Josephine Chen
December 2, 2014

Computational Drug Repositioning

Introduction

As a high school senior doing research at Stanford Medical School as part of the SIMR program, I remember listening to a seminar talk given by Professor Atul Butte about his lab's research in repurposing drugs for other diseases. It was such an obvious concept in a sense (the drug's already passed FDA regulations so why not see what other diseases it could treat?), yet it was not until computational methods were developed that the drug repurposing become a systematic, streamlined process rather than one of largely luck.

Background

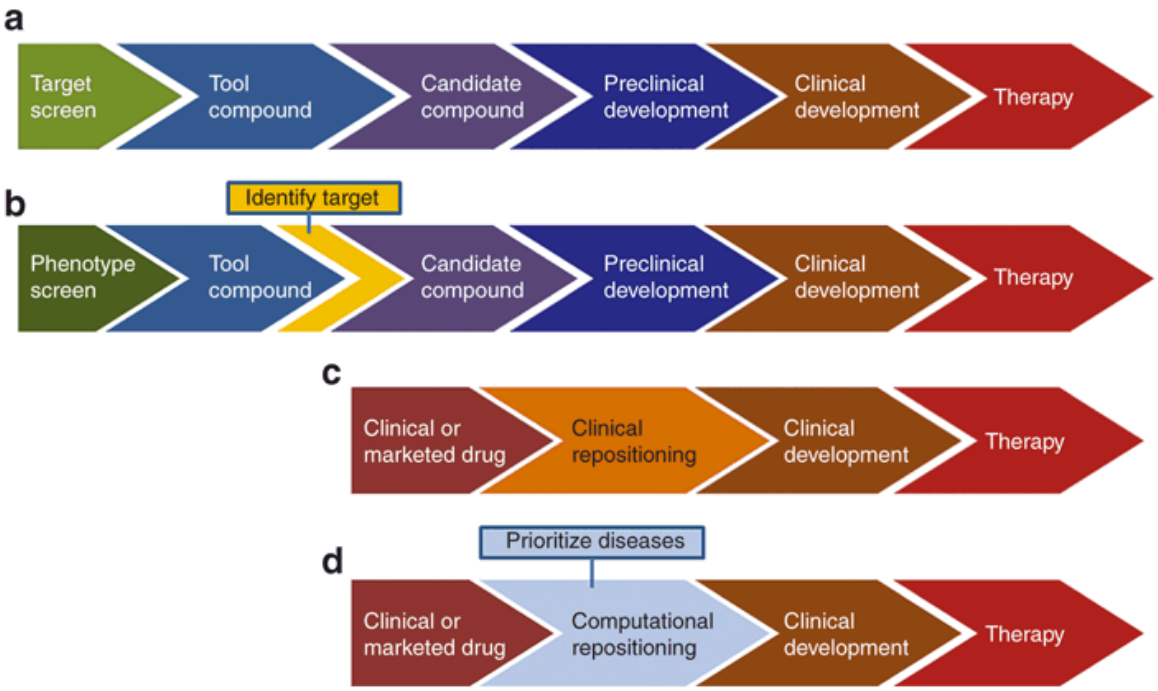
Since the development of genomics, drug discovery has largely been centered around the discovery of a new therapeutic target that acts through a specific mechanism. The drug target is sometimes genetically linked to the disease, or a biochemical assay screen is done of the target based on the predicted mechanism. Based on the results of these screens, compounds that affect the activity of the chosen target are then identified and optimized to reduce side effects due to off-target binding or unanticipated roles of the drug target (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html). This process tends to be slow, low-yield, and costly. The second common drug discovery process involves a phenotypic screen of the model system for compounds that could be efficacious. Through these two processes, only fifty "first-in-class small-molecule agents" were approved by the FDA (Food and Drug Administration) between 1999 and

2008, seventeen of which through the target-based discovery process and twenty-eight from the phenotypic screen process (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html). A single drug currently requires around 15 years and $800 million to $1 billion of research and development and clinical trials to bring to market (http://bib.oxfordjournals.org/content/12/4/303.long). The figures are intimidating.

While the processes described above are the most common ways pharmaceutical companies approach drug discovery, some very successful drugs currently on the market were not discovered strictly through those processes but through drug repurposing. These drugs were not developed for their current purposes but were repositioned for their use based on accidental discovery or unintended consequences. Examples include Viagra, which was originally meant to be a drug for angina but is now famously used to treat erectile dysfunction, and Avastin, which was studied for metastatic colon cancer and non-small-cell lung cancer but was approved for metastatic breast cancer (http://bib.oxfordjournals.org/content/12/4/303.long).

Drug repositioning or repurposing, which involves the discovery of new indications for existing drugs outside their original indications, is becoming an increasingly attractive method of drug discovery. It can renew a failed a drug or increase the number of indications for existing drugs; moreover, it can reduce the timeline for getting a new drug onto market, saving both time and money. The drug development timeline for a repurposed drug can be as short as 3 to twelve years, as several steps of the development pipeline can be skipped (http://bib.oxfordjournals.org/content/12/4/303.long). Another advantage of drug

repurposing is the fact that existing drugs have already been vetted by the FDA for safety and toxicity, so the repurposed drugs can enter clinical trials much more quickly (http://cancerdiscovery.aacrjournals.org/content/3/12/1364.long). The figure below illustrates the reduced timeline of drug repositioning compared to normal novel drug development.



(http://www.nature.com/clpt/journal/v93/n4/fig_tab/clpt20131f1.html#figure-title)

Drug repositioning techniques currently take the form of finding new targets for a known compound or finding new indications for known targets. The former approach leverages the fact that single molecules can act on multiple targets (in fact, off-target interactions are common among many approved drug compounds); in diseases where the additional targets are relevant, these molecules could serve as new potential therapeutics (http://bib.oxfordjournals.org/content/12/4/303.long). A second approach to drug repositioning takes advantage of the fact that many diseases or biological processes share

the same mechanisms and targets. As a result, drugs for a certain target could potentially be used for multiple indications.

However, while there are several clear advantages to drug repositioning and pharmaceutical and biotechnology companies have the knowledge that can enable it, finding new uses for existing drugs tends to be a rather haphazard, serendipitous process. The challenge in this procedure is in determining which new indications to test for a drug of interest; computational approaches developed in the last decade seem to hold promise in helping to guide and select these new indications for drugs (http://bib.oxfordjournals.org/content/12/4/303.long). Computational drug repositioning involves designing and validating automated workflows that will "generate hypotheses for new indications for a drug candidate" (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html). Because computational methods allow researchers to generate, evaluate, and prioritize data for many drugs and diseases simultaneously, they can amplify the productivity of traditional drug repositioning techniques (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html). These efforts are constantly enhanced by the expanding databases in the literature. In fact, these computational techniques can oftentimes be used for not just drug repositioning but also for finding the initial indications for a drug.

Computational Repositioning Methods

Many computational drug repositioning strategies have been developed over the years, and this paper will give an overview of a few of these methods. The strategies can

broadly be classified into two categories: drug-based strategies and disease-based methods. For the former, one starts from the chemical or pharmaceutical perspective in trying to find opportunities for repositioning in a drug. This method is preferred where rich pharmacological data for the drug is available. For the latter strategy, one starts by examining the symptoms, mechanism, or pathology of a disease, and this approach is favored in cases where efforts are focused on a specific disease or therapeutic category or when information about the drug is lacking (http://bib.oxfordjournals.org/content/12/4/303.long).

Drug-based Computational Approaches

Chemical Similarity

The efficacy of a drug compound as a therapeutic is largely tied to its structure and chemical properties; as a result, shared chemical characteristics between different drugs poses opportunities for drug repositioning. Although similar structures of compounds do not always result in the same behavior in a biological system, there are known quantitative relationships between chemical structures and biochemical activity (http://bib.oxfordjournals.org/content/12/4/303.long).

Finding drug repositioning opportunities based on chemical similarities involves extracting the chemical features for each drug in a set of drugs and then using the extracted features to relate the drugs to each other by clustering or constructing a network based on the features. Simple chemical associations and enriched biological features, such as drug targets, can then be identified in these clusters or networks (http://bib.oxfordjournals.org/content/12/4/303.long).

However, because this method requires full knowledge of the chemical properties of the compounds, it has several limitations. Many structures or chemical properties of drugs are proprietary information safeguarded by pharmaceutical companies; in addition, even the structures of disclosed drug compounds can sometimes contain errors. Lastly, because the way the body processes drugs and the metabolic and pharmacokinetic transformations of the drug as it is processed are largely uncharacterized, it is difficult to predict the effects of a drug based on its chemical properties alone (http://bib.oxfordjournals.org/content/12/4/303.long).

Molecular Activity Similarity

Computational methods can be used to assess the similarities in the molecular profiles of drugs to relate them to other drugs and diseases. While the precise mechanism of action is not understood for many drugs, high-throughput molecular measurement techniques, like gene microarrays, can be used to illustrate the effects of the molecular activity of a compound on a biological system. These molecular activity profiles can then be compared to establish therapeutic relationships between drugs and diseases (http://bib.oxfordjournals.org/content/12/4/303.long).

The Connectivity Map project (CMap) is one of the most comprehensive approaches to using transcriptomic data to connect expression profiles across conditions to propose drug repositioning targets (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html)–I actually used it myself when I was doing research in Professor Contag's lab! The Connectivity Map currently has gene expression profiles for more than a thousand compounds by exposing

these compounds to a few cancer cell lines and measuring the molecular activity response, which includes changes in transcriptional activity. Based on similarities in molecular activity shown in their CMap profiles, drugs can be connected to other drugs and diseases (http://bib.oxfordjournals.org/content/12/4/303.long). *In vivo* models have validated predictions made using this approach, including the use of topiramate, an anticonvulsant drug, for inflammatory bowel disease (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html).

One way the CMap database can be used to compare the molecular profiles of the drugs themselves to suggest how an indication for one drug can be an additional indication for a drug with a similar profile (http://bib.oxfordjournals.org/content/12/4/303.long). Another way the database can be used is the compare the molecular activity profiles of drugs with those of a disease state. If one thinks about a disease state as a perturbation of a normal biological state, we can apply the same method of creating drug molecular profiles to measure the genome-wide transcriptional changes of the disease condition to create its molecular activity profile. This disease profile or signature can then be used to compare with the profiles of drugs to identify therapeutic opportunities (http://bib.oxfordjournals.org/content/12/4/303.long). This method was used to identify rapamycin, an mTOR inhibitor, as a modulator of glucocorticoid resistance in acute lymphoblastic leukemia (http://bib.oxfordjournals.org/content/12/4/303.long).

One of main limitations of relying on the CMap database for drug repositioning is the way in which the molecular profiles are created. The molecular activity profiles are created by exposing the drug compound to various cancer cell lines, which may not

reflect the biological activity of the drug *in vivo*. In fact, many drugs undergo chemical transformations when they are metabolized, and these changes are neglected in the creation of the profiles. In addition, because many diseases affect multiple tissues and organ systems, it is difficult to represent them as single molecular activity profiles (http://bib.oxfordjournals.org/content/12/4/303.long).

Disease-based Computational Approaches

Side Effect Similarities

The side effects method involves discovering new disease indications for existing drugs by identifying drugs with similar side effects, which represent unintended consequences of drug action. Both therapeutic effects and side effects demonstrate the physiological consequences of a compound's biological activity; as a result, side effects can also potentially serve as phenotypic biomarkers for the disease treatment (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html). If therapeutics for the same disease work through different mechanisms but share the same (uncommon) side effects, one could hypothesize that there is an underlying pathway or mechanism that links the side effects and the treatment of the disease. (http://bib.oxfordjournals.org/content/12/4/303.long). This provides the foundation on which to relate drugs to other drugs or diseases through side effect profiles.

An example of this method is if many drugs addressing transplant rejection report increased cytomegalovirus infections as a side effect. This information could then be used to form the hypothesis that drugs that have increased cytomegalovirus infections, as a possible side effect may also be potential treatments for transplant reject. One such drug

is methotrexate, which, in fact, has been reported to have been used for transplant rejection (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html).

One of the main advantages of the side effect similarity method is the fact that there are no translational issues, with the side effects and therapeutic effects both observed in human patients rather than animal models. On the other hand, this method requires drugs having well-defined side effect profiles, which is oftentimes not the case. Side effects for many drugs are not fully fleshed out until years of being on the market and in clinical use. Moreover, because many drugs have so many side effects, there is oftentimes a lot of noise in determining unique, common side effects (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html). Lastly, similar phenotypic expression of a drug side effect does not necessarily always indicate common underlying pathways; many different underlying mechanisms, from hormonal to the immune system, could result in the same phenotypic side effect, such as hair loss (http://bib.oxfordjournals.org/content/12/4/303.long).

Genetic Similarities

Computational methods that assess molecular relationships between different diseases can serve as another approach for drug repositioning. Genetic studies can strongly connect specific genes with specific diseases. When diseases, even those that appear very different at the phenotypic or clinical level, are found to be similar at the genetic or molecular level, repositioning opportunities exist. Drugs can be evaluated for repurposing when their target is genetically associated with another disease the drug was not initially indicated for.

However, because disease pathology often involves many different organ systems, tissues, and molecules, it can be difficult to model the molecular basis of a disease in such a way that it can be easily compared with that of other diseases for computational purposes. Network-based approaches have been proposed as an alternative to address the difficulty of modeling complex molecular disease states (http://bib.oxfordjournals.org/content/12/4/303.long).

Genome-wide association studies (GWAS) have demonstrated the relationship between genetic variants and diseases, which has allowed for the linking of genes close to these variants with many complex diseases. Francis Collins found that 6 of the 44 GWAS loci for type 2 diabetes could be linked to drugs on the market, and this finding was broadened to hypothesize that GWAS-identified genes could more likely be targets for small molecules and biological (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html). While 155 GWAS genes have been identified as being targeted by at least one drug in the market or in development, 92 of those target genes are being addressed by drugs whose disease indication is different from the disease trait identified by GWAS. It would be interesting to see whether those drugs could be repurposed to address the disease traits indicated by GWAS.

However, while GWAS does allow for the identification of many gene-disease associations that present opportunities for drug repositioning, there are limitations to using GWAS data. For instance, it is different to determine from the GWAS data alone whether an activator or inhibitor is required to treat the disease (http://www.nature.com/clpt/journal/v93/n4/full/clpt20131a.html).

Conclusion

As a result of the slow pace in current drug discovery, the attractiveness of drug repositioning to find new indications for current drugs should only grow. While drug repurposing has formerly often involved much chance and luck, computational methods present a more efficient way to propose new disease indications for drugs. In addition, as molecular and genetic data for both diseases and drug mechanisms becomes more common, the power of these computational methods will only grow. However, it is important to remember that extensive clinical trials are still needed to fully demonstrate the efficacy of computational drug repositioning.