will act on several nodes of the disease network and to combine them in a network-guided rational manner. The repurposable drugs can either act on the same target which is part of two different diseases network (pleiotropy) or act on several targets, including one involved in the disease for which the drug was initially developed and one involved in the condition to which the drug is being repurposed (i.e. polypharmacology [120]).

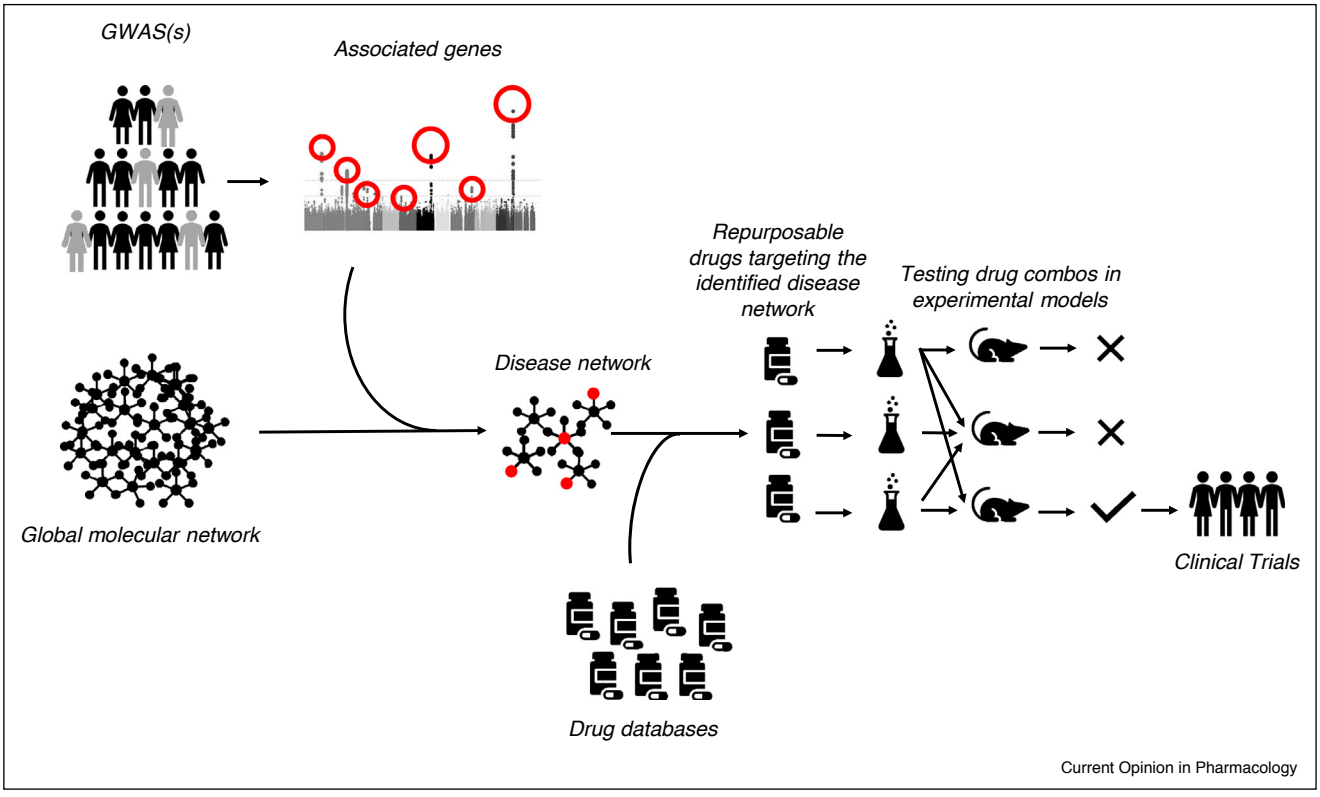
This proved to be an efficient approach in the treatment of the neuropathy Charcot-Marie-Tooth Type 1A (CMT1A), the most common subtype of a rare orphan genetic disease that affects an average of 1/2500 individuals. CMT1A results from a duplication of the gene *PMP22*, which leads to an overexpression-induced abnormal Schwann cell differentiation and dysmyelination [121]. Although a monogenic neuropathy, CMT1A involves a complex network of molecular interactions. *PMP22* was used as an initial seed of the network to derive a combination of three repurposed drugs: the $GABA_B$ agonist, baclofen, the opioid-receptor antagonist, naltrexone and D-sorbitol, a natural metabolite particularly known for its important role in energy metabolism, one of the processes deregulated in CMT1A [122]. A low-dose combination of these three drugs downregulated the over-expression of *PMP22* and improved the biological and behavioral phenotype of CMT1A rats [122]; the same low-dose combo also demonstrated a capacity to reduce CMT1A symptoms in human trials and no toxicity or side-effects when compared to a placebo [123]. A similar network-based combinational repurposing approach to Alzheimer's disease and Parkinson's disease, two complex neurological diseases, led to the identification of a synergic low-dose combination of repurposed baclofen and the acamprosate as highly promising based on experiments performed in cellular and animal models of those diseases [124–126]. Following a similar approach and using GWASs, Cheng *et al.* [127] recently focused on hypertension and cancer to predict efficient pairwise drug combinations based on systems biology analyses and especially on the overlap between protein-protein networks of disease genes and FDA-approved drug targets. Genes were linked to diseases using eight different sources, most of which rely on GWASs and other types of genetic studies, including the GWAS catalog, DisGeNET, a comprehensive database of disease-gene associations which compiles results from GWAS, model-organism studies and other inputs from the scientific literature (http://www.disgenet.org/) and the OMIM catalog. A key finding of Cheng *et al.* was that successful FDA-approved pairwise combinations are to be found when the subnetworks of both drugs overlap different parts of the disease subnetwork (which they term 'complementary exposure'), which increases therapeutic efficacy while limiting toxicity.

Collectively, these studies demonstrate the feasibility of a genetics-informed network-based repurposing and combinational repurposing approach in the treatment of a variety of diseases: pathogen-induced or not, rare or common, monogenic, or complex. This provides a global picture of a productive research model for combinational repurposing may look (Figure 3). From this, it is clear that a genetics and systems biology approach to combinational repurposing requires a combination of the power of big genetic data with network biology resources and methods, from protein to drug database inputs. The heuristic power of this approach is its ability to bridge *in silico* methodologies in a manner that accelerates and inform research via the use of specially designed validation experiments using animal models before bringing the candidate drug combinations to clinical trials. Having identified candidate combinations, the key remaining challenges are the selection of the doses to be used for the trials and the analyses of their efficacy. This can be especially challenging with low-dose repurposed combinations. However, several methods exist to maximize synergy while minimizing the side effects that take advantage of integrated development processes and their statistical analyses, using combinations of cell lines, model organisms, and patient studies [128,129].

Future aspects: combining GWAS, network biology, machine learning and artificial intelligence for the analysis of complex biological data

As machine learning comes of age [130–132], the GWAS plus network approach can be upgraded using several other inputs from typically 'big data' resources. Among them, deep learning-based natural language processing (NLP) method which allows better text mining of the scientific literature and patient medical records. As an example, Rajkomar *et al.* [133] analyzed more than 200 000 electronic health records from two US academic medical centers that represented more than 45 million data points. They predicted mortality, discharge diagnosis, length of stay, and readmission using deep learning NLP methods with high accuracy (i.e. an area under the operator curve between

Figure 3



Key steps in next-generation drug repurposing highlighted through one possible drug development scheme. A genome-wide association study (GWAS) or multiple GWASs is/are performed to find genes that are associated with the disease of interest. As explained, this approach is valid for both monogenic and polygenic diseases. In parallel, using other classes of functional data, especially protein–protein interactions, a global molecular network is built. Using GWAS associated genes (red circles), at a defined p-value threshold, possibly above the multiple correction adjusted value, together with the molecular network, the disease network can be built to identify druggable targets that are part of the same pathways as the genetically disease-associated genes. The disease network is then overlapped with databases of available drugs. When multiple repurposable drugs are identified, it is possible to combine them and test their synergistic as well as side effects in animal models. The animal studies-validated combo (compared to the drugs alone) can then be tested in human trials.

0.75 and 0.94). The public-private OpenTargets platform has recently released 'LINK', a knowledge graph tool of the Open Targets literature built using NLP (https://link.opentargets.io/). This AI-powered tool provides visually explorable links between genes, diseases, and drugs based on all available PubMed abstracts. For example, when the user inputs the gene, *BRCA1*, the tool outputs a graph with all genes, disease, drugs and key concepts related to *BRCA1* as nodes with a size-dependent indication of the importance of entity in the knowledge graph with the relationships between these entities shown as links between the nodes. This enables the interrogation of the details of each of the links to obtain the text on which the presented evidence can be based. Two options are particularly interesting in the perspective of combinational drug repurposing: i) the possibility to perform a search with multiple terms, for instance, a target, and two drugs; and ii) the possibility to initiate a bottom-up approach by searching all links to a specific paper.[12] Deep learning methods can also be used to explore pre-built disease networks. Recently, Zong *et al.* [134] adapted the *DeepWalk* algorithm, tested initially on social network data such as Flickr and YouTube, to predict novel associations between drugs and targets based on the exploration of a biological network built from drugs, diseases, and PPI databases. When the association between a target and a drug is unknown, for example, for new chemical entities, it is also possible to use deep learning in the form of a convolutional neural network to predict the docking of a protein with its ligand [135].

Conclusion

The current *Pharmacological Perspective* describes a drug discovery/drug repositioning methodology that integrates

[12] See http://blog.opentargets.org/2018/01/18/link/ for a detailed introduction.

Table 1

Four important features of next-generation drug repurposing. We present each feature next to the most relevant reference cited in this article

Feature	Relevance	Key references
Human Genetics + Network Biology	Statistical genetic association between genes and diseases yields a series of candidate genes that can be directly targeted by existing drugs or connected to druggable molecular-interaction partners in the same pathway. The approach can also be extended to diseases thought to be monogenic by identifying disease phenotypes or treatment response gene-modifiers.	Hopkins [69] Barabasi *et al.* [98] Okada *et al.* [99] Nelson *et al.* [41] Niemi *et al.* [62] Cheng *et al.* [127] Kim *et al.* [106]
Combinations	Combining repurposable drugs can increase their efficacy because several components of the disease network are targeted.	Lehar *et al.* [116] Sun *et al.* [113]
Synergy/Low-dose	When the effect of the drug combo on the disease is synergistic, e,g., greater than the sum of each individual drug alone (synergy) it is possible to use lower doses to minimize the side-effects while preserving efficacy.	Attarian *et al.* [123] Chumakov *et al.* [125] Prukop *et al.* [122] Hajj *et al.* [126]
Machine Learning/Artificial Intelligence	Machine learning and deep learning methods are transforming genetic analyses, the construction of biological networks for drug development and repositioning, as well as disease risk prediction using polygenic risk scores and natural language analyses of biological literature.	Zong *et al.* [134] Khera *et al.* [136] Vamathevan *et al.* [130]

human genetics, network biology, and network polypharmacology as a successful scheme for the next generation of drug repurposing efforts. Essential features for the next generation of drug repurposing approaches are leveraging human genetic data using GWASs to combine with the construction of the biological networks of human diseases to identify synergistic combinations of repurposable drugs. These can restore homeostasis to treat dysfunctional molecular pathways, which underpin most diseases while minimizing the side effects of these repurposed combinational treatments via the use of low doses, which remain efficient because of their synergy (Table 1). By relying on progress in big data analyses, machine learning and artificial intelligence methods, this ultimately allows for the identification of drug targets from human-specific disease pathways, increases the efficacy of clinical studies and is primed to meet the challenges of precision/personalized medicine, the development of preventive treatments and disease-modifying medications.

Conflict of interest statement

S.N., A.E.P, P.R., J.Y., R.H., and D.C. are employees of Pharnext S.A.

Author contribution

Serguei Nabirotchkin, Alex E. Peluffo, Philippe Rinaudo, Jinchao Yu, Rodolphe Hajj, and Daniel Cohen wrote the manuscript.

Acknowledgements

We thank Jan Bouaziz, Qibo Sun, Florian Leverve, Marina Ossokina, Sarra Seninet, Lydie Boussicault, Anthony Brureau, Mohamed Jabri, Fabrice Glibert, Célia Lokmane, Antoine Chouteau, Peter Schmitt, Julien Laffaire, as well as our colleagues at GeneNet Pharmaceuticals Co., Ltd and at Tasly Pharmaceutical Group Co., Ltd for fruitful discussions.

References

1. Munos B: **Lessons from 60 years of pharmaceutical innovation**. *Nat Rev Drug Discov* 2009, **8**:959-968.
2. Liu C, Constantinides PP, Li Y: **Research and development in drug innovation: reflections from the 2013 bioeconomy conference in China, lessons learned and future perspectives**. *Acta Pharm Sin B* 2014, **4**:112-119.
3. Wong CH, Siah KW, Lo AW: **Estimation of clinical trial success rates and related parameters**. *Biostatistics* 2019, **20**:273-286.
4. Kinch MS, Griesenauer RH: **2017 in review: FDA approvals of new molecular entities**. *Drug Discov Today* 2018, **23**:1469-1473.
5. Deverman BE, Ravina BM, Bankiewicz KS, Paul SM, Sah DWY: **Gene therapy for neurological disorders: progress and prospects**. *Nat Rev Drug Discov* 2018, **17**:641-659.
6. Yip A, Webster RM: **The market for chimeric antigen receptor T cell therapies**. *Nat Rev Drug Discov* 2018, **17**:161-162.
7. Garrison LP, Jackson T, Paul D, Kenston M: **Value-based pricing for emerging gene therapies: the economic case for a higher cost-effectiveness threshold**. *J Manag Care Spec Pharm* 2019, **25**:793-799.
8. Luzzatto L, Hyry HI, Schieppati A, Costa E, Simoens S, Schaefer F, Roos JCP, Merlini G, Kääriäinen H, Garattini S *et al.*: **Outrageous prices of orphan drugs: a call for collaboration**. *Lancet* 2018, **392**:791-794.
9. Prasad V: **Tisagenlecleucel - the first approved CAR-T-cell therapy: Implications for payers and policy makers**. *Nat Rev Clin Oncol* 2018, **15**:11-12.
10. Gilligan AM, Alberts DS, Roe DJ, Skrepnek GH: **Death or debt? National estimates of financial toxicity in persons with newly-diagnosed cancer**. *Am J Med* 2018, **131**:1187-1199.
11. Ghofrani HA, Osterloh IH, Grimminger F: **Sildenafil : from angina to erectile dysfunction to pulmonary hypertension and beyond**. *Nat Rev Drug Discov* 2006, **5**:689-702.

12. Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, Doig A, Guilliams T, Latimer J, McNamee C *et al.*: **Drug repurposing: progress, challenges and recommendations**. *Nat Rev Drug Discov* 2018, **18**:41-58
This thorough review on drug repurposing focuses on scientific approaches, including GWASs, but also financial, marketing and legal aspects.

13. Mullard A: **Drug repurposing programmes get lift off**. *Nat Rev Drug Discov* 2012, **11**:505-506.

14. Xue H, Li J, Xie H, Wang Y: **Review of drug repositioning approaches and resources**. *Int J Biol Sci* 2018, **14**:1232-1244
This review highlights interesting text mining (A.I methods) approaches to find repurposable drugs faster and more efficiently.

15. Granucci EJ, Griciuc A, Mueller KA, Mills AN, Le H, Dios AM, Mcginty D, Pereira J, Elmaleh D, Berry JD *et al.*: **Cromolyn sodium delays disease onset and is neuroprotective in the SOD1 ^{G93A} Mouse Model of amyotrophic lateral sclerosis**. *Sci Rep* 2019, **9**:17728.

16. Breckenridge A, Jacob R: **Overcoming the legal and regulatory barriers to drug repurposing**. *Nat Rev Drug Discov* 2018, **18**:1-2.

17. Hernandez JJ, Pryszlak M, Smith L, Yanchus C, Kurji N, Shahani VM, Molinski SV: **Giving drugs a second chance: overcoming regulatory and financial hurdles in repurposing approved drugs as cancer therapeutics**. *Front Oncol* 2017, **7**:1-8.

18. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J: **10 years of GWAS discovery: biology, function, and translation**. *Am J Hum Genet* 2017, **101**:5-22.

19. Roth BL, Sheffler DJ, Kroeze WK: **Magic shotguns versus magic bullets : selectively non-selective drugs for mood disorders and schizophrenia**. *Nat Rev Drug Discov* 2004, **3**:353-359.

20. Ramsay RR, Nikolic MRP, Nikolic K, Uliassi E, Bolognesi ML: **A perspective on multi - target drug discovery and design for complex diseases**. *Clin Transl Med* 2018, **7**:3 http://dx.doi.org/10.1186/s40169-017-0181-2 PMID: 29340951; PMCID: PMC5770353.

21. Cohen D, Chumakov I, Weissenbach J: **A first-generation physical map of the human genome**. *Nature* 1993, **366**:698-701.

22. Lander ESL, Birren LM, Nusbaum B, Zody C, Baldwin MC, Devon J, Dewar K, Doyle K, FitzHugh M, Funke W *et al.*: **Initial sequencing and analysis of the human genome**. *Nature* 2001, **409**:860-921.

23. Hebbring S: **Genomic and phenomic research in the 21st century**. *Trends Genet* 2019, **35**:29-41.

24. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, Wang D, Masys DR, Roden DM, Crawford DC: **PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations**. *Bioinformatics* 2010, **26**:1205-1210.

25. Albert FW, Kruglyak L: **The role of regulatory variation in complex traits and disease**. *Nat Rev Genet* 2015, **16**:197-212.

26. Spitz F, Furlong EEM: **Transcription factors: from enhancer binding to developmental control**. *Nat Rev Genet* 2012, **13**:613-626.

27. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gómez-Marín C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF *et al.*: **Obesity-associated variants within FTO form long-range functional connections with IRX3**. *Nature* 2014, **507**:371-375.

28. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, Hirschhorn JN: **Genome-wide association studies for complex traits: Consensus, uncertainty and challenges**. *Nat Rev Genet* 2008, **9**:356-369.

29. Bush WS, Moore JH: **Chapter 11: genome-wide association studies**. *PLoS Comput Biol* 2012, **8**.

30. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM: **A tutorial on conducting genome-wide association studies: quality control and statistical analysis**. *Int J Methods Psychiatr Res* 2018, **27**:e1608
This article provides a very thorough step-by-step guide to performing a GWAS with a simple laptop computer, assuming one already has genetic (SNPs) and phenotypic data. In addition, the paper details some approaches to genetics based disease risk prediction using Polygenic Risk Scores.

31. Khera AV, Kathiresan S: **Genetics of coronary artery disease: discovery, biology and clinical translation**. *Nat Rev Genet* 2017, **18**:331-344.

32. Gibbs RA, Boerwinkle E, Doddapaneni H, Han Y, Korchina V, Kovar C, Lee S, Muzny D, Reid JG, Zhu Y *et al.*: **A global reference for human genetic variation**. *Nature* 2015, **526**:68-74.

33. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K *et al.*: **A reference panel of 64,976 haplotypes for genotype imputation**. *Nat Genet* 2016, **48**:1279-9761283.

34. Sirugo G, Williams SM, Tishkoff SA: **The missing diversity in human genetic studies**. *Cell* 2019, **177**:26-31.

35. Hindorff LA, Bonham VL, Brody LC, Ginoza MEC, Hutter CM, Manolio TA, Green ED: **Prioritizing diversity in human genomics research**. *Nat Rev Genet* 2018, **19**:175-185.

36. Rosset S, Tzur S, Behar DM, Wasser WG, Skorecki K: **The population genetics of chronic kidney disease: insights from the MYH9-APOL1 locus**. *Nat Rev Nephrol* 2011, **7**:313-326.

37. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, Bjornsson S, Huttenlocher J, Levey AI, Lah JJ *et al.*: **Variant of *TREM2* associated with the risk of Alzheimer's disease**. *N Engl J Med* 2013, **368**:107-116.

38. Guo JF, Li K, Yu RL, Sun QY, Wang L, Yao LY, Hu YC, Lv ZY, Luo LZ, Shen L *et al.*: **Polygenic determinants of Parkinson's disease in a Chinese population**. *Neurobiol Aging* 2015, **36** 1765.e1-1765.e6.

39. Werness SA, Anderson DJ: **Lipoprotein-associated phospholipase A2 and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies**. *Lancet* 2010, **375**:1536-1544.

40. Millwood IY, Bennett DA, Walters RG, Clarke R, Waterworth D, Johnson T, Chen Y, Yang L, Guo Y, Bian Z *et al.*: **Lipoprotein-associated phospholipase A2 loss-of-function variant and risk of vascular diseases in 90 000 Chinese adults**. *J Am Coll Cardiol* 2016, **67**:230-000231.

41. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J *et al.*: **The support of human genetic evidence for approved drug indications**. *Nat Genet* 2015, **47**:856-860.

42. King EA, Davis JW, Degner JF: **Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval**. *PLoS Genet* 2019, **15**:e1008489.

43. Flannick J, Thorleifsson G, Beer NL, Jacobs SBR, Grarup N, Burtt NP, Mahajan A, Fuchsberger C, Atzmon G, Benediktsson R *et al.*: **Loss-of-function mutations in SLC30A8 protect against type 2 diabetes**. *Nat Genet* 2014, **46**:357-363.

44. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, Kuder JF, Wang H, Liu T, Wasserman SM *et al.*: **Evolocumab and clinical outcomes in patients with cardiovascular disease**. *N Engl J Med* 2017, **376**:1713-1722.

45. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, Gagnon DR, DuVall SL, Li J, Peloso GM *et al.*: **Genetics of blood lipids among ~300 000 multi-ethnic participants of the million veteran program**. *Nat Genet* 2018, **50**:1514-0001523
This article which involved the colloboration of public research labs, pharmaceutical industry leaders and the U.S. department of veterans affair, goes from GWAS to drug discovery and perfectly illustrates how human genetic data can serve as a starting point in pharmaceutical development.

46. Shinohara Y, Katayama Y, Uchiyama S, Yamaguchi T, Handa S, Matsuoka K, Ohashi Y, Tanahashi N, Yamamoto H, Genka C *et al.*: **Cilostazol for prevention of secondary stroke (CSPS 2): an aspirin-controlled, double-blind, randomised non-inferiority trial**. *Lancet Neurol* 2010, **9**:959-968.

47. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A *et al.*: **A**

genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science (80-)* 2006, **314**:1461-1463.

48. Feagan BG, Sandborn WJ, D'Haens G, Panés J, Kaser A, Ferrante M, Louis E, Franchimont D, Dewit O, Seidler U *et al.*: **Induction therapy with the selective interleukin-23 inhibitor risankizumab in patients with moderate-to-severe Crohn's disease: a randomised, double-blind, placebo-controlled phase 2 study**. *Lancet* 2017, **389**:1699-1709.

49. Krueger JG, Ferris LK, Menter A, Wagner F, White A, Visvanathan S, Lalovic B, Aslanyan S, Wang EEL, Hall D *et al.*: **Anti-IL-23A mAb BI 655066 for treatment of moderate-to-severe psoriasis: safety, efficacy, pharmacokinetics, and biomarker results of a single-rising-dose, randomized, double-blind, placebo-controlled trial**. *J Allergy Clin Immunol* 2015, **136**:116-124.e7.

50. Michailidou K, Lindström S, Dennis J, Beesley J, Hui S, Kar S, Lemaçon A, Soucy P, Glubb D, Rostamianfar A *et al.*: **Association analysis identifies 65 new breast cancer risk loci**. *Nature* 2017, **551**:92-94.

51. Medina-Gomez C, Kemp JP, Trajanoska K, Luan J, Chesi A, Ahluwalia TS, Mook-Kanamori DO, Ham A, Hartwig FP, Evans DS *et al.*: **Life-Course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects**. *Am J Hum Genet* 2018, **102**:88-102.

52. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, Zeng L, Ntalla I, Lai FY, Hopewell JC *et al.*: **Association analyses based on false discovery rate implicate new loci for coronary artery disease**. *Nat Genet* 2017, **49**:1385-1391.

53. Turner AW, Wong D, Dreisbach CN, Miller CL: **GWAS reveal targets in vessel wall pathways to treat coronary artery disease**. *Front Cardiovasc Med* 2018, **5**:1-10.

54. Bovijn J, Jackson L, Censin J, Chen CY, Laisk T, Laber S, Ferreira T, Pulit SL, Glastonbury CA, Smoller JW *et al.*: **GWAS identifies risk locus for erectile dysfunction and implicates hypothalamic neurobiology and diabetes in etiology**. *Am J Hum Genet* 2019, **104**:157-163.

55. Justice AE, Karaderi T, Highland HM, Young KL, Graff M, Lu Y, Turcot V, Auer PL, Fine RS, Guo X *et al.*: **Protein-coding variants implicate novel genes related to lipid homeostasis contributing to body-fat distribution**. *Nat Genet* 2019, **51**:452-469.

56. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA *et al.*: **The allelic landscape of human blood cell trait variation and links to common complex disease**. *Cell* 2016, **167**:1415-1429.e19.

57. Kim SK: **Identification of 613 new loci associated with heel bone mineral density and a polygenic risk score for bone mineral density, osteoporosis and fracture**. *PLoS One* 2018, **13**:1-20.

58. Sobacchi C, Frattini A, Guerrini MM, Abinun M, Pangrazio A, Susani L, Bredius R, Mancini G, Cant A, Bishop N *et al.*: **Osteoclast-poor human osteopetrosis due to mutations in the gene encoding RANKL**. *Nat Genet* 2007, **39**:960-962.

59. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL: **Leveraging polygenic functional enrichment to improve GWAS power**. *Am J Hum Genet* 2019, **104**:65-75.

60. Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R *et al.*: **Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci**. *Nat Genet* 2010, **42**:1118-1125.

61. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, Ripke S, Lee JC, Jostins L, Shah T *et al.*: **Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations**. *Nat Genet* 2015, **47**:979-986.

62. Niemi MEK, Martin HC, Rice DL, Gallone G, Gordon S, Kelemen M, McAloney K, McRae J, Radford EJ, Yu S *et al.*: **Common genetic variants contribute to risk of rare severe neurodevelopmental disorders**. *Nature* 2018, **562**:268-271

This landmark paper demonstrates the applicability of GWASs and quantitative genetics to diseases which were thought to be monogenic. Especially, the authors show that common genetic variation combines with severe mutations to modify the phenotypes of neurodevelopmental disorders. Although their work focused only on populations of European ancestry and on neurodevelopment disorders, it paves the way for additional studies on so called disease modifiers.

63. Tao F, Beecham GW, Rebelo AP, Blanton SH, Moran JJ, Lopez-Anido C, Svaren J, Morrow JM, Abreu L, Rizzo D *et al.*: **Variation in SIPA1L2 is correlated with phenotype modification in CMT Type 1A**. *Ann Neurol* 2019

This paper, and the next confirms, in CMT1A, the most common of rare diseases and a disease classically considered as monogenic (via a duplication of PMP22), the finding of Niemi *et al.*, Nature 2018, that monogenic diseases are impacted by more than one gene and that this partly explains why patients with monogenic diseases show a high variability in symptoms. These findings have important implications for drug development as they validate the use of GWASs and Network biology to develop combinatorial treatments.

64. Tao F, Beecham GW, Rebelo AP, Blanton SH, Moran JJ, Lopez-Anido C, Svaren J, Abreu L, Rizzo D, Kirk CA *et al.*: **Modifier gene candidates in charcot-marie-tooth disease Type 1A: a case-only genome-wide association study**. *J Neuromuscul Dis* 2019 http://dx.doi.org/10.3233/JND-190377.

65. Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, Fine RS, Lu Y, Schurmann C, Highland HM *et al.*: **Rare and low-frequency coding variants alter human adult height**. *Nature* 2017, **542**:186-190.

66. Timpson NJ, Greenwood CMT, Soranzo N, Lawson DJ, Richards JB: **Genetic architecture: the shape of the genetic contribution to human traits and disease**. *Nat Rev Genet* 2017, **19**:110-124.

67. O'Connor LJ, Schoech AP, Hormozdiari F, Gazal S, Patterson N, Price AL: **Extreme polygenicity of complex traits is explained by negative selection**. *Am J Hum Genet* 2019 http://dx.doi.org/10.1016/j.ajhg.2019.07.003

This important paper for the field of statistical genetics has important implications for drug discovery as it suggests that small phenotypic effects of most SNPs are not the result of a methodologial biais but the result of patterns of negative selection. In other words, we cannot expect to find one or a few associated genes for most diseases but rather always hundreds of them. This implies that although GWASs are very useful in prioritizing drug targets, they will continue to deliver a high number of candidates which will have to be filtered using, for example, the network biology approaches which are presented in this *Perspective*.

68. Sanseau P, Agarwal P, Barnes MR, Pastinen T, Richards JB, Cardon LR, Mooser V: **Use of genome-wide association studies for drug repositioning**. *Nat Biotechnol* 2012, **30**:317-320.

69. Hopkins AL: **Network pharmacology: the next paradigm in drug discovery**. *Nat Chem Biol* 2008, **4**:682-690.

70. Orgogozo V, Peluffo AE, Morizot B: **The "Mendelian Gene" and the "Molecular Gene: Two Relevant Concepts of Genetic Units"**. *Curr Topics Develop Biol* 2016, **119**:1-26.

71. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, Galver L, Kelley R, Karlsson A, Santos R *et al.*: **The druggable genome and support for target identification and validation in drug development**. *Sci Transl Med* 2017, **9**:1-16.

72. Visscher PM, Yang J: **A plethora of pleiotropy across complex traits**. *Nat Genet* 2016, **48**:707-708.

73. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA: **Detection and interpretation of shared genetic influences on 42 human traits**. *Nat Genet* 2016, **48**:709-717.

74. Liu CC, Kanekiyo T, Xu H, Bu G: **Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy**. *Nat Rev Neurol* 2013, **9**:106-118.

75. Yamazaki Y, Zhao N, Caulfield TR, Liu CC, Bu G: **Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies**. *Nat Rev Neurol* 2019, **15**:501-518.

76. Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, van der Sluis S, Andreassen OA, Neale BM,

Posthuma D: **A global overview of pleiotropy and genetic architecture in complex traits**. *Nat Genet* 2019 http://dx.doi.org/10.1038/s41588-019-0481-0.

77. Dina C, Meyre D, Gallina S, Durand E, Körner A, Jacobson P, Carlsson LMS, Kiess W, Vatin V, Lecoeur C *et al.*: **Variation in FTO contributes to childhood obesity and severe adult obesity**. *Nat Genet* 2007, **39**:724-726.

78. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JRB, Elliott KS, Lango H, Rayner NW *et al.*: **A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity**. *Science (80-)* 2007, **316**:889-894.

79. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orrú M, Usala G *et al.*: **Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits**. *PLoS Genet* 2007, **3**:1200-1210.

80. Pritchard JLE, O'Mara TA, Glubb DM: **Enhancing the promise of drug repositioning through genetics**. *Front Pharmacol* 2017, **8**:1-9.

81. Barabási AL, Oltvai ZN: **Network biology: understanding the cell's functional organization**. *Nat Rev Genet* 2004, **5**:101-113.

82. Simons M, Gordon E, Claesson-Welsh L: **Mechanisms and regulation of endothelial VEGF receptor signalling**. *Nat Rev Mol Cell Biol* 2016, **17**:611-625.

83. Weddell JC, Chen S, Imoukhuede PI: **VEGFR1 promotes cell migration and proliferation through PLCγ and PI3K pathways**. *NPJ Syst Biol Appl* 2018, **4**:1-11.

84. Apte RS, Chen DS, Ferrara N: **VEGF in signaling and disease: beyond discovery and development**. *Cell* 2019, **176**:1248-1264.

85. Liu L, Zuo X, Zhu Z, Wen L, Yang C, Zhu C, Tang L, Cheng Y, Chen M, Zhou F *et al.*: **Genome-wide association study identifies three novel susceptibility loci for systemic lupus erythematosus in Han Chinese**. *Br J Dermatol* 2018, **179**:506-508.

86. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Karlsson Linnér R *et al.*: **Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals**. *Nat Genet* 2018, **50**:1112-1121.

87. Yap CX, Sidorenko J, Wu Y, Kemper KE, Yang J, Wray NR, Robinson MR, Visscher PM: **Dissection of genetic variation and evidence for pleiotropy in male pattern baldness**. *Nat Commun* 2018, **9**:1-12.

88. Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, Chai RC, Vulpescu NA, Forgetta V, Kleinman A, Mohanty ST *et al.*: **An atlas of genetic influences on osteoporosis in humans and mice**. *Nat Genet* 2019, **51**:258-266.

89. Boyle EA, Li YI, Pritchard JK: **An expanded view of complex traits: from polygenic to omnigenic**. *Cell* 2017, **169**:1177-1186.

90. Wray NR, Wijmenga C, Sullivan PF, Yang J, Visscher PM: **Common disease is more complex than implied by the core gene omnigenic model**. *Cell* 2018, **173**:1573-1580
This important perspective paper by leaders of the GWASs field argues that polygenicity, that is, the fact that most phenotypes and diseases are underlied by numerous genes rather than one or a few 'core' genes, is a pervasive property of biological systems. The authors elegantly demonstrate, based not only on GWASs results but experimental and network biology studies, that it is a fallacy to consider genes of stronger phenotypic effect as being more important than genes with smaller effects in the etiology of a disease and that it would be a 'disservice' to clinical research to prioritize these supposedly 'core' genes.

91. Félix MA, Barkoulas M: **Pervasive robustness in biological systems**. *Nat Rev Genet* 2015, **16**:483-496.

92. Kim J, Vandamme D, Kim JR, Munoz AG, Kolch W, Cho KH: **Robustness and evolvability of the human signaling network**. *PLoS Comput Biol* 2014, **10**.

93. Wagner A: **Distributed robustness versus redundancy as causes of mutational robustness**. *BioEssays* 2005, **27**:176-188.

94. Gao Z, Waggoner D, Stephens M, Ober C, Przeworski M: **An estimate of the average number of recessive lethal mutations carried by humans**. *Genetics* 2015, **199**:1243-1254.

95. Narasimhan VM, Hunt KA, Mason D, Baker CL, Karczewski KJ, Barnes MR, Barnett AH, Bates C, Bellary S, Bockett NA *et al.*: **Health and population effects of rare gene knockouts in adult humans with related parents**. *Science* 2016, **352**:474-477.

96. MacArthur D, Balasubramanian S, Frankish A: **A systematic survey of loss-of-function variants in human protein-coding genes**. *Science* 2012, **335**:1-14.

97. Hopkins AL: **Network pharmacology**. *Nat Biotechnol* 2007, **25**:1110-1111.

98. Barabási AL, Gulbahce N, Loscalzo J: **Network medicine: a network-based approach to human disease**. *Nat Rev Genet* 2011, **12**:56-68.

99. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Yoshida S *et al.*: **Genetics of rheumatoid arthritis contributes to biology and drug discovery**. *Nature* 2014, **506**:376-381.

100. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D: **Benefits and limitations of genome-wide association studies**. *Nat Rev Genet* 2019, **20**.

101. Zipp F, Ivinson AJ, Haines JL, Sawcer S, Dejager P, Hauser SL, Oksenberg JR: **Network-based multiple sclerosis pathway analysis with GWAS data from 15 000 cases and 30 000 controls**. *Am J Hum Genet* 2013, **92**:854-865.

102. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al.*: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles**. *Proc Natl Acad Sci U S A* 2005, **102**:15545-15550.

103. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, Wadi L, Meyer M, Wong J, Xu C *et al.*: **Pathway enrichment analysis and visualization of omics data using g: Profiler, GSEA, Cytoscape and EnrichmentMap**. *Nat Protoc* 2019 http://dx.doi.org/10.1038/s41596-018-0103-9
This important paper offers a step-by-step guide on how to perform pathway enrichment analysis and other network biology methods over a list of genes derived from GWASs or other omics experiments for biologists with no prior bioinformatics training.

104. Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, Zhang R, Hartmann BM, Zaslavsky E, Sealfon SC *et al.*: **Understanding multicellular function and disease with human tissue-specific networks**. *Nat Genet* 2015, **47**:569-576.

105. Iotchkova V, Ritchie GRS, Geihs M, Morganella S, Min JL, Walter K, Timpson NJ, Dunham I, Birney E, Soranzo N: **GARFIELD classifies disease-relevant genomic features through integration of functional annotations with association signals**. *Nat Genet* 2019, **2019** http://dx.doi.org/10.1038/s41588-018-0322-6.

106. Kim SS, Dai C, Hormozdiari F, van de Geijn B, Gazal S, Park Y, O'Connor L, Amariuta T, Loh P-R, Finucane H *et al.*: **Genes with high network connectivity are enriched for disease heritability**. *Am J Hum Genet* 2019, **104**:896-913
This article analyzed 42 diseases and complex traits and finds that genes that contribute the most to the disease phenotype variation are genes that interact with numerous other genes at the gene regulatory level. Although these findings confirm the link between network connectivity and disease etiology, the authors caution that not accounting for gene regulatory annotation, that is, not considering which genes regulate which gene can yield spurious pathway enrichment results.

107. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, Lui JC, Vedantam S, Gustafsson S, Esko T *et al.*: **Biological interpretation of genome-wide association studies using predicted gene functions**. *Nat Commun* 2015:6.

108. Fine RS, Pers TH, Amariuta T, Raychaudhuri S, Hirschhorn JN: **Benchmarker: an unbiased, association-data-driven strategy to evaluate gene prioritization algorithms**. *Am J Hum Genet* 2019, **104**:1025-1039.

109. Atta MG, De Seigneux S, Lucas GM: **Clinical pharmacology in HIV therapy**. *Clin J Am Soc Nephrol* 2019, **14**:435-444.

110. Price RN, White NJ: **Drugs that reduce transmission of falciparum malaria**. *Lancet Infect Dis* 2018, **18**:585-586.

111. Wells TNC, Van Huijsduijnen RH, Van Voorhis WC: **Malaria medicines: a glass half full?** *Nat Rev Drug Discov* 2015, **14**:424-442.

112. Das P, Delost MD, Qureshi MH, Smith DT, Njardarson JT: **A survey of the structures of US FDA approved combination drugs**. *J Med Chem* 2019, **62**:4265-4311.

113. Sun W, Sanderson PE, Zheng W: **Drug combination therapy increases successful drug repositioning**. *Drug Discov Today* 2016, **21**:1189-1195.

114. Hung IFN, To KKW, Chan JFW, Cheng VCC, Liu KSH, Tam A, Chan TC, Zhang AJ, Li P, Wong TL *et al.*: **Efficacy of clarithromycin-naproxen-oseltamivir combination in the treatment of patients hospitalized for Influenza A(H3N2) infection: an open-label randomized, controlled, Phase IIb/III trial**. *Chest* 2017, **151**:1069-1080.

115. Tang J, Wennerberg K, Aittokallio T: **What is synergy? The Saariselkä agreement revisited**. *Front Pharmacol* 2015, **6**:1-5.

116. Lehár J, Krueger AS, Avery W, Heilbut AM, Johansen LM, Price ER, Rickles RJ, Short GF, Staunton JE, Jin X *et al.*: **Synergistic drug combinations tend to improve therapeutically relevant selectivity**. *Nat Biotechnol* 2009, **27**:659-666.

117. Cheng Y-S, Williamson PR, Zheng W: **Improving therapy of severe infections through drug repurposing of synergistic combinations**. *Curr Opin Pharmacol* 2019, **48**:92-98
This article clarifies the notion of synergistic combinations in the context of drug repurposing to multidrug resistant pathogens. The authors show that non-antibiotic compounds can be combined synergistically and repurposed to treat infections.

118. Mercorelli B, Palù G, Loregian A: **Drug repurposing for viral infectious diseases: how far are we?** *Trends Microbiol* 2018, **26**:865-876.

119. Riedel T, Demaria O, Zava O, Joncic A, Gilliet M, Dyson PJ: **Drug repurposing approach identifies a synergistic drug combination of an antifungal agent and an experimental organometallic drug for melanoma treatment**. *Mol Pharm* 2018, **15**:116-126.

120. Proschak E, Stark H, Merk D: **Polypharmacology by design: a medicinal chemist's perspective on multitargeting compounds**. *J Med Chem* 2019, **62**:420-444.

121. Chumakov I, Milet A, Cholet N, Primas G, Boucard A, Pereira Y, Graudens E, Mandel J, Laffaire J, Foucquier J *et al.*: **Polytherapy with a combination of three repurposed drugs (PXT3003) down-regulates Pmp22 over-expression and improves myelination, axonal and functional parameters in models of CMT1A neuropathy**. *Orphanet J Rare Dis* 2014, **9**:1-16.

122. Prukop T, Stenzel J, Wernick S, Kungl T, Mroczek M, Adam J, Ewers D, Nabirotchkin S, Nave K-A, Hajj R *et al.*: **Early short-term PXT3003 combinational therapy delays disease onset in a transgenic rat model of Charcot-Marie-Tooth disease 1A (CMT1A)**. *PLoS One* 2019, **14** e0209752
This article, together with a series of previously published papers by the same group (previous and next reference), demonstrates the efficacy of the low-dose, synergistic combination of baclofen, naltrexone and sorbitol in young CMT1A animals.

123. Attarian S, Vallat JM, Magy L, Funalot B, Gonnaud PM, Lacour A, Péréon Y, Dubourg O, Pouget J, Micallef J *et al.*: **An exploratory randomised double-blind and placebo-controlled phase 2 study of a combination of baclofen, naltrexone and sorbitol (PXT3003) in patients with Charcot-Marie-Tooth disease type 1A**. *Orphanet J Rare Dis* 2014, **9**:1-15.

124. Hajj R, Milet A, Toulorge D, Cholet N, Laffaire J, Foucquier J, Robelet S, Mitry R, Guedj M, Nabirotchkin S *et al.*: **Combination of acamprosate and baclofen as a promising therapeutic approach for Parkinson's disease**. *Sci Rep* 2015, **5**:16084.

125. Chumakov I, Nabirotchkin S, Cholet N, Milet A, Boucard A, Toulorge D, Pereira Y, Graudens E, Traore S, Foucquier J *et al.*: **Combining two repurposed drugs as a promising approach for Alzheimer's disease therapy**. *Sci Rep* 2015, **5**:1-12.

126. Hajj R: **Parkinson disease therapies and drugs**. In *Pathology, Prevention and Therapeutics of Neurodegenerative Disease.* Edited by Singh S, Joshi N. Singapore: Springer; 2019:151-158.

127. Cheng F, Kovács IA, Barabási AL: **Network-based prediction of drug combinations**. *Nat Commun* 2019, **10**:1197
This key paper demonstrates the possibility to combine genetic data (especially GWASs) with protein-protein interaction databases to identify novel drug combinations to treat complex diseases. The authors detail the concepts and methods behind the use of human genetic data together with network biology approaches and validate their work based on approved drug combinations for hypertension and cancer.

128. Pemovska T, Bigenzahn JW, Superti-Furga G: **Recent advances in combinatorial drug screening and synergy scoring**. *Curr Opin Pharmacol* 2018, **42**:102-110
This article is a great review of the statistical methods which exist to quantify drug synergy and how to use them accordingly to derive a combination of novel or repurposed drugs.

129. Foucquier J, Guedj M: **Analysis of drug combinations: current methodological landscape**. *Pharmacol Res Perspect* 2015, **3**: e00149.

130. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M *et al.*: **Applications of machine learning in drug discovery and development**. *Nat Rev Drug Discov* 2019 http://dx.doi.org/10.1038/s41573-019-0024-5
This important paper reviews the contribution of most machine learning approaches to drug discovery. It is a great map of the existing machine learning methods, where they can be used in the drug discovery process, and what are their limits.

131. Topol EJ: **High-performance medicine: the convergence of human and artificial intelligence**. *Nat Med* 2019, **25**:44-56.

132. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, Cui C, Corrado G, Thrun S, Dean J: **A guide to deep learning in healthcare**. *Nat Med* 2019, **25**:24-29.

133. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, Liu PJ, Liu X, Marcus J, Sun M *et al.*: **Scalable and accurate deep learning with electronic health records**. *NPJ Digit Med* 2018, **1**:1-10.

134. Zong N, Kim H, Ngo V, Harismendy O: **Deep mining heterogeneous networks of biomedical linked data to predict novel drug-target associations**. *Bioinformatics* 2017, **33**:2337-2344.

135. Ragoza M, Hochuli J, Idrobo E, Sunseri J, Koes DR: **Protein-ligand scoring with convolutional neural networks**. *J Chem Inf Model* 2017, **57**:942-957.

136. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT *et al.*: **Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations**. *Nat Genet* 2018, **50**:1219-1224.