Probing the Depths of Biological Diversity During the Second Century of *GENETICS*

Linnea Sandell and Sarah P. Otto[1]
Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada V6T1Z4

KEYWORDS diversity; evolution; genetics

After a century of *GENETICS*, we understand better than ever the diversity of life and its immense evolutionary history. Nevertheless, we are still gazing at the tip of the iceberg of biological complexity. For virtually every biological rule, an exception lies in some organism on some branch of the tree of life. Meiosis is fair. Mating is random. Chromosomes govern inheritance. The genetic code is universal. All have exceptions (meiotic drive: Buckler *et al.* 1999 and Didion *et al.* 2016; mating: Jiang *et al.* 2013; inheritance: Fang *et al.* 2012 and Houri-Ze'evi *et al.* 2016; genetic code: Saccone *et al.* 2000). In contemplating what is in store for the journal *GENETICS* in its second century, we argue that our vision of genetics will move increasingly away from trying to understand the general pattern of biology toward grappling with its variability and, in so doing, better reveal the depths of biological complexity.

Biological complexity is often discussed as the product of an evolutionary history spanning the ~4.1 billion years since the origin of life (Bell *et al.* 2015), but even this number is misleadingly small. Evolution is not linear: it branches into species, which explore, in parallel, different ways of surviving and reproducing. This exploration spans more than 10^{14} years of evolutionary discovery (Figure 1), a staggering number. By comparison, there are an estimated 10^{14} letters in total in all of the published books across human history (Urban 2014). It is no wonder that life is so variable.

Contemplating this breadth of evolutionary history is essential if one is to understand the richness of biology. Historically, we have done the reverse. We have stripped out the complexities, alternative contexts, and species interactions. This is a natural and necessary thing to do when first seeking out the impact of what a gene does or how a population evolves, but a full understanding of biology requires that we expand our viewpoint and consider alternative contexts and understudied organisms from across the tree of life. Why do genes not always perform in the same way? Why do populations not always evolve in the same direction? Although biologists, both theoreticians and empiricists, have been moving in this direction for decades, we argue that in the next century, our focus will shift from a search for general rules to a greater appreciation of biological variability. We describe ways in which we expect this shift to impact genetics and evolution.

Probing the Depths of Genetics

Over the past century, geneticists have identified possible functions of most genes in *Drosophila melanogaster*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana*, and in our own species, *Homo sapiens*—a tremendous feat. Furthermore, the commitment to sharing this knowledge through online databases has greatly enabled syntheses and has proven the value of community cooperation for scientific progress (*e.g.*, flybase.org, yeastgenome.org, arabidopsis.org, omim.org, and encodeproject.org). These compilations have significantly spurred scientific discovery, as evidenced by the numerous citations mentioning the "*Saccharomyces* Genome Database" (over 10,000 on Google Scholar) and "Online Mendelian Inheritance in Man" (over 45,200 on Google Scholar).

While genetic data are increasingly available, there are several reasons to believe that much remains unknown. Take for example, gene function. Genes are typically named for their inferred function, often based on mutant screens in a homogenous or inbred lineage. Arguably the best understood eukaryote genetically, *S. cerevisiae* has 5820 open reading frames (ORFs) (not considering dubious ORFs). The vast

doi: 10.1534/genetics.116.187625
[1]Corresponding author: Department of Zoology, University of British Columbia, Vancouver, BC V6T1Z4, Canada. E-mail: otto@zoology.ubc.ca

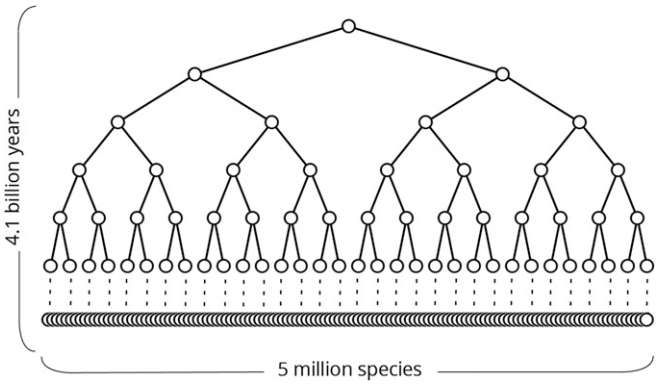

Figure 1 Summing the total evolutionary history of life. The tree of life is not completely resolved, so the sum total evolution that it spans is unknown. A simple birth model would start with a single species 4.1 billion years ago, undergoing $x = 22.25$ species doublings, to produce the ~5 million species alive today (Costello *et al.* 2013). With the length of each branch spanning $\tau = 4.1 \times 10^9/x$ billion years, the sum total amount of evolutionary time captured by the tree of life would be 1.5×10^{15} years $\left(\sum_{i=0}^{x} 2^x \tau\right)$. Account for extinction (using equation 29b in Nee *et al.* 1994) and assuming that extinction occurs at 90% the rate of speciation (*e.g.*, Rabosky *et al.* 2013), the span of evolutionary branches would total 4.0×10^{14} years.

majority of these ORFs (88.5%) have now been characterized to some extent (Figure 2). Yet our tendency to name and discuss genes as if they had one function is often unfounded. For example, even considering just coarse Gene Ontology (GO)-Slim categories (Figure 2), 35% of genes have more than one molecular function, 58% are involved in more than one biological process, and much remains to be discovered. The map from gene to function is, in many cases, multipronged and pleiotropic.

As another example, consider protein–protein interaction networks. The network for yeast was initially inferred in two different studies, which exhibited little overlap in the list of interactions (Uetz *et al.* 2000; Ito *et al.* 2001), indicating that many interactions remained undetected. Accounting statistically for the overlap, Grigoriev (2003) concluded that the average yeast protein was involved in five different protein–protein interactions. Consequently, for many genes, the most common of its interactions, as well as the best studied of its functions, may only be scratching the surface; the effects of the gene may be substantially broader. Similarly, in humans, Venkatesan *et al.* (2009) predicted ~130,000 protein–protein interactions, the vast majority of which remain undocumented. The fact that some genes, and their interactions, are better studied than others can generate biased conclusions about the biological connectedness of the genes that we study, for example, when comparing cancer-causing *vs.* control genes (Schaefer *et al.* 2015). Thus, in addition to being multipronged, the map from gene to function is webbed and contingent upon the state of many interacting genes (epistasis).

Even this webbed image of interacting genes is misleadingly static, given that gene expression varies among tissues, between the sexes, and over time (Kwekel *et al.* 2013; Gutierrez-Arcelus *et al.* 2015). The map from gene to function is dynamic and responsive to environmental (Scheiner and Lyman 1989; Lacaze *et al.* 2009) and social cues (Evans and Wheeler 2001; Toth *et al.* 2007).

The lack of full information about the various roles of a gene hampers our ability to predict its potential effects. For example, in stickleback fish, the signaling gene *Ectodysplasin* (*Eda*) is well known for its impact on armored plates, thought to be an important defense against predators (Colosimo *et al.* 2005). Yet, when a polymorphic population was tracked over time, substantial selection was found on plated and nonplated variants of *Eda* among juvenile fish, before the development of armor (Barrett *et al.* 2008). Exactly what role(s) the protein EDA plays in these juvenile fish is unknown, but there are many possibilities given that *Eda* is in a gene family known to impact the development of various ectoderm-derived body parts in mammals (*e.g.*, teeth, hair, and sweat glands; Colosimo *et al.* 2005).

Over the past century, we have attached an ever-growing stack of exceptions and anomalies into a core narrative of simple Mendelian inheritance and a one-gene, one-enzyme, one-function paradigm. It is time to recognize the exceptions for what they are: the outcome of an evolutionary process that need not stick to simple rules. A Gestalt switch, already under way, is needed that grapples with the tangled and dynamic web of genes and their interactions. As we move forward over the next century to understand the immensity of what remains unknown, open repositories of data, synthetic analyses, and probabilistic models will become increasingly important means to capture the diversity of genetic narratives. Analyses such as those mentioned above that compare the overlap among studies (Grigoriev 2003) or that quantify the likelihood of missing data based on how much different genes have been studied (Schaefer *et al.* 2015) are important ways of identifying gaps. Furthermore, predictive models that combine knowledge of gene functions across species with models of trait evolution across phylogenies would allow probabilistic statements to be made about the likelihood that a gene plays a given role in a given species (*e.g.*, Engelhardt *et al.* 2005).

Probing the Depths of Evolution

The last century has also seen a flourishing of evolutionary knowledge, with the development of a strong mathematical foundation and the establishment of a strong experimental and comparative toolkit. Yet evolution is often portrayed in a simple, hill-climbing fashion, according to the core premises of evolution by natural selection (Lewontin 1974): given that individuals vary in a trait, that some variants are more likely to survive and reproduce, and that variants can be inherited from parents to offspring, the offspring generation will differ from the parental generation. Evolution will have occurred.

In the simplest of models (in asexual populations, in sexual populations with one gene, or in sexual populations with an

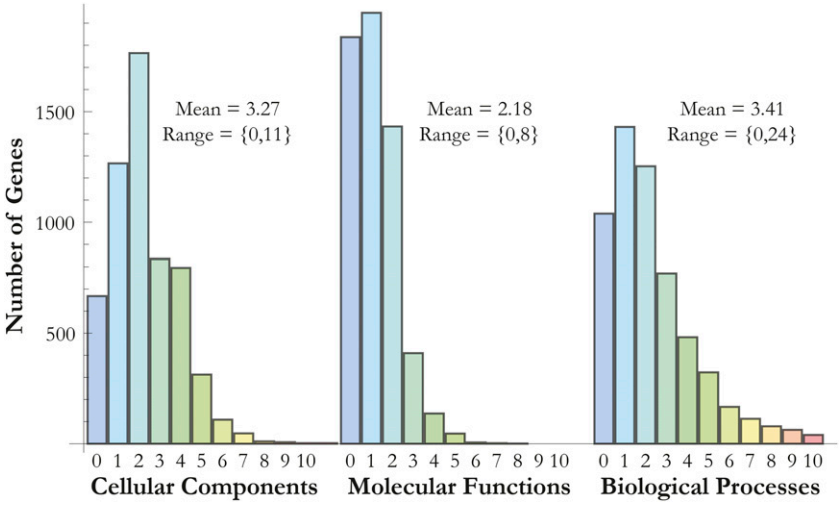

Figure 2 Most yeast genes have been well characterized with some information about the associated cellular component (88.52%), molecular function (68.44%), and biological process (82.15%). The figure charts the number of GO terms associated with each gene, focusing on "GO-Slim" terms, a broad-level categorization in the ontology (*e.g.*, "vitamin metabolic process" or "transmembrane transport"). Data are from the *Saccharomyces* Genome Database. Available at: http://downloads.yeastgenome.org/curation/literature/go_slim_mapping.tab.

infinite number of small-effect genes, all within a constant environment), evolution follows the trajectory that maximizes fitness: populations become successively better adapted to their current environment. By focusing on evolution as a hill-climbing mechanism, however, it becomes difficult to fathom why the biological world is so diverse. Surely, one might think, there would be an optimal genome structure or mating system, a hill that evolution climbed long ago. But this singular view belies reality, and much of evolutionary biology today is grappling with the proliferation of evolutionary trajectories. Why do some organisms reproduce sexually, others asexually (Simon *et al.* 2003; Whitton *et al.* 2008)? Why do some organisms separate male and female functions into different individuals, while others combine the sexes (Figure 3) (Tree of Sex Consortium 2014a)? Why do some organisms develop directly, while others proceed through larval stages (Gomez-Mestre *et al.* 2012)? Why do some organisms have small genomes, others large (Lynch 2007; Elliott and Gregory 2015)?

Consider, as a special case, the diversity of ploidy levels that underlies several of the quantum leaps in genome size taken by evolution. Dramatically, some organisms have recently experienced the addition of an entire genomic copy (polyploidization), including some animals and numerous plants; polyploidy is particularly common in domesticated crops such as coffee, sugar, wheat, and cotton (Otto and Whitton 2000). Another facet of diversity in ploidy arises from the alternation between haploid and diploid phases in sexually reproducing eukaryotes. Some organisms, including many fungi, algae, and bryophytes, spend the majority of their lives in the haploid phase; others, for instance most vascular plants and animals, in the diploid phase. However, this summary barely scratches the surface of ploidy diversity. Some species alternate between free-living haploid and diploid phases (*e.g.*, red algae). Some species have haploid males and diploid females (*e.g.*, Hymenoptera). Some species start their lives at one ploidy level and end at another (*e.g.*, paternal genome elimination; Gardner and Ross 2014). Why?

Empirically, evolutionary biologists examine environmental correlates for clues about the conditions that might favor one ploidy level over another. For example, Lewis (1985) proposed that nutrient limitation may favor haploids over diploids, due to their higher surface area-to-volume ratio. For similar reasons, it has been suggested that toxic environments may favor diploidy (Zörgö *et al.* 2013). One particularly striking pattern is that haploidy is rare among large-bodied organisms, suggesting that diploidy may have evolved to mask somatic mutations and reduce the risk of cancer among multicellular organisms (Orr 1995; Mable and Otto 1998). On the other hand, many organisms are not dominated by one ploidy phase, and several authors have explored the conditions that maintain both phases, including niche differentiation (Hughes and Otto 1999; Rescan *et al.* 2016) or contrasting dispersal *vs.* growth advantages (Bell 1997).

How does evolutionary theory move beyond the expected trajectory to grapple with the diversity of life that has evolved? One approach is to examine how genes evolve to alter characteristics of the biological system, such as its physical genome, the genetic map, the mutation rate, and the mode of reproduction. Such models can then be used to predict how different initial conditions (either biotic or abiotic) can generate different outcomes.

For example, evolutionary biologists have altered standard models of selection to include genes that modify the alternation of generations between haploid and diploid phases. These models can clarify the conditions under which evolution is predicted to take different paths. For example, selection is generally more efficient in haploid populations than in diploid populations, because alleles are fully expressed. On the other hand, individuals that bear deleterious mutations are much more likely to survive to reproduce if they are diploid and can mask the mutation with a second functional gene copy. These two opposing forces generate conflicting pressures on the life cycle. The evolutionary outcome depends, among other aspects, on the rate of sex and recombination within the population (see review by Mable and Otto 1998). Haploid life

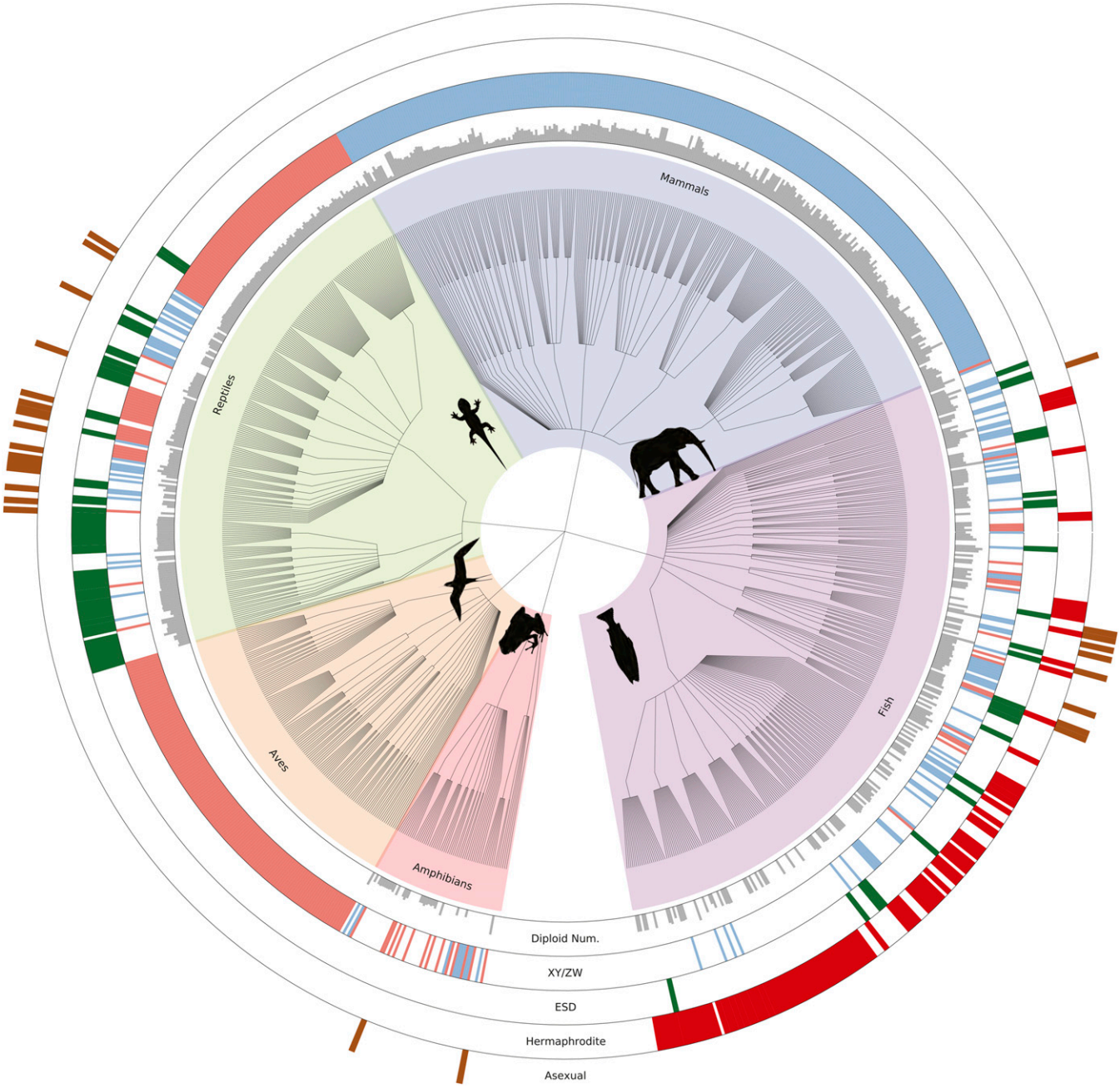

Figure 3 Variation in chromosome number ("Diploid Num.") and mode of reproduction across the vertebrate tree of life. XY/ZW, chromosomal sex determination (XY in blue; ZW in orange). ESD, environmental sex determination; hermaphrodite, individuals can serve as both males and females; asexual, reproduction without sexual mixing of genomes (including parthenogenesis, gynogenesis, and hybridogenesis). Modified from figure 3 of Tree of Sex Consortium (2014a).

cycles are favored in facultatively sexual organisms, when sex and recombination are rare, because genes modifying the life cycle gain from the long-term advantages of more efficient selection. By contrast, diploid life cycles are favored when sex and recombination are common because the modifier soon recombines away from its original lineage and benefits from masking.

Making further progress, however, requires that we be able to map life cycles across the tree of life and test predictions from such theories. Major progress has been made recently in the construction of large-scale phylogenies (*e.g.*, of birds, Jetz *et al.* 2012), progress made possible by the rise of genetic sequence data that are openly available from a large variety of taxa (Benson *et al.* 2013). Indeed, a first draft of the entire tree of life, combining classical taxonomic information and genetic data, has now been proposed (Hinchliff *et al.* 2015). Many uncertainties remain, and conflict among data sources is rife across this draft tree, but its existence makes it clearer

where the gaps lie, allowing us to target our efforts over the next century to improve the resolution of the tree of life.

Alongside phylogeny reconstruction, the development of trait databases [*e.g.*, the Chromosome Counts Database, Rice *et al.* 2015; genome-size databases, Gregory *et al.* 2007; the Plant Trait Database (TRY), Kattge *et al.* 2011; the Tree of Sex Database, Tree of Sex Consortium 2014b; the Encyclopedia of Life, Parr *et al.* 2014], make it increasingly possible to test predictions and evaluate theories about when and why evolutionary transitions occur. For example, in a recent study, we traced the evolution of sex chromosome fusions along phylogenies of fish and reptiles to evaluate different hypotheses for why fusions establish within populations (neutral drift, chance deleterious events, sexually antagonistic selection; Pennell *et al.* 2015).

Future work will increasingly reveal the context in which evolutionary transitions are favored. Where across the tree of algae do life cycle transitions occur and what features of the organism or its environment best predict when these transitions happen? What explains broad-scale variation in the number of chromosomes or in the genetic map length of chromosomes among taxa? For example, the social Hymenoptera exhibit 10 times higher rates of recombination than their solitary cousins (Ross *et al.* 2015); do such evolutionary transitions coincide best with changes in effective population size, colony size and structure, or mating system (*e.g.*, monogamy or polyandry)?

We are only now building the trees and the databases that will allow the next century of evolutionary genetics to synthesize the data to reveal when and where evolutionary transitions occur. This will not only allow evolutionary hypotheses to be tested, but they will undoubtedly reveal new patterns that require explanation.

Conclusion

The past century of genetics has focused on the main effects, the common patterns. The next century of genetics will be the century of diversity, of variance in function, and divergence in the paths taken by evolution. When we characterize genomes and their activity from more than one individual, from unusual species, in different environments, and at various points in development, we will increasingly unveil the unusual roles that genes play in different contexts. We will also come to better understand evolution, not as a singular process, but as a contingent process that can and has led to the endless forms most beautiful on earth.

Acknowledgments

We thank Heath Blackmon for sharing the source code to generate Figure 3; Ailene MacPherson for assistance in generating Figure 1; and Michael Scott, Gil Henriques, and Mark Johnson for helpful comments on the manuscript. We gratefully acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2016-03711 to S.P.O.) and the Department of Zoology at the University of British Columbia (to L.S.).

Literature Cited

Barrett, R. D. H., S. M. Rogers, and D. Schluter, 2008 Natural selection on a major armor gene in threespine stickleback. Science 322: 255–257.

Bell, G., 1997 The evolution of the life cycle of brown seaweeds. Biol. J. Linn. Soc. Lond. 60: 21–38.

Bell, E. A., P. Boehnke, T. M. Harrison, and W. L. Mao, 2015 Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. Proc. Natl. Acad. Sci. USA 112: 14518–14521.

Benson, D. A., M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman *et al.*, 2013 GenBank. Nucleic Acids Res. 41: 36–42.

Buckler, IV, E. S., T. L. Phelps-Durr, C. S. K. Buckler, R. K. Dawe, J. F. Doebley *et al.*, 1999 Meiotic drive of chromosomal knobs reshaped the maize genome. Genetics 153: 415–426.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. J. Villarreal, M. Dickson *et al.*, 2005 Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307: 1928–1933.

Costello, M. J., R. M. May, and N. E. Stork, 2013 Can we name Earth's species before they go extinct? Science 339: 413–416.

Didion, J. P., A. P. Morgan, L. Yadgary, T. A. Bell, R. C. McMullan *et al.*, 2016 R2d2 drives selfish sweeps in the house mouse. Mol. Biol. Evol. 33(6): 1381–1395.

Elliott, T. A., and T. R. Gregory, 2015 What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370: 20140331.

Engelhardt, B. E., M. I. Jordan, K. E. Muratore, and S. E. Brersfser, 2005 Protein molecular function prediction by Bayesian phylogenomics. PLoS Comput. Biol. 1: 0432–0445.

Evans, J. D., and D. E. Wheeler, 2001 Gene expression and the evolution of insect polyphenisms. Bioessays 23: 62–68.

Fang, W., X. Wang, J. R. Bracht, M. Nowacki, and L. F. Landweber, 2012 Piwi-interacting RNAs protect DNA against loss during oxytricha genome rearrangement. Cell 151: 1243–1255.

Gardner, A., and L. Ross, 2014 Mating ecology explains patterns of genome elimination. Ecol. Lett. 17: 1602–1612.

Gomez-Mestre, I., R. A. Pyron, and J. J. Wiens, 2012 Phylogenetic analyses reveal unexpected patterns in the evolution of reproductive modes in frogs. Evolution 66: 3687–3700.

Gregory, T. R., J. A. Nicol, H. Tamm, B. Kullman, K. Kullman *et al.*, 2007 Eukaryotic genome size databases. Nucleic Acids Res. 35: D332–D338.

Grigoriev, A., 2003 On the number of protein-protein interactions in the yeast proteome. Nucleic Acids Res. 31: 4157–4161.

Gutierrez-Arcelus, M., H. Ongen, T. Lappalainen, S. B. Montgomery, A. Buil *et al.*, 2015 Tissue-specific effects of genetic and epigenetic variation on gene regulation and splicing. PLoS Genet. 11: 1–25.

Hinchliff, C. E., S. A. Smith, J. F. Allman, J. G. Burleigh, R. Chaudhary *et al.*, 2015 Synthesis of phylogeny and taxonomy into a comprehensive tree of life. Proc. Natl. Acad. Sci. USA 112: 12764–12769.

Houri-Ze'evi, L., Y. Korem, H. Sheftel, L. Faigenbloom, I. A. Toker *et al.*, 2016 A tunable mechanism determines the duration of the transgenerational small RNA inheritance in *C. elegans*. Cell 165: 88–99.

Hughes, J. S., and S. P. Otto, 1999 Ecology and the evolution of biphasic life cycles. Am. Nat. 154: 306–320.

Ito, T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori *et al.*, 2001 A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. USA 98: 4569–4574.

Jetz, W., G. H. Thomas, J. B. Joy, K. Hartmann, and A. O. Mooers, 2012 The global diversity of birds in space and time. Nature 491: 444–448.
Jiang, Y., D. I. Bolnick, and M. Kirkpatrick, 2013 Assortative mating in animals. Am. Nat. 181: E125–E138.
Kattge, J., S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley *et al.*, 2011 TRY: a global database of plant traits. Glob. Change Biol. 17: 2905–2935.
Kwekel, J. C., V. G. Desai, C. L. Moland, V. Vijay, and J. C. Fuscoe, 2013 Sex differences in kidney gene expression during the life cycle of F344 rats. Biol. Sex Differ. 4: 14.
Lacaze, X., P. M. Hayes, and A. Korol, 2009 Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. Heredity 102: 163–173.
Lewis, W. M., 1985 Nutrient scarcity as an evolutionary cause of haploidy. Am. Nat. 125: 692–701.
Lewontin, R., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
Lynch, M., 2007 *The Origins of Genome Architecture*. Sinauer Associates, Sunderland, MA.
Mable, B. K., and S. P. Otto, 1998 The evolution of life cycles with haploid and diploid phases. BioEssays 20: 453–462.
Nee, S., R. M. May, and P. H. Harvey, 1994 The reconstructed evolutionary process. Philos. Trans. R. Soc. B Biol. Sci. 344: 305–311.
Orr, H. A., 1995 Somatic mutation favors the evolution of diploidy. Genetics 139: 1441–1447.
Otto, S., and J. Whitton, 2000 Polyploid incidence and evolution. Annu. Rev. Genet. 34: 401–437.
Parr, C. S., N. Wilson, P. Leary, K. S. Schulz, K. Lans *et al.*, 2014 The Encyclopedia of Life v2: providing global access to knowledge about life on earth. Biodivers. Data J. 2: e1079.
Pennell, M. W., M. Kirkpatrick, S. P. Otto, J. C. Vamosi, C. L. Peichel *et al.*, 2015 Y fuse? Sex chromosome fusions in fishes and reptiles. PLoS Genet. 11(5): e1005237.
Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas *et al.*, 2013 Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat. Commun. 4: 1958.
Rescan, M., T. Lenormand, and D. Roze, 2016 Interactions between genetic and ecological effects on the evolution of life cycles. Am. Nat. 187: 19–34.
Rice, A., L. Glick, S. Abadi, M. Einhorn, N. M. Kopelman *et al.*, 2015 The Chromosome Counts Database (CCDB): a community resource of plant chromosome numbers. New Phytol. 206: 19–26.
Ross, C. R., D. S. DeFelice, G. J. Hunt, K. E. Ihle, G. V. Amdam *et al.*, 2015 Genomic correlates of recombination rate and its variability across eight recombination maps in the western honey bee (Apis mellifera L.). BMC Genomics 16: 107.
Saccone, C., C. Gissi, C. Lanave, A. Larizza, G. Pesole *et al.*, 2000 Evolution of the mitochondrial genetic system: an overview. Gene 261: 153–159.
Schaefer, M. H., L. Serrano, and M. A. Andrade-Navarro, 2015 Correcting for the study bias associated with protein-protein interaction measurements reveals differences between protein degree distributions from different cancer types. Front. Genet. 6: 1–8.
Scheiner, S. M., and R. F. Lyman, 1989 The genetics of phenotypic plasticity I. Heritability. J. Evol. Biol. 2: 95–107.
Simon, J. C., F. Delmotte, C. Rispe, and T. Crease, 2003 Phylogenetic relationships between parthenogens and their sexual relatives: the possible routes to parthenogenesis in animals. Biol. J. Linn. Soc. Lond. 79: 151–163.
Tree of Sex Consortium, 2014a Sex determination: Why so many ways of doing it? PLoS Biol. 12: e1001899.
Tree of Sex Consortium, 2014b Tree of Sex: a database of sexual systems. Sci. Data 1: 140015.
Toth, A. L., K. Varala, T. C. Newman, F. E. Miguez, S. K. Hutchison *et al.*, 2007 Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. Science 318: 441–444.
Uetz, P., L. Giot, G. Cagney, and T. A. Mansfield, R. S. Judson *et al.*, 2000 A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. Nature 403: 623–627.
Urban, T., 2014 From 1,000,000 to Graham's Number. Blog: Wait but Why. Available at: http://waitbutwhy.com/2014/11/1000000-grahams-number.html. Accessed: August 2, 2016.
Venkatesan, K., J.-F. Rual, A. Vazquez, U. Stelzl, I. Lemmens *et al.*, 2009 An empirical framework for binary interactome mapping. Nat. Methods 6: 83–90.
Whitton, J., C. J. Sears, E. J. Baack, and S. P. Otto, 2008 The dynamic nature of apomixis in the angiosperms. Int. J. Plant Sci. 169: 169–182.
Zörgö, E., K. Chwialkowska, A. B. Gjuvsland, E. Garré, P. Sunnerhagen *et al.*, 2013 Ancient evolutionary trade-offs between yeast ploidy states. PLoS Genet. 9: e1003388.

Communicating editor: M. Johnston