Adam Poole

Final Paper
Drug Screening and Genomics

Beginning in the late 20th century, techniques for drug discovery have improved due to advancements in technology. One of the most practiced methods is drug screening. In this process, scientists can choose and test a new drug against a chosen target for a particular disease using only a computer. In this type of high-throughput screening (HTS), large libraries of chemicals are tested for their ability to modify the target. Computers go through these large databases of chemicals and select which chemicals might work best based on the chemical's characteristics; such as size and shape. Shape is a very important factor in signaling pathways because inhibitors and activators are specific to their corresponding binding sites. Because of this, the idea is to use a different chemical of similar structure to manipulate different pathways within the body; however, it is not always as simple as finding a matching structure to discover a new drug. There are a lot more factors that go into uncovering a functioning medicine using drug screening. The process from identifying the target to getting the drug actually approved for consumers is a very complicated process, but with the help of our growing knowledge of genomics, this process could soon be completed much easier.

Before explaining the processes behind drug screening, its important to know its origin and why it was needed in order to progress modern medicine. The history of drug discovery in the pharmaceutical industry and academic labs over the past 50 years shows a progression of discovery examples that began shortly after the production of penicillins became available to the public after World War II. That same decade also witnessed the

growth of synthetic organic chemistry, which had progressed to the point that the large-scale preparation of artificial drugs or drug candidates was economically feasible. Synthetic organic chemistry was a very important advance at the time, particularly because bacteria had begun to develop resistance to the natural penicillins. Synthetic chemistry provided the ability to prepare analogs that proved to have activity against resistant strains; however, there were still many diseases for which there were no effective therapeutic interventions, but scientists were sure synthetic chemistry offered a solution. Another revolution was beginning about that same time, which was sparked by the commercial availability of extraordinarily powerful spectrometers (especially NMR and MS) and separation techniques (HPLC) for determining the structures of minute quantities of biologically active natural products. These products had been isolated, identified, and then screened in panels of assays for various types of desired activity, such as toxicity against cancer cell lines. As attention began to shift away from random searches for active natural products to a new computational model for drug discovery called computer-aided drug design, driven by dramatic increases in computer power in the early 1980s and also by significant concurrent advances in structural biology that provided plenty of new protein structures, which were used to base computational drug design studies. It seems likely that integrating sophisticated new computational, bioinformatics, pharmacogenomics, engineering, and nanotechnology methods into the process will lead to the next stage of advances in drug discovery.

As computational drug screening evolved, so did the process's complexity. There are many steps that must be taken in order to actually make a new drug using this

method. One of the most important steps in developing a new drug is target identification and validation. A target is a term that can be applied to a range of biological entities, which may include proteins, genes, and RNA. A good target needs to be effective, safe, meet clinical and commercial needs, and be accessible to the putative drug molecule. Scientists estimate there are about 8,000 therapeutic targets that might provide a basis for new medicines. Most of these are proteins of various types; including enzymes, growth factors, cell receptors, and cell-signaling molecules. Because some targets are present in excess during disease, the goal is to block their activity. This can be done by a medicine that binds to the target to prevent it from interacting with other molecules in the body. In other cases, the target protein is deficient or missing, and the goal is to enhance or replace it in order to restore healthy function. Effective target identification and validation allows us to explore whether target modulation will lead to mechanism-based side effects. Biotechnology and advanced computers have made it possible to create therapies that are similar or even identical to the complex molecules the body relies on to remain healthy. This includes data mining of available biomedical data libraries, which have led to a significant increase in target identification. Data mining refers to the use of a bioinformatics approach to help in identify, select, and prioritize potential disease targets. The data that is available comes from a large variety of sources including publications and patent information, gene expression data, proteomics data, transgenic phenotyping and compound profiling data. Identification helps determine whether mRNA and proteins are expressed in disease and if they are correlated with disease exacerbation or progression. However, it is very challenging to choose good targets since human biology is so complex. It can take many years of research and clinical trials to learn that a new

target won't provide the desired results. To reduce that risk, scientists try to prove the value of targets through research experiments that show the target's role in the disease process. Once identified, the target then needs to be fully justified. Validation techniques range from *in vitro* tools through the use of whole animal models, to modulation of a desired target in disease patients.

Once the target has been set and validated, the next step is to identify a drug that impacts the target in the desired way. The aim is to find a molecule that will interfere with only the chosen target, but not other related targets. A hit molecule is a compound which has the desired activity in a compound screen and whose activity is confirmed upon retesting. A variety of screening paradigms exist to identify hit molecules. High throughput screening (HTS) involves the screening of the entire compound library directly against the drug target or in another assay system. These assay systems show activity that is dependent on the target but the process will also require secondary assays to confirm the site of action of the compounds. This screening paradigm uses complicated laboratory automation but assumes no prior knowledge of the characteristics of the chemotype that is likely to have activity in the target protein. There are multiple kinds of assays, like cell-based and biochemical assays, that are established for a drug to become certified for retail. Normally, cell-based assays have been used to target classes such as membrane receptors, ion channels and nuclear receptors. In contrast, biochemical assays, which have been applied to both receptor and enzyme targets, often simply measure the similarity of the test compound for the target protein. The relative advantages of biochemical and cell-based assays have been debated and reviewed extensively. Both

assay paradigms have been used successfully to identify hit and candidate molecules. The choice of assay format depends on the biology of the drug target protein, the equipment in the laboratory, the experience of the scientists in that laboratory, and whether an inhibitor or activator molecule is sought and the scale of the compound screen. Whatever the assay format that is selected, there are certain requirements and factors that are considered, including its pharmacological relevance, reproducibility, cost, quality, and its overall effectiveness.

Another type of screening, called focused or knowledge-based screening, involves selecting smaller subgroups of molecules from the chemical library that are likely to have activity at the target protein based on previous knowledge of the target protein from literature or patent precedents. From here, other screening runs will be made to see whether the hits against the chosen target will interfere with other related targets. This is known as the process of cross screening. Cross screening is important because the more unrelated targets that a compound hits, the more likely there will be off-target toxicity that will cause side effects. There is also fragment screening, which involves the production of very small molecular weight compound libraries. These libraries are screened at high concentrations and are typically followed by the generation of protein structures to enable compound progression. Finally, there is a more specialized focused screening approach that can also be taken called physiological screening. This is a tissue-based approach and looks for a response more aligned with the final desired *in vivo* effect as opposed to targeting one specific molecular component. All the different types of screening are done with automated systems and allow scientists to rapidly test thousands

of compounds to see which ones interfere with the target's activity. Then, potential compounds can be put through added tests to find a lead compound with the best potential to become a drug.

Once a solid number of hits have been obtained from virtual screening or HTS, the first role for the drug discovery team is to try to determine which compounds are the best to work on. This process of prioritizing the compounds is essential because a team will likely be left with many possible hits, which they will then need to reduce, confirm and cluster into series. After the best possible compound for the drug has been determined, the final step of the actual laboratory process is the figure out the dosage. This is done by obtaining a dose–response curve, which allows the generation of a half maximal inhibitory concentration to compare of the potencies of candidate compounds. This will decide what amount of the compound works best and is still safe for the body. General molecules need to be examined in models of genotoxicity such as the Ames test and in *in vivo* models of general behavior. High-dose pharmacology, dose linearity and repeat dosing PK looking for drug-induced metabolism and metabolic profiling all need to be carried out by the end of this stage. The chemical stability and salt selection for the accepted drug substance must also be taken into consideration. Finally, after thousands of tests, a candidate is selected to go to the market. However, the attrition rate of compounds entering the clinical phase is also high, so only about 1 in 10 candidates actually reach the market. Furthermore, at this stage the financial consequences of failure are much higher making it a risky process. There has been considerable debate in the drug industry as to how to improve the success rate of drugs. It is preferred to fail fast and cheap instead of

failing after an extended, expensive process. Once a candidate reaches the clinical stage, it can become increasingly difficult to halt the project, mainly because at this stage the project has become public and termination can influence shareholder value. Preforming additional studies prior to clinical development such as improved toxicology screens, establishing predictive translational models based on previous knowledge of the disease, and identifying biomarkers may help in this effort. It are these areas where academic-industry partnerships could really aid in the attempt to increase value preclinically and eventually help bring more effective drugs to patients.

While drug screening has greatly contributed the progression of drug production, it is still a lengthy and financially risky process. However, with our growing knowledge of genomics, this process has gradually begun to simplify. Genomics has significantly promoted the drug development progress by predicting characteristics of candidate compounds before testing even begin. Not only are we now able to identify targets much easier, but also we can identify a drug's side effects, and its toxicity before trails commence. This initial elimination of unusable compounds is very helpful because it saves a great deal of time and money, and also helps diminish the risk of an extended, expensive failure. Moreover, genomics has provided the knowledge to ensure that a higher fraction of leads will work as expected in actual, live biological systems. Many times there are false positives that will work in the labs but not in the human body, so this information will increase the percentage of functional drugs. Ultimately, genomics will expand our capabilities in drug discovery and development even further.

http://biotechnology.amgen.com/developing-biotech-medicines

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3058157/

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3716285/