EXPERT OPINION ON DRUG DISCOVERY, 2016
http://dx.doi.org/10.1080/17460441.2016.1236786

Taylor & Francis
Taylor & Francis Group

EDITORIAL

Should network biology be used for drug discovery?

Francisco Martínez-Jiménez [a,b,c] and Marc A. Marti-Renom [a,b,c,d]

[a]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain; [b]Gene Regulation, Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), Barcelona, Spain; [c]Universitat Pompeu Fabra (UPF), Barcelona, Spain; [d]ICREA, Pg. Lluís Companys, Barcelona, Spain

ARTICLE HISTORY Received 12 July 2016; accepted 12 September 2016; published online 23 September 2016
KEYWORDS *In-silico* drug discovery; network biology; computational target identification; hit identification

1. Introduction

According to The Tufts Center for the Study of Drug Development, the development and marketing approval for a new molecular entity takes more than 13 years and around US$2.6 billion. Moreover, the cost of putting a new drug into the market has dramatically increased since the 1970s [1]. This raise in the drug development cost has led to a dramatic shrinkage of the efficiency, which is measured in terms of the number of new approved drugs per billion US dollars of research and discovery spending [2]. Factors that have contributed to the raise of drug development costs include increased clinical trial complexity, larger clinical trial size, greater assessment of safety and toxicity drug profiles, or evaluation on equivalent drugs to accommodate payer demands for comparative effectiveness data [2]. Simultaneously, the emergence of high-throughput technologies such as high-throughput screenings (HTS) or next-generation sequencing has led into a drug discovery paradigm shift from the traditional single drug perspective to a more target-centric view [3]. The application of these technologies alongside the increasing complexity of the treating diseases and the growing intricacy of the mechanism of action of the drug has also significantly increased the costs of preclinical stages. Hence, there is an urgent need for readjusting the drug discovery process to tackle these new problems. More specifically, modern drug discovery programs should be able to deal with the massive amount of data generated in the initial stages of the drug discovery pipeline [4].

In this scenario, computational methods can play a significant role to decline the time and costs of preclinical stages. Over the last 30 years, computational methods have helped to the development of new therapeutics [5,6]. However, we are still far from extracting all their potential when applied to the drug discovery field. This review is oriented to present how computational methods in general, and network-based methods in particular, could be used to optimize the preclinical stages of the drug discovery process.

2. Right target and right drug

Wrong selection of the molecular target (i.e. weak association between protein targets and the treating disease) implies lack of the expected efficacy, which is the most important cause of failure in clinical trials [7,8]. The lack of efficacy problem is even more prominent when dealing with complex and multifactorial diseases such as cancer or Alzheimer's disease. Consequently, selecting the right target requires the complete understanding of the entire interconnected system where molecular targets play very specific role(s) on a large and precise machinery. That is the reason of why many novel computational methods for target identification are including molecular networks to better represent the biological system to intervene [9,10]. Moreover, network-based representation enables the integration of multiple sources of information. Information such as protein–protein interaction (PPI), target druggability assessment, gene–disease association, compound–protein interaction, or protein–side effects association that eventually resembles the reality in which the decision of the molecular target is based upon multiple distinct factors.

Frequently, the selected molecular target plays multiple functions in the cell. Hence, inhibition of the target can lead to severe side effects that do not compensate the positive ones [11]. A less harmful alternative consists on the specific regulation of the molecular interaction that is associated to the treating disease. Application of network biology can help to identify new PPIs amenable to be disrupted by small-molecule treatments. Unfortunately, targeting PPIs is a very challenging task [12], and the progress is still limited to certain classes of PPIs [13].

The multifactorial nature of some diseases prevents them from using a single pharmacological intervention [14]. In such cases, the combination of multiple therapeutic entities may be required for treating the disease. However, the development of multidrug treatments leads to a more complex scenario, where it is not only important the identification of the molecular targets, but also, whether the combination of drugs provides a synergistic effect. The combinatorial explosion, consequence of the larger number of variables, prompted the development of computational methods aiming to predict which of those combinations have the best pharmacological profile. In spite of the accuracy of such methods is far from optimal [15], computational prediction of compound-pair activity will be needed in the future where combinatorial drug regimes will be more frequent as a consequence to the increasing understanding of the target diseases.

CONTACT Francisco Martínez-Jiménez francisco.martinez@cnag.crg.eu

3. Networks to the service of mechanism of action identification

Once the association between the molecular target(s) and the treating disease is validated, the drug discovery process focuses on identifying molecules able to perform the desired pharmacological activity through the modulation of the candidate target(s) (i.e. the so-called hits). A variety of screening approaches exist to identify hits molecules [16]. Some modern drug discovery pipelines include an HTS of library of thousands or even millions of compounds against the candidate target. However, mining into the massive amount of generated data in the search for hits with the desired bioactivity can be a challenging task. Moreover, the expenses associated with application of HTS prevent them from a broader application. This problem is chiefly evident in the academia environment, where the resources are usually limited and where the search for molecules usually pursuers a proof-of-concept goal instead of the development of new pharmaceutics. In such cases, the application of *in silico* methods for ligand–target interaction prediction provides a cheaper and more accessible opportunity. Over the last years, there has been a significant increase in the number of publicly available computational methods for compound–target prediction [6]. This improvement is partly consequence of the creation of public bioassays and screenings databases such as ChEMBL [17] or PubChem [18] among others, which led to the development of more precise and wider applicable computational methods. Most of these methods leverage biological networks to represent the vast and heterogeneous public data [19]. More specifically, they build interconnected network where nodes represent pharmacological and epidemiological entities (e.g. compounds, drug cocktails, proteins, side effects, biological assays, and populations like counties) and edges represent any type of relationship among them (e.g. similarity, interaction, and coexistence). The final network is eventually used by a predictive model that forecasts new associations between the compounds and the candidate target [20–24]. Some of these networks also provide structural information of the protein–ligand interaction [25,26]. In fact, structural information of the compound–target interaction is essential for a hit bioactivity improvement, which combined with administration, distribution, metabolism, excretion, and toxicity (ADMET) optimization further develops hits towards the so-called leads molecules [27].

4. Expert opinion

The increasing growth in the cost of bringing a new drug into the market is not sustainable. Hence, modern drug discovery programs should readjust their pipelines to incorporate new strategies able to deal with the emerging problems of preclinical stages. The classic paradigm of one disease, one drug, and one target is no longer valid for most diseases. Alternatively, current drug discovery programs confront much more complex scenarios, including the modulation of multiple targets or the combination of multiple drugs. Moreover, diseases such as cancer cannot be treated as a single entity. Rather, each cancer type has a unique fingerprint that needs to be individually considered for more precise treatment development. To make things worse, the emergence of drug resistance reduces the duration of clinical benefits in many cancer therapies. Consequently, preclinical models should integrate multiple factors coming from different sources of information. Computational network-based approaches can face this challenge. We specifically discussed how target validation and hit identification processes benefit from the application of such approaches. We also trust that the predictive power of computational methods will significantly improve with the increase in publicly available drug discovery data. This is particularly necessary in fields like protein–protein inhibition, in which the discovery of a new PPI inhibitor can lead to the identification of new targetable structurally similar PPIs [13]. However, the application of computational models to the continuously increasing amount of biomedical data gives raise to new problems and challenges. Future network-based models will have to handle unprecedented amount of data. Moreover, the intrinsic noisy nature of biological data manifests the need for new algorithms able to find non-evident relationships from multiple-source data. Hence, future research should also focus on the development of large-scale and noise-free algorithms, which should replace those previously created to deal with less amount and more homogeneous data.

In summary, the integration of computational network-based methods would not only reduce the time and expenses of preclinical stages but also lead to more precise medicines, which eventually translates into lower drug attrition rates.

Funding

The authors are supported by the Spanish MINECO (The Ministry of Economy and Competitiveness) through grant BFU2013-47736-P.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

ORCID

Francisco Martínez-Jiménez http://orcid.org/0000-0001-7804-1134
Marc A. Marti-Renom http://orcid.org/0000-0002-0151-4279

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Paul SM, Mytelka DS, Dunwiddie CT, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov. 2010;9:203–214. DOI:10.1038/nrd3078
2. Scannell JW, Blanckley A, Boldon H, et al. Diagnosing the decline in pharmaceutical R&D efficiency. Nat Rev Drug Discov. 2012;11:191–200. DOI:10.1038/nrd3681

•• **Review analyzing the reasons of the decline of research and discovery (R&D) drug discovery efficiency.**

3. Drews J. Drug discovery: a historical perspective. Science. 2000;287:1960–1964.

- **Seminal review describing the changes of the drug discovery industry over the last century.**

4. Chen B, Butte AJ. Leveraging big data to transform target selection and drug discovery. Clin Pharmacol Ther. 2016;99:285–297.
5. Chou KC. Impacts of bioinformatics to medicinal chemistry. Med Chem. 2015;11:218–234.
6. Sliwoski G, Kothiwale S, Meiler J, et al. Computational methods in drug discovery. Pharmacol Rev. 2014;66:334–395.

- **Comprehensive review about the role of computational methods in the drug discovery field.**

7. Hutchinson L, Kirk R. High drug attrition rates – where are we going wrong? Nat Rev Clin Oncol. 2011;8:189–190.
8. Lindsay MA. Target discovery. Nat Rev Drug Discov. 2003;2:831–838.
9. Wang YC, Chen BS. A network-based biomarker approach for molecular investigation and diagnosis of lung cancer. BMC Med Genomics. 2011;4:2.

- **A study leveraging networks to highlight the pathways and mechanisms of the lung carcinogenic process.**

10. Csermely P, Korcsmaros T, Kiss HJ, et al. Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. Pharmacol Ther. 2013;138:333–408. DOI:10.1016/j.pharmthera.2013.01.016
11. Wang X, Thijssen B, Yu H. Target essentiality and centrality characterize drug side effects. PLoS Comput Biol. 2013;9:e1003119. DOI:10.1371/journal.pcbi.1003119
12. Arkin MR, Tang Y, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. Chem Biol. 2014;21:1102–1114. DOI:10.1016/j.chembiol.2014.09.001
13. Scott DE, Bayly AR, Abell C, et al. Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. Nat Rev Drug Discov. 2016;15:533–550. DOI:10.1038/nrd.2016.29

- **Review that describes the state-of-the-art of the protein–protein interaction inhibition challenge.**

14. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol. 2008;4:682–690. Epub 2008/10/22. DOI:10.1038/nchembio.118

- •• **Review that introduces network polypharmacology as a change in the paradigm of the drug discovery field.**

15. Bansal M, Yang J, Karan C, et al. A community computational challenge to predict the activity of pairs of compounds. Nat Biotechnol. 2014;32:1213–1222. DOI:10.1038/nbt.3052

- **Original article of the DREAM drug–drug interaction challenge.**

16. Hughes JP, Rees S, Kalindjian SB, et al. Principles of early drug discovery. Br J Pharmacol. 2011;162:1239–1249. DOI:10.1111/j.1476-5381.2010.01127.x
17. Bento AP, Gaulton A, Hersey A, et al. The ChEMBL bioactivity database: an update. Nucleic Acids Res. 2014;42:D1083–90. DOI:10.1093/nar/gkt1031
18. Wang Y, Suzek T, Zhang J, et al. PubChem bioassay: 2014 update. Nucleic Acids Res. 2014;42:D1075–82. DOI:10.1093/nar/gkt978.
19. Gonzalez-Diaz H, Herrera-Ibata DM, Duardo-Sanchez A, et al. ANN multiscale model of anti-HIV drugs activity vs AIDS prevalence in the US at county level based on information indices of molecular graphs and social networks. J Chem Inf Model. 2014;54:744–755. DOI:10.1021/ci400716y.
20. Campillos M, Kuhn M, Gavin AC, et al. Drug target identification using side-effect similarity. Science. 2008;321:263–266. DOI:10.1126/science.1158140

- •• **One of the most cited and read drug target identification methods.**

21. Cheng F, Liu C, Jiang J, et al. Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput Biol. 2012;8:e1002503. DOI:10.1371/journal.pcbi.1002503
22. Yamanishi Y, Kotera M, Moriya Y, et al. DINIES: drug-target interaction network inference engine based on supervised analysis. Nucleic Acids Res. 2014;42:W39–45. DOI:10.1093/nar/gku337
23. Woo JH, Shimoni Y, Yang WS, et al. Elucidating compound mechanism of action by network perturbation analysis. Cell. 2015;162:441–451.

- **Article describing a method leveraging network perturbation analysis to predict compound mode of action.**

24. Duran-Frigola M, Rossell D, Aloy P. A chemo-centric view of human health and disease. Nat Commun. 2014;5:5676.
25. Kalinina OV, Wichmann O, Apic G, et al. ProtChemSI: a network of protein-chemical structural interactions. Nucleic Acids Res. 2012;40:D549–D553. Epub 2011/11/24.
26. Martinez-Jimenez F, Marti-Renom MA. Ligand-target prediction by structural network biology using nAnnoLyze. PLoS Comput Biol. 2015;11:e1004157.
27. González-Díaz, H. ADMET-Multi-Output Cheminformatics Models for Drug Delivery, Interactomics, and Nanotoxicology. Current Drug Delivery. 2016;13:1.