



# On the Integration of *In Silico* Drug Design Methods for Drug Repurposing

Eric March-Vila<sup>1</sup>, Luca Pinzi<sup>1</sup>, Noé Sturm<sup>1</sup>, Annachiara Tinivella<sup>1</sup>, Ola Engkvist<sup>2</sup>, Hongming Chen<sup>2</sup> and Giulio Rastelli<sup>1\*</sup>

<sup>1</sup> Molecular Modelling & Drug Design Lab, Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy, <sup>2</sup> Discovery Sciences, Innovative Medicines and Early Development Biotech Unit, AstraZeneca R&D Gothenburg, Mölndal, Sweden

Drug repurposing has become an important branch of drug discovery. Several computational approaches that help to uncover new repurposing opportunities and aid the discovery process have been put forward, or adapted from previous applications. A number of successful examples are now available. Overall, future developments will greatly benefit from integration of different methods, approaches and disciplines. Steps forward in this direction are expected to help to clarify, and therefore to rationally predict, new drug–target, target–disease, and ultimately drug–disease associations.

## OPEN ACCESS

### Edited by:

Yuhei Nishimura,  
Mie University, Japan

### Reviewed by:

Antonio Macchiarulo,  
University of Perugia, Italy  
Yoshito Zamami,  
Tokushima University Graduate  
School of Medical Sciences, Japan

### \*Correspondence:

Giulio Rastelli  
giulio.rastelli@unimore.it

### Specialty section:

This article was submitted to  
Experimental Pharmacology and Drug  
Discovery,  
a section of the journal  
Frontiers in Pharmacology

**Received:** 07 April 2017

**Accepted:** 10 May 2017

**Published:** 23 May 2017

### Citation:

March-Vila E, Pinzi L, Sturm N,  
Tinivella A, Engkvist O, Chen H and  
Rastelli G (2017) On the Integration  
of *In Silico* Drug Design Methods  
for Drug Repurposing.  
Front. Pharmacol. 8:298.  
doi: 10.3389/fphar.2017.00298

**Keywords:** drug repurposing, drug discovery, molecular modeling, chemogenomics, structure-based drug design, ligand-based drug design, machine learning, transcriptomics

## INTRODUCTION

Drug repurposing (also known as drug repositioning) aims at identifying new uses for already existing drugs (Novac, 2013). In drug discovery, drug repurposing has gained an increasingly important role, because it helps to circumvent preclinical development and optimization issues, hence reducing time efforts, expenses and failures typically associated with the drug discovery process.

Over the years, biological and chemical information has been generated at an ever-increasing pace, marking the entrance in the so-called “big data” era (Costa, 2014). This offers the scientific community new opportunities to link drugs to diseases, although this relationship is indirect and relies on complex mechanisms of action. Therefore, a better understanding of the relationships between drugs and their targets, and between targets and diseases, is a key for drug repurposing. Unfortunately, we are still far from understanding the overall picture, partly due to the heterogeneity and incompleteness of the available data. However, computational methods offer valuable opportunities to create such links, as it will be illustrated below.

In this perspective, different computational methods and approaches are briefly presented, and their ability to complement and integrate each other in drug repurposing is discussed, which will certainly gain a foothold in the future.

## COMPUTATIONAL DRUG REPURPOSING STRATEGY BASED ON TRANSCRIPTIONAL SIGNATURES

Transcriptomic data can provide a list of over- and under-expressed genes in a biological system treated by a pharmacologically active compound. The perturbation of a biological system

can be measured from genome wide transcriptional responses, and the drug induced transcriptional responses represent the signature of the compound activity on biological systems. These molecular transcriptional signatures can then be compared to establish therapeutic relationships between known drugs and new disease indications.

One of the most comprehensive and systematic approaches toward leveraging the transcriptional signature approach for drug repositioning is the Connectivity Map project (Lamb et al., 2006). The publicly funded CMap database<sup>1</sup> initially contained profiles of 164 drugs and was later expanded to 1309 FDA-approved small molecules. These compounds are tested in five human cell lines, generating over 7000 gene expression profiles in the database. The cell perturbation profile of each drug in the reference collection contains, for each gene measured, a rank-based measure of the change in transcriptional activity after exposure to the drug compound, i.e., gene signatures. These signatures form the basis of comparing drugs mechanism of action at transcriptional level and have been successfully applied for drug repurposing in many examples. Chang et al. (2010) used CMap to identify new analgesic and antinociceptive properties of phenoxybenzamine, originally an anti-hypertensive drug. Subsequent testing using a rat inflammatory model validated the analgesic activity. In contrast with CMap gene signatures, biclustering methods were applied to CMap to group coregulated genes with the drugs they respond to Iskar et al. (2013). This led to the identification of vinburnine, a vasodilator, and sulconazole, a topical antifungal, as interesting cell cycle blockers for cancer therapy.

## NETWORK-BASED DRUG REPURPOSING

In recent years, network-based computational biology has attracted increasing attention. It aims at organizing the relationships among biological molecules in the form of networks to find newly emerged properties at a network level, and to investigate how cellular systems induce different biological phenotypes under different conditions. In the network pharmacology framework, a network can be depicted as a connected graph, where each node can represent either an individual molecular entity (e.g., a drug), its biological target, a modifier molecule within a biological process, or a target pathway, while an edge represents either a direct or indirect interaction between two connected nodes. Ultimately, both the efficacy and the toxicity of a drug are a consequence of the complex interplay among different cellular components. A system-scale perspective is therefore needed to aid modern drug discovery, especially for complex diseases, which are known to be caused by perturbation of biological networks.

Network-based analysis has become a widely used strategy for computational drug repositioning. Hu and Agarwal (2009) created a disease-similarity network using publicly available gene expression profiles from NCBI Gene Expression Omnibus

(GEO)<sup>2</sup> and integrated this network with molecular profiles and knowledge of drugs and drug targets to infer drug repositioning opportunities and suggest molecular targets and mechanisms underlying drug effects. Jin et al. (2012) developed a novel method to repurpose drugs for cancer therapeutics by leveraging off-target effects that may affect important cancer cell signaling pathways. The off-target effects of drugs on signaling proteins were identified by using a hybrid model composed of a network component called cancer-signaling bridges and Bayesian factor regression model.

## LIGAND-BASED APPROACHES IN DRUG REPURPOSING

Ligand-based approaches are based on the concept that similar compounds tend to have similar biological properties. In drug repurposing, these methods have been extensively used to analyze and predict the activity of ligands for new targets. Public databases of bioactive molecules, such as PubChem, ChEMBL, and DrugBank contain information retrieved and manually curated from literature data (Wishart et al., 2006; Gaulton et al., 2017; Wang et al., 2017). These databases represent a huge and ever-growing reservoir of chemical and biological information such as binding affinity, cellular activity, functional and ADMET data. Recent advances in drug repurposing include the release of databases focused on repurposed drugs, failed drugs, their therapeutic indications, and bioactivity data (Brown and Patel, 2017; Shameer et al., 2017).

One advantage of applying these approaches to drug repurposing is that the number of publicly accessible compound records (more than hundred millions provided only by PubChem) is far greater than the number of deposited protein crystal structures (as of today, less than 150,000 in the Protein Data Bank) (Berman et al., 2000; Wang et al., 2017). On the other hand, ligand-based methods obviously depend on the chemical space coverage of already known molecules. Moreover, a high overall similarity does not necessarily guarantee activity on a secondary target, since local structural divergences in chemical scaffolds can lead to “activity cliffs” (Stumpfe and Bajorath, 2012). This limitation, however, will eventually be overcome by the increase of structural diversity in bioactivity databases (Hu and Bajorath, 2013).

Recently, a 2D ligand-based similarity analysis of ChEMBL combined with support vector machine models and analysis of 3D structural information of protein–ligand complexes, identified a promising set of target combinations and associated ligands within the Hsp90 interactome, which are particularly suitable for multitarget drug design (Anighoro et al., 2015). Another ligand-based method correctly predicted 23 new drug–target associations using the similarity ensemble approach (Keiser et al., 2009). Pharmacophore screening has also been a

<sup>1</sup><http://www.broadinstitute.org/cmap>

<sup>2</sup><http://www.ncbi.nlm.nih.gov/geo/>

valuable strategy for drug repurposing (Liu et al., 2010). In this approach, a drug can be represented as a set of pharmacophoric features which can subsequently be used to interrogate chemical compound databases to provide compounds with different scaffolds.

Complementing different levels of ligand description increases the chances of identifying new repurposing possibilities. For example, Vasudevan et al. (2012) used 3D shape-based descriptors to compare approved drugs with a set of H1 receptor antagonists. Thirteen of the 23 tested drugs selectively inhibited histamine-induced calcium release by acting at the H1 receptor level. Interestingly, these drugs would not have been detected with 2D similarity searching (Vasudevan et al., 2012). Furthermore, Mervin et al. (2015) demonstrated how the inclusion of inactive data improved early recognition abilities in statistical prediction models. On different grounds, predictive models built upon disease feature descriptors, large-scale drug–target and target–disease associations showed performance improvements in predicting new drug–disease links (Iwata et al., 2015; Sawada et al., 2015). In particular, it was shown that chemical similarity and phenotypic similarity are complementary to each other, and that integrating predictions from both methods is beneficial (Sawada et al., 2015).

## LIGAND-BASED CHEMOGENOMICS AND MACHINE LEARNING IN DRUG REPURPOSING

A variety of *in silico* approaches have been applied in ligand-based chemogenomic campaigns (Mestres et al., 2006; Bender et al., 2007; Gregori-Puigjané and Mestres, 2008). During the last years, machine learning algorithms, which span from the older but still attractive Bayesian classifiers to the more advanced support vector machines, have become increasingly popular to assist the drug repositioning process (Bender et al., 2007). Methods such as deep learning and multi-task learning have been successfully used in chemogenomic benchmark studies (Unterthiner et al., 2014). Moreover, matrix factorization methods offer the opportunity to combine bioactivity data with other information, such as disease information, in one framework (Zhang et al., 2014). On a different line, other techniques inspired by e-commerce websites have shown interesting results in identifying new drug–target associations (Alaimo et al., 2016). In the study, the technique relies on a network-based inference algorithm and a drug–target bipartite graph extracted from DrugBank. It was shown that the algorithm performed better in predicting new drug–target associations when target and drug similarities are considered.

Given the versatility in their use and their computational efficiency, machine-learning approaches will likely continue to play a prominent role in *in silico* chemogenomics. Despite many papers have described test cases and various types of method development, there is still a lack of published success stories that employed ligand-based chemogenomics modeling in drug repurposing.

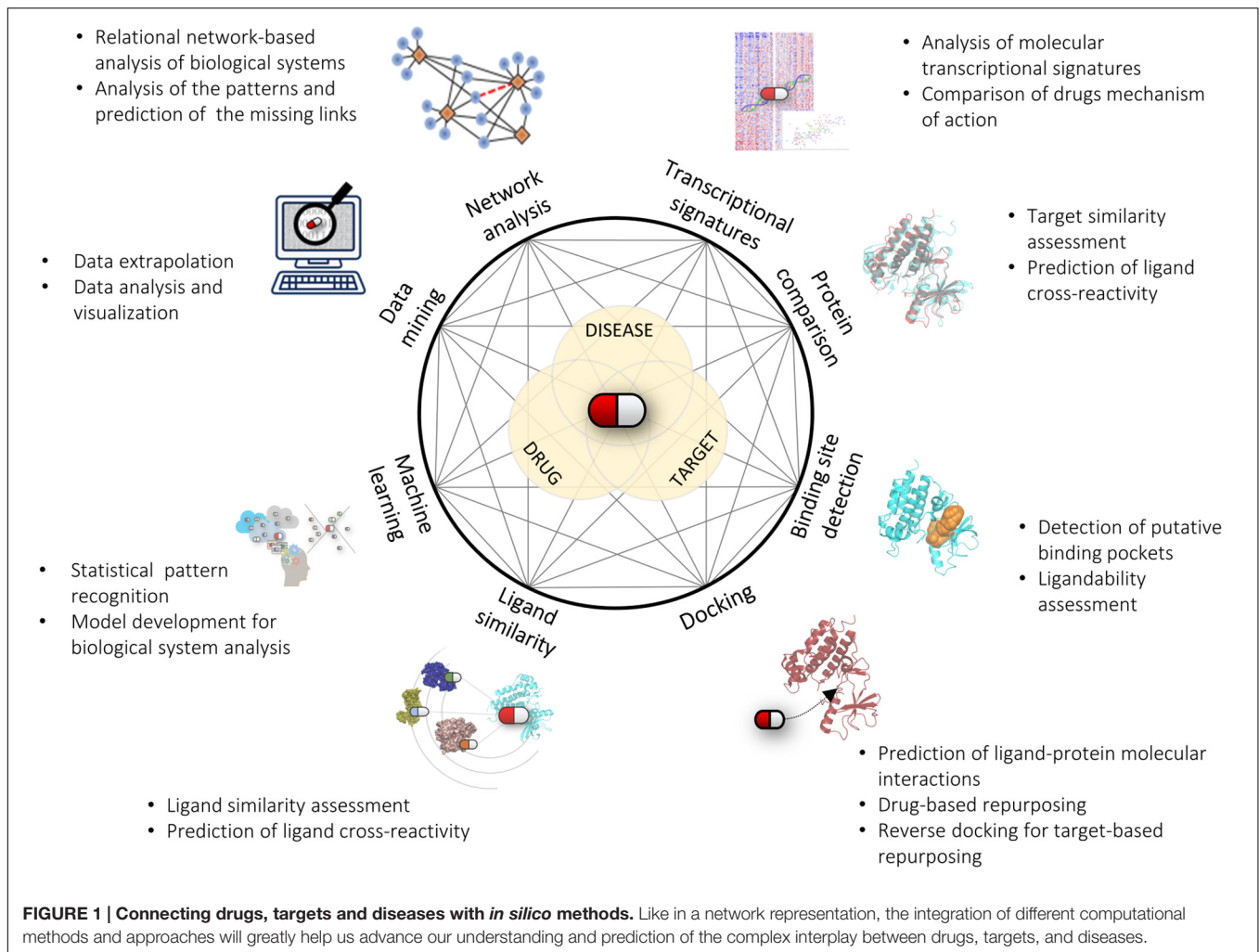
## STRUCTURE-BASED APPROACHES IN DRUG REPURPOSING

It is established that the similarity principle observed for ligands applies also to proteins. Proteins with similar structures are likely to have similar functions and to recognize similar ligands. In the field of drug repurposing, protein comparison is used as a method to identify secondary targets of an approved drug (Ehrt et al., 2016).

From a global point of view, proteins can be compared by sequence similarity. Protein sequences have been used to build phylogenetic trees, the most popular of which is represented by the kinome (Manning et al., 2002). In this tree, proteins of the same family are prone to have related functions and also to recognize related substrates or ligands, such as for example dual inhibitors of epidermal growth factor receptor (EGFR) and epidermal growth factor receptor B2 (ErbB2) (Zhang et al., 2004). Modern methods to perform multiple-sequence alignments, such as BLAST, are nowadays widely used and available through web-servers. It is important to note that small differences localized at key positions, such as those occurring in correspondence of the gatekeeper residue of protein kinases or of other oncogenic mutations, may have a huge impact on ligand binding (Huang and Fu, 2015). Hence, local differences in globally conserved protein sequences should be given careful consideration. Moreover, a study based on the similarity ensemble approach showed that similar ligands were able to bind proteins with distantly related sequences (Keiser et al., 2007). Overall, local binding site similarities can be more important than global similarities to determine polypharmacology and drug repurposing (Jalencas and Mestres, 2013b; Anighoro et al., 2015).

In identifying unknown targets of known ligands, sequence alignments perform well when proteins share a high degree of sequence identity, whereas local protein comparison performs better when proteins share low sequence identity (Chen et al., 2016). Detecting local similarities by comparing protein binding sites has become increasingly important (Ehrt et al., 2016). Binding site identification and comparison are commonly performed by scanning the protein surface in order to identify cavities (Laurie and Jackson, 2006) and then by calculating descriptors of different nature useful to derive a similarity score.

It is important to note that several approaches and algorithms for binding site comparison have been put forward, but none of them appears to be devoid of failures or limitations (Ehrt et al., 2016). Notwithstanding, binding site similarity has proven a valuable tool in a number of studies. For example, a study carried out by Defranchi et al. (2010) used a binding site comparison method to predict the cross-reactivity of four protein kinase inhibitors with Synapsin I. These discoveries were supported by sub-micromolar affinities of the kinase inhibitors for Synapsin I. Interestingly, binding site similarity and other molecular modeling techniques were used in combination to uncover new targets of the drugs entacapone and tolcapone (Kinnings et al., 2009). The study started from a large set of similar binding sites, which was further



finalized by simulating the binding mode of entacapone and tolcapone using docking. Proteins for which ligands gave the best docking scores were prioritized and further experimentally validated.

It is worth mentioning that ligand binding modes, when available, are a strong asset in the process of identifying new targets. One way to model the molecular recognition is to focus on target–ligand interactions. This can be achieved with various methods, such as structure-based pharmacophores or interaction fingerprints. When the structure of a protein–ligand complex is not available, one can use computational methods to predict hot spots in the binding site (Hall et al., 2015). Another approach joining ligand information to protein environments uses the concept of chemoisosterism (Jalencas and Mestres, 2013a). Chemoisosterism can be defined as the property of two protein environments to bind the same molecular fragment, and can shed light into the inherent cross-pharmacology between protein targets. The degree of chemoisosterism was found to be related to the polypharmacology of chemical fragments (Jalencas and Mestres, 2013b). This approach allows the creation of interaction

networks connecting chemical fragments to chemoisosteric protein environments. These networks, complemented with target–disease associations, constitute attractive starting points for drug repurposing efforts.

Based on similar concepts, a method for interrogating large data sets of proteins (as large as the PDB) with highly customizable geometric patterns as searching templates was recently described (Inhester et al., 2017). This method was able to identify chemoisosteric protein environments binding the uracil moiety of uridine diphosphate from a query built with deoxythymidine.

Structure-based methods are obviously dependent on the availability of crystallographic structures of protein–ligand complexes. Resolution and sensitivity to atomic coordinates impact the level of details that one can use to model a binding site. While crystallographic structures represent a static model of a protein, other pockets may appear upon conformational changes. Detecting those cryptic sites has become an emerging field of research, because it may provide additional options in drug repurposing. In fact, cryptic allosteric sites may be useful to gain selectivity, explore new chemical spaces for

drug design, and establish drug–target associations beyond the more commonly explored orthosteric site. For instance, Markov models have been applied in combination with experimental assays using a chemical probe to uncover cryptic allosteric sites of TEM-1  $\beta$ -lactamase (Bowman et al., 2015). Overall, uncovering new allosteric sites in proteins may provide far more opportunities to repurpose drugs than is currently recognized.

## MOLECULAR DOCKING

Molecular docking is a versatile tool used to predict the geometry and to score the interaction of a protein in complex with a small-molecule ligand (Kitchen et al., 2004). Therefore, these methods can be used to predict if a given drug is potentially able to bind other targets. Docking studies have been successfully exploited in drug repurposing, as reported in many recent studies (Kinnings et al., 2009; Li et al., 2011; Dakshanamurthy et al., 2012). In this context, virtual screenings can be performed either by docking a known drug into a large set of different target structures, or by docking a database of approved drugs into one intended specific target. Molecular docking is in fact a convenient and fast method to screen large libraries of both ligands and targets, with a full range of sampling options (Kitchen et al., 2004), and is obviously restricted to studies in which a 3D structure of the target is available through crystallography, nuclear magnetic resonance (NMR), or comparative models. It should be noted that docking methods still have drawbacks and limitations, mainly arising from the use of approximate scoring functions and imperfect binding mode placement algorithms. Often these problems can be overcome by post-processing docking results with more accurate scoring functions and/or other criteria (Sgobba et al., 2012).

In the study of Li et al. (2011), docking methods have been successfully exploited as a stand-alone method in drug repurposing, by docking the drugs of the DrugBank database into 35 crystal structures of MAPK14. The study identified the chronic myeloid leukemia drug nilotinib as a potential anti-inflammatory drug with an *in vitro* IC<sub>50</sub> of 40 nM (Li et al., 2011).

Docking is notably well suited for either drug-based and target-based drug repurposing, as reported in the results of the work of Dakshanamurthy et al. (2012), where an anti-parasitic drug was successfully tested as an anti-angiogenic vascular endothelial growth factor receptor 2 (VEGFR2) inhibitor, and a new connection was discovered between previously untargeted Cadherin-11, implied in rheumatoid arthritis, and cyclooxygenase-2 (COX-2) inhibitor celecoxib.

It is important to note that docking, despite its limitations, is a well-established and experimentally validated approach for predicting new drug–target associations. Once integrated with ligand-based methods and other available information about target–disease associations, it constitutes a powerful approach to repurpose (newly) targeted drugs for a specific disease.

## INTEGRATING DIFFERENT APPROACHES AND FUTURE DIRECTIONS

The goal of drug repurposing is to uncover new links between drugs and diseases, most commonly *via* targets. As illustrated in the previous sections, computational predictions followed by experimental assessment have been successfully used to identify new drug repurposing possibilities. As always, each computational method has its own field of applicability, drawbacks and limitations. One should be aware of the fact that none of these methods alone will be sufficiently able to disclose (or even model) the complex interplay between drugs, targets and diseases. Therefore, we are left with the possibility of using one or more computational approaches to “navigate” through the wealth of available information and hopefully find “clues” solid enough to justify a repurposing hypothesis worth of experimental investigation. The choice of the most appropriate method(s) will basically depend on the nature of the problem to solve and on the type, quality, and quantity of information available on that problem in the literature or in public or proprietary databases. Unfortunately, information is often fragmented, and generally reflects only a single or few aspects of a much more complicated story. Future efforts should be more thoroughly directed toward disclosing hubs and links of the complex network that relates drugs, targets and diseases. Integrating the huge and heterogeneous amount of available data (chemical, biological, structural, clinical) into a unified workflow is obviously a challenging task. In this respect, the integration and use of different computational methods as shown above will provide valuable opportunities to extend the domain of applicability of each method and more thoroughly exploit information coming from different sources (Figure 1). Likewise, this will greatly benefit from better integration of multidisciplinary work. A network-based approach built upon these considerations will likely provide new routes to navigate through all the potential links between drugs and diseases, thus creating new opportunities for drug repurposing and drug discovery in general.

## AUTHOR CONTRIBUTIONS

All authors contributed in writing and editing the manuscript. EM-V, LP, NS and AT contributed equally. GR conceived the study and coordinated the writing.

## FUNDING

The project leading to this article has received funding (for EM-V, OE, HC, and GR) from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 676434, “Big Data in Chemistry” (“BIGCHEM”, <http://bigchem.eu>). The article reflects only the authors’ view and neither the European Commission nor the Research Executive Agency are responsible for any use that may be made of the information it contains.

## REFERENCES

- Alaimo, S., Giugno, R., and Pulvirenti, A. (2016). Recommendation techniques for drug-target interaction prediction and drug repositioning. *Methods Mol. Biol.* 1415, 441–462. doi: 10.1007/978-1-4939-3572-7-23
- Anighoro, A., Stumpfe, D., Heikamp, K., Beebe, K., Neckers, L. M., Bajorath, J., et al. (2015). Computational polypharmacology analysis of the heat shock protein 90 interactome. *J. Chem. Inf. Model.* 55, 676–686. doi: 10.1021/ci5006959
- Bender, A., Scheiber, J., Glick, M., Davies, J. W., Azzaoui, K., and Hamon, J. (2007). Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure. *ChemMedChem* 2, 861–873. doi: 10.1002/cmdc.200700026
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The protein data bank. *Nucleic Acids Res.* 28, 235–242. doi: 10.1093/nar/28.1.235
- Bowman, G. R., Bolin, E. R., Hart, K. M., Maguire, B. C., and Marqusee, S. (2015). Discovery of multiple hidden allosteric sites by combining Markov state models and experiments. *Proc. Natl. Acad. Sci. U. S. A.* 112, 2734–2739. doi: 10.1073/pnas.1417811112
- Brown, A. S., and Patel, C. J. (2017). A standard database for drug repositioning. *Sci. Data* 4, 170029. doi: 10.1038/sdata.2017.29
- Chang, M., Smith, S., Thorpe, A., Barratt, M. J., and Karim, F. (2010). Evaluation of phenoxybenzamine in the CFA model of pain following gene expression studies and connectivity mapping. *Mol. Pain* 6:56. doi: 10.1186/1744-8069-6-56
- Chen, Y.-C., Tolbert, R., Aronov, A. M., McGaughey, G., Walters, W. P., and Meireles, L. (2016). Prediction of protein pairs sharing common active ligands using protein sequence, structure, and ligand similarity. *J. Chem. Inf. Model.* 56, 1734–1745. doi: 10.1021/acs.jcim.6b00118
- Costa, F. F. (2014). Big data in biomedicine. *Drug Discov. Today* 19, 433–440. doi: 10.1016/j.drudis.2013.10.012
- Dakshnamurthy, S., Issa, N. T., Assefnia, S., Seshasayee, A., Peters, O. J., Madhavan, S., et al. (2012). Predicting new indications for approved drugs using a proteo-chemometric method. *J. Med. Chem.* 55, 6832–6848. doi: 10.1021/jm300576q
- Defranchi, E., De Franchi, E., Schalon, C., Messa, M., Onofri, F., Benfenati, F., et al. (2010). Binding of protein kinase inhibitors to synapsin I inferred from pairwise binding site similarity measurements. *PLoS ONE* 5:e12214. doi: 10.1371/journal.pone.0012214
- Ehrt, C., Brinkjost, T., and Koch, O. (2016). Impact of binding site comparisons on medicinal chemistry and rational molecular design. *J. Med. Chem.* 59, 4121–4151. doi: 10.1021/acs.jmedchem.6b00078
- Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., et al. (2017). The ChEMBL database in 2017. *Nucleic Acids Res.* 45, D945–D954. doi: 10.1093/nar/gkw1074
- Gregori-Puigjané, E., and Mestres, J. (2008). A ligand-based approach to mining the chemogenomic space of drugs. *Comb. Chem. High Throughput Screen.* 11, 669–676.
- Hall, D. R., Kozakov, D., Whitty, A., and Vajda, S. (2015). Lessons from hot spot analysis for fragment-based drug discovery. *Trends Pharmacol. Sci.* 36, 724–736. doi: 10.1016/j.tips.2015.08.003
- Hu, G., and Agarwal, P. (2009). Human disease-drug network based on genomic expression profiles. *PLoS ONE* 4:e6536. doi: 10.1371/journal.pone.0006536
- Hu, Y., and Bajorath, J. (2013). Compound promiscuity: what can we learn from current data? *Drug Discov. Today* 18, 644–650. doi: 10.1016/j.drudis.2013.03.002
- Huang, L., and Fu, L. (2015). Mechanisms of resistance to EGFR tyrosine kinase inhibitors. *Acta Pharm. Sin. B* 5, 390–401. doi: 10.1016/j.apsb.2015.07.001
- Inhester, T., Bietz, S., Hilbig, M., Schmidt, R., and Rarey, M. (2017). Index-based searching of interaction patterns in large collections of protein-ligand interfaces. *J. Chem. Inf. Model.* 57, 148–158. doi: 10.1021/acs.jcim.6b00561
- Iskar, M., Zeller, G., Blattmann, P., Campillos, M., Kuhn, M., Kaminska, K. H., et al. (2013). Characterization of drug-induced transcriptional modules: towards drug repositioning and functional understanding. *Mol. Syst. Biol.* 9, 662. doi: 10.1038/msb.2013.20
- Iwata, H., Sawada, R., Mizutani, S., and Yamanishi, Y. (2015). Systematic drug repositioning for a wide range of diseases with integrative analyses of phenotypic and molecular data. *J. Chem. Inf. Model.* 55, 446–459. doi: 10.1021/ci500670q
- Jalencas, X., and Mestres, J. (2013a). Chemoisosterism in the proteome. *J. Chem. Inf. Model.* 53, 279–292. doi: 10.1021/ci3002974
- Jalencas, X., and Mestres, J. (2013b). Identification of similar binding sites to detect distant polypharmacology. *Mol. Inform.* 32, 976–990. doi: 10.1002/minf.201300082
- Jin, G., Fu, C., Zhao, H., Cui, K., Chang, J., and Wong, S. T. C. (2012). A novel method of transcriptional response analysis to facilitate drug repositioning for cancer therapy. *Cancer Res.* 72, 33–44. doi: 10.1158/0008-5472.CAN-11-2333
- Keiser, M. J., Roth, B. L., Armbruster, B. N., Ernsberger, P., Irwin, J. J., and Shoichet, B. K. (2007). Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* 25, 197–206. doi: 10.1038/nbt1284
- Keiser, M. J., Setola, V., Irwin, J. J., Laggner, C., Abbas, A. I., Hufeisen, S. J., et al. (2009). Predicting new molecular targets for known drugs. *Nature* 462, 175–181. doi: 10.1038/nature08506
- Kinnings, S. L., Liu, N., Buchmeier, N., Tonge, P. J., Xie, L., and Bourne, P. E. (2009). Drug discovery using chemical systems biology: repositioning the safe medicine Comtan to treat multi-drug and extensively drug resistant tuberculosis. *PLoS Comput. Biol.* 5:e1000423. doi: 10.1371/journal.pcbi.1000423
- Kitchen, D. B., Decornez, H., Furr, J. R., and Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* 3, 935–949. doi: 10.1038/nrd1549
- Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., et al. (2006). The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, 1929–1935. doi: 10.1126/science.1132939
- Laurie, A. T. R., and Jackson, R. M. (2006). Methods for the prediction of protein-ligand binding sites for structure-based drug design and virtual ligand screening. *Curr. Protein Pept. Sci.* 7, 395–406.
- Li, Y. Y., An, J., and Jones, S. J. M. (2011). A computational approach to finding novel targets for existing drugs. *PLoS Comput. Biol.* 7:e1002139. doi: 10.1371/journal.pcbi.1002139
- Liu, X., Ouyang, S., Yu, B., Liu, Y., Huang, K., Gong, J., et al. (2010). PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach. *Nucleic Acids Res.* 38, W609–W614. doi: 10.1093/nar/gkq300
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* 298, 1912–1934. doi: 10.1126/science.1075762
- Mervin, L. H., Afzal, A. M., Drakakis, G., Lewis, R., Engkvist, O., and Bender, A. (2015). Target prediction utilising negative bioactivity data covering large chemical space. *J. Cheminformatics* 7, 51. doi: 10.1186/s13321-015-0098-y
- Mestres, J., Martín-Couce, L., Gregori-Puigjané, E., Cases, M., and Boyer, S. (2006). Ligand-based approach to in silico pharmacology: nuclear receptor profiling. *J. Chem. Inf. Model.* 46, 2725–2736. doi: 10.1021/ci600300k
- Novac, N. (2013). Challenges and opportunities of drug repositioning. *Trends Pharmacol. Sci.* 34, 267–272. doi: 10.1016/j.tips.2013.03.004
- Sawada, R., Iwata, H., Mizutani, S., and Yamanishi, Y. (2015). Target-based drug repositioning using large-scale chemical–protein interactome data. *J. Chem. Inf. Model.* 55, 2717–2730. doi: 10.1021/acs.jcim.5b00330
- Sgobba, M., Caporuscio, F., Anighoro, A., Portioli, C., and Rastelli, G. (2012). Application of a post-docking procedure based on MM-PBSA and MM-GBSA on single and multiple protein conformations. *Eur. J. Med. Chem.* 58, 431–440. doi: 10.1016/j.ejmech.2012.10.024
- Shameer, K., Glicksberg, B. S., Hodos, R., Johnson, K. W., Badgeley, M. A., Readhead, B., et al. (2017). Systematic analyses of drugs and disease indications in RepurposeDB reveal pharmacological, biological and epidemiological factors influencing drug repositioning. *Brief. Bioinform.* doi: 10.1093/bib/bbw136 [Epub ahead of print].
- Stumpfe, D., and Bajorath, J. (2012). Exploring activity cliffs in medicinal chemistry. *J. Med. Chem.* 55, 2932–2942. doi: 10.1021/jm201706b
- Unterthiner, T., Mayr, A., Klambauer, G., Steijaert, M., Wegner, J. K., Ceulemans, H., et al. (2014). *Deep Learning as an Opportunity in Virtual Screening*. Available at: <http://www.datascienceassn.org/sites/default/files/Deep%20Learning%20as%20an%20Opportunity%20in%20Virtual%20Screening.pdf> [accessed March 20, 2017].

- Vasudevan, S. R., Moore, J. B., Schymura, Y., and Churchill, G. C. (2012). Shape-based reprofiling of FDA-approved drugs for the H1 histamine receptor. *J. Med. Chem.* 55, 7054–7060. doi: 10.1021/jm300671m
- Wang, Y., Bryant, S. H., Cheng, T., Wang, J., Gindulyte, A., Shoemaker, B. A., et al. (2017). PubChem BioAssay: 2017 update. *Nucleic Acids Res.* 45, D955–D963. doi: 10.1093/nar/gkw1118
- Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., et al. (2006). DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 34, D668–D672. doi: 10.1093/nar/gkj067
- Zhang, P., Wang, F., and Hu, J. (2014). Towards drug repositioning: a unified computational framework for integrating multiple aspects of drug similarity and disease similarity. *AMIA Annu. Symp. Proc. AMIA Symp.* 2014, 1258–1267.
- Zhang, Y.-M., Cockerill, S., Guntrip, S. B., Rusnak, D., Smith, K., Vanderwall, D., et al. (2004). Synthesis and SAR of potent EGFR/erbB2 dual inhibitors. *Bioorg. Med. Chem. Lett.* 14, 111–114.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 March-Vila, Pinzi, Sturm, Tinivella, Engkvist, Chen and Rastelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.