A better way to discover gene function in the social amoeba *Dictyostelium discoideum*

William F. Loomis[†]

Division of Biology, University of California San Diego, La Jolla, California 92093, USA

For many years, the social amoeba *Dictyostelium discoideum* has served as a system to understand what it takes to generate a multicellular organism. Biochemistry and biophysics have defined many of the underlying interactions of protein and cytoskeletal complexes, and sequence analyses have quantitatively defined the changes in gene expression programs. However, the dearth of genetic studies that offer insight into the roles of specific genes has prevented the more widespread use of this fascinating model organism. The eukaryotic amoebas of *Dictyostelium* provide all the advantages of growth as a microorganism and development as a multicellular organism. As a consequence, they can be analyzed using the techniques of microbial genetics that rely on having millions, even billions, of cells from which to select rare mutant strains; at the same time, they present a myriad of multicellular phenotypes from which to pick. The fact that *Dictyostelium* can grow and develop just as well as a haploid or a diploid makes the initial screening for mutants fast and simple and allows straightforward mapping by complementation. Dominance and recessivity can be easily determined in diploids formed with wild-type strains.

These characteristics were exploited from 1950 to 1990, initially by Maurice Sussman at Brandeis University in Massachusetts, and then by John Asworth, Bill Loomis, Peter Newell, Keith Williams, Kai Yanagisawa, and their students all over the world, who collectively formed the *Dictyostelium* community (Newell 1978; Loomis 1987). *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (NTG) was soon found to be the mutagen of choice that allowed a large number of morphological and temperature-sensitive mutants to be isolated. Loci were recognized by failure of alleles to complement in diploids and then mapped to specific chromosomes by parasexual genetics. The developmental genes could be formally arranged in dependent hierarchies, but there was no way to know what the genes were doing in detail.

When the techniques of molecular genetics became universally available in the 1990s, genes could be cloned, modified in vitro, and returned to their original location in the chromosomes using the highly efficient homologous recombination provided by *Dictyostelium*. Sometimes homologs of genes that had been characterized in other organisms were cloned from *Dictyostelium* and used to inactivate the endogenous genes by reverse genetics. The resulting phenotypes shed light on the functions in *Dictyostelium*. However, although it was possible to go from gene to phenotype, there was no convenient way to go from phenotype to gene.

Finally, in 1992, along came a technique that allowed mutated genes to be recovered and sequenced—Restriction Enzyme–Mediated Integration (REMI) (Kuspa and Loomis 1992). REMI mutagenesis works by introducing restriction enzymes into cells along with DNA from a plasmid carrying a selectable drug resistance gene. The restriction enzyme increases the rate of insertion of the plasmid into cognate restriction sites in the genome so that random mutants can be easily recovered. The integrated plasmids can then be recovered along with flanking sequences that define the mutated gene. Hundreds of REMI mutants with morphological defects have been selected and analyzed in the last 25 years that probably include most of the developmental genes. Molecular genetics allows for targeted mutagenesis as well as ectopic expression of each of these genes, which has shed light on subtle aspects of the underlying circuitry. Clearly, *Dictyostelium* has proven to be the wonderful genetic system that was hoped for 50 years ago, but still, there was a tool missing for the full exploitation of the genetic approach to understanding multicellular morphogenesis. That is, not all developmental genes can be discovered in REMI knockout mutants because some are expected to be vital for growth as well as development. In this issue of *Genome Research*, Gad Shaulsky and his colleagues describe a next-generation technique that fills this gap (Li et al. 2016).

The advantages of chemical mutagenesis

Li et al. (2016) show that the critical base changes in *Dictyostelium discoideum* mutants generated by chemical mutagenesis can be pinpointed by whole-genome sequencing as long as a sufficient number of mutants with the same phenotype are analyzed, an approach that has been used in different model species (Smith et al. 2008; Irvine et al. 2009; Haelterman et al. 2014; Schneeberger 2014). Although random mutagenesis with NTG will affect many genes that have nothing to do with the selected trait, any gene directly responsible for the phenotype will be found to have been mutated in multiple independent mutant strains. When a significant number of mutant strains carry base changes in the same gene, that gene is likely to be responsible for the trait. However, calculating significance might not be straightforward and requires full knowledge of mutation preferences of the mutagen, because false positive genes with more mutations than expected are observed as well. The mutations do not have to be exactly the same as long as they are all in the same gene. The individual base changes can be inspected to predict whether the mutations will result in loss-of-function, partial loss-of-function, or gain-of-function.

The power of forward genetics is that alleles that generate the chosen phenotype are selected regardless of whether the gene is

essential or pleiotropic. REMI mutagenesis can tag genes with specific roles, but it usually generates null mutations because insertion of the plasmid into the coding region of a gene results in a truncated protein product, which is likely to be rapidly degraded. Partial loss-of-function and gain-of-function mutations could illuminate whole new sets of genes involved in interesting developmental function, but plasmid insertion is unlikely to generate such mutations. REMI mutagenesis can only generate a suppressor of a developmental block if there is a dispensable gene that acts in a negative fashion downstream from the disrupted process. Null mutations that knock out this gene will suppress the original phenotype. However, loss-of-function mutations in a gene that acts in a positive fashion downstream from the disrupted process will never be detected. On the other hand, the technique described by Li et al. (2016) can result in gain-of-function mutations in positively acting genes that suppress the original phenotype. The random mutations resulting from treatment with NTG can change amino acid codons, generate new start or stop signals, or alter regulation. When a gene is found to be mutated in multiple isolates from a genetic selection, each mutation will shed light on the biological process.

Decades ago, and soon after it was recognized that NTG was a wonderful mutagen, mutants and their suppressors were isolated in a wide range of species. However, in many instances, the exact base changes responsible for the mutant phenotype remained a mystery. In the last 10 years, next-generation whole-genome sequencing techniques promised to discover what was mutated in a rapid and affordable manner but was held up by encountering many different genes with base changes in the same strain. There was no simple way to know which base changes were important and which were just genetic noise to be expected in genomes with more than 5000 genes. One way around this was to genetically map the mutated gene based on the recombination frequency with chromosomal markers. In most cases, this approach was way too much work and expense. In the few instances in which it was possible to map the mutation within a centimorgan (1% recombination with a marker), only the base changes within 1 Mb of the marker had to be considered. Such fine mapping reduces the number of confusing base changes at least 10-fold but still leaves some uncertainty. Ultimate proof of causation still required complementation with a known allele or recreation of the allele using reverse genetic approaches followed by recapitulation of the phenotype.

The approach described by Li et al. (2016) obviates genetic mapping by exploring a sufficiently large mutagenized population. When a gene is found to be mutated in several independently isolated mutant strains, it becomes a candidate for the function under scrutiny. The functional role can be confirmed by reverse genetic modification of the gene. This simple and straightforward technique can be applied to any organism of interest but is easier with haploid microorganisms, in which large populations can be efficiently screened to generate a sufficiently large number of candidate isolates. For instance, other social amoebas such as *Dictyostelium purpureum* or *Polysphondylium pallidum*, in which the genome sequences are available, could be analyzed irrespective of their ability to grow axenically in defined medium. NTG mutagenesis works just as well with cells feeding on bacteria as it does with cells grown axenically. Multiplexed DNA from dozens of isolates can be rapidly and affordably sequenced using next-generation techniques. Using short read alignments against the 34-Mb reference sequence of the organism, significant base changes have to be flagged and inspected in the context of the affected

gene. This process has been automated for *Dictyostelium*, and the code is available online (Li et al. 2016).

Developmental signal transduction pathways uncovered by genetic suppressors

To demonstrate the effectiveness of the technique, Li et al. (2016) sequenced and compared the genomes of 63 suppressor strains that had been selected to bypass a block to mound formation. Development was arrested in the parental strain because the TgrB1 and TgrC1 signaling proteins were unable to form effective heterodimers with each other. When adjacent cells have compatible TgrB1 and TgrC1, the membrane proteins form a heterodimer that stimulates transcriptional responses that are necessary for completion of morphogenesis. Suppressor strains were isolated that proceeded beyond the mound stage and formed spores as well as normal fruiting bodies. Not surprisingly, many mutations were found to affect the *tgrB1* gene to make it compatible with the *tgrC1* allele. However, it was a welcome surprise to find that a high proportion of the mutations (19/63) fell in the *rapgapB* gene, which regulates the activity of the small GTPase Rap1. Other genes discovered in this screen are further defining the signal transduction pathway that mediates TgrB1/C1 signaling.

This approach to discovering gene functions works best when mutations in a relatively small number of genes result in the selected phenotypes (Shaulsky et al. 1996). That way, the number of independent mutants that have to be analyzed is kept small. For example, selecting for "stalky" mutants that develop well but make long thin stalks and very few spores might discover genes in addition to the one encoding a GATA-type transcription factor that was found in a rare REMI mutant (Chang et al. 1996). New "stalky" genes might turn out to be regulatory targets of the GATA transcription factor or might define an independent pathway to the phenotype. Suppressor mutations that restore sporulation in the stalky strain might define the signal transduction pathway downstream from the GATA transcription factor.

Likewise, a screen for rapidly developing mutants has already discovered three very interesting genes the old-fashioned way (Abe and Yanagisawa 1983; Loomis 2014, 2015) and might uncover surprising new genes now that whole-genome sequencing is possible. NTG mutagenesis might perturb critical genes such that sporulation occurs precociously. New genetic regulatory networks might be uncovered by this approach.

Discovering gene functions in yeast, flies, worms, plants, and fish

In the last 10 years, whole-genome sequencing in a variety of organisms has uncovered their full coding capacity and increased the resolution for evolutionary inference of gene function. Sequencing combined with computational techniques have been able to define most genes in bacteria, plants, amoebozoa, yeast, flies, worms, fish, birds, and mammals. However, the challenge remains to determine what each gene can do and how its product integrates into the complex cellular circuitry. Analyses of mutant strains can provide this functional information.

Every one of the approximately 6000 genes in the budding yeast *Saccharomyces cerevisiae* has been knocked out, and the resulting phenotype carefully characterized (Boone 2014). Nevertheless, new assays may uncover new phenotypes, particularly when

genetic mutations are combined systematically (Costanzo et al. 2010).

Many of the early gene hunts were carried out in flies, worms, or plants. Genes can be discovered in the fly *Drosophila melanogaster* by chemical mutagenesis or P-element mutagenesis. The latter is a strategy much like REMI in *Dictyostelium*. Many of the genes of *Drosophila* have been disrupted, one at a time, and the phenotypes of allelic series used to define the function (dos Santos et al. 2015). Mutations in many genes give pleiotropic phenotypes that can indicate the molecular function. Moreover, as in yeast, combining known mutations in individual flies has uncovered surprising interactions (Dobzhansky 1946). It is very likely that unbiased forward genetic screens followed by sequencing will continue to define gene functions.

The round worm *Caenorhabditis elegans* is a self-fertilizing hermaphrodite that rapidly gives rise to homozygous offspring. Phenotypes of recessive mutations can be easily recognized in the transparent embryos. Screens for mutations affecting the patterns of cell division and cell fate during early development uncovered hundreds of genes and allowed them to be mapped. Many developmental genes were cloned and characterized by conventional mapping approaches. Whole-genome sequencing of a collection of several thousand mutagenized strains has recently recognized mutant alleles in each of the 20,000 genes (Thompson et al. 2013). Genetic studies on these strains should be able to sort out mutant genes that generate specific phenotypes.

During the second half of the last century, the self-fertilizing, diploid plant *Arabidopsis thaliana* became a major model organism for plant biology. Radiation and later chemical mutagen screens revealed a wide range of physiological and morphological phenotypes. Similar to worms, many genes underlying these traits were cloned by classical mapping efforts. Whole-genome sequencing of mutant plants is challenged by the typically large number of mutated genes within a single mutant genome, which hampers immediate identification of the causal mutation. Nevertheless, the first report on identification of a gene by direct sequencing of mutants of an allelic series was performed in plants (Schneeberger 2014), and even sequencing of entire mutant populations of crop plants with large genomes like rice or wheat is now possible.

The zebrafish *Danio rerio* is a more recently exploited model vertebrate in which mutational genetics has been extensively used to study embryogenesis (Nusslein-Volhard 2012). It is also a wonderful model system for the study of vertebrate physiology, pigmentation, regeneration, aging, and many other biological processes. However, the genome is 50 times larger than the *Dictyostelium* genome, which increases the difficulty and expense of NTG/comparative genome techniques.

Discovering gene functions in humans

Whole-genome sequencing (WGS) or whole-exome sequencing (WES) of patients and their parents have been successful at identifying the causative genes of numerous rare diseases (Lupski et al. 2010). This is conceptually different from the approach of Li et al. (2016) because they had to mutagenize their populations of *Dictyostelium* in order to recover a sufficient number of mutants with the same phenotypes, whereas human patients who show up at a clinic anywhere in the world can be assigned to a medical specialist for diagnosis. However, with a population of 7 billion individuals, many spontaneous mutations occur every generation. If some of those with shared symptoms are found to carry mutations

in the same gene, it is also likely that the product of that gene is involved in the specific process that causes the disease.

The NTG/comparative genome techniques will certainly be given a warm welcome by those who work with *Dictyostelium*. Mutants can give novel insights into cellular functions that would not necessarily come from simple linear thinking about physiology or signaling, and Li et al. (2016) give us a new method to come by such insights.

Acknowledgments

References

Abe K, Yanagisawa K. 1983. A new class of rapidly developing mutants in *Dictyostelium discoideum*: implications for cyclic AMP metabolism and cell differentiation. *Dev Biol* **95:** 200–210.

Boone C. 2014. Yeast systems biology: our best shot at modeling a cell. *Genetics* **198:** 435–437.

Chang WT, Newell PC, Gross JD. 1996. Identification of the cell fate gene *stalky* in Dictyostelium. *Cell* **87:** 471–481.

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, et al. 2010. The genetic landscape of a cell. *Science* **327:** 425–431.

Dobzhansky T. 1946. Genetics of natural populations. Xiii. Recombination and variability in populations of *Drosophila pseudoobscura*. *Genetics* **31:** 269–290.

dos Santos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond J, Emmert DB, Gelbart WM; FlyBase Consortium. 2015. FlyBase: introduction of the *Drosophila melanogaster* Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Res* **43:** D690–D697.

Haelterman NA, Jiang L, Li Y, Bayat V, Sandoval H, Ugur B, Tan KL, Zhang K, Bei D, Xiong B, et al. 2014. Large-scale identification of chemically induced mutations in *Drosophila melanogaster*. *Genome Res* **24:** 1707–1718.

Irvine DV, Goto DB, Vaughn MW, Nakaseko Y, McCombie WR, Yanagida M, Martienssen R. 2009. Mapping epigenetic mutations in fission yeast using whole-genome next-generation sequencing. *Genome Res* **19:** 1077–1083.

Kuspa A, Loomis WF. 1992. Tagging developmental genes in *Dictyostelium* by restriction enzyme-mediated integration of plasmid DNA. *Proc Natl Acad Sci* **89:** 8803–8807.

Li CL, Santhanam B, Webb AN, Zupan B, Shaulsky G. 2016. Gene discovery by chemical mutagenesis and whole-genome sequencing in *Dictyostelium*. *Genome Res* (this issue) **26:** 1268–1276.

Loomis WF. 1987. Genetic tools for *Dictyostelium discoideum*. *Meth Cell Biol* **28:** 31–65.

Loomis WF. 2014. Cell signaling during development of *Dictyostelium*. *Dev Biol* **391:** 1–16.

Loomis WF. 2015. Genetic control of morphogenesis in *Dictyostelium*. *Dev Biol* **402:** 146–161.

Lupski JR, Reid J, Gonzaga-Jauregui C, Rio Deiros D, Chen D, et al. 2010. Whole-genome sequencing in a patient with Charcot–Marie–Tooth neuropathy. *N Engl J Med* **362:** 1181–1191.

Newell PC. 1978. Genetics of the cellular slime molds. *Annu Rev Genet* **12:** 69–93.

Nusslein-Volhard C. 2012. The zebrafish issue of development. *Development* **139:** 4099–4103.

Schneeberger K. 2014. Using next-generation sequencing to isolate mutant genes from forward genetic screens. *Nat Rev Genet* **15:** 662–676.

Shaulsky G, Escalante R, Loomis WF. 1996. Developmental signal transduction pathways uncovered by genetic suppressors. *Proc Natl Acad Sci* **93:** 15260–15265.

Smith DR, Quinlan AR, Peckham HE, Makowsky K, Tao W, Woolf B, Shen L, Donahue WF, Tusneem N, Stromberg MP, et al. 2008. Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res* **18:** 1638–1642.

Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, Adair R, Au V, Chaudhry I, Fernando L, Hutter H, et al. 2013. The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res* **23:** 1749–1762.

Editor's note

The Editors are grateful to the reviewers of this Perspective who exceeded reviewer responsibilities in helping to revise the final version. We also thank Dr. Adam Kuspa, Baylor College of Medicine, for his words of remembrance.

William F. Loomis 1940–2016

The author, Distinguished Professor Emeritus of the University of California San Diego (UCSD) Division of Biology, passed away on June 30, 2016, while finalizing this article. After training at Harvard, MIT, and Brandeis, Bill Loomis served for 50 years as a faculty member at UCSD. Bill made major contributions to science by elucidating developmental mechanisms of the social amoeba *Dictyostelium discoideum*, and he also developed genetic and genomic methods that benefited the field enormously. Bill published more than 240 articles, books, and book chapters, was an American Cancer Society Scholar, served as president for the Society for Developmental Biology, and was an elected fellow of the American Association for the Advancement of Science. He was a passionate and erudite scientist, his enthusiasm and love of knowledge were infectious, and his devotion to rigorous biological research and education was unwavering. Bill was a teacher, a mentor, and a friend to many in the field and will be dearly missed.