

shortly after they got one of two different cues—but not if they got both cues at once. “To predict mealtimes accurately in this case, the microbes would have to solve a complex logic problem,” says Tagkopoulos, an electrical engineer associated with the Lewis Sigler Institute for Integrative Genomics. Sure enough, after a few thousand generations, a gastronomically savvy—and ecologically fit—strain of microbe emerged. The feeding response of such a fit bug (see figure) illustrates how interacting genes and proteins could evolve complex behavior.

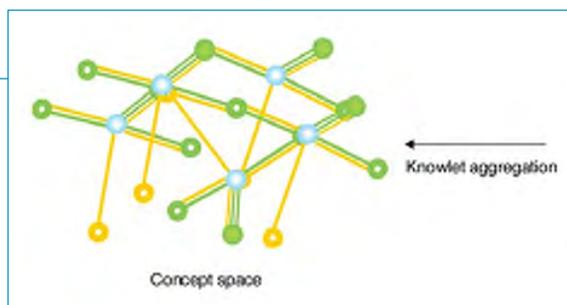
According to **David Reiss, PhD**, a computational biologist at the Institute for Systems Biology in Seattle, the researchers’ computational framework is notable for incorporating more biological mechanisms than prior models did. He cautions, however, that even this model oversimplifies the behavior of real microbes. Nevertheless, Reiss says, the study is interesting and novel for showing that anticipatory behavior is not restricted to higher systems with decision-making capability.

—By **Chandra Shekhar**

## Molecular Biology Wikis Launched

If you build it, will they come? That’s the question on everyone’s mind after the launch of two pioneering initiatives in community annotation: WikiProteins and Gene Wiki, announced, respectively, in the May 28 issue of *Genome Biology* and the July 8 issue of *PLoS Biology*. The efforts create a central repository of information on genes and proteins and call on the scientific community to keep it up-to-date and accurate.

“There’s no way we can handle the current growth of knowledge with central annotation only,” says **Barend Mons, PhD**, who leads the WikiProteins effort. “I’m a big fan of the authoritative databases like UniProt, but we have to make them grow faster. So what we need is a shell around them of community annotation.” Mons is associate professor of human genetics at the Leiden University Medical Centre and



*Each unique biomedical concept in WikiProteins is attached to a “knowlet” or concept cloud, illustrated here. A concept (depicted as a solid blue ball) is associated with other concepts through facts (established relationships, depicted as solid green balls), co-occurrences (co-occurrences in sentences in PubMed, depicted as green rings), or implicit associations (overlapping concepts in their Knowlets, depicted as yellow rings). Reprinted from Barend Mons, et al., *Calling on a million minds for community annotation in WikiProteins*, in *Genome Biology* 2008, 9:R89.*

of medical informatics at Erasmus University, both in the Netherlands.

“WikiProteins is more than just a Wiki; it has the whole knowledge space hovering over it,” Mons says. Using text mining, WikiProteins imported structured content (adhering to computer-readable, controlled vocabularies) on 1.2 million unique biomedical concepts from existing databases, such as PubMed, Swiss-Prot, and Gene Ontology. The system also created profiles for about 1.6 million authors in PubMed, who are expected to serve as the knowledge guardians. “If you have 1.6 million people in PubMed publishing today and you have 1.2 million concepts in the Wiki, then roughly everyone could take one concept and make sure the page on that concept is correct. That’s doable,” Mons says.

Gene Wiki operates within Wikipedia and, in contrast to WikiProteins, emphasizes unstructured content, such as free text and images, “more akin to a review article,” says **Andrew Su, PhD**, of the Genomics Institute of the Novartis Research Foundation, who leads the effort. Using data from Entrez Gene, the system added or amended about 9000 Wikipedia “stub” entries on human genes, which anyone can edit. “Being part of the larger Wikipedia community is certainly an advantage of this system. The people there are experts at welcoming newcomers, fighting vandalism, and formatting things correctly,” Su says.

Su and Mons have plans to collaborate. WikiProtein and Gene Wiki entries will be linked through a common “entry page” (likely hosted in WikiProteins),

making it easy to navigate between the systems. “This will allow users to take advantage of whichever system they feel comfortable with,” Su says.

Getting bench scientists to participate will be a challenge, Mons says, but he believes the incentives are high. The WikiProteins system

mines PubMed for new information daily, finds new explicit and implicit associations—such as predicting protein-protein interactions—and alerts scientists of all edits and updates to concepts in their purview. “I hope it becomes a daily part of their knowledge

“I’m a big fan of the authoritative databases like UniProt, but we have to make them grow faster. So what we need is a shell around them of community annotation,” says Barend Mons.

discovery process,” Mons says. Since its launch, WikiProteins has also received requests to enable users to enter data as unstructured, free text, which should lower the barrier to participation.

Another factor that may boost participation is the development of ways to trace authorship for each entry, so that authors can get credit for their work and readers can assess the reliability of content. A recent proof of this possibility was demonstrated in

“WikiGenes,” (not to be confused with the GeneWiki!) a project described in *Nature Genetics* in September 2008. WikiGenes was developed by **Robert Hoffmann, PhD**, at the Massachusetts Institute of Technology. It’s part of his Memoir project, which has, he says, “the ambitious goal to create a free collaborative knowledge base for all of science—where authorship matters.”

Though the creators of the various Wikis have not yet formally quantified participation, Su says that there’s been an uptick in Gene Wiki activity since the *PLoS Biology* paper came out. “It gives me hope that the system is right and that the framework is there, so if we are tapping into a desire in the community to share knowledge and harness community intelligence, then we have the structure to do it now.”

More information is available at: [www.wikiprofessional.org](http://www.wikiprofessional.org) (WikiProteins) and [http://en.wikipedia.org/wiki/Portal:Gene\\_Wiki](http://en.wikipedia.org/wiki/Portal:Gene_Wiki) (Gene Wiki).

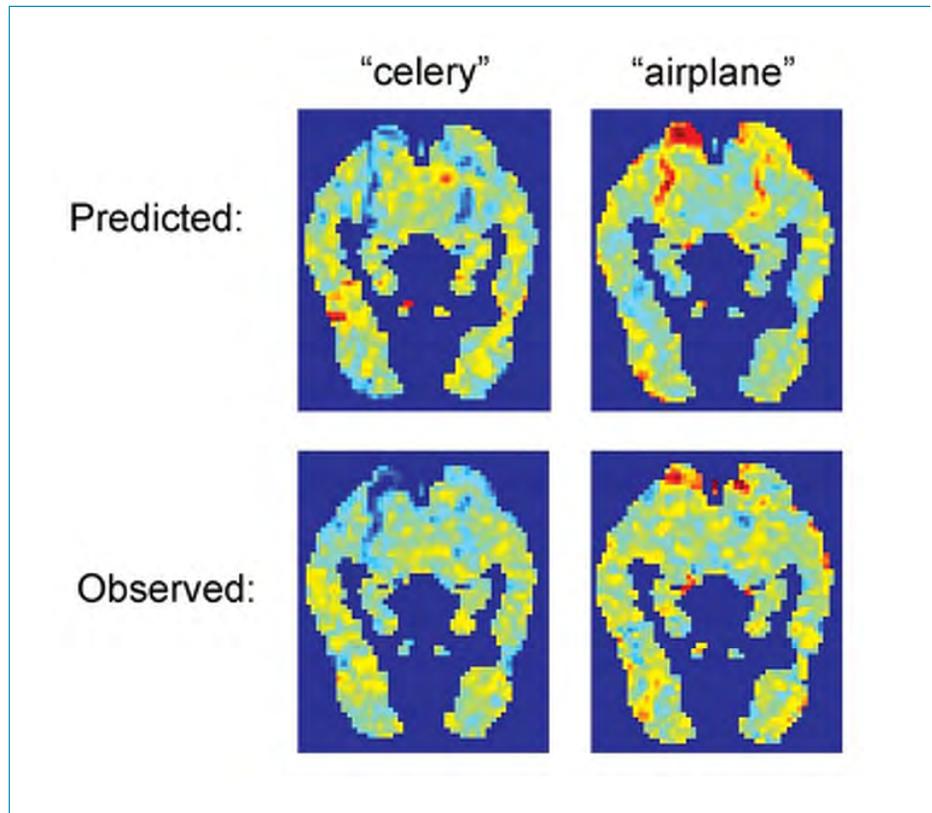
—By *Kristin Sainani, PhD*

## Predicting Brain Response To Nouns

Thinking of a noun—a peach, train, or bird, for example—activates specific parts of the brain. Now, scientists have trained a computer to predict such activation patterns. The achievement represents a step toward understanding language processing and could one day contribute to treatments for cognitive decline.

“If we had a better model of how the

brain represents language, we’d be better able to make sense of disorders like dementia,” says **Tom Mitchell, PhD**, a professor of computer science at Carnegie Mellon in Pittsburgh and lead author of the research published in the



*Brain activation patterns in response to nouns: The computer algorithm predicted the response to newly encountered words with 77% accuracy. Courtesy of Tom Mitchell. From Mitchell, TM, et al., Predicting Human Brain Activity Associated with the Meanings of Nouns, Science, 320 (5880): 1191 (2008 ) DOI: 10.1126/science.1152876. Reprinted with permission from AAAS.*

May 30 issue of *Science*.

Functional magnetic resonance imaging, or fMRI, registers changes in blood flow within peoples’ brains as they are asked to do a specific task—

computer to produce fMRI images like those generated by humans. The training process uses two sources of data: fMRI images collected from nine people viewing 60 nouns; and a database

The computer model was able to produce a pattern of brain activity in response to words it had never before encountered with greater than 70 percent accuracy.

such as thinking of a specific word. Since 2000, Mitchell and **Marcel Just, PhD**, professor of psychology at Carnegie Mellon and co-director of the Pittsburgh Brain Imaging Research Center have collaborated to train a

(derived from a trillion words of text from the Internet) describing pairings of nouns and the verbs that accompany them most frequently in written text. Noun-verb pairings are the basis of language, as anyone knows who has