



2008 TAIR Community Survey Results

Robert J. Muller

May 14, 2008

CARNEGIE
INSTITUTION FOR
SCIENCE

Department of Plant Biology/TAIR
260 Panama Street
Stanford, CA 94305

Introduction

The 2008 TAIR Community survey polled the TAIR user community to find out how well the TAIR web site is serving the needs of the community of users. It also surveyed the community on preferences for future tools and data.

The survey itself is an online survey that TAIR built on the surveymonkey.com web site and presented to the community through the TAIR web site. The survey had two separate sets, one a self-selected group of community users accessing the survey as a result of web site visits, the other a random sample of 300 2007 ICAR submitters contacted through email. The former set consists of 292 respondents, while the latter consists of 95 respondents (an excellent sample response rate of 32%) for a total response set of 387 respondents. TAIR tested the random sample approach to see if such a sample would yield sufficient response for statistical analysis. The approach was very successful. The statistical analysis in this report uses the full response set where random sampling is not required; where it is required by the statistical techniques, the report uses the random sample only and notes that fact. Where the samples differ, the analysis characterizes the differences and explains them.

The report first presents tables and histograms of the data along with summary statistics that describe the basic characteristics of the response. Each section interprets the results in a summary sentence at the beginning of the section.

The report then addresses several questions of interest:

1. Is there any significant difference on the major questions between the USA-based respondents and community members based elsewhere?
2. Do crop researchers have different needs than the general user population?
3. Do computational biologists and bioinformaticists have different needs than the general user population?
4. Do frequent users of the TAIR web site have different opinions on any question?

Executive Summary of Results

TAIR is essential (83%) or very useful (13%) to most survey respondents.

Most respondents are very satisfied or satisfied with all aspects of TAIR, and most are very satisfied with availability of the web site. Most respondents have never contacted TAIR for support; those that do have mostly found TAIR's response very helpful.

Most survey respondents work primarily on *Arabidopsis thaliana*. Survey respondents also worked on a diverse variety of different organisms (138 total) in addition to working on *Arabidopsis thaliana*, the top three being yeast, tobacco, and *E. coli*.

Most respondents felt that correct gene structures were extremely important, followed closely by experimentally verified gene function data, seed and DNA stock information, gene expression, mutant phenotypes, and polymorphism-related information. Other types of information were still important, however; most respondents did not mark any type of information as unimportant or somewhat important.

Respondents would most like to see promoters/cis elements, regulation of transcription and orthologous genes from other plants and model organisms added to TAIR. Survey respondents expressed a decided preference for protein-related tools (interaction, domain and modification site, and alignment viewers) as future additions to TAIR.

Survey respondents expressed a clear preference for integration of TAIR data submission into the publication process, followed closely by improved data submission forms.

Most survey respondents felt well informed with no additional community outreach needed. A significant portion would like to see RSS news feeds or email notification of changes to genes of interest.

There were two significant differences between USA and other respondents, on performance and data correctness satisfaction. USA respondents were moderately more satisfied on both counts.

There were several significant differences between crop researchers and other respondents, on data completeness and on promotor/cis elements and regulation of transcription data. Crop researchers were more dissatisfied than other respondents with data completeness, though not by much. Also, crop scientists were moderately less interested in promotor/regulation of transcription data.

Bioinformaticists had several significant but moderate differences from other respondents:

- Felt less satisfied and more dissatisfied with data completeness
- Felt splice variants, polymorphisms, stock information, and mutant phenotypes were less important than other respondents did
- Felt experimentally verified gene function data was more important than other respondents did
- Chose complete ecotype sequences less often as a future type of data
- Chose customizable bulk data retrieval tools slightly more often than others did
- Chose protein alignment and interaction viewers slightly less often than others did

Frequent TAIR users (daily or weekly) had minor differences from infrequent users. Frequent users tended toward being satisfied with availability instead of very or partially satisfied, a mixed result. Infrequent users tended to be less satisfied with the organization of the web site. Frequent

users tended to be very satisfied with data correctness while infrequent users were just satisfied. There was a slight difference between frequent and infrequent users in choosing orthologous genes as a future data type, but it wasn't strongly significant.

The survey respondents have labs that mostly range from 1-10 people, with labs in the 6-10 people range being the median. Most of the survey respondents spend their days carrying out or supervising lab or field experiments. Most survey respondents were professors, graduate students, or postdocs. Most respondents belonged to academic institutions. Most survey respondents were from the USA or Europe with a significant portion in Asia.

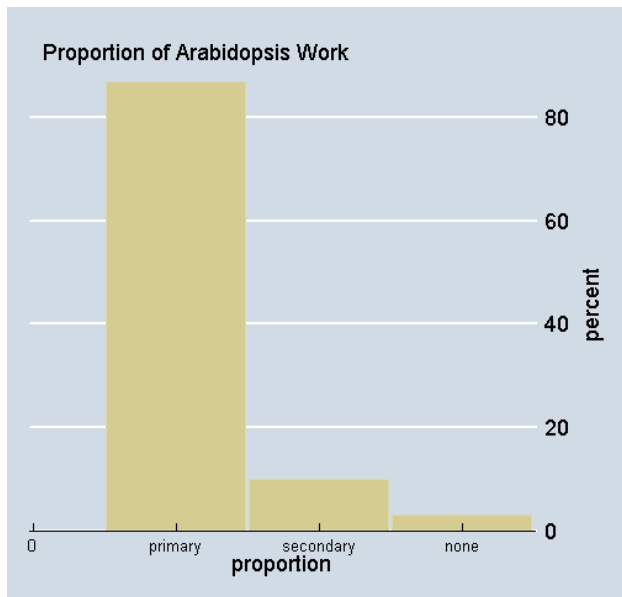
Arabidopsis Focus

Most survey respondents work primarily on *Arabidopsis thaliana*.

Question: What proportion of your current work or study is focused on Arabidopsis?

Responses:

Arabidopsis is the primary organism for my current work	336	87%
Another organism is my main focus but I also work on Arabidopsis	39	10%
I don't use Arabidopsis in my current work	12	3%



The median for the proportion of Arabidopsis work is 1 (primary). There is no significant difference between the random sample and the self-selected sample median.

Other Organisms Worked On

Survey respondents worked on a diverse variety of different organisms (138 total) in addition to working on *Arabidopsis thaliana*, the top three being yeast, tobacco, and *E. coli*.

Question: Which other organisms do you work on? Please list up to five species in order of importance to your work (you may include plants, animals, fungi, bacteria, or any other species of relevance to your work)

Here is the top-ranked group of organisms:

<i>Saccharomyces cerevisiae</i>	49
<i>Oryza sativa</i>	47
<i>Escherichia coli</i>	44
<i>Nicotiana tabacum</i>	42
<i>Arabidopsis thaliana</i>	24
<i>Lycopersicon esculentum</i>	24

Please see [Appendix B: Organisms](#) for a complete list of other organisms. There is no significant difference between the random sample and the self-selected sample in terms of working on other organisms: the diversity is large in both cases.

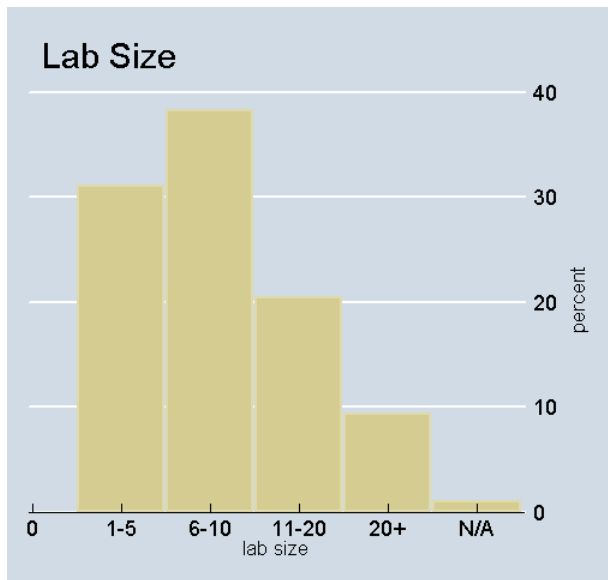
Lab Size

The survey respondents have labs that mostly range from 1-10 people, with labs in the 6-10 people range being the median.

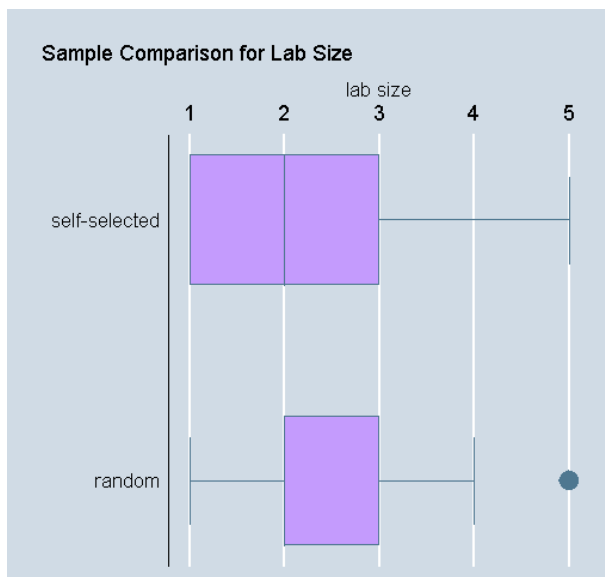
Question: Number of people in your lab group or research unit?

Responses:

1-5	120	31%
6-10	148	38%
11-20	79	21%
20+	36	9%
Does not apply	4	1%



The median lab size is 6-10. The samples differ slightly in distribution of responses; the random sample comes from the population of ICAR submitters, a more uniform group of researchers.



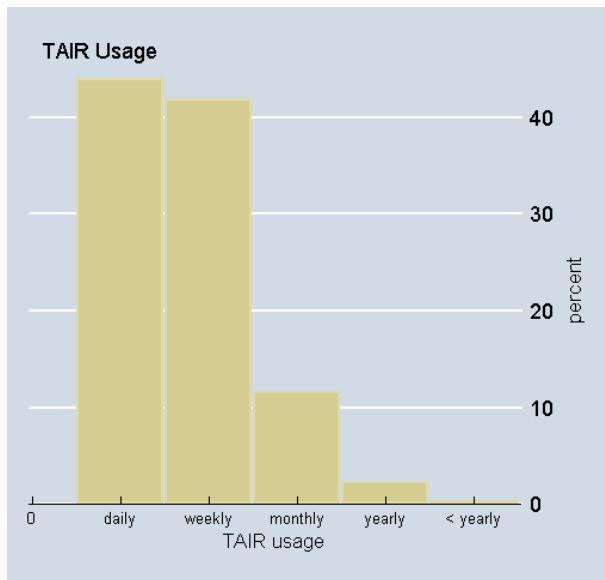
TAIR Usage

The survey respondents have labs that mostly range from 1-10 people, with labs in the 6-10 people range being the median.

Question: How often do you use TAIR on average?

Responses:

Most workdays	170	44%
Once a week or more	162	42%
Once a month or more	45	12%
Once a year or more	9	2%
Less than once a year	1	0%



The median usage level is weekly. There is no significant difference between the self-selected and random samples.

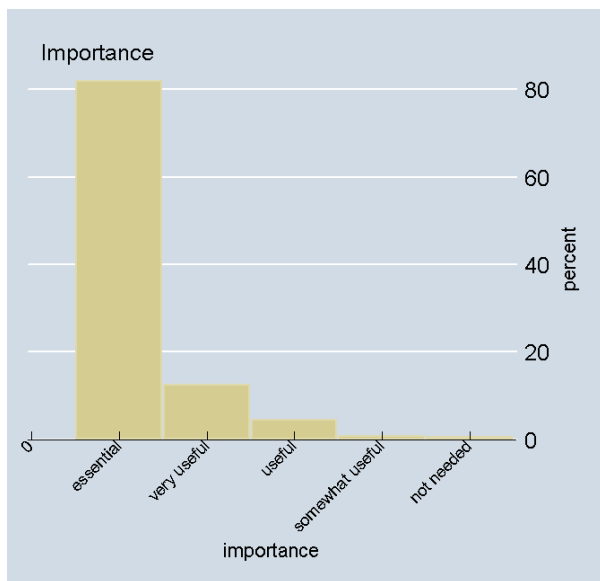
TAIR Importance

TAIR is essential (83%) or very useful (13%) to most survey respondents.

Question: How necessary is TAIR and the data it contains for your research or teaching?

Responses: 1-5, where 1 is "essential" and 5 is "not needed at all for my work"

essential	316	82%
very useful	48	12%
useful	17	4%
somewhat useful	3	1%
not needed at all	2	1%



The median value is "essential." There is no significant difference between the self-selected and random samples, though the latter has even fewer choices other than "essential."

Current Position

Most survey respondents were professors, graduate students, or postdocs.

Question: What is your current position?

Responses:

Professor (assistant, associate, or full)	79	21%
Lecturer (college or university)	1	-
Research group leader	33	9%
Postdoc	127	34%
Other researcher	21	5%
Graduate student (Master's or PhD degree candidate)	118	31%
Undergraduate student (working toward Bachelor's degree)	3	-
Teacher (elementary, middle, or high school)	0	-
Other (please specify)	5	-

The top three responses accounted for 86% of the total:

1. Postdoc
2. Graduate student
3. Professor

Other specified positions included (one each):

- software developer
- scientist
- stock center
- graduate student (Graduate Diploma in Science)
- database curator

There is no difference between the self-selected and random samples for position.

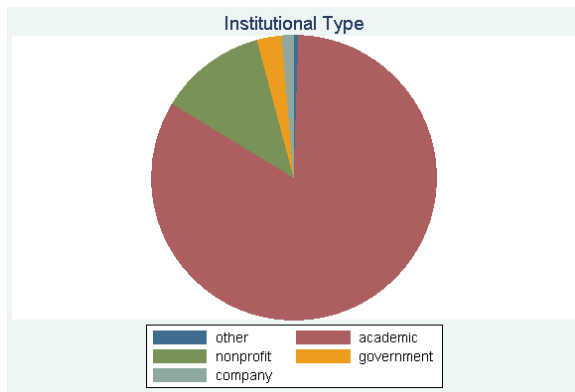
Institutional Type

Most respondents belonged to academic institutions.

Question: What type of institution are you at?

Responses:

Academic institution	322	83%
Nonprofit research institution	47	12%
Government agency	11	3%
Company	5	1%
Other (please specify)	2	1%



There was no significant difference between the self-selected and random samples for institutional type; the random sample had a slightly greater concentration in academic institutions.

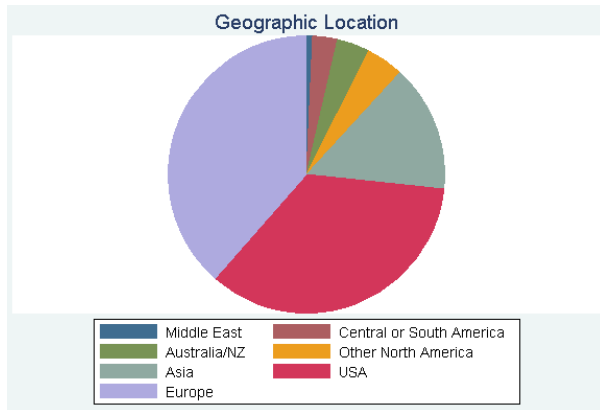
Geographic Location

Most survey respondents were from the USA or Europe with a significant portion in Asia.

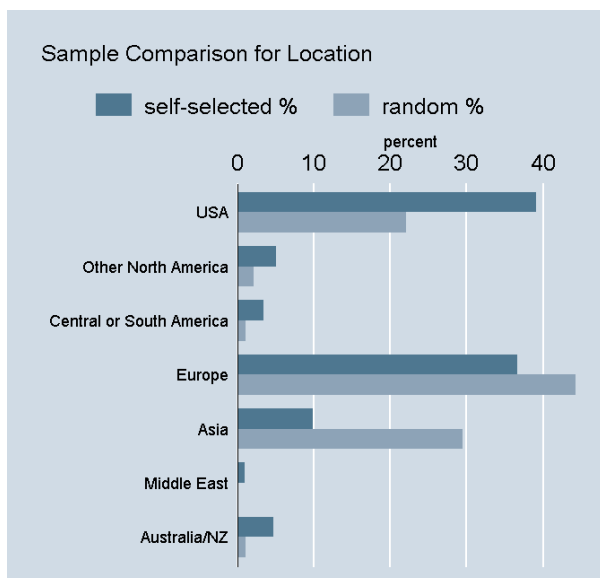
Question: What is your geographic location:

Responses:

USA	135	35%
Other North America	17	4%
Central or South America	11	3%
Europe	149	38%
Asia	57	15%
Africa	-	-
Middle East	3	1%
Australia/New Zealand	15	4%
Other (please specify)	-	-



Most of the respondents were from the USA or Europe. There is a significant difference in geographic distribution between the random sample and the self-selected sample; there are more European and Asian respondents in the random sample than in the self-selected one, probably because the ICAR conference providing the population was in Beijing in 2007.



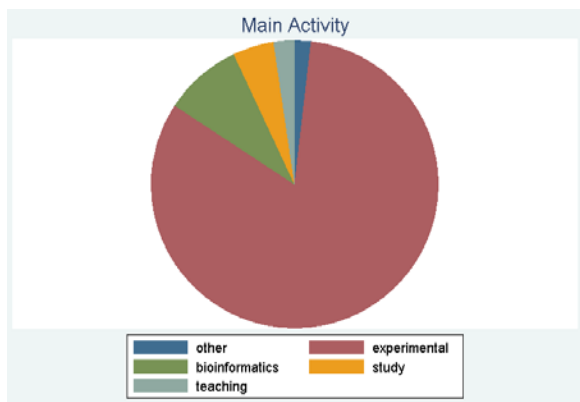
Main Activity

Most of the survey respondents spend their days carrying out or supervising lab or field experiments.

Question: Which best describes your main daily activity?

Responses:

Carry out or supervise laboratory or field experiments	319	82%
Carry out or supervise bioinformatics or computational biology work	34	9%
Coursework and study	18	5%
Teaching	9	2%
Other (please specify)	7	2%



The Other category had 7 responses, which were mostly combinations of the categories:

- Research and administration
- Carry out or supervise both laboratory and bioinformatics work
- Field experiments, coursework and study
- Teaching and research are equal
- Bioinformatics AND laboratory
- Supervise bioinformatics and molecular laboratory work
- PhD research work

Customer Satisfaction

Most respondents are very satisfied or satisfied with all aspects of TAIR, and most are very satisfied with availability of the web site.

Question: How satisfied are you with the following aspects of TAIR?

Responses: a scale of 1-5 where 1 is "very satisfied" and 5 is "very dissatisfied"

Aspects	very satisfied	satisfied	partially satisfied	dissatisfied	very dissatisfied
Performance (speed)?	138	154	60	25	9
Availability of web site (is it always accessible)?	207	128	37	10	3
Organization of web pages (easy to find information)?	87	172	99	24	3
Completeness of data?	100	195	75	14	1
Correctness of data?	102	215	49	15	1



The median for availability satisfaction is "very satisfied." the medians for the other types of satisfaction are all "satisfied." There is no difference between the self-selected and random samples.

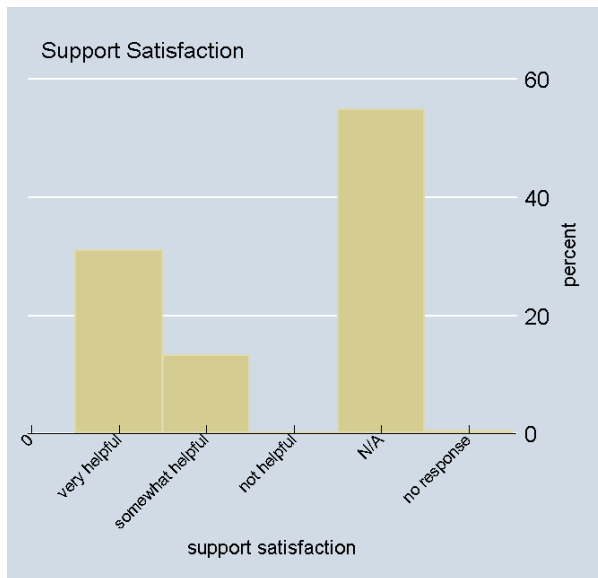
Support Satisfaction

Most respondents have never contacted TAIR for support; those that do have mostly found TAIR's response very helpful.

Question: If you have every contacted TAIR curators, how helpful was the response?

Responses:

Very helpful	114	29%
Somewhat helpful	49	13%
Not helpful	1	-
Never contacted TAIR	202	52%
Contacted but got no response	2	1%



Information Types

Respondents felt that correct gene structures were extremely important, followed closely by experimentally verified gene function data, seed and DNA stock information, gene expression, mutant phenotypes, and polymorphism-related information. Other types of information were still important, however; most respondents did not mark any type of information as unimportant or somewhat important.

Question: How important are the following types of information to you?

Response: Scale of 1-5 with 1 being "extremely important" and 5 being "unimportant"

Information Types	extr important	very important	important	somewhat important	unimportant
Correct gene structures	304	66	8	3	1
Alternate isoforms (splice variants)	142	132	72	24	6
Polymorphisms/SNPs/T- DNA and Tn insertions	234	84	32	19	8
Experimentally verified gene function data	249	99	28	3	1
Computationally predicated gene function data	127	156	75	15	6
Gene families	144	165	59	9	4
Seed and DNA stock information	276	65	23	12	6
Protein-protein interactions	147	125	70	26	12
Gene expression	229	112	33	7	1
Metabolic pathways	116	123	79	46	16
Mutant phenotypes	231	103	31	11	5

5 respondents did not evaluate information types (total respondents was 382). There was no obvious pattern of clustering of information type importance from an exploratory classification analysis.

There was little difference between the self-selected and random samples. Seed and DNA stock information was extremely important in the random sample but only very important in the self-selected sample.

The following table ranks the types of information by the sum of the extremely important and very important counts and shows the median importance:

Information Types	extr important	very important	both	median
Correct gene structures	304	66	370	extremely
Experimentally verified gene function data	249	99	348	extremely
Seed and DNA stock information	276	65	341	extremely
Gene expression	229	112	341	extremely
Mutant phenotypes	231	103	334	extremely
Polymorphisms/SNPs/T-DNA and Tn insertions	234	84	318	extremely
Computationally predicated gene function data	127	156	283	very
Alternate isoforms (splice variants)	142	132	274	very
Protein-protein interactions	147	125	272	very
Gene families	144	165	269	very
Metabolic pathways	116	123	239	very

There is a clear break between the first six information types and the last five, but those last five are still very important to most respondents.

Future Data Types

Respondents would most like to see promoters/cis elements, regulation of transcription and orthologous genes from other plants and model organisms added to TAIR.

Question: Which of the following types of data would you most like to have available in TAIR? Please select up to three answers.

Responses:

Promoters/cis elements, regulation of transcription	219
Orthologous genes from other plants and model organisms	204
Complete ecotype sequences	155
Regulatory interactions	144
Genetic interactions (e.g., epistatic, additive)	140
Protein modification data	135
Small RNAs	134
Other Brassicaceae genes/genomes	96
Arabidopsis lyrata and Capsella rubella genes/genomes	93
DNA methylation patterns	88
QTL data	57
Other (please specify)	14

5 respondents did not answer this question (total respondents was 383). The first group of future data types ranked by vote was

1. Promoters/cis elements, regulation of transcription
2. Orthologous genes from other plants and model organisms

The following five data types were in a second group in this order:

3. Complete ecotype sequences
4. Regulatory interactions
5. Genetic interactions
6. Protein modification data
7. Small RNAs

A third group included:

8. Other Brassicaceae genes/genomes
9. Arabidopsis lyrata and Capsella rubella genes/genomes
10. DNA methylation patterns

Last came QTL data, chosen by only 15% of the respondents, in a clear last group.

There were 14 "other" data types entered:

- data about nicotiana genes related to Arabidopsis
- Comprehensive microarray data/Metabolite profiling dataset downloads divided by tissue type
- Protein sub-cellular localization
- close/functional links to other model organisms' sites (equivalent to TAIR)
- Histone modification patterns
- an interactome like Nick Provar has on the BAR (Plant Physiology 2007 145:317-329), also, a way to submit a long list of AGIs and get the latest annotation, suitable for import into multiple Excel columns
- The best regularly updated bioinformatics tools
- histones modifications
- Paralogous genes or gene groups accompanied by phylogenetic analysis
- a more complete and better curated GO annotation for the Arabidopsis genome
- Complete ecotype habitat, latitude, and longitude information.
- Modified histone pattern
- Link to data published in the primary literature, compiled in a machine readable format
- protein interactions

There was little or no difference between the self-selected and random samples on future data types. The brassica data had broader support from the self-selected respondents, but fewer than half of those respondents chose that option. The methylation data had broader support from the random sample, but fewer than half of those respondents chose that option.

Future Tools

Survey respondents expressed a decided preference for protein-related tools (interaction, domain and modification site, and alignment viewers) as future additions to TAIR.

Question: Which of the following tools would you most like to have available in TAIR? Please select up to three answers.

Responses:

Protein interaction viewer	211
Protein domain and modification site viewer	206
Protein alignment viewer	199
Customizable bulk data retrieval and download tool (e.g., Biomart)	161
SNP viewer	102
Synteny viewer	75
QTL viewer	53
Other (please specify)	10

17 respondents did not answer this question (total respondents = 371).

The first three future tools ranked by vote form the first group, protein-related viewers:

1. Protein interaction viewer
2. Protein domain and modification site viewer
3. Protein alignment viewer

The second group consists of one data retrieval tool:

4. Customizable bulk data retrieval and download tool

The third group, with a lot fewer votes, consists of three viewers:

5. SNP viewer
6. Synteny viewer
7. QTL viewer

There were 10 "other" tools entered:

- complete analysis tools
- view genomic data by different sections (promoter, repeat, coding etc.)
- small RNA data base for find potential target gene
- Please no BioMart! The UI is dreadful. Please emulate the UCSC Table Browser, which is much better!
- sequence viewer
- protein 3D structure software
- ABILITY TO SEARCH FOR MORE THAN ONE ATG NUMBER AT A TIME
- Primers design with specific option as RT-PCR (across 2 exons) ...
- Virtual PCR tool for the Arabidopsis genome
- Microarray analysis tools

There was little or no difference between the self-selected and random samples on future tools. The SNP viewer had broader support from the random sample, but fewer than half of those respondents selected that option.

Data Submission

Survey respondents expressed a clear preference for integration of TAIR data submission into the publication process, followed by improved data submission forms.

Question: Which of the following methods would most strongly encourage you to submit data to TAIR? Please select up to three answers.

Responses:

Integration of TAIR data submission into the publication process for research articles	195
Improved data submission forms	127
Ability to directly edit TAIR pages if logged in	99
Ability to volunteer as primary editor for one or more genes	72
More visible linking of my name to the data I submitted	64
Free publicity for my lab on the TAIR home page if I submit data	63
Annotation booth at conferences where I can add data with a curator's help	46
Other (please specify)	11

61 respondents did not answer this question (total respondents = 327).

By far the most votes were for the integration of TAIR data submission into the publication process.

1. Integration of data submission into the publication process

This was followed at some distance by:

2. Improved data submission forms

The rest of the choices were much less preferred:

3. Ability to directly edit TAIR pages if logged in
4. Ability to volunteer as primary editor
5. More visible linking of my name to the data I submitted
6. Free publicity for my lab on the TAIR home page if I submit data
7. Annotation booth at conferences

There were 11 "other" methods entered:

- I would rather have a curator handle data entry to maintain a uniform standard of quality
- direct request from TAIR staff by email
- it is not a matter for me
- won't submit any data because of lack of generating data in my work
- less intimidating to edit / add
- No opinion
- Ability to withdraw or edit comments after submission
- I work on community need basis. Data submission is a natural thing to do. No further encouragement needed.
- don't generate data of interest at the moment
- I'm working on bacteria - so I have no interesting data for TAIR
- no opinion

There was little or no difference between the self-selected and random samples on data submission techniques. There was more support in the self-selected sample for directly editing submissions, but fewer than half of those respondents selected that option. There was more support in the random sample for free lab publicity, but fewer than half of those respondents selected that option.

Community Outreach

Most survey respondents felt well informed with no additional community outreach needed. A significant portion would like to see RSS news feeds or email notification of changes to genes of interest.

Question: How can TAIR keep you better informed on the availability of new tools an data? Please select all that apply.

Responses:

I feel well informed already, nothing more is needed	171
RSS news feed or email notification of changes to my genes of interest	138
I would like more TAIR workshops to be offered at plant biology conferences	85
I would like to receive more frequent TAIR newsletters via email	81
Other (please specify)	4

17 respondents did not answer this question (total respondents = 371).

Most respondents felt that nothing more was needed:

1. I feel well informed already, nothing more is needed

This was followed at some distance by:

2. RSS news feed or email notification of changes to my genes of interest

The rest of the choices were about equally preferred:

3. I would like more TAIR workshops to be offered at plant biology conferences
4. I would like to receive more frequent TAIR newsletters via email

There were 4 "other" methods entered:

- Links or information about such new tools on TAIR home (may already exist)
- A walk-through on-line tutorial for different sections--so if I am trying to do something new I can teach myself how and what is available at TAIR (I know there are some, but more would be nice).
- TAIR could facilitate coordination with other researchers working on the same or related mutants and genes.
- Online guide and walkthrough, with do-it-yourself exercises in how to perform some of the most common tasks (i.e. so I can send students to that page and train them)

There was little or no difference between the self-selected and random samples on community outreach. There was more support in the random sample for newsletters and workshops, but fewer than half of those respondents selected those options.

Suggestions

Question: General suggestions and feedback for TAIR

93 respondents gave feedback:

- Keep up the great work
- Excellent resources! Very well maintained, easy to use, easy for students to learn how to access data.
- TAIR website is already easy to find useful information about my genes of interest. However, I'm interested in epigenetic modifications and protein-protein/DNA interactions, so I hope that TAIR gives some information about those molecular events.
- Overall TAIR does a great job. The more information that can be integrated to TAIR the better as this saves having to know the websites to source various datasets from. I particularly like the idea of having better ways to capture information as it is published, perhaps this could go as far as integrating data such as figures from papers into the sections for a particular gene. Making it as easy as possible to put data into TAIR is another important factor as good intentions to do this are often defeated by not having the time to actually do it.
- Great Job so far. Keep it up. Thanks
- I think TAIR is the best website of Arabidopsis research. I wish it can be the best 4ever! If it can be much easier to use, it will be great!
- I think TAIR should make an extra effort to better annotate the genes, because I have the feeling the annotation lacks behind what is known in literature. As I study protein-protein interactions, I find this annotation extremely important!!!
- I think TAIR should make an extra effort to better functional annotate the genes/proteins, because I have the feeling the annotation lacks behind what is known in literature. As I study protein-protein interactions, I find this annotation extremely important!!! Maybe you should encourage researchers more to send their annotations plus the paper!
- I would like to have more tutorials or better tools explanations. If you are new to TAIR sometimes it is difficult to get advance of all the possibilities that this database has.
- Even though working with other organism than Arabidopsis the site has been essential. Thank you!
- In general TAIR is an invaluable tool. Currently I face troubles when searching for a specific gene and the search program does not always give the right specific results
- The site is very good : a well of knowledge. It's easy to use. Good job guys !
- Search engines in Tair should be improved.
- You are a great resource (but there's always room for improvement/expansion)!
- tair is a very nice resource. minor comments: The dna stock form could be updated to make it more user friendly (such as pasting single and multiple agi numbers aka the very nice sequence download page). Some of the gene function info taken from publications on tair refer to yeast genes with the same name. less effort should be spent on things

like microarrays (can they improve on genevestigator and mapman?) and aracyc. nasc have wasted loads of time on second-rate databases and tools -don't do the same.

- I always get confused when I search for genes, do I have to click gene model or locus to get the data. Sometimes TAIR did not seem to work well in Firefox on my Computer but that might be just my misconfiguration. Also I just want to mention that TAIR does a great job in annotating genes and offering superb gene models! Compared to other so-called sequenced model plants TAIR is by far superior.
- Needs better linking of different types of data (probably at the locus pages), should not have to spend half an hour trying to find something.
- Overall, the information available on TAIR is extensive and well organized. I especially like the addition of annotation details, which allow one to see whether genetic, computational, or other evidence supports a particular annotation. There is, however, one minor inconvenience on the SequenceViewer page, with regard to the pop-up windows. For some reason, when I try to move my mouse to click on the "nucleotide sequence view" link, the pop-up window often moves or disappears. Of course, I can access the sequence through the "10 kb" view, but it would be great if I could consistently and quickly focus on the gene of interest via "nucleotide sequence view." Thanks for putting together an extremely useful website!
- The information at TAIR has always been pretty good and continues to get better. The curators have also always been very helpful to me. My primary complaints regarding TAIR are the incredibly slow page loading speeds (VERY frustrating) and the structure of the information. I use TAIR quite frequently, but I still struggle to remember where to go to find relevant information - it's just not intuitive and there are often way too many button clicks to get to the relevant information. I have felt the site is due for a major interface re-write for years. While many new and very useful tools have been added, the basic site structure hasn't changed much. I have spoken with many others with similar complaints. Regardless, this site is indispensable for my lab's research and I greatly appreciate the efforts of the TAIR staff.
- I hate, and I really mean hate, the fact that so many windows are launched as I migrate about the site. It seems to take too many clicks to just get to a genome sequence view of my gene of interest. Please, abandon multiple windows.
- TAIR is great. Keep up the good work!
- you guys are doing a great job. thank you
- TAIR does a superb job for the US and international community. I hope they can secure funding for interaction data. Well-curated interaction data is essential for systems research.
- I feel ok about TAIR but i think that the annotation process should be improved and curated process carry out by more than one lab if possible.
- If I wish to obtain gene expression data from specific tissue types or developmental stages I need to search all NCBI GEO, Array Express, and NASC array databases to find experiments. Then I need to manually look through all the experiments to find the tissue type and experimental conditions performed. Wouldn't it be nice to have this data all

deposited in one place and curated so it could be easily searchable. (For example: One could query the database and select all experiments performed on the ATH1 chip performed in tissue types "leaf", "aerial tissue", "green tissue" etc.) If this was done correctly in the public domain nobody would ever have to do it again.

- I am very happy with the TAIR website. Please continue!
- You may want to consider introducing *RNA-level* information, which is now becoming much more available, for example microarray data on -RNA stability (Whelan Plant Cell 2007) -exosome targeting (Belostotsky Cell 2008) -polysome loading state (Bailey Serres lab and others) etc. I would also be in favor of a renewed effort to include basic RNA transcript-level information. We are also in need of a 'Maniatis' of Arabidopsis bioinformatics. Perhaps you have ideas of how to implement this. You did not ask how satisfied we are with TAIR :-). I'd like to say that TAIR is providing a tremendous service and is very user-friendly. Compared to the disarray that all? other plant genomes are in you are way ahead of the curve. Keep up the great work.
- TAIR plays an essential role in moving the rate of plant discovery forward, and is an outstanding resource.
- I hope to get definite data about SNP position and RS series number of SNP.
- Since numerous interference RNA have been sequenced in Arabidopsis, it should be a database for find potential target gene.
- I use TAIR extensively and would be lost without it. I find the list of publications about each gene useful. I also find the sequence viewer function to be extremely valuable. Thanks!
- TAIR has been a very useful research resource. I hope that the communication between TAIR curators and other researchers could be further enhanced.
- TAIR has done a great job. It is essential for my research.
- no comment
- TAIR is very important for us, thank you very much.
- TAIR is my best favorite website! Hope it better and better!
- It would be great if you can increase the speed of sending materials (e.g. seeds from stocks) upon request. Frequently (at least for us), it takes several weeks, sometimes two months, to get them in the lab.
- please refine protocol part. Searching with key words does not work efficiently. Rather, please make the whole list and web links of all protocols (downloading pdfs without knowing the contents in advance is inconvenient and useless).
- I am surprised when I try to order a seed stock from the link you make that it does not keep in memory the line I wanted to order. I don't know if this is your responsibility or the one of your partners like NASC. Thanks for your help.
- Overall a WONDERFUL site. I use this all the time, I think it, of course, could be improved. Linking out to other websites/ making it easier for other websites to link to TAIR. For example, the GO packages for many web-based analysis programs are all way

out of date for Arabidopsis and when I try to fix the link or ask them to update them- the links to TAIR are not working / in the wrong format compared to all other GO annotations. The GO package in TAIR works fine but often I would like to streamline things and not have to make a special exception for Arabidopsis. I think in general linking out more / integrating other programs will ease the load on the TAIR employees and make more tools available to TAIR users. I know there is the issue of quality and TAIR does an excellent job maintaining this, but if it could be noted that this is not supported by TAIR or something- but still a link provided. One example of something I would like to see linked to is IHOP- so if I find a gene of interest in Arabidopsis I can look and see if it's function is known in other organisms. Of course, these are all just suggestions, I really do think TAIR is exceptional compared to other organisms, but it never hurts to improve. Thanks!

- I usually face problem in downloading some of the software related to promoter element prediction and analysis in a large scale.
- TAIR is a very well informed site, the data presentation and retrieval system is real well. whenever contacted the curator has been real fast in his/her response
- Good job, keep going. thanks, andrea
- Your new GBrowse browser is brilliant. I would love to see QTL information integrated into this browser as well as protein location information pictures (as in Wormbase).
- Thanks for the great work. I think the site could have been faster, and opening of the links in new windows could be optimized. Links to orthologs in other model organisms would be very very useful. Thanks again.
- Nice job, guy!
- good and can be better
- I would like to see the database for Rice on TAIR site.
- We were just discussing today at a European project meeting with Koornneef, Weigel, etc... that the community really had to find a way to store and make available phenotypes, genotypes and QTL data from the massive quantitative analysis generated by many groups on RIL sets and accessions.
- I very much appreciate the TAIR service. Maybe you can include also virtual restriction digests and in silico cloning tools and a site where published constructs can be looked up and requested. With kind regards Marie-Theres Hauser
- Fix the searches! My most common use of TAIR is to find information about a gene or stock. When I enter a gene name I want to go straight to the locus page, I don't want to have to search through irrelevant answers and have to click to get there. As an example try putting 'pin1' in the gene search - the correct result is the fourth one! A better stock search would be very helpful too - it is often very hard to find what I am looking for. Keep up the great work.
- Great resources
- some mis annotations and position of insertion seriously throws one off the track while trying to find a homozygous line for a gene.

- An extremely valuable resources even for researchers like us who work on related but Non-Arabidopsis plants.
- More speedy..
- You do a great job! TAIR Web pages sometimes seem a little slow, but well worth the wait.
- When I become depending on TAIR on daily work, I just wish that TAIR will provide the current service as long as possible. I can not imagine my rearch without TAIR
- I find TAIR to be a great website in general. As a computational biologist, I mainly use the "download" section. I would like to see more explanations and details in the "READMEs" there (what the columns of each files are, the date, citation and abstract of the paper that published this data etc..).
- When searching for genes with the locus code, it would be helpful if the gene searched for is at the top of the list with the other related genes below it. If many other genes have the gene of interest mentioned then it is bothersome to search the list. Otherwise a top piece of work and the envy of other-organism labs!
- For the relevancy and importance of the data, this site needs to move quicker.
- 1) Annotate protein sequences with protein modifications (e.g., highlight phosphorylated Ser/Thr residues, show protein cleavage sites -SEC//AED, etc) and underline domains with different colour underlines to delineate domains 2) I really like the TAIR system and keep saying to myself that other genomes' sites should be organized like TAIR's- especially when I use the JGI poplar database!
- This is an incredibly valuable resource. Unfortunately I don't think I take full advantage of it, but encourage my students to do so. The workshops at meetings are great! Keep up the great work.
- it is very best site to Asses new information about the plant science research . very good site for jobs also we can find jobs easily which are updated daily.
- the backbone of Arabidopsis research! Thanks...
- Since TAIR is specific for Arabidopsis (and at least I use it exactly because of that) it should be very carefully thought through whether the addition of data from other organisms is wise. It might be useful to have links to for example orthologous genes from other organisms, but it should not blur the Arabidopsis thaliana information. I value smaller amount of well-curated information over wealth of automatically generated data.
- I love TAIR, it is my primary source for Arabidopsis information.
- Microarray tools are extremely slow to download and don't have a lot of variety
- If TAIR could provide GABI-kat T-DNA insertion lines, that would be great. It is very inconvenient for us researchers in North America to obtain them through NASC.
- excellent service - I find it the best for Arabidopsis stuff
- - I often wished to have a more direct access to the cds or protein sequence from the gene accessions site. Currently, they are only linked and more than two clicks away. - search options for microarray expression data are manifold, but I stopped searching with TAIR

for "stress" or "mutant name" as I never returned the results I wanted. That works much better via NASC.

- IT WOULD BE NICE TO BE ABLE TO SEARCH FOR MORE THAN ONE ATG NUMBER AT A TIME. WHEN SEARCHING FOR MICROARRAY ELEMENTS OR ATG LOCUS, IT WOULD BE NICE THAT THE RESULTING INFORMATION COULD BE PRESENTED IN A WAY THAT WOULD ALLOW US TO COPY ONLY SELECTED INFORMATION (SUCH AS ANNOTATION) RATHER THAN THE WHOLE TABLE, TO BE ABLE TO COPY INTO EXCEL, FOR EXAMPLE.
- It's already almost perfect. TAIR is the BEST
- keep up the very good work
- I would find a FAQ page or a forum helpful, to be able to discuss problems e.g. concerning tool with other users.
- Thank you, TAIR, for taking on this monumental task and overall doing an admirable job. Your expertise, dedication, and perseverance are greatly appreciated. One complaint I have is that sometimes the most basic information about a gene (in summarized form) seems to be lost in the detail. Often I want to get a quick feeling for what is known about a gene and I find myself going through all kinds of different links and websites to get my desired "snapshot". Have you thought about the "snapshot" idea for simplifying and consolidating many of the essential details into one place for easy reading? For each gene, it would be wonderful to know, for example, if a full-length cDNA and/or splice variants had been identified, a mutant phenotype associated with a loss of function had been described, a protein interactor had been characterized, a cellular function had been experimentally determined, and a putative ortholog in a non-plant model system had been found, just to name a few that come to mind. I realize that different people will come up with different prioritized lists of what to include in such a snapshot, but I encourage you to at least consider and discuss the concept. Another complaint is that the gene annotation (predicted and/or confirmed cellular function of a gene product) is often confusing, complex, and outdated. TAIR should be the most authoritative place to turn when I want to understand what a particular gene does or appears to do. In reality, that is not the case. And I find that unsettling. Perhaps in your effort to build a comprehensive database you have lost track somewhat of being an authoritative resource for the most basic and critical information. I urge you to decide what things you want to master and excel at and what other things you can let other people provide. I don't believe TAIR needs to be all things to all people, provided you can direct people to other sites that complement your own. But I think you must be the definitive source of baseline information for Arabidopsis, and that information needs to be exceptionally well curated and frequently updated.
- Keep up your great work. Thanks!
- Getting gene sequence information is still somewhat cumbersome and circuitous to find on the site. I would like to be able to type in a gene identifier (At xxx) and directly go to the sequence, have it highlighted in color for introns/exons, have it cut-and-paste-able into other applications (such as a primer design function) and see more than 10kb at a

time. And it would be great to have a sliding window to "scroll" up or down 5' or 3' without having to jump in 5kb bits.

- Registration is very complicated. Also payment via credit card is not so popular in Germany. It would be easier for a company like us to use wire transfer and pay a bit more. Best regards, Frank
- Thank you for your work, it's really essential The updating with the published work should be improved as much as possible
- Fewer tools done better.
- Refine the seed stock donation forms, perhaps separate forms for seeds, DNA, etc.
- Sometimes when we blast an intergenic sequence (for example, a sequence obtained by TAIL-PCR from an activation tagging mutant) we get the hit to the whole BAC, but it is hard to position the sequence with reference to the genes in the BAC. We go then to sequence viewer and solve it. I don't know whether we are not using the blast tool properly, but it would be useful to be able to see the position (coordinates) of the genes in a BAC. For the rest, thanks!! The site is extremely useful for our research.
- very good
- I have been using TAIR website for the past one year. It has been useful a lot for my research. Some time I needed NCBI for better diagrams for my presentation, which are friendlier for audience in a presentation. As a graduate student I think there must be more help material (PowerPoint tutorials) to utilize TAIR site more efficiently.
- TAIR can do a lot but cannot do it all. I hope that TAIR advocates for the creation of foundational tools to be created that benefit the Arabidopsis as well as the plant community as a whole. For example, the ability for users to enrich existing knowledge bases that curate published papers, with the support of existing tools (the collaboration with textpresso is a step in that direction); the integration of the many different formats in which metabolic pathways are currently compiled.
- Virtual PCR tool for the Arabidopsis genome to improve the reliability of generated primers and reduce the risk of unspecific primers.
- nice work, keep on going
- My lab is very happy and pleased with the opportunities your people give us. It is a great help in our daily lab work. Thank you very much!!!
- The literature linked to genes for example very incomplete. It would be useful to have better links with CATMA probes Network building information would be useful - prior information available on each gene - what it interacts with, what activates it etc.
- just don't let it slow down to the speed it used to be. :)
- TAIR is already a pretty good ONE-STOP-SHOP for all Arabidopsis related work, however, linking to other organism-databases would greatly improve the usefulness
- The search engine is stymied sometimes by extraneous characters. Also, it would be nice to have a broad search function (e.g., "starts with")
- You are doing a fantastic job! The new changes to TAIR are greatly appreciated.

Analysis

This section addresses the following research questions:

1. Is there any significant difference on the major questions between the USA-based respondents and community members based elsewhere?
2. Do crop researchers have different needs than the general user population?
3. Do computational biologists and bioinformaticists have different needs than the general user population?
4. Do frequent users of the TAIR web site have different opinions on any question?

In analyzing these relationships, I use n-way contingency table analysis with various parametric and non-parametric statistics such as Pearson's χ^2 test, Cramér's V, and proportion difference prob-values to assess statistical significance and degree of association between the two variables I am comparing. Consult any good introductory statistics text on non-parametric statistics for details, or try http://www.unesco.org/webworld/idams/advguide/Chapt4_2.htm for an excellent summary.

The analysis evaluates significance based on a χ^2 statistic, reporting prob values at or below .05 (95% confidence in rejecting the null hypothesis). Each section reports results for any significant relationship; anything not mentioned is not significant at the 95% confidence level.

The analysis evaluates the magnitude of differences based on Cramér's V, a nonparametric association measure which varies between 0 and 1. Many of the reported magnitudes are in the 0.1-0.2 range, which is relatively small; none of the reported statistics are greater than 0.5, or moderate. The analysis reports the major value differences in percentage terms for illustration of the magnitude. For proportional data (0-1 values, future data types and future tools), each section also evaluates the difference between the proportions between groups using standard prob-values.

Finally, the sections for satisfaction measures evaluate the difference between the medians as a non-parametric central tendency test, the K-sample equality of medians test. The χ^2 statistic prob-value for this test evaluates the null hypothesis that the samples were drawn from populations with the same median.

Most of the analysis was done with both the full sample (387 self-selected and random responses) and the random sample for comparison, and there were only a few minor differences in terms of statistical significance, so the analysis uses the random sample as the more appropriate data set for this kind of analysis.

US-Foreign Differences

This section analyzes several questions with respect to the difference between US-based respondents and other respondents. It discusses differences with respect to the 5 satisfaction questions, the 11 importance-of-types-of-information questions, the 11 future-data-types questions, and the 7 future-tools questions.

Satisfaction

Performance satisfaction shows a significant difference (χ^2 prob-value 0.027) that is moderate (Cramér's V 0.3399). 67% of USA respondents were very satisfied while 30% of other respondents were very satisfied. The medians are not significantly different (prob-value 0.512).

Data correctness satisfaction shows a significant difference (χ^2 prob-value 0.051) that is moderate (Cramér's V 0.2875). 52% of USA respondents were very satisfied while 25% of other respondents were very satisfied, while 10% of USA respondents were partially satisfied versus 18% of other respondents. The medians are not significantly different (prob-value 0.186).

Information Types

There were no significant differences between USA and other respondents with respect to the importance of various information types in TAIR.

Data Types

There were no significant differences between USA and other respondents with respect to the choice of future data types in TAIR.

Tools

There were no significant differences between USA and other respondents with respect to the choice of future tools in TAIR.

Crop Researcher Differences

This section analyzes several questions with respect to the difference between crop-researcher respondents and other respondents. TAIR curators classified the respondents as crop researchers based on the organisms they listed as ones on which they worked. This section discusses differences with respect to the 5 satisfaction questions, the 11 importance-of-types-of-information questions, the 11 future-data-types questions, and the 7 future-tools questions.

Satisfaction

Data completeness satisfaction shows a significant difference (χ^2 prob-value 0.020) that is moderate (Cramér's V 0.3226). 49% of crop-scientist respondents were satisfied (level 2) while 43% of non-crop-scientist respondents felt that way, while 9% of crop-scientist respondents were dissatisfied versus 0% of non-crop-scientist respondents. The medians are not significantly different (prob-value 0.227).

Information Types

There were no significant differences between crop scientists and other respondents with respect to the importance of various information types in TAIR.

Data Types

Promoters/cis elements, regulation of transcription shows a significant difference (χ^2 prob-value 0.038, proportion difference prob-value 0.38) that is small (Cramér's V 0.2127). 47% of crop-scientist respondents chose this data type while 68% of non-crop-scientist respondents chose it. The medians are not significantly different (prob-value 0.512).

Tools

There were no significant differences between crop scientists and other respondents with respect to the choice of future tools in TAIR.

Bioinformaticist Differences

This section analyzes several questions with respect to the difference between bioinformaticists and other respondents. Bioinformaticists are those who responded to the main-daily-work question with the response "Carry out or supervise bioinformatics or computational biology work." It discusses differences with respect to the 5 satisfaction questions, the 11 importance-of-types-of-information questions, the 11 future-data-types questions, and the 7 future-tools questions.

This section uses the full survey, as there were only 3 bioinformaticists in the random sample, a very skewed sample. No proportion difference prob-values appear, as this depends on having a random sample.

Satisfaction

Data completeness satisfaction was significantly different (χ^2 prob-value 0.022) to a small degree (Cramér's V 0.1721). 24% of bioinformaticist respondents were very satisfied (level 1) while 26% of non-bioinformaticist respondents felt that way, while 3% of bioinformaticist respondents were dissatisfied versus 0% of non-bioinformaticist respondents. The medians are not significantly different (prob-value 0.982).

Information Types

Alternate isoforms (splice variants) information importance was significantly different (χ^2 prob-value 0.010) to a small degree (Cramér's V 0.1878). 18% of bioinformaticists thought splice variants were important (level 3) as opposed to 37% of other respondents.

Polymorphisms/SNPs/T-DNA and Tn insertions information importance was significantly different (χ^2 prob-value 0.000) to a moderate degree (Cramér's V 0.2850). 30% of bioinformaticists thought this kind of information was extremely important (level 1) as opposed to 65% of other respondents. 18% of bioinformaticists thought this kind of information was important (level 3) as opposed to 8% of other respondents. Finally, 9% of bioinformaticists thought this kind of information was unimportant (level 5) as opposed to 1% of other respondents.

Experimentally verified gene function data importance was significantly different (χ^2 prob-value 0.018) to a small degree (Cramér's V 0.1767). 70% of bioinformaticists thought this kind of data were extremely important (level 1) as opposed to 65% of other respondents.

Seed and DNA stock information importance was significantly different (χ^2 prob-value 0.000) to a moderate degree (Cramér's V 0.4493). 36% of bioinformaticists thought this kind of data were extremely important (level 1) as opposed to 76% of other respondents. 15% of bioinformaticists thought this kind of data were unimportant as opposed to less than 1% of other respondents.

Mutant phenotypes information importance was significantly different (χ^2 prob-value 0.000) to a moderate degree (Cramér's V 0.2806). 40% of bioinformaticists thought this kind of data were

extremely important (level 1) as opposed to 63% of other respondents. 6% of bioinformaticists thought this kind of data were unimportant as opposed to less than 1% of other respondents.

Data Types

Complete ecotype sequences as future data was significantly different (χ^2 prob-value 0.015) to a small degree (Cramér's V 0.1233). 21% of bioinformaticists chose this kind of data as opposed to 42% of other respondents.

Tools

Customizable bulk data retrieval and download tool (e.g. BioMart) as a future tool was significantly different (χ^2 prob-value 0.033) to a small degree (Cramér's V 0.1084). 59% of bioinformaticists chose this tool as opposed to 40% of other respondents.

Protein alignment viewer as a future tool was significantly different (χ^2 prob-value 0.007) to a small degree (Cramér's V 0.1367). 29% of bioinformaticists chose this tool as opposed to 54% of other respondents.

Protein interaction viewer as a future tool was significantly different (χ^2 prob-value 0.046) to a small degree (Cramér's V 0.1015). 38% of bioinformaticists chose this tool as opposed to 56% of other respondents.

Frequent Use of TAIR

This section analyzes several questions with respect to the difference between frequent TAIR users and other respondents. A frequent TAIR user is one who uses TAIR daily or weekly. This section discusses differences with respect to the 5 satisfaction questions, the 11 importance-of-types-of-information questions, the 11 future-data-types questions, and the 7 future-tools questions.

Satisfaction

Availability satisfaction shows a significant difference (χ^2 prob-value 0.041) that is moderate (Cramér's V 0.3238). 50% of frequent users were very satisfied (level 1) while 58% of infrequent users felt that way, while 42% of frequent users were satisfied (level 2) versus 16% of infrequent users and 5% of frequent users were partially satisfied (level 3) versus 11% of infrequent users. In other words, the frequent users tended toward satisfied rather than very or partially satisfied. The medians are not significantly different (prob-value 0.538).

Organization of the web site satisfaction shows a significant difference (χ^2 prob-value 0.037) that is moderate (Cramér's V 0.3274). 33% of frequent users were very satisfied (level 1) while 21% of infrequent users felt that way, while 5% of frequent users were dissatisfied (level 4) versus 21% of infrequent users. Infrequent users were thus less satisfied overall with site organization. The medians are significantly different (prob-value 0.020), with the median for frequent users being satisfied and the median for infrequent users tending toward being partially satisfied.

Data correctness satisfaction shows a significant difference (χ^2 prob-value 0.023) that is moderate (Cramér's V 0.3184). 35% of frequent users were very satisfied (level 1) while 16% of infrequent users felt that way, while 44% of frequent users were satisfied (level 2) versus 63% of infrequent users and 19% of frequent users were partially satisfied (level 3) versus 5% of

infrequent users. In other words, the frequent users tended toward very satisfied while infrequent users tended toward satisfied. The medians are not significantly different (prob-value 0.979).

Information Types

There were no significant differences between frequent and infrequent TAIR users with respect to the importance of various information types in TAIR.

Data Types

Orthologous genes from other plants and model organisms have significantly different proportions (prob-value .049, 95% confidence interval 49.3%-0.7% for the difference) but the contingency table differences are not significant. Frequent users choose this future data type 46% of the time, while infrequent users choose it 37% of the time.

Tools

There were no significant differences between frequent and infrequent TAIR users with respect to the choice of future tools in TAIR.

Appendix A: Data Quality

This appendix contains the details on data quality issues in the survey responses and their resolution.

There were 26 responses that came from the same IP address and might be considered duplicate submissions. Most such groupings had two responses from the same address; one had four. All of these responses look different and are, in our judgment, different and legitimate submissions from the same computer, probably a shared computer.

The analysis methods ignore null answers to questions, removing them from the analysis and lowering the number of responses (n).

One response (respondent id 593209919) was completely null and was removed.

The question on working on other organisms was free text entry. We standardized the responses to eliminate duplicate spellings and common names where appropriate. See [Appendix B: Organisms](#) for the complete list of organisms entered.

The entries for Institutional Type--Other contained one entry "University" which was recoded to "academic."

The entries for Geographic Location--Other contained one entry "China" which was recoded to "Asia."

The questions on future data, future tools, and data submission instructed the respondent to choose "up to three" responses. Many (>90) chose more than that, as the survey instrument did not enforce the restriction. It was decided to ignore the restriction for purposes of analysis.

Appendix B: Organisms

This is a complete list of the standardized names for organisms entered by respondents in answer to the question about other organisms on which they worked. The list is ordered by count of occurrence in the survey, then alphabetically for tied counts.

<i>Saccharomyces cerevisiae</i>	49
<i>Oryza sativa</i>	47
<i>Escherichia coli</i>	44
<i>Nicotiana tabacum</i>	42
<i>Arabidopsis thaliana</i>	24
<i>Lycopersicon esculentum</i>	24
<i>Zea mays</i>	16
<i>Brassica napus</i>	13
<i>Nicotiana benthamiana</i>	13
<i>Populus</i>	13
<i>Triticum aestivum</i>	13
<i>Pseudomonas syringae</i>	12
<i>Medicago truncatula</i>	11
<i>Physcomitrella patens</i>	10
bacteria	9
<i>Glycine max</i>	9
<i>Hordeum vulgare</i>	9
<i>Agrobacterium tumefaciens</i>	8
<i>Pisum sativum</i>	7
<i>Brachypodium distachyon</i>	5
<i>Brassica</i>	5
<i>Caenorhabditis elegans</i>	5
<i>Gossypium hirsutum</i>	5
<i>Vitis vinifera</i>	5
<i>Alternaria brassicicola</i>	4
<i>Antirrhinum majus</i>	4
<i>Arabidopsis lyrata</i>	4
<i>Brassica oleracea</i>	4
<i>Brassica rapa</i>	4
mouse	4
Plants	4
<i>Solanum tuberosum</i>	4
algae	3
bean	3
<i>Chlamydomonas reinhardtii</i>	3
<i>Picea abies</i>	3
<i>Pichia pastoris</i>	3
bright yellow 2 (BY2) cells	2
<i>Drosophila melanogaster</i>	2
<i>Homo sapiens</i>	2
<i>Lactuca sativa</i>	2
<i>Linum usitatissimum</i>	2
Mammalia	2
moss	2
onion	2
<i>Panicum virgatum</i>	2

Phaseolus vulgaris	2
powdery mildew	2
Selaginella	2
Adathoda vasica	1
Andrographis paniculata	1
Arabidopsis cebennensis	1
Arabidopsis halleri	1
Arabidopsis petraea	1
Artemisia annua	1
Avena sativa	1
Bacillus subtilis	1
bacteria/rhizobacteria	1
Betula	1
Betula papyrifera	1
Betula pendula	1
Boechera divaricarpa	1
Botrytis cinerea	1
Burkholderia malei	1
Capsella	1
Cardamine pratensis	1
Catharanthus roseus	1
Centella asiatica	1
Ceratopteris	1
chrysanthemum	1
Citrus sinensis	1
Coffea canephora	1
cucumber	1
Cyanidioschyzon merolae	1
Cyanobacteria	1
Daucus carota	1
fission yeast	1
Fragaria vesca	1
Fungi	1
Geranium lucidum	1
Glycine tomentella	1
Glycyrrhiza	1
Heterodera schachtii	1
Hyaloperonospora arabidopsis	1
Hyaloperonospora parasitica	1
insect like bollworm	1
legumes	1
Lemna minor	1
Lilium longiflorum	1
Lotus japonicus	1
Maesa lanceolata	1
mammalian cell culture	1

MDCK cells	1
Meloidogyne incognita	1
miscanthus	1
mugwort	1
Musa acuminata	1
Myzus persicae	1
nematode	1
Neurospora	1
Orchidaceae	1
parsley	1
perennial fruit trees	1
Peronospora	1
Pinus taeda	1
Plectosphaerella cucumerina	1
Populus trichocarpa	1
Prunus persica	1
ragweed	1
Ralstonia solanacearum	1
Rattus	1
Saccharomyces paradoxus	1
Salix	1
scarlet runner bean	1
Solanum commersonii	1
Solanum xanthocarpum	1
sugarcane	1
Synechocistis PCC6803	1
Tetranychus urticae	1
Thalictrum	1
Thlaspi caerulescens	1
various plant species	1
Verticillium	1
Vicia faba	1
Withania somnifera	1
woad	1