Back to the Future: Multiparent Populations Provide the Key to Unlocking the Genetic Basis of Complex Traits

In the past decade, the ability to generate whole-genome sequences has provided geneticists with a view of the astonishing breadth of genetic variation. This, in theory, means we should be able to identify the specific differences in DNA sequence that lead to an inherited phenotype, including disease states. But this wealth of new information has revealed perhaps the most fundamental challenge for geneticists since Mendel. While we understand that phenotypes are influenced by genetic variation, we do not yet know how to interpret individual genome sequences, and, therefore, we cannot predict which genetic variants are linked to which phenotypes. Indeed, the term "missing heritability" was coined to highlight the fact that in natural populations the genes or genetic elements associated with complex traits explain only a small proportion of the phenotypic variation in these traits.

In stark contrast, controlled crosses of model organisms have generated a wealth of information about the genetic basis of phenotypes. From broad associations of genomic regions with traits, to individual polymorphisms that act by well understood mechanisms, geneticists have been remarkably successful in revealing the impact of genetic variation on phenotype. Applications as diverse as targeted drug therapy and dramatic improvements in agricultural output have been enabled by our understanding of genetics. But it remains a significant challenge to transfer this understanding to natural populations.

To bridge the gap between natural populations and experimental systems, experimental systems need to incorporate more of the complexity of natural populations. This has given rise to a burst of creativity in the design of genetic reference populations. The basic idea is simple: combine the strength of the experimental system, where the genetic composition can be replicated, with the genetic diversity of the target population. Rather than choose two inbred lines or two phenotypically divergent individuals as founders of a genetic reference panel (recombinant inbreds), choose eight, or 25. Using multiple lines as founders of a set of inbred lines whose haplotypes can be replicated has been referred to as Interconnected populations multi-parent, advanced-generation inter-cross design, Complex Cross, and multi-parental RIL. We are choosing to refer to this broad set of genetic reference panels as multi-parent populations (MPP).

Fifteen years ago, the mouse genetics community embraced the challenge of creating strains that would represent the diversity of natural variation in mice, thereby improving the utility of the organism for exploring complex human disease. Eight founder mouse strains were selected, and offspring populations with all eight haplotypes were developed in a funnel mating scheme (Figure 1, Collaborative Cross Consortium 2012). The first set of papers describing these strains was published in *GENETICS* and *G3* in 2012 (http://www.g3journal.org/content/mpp_mouse#cc). Systematic monitoring of progress with the mouse collaborative cross has provided a window into the impact of drift on the genomes (Srivastava *et al.* 2017), a startling insight into the genetic basis of male sterility (Odet *et al.* 2015; Shorter *et al.* 2017), the impact of structural variation (Morgan *et al.* 2017), and a new method for estimating haplotypes and preserving uncertainty (Oreper *et al.* 2017). The resources developed for mouse enable detection of many types of loci, from those associated with SARS (Gralinski *et al.* 2017) and West Nile (Green *et al.* 2017) virus infections to those associated with fertility (Shorter *et al.* 2017) allergens Kelada (2016). Morgan *et al.* (2016) and Dumont *et al.* (2017) also provide insights into genome structure.

This large effort in mouse is matched by ambitious projects on a plethora of organisms. MPPs have been created in plants [*Arabidopsis* (Kover *et al.* 2009), Maize (Yu *et al.* 2008), wheat (Mackay *et al.* 2014), sunflower (Bowers *et al.* 2012), and other crops (Brenton *et al.* 2016; Nice *et al.* 2016)], in animals [*Drosophila* (Mackay *et al.* 2012; King *et al.* 2012)], and in yeast (Cubillos *et al.* 2013). In 2014, we highlighted the diversity of MPPs in *GENETICS* and *G3*

Copyright © 2017 de Koning and McIntyre
doi: https://doi.org/10.1534/genetics.117.203265
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

with articles on maize, sorghum, wheat, triticale, *Arabidopsis*, *Drosophila*, and mouse (http://www.genetics.org/content/multiparental_populations). These issues of *GENETICS* and *G3* feature MPPs of sorghum (Bouchet *et al.* 2017), strawberry (Mangandi *et al.* 2017), rice (Raghavan *et al.* 2017), oil palm (Tisné *et al.* 2017), yeast (Cubillos *et al.* 2017), *Drosophila* (King and Long 2017; Najarro *et al.* 2017; Stanley *et al.* 2017) and mouse (Gralinski *et al.* 2017; Green *et al.* 2017; Morgan *et al.* 2017; Oreper *et al.* 2017; Shorter *et al.* 2017; Srivastava *et al.* 2017; Tyler *et al.* 2017).

GENETICS and *G3* are committed to fostering discussion about the genetic inferences made from MPPs, as well as the best ways to analyze the data, and to extending inferences to natural populations. Projects that rely on a common set of germplasm (or set of strains) rely on data sharing. One of the benefits to working with a reference panel is the ability to leverage data collected in different ways, for different purposes. Our journals have long had policies for reagent and data sharing that reflect the values of our community, and this is evident in these articles on MPPs. Each MPP paper in these issues has the *Data availability* section that is standard for all GSA publications, as well as a one-page guide to the data that makes it easier to browse the data behind the papers.

In recognition of the ongoing importance of MPPs for understanding fundamental questions in genetics, *G3* and *GENETICS* have designed a special web resource for MPPs. Papers are organized in a special collections page, with subheaders that help navigate the growing literature. Our journals have long partnered with model organism databases FlyBase, SGD, WormBase, and others, and we now incorporate news, blogs, tips, and protocols directly on our webpage to help geneticists interested in MPPs get a handle on this topic. Tweet your insights to #MPP #GSAjournals, and use MPP as a keyword of your MPP papers to enable text search engines to collate this literature. The GSA journals are committed to creating a community platform that spans species and disciplines, yet remains focused on common research questions. We thank the authors, referees and editors for making this resource a reality!

Literature Cited

Bouchet, S., M. O. Olatoye, S. R. Marla, R. Perumal, T. Tesso *et al.*, 2017 Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. Genetics 206: 573–585.

Bowers, J. E., E. Bachlava, R. L. Brunick, L. H. Rieseberg, S. J. Knapp *et al.*, 2012 Development of a 10,000 locus genetic map of the sunflower genome based on multiple crosses. G3 2: 721–729.

Brenton, Z. W., E. A. Cooper, M. T. Myers, R. E. Boyles, N. Shakoor *et al.*, 2016 A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. Genetics 204: 21–33.

Cubillos, F. A., L. Parts, F. Salinas, A. Bergström, E. Scovacricchi *et al.*, 2013 High-resolution mapping of complex traits with a four-parent advanced intercross yeast population. Genetics 195: 1141–1155.

Cubillos, F. A., J. Molinet, C. Brice, S. Tisné, V. Abarca, S. M. Tapia *et al.*, 2017 Identification of nitrogen consumption genetic variants in yeast through QTL mapping and Bulk segregant RNA-seq analyses. G3 (Bethesda) 7: 1693–1705.

Dumont, B. L., 2017 Meiotic consequences of genetic divergence across the murine pseudoautosomal region. Genetics 205: 1089–1100.

Gralinski, L. E., V. D. Menachery, A. P. Morgan, A. Totura, A. Beall *et al.*, 2017 Allelic variation in the toll-like receptor adaptor protein ticam2 contributes to SARS-coronavirus pathogenesis in mice. G3 (Bethesda) 7: 1653–1663.

Green, R., C. Wilkins, S. Thomas, A. Sekine, D. M. Hendrick *et al.*, 2017 Oas1b-dependent immune transcriptional profiles of West Nile virus infection in the collaborative cross. G3 (Bethesda) 7: 1665–1682.

King, E. G., and A. D. Long, 2017 The Beavis effect in next-generation mapping panels in *Drosophila melanogaster*. G3 (Bethesda) 7: 1643–1652.

King, E. G., S. J. Macdonald, and A. D. Long, 2012 Properties and power of the *Drosophila* synthetic population resource for the routine dissection of complex traits. Genetics 191: 935–949.

Kover, P. X., W. Valdar, J. Trakalo, N. Scarcelli, I. M. Ehrenreich *et al.*, 2009 A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. PLoS Genet. 5: e1000551.

Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The *Drosophila melanogaster* genetic reference panel. Nature 482: 173–178.

Mangandi, J., S. Verma, L. F. Osorio, N. A. Peres, E. van de Weg *et al.*, 2017 Pedigree-based analysis in a multiparental population of octoploid strawberry reveals QTL alleles conferring resistance to *Phytophthora cactorum*. G3 (Bethesda) 7: 1707–1719.

Morgan, A. P., J. M. Holt, R. C. McMullan, T. A. Bell, A. M.-F. Clayshulte *et al.*, 2016 The evolutionary fates of a large segmental duplication in mouse. Genetics 204: 267–285.

Morgan, A. P., D. M. Gatti, T. M. Keane, R. J. Galante, A. I. Pack *et al.*, 2017 Structural variation shapes the landscape of recombination in mouse. Genetics 206: 603–619.

Najarro, M. A., J. L. Hackett, and S. Macdonald, 2017 Loci contributing to boric acid toxicity in two reference populations of *Drosophila melanogaster*. G3 (Bethesda) 7: 1631–1641.

Nice, L. M., B. J. Steffenson, G. L. Brown-Guedira, E. D. Akhunov, C. Liu *et al.*, 2016 Development and genetic characterization of an advanced backcross-nested association mapping (AB-NAM) population of wild × cultivated Barley. Genetics 203: 1453–1467.

Odet, F., W. Pan, T. A. Bell, S. G. Goodson, A. M. Stevans *et al.*, 2015 The founder strains of the collaborative cross express a complex combination of advantageous and deleterious traits for male reproduction G3 (Bethesda) 5: 2671–2683.

Oreper, D. G., Y. Cai, L. M. Tarantino, F. Pardo-Manuel de Villena, and W. Valdar, 2017 Inbred strain variant database (ISVDB): a repository for probabilistically informed sequence differences among the collaborative cross strains and their founders. G3 (Bethesda) 7: 1623–1630.

Raghavan, C., R. P. Mauleon, V. L. Apostol, M. L. S. Jubay, H. Zaw, J. B. Bonifacio *et al.*, 2017 Approaches in characterizing genetic structure and mapping in a rice multi-parental population. G3 7. 1721–1730.

Shorter, J. R., F. Odet, D. L. Aylor, W. Pan, C.-Y. Kao *et al.*, 2017 Male infertility is responsible for nearly half of the extinction observed in the collaborative cross. Genetics 206: 557–572.

Srivastava, A., A. P. Morgan, M. Najarian, V. K. Sarsani, J. S. Sigmon *et al.*, 2017 The genomes of the collaborative cross. Genetics 206: 537–556.

Stanley, P. D., E. Ng'oma, S. O'Day, and E. G. King, 2017 Genetic dissection of nutrition-induced plasticity in insulin/insulin-like growth factor signaling and median lifespan in a *Drosophila* multiparent population. Genetics 206: 587–602.

Tisné, S., V. Pomiès, V. Riou, I. Syahputra, B. Cochard *et al.*, 2017 Identification of Ganoderma disease resistance loci using natural field infection of an oil palm multiparental population. G3 (Bethesda) 7: 1683–1692.

Tyler, A. L., B. Ji, D. M. Gatti, S. C. Munger, G. A. Churchill *et al.*, 2017 Epistatic networks jointly influence phenotypes related to metabolic disease and gene expression in diversity outbred mice. Genetics 206: 621–639.

Yu, J., J. B. Holland, M. D. McMullen, and E. S. Buckler, 2008 Genetic design and statistical power of nested association mapping in maize. Genetics 178: 539–551.

Dirk-Jan de Koning, Deputy Editor, G3: Genes | Genomes | Genetics
Swedish University of Agricultural Sciences

Lauren M. McIntyre, Senior Editor, GSA Journals
University of Florida

@DJ_de_Koning
@LaurenMMcIntyr1