NEWS AND VIEWS

PERSPECTIVE

Standing and flowing: the complex origins of adaptive variation

JOHN J. WELCH* and CHRIS D. JIGGINS†
*Department of Genetics, University of Cambridge, Downing St., Cambridge CB23EH, UK; †Department of Zoology, University of Cambridge, Downing St., Cambridge CB23EJ, UK

A population faced with a new selection pressure can only adapt if appropriate genetic variation is available. This genetic variation might come from new mutations or from gene exchange with other populations or species, or it might already segregate in the population as standing genetic variation (which might itself have arisen from either mutation or gene flow). Understanding the relative importance of these sources of adaptive variation is a fundamental issue in evolutionary genetics (Orr & Betancourt 2001; Barrett & Schluter 2008; Gladyshev et al. 2008) and has practical implications for conservation, plant and animal breeding, biological control and infectious disease prevention (e.g. Robertson 1960; Soulé & Wilcox 1980; Prentis et al. 2008; Pennings 2012). In this issue of Molecular Ecology, Roesti et al. (2014) make an important contribution to this longstanding debate.

Keywords: adaptation, development and evolution, ecological genetics, fish, genomics, hybridization

Received 20 May 2014; revised 1 July 2014; accepted 10 July 2014

The authors simulate replicated colonization events of a novel habitat type, each from a common source population, and compare genetic differentiation at a locus under divergent selection, both between the source and derived populations and between pairs of derived populations. Simulations show a peak of FST between source and derived populations, which arises from the barrier to gene flow caused by locally adapted alleles (Petry 1983; Charlesworth et al. 1997). In contrast, between pairs of derived populations, there is a bimodal peak-valley-peak pattern. A related pattern has been noted before and shown to arise as a transient effect of a selective sweep in a structured population (Slatkin & Wiehe 1998; Bierne 2010; Kim & Maruki 2011). Roesti et al. show that these two effects interact and that the maintenance of local adaptation enhances and prolongs the twin peaks of differentiation between the derived populations—even though these share the same adaptive variant.

Roesti et al. next apply their theoretical insight to interpret genomic data from several contiguous populations of three-spine sticklebacks. Comparisons of marine and freshwater populations, and of pairs of freshwater populations, show the same patterns of differentiation that were observed in the simulations. This is consistent with the hypothesis that the freshwater populations adapted in parallel to fresh water, using alleles introgressed from the marine population.

One great strength of the study is that the long history of work in this system provides candidate loci whose putative role in freshwater adaptation is based on functional evidence. A danger of genome-wide studies is the ease with which we can generate post hoc explanations for any anomalous region. In a large genome, many loci will show unusual patterns, even under a null model (e.g. Teshima et al. 2006; Hermisson 2009; Pavlidis et al. 2012). This can be mitigated by focusing on loci with known fitness consequences. Although strong candidate loci will not be available for all systems, Roesti et al. (2014) show the benefits of studying genomic patterns in systems where such candidates do exist.

Roesti et al.’s study also shows clearly the benefits of studying adaptation in structured populations, with recurrent habitat types. Unlike isolated populations, such systems provide natural replication and the potential for observing the ancestral state, alongside the novel adaptation. Furthermore, such systems can have richer dynamics than single populations or simple two-deme models. However, this very richness also makes the data more difficult to interpret in terms of the classic debates about the sources of adaptive variation.

For example, we can imagine four distinct evolutionary scenarios that might have led to the same adaptive alleles being present in each of the freshwater habitats (Fig. 1). The first scenario is adaptation from new mutations arising independently in the freshwater habitats and so represents parallel evolution in the most straightforward sense (Arendt & Reznick 2007). This scenario could generate the same peaks of differentiation between marine and freshwater populations, but would not be expected to lead to a valley of differentiation between the freshwater populations, because different freshwater populations are unlikely to have fixed the same haplotype if their adaptive mutations arose independently. The second scenario is secondary contact involving multiple freshwater refugia. This could also produce the peaks of FST between marine and freshwater habitats, because recurrent migration after the contact might equalize allele frequencies in regions of the genome not involved in local adaptation. However, this scenario seems unlikely to lead to the peak-valley-peak pattern of differentiation between the
freshwater populations. These two scenarios highlight how the comparison between pairs of derived populations can be informative regarding the source of adaptive variants.

The third and fourth scenarios have some important features in common. Both involve locally adapted alleles passing from the marine habitat (where they are deleterious) to the novel freshwater habitats (where they are beneficial), and so, both involve adaptation via shared variation (Barrett & Schluter 2008). From this perspective, the sole difference is whether the variation was introduced to the marine populations by recurrent mutation (scenario no. 3; Kimura 1965) or recurrent migration from another freshwater population (scenario no. 4; Gillespie 1973), and furthermore, these processes have some formal similarities (Pennings & Hermisson 2006).

But despite these similarities, there are also important differences between the scenarios. Most importantly, in scenario no. 3, the selection for freshwater habitation might be truly novel (and so this scenario represents ‘adaptation from standing genetic variation’ in the most common sense). In contrast, scenario no. 4 requires a population that was already adapted to fresh water. For this reason, scenario no. 4 is sometimes called ‘the transporter hypothesis’ (Schluter & Conte 2009; Bierne et al. 2013), as it involves adaptive gene flow from the freshwater refugium to other freshwater habitats. The marine population acts less as a source population and more as a conduit for pre-existing adaptive variation. The consequence is that in scenario no. 4—but not scenario no. 3—the freshwater-adapted alleles may be relatively ancient and have had time to build up complex co-adapted haplotypes involving multiple mutational steps (Bierne et al. 2013; Stern 2013).

To match their stickleback data, the simulations of Roesti et al. assume invasion of a single freshwater-adapted allele found on a haplotype linked to many unique neutral variants. This haplotype is therefore not representative of the remaining standing variation at this locus in the marine population. In this way, the simulated data are more representative of the gene-flow scenario (no. 4) than the recurrent mutation scenario (no. 3), and we agree with the authors that this must be the most plausible explanation of the stickleback data. However, distinguishing with confidence between these scenarios will be challenging in some other cases. Currently, methods for structured populations lag behind the state of art in detecting selective sweeps in single populations (e.g. Günther & Schmid 2011). One potential solution is to use other properties of the allelic genealogies, which can contain more information about past selection than pairwise differentiation alone (Barton 1998; Le Corre & Kremer 2012). Roesti et al. (2014) complement their analyses of $F_{ST}$ with the Genealogical Sorting Index, a measure of the reciprocal monophyly in their genealogies, and approaches such as this should help in the future to identify with confidence the sources of adaptive genetic variation.
References


J.J.W. and C.D.J. contributed equally to writing this article.

doi: 10.1111/mec.12859