Yeast Systems Biology: Our Best Shot at Modeling a Cell

Charles Boone

Banting and Best Department of Medical Research, Donnelly Center for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada M5S 3E1

THE Genetics Society of America's Edward Novitski Prize recognizes an extraordinary level of creativity and intellectual ingenuity in the solution of significant problems in genetics research. The 2014 recipient, Charles Boone, has risen to the top of the emergent discipline of postgenome systems biology by focusing on the global mapping of genetic interaction networks. Boone invented the synthetic genetic array (SGA) technology, which provides an automated method to cross thousands of strains carrying precise mutations and map large-scale yeast genetic interactions. These network maps offer researchers a functional wiring diagram of the cell, which clusters genes into specific pathways and reveals functional connections.

It is safe to say that yeast is better understood than any other cell. Hundreds of labs worldwide have for decades been studying this powerful genetic model from various perspectives, and we have made spectacular advances in understanding most pathways and cellular functions. Nevertheless, none of us would claim to know how the yeast cell really works.

The major problem is that the ever-growing mass of detailed biological information has not yet been assembled into a complete and integrated picture. Ultimately, if we are going to model life on a whole-cell level, we must understand how all of its components are connected and coordinated. Only then will we be able to predict the physiological responses to a specific genetic or environmental perturbation. Although a daunting task, I think that modeling the cell is well within our grasp and, while I may be biased, I believe the budding yeast offers our best shot at realizing this challenge.

The landmark first step toward a comprehensive understanding of a cell was made by Andre Goffeau's international team, who assembled the first sequence of a eukaryote (Goffeau *et al.* 1996). Having a complete picture of the yeast genome opened the door to functional genomics approaches that had previously been barely imaginable. As a postdoctoral fellow at the University of Oregon, I remember being blown away by what then seemed an absolutely wild idea. Over dinner, our visiting speaker, Stan Fields, described how he was planning to clone every yeast gene

in an attempt to test all possible yeast protein pairs, covering an entire 6000×6000 matrix, for physical interactions using his two-hybrid assay (Uetz *et al.* 2000). At the same time, several different yeast groups, both academic (Derisi 1997; Eisen *et al.* 1998) and commercial enterprises (Dimster-Denk *et al.* 1999; Hughes *et al.* 2000), were pioneering genome-wide gene expression analysis to reveal global transcriptional responses. Meanwhile, systematic phenotypic screens were enabled by genome-scale mutant collections, assembled in the form of transposon mutagenesis libraries (Ross-Macdonald *et al.* 1999) and deletion collections (Winzler *et al.* 1999; Giaever *et al.* 2002). Taken together, these systems-level approaches provided a new integrated way of thinking about science, one that inspired our automated form of yeast genetics called synthetic genetic array (SGA) analysis (Tong *et al.* 2001), a methodology designed to map genetic interactions on a genome-wide scale (Tong *et al.* 2004).

Harnessing the expertise and power of the entire yeast research community in a coordinated manner would represent the ultimate systems level approach; biology's version of sophisticated CERN-like science.

—C.B.

Success in functional genomics depends on multidisciplinary teams that can design, implement, and interpret large-scale experimental strategies. An integral part of this process is the computational analysis required to process and quantify the

emerging data. I think the yeast community's culture of open sharing of reagents and ideas prepared us all to embrace this new style of widely collaborative science. The development of the *Saccharomyces* Genome Database (SGD) (Cherry 1998) also played an important role in building this open-access culture by assembling and organizing data from both focused and large-scale studies. Through its team of experts, SGD also curates the data derived primarily from focused studies to generate machine-readable Gene Ontology (GO) annotations for yeast genes (Ashburner *et al.* 2000). While this detailed annotation is critical for communicating our understanding of gene function, it also provides a gold standard for quantifying the functional information derived from large-scale studies, which may vary in quality and breadth. Thus, SGD bridges a gap between highly accurate, but biased, focused studies, and global studies, with the broad potential to both address the roles of previously uncharacterized genes and to map novel functional connections between seemingly unrelated processes.

Precisely because it coordinates all yeast experimental data and makes it generally available, SGD has become the centerpiece of our field. Perhaps most importantly, through SGD, the yeast community has mapped a highly successful model for tackling the functional annotation of a genome. In fact, if I were directing major sources of funding, I would invest heavily in the implementation of a similar SGD strategy for the human genome (*i.e.*, HGD). Having visited the SGD website almost every day since its inception, I can only imagine that an HGD counterpart would have an immeasurable impact on human genetics and our understanding of the human genome.

In the aftermath of the recent financial crisis, our governments are cutting back on basic science funding and, unfortunately, support for a project like HGD is unlikely. Ironically, just at a time when we are beginning to make real headway toward a mechanistic understanding of how life works, the resources dedicated to basic research are shrinking. With this in mind, it seems obvious that HGD would fit neatly into a private information corporation's portfolio. Assembling HGD is bound to be profitable because it ultimately constitutes the basis for precision medicine. In the past, we all worried about corporate interests owning the sequence of the human genome and, ironically, it seems that in today's environment, the private sector may be our only hope to fund a project like HGD.

SGA genetic network analysis has provided our research group with an opportunity to interface with SGD, BioGRID (Stark *et al.* 2006), and other databases to contribute to the functional annotation of the yeast genome. Brenda Andrews, Michael Costanzo, and I have worked together to assemble and implement the methodology and reagents necessary to map a complete genetic interaction network for yeast. Our major computational collaborators, Chad Myers and his team, are largely responsible for figuring out how to extract meaningful information from our large-scale, but relatively noisy, dataset. With SGA analysis, we quantify both negative and positive genetic interactions, where double mutants are scored as growing

worse or better than expected, respectively (Baryshnikova *et al.* 2010). Our assembly of a global genetic network composed of hundreds of thousands of genetic interactions highlights the power of combinatorial genetics for identifying pathways, delineating how they work together to control essential cellular functions, and mapping a functional wiring diagram for the cell (Costanzo *et al.* 2010).

More than 10 years ago, Lee Hartwell and colleagues suggested that genetic interactions may play a key role in our ability to interpret the genotype–phenotype relationship for an individual (Hartman *et al.* 2001). This idea is now gaining traction with both yeast (Bloom *et al.* 2013) and human geneticists (Zuk *et al.* 2012) and will likely become more and more relevant with the imminent sequencing of millions of individual human genomes. While it remains to be proven, given the scope and breadth of the global yeast genetic network, we can most certainly anticipate that genetic interactions and their networks must underlie a significant proportion of human disease phenotypes. Fortunately, there are some simple rules associated with the structure and topology of genetic networks that appear to be generally conserved. In particular, the genes within pathways often behave in a coherent manner, connected to other pathways through a consistent set of positive or negative interactions. Indeed, by extrapolating these rules from the yeast genetic network, Chad's computational team appears to have developed methods with enough statistical power to detect significant genetic interaction signals from genome-wide association studies (GWAS) in humans. Thus, the fundamental properties of genetic networks we learn from yeast may be critical for the interpretation of our own genomes.

The methods pioneered by Charlie Boone have proven to be the richest source of biological interactions known to date, and are foundational for the 'interactomes' that drive much of contemporary genetic experimentation and thought.

—Jasper Rine, University of California, Berkeley

Systems-level technologies developed in yeast and the resultant genome-scale data have fueled the field of bioinformatics and computational biology. Indeed, it is only with detailed computational processing of functional genomics data that we can realize its full potential. However, to build an accurate and comprehensive model of the cell, we must keep pushing the boundaries of both functional genomics and its bioinformatics. I think the next step requires a new operative mode where data are collected at the community level rather than by individual labs. With all of our research enterprises working on the exact same cell, this next level of highly coordinated science is entirely feasible. Our reference strain, S288c, provides us with genetic continuity, which means that quantitative genome-scale data derived from different labs all around the world can be compiled and assembled in a unified

format. Given that each lab has expertise in specific pathways and thus can design exquisite pathway-specific readouts, our community has the potential to coordinate a quantitative analysis of most pathways under the influence of a genome-wide set of genetic perturbations and a standardized set of environmental conditions. It is not clear exactly how to do this, but I suspect this mode of analysis could be achieved through a number of different types of experiments; one obvious approach involves combining automated SGA yeast genetics with cell sorting or high-content screening to quantify the activity of diagnostic reporters in a comprehensive set of mutants (Jonikas *et al.* 2009; Vizeacoumar *et al.* 2010). Although there will surely be technical challenges to overcome, harnessing the expertise and power of the entire yeast research community in a coordinated manner would represent the ultimate systems level approach; biology's version of sophisticated CERN-like science.

The field of yeast systems biology has come a long way (Botstein and Fink 2011), but perhaps now is the time for us to take the next big step. The yeast community should be able to deliver the kind of data that both biologists and theorists require to realize the modeling of a eukaryotic cell.

Literature Cited

- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler *et al.*, 2000 Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25: 25–29.
- Baryshnikova, A., M. Costanzo, Y. Kim, H. Ding, J. Koh *et al.*, 2010 Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. *Nat. Methods* 7: 1017–1024.
- Bloom, J. S., I. M. Ehrenreich, W. T. Loo, T.-L. V. Lite and L. Kruglyak, 2013 Finding the sources of missing heritability in a yeast cross. *Nature* 494: 234–237.
- Botstein, D., and G. R. Fink, 2011 Yeast: an experimental organism for 21st Century biology. *Genetics* 189: 695–704.
- Cherry, J., 1998 SGD: Saccharomyces Genome Database. *Nucleic Acids Res.* 26: 73–79.
- Costanzo, M., A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear *et al.*, 2010 The genetic landscape of a cell. *Science* 327: 425–431.
- DeRisi, J. L., 1997 Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278: 680–686.
- Dimster-Denk, D., J. Rine, J. Phillips, S. Scherer, P. Cundiff *et al.*, 1999 Comprehensive evaluation of isoprenoid biosynthesis regulation in *Saccharomyces cerevisiae* utilizing the Genome Reporter Matrix. *J. Lipid Res.* 40: 850–860.
- Eisen, M. B., P. T. Spellman, P. O. Brown and D. Botstein, 1998 Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95: 14863–14868.
- Giaever, G., A. M. Chu, L. Ni, C. Connelly, L. Riles *et al.*, 2002 Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418: 387–391.
- Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon *et al.*, 1996 Life with 6000 genes. *Science* 274(546): 563–567.
- Hartman, J. L., B. Garvik and L. Hartwell, 2001 Principles for the buffering of genetic variation. *Science* 291: 1001–1004.
- Hughes, T. R., M. J. Marton, A. R. Jones, C. J. Roberts, R. Stoughton *et al.*, 2000 Functional discovery via a compendium of expression profiles. *Cell* 102: 109–126.
- Jonikas, M. C., S. R. Collins, V. Denic, E. Oh, E. M. Quan *et al.*, 2009 Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 323: 1693–1697.
- Ross-Macdonald, P., P. S. Coelho, T. Roemer, S. Agarwal, A. Kumar *et al.*, 1999 Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* 402: 413–418.
- Stark, C., B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz *et al.*, 2006 BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 34: D535–D539.
- Tong, A. H., M. Evangelista, A. B. Parsons, H. Xu, G. D. Bader *et al.*, 2001 Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294: 2364–2368.
- Tong, A. H. Y., G. Lesage, G. D. Bader, H. Ding, H. Xu *et al.*, 2004 Global mapping of the yeast genetic interaction network. *Science* 303: 808–813.
- Uetz, P., L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson *et al.*, 2000 A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403: 623–627.
- Vizeacoumar, F. J., N. van Dyk, F. S. Vizeacoumar, V. Cheung, J. Li *et al.*, 2010 Integrating high-throughput genetic interaction mapping and high-content screening to explore yeast spindle morphogenesis. *J. Cell Biol.* 188: 69–81.
- Winzeler, E. A., D. D. Shoemaker, A. Astromoff, H. Liang, K. Anderson *et al.*, 1999 Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285: 901–906.
- Zuk, O., E. Hechter, S. R. Sunyaev and E. S. Lander, 2012 The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. USA* 109: 1193–1198.