

The Arabidopsis Information Resource (TAIR)

<https://www.arabidopsis.org/>

Leonore Reiser

lreiser@arabidopsis.org

Phoenix Bioinformatics



with input from Tanya Berardini, tberardini@arabidopsis.org; Sabarinath Subramaniam, shabari@arabidopsis.org, Eva Huala, huala@arabidopsis.org

Recent activities and newly developed tools and/or resources

In 2020 we released or updated tools for comparative genomics, genome visualization and functional annotation. We released a new Synteny Viewer. This tool uses pre-computed syntenic comparisons between *A.thaliana* and genomes from 36 other plant species. Users can query for syntelogs using either AGI locus IDs (e.g., AT1G68640), AGI gene model IDs (e.g., AT1G68640.1) or chromosomal location (e.g.1:25769576..25772650). Syntenic comparisons were generated using the SynMap tool at CoGE and syntenic regions are displayed using the CoGE GEvo tool. After taking over the updating and maintenance of JBrowse from the Araport project, we continue to add new data tracks on a regular basis. New/updated data includes TIF-seq and plaNET-seq Data from the Marquardt lab, and Ws/Ws-2 polymorphism data from GBrowse. Please see the JBrowse release notes for more information on the specific updates and changes.

Finally, we released a new tool to enable community curation of gene function called the Generic Online Annotation Tool (GOAT, <https://goat.phoenixbioinformatics.org/>). GOAT replaces the Arabidopsis- and TAIR-centric TOAST tool for collecting Gene Ontology and Plant Ontology annotations from the research community and offers some significant improvements including the ability to save a submission in progress, sign in using ORCID, and the ability to annotate any gene from any organism. Future updates will further improve the functionality of the tool. Other activities in the past year include outreach in the form of online webinars, scheduled public data releases and data curation. We took advantage of 'captive' home audiences to launch a webinar series using Zoom and Voov (for users in China) platforms. The recorded webinars are available at our YouTube Channel and on Bilibili (for users in China).

In 2020 as part of our long-term goal to produce a 'gold standard' annotated reference genome, we continued to produce and incorporate new Arabidopsis gene function annotations made by curators and members of our community. As of March 2021, the percentages of protein-coding Arabidopsis genes with GO annotations are as follows: For GO molecular function, 29% have at least one experimentally based annotation, another 30% have no experimental annotation but at least one based on other evidence. For GO biological process 31% have experimentally based annotations and another 26% have annotations based on other evidence. For GO cellular component 35% have experimentally supported annotations, another 55% have annotations based on other types of evidence.

Planned future activities

In 2021 we will continue our functional annotation efforts, adding new data to our database and tools, enhancing the usability of our tools and website, and outreach to broaden access. We continue to prioritize curation of previously 'unknown' genes and capturing 100% of novel gene functions as they are published. As time permits, we are delving into our backlog to identify and capture information that were unable to in previous years.

We plan improvements to the GOAT community curation tool to improve usability. To comply with the directives from the Gene Ontology Consortium, GO annotations from GOAT and TAIR will now be exported in GAF2.2 format. For JBrowse, we plan to continue adding new data including transcriptomic, proteomic and community submitted tracks. Please contact us if you have data that you think would be appropriate to share via JBrowse. We also plan to integrate the new JBrowse2 into TAIR. Planned updates for Synteny Viewer include the addition of viewing options for JBrowse2 and the NCGR Genome Context Viewer. We hope to complete a long-planned overhaul of our Locus page user interface to improve data visibility.

Please provide a paragraph describing the general impact of the COVID19 pandemic on your activities

In March 2020, in response to the global shelter in place order, TAIR lifted monthly limits until June 30, 2020 to ensure that researchers working from home were able to access TAIR. Due to the ongoing impacts of COVID19 on the academic community TAIR has experienced downstream effects of reduced university budgets on subscriptions. To reduce costs, Phoenix Bioinformatics, where TAIR is maintained, has made remote work permanent for all employees.