Published in final edited form as:

Science. 2010 April 23; 328(5977): 469. doi:10.1126/science.1189015.

Genotype to Phenotype: A Complex Problem

Robin D. Dowell[1,*,†], **Owen Ryan**[2,3,4,†], **An Jansen**[5], **Doris Cheung**[2,3,4], **Sudeep Agarwala**[5], **Timothy Danford**[1], **Douglas A. Bernstein**[5], **P. Alexander Rolfe**[1], **Lawrence E. Heisler**[2,3,4], **Brian Chin**[5], **Corey Nislow**[2,3,4], **Guri Giaever**[3,4,6], **Patrick C. Phillips**[7], **Gerald R. Fink**[5,8], **David K. Gifford**[1,5,8,‡], and **Charles Boone**[2,3,4,‡]

[1]Computer Science and Artificial Intelligence Laboratory, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA

[2]Banting and Best Department of Medical Research, University of Toronto, Toronto, ON M5S 3E1, Canada

[3]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 3E1, Canada

[4]Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON M5S 3E1, Canada

[5]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

[6]Department of Pharmacy University of Toronto, Toronto, ON M5S 3E1, Canada

[7]Center for Ecology and Evolutionary Biology, 5289 University of Oregon, Eugene, OR 97403, USA

[8]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA

Rapid genome-sequencing methods coupled with whole-genome transcription profiling suggests that it may be possible to predict phenotype from a genotype. Human genetic association studies of common single-nucleotide polymorphisms (SNPs) explain only a fraction of phenotypic variation among individuals (1). This may be due to rare SNPs (1), structural (2) and epigenetic variants, or multiple alleles with additive effects or synergistic genetic interactions associated with complex combinations of genetic variation (3).

To address the genotype-to-phenotype problem, we developed a simple comparative model for the budding yeast *Saccharomyces cerevisiae* that enables a comprehensive assessment of the genetic mechanisms leading to different phenotypes for the same mutation in two different genetic backgrounds. The strain Σ1278b mates and forms viable meiotic progeny

‡To whom correspondence should be addressed. E-mail: charlie.boone@utoronto.ca (C.B.); gifford@mit.edu (D.K.G.).
*Present address: Molecular, Cellular, and Developmental Biology, 347 UCB, University of Colorado, Boulder, CO 80309, USA.
†These authors contributed equally to this work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5977/469/DC1
Figs. S1 and S2
Tables S1 to S3
References

with the reference strain, S288c, and the divergence between the two strains is roughly equal to the divergence between the genomes of two humans (4, 5).

We sequenced and assembled the 12-Mb Σ1278b genome, annotating 6923 open reading frames (ORFs) and RNAs, of which 6848 have orthologs within S288c (5). The order of genes between the strains was the same (except in the highly variable subtelomeric regions), and the sequence of 46% of the Σ1278b ORFs was identical to those in S288c. Differences between the strains were largely due to small insertions and deletions or SNPs, with an average SNP density of 3.2 per kilobase.

We deleted ~5100 genes within Σ1278b (5) to systematically compare identical deletion mutants (6). In particular, we identified "conditional essentials," those genes required uniquely for viability in either strain (fig. S3 and table S1). We scored colonies as dead or alive and surveyed all vital pathways for individual-specific genetic interactions. We expected such conditional essential genes to be rare because the genomes of Σ1278b and S288c are nearly identical.

Although 894 genes were essential in both S288c and Σ1278b, 44 genes were essential only in Σ1278b and 13 genes were essential only in S288c (Fig. 1A). The conditional subsets included genes of various functions; however, the Σ1278b subset was enriched for genes involved in mRNA metabolic process, whereas the S288c set was enriched for genes annotated to SRP-dependent co-translational targeting. These biological biases suggest that these phenotypes result from genetic interactions associated with an individual genotype.

Hybrid strain crosses and tetrad analysis focusing on 18 mutants that were lethal in Σ1278b with wild-type levels of fitness when deleted in S288c were used to investigate conditional essentiality. We mated viable haploid S288c deletion mutants to wild-type Σ1278b and analyzed the hybrid diploid progeny by tetrad analysis. The number of viable meiotic progeny carrying the deletion allele is related to the number of unlinked background-specific modifiers (5) that contribute to the genetic interaction. In all 18 cases, the conditional phenotype was associated with numerous modifier genes that differ between strains. The simplest cases, *SKI7* and *BEM1*, are likely due to a genetic interaction with at least two or more modifiers, but all other cases were more complex (Fig. 1B). Thus, our analysis showed that conditional essentiality is almost always a consequence of complex genetic interactions involving multiple modifiers associated with strain-specific genetic variation rather than classic digenic synthetic lethality (6, 7).

Our genome-wide survey of conditionally essential genes demonstrates that in most cases a complex set of background-specific modifiers influence a mutation whose phenotype differs between individuals. These results raise the possibility that similar complex modifiers may largely explain the difficulty in identifying the genetic basis for individual phenotypes. The potential for genetic interactions to control individual phenotypes becomes even more important if different combinations of alleles can lead to the same physiological state. The ability to identify these conditional essential phenotypes in yeast provides a framework to unravel the fundamental principles of genetic networks resulting from natural variation, including those that underlie human disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by NIH grants DK076284 (R.D.D.), GM035010 (G.R.F.), GM069676 (D.K.G.), P01 NS055923 (D.K.G.), ACS PF-09-072-01-MBC (D.A.B.); NSF grant DEB-0614588; a Senior Scholar Award in Aging from the Ellison Medical Foundation (P.C.P.); the Natural Sciences and Engineering Research Council (NSERC) of Canada; and Howard Hughes Medical Institute (HHMI) (C.B.). Sequence and assembly data has been deposited at DNA Data Bank of Japan/European Molecular Biology Laboratory/GenBank, project accession ACVY00000000.

References and Notes

1. Dickson SP, et al. PLoS Biol. 2010; 8:e1000294. [PubMed: 20126254]
2. Korbel JO, et al. Science. 2007; 318:420. published online 27 September 2007. 10.1126/science. 1149504 [PubMed: 17901297]
3. Hartman JL 4th, Garvik B, Hartwell L. Science. 2001; 291:1001. [PubMed: 11232561]
4. Wang J, et al. Nature. 2008; 456:60. [PubMed: 18987735]
5. Materials and methods are available as supporting information on *Science* Online.
6. Giaever G, et al. Nature. 2002; 418:387. [PubMed: 12140549]
7. Costanzo M, et al. Science. 2010; 327:425. [PubMed: 20093466]

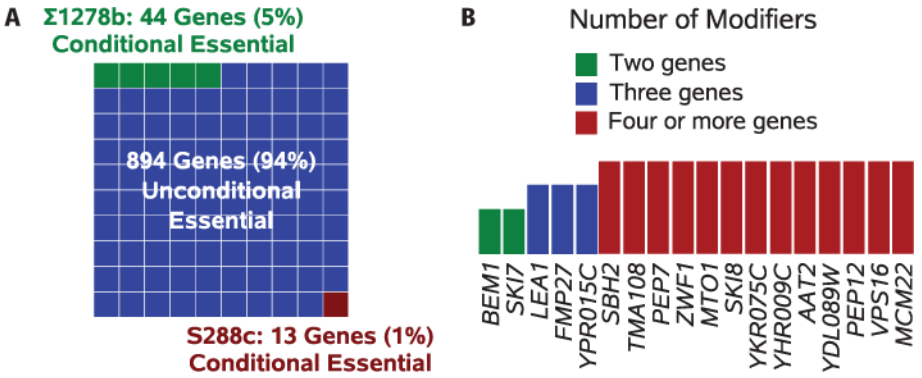

Fig. 1.

(**A**) Most S288c essential genes are also essential in Σ1278b (94%); however, ~5% are essential only in the Σ1278b genetic background, whereas ~1% are essential only in S288c. (**B**) Conditional essential genes in Σ1278b are the consequence of complex genetics. χ^2 tests indicated the number of modifiers associated with conditional essentiality (5).