

Popular plant database set to charge users

Loss of grant support leaves resource with no other option.

Erika Check Hayden

31 August 2013



Getty

A database that provides curated data on the model plant *Arabidopsis* is introducing subscription fees.

As funds dwindle, scientists running [The Arabidopsis Information Resource](#) (TAIR) have made the decision to ask users to pay to access the database.

The move is the latest example of researchers charging users for access to a previously free data resource. It comes four years after the US National Science Foundation (NSF) said that it would [wind down its \\$1.6 million in annual funding for TAIR](#), which provides genetic and molecular-biology data on the model plant thale cress (*Arabidopsis thaliana*) to 40,000–60,000 users per month.

Eva Huala, director of TAIR, which is based in the Department of Plant Biology at the Carnegie Institution for Science in Stanford, California, said today in a statement e-mailed to users that companies will be asked to begin paying for access in October, followed by academics next year, and that a new non-profit entity will collect the fees. She initially resisted charging subscription fees but has now exhausted alternative options. Other biologists lament NSF's decision not to provide continuing support for TAIR, saying that it puts the widely used resource in jeopardy.

"It is a huge mistake to completely cut funding for curation by TAIR," says Michael Cherry, a geneticist at Stanford School of Medicine who founded a precursor to TAIR in the early 1990s with NSF support, and is not associated with the current project. "I fear it will be a very difficult road ahead for TAIR. Research funds are tighter than ever and this will pressure labs and institutions into a difficult decision."

NSF has provided initial funding to a group called the [International Arabidopsis Informatics Consortium](#) to take on some of TAIR's functions, but the consortium may not be funded to provide what TAIR investigators say is the resource's most valuable function: human curators who update details on gene function with findings from the published literature. Huala says that the NSF is reluctant to pay for this curation or to foot the entire bill for the resource when two-thirds of TAIR's users are based outside the United States.

"There has been kind of a reluctance that has never been publicly stated, but is behind the scenes on NSF's part — that data curation of this sort is too expensive, not scalable and not something that they want to fund," Huala says.

No maintenance

Lily Whiteman, senior public-relations officer at the NSF, responds that the Advances in Biological Informatics program that initially funded TAIR typically funds “development and improvement” of information resources in biology, but not continuing operations or maintenance of these resources, such as curation.

Cherry says this is a short-sighted policy. “While we would all like there to be automatic methods to process the literature, there has been nothing yet created that is close to what biocurators provide.”

But Eric Lyons, a plant biologist at the University of Arizona who helped develop iPlant, an NSF-funded plant computational biology project, says that if this curation service is valuable enough to the community, scientists will support it.

“If the resource is as essential as I assume it to be, making it a fee-for-service makes sense,” Lyons says.

Lyons notes that other data resources have navigated the transition away from government funding as [agencies have ratcheted back support](#) for a first generation of bioinformatics, genetics and genomics data resources.

In 2011, for example, the [Kyoto Encyclopedia of Genes and Genomes](#) (KEGG) — which integrates genomic and molecular data into widely used pathway maps that have been cited in research papers some 8,000 times — began charging academic users to download its data after losing support from the Japan Science and Technology Agency, although it remains free through a web portal. It is managed by a non-profit organization, [NPO Bioinformatics Japan](#).

Before that, [Genevestigator](#), a resource started by academics at the Swiss Federal Institute of Technology (ETH) in Zurich that allows users to mine transcriptome data from certain plants and mammals, instituted a ‘freemium’ model operated by a for-profit company, [Nebion](#). It charges a range of prices for different users and functions of Genevestigator.

Plant geneticist Wilhelm Gruissem, one of the founders of Genevestigator, says that he could not have supported the resource with payments from the academics who study *Arabidopsis*. It is also difficult to collect payments from academic customers unless certain conditions apply, he says — for instance, there is no alternative to the paywalled resource, it provides strong customer support and it charges a very low price.

“TAIR would need to be massively improved in format and speed to make its customers ready to pay for it,” Gruissem says.

Huala responds that TAIR hopes to make improvements as soon as it can find enough money to do so. “We believe that the community understands the difference between a for-profit database that will charge a few well-funded people a high rate for access to a very polished site and a nonprofit project like ours that will continue to keep service to the community and maximum data availability as its highest goal,” Huala says.

Ultimately, scientists and funding agencies need to change their attitudes and policies if databases are to continue to operate, say Gruissem and bioinformatician Minoru Kanehisa at Kyoto University, who created KEGG. Kanehisa advocates a cultural change in biology that will lead to greater acceptance of licence or subscription fees to databases; these are prohibited by some funders and viewed with scepticism by biologists.

“As the biological science matures and biological databases mature, I hope the biological community will also change,” he says.

Nature | doi:10.1038/nature.2013.13642