approximately 86 000 available georeferenced mitochondrial sequences from more than 4500 amphibian and mammal species worldwide, to provide the first global map of genetic diversity. This global mapping revealed clear latitudinal gradients (as observed traditionally for species diversity), with genetic diversity decreasing from the tropics to the poles (Figure 1C). In addition, the authors identified significant anthropogenic impacts on the genetic diversity of amphibians, pinpointing urgent biodiversity conservation challenges at the global scale. As a result of its unprecedented spatial scale, this work identified hotspots of genetic diversity across the planet, and paves the way for novel avenues into both theoretical and applied evolution research.

Third, Chen *et al.* [6] performed a 20-year field survey of the endangered Florida scrub jay (*Aphelocoma coerulescens*) to describe temporal patterns of genomic diversity in a focal population. Through detailed temporal screening of changes in genome-wide diversity, the authors identified a rapid decline in the number of individuals immigrating from peripheral populations to the focal population, likely due to increased anthropogenic fragmentation of the surrounding landscape. They further demonstrated that immigrants had a vital role in maintaining low levels of inbreeding in the focal population (Figure 1D). Inbreeding negatively influences traits related to the fitness of plants and animals, leading to inbreeding depression [7]. Accordingly, the authors showed that the reduction in the number of immigrants over time due to human activities was associated with strong inbreeding depression on several fitness-related traits, including reduced hatching success and survival rates. This study demonstrates the strength of long-term genetic surveys to illuminate harmful effects of habitat fragmentation on the well-being of wild populations, and to uncover 'early signals' of population decline that can be efficiently used to preclude population extinctions.

Ongoing data generation and accumulation clearly is initiating a major shift towards unraveling broad-scale patterns of genetic diversity and their underlying processes, and, hence, to resolving open and integrative questions in evolutionary biology. This shift to a macrogenetics view of population genetics goes beyond comparative population genetics by embracing all scales of variation (i.e., taxonomic, spatial, and temporal scales), which is a necessary step to better appraise how underlying processes interact across scales. To further exemplify this perspective, we highlight how this shift toward macrogenetics will improve long-term species conservation, provided that massive data generation is accompanied by parallel computational developments to reduce both data and systematic errors that may drastically slow down the achievements of these fascinating goals. First, understanding the role of genetic diversity in structuring ecological communities has been the focus of many studies, yet we still do not know the large-scale impacts of genetic diversity on the assemblages of species to recommend management of high genetic diversity as a means of preserving ecosystem functioning [8]. Second, a more holistic perspective on population genetic patterns associated with the interactions of species (e.g., host–parasite or plant–pollinator dynamics) could highlight the crucial role of 'genetic synchronism' in rendering specialized ecological networks of interacting species more vulnerable to environmental change [9]. Third, it is widely accepted that both nonadaptive and adaptive (i.e., driven by natural selection) components of genetic diversity govern the ecoevolutionary dynamics underlying the long-term survival of species. However, we still question whether these distinct, yet not mutually exclusive, components follow similar temporal and spatial patterns and to what extent they contribute to the ability of species to cope with environmental changes [10]. We argue that the opportunities provided by recent sequencing approaches and decades of genetic diversity data accumulation should be exhaustively exploited to boost our understanding of the broad-scale processes driving macrogenetic diversity in light of the sustainable conservation of ecosystems and their services.

[1]Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), Station d'Ecologie Théorique et Expérimentale, UMR-5321, F-09200 Moulis, France

*Correspondence:
simon.blanchet@sete.cnrs.fr (S. Blanchet).

http://dx.doi.org/10.1016/j.tig.2017.06.007

References

1. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution,* Cambridge University Press
2. Ellegren, H. and Galtier, N. (2016) Determinants of genetic diversity. *Nat. Rev. Genet.* 17, 422–433
3. Lawton, J.H. (1996) Patterns in ecology. *Oikos* 75, 145
4. Romiguier, J. *et al.* (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515, 261–263
5. Miraldo, A. *et al.* (2016) An Anthropocene map of genetic diversity. *Science* 353, 1532–1535
6. Chen, N. *et al.* (2016) Genomic consequences of population decline in the endangered Florida scrub-jay. *Curr. Biol.* 26, 2974–2979
7. Charlesworth, D. and Charlesworth, B. (1987) Inbreeding depression and its evolutionary consequences. *Ann. Rev. Ecol. Syst.* 18, 237–268
8. Whitlock, R. (2014) Relationships between adaptive and neutral genetic diversity and ecological structure and functioning: a meta-analysis. *J. Ecol.* 102, 857–872
9. Strona, G. and Lafferty, K.D. (2016) Environmental change makes robust ecological networks fragile. *Nat. Commun.* 7, 12462
10. Mittell, E.A. *et al.* (2015) Are molecular markers useful predictors of adaptive potential? *Ecol. Lett.* 18, 772–778

Spotlight

CRISPRing the Regulatory Genome, the Challenge Ahead

Stephanie E. Sansbury,[1,2,3] Katelyn M. Sweeney,[1,2,3] and Ophir Shalem[1,2,*]

CRISPR saturation mutagenesis has the potential to dissect the functional landscape of

noncoding regions, but is highly susceptible to false discovery and misinterpretation. As recently published, Canver *et al.* have now taken the first steps towards addressing these issues by increasing screening resolution and analyzing the effects of off targets on hit calling.

Forward genetic screens using CRISPR-associated nucleases are emerging as a prominent tool for functional genomics due to high target specificity, simple programmability, and incredible versatility achieved by fusing different protein domains to an inactivated nuclease. CRISPR-based screens targeting annotated coding sequences (CDS) have displayed remarkable results, including high concordance between unique reagents targeting the same gene and high perturbation efficiency. Most importantly, they continue to reveal a large number of previously uncharacterized gene hits associated with multiple cellular phenotypes. Given that RNA-guided CRISPR nucleases target the DNA directly, they can be used for not only performing functional studies of annotated genetic elements, but also for the discovery of genetic elements through unbiased indel saturation targeting of every possible site within a large genomic region.

Early noncoding CRISPR screens demonstrated the strength and promise of such an approach for the discovery of regulatory elements and for dissecting the functional landscape of genomic regions harboring disease-associated single nucleotide polymorphisms (SNPs). In a previous paper, Canver *et al.* performed single-guide (sg)RNA saturating mutagenesis on the human and mouse BCL11A erythroid enhancer [1]. These experiments revealed the functional organization of these enhancers, key differences between the human and mouse sequences, and an *in vivo* effect of the discovered regulatory regions on hemoglobin switching. Other CRISPR-based studies also showed precise identification of novel regulatory elements using either a growth- [2,3] or fluorescence-based phenotype [4].

A critical feature of large-scale screens using targeting reagents designed *in silico* is the ability to exclude false positives and estimate the rate of false negatives. Reagents that perturb unintended targets in the genome give rise to false positives, while false negatives may result from inefficient targeting, or from efficient reagents that induce a second phenotype, preventing accurate measurement of the screened phenotype. The accuracy of CDS-targeting screens benefits from the availability of multiple precisely annotated exons within a gene, such that sgRNAs can be designed with sufficient flexibility to optimize library target specificity and retain multiple sgRNAs per element [5]. While CDS-targeting screens leverage this redundancy to exclude false positives and reduce the likelihood of false negatives [5], noncoding screens are handicapped by the poorly defined sizes and boundaries of noncoding genetic elements. In a recent paper, Canver *et al.* take the first steps in addressing these issues by increasing mutagenesis resolution, incorporating known genetic variation into sgRNA library design, and carefully analyzing the effects of sgRNA off-target activity on hit identification [6].

Genetic variation within the HBS1L-MYB intergenic region was found to be associated with erythroid traits and fetal hemoglobin (HbF) levels, which is likely due to changes in MYB expression. In their recently published paper, Canver *et al.* interrogated the regulatory landscape of the HBS1L-MYB intergenic region using an unbiased CRISPR saturation mutagenesis library targeting 98 DNase I-hypersensitive sites (DHSs) in that region. They improved the detection sensitivity of functional noncoding elements using two complementary strategies. First, they increased the density of target sites by using two nucleases, effectively doubling the number of nuclease-specific PAM sequences that anchor sgRNAs. Second, they incorporated information on existing genetic variation (from the 1000 Genomes Project) into the sgRNA library design, adding sgRNAs that accounted for documented variation in the target sequence as well as novel sgRNAs made possible by variants that generate a PAM site. This two-pronged strategy improved detection rates by expanding the pool of targetable elements and by increasing the number of consecutive sgRNAs targeting any one element, thus empowering HMM-based sliding window statistical analysis to reveal true positive functional regions [1,6].

sgRNAs that target additional sites in the genome (e.g., repetitive regions) are problematic not only because multiple edits confound the identification of the 'real' functional element, but also because runaway editing may result in a drop-out phenotype, probably due to activation of DNA damage pathways [7,8]. Given that MYB dysregulation is known to prevent proliferation of erythroid lineages, Canver *et al.* rely on a drop-out phenotype to recover DHSs with regulatory potential. To disentangle 'true' drop-outs resulting from the disruption of key regulatory elements from those caused by nonspecific targeting, Canver *et al.* stratified sgRNAs by off-target scores. This significantly reduced the number of hits, emphasizing the ease with which false positives are mistakenly identified if such an analysis is not performed, and calling attention to how the characteristics of noncoding regions (which are commonly enriched for repetitive and low-complexity sequences) may limit the utility of CRISPR-based screens to reveal their functional makeup.

Altogether, three DHSs were identified with predicted regulatory potential.

Follow-up with more in-depth analysis highlighted cases where there is a discrepancy between annotated transcription factor sites and the functional consequence of mutating that site. For example, a highly specific sgRNA in DHS −36, which was lacking an annotated regulatory sequence, produced a strong phenotype, while a previously implicated SNP in a GATA1 site did not produce any cellular phenotype. These examples highlight the importance of functional genetic analysis to verify and elaborate on the results of correlative studies. As next generation sequencing-based technologies continue to generate an unprecedented view of genome organization [9], the interactions they uncover will need to be interrogated by perturbation studies on a genomic scale. Thus, functional data produced from CRISPR-based unbiased noncoding screens followed by careful analyses will undoubtedly lead to a more complete understanding of the noncoding genome.

[1]Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[2]Department of Genetics, Perlman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[3]These authors contributed equally

*Correspondence: shalemo@upenn.edu (O. Shalem).

http://dx.doi.org/10.1016/j.tig.2017.07.001

References

1. Canver, M.C. *et al.* (2015) BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527, 192–197
2. Sanjana, N.E. *et al.* (2016) High-resolution interrogation of functional elements in the noncoding genome. *Science* 353, 1545–1549
3. Korkmaz, G. *et al.* (2016) Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat. Biotechnol.* 34, 192–198
4. Rajagopal, N. *et al.* (2016) High-throughput mapping of regulatory DNA. *Nat. Biotechnol.* 34, 167–174
5. Doench, J.G. *et al.* (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191
6. Canver, M.C. *et al.* (2017) Variant-aware saturating mutagenesis using multiple Cas9 nucleases identifies regulatory elements at trait-associated loci. *Nat. Genet.* 49, 625–634
7. Munoz, D.M. *et al.* (2016) CRISPR screens provide a comprehensive assessment of cancer vulnerabilities but generate false-positive hits for highly amplified genomic regions. *Cancer Discov.* 6, 900–913
8. Aguirre, A.J. *et al.* (2016) Genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 targeting. *Cancer Discov.* 6, 914–929
9. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74