

# Major Update of LINCS Research Database Makes Searches Simpler and Faster

Searching an extensive research database at the University of Miami just got simpler and speedier – potentially expediting the work of investigators looking for properties of promising compounds, the viability of cell lines, the expression of certain genes, and more.

“Researchers can now search all data easier and faster,” said Stephan Schürer, Ph.D., program director for drug discovery at the University of Miami Center for Computational Science, and professor of molecular and cellular pharmacology at the Miller School of Medicine.



Driven by the goal of providing easier access to the most relevant data, Dr. Schürer and team [announced the new features](#) of their revised Library of Integrated Network-Based Cellular Signatures (LINCS) Data Portal 2.0 in the January issue of the journal *Nucleic Acids Research*.

“The changes were made based on user feedback and our own

analysis,” he said. He is most proud of “the ability to search across all metadata at the same time.”

The [LINCS Data Portal 2.0](#) features a “better user interface and user experience design,” Dr. Schürer said.

The simplified home page now features three categories – Perturbations, Model Systems and Signatures.

Small molecules, gene knockdown and microenvironments are examples of perturbations. Cell lines, embryonic stem cells and primary cells fall under the model systems heading, while gene expression, cell phenotype and protein binding are types of signatures.

In fact, Dr. Schürer and colleagues point out in the report that “the most important change in LDP 2.0 was the transition from the LINCS dataset packages to computable LINCS signatures.”

In the past, searching the database might produce too much of a good thing. In other words, downloading a dataset package was analogous to getting a whole folder of information off a computer when someone only needs one file within the folder. With a data package, “You have all you need there, but you need the skills to manage it.”

Furthermore, if a search in the past produced results across multiple datasets, it meant more work for the user. For example, a researcher would have to download each dataset separately, then filter and aggregate the information to narrow it down to what they need.

Now one search can yield results based on a small molecule or

cell line name, target, associated disease, tissue, or mechanism of action, etc.

The portal thus identifies relevant query results based on extensive and standardized metadata annotations.

For example, if a researcher searches the database for ‘EGFR’ (epidermal growth factor receptor), they can get small molecule drugs that work via EGFR inhibition as their mechanism of action or other molecules that inhibit EGFR kinase and all the signatures that are associated with these molecules. The query would also return other perturbations, such as EGFR knockdown (via single guided RNA or CRISPR) and all signatures that quantify EGFR (e.g., gene expression or kinase inhibition). If a user typed in “prostate” the system would return cell lines that are associated with prostate cancer or prostate organ/tissue.

“Bringing all these different contexts under the same user interface [UI] was a complicated endeavor which we solved by our unique home page UI,” the researchers wrote.

The overarching goal remains simplifying a complex database. “To make the portal easy to use across both computational and experimental researchers, it was important to design a simple but informative web interface that would require a minimal learning curve for users,” they add.

The LINCS 2.0 enhancements reinforce the University of Miami’s position on the forefront of metadata management. LINCS is funded by the NIH through its Common Fund Project and [UM serves as a Data Coordination Center](#).

The updated data portal is not only designed to save time for research searches, it could expedite basic research and translation of findings to the clinical setting as well.

For example, the portal includes data for more than 20,000 small molecules. The information includes approved drugs, compounds in clinical trials, and tool compounds used in research. The updated LINCS Data Portal 2.0 also features extensive properties and target annotations of compounds and molecules, helping researchers improve their experimental designs. The comprehensive information also helps users identify, in advance, characteristics of small molecules or drugs such as absorption, distribution, metabolism and excretion.

The information “could suggest if it’s a good idea to develop this drug further – would it be orally available?” Dr. Schürer said. “You can also find a compound and want to know if it will work or is potentially problematic.” Providing such information also could ultimately reduce the probability of early failure in clinical trials.

Despite the new updates just announced, the LINCS project will continue to evolve. One future goal is to enhance interoperability of the data so it can be shared among external systems, further boosting its usefulness to researchers.

“That is a big challenge,” Dr. Schürer said. Connecting the data to other, external data means “everything needs to be standardized – using ontologies, reference IDs, and normalization to make data comparable at the signature level.” It is one of the components in the [FAIR guiding principles](#)

(Findable, Accessible, Interoperable, and Reusable) espoused by the NIH within the Common Fund initiative.

“Through the new scalable data infrastructure and modular UI design,” the researchers note, “we are ensuring the longevity of the data portal and its positioning as a central analytical hub that will grow as more signatures, methodologies, and tools become availa