Summary
While small tweaks can make incremental improvements, no amount of drug candidate optimization can completely overcome clinical trial failures due to poor or limited efficacy, or safety issues if they stem from modulating a target with limited prevalence in the patient population or with high probability of side effects. Choosing the best target early in the discovery process is key to increasing the chances of regulatory approval success with minimal safety issues and maximum efficacy. This paper discusses a workflow that leverages the Pathway Studio knowledgebase to build a new disease model which enables the identification and prioritization of novel drug targets using published literature and experimental data. This approach potentially saves weeks or months of research time.
Case Study: The importance of selecting the most promising drug target

“The most promising targets are tightly connected to the disease of interest, have a proven function in the underlying pathology, and are observed with high frequency in the patient population.”

Increasing economic pressure due to late-stage drug approval failures is forcing drug discovery and development scientists to adapt their strategies to increase their chances of success (Fig 1). While focusing on optimizing drug candidates with respect to pharmacokinetics, efficacy and toxicity is one important aspect, perhaps even more critical to a drug’s ultimate success is ensuring that it modulates the right target in the first place. Most drugs end up failing due to poor efficacy in large populations or problems with toxicity. While small tweaks can make incremental improvements, no amount of drug candidate optimization can completely overcome these issues if they stem from modulating a target with limited prevalence in the patient population or with high probability of side effects.

Choosing the best target early in the discovery process is key to increasing the chances of regulatory approval success. The most promising targets are tightly connected to the disease of interest, have a proven function in the underlying pathology, and are observed with high frequency in the patient population. Optimal target modulation should affect the desired disease pathway more potently than any other pathway in which the target is involved, so that the predictable side effects are few or none. Thoroughly evaluating potential drug targets with respect to these properties requires scientists to develop a complete picture of the underlying disease biology and to understand a large number of intricate interactions involved in the related pathways.

Pathway Studio® enables scientists to quickly and easily gather this information and unravel these numerous and complex biological interactions, by allowing them to construct pathways based on literature data, imported experimental data or a combination of both. Derived using Elsevier’s proprietary deep reading natural language processing technology, this ever-growing knowledgebase contains molecular relationships from 4.3 million full-text scientific articles, 26 million PubMed abstracts and more than 325,000 clinical trials. These are among a list of source documents that provide a comprehensive dataset which helps enable more informed hypotheses generation and research decisions. Here we describe an approach that uses Pathway Studio powerful tools to identify and prioritize drug targets related to insulin resistance based on the targets’ potency in the insulin signaling pathway and experimental data from type 2 diabetic patients. This approach can potentially save weeks or months of research time in understanding disease biology and that is critically important to novel target discovery.
“Optimal target modulation should affect the desired disease pathway more potently than any other pathway in which the target is involved, so that the predictable side effects are few or none.”

A multi-faceted approach anchored in Pathway Studio

This multi-faceted approach uses Pathway Studio to identify potential drug targets that might be used to increase insulin signaling and then prioritizes these targets based on their proven linkage to insulin resistance pathology, as illustrated in the workflow in Figure 2. The first aspect involves developing a comprehensive and current overview of the insulin response (InsR) pathway and identifying significant pathway regulators based on Pathway Studio’s extensive molecular interaction database extracted from published literature. Many researchers use a limited set of literature to complete this step because it is very time consuming to read and annotate important relationships and because access to large corpuses of literature can be expensive. These restrictions may increase risks in experimental designs, hypothesis generation and clinical development.

The second aspect involves assessing experimental gene expression datasets to identify additional, potentially novel, targets that are differentially up- or down-regulated in insulin resistance most frequently among patients. The third aspect prioritizes the most promising regulator targets based on each target’s linkage to insulin resistance pathology in the scientific literature and using the Sub-Network Analysis tool in Pathway Studio. The scores are combined to enable a final target ranking and prioritization.

Figure 2. Overview of the target identification and prioritization strategy described here. Pathway Studio provides a comprehensive picture of the insulin response based on published literature and existing curated pathways. Using the tools in Pathway Studio, key regulators are identified using two methods: the advanced search query of literature data and analysis of experimental patient gene expression dataset. The insulin response pathway is expanded to also include glucose uptake and transportation. The potential targets are assessed and scored based on linkage to the InsR-glucose uptake and transportation pathway and how well they have been studied in the literature, as well as the Pathway Studio Sub-Network Enrichment Analysis. The targets are then ranked according to their compiled global scores.
**Insulin response and resistance**

Normally, after eating carbohydrates, the human body produces the hormone insulin, which signals cells to take up glucose from the blood and use it for energy. Insulin Resistance (IR) is a metabolic disorder in which cells cannot properly respond to insulin, resulting in high blood sugar that further increases insulin production. While this condition can be inherited, it is commonly associated with environmental factors, such as obesity and sedentary lifestyle, which are increasingly common around the world. Undetected insulin resistance can eventually lead to type 2 diabetes and cardiovascular disease, so identifying promising drug targets to increase insulin signaling is an area of great interest.

**Building the insulin response pathway by integrating multiple data sources**

To easily analyze and visualize the disease biology behind insulin resistance, we used Pathway Studio to construct a pathway based on primary literature data, detailing the proteins involved with insulin response and their particular roles. This analysis approach offers an unbiased perspective which often leads to unique insights based on the biology behind these molecular interactions. The resulting insulin response (InsR) pathway includes proteins responsible for insulin receptor activation and glucose uptake by cells.

To start this project, the search and pathway building functions in Pathway Studio were used to find over 400 proteins linked to both insulin resistance and glucose uptake and transportation. Review of the literature evidence supporting each interaction in Pathway Studio is enhanced as the actual sentence containing the extracted relationships is available for review. In addition, scientists can keep or remove a relationship from further analysis. Both capabilities allow for quick identification of the role of each protein in these processes based on the research scope – for example, analysis by tissue type, organ system, physiological response or other parameters. This enabled investigation on how these proteins interact with one another to transmit signals from the insulin receptor to proteins involved with glucose uptake. Taken together these data helped to inform the creation of the new insulin response-glucose uptake and transportation pathway.

Manual review of relations and corresponding literature allowed us to classify all proteins into three classes: proteins directly involved in regulation of glucose transport by insulin, protein regulators of insulin receptor-glucose uptake and transporting pathway, and remaining proteins that regulate insulin resistance in other human tissues. Finally, the review of the literature describing proteins directly involved in regulation of glucose transporting and trafficking by insulin in the target tissue provided additional insights to complete the main target pathway, InsR-Glucose uptake and transportation pathway (Fig 3).

Next, we used the “DirectRegulation” and “ProteinModification” relations in Pathway Studio to find out which proteins in the InsR-Glucose uptake and transportation pathway are downstream targets of the identified regulators. Proteins that could not be confirmed to be involved with insulin receptor regulation, as well as those that regulate PI3K and AKT1 in other pathways not related to InsR were removed. PI3K and AKT1 are very common signal transduction molecules that participate in many signaling pathways, and modulating these regulators is likely to cause unwanted side effects.

The remaining 131 regulators were ranked according to their ability to affect the InsR-glucose uptake and transportation pathway, based on the number of entities in the pathway they regulate in different directions as well as the significance of their targets in the InsR-glucose uptake and transporting pathway.

**Identifying upstream regulators from patient data**

In order to identify significant upstream expression regulators of the InsR-glucose uptake and transporting pathway in experimental data from diabetic patients, two public human gene expression datasets were assessed. Analyzing patient data in Pathway Studio makes it possible to easily identify regulators responsible for the differential expression between healthy and diseased individuals and determine which pathways are enriched with the most active expression regulators. This analysis helps confirm the findings from published literature, and can also identify potential novel drug targets that have not been previously described in the literature. In addition, this data further strengthens the prioritization ranking of selected targets.
Case Study: The importance of selecting the most promising drug target

This analysis used the GSE13070 and GSE59363 datasets downloaded from the GEO database. The fold changes for each gene and the corresponding p-values were imported into Pathway Studio and then examined using the Sub-Network Enrichment Analysis (SNEA) function to identify the major expression regulators and rank them according to their statistical significances. Additional experimental data sets can add greater confidence in the results.

Scoring each target based on the bibliometric evidence

We assigned a score to all of the identified regulators according to their linkage to insulin resistance in the scientific literature. Pathway Studio was used as a first order ranking to determine how many scientific references linked each target to insulin resistance. These results were then enriched using Elsevier’s Text Mining technology to measure the number of relationships available for each regulator in Pathway Studio, as an indicator of how well each regulator has been studied in the literature. Thus, each target was assigned three quantitative scores – two bibliometric scores and one patient-based score and the p-values from these were combined into a global p-value, on which the final prioritization was based. The next steps involve conducting experiments to confirm these findings on the selected targets using the appropriate model system, which are outside the scope of Pathway Studio.

Conclusion

Selecting the most promising drug targets is a critical first step to ensuring development of safe and effective drugs, eventual regulatory approval success, and suitability of a drug for the large market. Modulating a target that is poorly connected to disease pathology in the patient group can result in an ineffective drug, while modulating a target involved in multiple pathways can lead to unwanted side effects. Determining which targets are the best ones to pursue can be a daunting task, due to the numerous, complex interactions that must be considered. However, Pathway Studio and Elsevier’s Text Mining technology can now enable scientists to take advantage of vast knowledgebases and ways of prioritizing targets to easily approach this challenge in just a simple workflow that could save weeks or months of research effort. The resulting comprehensive overview of the disease pathway and a quantitative measure of how strongly each target is linked to the disease equips them to move their discovery and development efforts in the right direction with more informed research and better positions them for success.

Figure 3. This pathway illustrates the entities involved with glucose uptake and vacuolization, transport and release through the plasma membrane. Blue arrows indicate the direction of processing.
Pathway Studio helps customers streamline drug target and biomarker discovery and supports drug repurposing projects by providing biological researchers with greater insights into the mechanism of disease and facilitating more comprehensive interpretation of experimental data.

LEARN MORE

To request information or a product demonstration, please contact us at www.elsevier.com/pathwaystudio/contact-sales.

Visit www.elsevier.com/rd-solutions or contact your nearest Elsevier office.