

Opinion piece

Landscape evolutionary genomics

Tremendous advances in genetic and genomic techniques have resulted in the capacity to identify genes involved in adaptive evolution across numerous biological systems. One of the next major steps in evolutionary biology will be to determine how landscape-level geographical and environmental features are involved in the distribution of this functional adaptive genetic variation. Here, I outline how an emerging synthesis of multiple disciplines has and will continue to facilitate a deeper understanding of the ways in which heterogeneity of the natural landscapes mould the genomes of organisms.

Keywords: population genetics; adaptation; landscape ecology; genomics; natural selection; GIS

1. INTRODUCTION

In 2003, three landmark papers envisioned an emerging integration of ecology, evolution and population genetics. Luikart *et al.* (2003) defined the field of *Population Genomics* as the ‘simultaneous study of numerous loci or genome regions to better understand the roles of evolutionary processes that influence variation across genomes and populations’. Feder & Mitchell-Olds (2003) recognized the synthetic discipline *Ecological and Evolutionary Functional Genomics* or EEFG. The main goal of EEFG was to use all the genetic and genomic tools available to determine the exact functional genetic changes involved in the evolution of adaptations. A third field, *Landscape Genetics*, was born out of the fusion of population genetic techniques and landscape ecology’s layered geographical information system (GIS) maps (Manel *et al.* 2003). Landscape Genetics has thus far primarily focused on how various landscape features affect gene flow of neutral genetic variation, usually with the goal of identifying threatened or endangered populations for conservation purposes.

In this piece, I will briefly outline the current states of the fields of Population Genomics, EEFG and Landscape Genetics. I then discuss how a further synthesis of these fields has and will continue to facilitate a better understanding of the nature of adaptive genetic variation.

2. THE GENOMIC SCAN GOLD RUSH

Genomic scans are a major hallmark of Population Genomics. The last few years have seen an expansion of scans focused on genomic heterogeneity across habitats in a plethora of biological systems (Nosil *et al.* 2009). Here, a number of individuals from populations located in distinct habitats or across an ecological cline are genotyped for multiple markers. The logic behind

such genomic scans is that neutral regions of the genome will freely move between populations via gene flow while loci under selection will show higher genomic divergence across habitats. Genomic scans can range in size from couple of hundred markers to true population genomics through resequencing of the whole genomes with the aid of tiling arrays or next generation technologies (e.g. Turner *et al.* 2005).

Genetic differentiation resulting from habitat-mediated selection can result in divergence of neutral markers linked to locus under selection for many centimorgans. For example, a recent study claimed to find very long-distance genetic differentiation in the vicinity of quantitative trait loci (QTLs) for divergently selected traits in pea aphids (Via & West 2008). However, there are many instances of genetic differentiation extending only a few kilobases around a selected gene and even being limited to a single exon (Storz & Kelly 2008). Ultimately, the ratio of the selection coefficient to recombination rate determines the width of elevated divergence along a chromosome. Yet, expectations differ depending on whether the region was the subject of a recent selective sweep (Slatkin & Wiehe 1998) or long-term habitat-mediated balancing selection (Charlesworth *et al.* 1997). Unfortunately, if regions of elevated molecular divergence are small, any genomic scan with less than hundreds of thousands of markers will miss most important loci involved in adaptation. On the other hand, if the region of divergence is large, fewer markers will be required. Even so, determination of the ultimate cause of why any particular region is distorted and the extent to which a given locus contributes to adaptation will still require forward genetic approaches.

Beyond the difficulty in determining the causal mutations involved in adaptation exclusively through genomic scans, there are some fundamental problems with genomic scans that are often ignored. Population structure is a major challenge. When population structure is high, as is often the case for sessile organisms with discrete populations, it may be very difficult to detect outlier loci above the cloud of the high F_{ST} null distribution. Further, demographic histories are very difficult to determine. Past population bottlenecks and hierarchical population structure can contribute to high genome-wide variances in summary statistics (Excoffier *et al.* 2009). As a result, genomes can be extremely heterogeneous, which can lead to a high rate of false positives. Thus, it is possible that insufficient modelling of demographic history and not rampant selection may be the cause of the 5–10% rate of outlier loci found in a recent review of genomic scans studies (Nosil *et al.* 2009).

3. THE GENE-FIRST APPROACH

The best landscape-scale EEFG studies have first identified the genes involved in adaptive divergence and then established the spatial distribution of functional allelic variation through multi-population resequencing. The greatest of these successes have arguably come from studies of stickleback fish (Colosimo *et al.* 2005; Barrett *et al.* 2008) and *Peromyscus* mice (Steiner *et al.* 2007; Storz & Kelly 2008). In both systems, genes that are involved in adaptations to very divergent habitats have been cloned by forward genetic

techniques in conjunction with knowledge of candidate genes. After gene identification, population genetic analysis was conducted to determine the geographical distribution of alleles involved in ecotype-defining traits. This approach allowed the researchers to distinguish between phenotypes that result repeatedly from standing genetic variation and parallel phenotypes arising from new mutation.

Critically, field experimentation after gene identification can be used to confirm the adaptive significance of particular phenotypes. In the case of sticklebacks, field experiments with natural mutants of the armour control gene *eda* allowed researchers to test whether particular alleles are favoured in freshwater habitats (Barrett *et al.* 2008).

The gene-first approach is definitely more rigorous than genomic scans in terms of ability to identify novel gene functions and understand the forces involved in the geographical distribution of adaptive genetic variation. However, the cloning of genes remains an expensive and labour-intensive bottleneck in the process. Further, the difficulty of fine mapping and cloning adaptive genes means that they have for the most part been biased towards large-effect loci underlying discrete phenotypic traits.

Incorporation of QTL analysis into reciprocal transplant experiments may also be effective in determining the factors governing the spatial distribution of adaptive alleles, such as whether trade-offs at individual loci (i.e. antagonistic pleiotropy) underlie habitat-mediated adaptation. Recently, a study used field QTL analysis to determine the fitness effects of loci across habitats for plant ecotypes known to be locally adapted to coastal and inland habitats (Lowry *et al.* 2009). Here, three salt tolerance QTLs, previously identified in the laboratory, were found to have fitness effects in coastal but not inland habitat. This result may suggest that different sets of loci are responsible for adaptation to each habitat. Further, if adaptive alleles are indeed conditionally neutral, then they could diffuse unidirectionally by gene flow between habitats. More field studies are necessary to determine the extent to which trade-offs determine the spatial distribution of adaptive alleles among natural populations.

4. ADAPTIVE ALLELES AS A GIS LAYER

Since Manel *et al.* (2003), much thought has been put into how to combine multivariate-layered GIS maps with population genetic data. Many methods have been developed to assess population genetic structure (reviewed in Holderegger & Wagner 2008), and have been used to determine how landscape features contribute to the structuring of what is presumed to be neutral genetic variation. While exploring the distribution of neutral genetic variation can definitely inform us about the patterns and processes that limit gene flow, landscape genetics has yet to develop a framework to understand how landscape features contribute to the distribution of adaptive genetic variation.

Taking a landscape perspective could have huge implications for evolutionary biology. Studies of the

genetics of adaptation commonly focus on a single environmental factor as it is distributed across a cline or compare phenotypes across binomial habitats (e.g. coast versus inland). Natural landscapes are much more heterogeneous. Further, the distribution of adaptive alleles can be influenced by multiple environmental factors.

Landscape genetics is a maturing field that incorporates many types of data collected through remote sensing, weather stations and geological maps. These multivariate data are layered on top of each other and subsequent analyses are conducted. Genetic data can also be incorporated as a layer that can be used to understand the distribution of neutral genetic variation and gene flow (Kozak *et al.* 2008). Comparisons between the geographical distributions of neutral alleles and alleles thought to be involved in local adaptation could also be used to test for selection.

Joost *et al.* (2007) recently developed a methodology that uses GIS to compare geographical and genetic data to detect alleles associated with particular environmental factors. While this is a significant step forward, comprehensive analysis of the spatial distribution of alleles with regard to the distribution of environmental heterogeneity and barriers to gene flow has yet to be developed. The great hope is that multivariate geographical information could be incorporated with population genetic models to create more robust analyses of landscape-level natural selection. Further, field experimentation to 'ground truth' hypotheses as well as sampling design is very important with any landscape study and should be carefully considered before populations are selected for analysis.

5. FUTURE DIRECTIONS

As evolutionary biologists begin to get a better handle on what loci are involved in adaptations to different habitats, a new set of questions is likely to emerge. For example, it is currently unknown the extent to which fitness trade-offs at individual loci occur across the landscape, how geographical barriers influence the spread of adaptive versus neutral alleles and whether ecotypic divergence is due to the fixation of adaptive alleles or small shifts in allele frequencies at many loci. Current genome scans and gene-first approaches may not be representative of the complexity of landscape-scale adaptations as they are biased towards finding large-effect alleles that are fixed among taxonomic groups.

Recent studies on human population genetics provide a glimpse into what lies ahead for landscape evolutionary genomics. Coop *et al.* (2009) examined global allele frequencies across numerous populations at hundreds of thousands of single nucleotide polymorphisms to search for loci under selection. Overall, very few genes in the human genome had extreme allele frequency differences among populations. This may indicate that selection has only acted on a few loci. Alternatively, local selection may have been more widespread, but adaptive phenotypic change was achieved through small allelic changes at multiple loci. With ongoing improvements and decreased costs of genome sequencing technologies,

much broader analyses will soon be possible in many other systems. It will be important that these data be viewed in a landscape ecological context to better understand factors contributing to the geographical distribution of adaptive alleles.

6. CONCLUSIONS

Indeed, fully understanding adaptation on landscape scale is a monumental task even for one system. Habitat-mediated adaptation almost invariably involves multiple phenotypic changes each of which have a complex genetic basis. The complexity of this pursuit becomes multiplicative when landscape level environmental variation is added to the equation. Understanding adaptation at the level of the natural landscape may be especially difficult for evolution of polygenic traits, where adaptation has occurred through small allelic shifts across loci. Even so, there are now a few good examples of successfully connecting the distribution of functional genetic variation to coarse landscape features (Colosimo *et al.* 2005; Steiner *et al.* 2007; Storz & Kelly 2008). As more systems enter the genomic era, we will gain greater insight into how the mosaic of the natural landscape moulds the genomes of the organisms distributed across its vastness.

I would like to thank members of the T. Mitchell-Olds, M. Noor, M. Rausher, J. Willis and G. Wray labs at Duke University, B. Charlesworth and two anonymous reviewers for helpful conversations and suggestions. Funding was provided by NSF grants (EF-0328636, EF-0723814 and DEB-0710094) and an NIH graduate student fellowship.

David B. Lowry*

*University Program in Genetics and Genomics,
Box 3565 Duke University Medical Center,
Durham, NC 27710, USA*

*davidbryantlowry@gmail.com

- Barrett, R. D. H., Rogers, S. M. & Schluter, D. 2008 Natural selection on a major armor gene in threespine stickleback. *Science* **322**, 255–257. (doi:10.1126/science.1159978)
- Charlesworth, B., Nordborg, M. & Charlesworth, D. 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* **70**, 155–174. (doi:10.1017/S0016672397002954)
- Colosimo, P. F. *et al.* 2005 Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin

- alleles. *Science* **307**, 1928–1933. (doi:10.1126/science.1107239)
- Coop, G. *et al.* 2009 The role of geography in human adaptation. *PLoS Genet.* **5**, e1000500. (doi:10.1371/journal.pgen.1000500)
- Excoffier, L., Hofer, T. & Foll, M. 2009 Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285–298. (doi:10.1038/hdy.2009.74)
- Feder, M. E. & Mitchell-Olds, T. 2003 Evolutionary and ecological functional genomics. *Nat. Rev. Genet.* **4**, 651–657. (doi:10.1038/nrg1128)
- Holderegger, R. & Wagner, H. H. 2008 Landscape genetics. *Bioscience* **58**, 199–207. (doi:10.1641/B580306)
- Joost, S., Bonin, A., Bruford, M. W., Despres, L., Conord, C. & Erhardt, G. 2007 A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Mol. Ecol.* **16**, 3955–3969. (doi:10.1111/j.1365-294X.2007.03442.x)
- Kozak, K. H., Graham, C. H. & Wiens, J. J. 2008 Integrating GIS-based environmental data into evolutionary biology. *Trends Ecol. Evol.* **23**, 141–148.
- Lowry, D. B., Hall, M. C., Salt, D. E. & Willis, J. H. 2009 Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland *Mimulus guttatus*. *New Phytol.* **183**, 776–788. (doi:10.1111/j.1469-8137.2009.02901.x)
- Luikart, G., England, P. R., Tallmon, D., Jordan, S. & Taberlet, P. 2003 The power and promise of population genomics: from genotyping to genome typing. *Nat. Rev. Genet.* **4**, 981–994. (doi:10.1038/nrg1226)
- Manel, S., Schwartz, M. K., Luikart, G. & Taberlet, P. 2003 Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol. Evol.* **18**, 189–197. (doi:10.1016/S0169-5347(03)00008-9)
- Nosil, P., Funk, D. J. & Ortiz-Barrientos, D. 2009 Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375–402. (doi:10.1111/j.1365-294X.2008.03946.x)
- Slatkin, M. & Wiehe, T. 1998 Genetic hitch-hiking in a subdivided population. *Genet. Res.* **71**, 155–160. (doi:10.1017/S001667239800319X)
- Steiner, C. C., Weber, J. N. & Hoekstra, H. E. 2007 Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol.* **5**, e219. (doi:10.1371/journal.pbio.0050219)
- Storz, J. F. & Kelly, J. K. 2008 Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: insights from deer mouse globin genes. *Genetics* **180**, 367–379. (doi:10.1534/genetics.108.088732)
- Turner, T. L., Hahn, M. W. & Nuzhdin, S. V. 2005 Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol.* **3**, 1572–1578. (doi:10.1371/journal.pbio.0030285)
- Via, S. & West, J. 2008 The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Mol. Ecol.* **17**, 4334–4345. (doi:10.1111/j.1365-294X.2008.03921.x)