

# Modeling Transcriptome Dynamics in a Complex World

Philipp A. Jaeger,<sup>1</sup> Colleen Doherty,<sup>2</sup> and Trey Ideker<sup>1,\*</sup>

<sup>1</sup>Departments of Medicine and Bioengineering

<sup>2</sup>Division of Biological Sciences

University of California San Diego, La Jolla, CA 92093, USA

\*Correspondence: [trey.ideker@gmail.com](mailto:trey.ideker@gmail.com)

<http://dx.doi.org/10.1016/j.cell.2012.11.031>

**An accurate prediction of how extrinsic stimuli influence changes in gene expression has been challenging. In this issue, Nagano and colleagues successfully model genome-wide mRNA expression changes under variable environmental conditions in rice, raising hopes that scientists will soon be able to predict genome-wide transcriptional responses in a variety of organisms in uncontrolled real-world settings.**

A longstanding challenge in biology has been to predict gene expression changes in response to a varying environment. In model species such as *Arabidopsis* or yeast, a popular approach has been to model gene expression patterns in response to a panel of perturbations to individual environmental factors in otherwise controlled conditions. Ultimately, however, one needs to extrapolate these models to “real-world” scenarios, such as predicting the gene expression response to the complex conditions seen in healthy or diseased tissues or in the field.

In this issue, Nagano et al. (2012) take an entirely different approach by building a model of the transcriptional response of rice grown directly in the field under real-world conditions. Climate, developmental age, and genotype are used as input to a “simple” linear model to predict the genome-wide transcriptional response. Surprisingly, the model accurately predicts most expression changes in the rice plants based on atmospheric data and developmental age alone. This study successfully demonstrates that models incorporating relevant information from the complex surrounding environment can yield improvements over the more traditional controlled environment approaches. Nagano and colleagues thus infuse renewed vigor into attempts to model other organisms’ transcriptome changes under circumstances that cannot be modeled in the laboratory, such as changes in human gene expres-

sion patterns that occur during development or disease.

In recent years, major advances in high-throughput technology have vastly increased our ability to measure changes in the “omics” world. We can now routinely measure DNA sequences, epigenetic changes, transcription factor binding, mRNA expression, protein production and decay, and numerous metabolic changes (Chen et al., 2012). Transcriptome changes in response to environmental or developmental stimuli have been studied in a variety of model organisms, for example, in single-cell bacteria (Bonneau et al., 2007), yeast (Nagalakshmi et al., 2008), maize (Li et al., 2010), fruit flies (Graveley et al., 2011), mice (Okazaki et al., 2002), and even humans (Kang et al., 2011; Djebali et al., 2012). However, more important than monitoring static transcriptome status is our ability to develop robust models to predict dynamic transcriptome changes in response to complex and ever-changing stimuli. Here, rice provides an excellent model organism, as it has a well-annotated genome and available high-throughput technologies, and it can be characterized easily with high temporal resolution across key environmental stimuli and developmental stages.

Previous attempts in plant biology, for example in *Arabidopsis* or rice (Opgen-Rhein and Strimmer, 2007), have successfully generated gene regulatory networks for a limited number of genes based on expression patterns (Walley

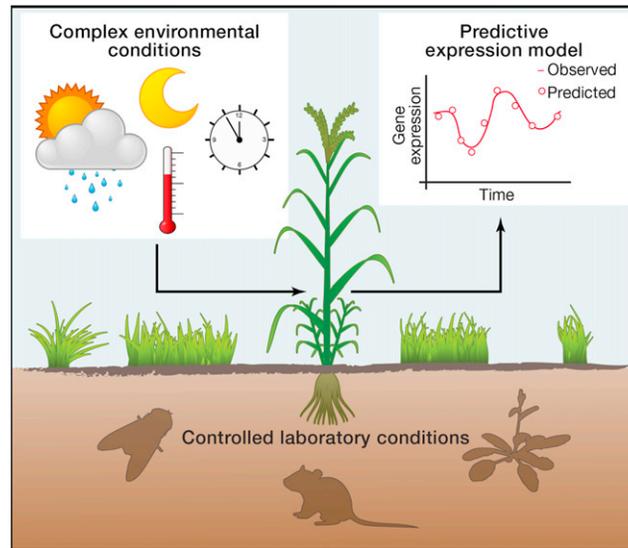
and Dehesh, 2010). These have provided targets and markers for investigations in crop plants (Ferrier et al., 2011). However, few such approaches have incorporated detailed growth and environmental conditions that replicate actual field conditions, and of those that have, it remains unclear how successful the translation from controlled growth chambers to actual field conditions will be.

Taking advantage of these benefits, Nagano and colleagues sampled mRNA expression values from field-grown rice leaves at various seasonal, diurnal, and developmental time points. Continuous monitoring of the atmospheric conditions in the field enabled them to associate changes in expression with environmental conditions. They tested these models on plants grown the following year, in which the environment became substantially warmer, and found that the models were nonetheless highly predictive for the expression of the majority of the genome. This surprising success is perhaps due to a unique incorporation of the biological understanding of the effects of environment and of how variables were incorporated into their model. For example, the effects of the environment were allowed to impact the expression models either directly or through a gating (time-of-day-dependent) mechanism imposed by the circadian clock.

Intriguingly, although the Nagano et al. model performs well in the field, it performs less well on predicting expression changes in plants grown under controlled

laboratory conditions. It will be interesting to investigate whether this reduced predictive ability is an artifact of the modeling approach or whether it reflects a genuine need of the plants for environmental variability to establish stable expression patterns. The latter possibility might have far-reaching consequences for other modeling attempts. Hypothesis generation in biology is often based on highly controlled laboratory organisms; however, the transcriptome or any other omics-derived network models based on these organisms might be “unstable” due to a lack of complex environmental inputs. This idea is supported by observations in yeast that only about 60% of genes show epistatic interactions under standard growth conditions, whereas the other 40% contribute to the genetic network only once the cell is stressed (Hillenmeyer et al., 2008; Dowell et al., 2010). This finding suggests that a complex environment with changing conditions could substantially improve network connectivity, reduce random network noise, and ultimately improve the model’s predictive power.

For the plant biology community, this research represents a leap forward in understanding which environmental variables are predictive of gene expression. However, possibly the most significant impact will be to identify the links between the predictive atmospheric and developmental factors and the regulatory processes and signaling pathways that mediate this response. Such mapping will provide molecular targets for breeding and future research to improve crop yield



**Figure 1. Breakthrough into a Complex World**

Successful modeling and prediction of gene expression changes in rice under variable field conditions leads model organisms into a new era of real-world “omics.”

and performance. Extending this in-the-field modeling approach to include the effects of stress factors such as pathogens or drought may offer insights into molecular mechanisms that can circle back into basic research (Figure 1). Moreover, determining the effects that stresses such as high CO<sub>2</sub> or increased temperature have on gene expression will have significant impact for understanding the effects of climate change on the molecular networks of plants.

#### ACKNOWLEDGMENTS

This work was supported by NIH grants GM084279 and GM085764.

#### REFERENCES

Bonneau, R., Facciotti, M.T., Reiss, D.J., Schmid, A.K., Pan, M., Kaur, A., Thorsson, V., Shannon,

P., Johnson, M.H., Bare, J.C., et al. (2007). *Cell* 131, 1354–1365.

Chen, R., Mias, G.I., Li-Pook-Than, J., Jiang, L., Lam, H.Y.K., Chen, R., Miriami, E., Karczewski, K.J., Hariharan, M., Dewey, F.E., et al. (2012). *Cell* 148, 1293–1307.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). *Nature* 489, 101–108.

Dowell, R.D., Ryan, O., Jansen, A., Cheung, D., Agarwala, S., Danford, T., Bernstein, D.A., Rolfe, P.A., Heisler, L.E., Chin, B., et al. (2010). *Science* 328, 469.

Ferrier, T., Matus, J.T., Jin, J., and Riechmann, J.L. (2011). *Curr. Opin. Biotechnol.* 22, 260–270.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). *Nature* 471, 473–479.

Hillenmeyer, M.E., Fung, E., Wildenhain, J., Pierce, S.E., Hoon, S., Lee, W., Proctor, M., St Onge, R.P., Tyers, M., Koller, D., et al. (2008). *Science* 320, 362–365.

Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). *Nature* 478, 483–489.

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., Patel, R., Myers, C.R., et al. (2010). *Nat. Genet.* 42, 1060–1067.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). *Science* 320, 1344–1349.

Nagano, A.J., Sato, Y., Mihara, M., Antonio, B.A., Motoyama, R., Itoh, H., Nagamura, Y., and Izawa, T. (2012). *Cell* 151, this issue, 1358–1369.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., et al; FANTOM Consortium; RIKEN Genome Exploration Research Group Phase I & II Team. (2002). *Nature* 420, 563–573.

Opgen-Rhein, R., and Strimmer, K. (2007). *BMC Syst. Biol.* 1, 37.

Walley, J.W., and Dehesh, K. (2010). *J. Integr. Plant Biol.* 52, 354–359.