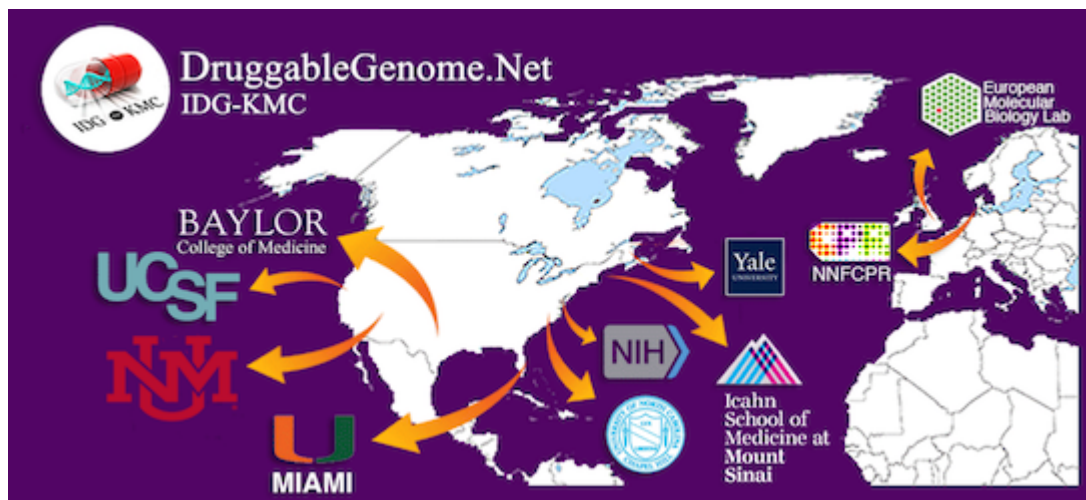


# Scientists to Explore the ‘Dark Genome’ for New Drug Targets

Scientists in the international Illuminating the Druggable Genome (IDG) project, including the University of Miami Miller School of Medicine, are taking an unprecedented approach to discovering new medical therapies. They’re exploring the 38 percent of the human genome that is considered “dark” or largely untapped – so far -- for potential new drug targets.



Traditionally, investigators focus their attention and resources on the less than

10 percent of the human genome associated with proteins that are already very well studied. These genetic regions code for well-known proteins already involved in approved therapies, feature known mechanisms of actions, or strongly bind to small molecules – a characteristic that raises their prospects for new medication development. Now an international collaboration of researchers is changing the paradigm, as reported in [“Unexplored therapeutic opportunities in the human genome”](#)

in the February 23 issue of *Nature Reviews Drug Discovery*.

Turning instead toward the “dark genome” for promising drug targets is a bold step. Researchers are motivated by the potential to discover novel therapies that ultimately improve the lives of people living with a number of diseases and conditions. The initiative also relies on new technologies and ways of thinking – including that a vast amount of data could become [more manageable and meaningful when classified into different ontologies](#), or related categories.

“Our group primarily contributed the Drug Target Ontology [DT0],” said Stephan Schürer, Ph.D., Program Director, Drug Discovery, University of Miami Center for Computational Science, and associate professor of molecular and cellular pharmacology at the Miller School of Medicine. In other words, Schürer and his team are contributing to creating the data infrastructure and organization system to accelerate this avenue of research, not just for the dark genome but eventually for the entire proteome as well. The scope and components of the DT0 are illustrated in an [easy-to-understand, interactive online visualization](#).

One goal of the DT0 component is standardizing how knowledge about drug targets is collected, stored and accessed by researchers. In contrast with traditional technologies to capture and organize information, Schürer’s team applies semantic web technologies, including a formal language -- known as OWL, or web ontology language – to formalize knowledge and simplify the ability of researchers to mine drug target data effectively.

Their efforts are part of the [NIH’s ‘Illuminating the](#)

[Druggable Genome' \(IDG\) initiative](#). As part of this NIH Common Fund initiative, Schürer has been appointed to the IDG's Resource Dissemination and Outreach Center (RDOC) to organize the incoming data and coordinate all the project's resources.

Along with researchers at other institutions in the U.S. and Europe, the UM team will help group proteins into one of four "target development level" categories. In addition to the 38 percent of the proteome containing proteins from the "dark genome," the biology group contains the 53 percent of human proteins with some information on structure and function. The chemical category refers to the 6 percent of the proteins with a high affinity for binding to small molecules. Interestingly, the clinical group – the most thoroughly studied proteins to date -- represents only 3 percent of the total.