

EDITORIAL

Large-Scale Biology

This month, *The Plant Cell* introduces a Special Series on Large-Scale Biology comprised of Current Perspective Essays commissioned and edited by Coeditor Sarah Assmann and News and Reviews Editor Nancy Eckardt. A broad range of topics will be covered, including transcriptomics, proteomics, metabolomics, epigenomics, genetic association studies, statistical approaches, network analysis, and systems biology. Presumably, all plant biologists are now aware that these new fields of research are going to have a tremendous impact on the whole of biology. Our purpose in presenting this new series is to make large-scale biology more accessible to all plant biologists and to explain and discuss the new insights into biological processes of plants that are obtained from large-scale studies.

Although large-scale studies will contribute significantly to systems biology, it is worth noting that we have not entitled this series “systems biology.” This is intentional. In our view, it is important that the definition of systems biology be consistent with formal dynamic systems theory. Elucidation and description of the basic elements of a system, such as biochemical pathways and molecular interactions, although perhaps necessary for systems biology, are not sufficient for it. True systems biology requires mathematical modeling and simulations of dynamic networks. Although large-scale investigations will be an important contributor to systems biology, they will have to be combined with formal mathematical approaches to produce a true systems perspective. If all goes well, systems biology will lead to the discovery of novel, emergent properties of the molecules and interactions that drive network behavior as well as new higher-order principles of biology. We plan to address the nature of systems biology in some detail in future issues.

The first installment in the Special Series on Large-Scale Biology presents essays on

transcriptome dynamics from two complementary scientific perspectives: that of the biologist and that of the statistician. In the first article, Philip Benfey and colleagues offer their perspective on the current state of transcriptomics research, its possibilities, and the likely near-term direction for describing dynamic changes in the transcriptome. In the second article, Dan Nettleton explains why and how good statistical design and analysis principles are an essential component of transcriptome biology.

Importantly, these two points of view are interdependent. Just as molecular biologists cringe when biologists who lack molecular training misuse a method such as the polymerase chain reaction without controlling for common artifacts—a prominent example being the purported detection of transgenes in Mexican farmers' maize fields (Quist and Chapela, 2001; Ortiz-Garcia et al., 2005)—statisticians cringe as well whenever biologists attempt to interpret large data sets using arbitrary cutoffs and to draw conclusions that have no statistical support. Even if arbitrary fold changes may have been considered acceptable by biologists in the early days of microarray analysis, the fact is they fail to provide any useful information about the statistical significance of the results. As discussed by Nettleton in this issue, good statistical analysis always begins with proper experimental design. Applying statistics to an improperly designed experiment is about as likely to succeed as trying to reflate a soufflé after it has fallen. This is why *The Plant Cell* has long recommended that authors involve a statistician from the outset, rather than waiting to call one in only after the data have been collected, in the hope that an expensive experiment might somehow be salvaged.

The Current Perspective Essay on statistical methods by Nettleton in this issue provides an excellent overview of the concepts, principles, and methods for the proper design and analysis of microarray

experiments that we hope will help readers understand the need for and nature of statistical approaches in large-scale biology. However, as good as the article is, it is important to recognize that no single article could provide all the information that a biologist might need to properly apply statistics to the interpretation of large-scale experiments. That would require at least a book, a course, or more likely, an academic degree. Thus, the most effective large-scale experimentalists are likely to be those who know that “a little knowledge can be a dangerous thing” and so choose to collaborate with trained statisticians who possess the requisite expertise to help them properly design and analyze their experiments.

The Plant Cell takes pride in its leading role in establishing and maintaining high standards for the design and analysis of large-scale data sets, standards that are being emulated by other journals. The individual primarily responsible for the journal's leadership in this area is Coeditor Rebecca Doerge, a statistician and statistical geneticist with an extensive publication record in this area. Doerge's expertise has been indispensable for developing and implementing the journal's rigorous and flexible approach for evaluating and approving manuscripts that present large-scale experiments. We sincerely thank her for her unselfish contributions.

How does *The Plant Cell* evaluate manuscripts with large-scale data sets? As described in our Instructions for Authors, the criteria for assessing microarray experiments include the following: whether the description of the experiment is sufficiently clear and complete so that the reader can understand the experimental approach and assess the validity of the design; whether biological and/or technical replicates should have been used and are appropriately described; what statistical analysis has been performed; whether a multiple comparison correction, such as a false discovery rate calculation, has been used

to control the type I family-wise error rate; and whether the conclusions gained from interpreting the large-scale data set should be independently validated.

To be scientifically consistent, the journal does not allow statistically unsupported conclusions to be drawn from other articles that have been published elsewhere, irrespective of the reputation of the journal; as every good scientist knows, publication cannot magically make an unsound conclusion sound. On the other hand, it is also

the case that large-scale experiments that are used only as screens for candidate genes do not require statistical analysis to be published in *The Plant Cell*, as long as the manuscript makes this purpose clear and does not attempt to draw conclusions (such as that a certain fraction of genes responds in a given manner).

Rich Jorgensen
Editor
raj@ag.arizona.edu

REFERENCES

- Ortiz-Garcia, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberon, J., and Snow, A.A.** (2005). Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003–2004). *Proc. Natl. Acad. Sci. USA* **102**, 12338–12343.
- Quist, D., and Chapela, I.H.** (2001). Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* **414**, 541–543.

Large-Scale Biology
Rich Jorgensen
Plant Cell 2006;18;2095-2096
DOI 10.1105/tpc.106.180980

This information is current as of February 26, 2021

References	This article cites 2 articles, 1 of which can be accessed free at: /content/18/9/2095.full.html#ref-list-1
Permissions	https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&issn=1532298X&WT.mc_id=pd_hw1532298X
eTOCs	Sign up for eTOCs at: http://www.plantcell.org/cgi/alerts/ctmain
CiteTrack Alerts	Sign up for CiteTrack Alerts at: http://www.plantcell.org/cgi/alerts/ctmain
Subscription Information	Subscription Information for <i>The Plant Cell</i> and <i>Plant Physiology</i> is available at: http://www.aspb.org/publications/subscriptions.cfm