

THE AUTHOR FILE

Jeanne Loring and Franz-Josef Müller

An open-access tool tells researchers whether a human cell line is pluripotent.

In 2008, Jeanne Loring and Franz-Josef Müller showed they could teach a computer to distinguish between different sorts of human stem cells¹, but what they really wanted was a program that could report whether or not a given cell line was pluripo-



Jeanne Loring and Franz-Josef Müller.

tent. Now they have built such a program. The bioinformatic assay called PluriTest allows researchers to upload microarray data and get an assessment of a cell line's pluripotency in about ten minutes².

The goal is a cheaper, quicker alternative to the much-hated but currently indispensable teratoma assay, in which putative pluripotent cells are

injected into a mouse lacking an immune system. If, over six to eight weeks, the cells grow into a bizarre tumor containing cells representing major tissue types, the cell line is declared pluripotent. "If you get a whole bunch of [stem cell scientists] in a room and ask who thinks a teratoma is necessary, no one will raise a hand. If you ask whether it's necessary to publish, everyone will," says Loring, a biologist at The Scripps Research Institute. And though teratomas vary considerably, detailed information is rarely accessible outside the laboratory where the assay was conducted.

To build the dataset on which PluriTest is based, Loring, Müller and colleagues collected samples from all over the globe, generating gene expression data from hundreds of human embryonic stem cell lines, dozens of human induced pluripotent stem cell lines plus hundreds of more-differentiated cell types. Although a separate pluripotency assessment based on more comprehensive profiling of fewer stem cell lines has also been published³, PluriTest was designed to accommodate data that researchers could collect readily in their own labs.

Müller, a bioinformatician at the Zentrum für Integrative Psychiatrie in Kiel, Germany, says the main barrier was making a model that could withstand the specter of the 'black swan', in this case, a cell that is both clearly pluripotent and clearly abnormal. "With the classical [machine-learning] approach, the predictor will either say this is a perfect swan or that this is a different type of bird altogether."

The researchers spent a year stumped for a solution, and then coauthor Bernhard Schuldt called Müller with an idea. Müller was at a conference and the call came very early in the morning. Müller recalls pacing an empty hotel lobby on his cell phone saying "yes, yes, that will work." Schuldt had proposed calculating a 'novelty score', which involves deconstructing pluripotency into genes that are regulated together. "We can tell the researcher that [a cell] looks in all these parts like a pluripotent cell, but this [other] part is off the charts," explains Müller.

The researchers fed black swans into the model in the form of data from teratocarcinomas and parthenogenetic cell lines, and were satisfied with PluriTest's assessment: pluripotent, yes, but unusual, too. PluriTest also distinguishes partially from fully reprogrammed induced pluripotent stem cells and even flags abnormalities that can be hard to detect. "If there is a genetic mutation that has an effect, we won't know what it is, but we can identify [the line] as abnormal," says Loring.

Müller and Loring plan to extend PluriTest in several ways: for example, to assess differentiation, perhaps a NeuroTest for neural cells or CardioTest for cardiomyocytes could be created. As for PluriTest itself, the hope is that researchers who use the test will also share data, making PluriTest more robust. Moreover, Müller is working on the best way to fold in data from additional microarray platforms as well as sequencing data and information about genome methylation. Eventually, Müller hopes, PluriTest will become a searchable, Google-like repository of stem cell information. "We're going to go the Facebook direction, where people can talk to each other about their data and their findings."

Monya Baker

**"How do you predict something you cannot anticipate?"
—Franz-Josef Müller**

1. Müller, F.J. *et al.* Regulatory networks define phenotypic classes of human stem cell lines. *Nature* **455**, 401–405 (2008).
2. Müller, F.J. *et al.* A bioinformatic assay for pluripotency in human cells. *Nat. Methods* **8**, 315–317 (2011).
3. Bock, C. *et al.* Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439–452 (2011).