

air pressure, than moving around fluid that clogs a partially collapsed airway.

However, the researchers acknowledge shortcomings in their model. Airways are not infinitely long tubes, as Heil and Hazel assumed. “The airway branches are relatively short before they branch again,” says Heil. “You have this tree structure, and that is something we do not take into account.”

Oliver Jensen, PhD, professor of applied mathematics at the University of Nottingham, United Kingdom, says the new model is a step forward compared to older ones. “They’ve developed an absolutely wonderful tool,” says Jensen. He notes the model also should apply to other systems with fluid-lined tubes, such as blood vessels.

For doctors who treat collapsed airways, Heil and Hazel’s work eventually could lead to fine-tuned air pressure for different patients. For now, when doctors sit down to use a ventilator, says Jensen, “It’s nice if they can at least understand what is happening in the airways.”

—*Sarah CP Williams*

New Algorithm Finds Stories in Biomedical Literature

A good story ties up all the loose ends. A new data-mining tool takes a stab at doing the same. Dubbed *storytelling*, the algorithm may make it easier to unearth unexpected connections in the avalanche of freshly published research, or among high-throughput datasets. For example, *storytelling* can sift through tens of thousands of PubMed abstracts to discover scientific links between two apparently unrelated topics; or draw connections across a knowledge structure such as the Gene Ontology.

“What we are trying to do is link data sets very far apart,” says **Richard Helm, PhD**, associate professor of biochemistry at Virginia Polytechnic Institute. “In the end, we link data set A with data set Z in the form of a story.” The work was presented at the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2006), in August 2006.

Researchers might not have time to find complex relationships waiting for discovery in the literature, but *storytelling* does. It finds key documents bridging from one research publication (the starting point) to another (the end point). Using System X, a Virginia Tech supercomputer, the algorithm first classifies each PubMed article’s abstract into an organized branched set of terms. It can then make thousands of comparisons and join related publications into a chain connecting start to finish.

For example, Helm and his colleagues, used *storytelling* to dig through the literature seeking ties between two remotely related papers: one on tomato genes expressed in yeast and another on how chemical stress affects yeast gene expression. The supercomputer boiled down 140,000 yeast publications to nine abstracts—stepping stones from paper one to paper two. The results included a paper that identified a novel protein, expressed only when yeast cells are exposed to cadmium, which researchers might not have immediately connected with the first two papers. Although the paper might have surfaced in an ordinary PubMed search, it would have required much sifting to find it. While not every search will yield treasures,

hopefully most results from *storytelling* will provide new insights and hypotheses that researchers can test at the bench, Helm says.

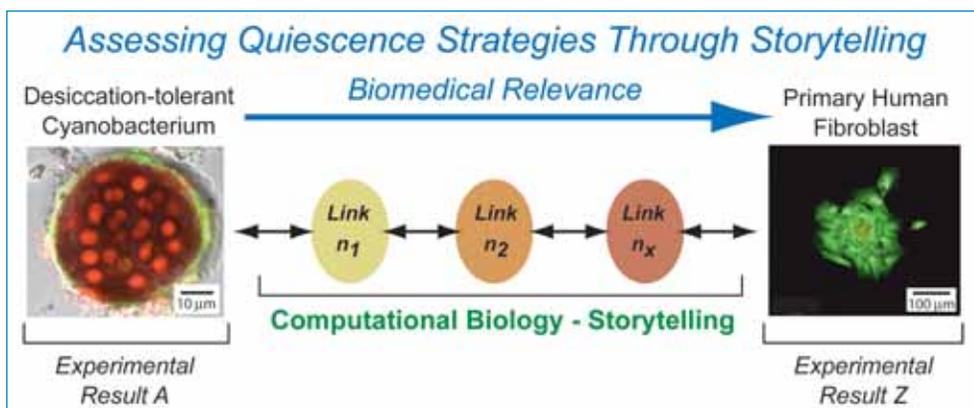
Bud Mishra, PhD, professor of computer science and cell biology at New York University, thinks *storytelling* can help biologists make new connections. “In some sense it closely resembles what biologists do, and it works in the same way that biologists think,” says Mishra.

—*Brian Lee*

A Fast Lane Through the Stomach

What goes into the stomach must come out, but perhaps not in the same order in which it entered, as gastroenterologists have long assumed. A two-dimensional computer model of human stomach digestion reveals a previously unknown narrow pathway that can funnel liquids from the top of the stomach to the intestines within 10 minutes.

“There are very few times you discover something that you weren’t expecting to find when you designed the experiment,” says **James G. Brasseur, PhD**, professor of mechanical engineering, bioengineering and mathematics at Penn State University and one of the researchers



As shown here, one might use storytelling to understand the pathways into and out of a quiescent state. Datasets evaluating desiccation-tolerant cyanobacteria (left, coccoid cells stained red and encased in an extracellular matrix, the exterior of which is stained green) could potentially be linked to studies involving the metabolic arrest and recovery of primary human fibroblasts (right, Live/Dead stain; image taken 72 hours after a two week metabolic arrest). Storytelling allows the biologist to link disparate datasets, allowing for the development of new hypotheses that can be tested at the bench and re-evaluated within the algorithm, ultimately resulting in new insights into the process of interest. Courtesy of Richard Helm.