MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 2144–2164 doi: 10.1111/mec.13606

INVITED REVIEWS AND SYNTHESES

Genomics of local adaptation with gene flow

ANNA TIGANO and VICKI L. FRIESEN

Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada

Abstract

Gene flow is a fundamental evolutionary force in adaptation that is especially important to understand as humans are rapidly changing both the natural environment and natural levels of gene flow. Theory proposes a multifaceted role for gene flow in adaptation, but it focuses mainly on the disruptive effect that gene flow has on adaptation when selection is not strong enough to prevent the loss of locally adapted alleles. The role of gene flow in adaptation is now better understood due to the recent development of both genomic models of adaptive evolution and genomic techniques, which both point to the importance of genetic architecture in the origin and maintenance of adaptation with gene flow. In this review, we discuss three main topics on the genomics of adaptation with gene flow. First, we investigate selection on migration and gene flow. Second, we discuss the three potential sources of adaptive variation in relation to the role of gene flow in the origin of adaptation. Third, we explain how local adaptation is maintained despite gene flow: we provide a synthesis of recent genomic models of adaptation, discuss the genomic mechanisms and review empirical studies on the genomics of adaptation with gene flow. Despite predictions on the disruptive effect of gene flow in adaptation, an increasing number of studies show that gene flow can promote adaptation, that local adaptations can be maintained despite high gene flow, and that genetic architecture plays a fundamental role in the origin and maintenance of local adaptation with gene flow.

Keywords: adaptive introgression, new mutation, recombination, review, selection–migration balance, standing genetic variation

Received 7 December 2015; revision received 22 February 2016; accepted 1 March 2016

Introduction

That gene flow is a fundamental evolutionary force in speciation and adaptation is now widely accepted. Although the role of gene flow was addressed in many theoretical models since the early 1930s (e.g. Haldane 1930; Wright 1931; Slatkin 1987), empirical studies explicitly addressing the role of gene flow began with the advent of genetic markers, which allowed its quantification (Ellstrand 2014). Then, the development of new sequencing technologies and methods to estimate gene flow (e.g. Hey 2006) opened the door to many previously intractable questions in evolution, especially regarding speciation and adaptation with gene flow. The attention that gene flow has received in speciation theory caused a shift from the traditional geographical models of speciation (e.g. sympatric vs. allopatric speciation; Mayr 1942) to a framework centred on the role of gene flow (Wolf *et al.* 2010; Feder *et al.* 2012a). This shift, coupled with an unprecedented availability of genomic data, resulted in a bloom of theoretical and empirical studies on the genomics of speciation with gene flow (Feder *et al.* 2012a). Due to the general relationship between speciation and adaptation, some of the findings from these new studies also affected the way we perceive and study adaptation (Savolainen *et al.* 2013). However, theoretical models addressing specific questions about adaptation with gene flow at the genome scale have begun to be developed only very recently. Most of what we know about the genomics of adaptation with gene flow comes from studies where adaptation was investigated in a speciation framework, mostly as a means for reproductive isolation. However, understanding the effect of gene flow in the origin and

Correspondence: Anna Tigano, E-mail: anna.tigano@queensu.ca

maintenance of adaptation poses further challenges in that it requires that the genomic basis of the adaptive trait be characterized and that a connection is made among the putative adaptive trait, its genotype and its fitness effect (Barrett & Hoekstra 2011). In this respect, **QTL** (see Box 1 for definitions of terms in bold) studies and whole genome sequencing and resequencing provide remarkable tools to identify not only the genetic basis of adaptive traits but also their **genetic** and **genomic architectures**. Investigating the role of gene flow in adaptation is especially important as humans are rapidly changing both the natural environment (local and global habitats) and natural levels of gene flow (both increasing and decreasing them, for example creating new barriers to gene flow among previously interconnected populations or causing range shifts that connect differentiated populations and species; Crispo *et al.* 2011). Several reviews have recently focused on genomics of adaptation, genomics of speciation and speciation with gene flow (Barrett & Hoekstra 2011; Feder *et al.* 2012a; Olson-Manning *et al.* 2012; Savolainen *et al.* 2013; Seehausen *et al.* 2014). Despite the key role of gene flow in evolution, a thorough synthesis on the role of gene flow in the genomics of local adaptation is lacking. With this review on the genomics of local adaptation with gene flow, we aim to fill this gap.

'Adaptation with gene flow' includes three main scenarios. It refers to situations where (i) gene flow occurs between populations adapting to local conditions; and (ii) gene flow is re-established between differentiated populations following secondary contact. 'Adaptation with gene flow' can also apply to (iii) a single population where a stable adaptive polymorphism is maintained despite free interbreeding among different morphs (e.g. through **supergenes** as in colour mimicry in *Heliconius* butterflies; Joron *et al.* 2011). Theory proposes a multifaceted role for gene flow in adaptation. Basic theory predicts that gene flow disrupts the adaptation process if selection is not strong enough to prevent the loss of advantageous alleles (Haldane 1930). However, empirical evidence shows not only that gene flow can promote local adaptation but also that adaptive polymorphisms can be maintained within populations despite high gene flow (e.g. Joron *et al.* 2011; Comeault *et al.* 2015; Laurent *et al.* 2016). In this review, we first investigate selection on migration and gene flow and develop hypotheses about the effect of genetic architecture on the persistence of gene flow. Then, we explore the role of gene flow in the two stages of local adaptation: origin and maintenance of adaptation. We review the three main sources of genetic variation addressing specific questions related to how gene flow affects adaptation from new mutations and standing genetic variation, and the importance of gene flow as a source of adaptive variation, i.e. adaptive introgression. We then provide a synthesis of theoretical and simulation-based models of adaptive evolution with gene flow with a focus on clustering of adaptive loci. We review the mechanisms potentially maintaining adaptations when gene flow is high, and provide empirical evidence for adaptation with gene flow. Although gene flow among genetically differentiated populations and species can lead to the formation of hybrids, we do not discuss hybridization per se here and refer the reader to the rich body of literature on the topic (Bullini 1994; Seehausen 2004; Mallet 2007; Abbott *et al.* 2013). We also do not review genomic methods and analyses, but the reader will find explanation of methods in the studies cited in this review.

In the literature on adaptation, and especially in population genetics, '**migration**' and '**gene flow**' are often used interchangeably. In this study, we use 'migration' to refer to the movement and dispersal of individuals or gametes, and 'gene flow' for the movement of alleles, and eventually their establishment, into a genetic pool different from their genetic pool of origin (Endler 1977). That gene flow does not necessarily follow migration is important to remember. Additionally, we specify the intended meaning of 'migration' and 'gene flow' when needed, but do not replace 'migration-selection' with 'gene flow-selection', although the latter would be more appropriate.

Selection on migration and gene flow

When selection is spatially heterogeneous but temporally constant, gene flow can erode local adaptation by swamping local adaptive alleles, and/or impose a fitness cost on immigrants (Balkau & Feldman 1973; Lenormand 2002; Blanquart & Gandon 2011). For example, in a two alleles-two demes model, **gene swamping** is predicted to occur when $m/s > \alpha/(1\text{-}\alpha)$, where m is gene flow, s is selection in one deme and α is the ratio of selection coefficients between two demes (Bulmer 1972).

Selection can act directly on immigrants that disperse to an unfavourable environment, and/or on their ability or propensity to disperse. Several models of the evolution of **migration** with local adaptation, where migration refers to the heritable ability or propensity to disperse, have been developed (e.g. Billiard & Lenormand 2005; Blanquart & Gandon 2011). These are primarily two-locus models, where one locus is under selection and the other locus determines migration rate. Billiard & Lenormand (2005) discussed the importance of linkage and recombination between the two loci and suggested that migration rate can depend on the genetic architecture and effect

Box 1 Glossary.

Conditionally neutral selection: when two alleles do not confer fitness advantage in one environment but differ in their fitness effect in another environment.
CpG fraction: regions of the genome characterized by CpG sites, which are sites where the nucleotide cytosine is followed by the nucleotide guanine on a linear genetic sequence.
End-to-end fusions: fusions of chromosomes from their telomeric ends.
Environmentally antagonistic selection: given environment *A* and environment *B* where allele *a* is locally adapted in environment *A* and allele *b* is locally adapted in environment *B*, selection is environmentally antagonistic when allele *a* confers higher fitness than allele *b* in environment *A*, and allele *b* confers higher fitness than allele *a* in environment *B*.
Epistasis: interactions between different genes. The phenotypic effect of one allele at one gene depends on the effect of another allele at a second gene.
Fusions in holocentric chromosomes: fusions of chromosomes that result in a chromosome that lacks a defined centromere and the whole length functionally acts as a centromere.
G-matrix: matrix of additive genetic variances and covariances that describes the evolutionary trajectory of phenotypic traits based on their genetic relationships.
Gene flow: the movement, and eventual establishment, of alleles from one genetic pool to another that occur after successful mating. Gene flow follows migration (see definition), but not necessarily.
Gene swamping: the substitution of a locally adapted allele with a maladaptive allele from an immigrant individual that occurs when gene flow is stronger than selection.
Genetic architecture: the genetic basis of a trait and the interactions among alleles underlying the trait (dominance, epistasis, pleiotropy, polygeny).
Genomic architecture: genomic arrangements that characterize a trait (chromosomal rearrangements such as inversions, fusions, fissions, etc... and position in the genome).
Large-effect and small-effect alleles: alleles that have a high and low effect on fitness, respectively. Effect size can also refer to the contribution of an allele to a trait, independent from its fitness effect. In this review, we always refer to the fitness effect of a locus.
Migration: we refer to migration as the act of moving and dispersing by individuals or gametes. Migration could also refer to the heritable ability or propensity to disperse (*sensu* Billiard & Lenormand 2005). A more appropriate term to use for 'migration' as we mean it in this review would be 'dispersal', as the ecological meaning of 'migration' is the seasonal movement of animals from one location to another generally followed by a return migration. Nonetheless, we use the term 'migration' to maintain continuity with the large body of quantitative and population genetics studies that span over a century of evolutionary theory.
Phenotypic plasticity: ability of an individual to adjust its phenotype in response to changes in the environment, without genetic changes.
Pleiotropy: one or more variants in one gene affect several traits, either directly or indirectly (see Solovieff *et al.* 2013).
QTL: quantitative trait locus is a portion of the genome that correlates with variation at a polygenic phenotypic trait. A QTL can contain the gene underlying the trait, or be genetically linked to it.
RAD-seq: restriction site associated DNA sequencing is a genomic method to genotype DNA fragments associated with restriction sites that are cut using enzymatic digestion.
Rb translocations: Robertsonian (Rb) translocations are chromosomal rearrangements formed by the fusion of the long arms of acrocentric chromosomes (chromosomes where the centromere is situated in proximity of the telomere).
Reciprocal translocations: reciprocal exchange of portions of chromosomes.
RNA-seq: genomic method to sequence the RNA fragments expressed in a certain tissue at a given moment in time.
Supergene: 'a genetic architecture involving multiple linked functional genetic elements that allow switching between discrete complex phenotypes maintained in a stable local polymorphism' (Thompson & Jiggins 2014).
Tandem fusions: fusions on two chromosomes where the end of a chromosome fused with the end or the centromere of another chromosome.

size of the loci controlling both the migration rate and the adaptive trait.

We argue that the genetic architecture of locally adapted traits can also affect the strength of selection on gene flow. We distinguish 'selection on gene flow' from 'selection on migration', first, to differentiate gene flow from migration, both as the act of dispersing, and as the heritable ability to disperse (*sensu* Billiard & Lenormand 2005; see Box 1); second, because we are interested in the interactions that occur between locally adaptive and invading alleles within an individual genome after mating; and third, because of the temporal gap that potentially occurs between moving to a different environment and exchanging alleles with the local population.

An example of a study system where genetic architectures of adaptive traits affect selection on gene flow is the stick insect *Timema cristinae*, which has evolved different colour patterns (striped and unstriped green) to avoid predation on each of two plant hosts (Nosil *et al.* 2002). A third melanistic brown form, cryptic on woody parts of both plants, occurs across the range and was found to act as a 'genetic bridge' that maintains connectivity, i.e. gene flow, between the plant host specialized forms (Comeault *et al.* 2015). A combination of hierarchical dominance (unstriped, dominant over striped at the pattern locus; and green, dominant over brown at the colour locus) and epistasis (striped is not expressed in brown forms) between the loci controlling pattern and colour, respectively, relaxes selection against gene flow.

We propose a simple two-alleles two-environments model (as shown in Fig. 1) where the red allele is adaptive in the red environment and disfavoured in the blue environment, and the blue allele is adaptive in the blue environment and disfavoured in the red environment (i.e. **environmentally antagonistic selection**; Fig. 1a–c), and we assume equivalent dispersal potential and no variation in dispersal ability. We consider a simple genetic architecture such as dominance, and show that red phenotypes are disfavoured when they migrate to the blue environment regardless of the dominance pattern of the red allele, and that the genotype does not affect selection on immigrants (Fig. 1b). However, dominance can affect selection on gene flow. In fact, if red immigrants survive temporarily in the blue environment and successfully mate and exchange alleles with blue locals, selection on the first filial generation (F1) will depend on the dominance of the immigrant allele. If the blue and the red alleles are codominant, F1 individuals express an intermediate phenotype RB; if red is dominant, F1 individuals with genotype bR express the red phenotype and are disfavoured in the blue environment; if red is recessive, the genotype Br expresses the locally adapted phenotype making the red allele not visible to selection (Fig. 1d). Selection on F1 is highest if red is dominant, intermediate if red and blue are codominant, and null if red is recessive (Fig. 1d). Therefore, high levels of gene flow could be maintained if the immigrant allele is recessive, but gene flow would be selected against and therefore limited if the immigrant allele was dominant, thus potentially building up divergence between populations. In stick insects, in the absence of a genetic bridge based on the interaction between dominance and epistasis between two adaptive loci, gene flow between populations would have been limited and genetic differences would have accumulated and eventually led to speciation (Comeault *et al.* 2015).

This simple model of selection on gene flow assumes environmentally antagonistic selection, such that selection on colour is opposite in the two environments ($s_{red} = -s_{blue}$), and would apply anytime gene flow is asymmetrical, such as in continent-island scenarios, at range limits and/or in source-sink dynamics. The stronger the asymmetry the stronger the selection against gene flow is. In stick insects, gene flow is symmetrical due to the presence of a third generalistic phenotype (the melanistic morph) that masks the expression of locally adapted phenotypes (striped and unstriped green).

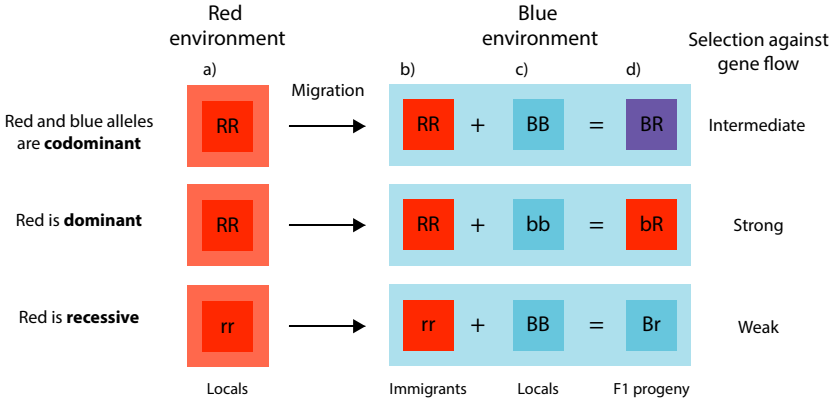


Fig. 1 Genetic architecture and selection on migration vs. gene flow. (a) Red individuals adapted to local conditions in the red environment (b) migrate to the blue environment, where selection against them is high, regardless of the genetic architecture (dominance) of their phenotype. (c) If locally adapted blue individuals mate with surviving immigrants (d) selection on F1 progeny, and therefore on gene flow, depends on the genetic architecture of the locally adaptive trait.

Although our hypothesis has not been formally tested either theoretically or empirically so far, the prediction that the genetic architecture of locally adaptive traits affects the strength of selection on gene flow could be extended to polygenic adaptive traits (i.e. quantitative traits) and other genetic architectures that are more resistant to gene flow, such as physical linkage and chromosomal rearrangements. Testing our hypothesis requires knowledge of the genetic architecture of a locally adaptive trait, the strength of selection on the trait, and levels of genetic differentiation and gene flow between two populations. Stick insects provide an example of how genetic architecture allows the maintenance of gene flow between locally adapted populations (Comeault *et al.* 2015), but more studies of this kind are necessary for rigorous testing.

Origin of adaptation

Three main sources of variation for genetic adaptation have been proposed: new mutations, standing genetic variation and adaptive introgression. Although the three sources have been reviewed elsewhere (Barrett & Schluter 2008; Hedrick 2013), here we use theoretical and empirical studies to address specific questions related to how gene flow affects adaptation from new mutations and standing genetic variation, and to the importance of gene flow as a source of adaptive variation, i.e. adaptive introgression. We treat each of the three categories separately for clarity, even though they are not independent (Fig. 2).

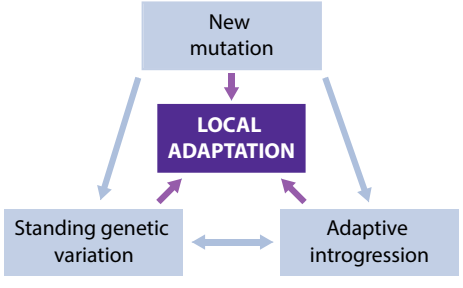


Fig. 2 Sources of adaptive variation and potential interactions among them. New mutations are the ultimate source of adaptive variation. They can either promote adaptation directly or contribute beneficial alleles to the pool of standing genetic variation in the same population, or in another population or species, which could potentially be transferred and be beneficial in another population (adaptive introgression). Introgressed adaptive alleles with a high fitness benefit can be quickly brought to fixation or be maintained within the pool of standing genetic variation.

How gene flow affects establishment of new beneficial mutations

In the transition from genetic to genomic studies of adaptation, theory has shifted from 'beanbag thinking' based on the study of individual genes (highly criticized by Mayr, who acknowledged the importance of **epistasis** and genetic linkage in speciation; Slatkin 1987), to models that included the interactions among genes, and between genes and the genomic landscape with particular attention to genetic architecture. In fact, recent studies have stressed the importance of genetic architecture in the process of divergence with gene flow (e.g. Bürger & Akerman 2011; Yeaman & Whitlock 2011; Feder *et al.* 2012b; Yeaman 2013).

For example, how does gene flow into a population affect the probability of establishment of new mutations under different genetic architectures? Whether interactions among genes are considered or not (single- vs. multi-locus models), simulation-based studies show that the interplay between selection (s) and rate of gene flow (m) is the strongest determinant of whether a beneficial new mutation establishes in a population: a new mutation is most likely to establish when s is greater than m (Haldane 1930; Wright 1931; Lenormand 2002; Yeaman & Otto 2011; Feder *et al.* 2012b), or otherwise it will be swamped from the local genetic pool. Linkage with an already differentiated locus does not increase the probability of establishment in cases of strong selection and low gene flow, but the role of linkage becomes appreciable when selection on the new mutation is below the rate of gene flow, and selection at the already differentiated locus is strong ($s > 2m$; Feder *et al.* 2012b). With an increasing number of loci under strong divergent selection, the probability of a new mutation being linked with one of these loci increases, and so does its probability of establishment. When genome-wide divergence increases (generally at a later stage in the divergence process) effective gene flow at the genome level decreases, thereby increasing the chances of a new mutation to establish, even with weak selection, so that the probability of establishment (P) depends solely on selection on the new mutation ($P = s/2$; Feder *et al.* 2012b).

Several approaches have been developed to determine whether an adaptive trait evolved from new mutations vs. standing genetic variation (Barrett & Schluter 2008). However, the key to understanding adaptation from new mutations is to compare the estimated date of the mutation event with the date of the environmental change that drove the evolution of the adaptation (Stapley *et al.* 2010). One example involves deer mice (*Peromyscus maniculatus*) in the Sand Hills of Nebraska, where light coat colour evolved to match the

lighter soil compared to the surrounding area (Linnen *et al.* 2009) in response to avian predation (Linnen *et al.* 2013). Linnen *et al.* (2009, 2013) demonstrated that the lighter colour has adaptive value and that the alleles conferring lighter coat colour arose *de novo* only after formation of the light soil Sand Hills. This system is useful for learning about the role of gene flow in the establishment of new mutations. Gene flow between light and dark mice is very high: the two morphs freely interbreed. Mice at the juncture of the dark- and light-soil habitats are phenotypically diverse, showing many intermediate phenotypes and low levels of linkage disequilibrium, which suggests that high recombination rates combined with large effective population sizes and gene flow maintained a diversity of alleles. In this study system, the establishment of the light derived alleles seems to have been driven by selection strong enough to overwhelm the homogenizing effect of gene flow. Additionally, the light phenotype is a combination of multiple pigmentation traits, mostly genetically independent, so that high gene flow and high recombination rates provide a plausible mechanism of the fine phenotypic tuning. The many mutations that fine-tune coloration in the deer mice now contributed greatly to the pool of standing genetic variation (Fig. 2).

Studies where adaptation is caused by new mutations are scarce (Table 1). In deer mice, coloration is the result of epistatic interactions between two genes, *Mc1r* and *Agouti*, but the detailed genetic architecture of the trait, in terms of the dominance hierarchy of individual SNPs (single nucleotide polymorphisms) for example, has not been described yet. More empirical studies are necessary to understand better the role of genetic architecture in the establishment of new beneficial mutations.

Gene flow in spatially and temporally variable environments can augment standing genetic variation

In addition to a century of quantitative genetics studies stressing the role of standing genetic variation in adaptation, genomic studies are increasingly demonstrating standing genetic variation to be the main source of adaptive variation (reviewed in Barrett & Schluter 2008; but see Karasov *et al.* 2010). Adaptation from standing genetic variation poses three main advantages compared to adaptation from new mutations (Barrett & Schluter 2008). First, it is faster because it evolves from alleles already available in the population. Second, due to their occurrence at higher frequencies, beneficial alleles not only spread faster in the population but also are less likely to be lost by drift. Third, the adaptive value of pre-existing alleles may have already been tested in the past or in other parts of the species' range.

The *Rhagoletis pomonella* study system provides clear insights into the advantages of evolving adaptations from standing genetic variation. *Rhagoletis pomonella* is a species complex of parasitic fruit flies, known to have recently (~150 years ago) shifted hosts after the introduction of apple (*Malus pumila*) in the United States. What seemed to be a classic case of fast host race formation and sympatric speciation revealed a history of ancient gene flow and introgression (Feder *et al.* 2003, 2005): marked differences between gene trees based on neutral markers (including mtDNA) vs. loci associated with diapause adaptation to different plant hosts suggested that alleles associated with the host shift not only were present well before apple was introduced (more than 1.5 Mya) but also derived from a currently allopatric Mexican fly population adapted to warmer conditions with different diapause cycles. Genetic

Table 1 Examples of studies in which the origin of adaptation is identified, and the role of gene flow is addressed

Species	Trait	Source	Reference
Deer mouse *Peromyscus maniculatus*	Cryptic coloration	New mutation → standing genetic variation	Linnen *et al.* (2009, 2013)
White Sands lizards *Sceloporus cowlesi* and *Aspidoscelis inornata*	Cryptic coloration	New mutation	Laurent *et al.* (2016)
Fruit fly *Rhagoletis pomonella*	Diapause	Standing genetic variation	Feder *et al.* (2003, 2005)
Yellow monkeyflower *Mimulus guttatus*	Morphology and phenology	Standing genetic variation	Monnahan *et al.* (2015)
Threespine stickleback *Gasterosteus aculeatus*	Marine vs. freshwater ecotypes	Standing genetic variation	Jones *et al.* (2012a)
Malaria mosquito *Anopheles coluzzii*	Insecticide resistance	Adaptive introgression	Clarkson *et al.* (2014); Norris *et al.* (2015)
Domestic mouse *Mus musculus domesticus*	Rodenticide resistance	Adaptive introgression	Song *et al.* (2011); Liu *et al.* (2015)
Human *Homo sapiens*	High altitude adaptation	Adaptive introgression	Huerta-Sánchez *et al.* (2014)

variation in diapause introduced by gene flow from Mexico into the United States seems to have allowed the host shift to apple and other plant hosts (Feder *et al.* 2003, 2005), making this study system an example of how alleles tested in another place or time can provide ready-to-use genetic material for adaptation to a new situation.

In fact, environmental heterogeneity appears to be one of the key sources for the maintenance of genetic variation (reviewed in Felsenstein 1976). For example, Yeaman & Jarvis (2006) used an extensive 20-year dataset of common garden experiments to test growth response to climate in lodgepole pines (*Pinus contorta*) and observed that the combination of gene flow and environmental heterogeneity promotes and maintains high levels of standing genetic variation in this species. The authors proposed two general conditions for environmental heterogeneity and gene flow to support genetic variation: strong selection to maintain phenotypic variation and counteract the homogenizing effect of gene flow, and substantial spatial variation. Similar results can be obtained in temporally variable environments. Blanquart *et al.* (2013) used evolutionary simulations to test the effect of gene flow in temporally variable and temporally stable environments and found that, while local adaptation is eroded by gene flow in temporally stable environments, intermediate levels of gene flow maximize local adaptation in the presence of temporal variation (Fig. 3). Therefore, temporal fluctuations seem to play a role similar to spatial heterogeneity (Gandon 2002; Blanquart & Gandon 2011; Fig. 4a). Regardless of whether environmental variation is spatial or temporal, the genetic architecture of a given adaptive trait, for example if selection is environmentally antagonistic rather that **conditionally neutral**, can affect the relationship between spatial/temporal variation and genetic diversity (Huang *et al.* 2015).

High levels of genetic variation can also alleviate gene swamping in local adaptation with gene flow. **Large-effect alleles** are predicted to be more resistant to swamping and more likely to contribute to local adaptation than **small-effect alleles** (Yeaman & Otto 2011). However, a simulation-based study (Yeaman 2015) suggested that high genetic variation can contribute to local adaptation through small-effect alleles in two main ways. First, alleles that are prone to swamping can contribute temporarily to local adaptation, before being swamped. Second, considering that a quantitative trait can be expressed by numerous combinations of several small-effect alleles, high genetic variation provides a pool of potential combinations from which the population can draw. An example of how high genetic variation can provide different combinations of alleles for the same phenotypic expression of a trait comes from colour mimicry in *Heliconius* butterflies, where different species have different genomic bases for wing colour patterns. In *Heliconius numata*, the 'supergene' *P* is most strongly associated with adaptive colour polymorphism, but weaker associations are also found with other loci, some of which underpin similar colour mimicry patterns in other *Heliconius* species (Jones *et al.* 2012b). In fact, the same phenotype can be obtained by different combinations of alleles at different loci on the same or different chromosomes, and this high genetic variation increases the number of available adaptive combinations of loci and alleles. Nevertheless, theoretical predictions about small-effect alleles are difficult to test empirically using genome scans because most currently available methods to detect signatures of selection have limited power when selection is weak and/

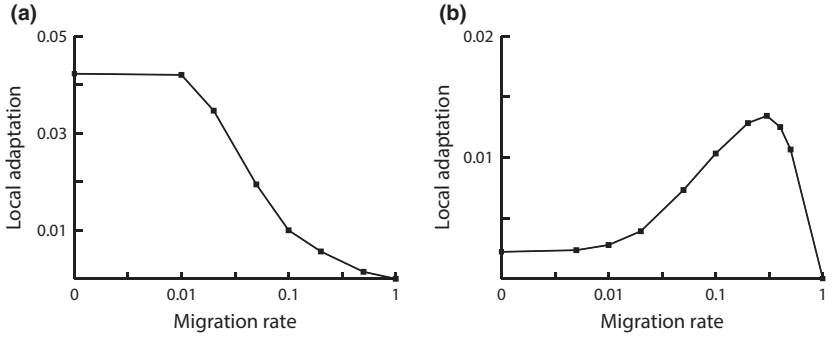


Fig. 3 Effect of gene flow on local adaptation in (a) temporally stable vs. (b) temporally fluctuating environments as inferred by evolutionary simulations (Blanquart *et al.* 2013). Each point represents the difference in fitness between local and migrant individuals. Panel (b) shows that in temporally variable environments intermediate levels of gene flow maximize local adaptation. Figure reproduced with permission of F. Blanquart and co-authors.

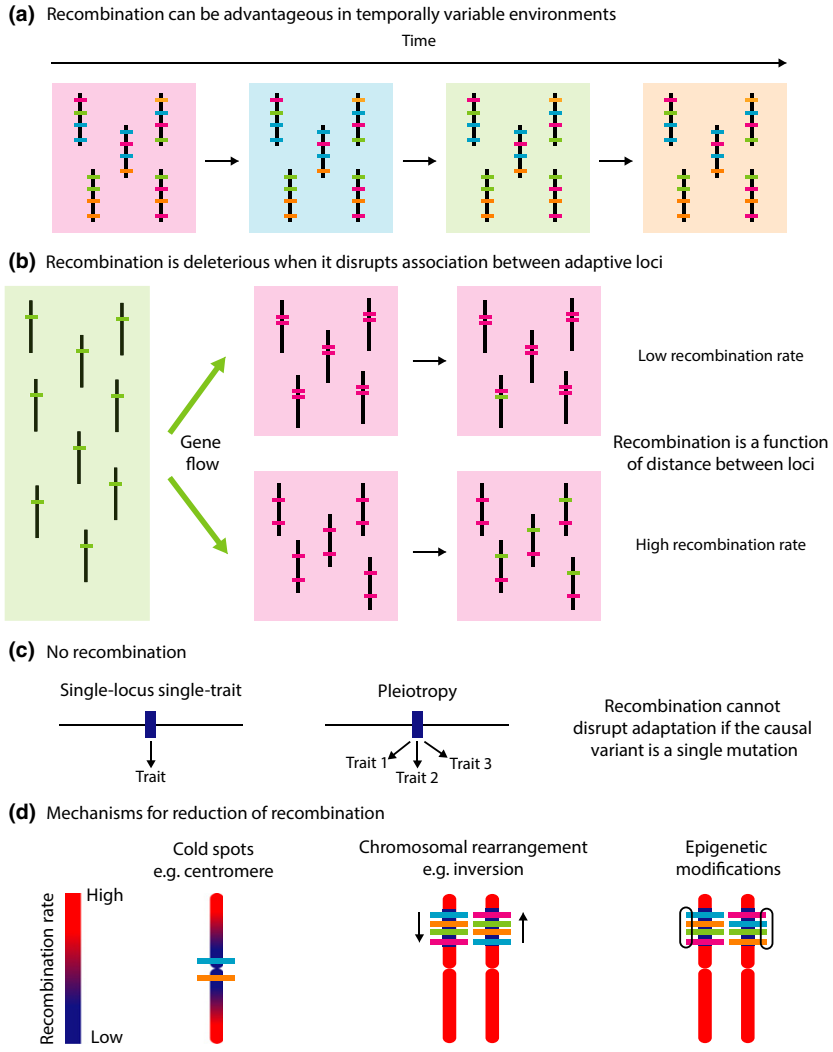


Fig. 4 The role of recombination in the genomics of local adaptation with gene flow. In (a) and (b) coloured boxes represent different environments, black vertical lines individual genomes, coloured horizontal lines different alleles. The allele is locally adaptive when it matches the colour of the environment it is found in. (a) High recombination rates and gene flow can be beneficial in temporally variable environments. High recombination among diverse loci increases the chances that some individuals in a population will have a combination that results in the survival of the population as a whole. (b) In a simple continent-island model (continent = green box, island = pink box), where local adaptation in the island is based on two loci (pink), the spread of a maladaptive allele (green) from the continent to the island depends on the distance between the two adaptive alleles in the island, which in turn affects recombination rates between continent and island alleles. (c) If the genetic basis of the trait is a single locus, recombination cannot disrupt adaptation. (d) Mechanisms that reduce recombination among clusters of locally adapted alleles. Chromosomes are shaded based on relative levels of recombination (blue = low, red = high).

or the trait of interest is polygenic (Yeaman 2015; but see Yang *et al.* 2010).

Studies that specifically address the role of gene flow in the origin and maintenance of standing genetic variation are rare (but see Feder *et al.* 2003, 2005). Gene flow following secondary contact due to climatic oscillations, however, is common for many temperate species in the Northern Hemisphere. What is the role of gene flow in adaptation in these situations? Analyses of 70 microsatellite loci in the European aspen (*Populus tremula* L.) revealed that postglacial admixture following secondary contact increased variance in phenotypic traits involved in adaptation, such as phenology of bud set, and suggested that selection acted on standing genetic variation, whose high levels were due to gene flow among populations (De Carvalho *et al.* 2010). A study on migration-selection balance in a recently established population of *Mimulus guttatus*, resembling a continent-island model, based on **RAD-seq** and whole genome data showed that the role of gene flow in adaptation can change with time: in a first phase gene flow augmented standing genetic variation on which selection acted to optimize local adaptation, but once the population was established and locally adapted, further gene flow introduced maladapted alleles (Monnahan *et al.* 2015).

Recently, whole genome resequencing data have provided insights into levels and features of standing genetic variation in natural populations of sticklebacks (Feulner *et al.* 2013; Chain *et al.* 2014), cichlids (Fan & Meyer 2014), pea aphids (Duvaux *et al.* 2015), and monkeyflowers *Mimulus guttatus* (Flagel *et al.* 2014). These studies suggest that even if SNPs are the most common form of genetic variation, other types of genetic variants, including structural variants (insertions, deletions, inversions and transpositions) and copy number variation, are widespread in the genome and might have an important role in adaptation. In

cichlids, large amounts of shared variation among five African species (Fan & Meyer 2014) were attributed to retained ancestral variation and/or gene flow, which are two factors hard to distinguish (Nachman & Payseur 2012). We encourage more research in this direction to gain a better understanding of the importance that gene flow has in contributing to levels of standing genetic variation, and how this affects the potential of populations to adapt.

Gene flow can promote adaptation via adaptive introgression

Although adaptive introgression generally refers to the movement of alleles from one species to another (Hedrick 2013; Hamilton & Miller 2016), the actual introgression of adaptive alleles through interbreeding and backcrossing applies to the population level as well, even when reproductive barriers are weak or nonexistent.

Adaptive introgression has been documented in several plant species (Martin 2005; Whitney *et al.* 2006, 2015; Kim *et al.* 2008), with the strength of evidence growing as sequencing technology and genomic coverage improve. Similarly, lateral gene transfer can spread antibiotic resistance among species of enteric and other bacteria (Ochman *et al.* 2000) and is *de facto* adaptive introgression. The frequency of hybridization and adaptive introgression was overlooked in animals until recently, mainly due to the prevalence of the biological species concept and low fitness observed in many animal hybrids (Hedrick 2013). Testing adaptive introgression requires identification of the donor species, characterization of the genetic basis of the trait in the donor and recipient species, and evidence of a significant fitness effect on the recipient species (Rieseberg 2011; Hedrick 2013). Additionally, the evolutionary history of the trait transfer must be observed or reconstructed (Hedrick 2013). Power in detecting events of adaptive introgression has been improved recently by increased genomic coverage made possible by high throughput sequencing technologies and new inference approaches based on phylogenetic networks and hidden Markov models (Liu *et al.* 2014), and association mapping methods (Hejase & Liu 2016).

The best-characterized example of adaptive introgression in animals is insecticide resistance in mosquitoes. Whole genome sequencing of *Anopheles gambiae* and *Anopheles coluzzii* showed the introgression of a whole island of divergence carrying an insecticide-resistance mutation (*Vgsc-1014F kdr*) from *A. gambiae* to *A. coluzzii* (Clarkson *et al.* 2014), despite strong reproductive isolation and low fitness in hybrids (Lee *et al.* 2013). Temporal transects pinpointed the introgression event to a massive insecticide-treated bed net campaign in Mali (Norris *et al.* 2015), which acted as a strong selective pressure on *A. coluzzii*, which previously lacked the beneficial allele. This study system also offered the opportunity to reconstruct, in almost real time, the mechanisms of adaptive introgression without formation of persistent hybrids (Box 2; Table 1).

Maintenance of adaptation

Theoretical framework

In the absence of gene flow, locally adapted alleles follow independent evolutionary trajectories and their persistence in the population is the result of the interaction between selection and drift (Wright 1931). In isolation, genetic differences, both neutral and adaptive, will accumulate and eventually lead to speciation (Coyne & Orr 2004). In contrast, with very high levels of gene flow, neutral and adaptive variations are homogenized and locally adapted alleles can be swamped (Bulmer 1972; Lenormand 2002). With intermediate levels of gene flow, however, locally adapted alleles can be maintained if gene flow is below a certain critical threshold (see gene swamping above; Haldane 1930; Blanquart & Gandon 2011; Bürger & Akerman 2011; Yeaman & Whitlock 2011; Blanquart *et al.* 2013). In this scenario, genetic architecture assumes a pivotal role in the preservation of local adaptation (see below; Bürger & Akerman 2011; Yeaman & Whitlock 2011; Yeaman 2013; Akerman & Bürger 2014).

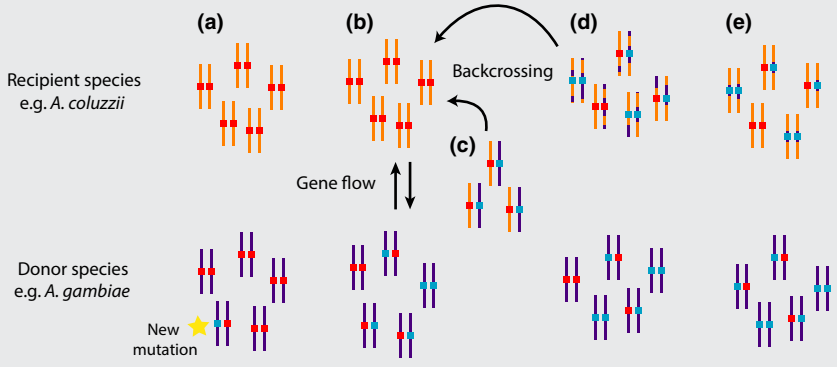
For decades, the theory on the evolution of adaptation with gene flow was based on either single-locus two-alleles models (Haldane 1930; reviewed in Lenormand 2002), where the interactions among loci in the genome were not considered, or two-locus models (Felsenstein 1976). However, adaptive traits are often polygenic and several studies have shown that predictions based on single-locus and two-locus models cannot be extrapolated to polygenic traits and the whole genome (Spichtig & Kawecki 2004; Yeaman 2015 and references therein). We reserve the term 'genomic models' for those models that consider linkage disequilibrium, epistasis and/or recombination among loci, which are pivotal in the maintenance of adaptation with gene flow. Among the two- and multilocus models that are available to date, the majority do not adopt a genomic view and tend to focus on very specific case scenarios, probably due to the modelling complications associated with the interactions among loci (Akerman & Bürger 2014; but see Barton 1983; Barton & Bengtsson 1986).

Single-locus theory predicts that in a continent-island model a locally adapted allele will be lost in the island

Box 2 The malaria mosquito as a model of genomics of adaptation with gene flow.

The evolution of insecticide resistance in the African malaria mosquitos provides an exemplary system to study genomics of adaptation with gene flow. Not only is it the best documented case of adaptive introgression, but it also shows the potential interplay between different sources of adaptive variation (Fig. 2). The insecticide-resistance mutation *Vgsc-1014F kdr* arose as a new mutation in *Anopheles gambiae* (Box Fig. 1a) adding to the standing genetic variation in the species (Box Fig. 1b). Thanks to its selective advantage, the mutation then steadily increased in frequency in several populations, at a rate that depended on the local use of insecticides (Tripet *et al.* 2007; Box Fig. 1b). The mutation eventually introgressed from *A. gambiae* to *Anopheles coluzzii* in an area of sympatry, and then spread to other populations of *A. coluzzii* through gene flow (Norris *et al.* 2015).

The short generation time of mosquitos in combination with strong human-induced selection allowed researchers to follow evolution in real time in a natural system, to observe the progress of adaptive introgression, and to understand how reproductive isolation between young species can be overcome. *A. gambiae* and *A. coluzzii* were recently recognized as different species, and although short bouts of hybridization occur, reproductive isolation is strong due to fitness reduction in hybrids (Lee *et al.* 2013). Whole genome sequences across a temporal transect indicated when introgression occurred, and when the fitness effect of *kdr* overcame the hybrid disadvantage and allowed backcrossing to *A. coluzzii* (Norris *et al.* 2015; Box Fig. 1c). Continuous backcrossing into the parental population eventually eroded differences between the two species (Box Fig. 1d), including even the island of divergence that contains *kdr*. The only exception to this pattern was the *kdr* mutation and surrounding sequence (Norris *et al.* 2015; Box Fig. 1e). Although it is based on the study of adaptive introgression in *Anopheles mosquitos*, we believe that Box Fig. 1 could be generalized to other cases of adaptive introgression, for example warfarin resistance in the house mouse (Song *et al.* 2011; Liu *et al.* 2015).



Box Fig. 1 Graphical representation of the process of adaptive introgression. Each line represents a DNA copy. Orange lines represent the recipient species DNA (e.g. *Anopheles coluzzii*), purple lines the donor species DNA (e.g. *Anopheles gambiae*), red squares the ancestral allele, and blue squares the derived adaptive allele (e.g. insecticide-resistant mutation *kdr*). Letters indicate the different steps of the process of adaptive introgression. (a) A new beneficial mutation arises in the donor species; (b) gene flow occurs between the donor and the recipient species (c) creating hybrids. (d) Hybrids backcross with the recipient parental species (e) until all donor species DNA is lost and only the beneficial mutation is retained.

if gene flow m overwhelms selection s (Haldane 1930). This prediction highlights the importance of the effect size of a beneficial allele: alleles with larger effects are less likely to be lost (Yeaman & Otto 2011). What happens when multiple loci are involved in adaptation? Genomic migration-selection models address how interactions among loci change the net effect of the tension between m and s (termed the diversification coefficient δ by Yeaman & Otto 2011), and therefore the critical migration (i.e. gene flow) threshold m_{crit} above which locally adapted alleles are lost (Bürger & Akerman 2011; Yeaman & Whitlock 2011; Aeschbacher & Bürger 2014). Because adaptation with gene flow favours large-effect alleles, multilocus phenotypes should be based on few large-effect alleles or a group of small-effect alleles in tight linkage that *de facto* act as a single large-effect locus (Griswold 2006; Yeaman & Otto 2011). Although tight genetic architectures are favoured, less-linked architectures can still persist, transitorily or stably, especially when selection is strong. Yeaman & Whitlock

(2011) addressed the role of genetic architectures in maintaining multilocus adaptive phenotypes with gene flow using individual-based evolutionary simulations, and explored the effects of gene flow, selection, recombination, mutation and effect size on architectures that would develop after long periods of evolution. In summary, simulations indicate that (i) the number of loci contributing to divergence, and the distance among them tends to decrease with increasing gene flow, (ii) the number of loci contributing to divergence decreases with increasing recombination, and effect size tends to increase, (iii) clusters of loci experiencing a benefit of linkage tend to be spread over larger distances with stronger selection, (iv) individual mutations of small effect tend to cluster in groups of loci with larger effect and (v) the effect of mutation rate depends on effect size. Therefore, genetic architectures based on clusters of few loci of large effect are favoured when adaptation occurs with gene flow, as suggested by empirical work on hybridization in sunflowers (Rieseberg *et al.* 1999) and *Drosophila* (Noor *et al.* 2001), and inferred by previous two-locus theory on selection-migration-recombination (Lenormand & Otto 2000 and references therein). Bürger & Akerman (2011) and Akerman & Bürger (2014) supported these predictions using different models and assumptions, and additionally emphasized the importance of suppression of recombination in the maintenance of adaptive polymorphisms (Kirkpatrick & Barton 2006).

Most theoretical work revolves around island and continent-island models of migration and gene flow, and fewer studies have focused on adaptation with gene flow in continuous space. Although the balance between selection and gene flow remains the strongest determinant of whether an adaptive variant will be maintained in a population, in continuous space, such as along clines, gene flow decreases with distance among individuals, and this difference can affect the migration-selection balance. Endler (1977) summarized the main factors promoting maintenance of differentiation with gene flow along a cline: selection gradients or ecotones, short clines, dominance, selective differences among morphs but low competitive advantage of one morph over another and assortative mating. Early single-locus models (reviewed in Felsenstein 1976) often considered one aspect of genetic architecture, dominance, but neglected the relationship of a locus with its genomic background. Barton (1983) used a multilocus approach for the first time and found that with n number of loci under selection, linkage among loci strengthens selection and its opposition to gene flow by a factor of $\sqrt{n}$, which is in line with predictions based on island and continent-island models.

Genomic mechanisms maintaining adaptation despite gene flow

To understand how clusters of adaptive loci are maintained despite gene flow, the mechanisms underpinning their evolution, which are based on protection from the disrupting effects of recombination, are fundamental to identify. These mechanisms can be classified into four main categories (Yeaman 2013): (i) linkage with an already diverged locus, (ii) increased resistance to gene flow following secondary contact, (iii) competition among genetic architectures and (iv) competition among genomic architectures (including mechanisms that reduce or suppress recombination, see below). As previously discussed, linkage with an already diverged locus can favour the establishment of new advantageous mutations if they occur in the vicinity of a site already under selection (Bürger & Akerman 2011; Yeaman & Whitlock 2011; Feder *et al.* 2012b). Distance between loci is pivotal because it affects both the probability of establishment, through linkage, and the probability of resistance to gene flow, given that the rate of recombination increases with distance between loci (Fig. 4b). However, linkage disequilibrium (LD) depends on more than just distance between loci and can vary greatly across species, populations, and across regions of one individual genome (Reich *et al.* 2001; Flint-Garcia *et al.* 2003). For example, in certain regions of the genome LD in sticklebacks can extend to distances longer that 20 Mb (Hohenlohe *et al.* 2012), whereas in the collared (*Ficedula albicollis*) and the pied flycatchers (*F. hypoleuca*) average genome-wide LD is 17 kb, with peaks of LD corresponding to areas of high differentiation (Kawakami *et al.* 2014). Assuming equivalent phenotypic outcomes, selection will favour combinations of alleles that are more tightly linked (genetic architecture) because of their higher resistance to introgression of unfavourable alleles (Yeaman & Whitlock 2011; Fig. 4b). Similarly, genomic architectures that bring advantageous alleles together and are intrinsically characterized by low recombination rates will be favoured by selection (Kirkpatrick & Barton 2006).

In a simple single-trait single-locus scenario, the fate of an adaptation depends mostly on the tension between gene flow and selection (and drift), and recombination is not involved (Fig. 4c). If one locus controls pleiotropically several traits, the association among traits cannot be broken by recombination, and phenotypic differences are inherited as a single fixed module (see Solovieff *et al.* 2013 for a review on **pleiotropy**; Fig. 4c). Many studies have focused on the maladaptive effects of antagonistic pleiotropy that occur when a single locus controls two or more phenotypic traits with opposing fitness effects, thus hindering adaptation (e.g.

Otto 2004). If the co-varying traits lie on the same adaptive axis though (i.e. the slope of the ellipse describing the **G-matrix** is positive), pleiotropy can be beneficial in that it can even accelerate the adaptive response to selection. Adaptive pleiotropy was observed in *Arabidopsis thaliana* at the locus *FRI* (*FRIGIDA*), which is involved in local adaptation to drought by controlling flowering time, growth rate and water use efficiency (Lovell *et al.* 2013). Whether pleiotropy is antagonistic or adaptive, the association among the traits the gene controls cannot be broken by recombination, and the maintenance of a complex pleiotropic phenotype in a population follows predictions of single-locus theory (Lenormand 2002).

Supergenes are the most compelling example of genetic architectures resembling the behaviour of a pleiotropic gene. Although their existence has long been debated, the recent molecular characterization of supergenes proved them important in both adaptation and speciation (reviewed in Schwander *et al.* 2014; Thompson & Jiggins 2014). A supergene is fundamentally a cluster of adaptive loci that is inherited as though it was a single locus, and whose polymorphism is maintained in a population via negative frequency-dependent selection and/or spatially and temporally heterogeneous selection.

Thanks to the rapid and constant development of genomic methods, empirical studies in the area of speciation and adaptation with gene flow are perhaps ahead of theoretical predictions, and are offering precious insights into the genomic mechanisms and architectures that maintain adaptation with gene flow. In particular, these studies are unravelling the mechanisms that reduce or suppress recombination, including (i) cold spots of recombination, (ii) chromosomal rearrangements and (iii) epigenetic modifications (Fig. 4d).

Cold spots of recombination. Recombination rates are not even across genomes and depend on many factors. Cold spots are regions in the genome where the recombination rate is lower than average. What determines variation in recombination rates across the genome? First, recombination rates are a function of chromosome length, in that shorter chromosomes usually experience more recombination events per base pair than longer ones (Kaback *et al.* 1992; Kong *et al.* 2002). Second, recombination varies with genome sequence: recombination is lower in heterochromatic regions, areas of high GC content and stretches of polyA/polyT, and it is positively correlated with **CpG fraction** (Kong *et al.* 2002). Third, position on the chromosome matters. Heterochromatic regions are found in centromeres and telomeres and are characterized by low gene content, AT-rich sequences and high density of interspersed repeats. Centromeres in particular are apparently areas of low recombination and as such resistant to gene flow (Fig. 4d). For example, levels of introgression between the two subspecies of European rabbit (*Oryctolagus cuniculus cuniculus* and *O. c. algirus*) were investigated using four X-linked loci: two centromeric and two telomeric (Geraldes *et al.* 2006). High linkage disequilibrium, low levels of variability and little introgression between subspecies characterized the centromeric loci, whereas the telomeric loci showed opposite characteristics, suggesting reduced recombination and resistance to gene flow in proximity to the X chromosome centromere. Further support is provided by the two African malaria mosquito species *A. gambiae* and *A. coluzzii* (previously known as *M* and *S* forms), where gene flow is high despite apparent reproductive isolation. Turner *et al.* (2005) used genome-wide markers to identify the genomic areas of high differentiation potentially involved in reproductive isolation between the two species. They found only three differentiated regions between the two genomes, each containing fixed differences and no shared variation: of these, two lay in the proximity of centromeres. Other cold spots of recombination are gene-poor regions, which often occur in proximity to centromeres (Fig. 4d).

Chromosomal rearrangements. According to underdominance models, karyotypic differences are a major driver in speciation due to infertility or inviability of hybrids that inherit chromosomes differing in number or structure (White 1978). Rieseberg (2001) reviewed evidence undermining this view, including the observation that many chromosomal rearrangements, such as paracentric inversions, do not reduce fitness in hybrids, and proposed that chromosomal rearrangements promote differentiation more often by reducing gene flow through the suppression of recombination. This led to the development of the so called 'genic model' of speciation due to the emphasis given to the genes affected by the chromosomal arrangement rather than the arrangement per se. Chromosomal rearrangements reducing recombination include inversions, reciprocal translocations, fusions and fissions.

Chromosomal inversions could be easily investigated using cytological methods well before PCR and other genetic methods were developed. Extensive work conducted on *Drosophila* since the beginning of the 20th century, where several inversions were detected in many species, showed the importance of inversions in adaptation and promoted the development of several theories on the establishment and maintenance of inversions and inversion polymorphisms (reviewed in Hoffmann *et al.* 2004; Hoffmann & Rieseberg 2008). The first indication of a role for inversions in adaptation came

from the observation of seasonal changes in inversion polymorphism clines in *Drosophila* (Dobzhansky 1943). Later work reported the repeated and independent evolution of inversion polymorphism clines on three continents (Krimbas & Powell 1992). Additionally, research on *Drosophila* revealed that sympatric species show more differences involving inversions than do allopatric species (Noor *et al.* 2001), supporting that inversions play an important role in maintaining genetic differentiation in the face of gene flow. In reef fishes (Martinez *et al.* 2015) and estrildid finches (Hooper & Price 2015), chromosomal rearrangements fix faster in lineages with higher dispersal potential and gene flow, which is consistent with the prediction that gene flow favours fixation of chromosomal rearrangements that create and maintain associations among locally adapted loci (Kirkpatrick & Barton 2006). The reduction of recombination associated with inversions is what links inversions to adaptation with gene flow (Fig. 4d). However, low levels of recombination could still occur because recombination is really suppressed only in heterozygotes, and double crossovers and gene conversions that could potentially modify linkage disequilibrium among loci within an inversion, although rare, can occur (Andolfatto 2001). Factors other than reduced recombination can help explain the role of inversions in facilitating maintenance of genetic differences (Feder & Nosil 2009). Inversions can be important (i) when phenotypic differences are due to the break point rather than the sequence itself; (ii) when they capture many advantageous alleles together; and (iii) when gene flow following secondary contact is recent. Lastly, the apparent importance of inversions could be biased by the ease with which they are detected with cytological and genomic methods compared to other mechanisms.

While our focus is to understand the genomic mechanisms maintaining local adaptation with gene flow, and we provide evidence for an important role of inversions, Kirkpatrick & Barton (2006) addressed the relationship between inversions and local adaptation in the reverse manner, that is they investigated the mechanisms maintaining inversion polymorphisms. In their study, local adaptation seems the most plausible scenario among many examined. Most of the alternative scenarios focused on the balance among the fitness effects on the co-adapted alleles captured within an inversion (e.g. Dobzhansky 1947, 1951, 1970; Haldane 1957; Wasserman 1968), while other hypotheses were based on frequency-dependent and/or fluctuating selection (Wright & Dobzhansky 1946; Lewontin & White 1960; Álvarez-Castro & Alvarez 2005). Nonetheless, however it is investigated, the tight relationship between inversions and adaptation is widely and strongly supported.

Compared to the extensive literature on inversions, little is known about other chromosomal rearrangements such as fusions. For simplicity, we group all the chromosomal rearrangements that bring together previously unlinked loci and reduce recombination rates among them as 'fusions' (as in Guerrero & Kirkpatrick 2014); these include **Rb translocations**, **end-to-end fusions**, **tandem fusions**, **reciprocal translocations** and **fusions in holocentric chromosomes**. Fusions are different from inversions in that recombination is reduced not only in heterozygotes but also in homozygotes, and they are the chromosomal rearrangements that are more likely to cause sterility in F1 hybrids. If local adaptation favours the reduction of recombination rates, fusions should be favoured as much as inversions, if not more. Guerrero & Kirkpatrick (2014) addressed this question using population genetics simulations based on a two-locus continent-island model and a two-locus two-deme model. Their results are slightly different depending on the model used: under a continent-island model fusions evolve when recombination in heterozygotes is weaker than selection, whereas under a two-deme model fusions are always favoured to spread. The maintenance of fusions and fusion polymorphisms depends on recombination rates in homozygotes and heterozygotes, and on linkage disequilibrium among loci. As with inversions, the maintenance of fusions seems to be favoured by local adaptation, and local adaptation in turn could be maintained despite gene flow due to the lower recombination rates found in fusions.

In contrast, chromosomal rearrangements that increase recombination rates, such as chromosomal fissions, can be favoured in stressful or variable environments (Fig. 4a). This has been observed in at least two study systems, each involving a different chromosomal rearrangement: Rb translocation polymorphisms in grasshoppers (*Dichroplus pratensis*) and chromosomal fissions in blind mole rats (*Spalax* spp.). In South American acridid grasshoppers Rb translocation polymorphisms are frequent in central, stable and highly productive environments, and decline to monomorphism towards the southern extreme of the range, which is characterized by harsh, highly seasonal and variable habitats (Bidau & Martí 2002; Bidau *et al.* 2012). Several hypotheses have been suggested to explain this pattern, including the central-marginal hypothesis that predicts lower diversity at the margins of species' distributions (Kirkpatrick & Barton 1997). In *D. pratensis* translocation monomorphism at the range margin arguably allowed higher recombination rates, which in turn provided higher probabilities to draw an advantageous combination of loci for the local variable environmental conditions (Bidau & Martí 2002; Fig. 4a). In fact, with decreasing frequency of Rb translocations,

variability in several morphological traits increases and, with it, the chances to express a fit phenotype. In the blind mole rat, increasing numbers of chromosomes resulting from chromosomal fissions are correlated with ecological stress, in terms of aridity and climatic unpredictability, and have evolved independently in different geographic regions along environmental clines (Nevo *et al.* 1994; Nevo 2013). Although based on a different mechanism, an increase in recombination rates could potentially contribute to adaptation in the blind mole rat. Experimental tests and detailed information on levels of gene flow, the type of environmental heterogeneity and epistasis are necessary to reliably predict when recombination is beneficial (Lenormand & Otto 2000; Bürger & Akerman 2011).

Epigenetic modifications. Among the many mechanisms reducing recombination, epigenetic modifications such as DNA methylation, histone modifications and noncoding RNAs are the least understood, partly because the study of epigenetics only formally began in the early 1970s (Holliday 2006). Different organisms exhibit different levels of DNA methylation, from little or none in *Drosophila melanogaster* to very high levels in many plant genomes (Deaton & Bird 2011). Therefore, plants are good candidates to investigate epigenetics and its role in recombination reduction. In the plant model species *Arabidopsis thaliana* epigenetic modifications, especially DNA methylation, can explain differences in recombination in different parts of the genome (Mirouze *et al.* 2012; Yelina *et al.* 2012; Fig. 4d), or why centromeres are more resistant to recombination (reviewed in Henderson 2012; Fig. 4d). However, to the best of our knowledge, no study has addressed the specific role of epigenetic modifications in maintaining clusters of adapted alleles in the presence of gene flow. Although not focused on the maintenance of adaptation with gene flow, a recent study on migratory behaviour in the trout *Oncorhynchus mykiss* showed that a combination of epigenetic modifications and changes in sequence could explain phenotypic differences between migratory and nonmigratory individuals even though the genes underpinning such differences are not found in clusters (Baerwald *et al.* 2015; Table 2). These findings suggest that epigenetic modifications may have an important role in adaptation with gene flow, other than reducing recombination among adaptive loci.

Empirical evidence

Empirical studies on the genomics of adaptation followed a similar trend as the theory, although at a faster pace, moving from a few candidate genes, through hundreds of thousands of genome-wide markers (e.g. RNA-seq, RAD-seq), to whole genome sequencing and resequencing. Increasing genomic coverage allows analysis of the relationship among genes and genetic architecture, to test theoretical predictions and finally to understand the genomic mechanisms initiating and maintaining adaptation with gene flow: as previously discussed, the position in the genome, the distance among loci and recombination rates are important factors in the evolution of adaptation, and increasing genomic coverage helps unravel their relative roles. Due to decreasing costs of sequencing, the study of the genomics of adaptation has flourished in recent years, and with it more and more studies are addressing the role of gene flow in adaptation at the genomic level (Tables 1 and 2).

Marine organisms provide good systems to investigate the genomics of adaptation with gene flow. Physical barriers to gene flow are minimal within ocean basins: larvae are often passively dispersed by currents, communal areas of spawning and/or foraging facilitate gene flow among organisms from different geographic origins, and the effect of genetic drift is negligible because marine organisms generally have large population sizes. On the other hand, marine organisms often occupy wide geographical ranges and are therefore exposed to strong differences in environmental selective pressures. The purple sea urchin (*Strongylocentrotus purpuratus*), for example, is a highly dispersing marine invertebrate that occupies a wide latitudinal cline in the Pacific Ocean. Genome-wide scans (Pespeni *et al.* 2010) and analyses of putative adaptive genes along the cline (Pespeni & Palumbi 2013) showed that allele frequencies vary with temperature, indicating that spatial or temporal balancing selection has an important role in promoting and maintaining local adaptation despite high levels of gene flow (Table 2). Similar results were found in the Atlantic herring (*Clupea harengus*; Limborg *et al.* 2012), where spatially varying selection seems to maintain local adaptation despite high dispersal and gene flow likely occurring in communal feeding areas. However, in the European eel (*Anguilla anguilla*), spatially varying selection expected across subarctic and subtropical waters does not seem to counteract the homogenizing effect of gene flow (Pujolar *et al.* 2014): although selection is strong and acts within a single generation, panmixia at the spawning grounds nullifies any adaptation evolved in the previous generation. Pujolar *et al.* (2014) tested theoretical predictions regarding migration-selection balance and genetic architecture (Yeaman & Whitlock 2011) and found that loci putatively under selection were randomly scattered across the genome, a pattern that is consistent with panmixia and selection acting on single generations. Scattering of loci under selection across the genome rather than clus-

Table 2 Representative studies on the genomics of adaptation with gene flow. 'pops' stands for populations; 'genomic mechanism' refers to the mechanism that maintains local adaptation despite gene flow

Species	Trait	Genetic basis	Within or between pops?	Genomic mechanism	Reference
Steelhead/rainbow trout *Oncorhynchus mykiss*	Migratory vs. nonmigratory	Multiple loci across genome	Within and between pops	Sequence and epigenetics	Baerwald *et al.* (2015)
Malaria mosquito *Anopheles gambiae* S form	Adaptation to aridity	Several candidate genes associated with inversions	Clinal variation in inversions frequency	Inversion	Cheng *et al.* (2012)
Purple sea urchin *Strongylocentrotus purpuratus*	Adaptation to temperature	Several candidate genes	Clinal variation	Selection > migration	Pespeni *et al.* (2010); Pespeni & Palumbi (2013)
Arabidopsis lyrata	Adaptation to serpentine soils	Several candidate genes	Between pops	Selection > migration	Turner *et al.* (2010)
Butterfly *Heliconius numata*	Batesian mimicry	Supergene	Within and between pops	Inversions and tight linkage	Joron *et al.* (2011)
Yellow monkeyflower *Mimulus guttatus*	Different life-history strategies in response to soil moisture	Unknown	Between pops	Inversion	Lowry & Willis (2010); Twyford & Friedman (2015)
Threespine stickebacks *Gasterosteus aculeatus*	Marine vs. freshwater adaptation	Several traits	Between pops	Inversions Selection > migration	Jones *et al.* (2012a)
Arabidopsis thaliana	Drought resistance strategies (flowering time, growth rate, water use efficiency)	Locus *FRIGIDA*	Between pops	Pleiotropy	Lovell *et al.* (2013)
Deer mouse *Peromyscus maniculatus*	Cryptic coloration	Epistasis between *Mc1r* and *Agouti*	Within pops	Selection > migration Epistasis	Linnen *et al.* (2009, 2013)
Lizard *Sceloporus cowlesi*	Cryptic coloration	Single mutation in *Mc1r* + more unknown loci	Between pops	Selection > migration Dominance?	Rosenblum *et al.* (2004, 2010); Laurent *et al.* (2016)
Snail *Cepaea nemoralis*	Cryptic coloration	Unknown	Within pops	Tight linkage Supergene?	Richards *et al.* (2013)
Stick insect *Tinema cristinae*	Cryptic coloration	2 loci, unknown function	Within pops	Tight linkage Epistasis	Comeault *et al.* (2015)

tering however can also be observed when gene flow is low but selection is very strong, as in populations of *Arabidopsis lyrata* adapted to serpentine soils (Turner *et al.* 2010; Table 2).

An extreme case of adaptation with gene flow occurs when two or more morphs that mate randomly are maintained in a single population, and supergenes are the underlying genetic architecture of the adaptive trait. In the butterfly *Heliconius numata* for example, the supergene that controls wing mimicry seems to be maintained by both tight linkage among modules and a series of inversions that suppress recombination and thus prevent the expression of the maladapted combinations that would not be cryptic (Joron *et al.* 2011; Jones *et al.* 2012b; Table 2).

Empirical studies indicate that inversions are a common mechanism to maintain associations among adaptive alleles despite gene flow, and not only in supergenes (Hoffmann *et al.* 2004; Hoffmann & Rieseberg 2008; Schwander *et al.* 2014; Table 2). Although inversions are common, the loci captured by inversions are mostly yet to be determined, and hence the physiological mechanisms underpinning local adaptation are often unknown. An example of a well-characterized inversion, although associated with sexual selection and speciation rather than adaptation, is provided by Poelstra *et al.* (2014), who used whole genome resequencing to study hybridization between two crow subspecies (*Corvus (corone) corone and C. (corone) cornix*) that came into secondary contact after differentiating in separate glacial refugia. The interesting question is how the subspecies maintained phenotypic differences in coloration despite gene flow and lack of neutral genetic differentiation in the hybrid zone. The authors found a genomic region where fixed differences in genes mostly involved in pigmentation and visual perception were tightly linked, most likely due to an inversion in the region, corroborating a role for colour-mediated sexual selection (Poelstra *et al.* 2014).

The direct contribution of an inversion to local adaptation was determined for the first time using a field experiment and QTL mapping in the monkeyflower (*Mimulus guttatus*), even though the genes captured by the inversion were not characterized (Lowry & Willis 2010; Table 2). In North America, perennial and annual ecotypes of the monkeyflower evolved in response to differences in soil moisture. Twyford & Friedman (2015) demonstrated that high gene flow between ecotypes homogenized the collinear part of the genome, except in a chromosomal inversion responsible for the two life-history strategies. Similarly, several inversions potentially involved in adaptation with gene flow were repeatedly found among marine-freshwater population pairs of threespine sticklebacks (*Gasterosteus aculeatus*) using whole genome resequencing (Jones *et al.* 2012a; Table 2). One of the inversions identified in this study contains a voltage-gated potassium channel gene, *KCNH4*. This gene is potentially involved in the transition between marine and freshwater environments, and its transcription seems to be affected by the inversion orientation and breakpoint. These studies provide empirical evidence for two mechanisms that render inversions important in local adaptation: capture of two or more adapted, but not necessarily co-adapted, alleles together (Kirkpatrick & Barton 2006), or change in phenotype due to inversion breakpoints rather than differences in sequence (Feder & Nosil 2009).

Future directions

In 1984, Endler wrote 'A major problem in this subject [natural selection] is that there is multiplicity of meanings for the same terms, and the same term means different things to different people.' (Endler 1986). After more than 30 years, we find that this still holds true. For example, although we did our best to guide the reader through the many uses of 'migration' in the scientific literature, it is evident that some confusion could be avoided, especially considering that gene flow, dispersal and migration (as in the movements that many animals undertake seasonally or periodically), are different phenomena that can co-occur in the same study system.

Several aspects of the genomics of local adaptation with gene flow will see or need a better understanding in the near future.

- Regarding the role of gene flow in the origin of adaptation we encourage the adoption of a genomic approach in studying adaptation from new mutations, and in assessing levels of standing genetic variation. For example, candidate gene approaches, where relatively short regions of the genome associated with an adaptive trait are analysed, can potentially give an either incomplete or incorrect picture of the process of adaptation. Also, assessing genome-wide levels of standing genetic variation with a distinction of putatively neutral and putatively adaptive variation could be used as a monitoring tool to study and predict the effect of environmental change on levels of standing genetic variation.
- We have addressed here for the first time that the strength of selection against gene flow can depend on the genomic architecture of adaptive traits. In the future, we will develop a quantitative model expanding on the hypotheses here presented and, at the same time, hope for empirical studies explicitly

addressing selection on gene flow and the role of genetic architectures.

- With increasing computational resources, more complex models and simulations that include more parameters and span entire chromosomes and genomes could be run in the near future. One potential improvement of theoretical models entails the inclusion of LD: even though currently available models stress the importance of recombination and distance among loci in the establishment and maintenance of adaptive alleles in a population (Bürger & Akerman 2011; Yeaman & Whitlock 2011; Feder *et al.* 2012b), the explicit parameterization of LD calculated from genomic data could help understand the role of genetic and genomic architectures in adaptation.
- We found that in some genomic studies of adaptation gene flow was present but hardly discussed, even when it seemed to exert a non negligible role in the pattern observed. We warn that to neglect the role of gene flow could cause a misinterpretation of results.
- Despite widespread evidence for a role of chromosomal rearrangements in adaptation since the beginning of the 20th century, the genes potentially underlying adaptations within these regions remain poorly characterized (but see Poelstra *et al.* 2014). Unveiling the contents of rearranged genomic regions would deepen our understanding of how genes get 'captured' together, and of the relationship between adaptive and nonadaptive genes therein. We encourage research in this direction, even though we are aware that this is happening as we write.
- In contrast to the wealth of studies on chromosomal rearrangements, such as inversions, that suppress recombination, little is known about the effects of increased recombination, which, as we have discussed, can be beneficial in stressful environments. Studies addressing this gap are needed.
- The study of epigenetics is relatively new but an area of active research as suggested by a special issue of Molecular Ecology (2016) on 'Epigenetic Studies in Ecology and Evolution'. To comprehend its role in the genomics of local adaptation with gene flow, we need a better understanding of epigenetic effects on recombination rates and on the evolution of adaptive traits.

Conclusions and broader implications

Theoretical models and empirical evidence show that gene flow does not always disrupt adaptation. The main determinants of whether a new beneficial allele is established and maintained are strength of selection and gene flow, but the genetic architecture of a trait and the genomic landscape can decisively alter the migration-selection balance. Gene flow can even enhance adaptation by augmenting genetic variation, especially if selection is spatially and/or temporally variable.

Understanding how adaptation with gene flow occurs can help predict how organisms will cope with ongoing and future environmental change. When faced with environmental change organisms can move to better conditions, adjust in the short term via **phenotypic plasticity**, or evolve genetic adaptations. Ultimately, genetic change is what grants long-term survival in populations facing a changing environment. Because the extent of phenotypic plasticity can have a genetic component, genetic change is also important for short-term responses to environmental fluctuations. Dispersal as a strategy to escape local unfavourable conditions can re-connect populations and species differentiated in allopatry, potentially causing gene flow among them and thus affecting patterns of local adaptation. Therefore, the interplay of genetic variation and gene flow is fundamental when predicting the potential of a species to adapt. This is particularly important for species conservation in relation to anthropogenic change: habitat alteration, climate change, direct introductions and modification of reproductive barriers can change barriers to gene flow by either creating new obstacles or destroying existing ones (reviewed in Crispo *et al.* 2011). Hamilton & Miller (2016) recently discussed examples where hybridization and introgression between populations and species facilitated an adaptive response to environmental change, arguing that human-mediated gene flow may be a practice that conservation managers need to consider, especially in a climate change scenario. Based on existing knowledge however, predicting how increasing and decreasing gene flow will affect local adaptation and the potential to evolve in response to environmental change is not easy. Nevertheless, the development of more sophisticated simulation-based studies and a focus on gene flow in research on the genomics of adaptation would provide valuable insights into adaptation and speciation, and therefore conservation as well.

Acknowledgements

We thank Rob Colautti, Sarah Yakimowski and Sam Yeaman for valuable discussions that greatly improved an earlier version of the manuscript. We also thank Haley Kenyon, Dave Ensing and the members of the Friesen Lab and of the Molecu-

lar Ecology journal group for helpful comments and discussions on the manuscript and/or on studies discussed in this review. Special thanks to François Blanquart for sharing data for Fig. 3. Thanks to the Review Editor Louis Bernatchez for inviting this review, and to Nick Barton and two anonymous referees for helpful and constructive comments. This work was supported by Queen's University (Queen's Graduate Research and International Tuition Awards to AT) and by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant 203–320 to VLF).

References

Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Aeschbacher S, Bürger R (2014) The effect of linkage on establishment and survival of locally beneficial mutations. *Genetics*, **197**, 317–336.

Akerman A, Bürger R (2014) The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model. *Journal of Mathematical Biology*, **68**, 1135–1198.

Álvarez-Castro JM, Alvarez G (2005) Models of general frequency-dependent selection and mating-interaction effects and the analysis of selection patterns in *Drosophila* inversion polymorphisms. *Genetics*, **170**, 1167–1179.

Andolfatto P (2001) Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. *Molecular Biology and Evolution*, **18**, 279–290.

Baerwald MR, Meek MH, Stephens MR *et al.* (2015) Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. *Molecular Ecology*, doi:10.1111/mec.13231.

Balkau BJ, Feldman MW (1973) Selection for migration modification. *Genetics*, **74**, 171–174.

Barrett R, Hoekstra H (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767–780.

Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, **23**, 38–44.

Barton NH (1983) Multilocus clines. *Evolution*, **37**, 454–471.

Barton NH, Bengtsson BO (1986) The barrier to genetic exchange between hybridising populations. *Heredity*, **57**, 357–376.

Bidau CJ, Martí DA (2002) Geographic distribution of Robertsonian fusions in *Dichroplus pratensis* (Melanoplinae, Acrididae): the central-marginal hypothesis reanalysed. *Cytogenetic and Genome Research*, **96**, 66–74.

Bidau CJ, Miño CI, Castillo ER, Martí DA (2012) Effects of abiotic factors on the geographic distribution of body size variation and chromosomal polymorphisms in two neotropical grasshopper species (Dichroplus: Melanoplinae: Acrididae). *Psyche: A Journal of Entomology*, Article ID 863947, **2012**, 1–11.

Billiard S, Lenormand T (2005) Evolution of migration under kin selection and local adaptation. *Evolution*, **59**, 13–23.

Blanquart F, Gandon S (2011) Evolution of migration in a periodically changing environment. *American Naturalist*, **177**, 188–201.

Blanquart F, Kaltz O, Nuismer SL, Gandon S (2013) A practical guide to measuring local adaptation. *Ecology Letters*, **16**, 1195–1205.

Bullini L (1994) Origin and evolution of animal hybrid species. *Trends in Ecology and Evolution*, **9**, 422–426.

Bulmer MG (1972) The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genetical Research*, **19**, 17–25.

Bürger R, Akerman A (2011) The effects of linkage and gene flow on local adaptation: a two-locus continent-island model. *Theoretical Population Biology*, **80**, 272–288.

Chain FJJ, Feulner PGD, Panchal M *et al.* (2014) Extensive copy-number variation of young genes across stickleback populations. *PLoS Genetics*, **10**, e1004830.

Cheng C, White BJ, Kamdem C *et al.* (2012) Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach. *Genetics*, **190**, 1417–1432.

Clarkson CS, Weetman D, Essandoh J *et al.* (2014) Adaptive introgression between *Anopheles* sibling species eliminates a major genomic island but not reproductive isolation. *Nature Communications*, **5**, 1–10.

Comeault AA, Flaxman SM, Riesch R *et al.* (2015) Selection on a genetic polymorphism counteracts ecological speciation in a stick insect. *Current Biology*, **25**, 1975–1981.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Crispo E, Moore J-S, Lee-Yaw JA, Gray SM, Haller BC (2011) Broken barriers: Human-induced changes to gene flow and introgression in animals. *BioEssays*, **33**, 508–518.

De Carvalho D, Ingvarsson PK, Joseph J *et al.* (2010) Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molecular Ecology*, **19**, 1638–1650.

Deaton A, Bird A (2011) CpG islands and the regulation of transcription. *Genes & Development*, **25**, 1010–1022.

Dobzhansky T (1943) Genetics of natural populations IX. Temporal changes in the composition of populations of *Drosophila pseudoobscura*. *Genetics*, **28**, 162–186.

Dobzhansky T (1947) Genetics of natural populations. XIV. A response of certain gene arrangements in the third chromosome of *Drosophila pseudoobscura* to natural selection. *Genetics*, **32**, 142–160.

Dobzhansky T (1951) *Genetics and the Origin of Species*, 3 edn. Columbia University Press, New York, New York.

Dobzhansky T (1970) *Genetics of the Evolutionary Process*. Columbia University Press, New York, New York.

Duvaux L, Geissmann Q, Gharbi K *et al.* (2015) Dynamics of copy number variation in host races of the pea aphid. *Molecular Biology and Evolution*, **32**, 63–80.

Ellstrand NC (2014) Is gene flow the most important evolutionary force in plants? *American Journal of Botany*, **101**, 737–753.

Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ.

Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ.

Fan S, Meyer A (2014) Evolution of genomic structural variation and genomic architecture in the adaptive radiations of African cichlid fishes. *Frontiers in Genetics*, **5**, 1–9.

Feder JL, Nosil P (2009) Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? *Evolution*, **63**, 3061–3075.

Feder JL, Berlocher SH, Roethele JB *et al.* (2003) Allopatric genetic origins for sympatric host-plant shifts and race

formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 10314–10319.
Feder JL, Xie X, Rull J *et al.* (2005) Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in Rhagoletis. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(suppl 1), 6573–6580.
Feder JL, Egan SP, Nosil P (2012a) The genomics of speciation-with-gene-flow. *Trends in Genetics*, **28**, 342–350.
Feder JL, Gejji R, Yeaman S, Nosil P (2012b) Establishment of new mutations under divergence and genome hitchhiking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 461–474.
Felsenstein J (1976) The theoretical population genetics of variable selection and migration. *Annual Review of Genetics*, **10**, 253–280.
Feulner PGD, Chain FJJ, Panchal M *et al.* (2013) Genome-wide patterns of standing genetic variation in a marine population of three-spined sticklebacks. *Molecular Ecology*, **22**, 635–649.
Flagel LE, Willis JH, Vision TJ (2014) The standing pool of genomic structural variation in a natural population of *Mimulus guttatus*. *Genome Biology and Evolution*, **6**, 53–64.
Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, **54**, 357–374.
Gandon S (2002) Local adaptation and the geometry of host-parasite coevolution. *Ecology Letters*, **5**, 246–256.
Geraldes A, Ferrand N, Nachman MW (2006) Contrasting patterns of introgression at X-linked loci across the hybrid zone between subspecies of the European Rabbit (*Oryctolagus cuniculus*). *Genetics*, **173**, 919–933.
Griswold CK (2006) Gene flow's effect on the genetic architecture of a local adaptation and its consequences for QTL analyses. *Heredity*, **96**, 445–453.
Guerrero RF, Kirkpatrick M (2014) Local adaptation and the evolution of chromosome fusions. *Evolution*, **68**, 2747–2756.
Haldane JBS (1930) Theoretical genetics of autopolyploids. *Journal of Genetics*, **22**, 359–372.
Haldane JBS (1957) The cost of natural selection. *Journal of Genetics*, **55**, 511–524.
Hamilton JA, Miller JM (2016) Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology*, **30**, 33–41.
Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, **22**, 4606–4618.
Hejase HA, Liu KJ (2016) Mapping the genomic architecture of adaptive traits with interspecific introgressive origin: a coalescent-based approach. *BMC Genomics*, **17**, 41–57.
Henderson IR (2012) Control of meiotic recombination frequency in plant genomes. *Current Opinion in Plant Biology*, **15**, 556–561.
Hey J (2006) Recent advances in assessing gene flow between diverging populations and species. *Current Opinion in Genetics & Development*, **16**, 592–596.
Hoffmann AA, Rieseberg LH (2008) Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics*, **39**, 21–42.
Hoffmann A, Sgro C, Weeks A (2004) Chromosomal inversion polymorphisms and adaptation. *Trends in Ecology and Evolution*, **19**, 482–488.
Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **367**, 395–408.
Holliday R (2006) Epigenetics: a historical overview. *Epigenetics*, **1**, 76–80.
Hooper DM, Price TD (2015) Rates of karyotypic evolution in estrildid finches differ between island and continental clades. *Evolution*, **69**, 890–903.
Huang Y, Stinchcombe JR, Agrawal AF (2015) Quantitative genetic variance in experimental fly populations evolving with or without environmental heterogeneity. *Evolution*, **69**, 2735–2746.
Huerta-Sánchez E, Jin X, Bianba Z *et al.* (2014) Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, **512**, 194–197.
Jones FC, Grabherr MG, Chan YF *et al.* (2012a) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.
Jones RT, Salazar PA, ffrench-Constant RH, Jiggins CD, Joron M (2012b) Evolution of a mimicry supergene from a multilocus architecture. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 316–325.
Joron M, Frezal L, Jones RT *et al.* (2011) Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature*, **477**, 203–206.
Kaback DB, Guacci V, Barber D, Mahon JW (1992) Chromosome size-dependent control of meiotic recombination. *Science*, **256**, 228–232.
Karasov T, Messer PW, Petrov DA (2010) Evidence that adaptation in *Drosophila* is not limited by mutation at single sites. *PLoS Genetics*, **6**, e1000924.
Kawakami T, Backström N, Burri R *et al.* (2014) Estimation of linkage disequilibrium and interspecific gene flow in Ficedula flycatchers by a newly developed 50k single-nucleotide polymorphism array. *Molecular Ecology Resources*, **14**, 1248–1260.
Kim M, Cui ML, Cubas P *et al.* (2008) Regulatory genes control a key morphological and ecological trait transferred between species. *Science*, **322**, 1116–1119.
Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist*, **150**, 1–23.
Kirkpatrick M, Barton NH (2006) Chromosome inversions, local adaptation and speciation. *Genetics*, **173**, 419–434.
Kong A, Gudbjartsson DF, Sainz J *et al.* (2002) A high-resolution recombination map of the human genome. *Nature Genetics*, **31**, 241–247.
Krimbas CB, Powell JR (1992) *Drosophila Inversion Polymorphism*. CRC Press, London.
Laurent S, Pfeifer SP, Settles M *et al.* (2016) The population genomics of rapid adaptation: disentangling signatures of selection and demography in white sands lizards. *Molecular Ecology*, **25**, 306–323.
Lee Y, Marsden CD, Norris LC *et al.* (2013) Spatiotemporal dynamics of gene flow and hybrid fitness between the M and S forms of the malaria mosquito, *Anopheles gambiae*. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 19854–19859.
Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, **17**, 183–189.

Lenormand T, Otto SP (2000) The evolution of recombination in a heterogeneous environment. *Genetics*, **156**, 423–438.

Lewontin RC, White MJD (1960) Interaction between inversion polymorphisms of two chromosome pairs in the grasshopper, *Moraba scurra*. *Evolution*, **14**, 116–129.

Limborg MT, Helyar SJ, De Bruyn M *et al.* (2012) Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Molecular Ecology*, **21**, 3686–3703.

Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE (2009) On the origin and spread of an adaptive allele in deer mice. *Science*, **325**, 1095–1098.

Linnen CR, Poh Y-P, Peterson BK *et al.* (2013) Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, **339**, 1312–1316.

Liu KJ, Dai J, Truong K, Song Y, Kohn MH, Nakhleh L (2014) An HMM-based comparative genomic framework for detecting introgression in eukaryotes. *PLOS Computational Biology*, **10**, e1003649.

Liu KJ, Steinberg E, Yozzo A, Song Y, Kohn MH, Nakhleh L (2015) Interspecific introgressive origin of genomic diversity in the house mouse. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 196–201.

Lovell JT, Juenger TE, Michaels SD *et al.* (2013) Pleiotropy of *FRIGIDA* enhances the potential for multivariate adaptation. *Proceedings of The Royal Society B: Biological Sciences*, **280**, 20131043.

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, **8**, e1000500.

Mallet J (2007) Hybrid speciation. *Nature*, **446**, 279–283.

Martin NH (2005) Detecting adaptive trait introgression between *Iris fulva* and *I. brevicaulis* in highly selective field conditions. *Genetics*, **172**, 2481–2489.

Martinez PA, Zurano JP, Amado TF *et al.* (2015) Chromosomal diversity in tropical reef fishes is related to body size and depth range. *Molecular Phylogenetics and Evolution*, **93**, 1–4.

Mayr E (1942) *Systematics and the Origin of Species, From the Viewpoint of a Zoologist*. Columbia University Press, New York, New York.

Mirouze M, Lieberman-Lazarovich M, Aversano R *et al.* (2012) Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 5880–5885.

Monnahan PJ, Colicchio J, Kelly JK (2015) A genomic selection component analysis characterizes migration-selection balance. *Evolution*, **79**, 1713–1727.

Nachman MW, Payseur BA (2012) Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 409–421.

Nevo E (2013) Stress, adaptation, and speciation in the evolution of the blind mole rat, *Spalax*, in Israel. *Molecular Phylogenetics and Evolution*, **66**, 515–525.

Nevo E, Filippucci MG, Redi C, Korol A, Beiles A (1994) Chromosomal speciation and adaptive radiation of mole rats in Asia Minor correlated with increased ecological stress. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 8160–8164.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 12084–12088.

Norris LC, Main BJ, Lee Y *et al.* (2015) Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 815–820.

Nosil P, Crespi BJ, Sandoval CP (2002) Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature*, **417**, 440–443.

Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299–304.

Olson-Manning CF, Wagner MR, Mitchell-Olds T (2012) Adaptive evolution: evaluating empirical support for theoretical predictions. *Nature Reviews Genetics*, **13**, 867–877.

Otto SP (2004) Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proceedings of the Royal Society of London B: Biological Sciences*, **271**, 705–714.

Pespeni MH, Palumbi SR (2013) Signals of selection in outlier loci in a widely dispersing species across an environmental mosaic. *Molecular Ecology*, **22**, 3580–3597.

Pespeni MH, Oliver TA, Manier MK, Palumbi SR (2010) Restriction site tiling analysis: accurate discovery and quantitative genotyping of genome-wide polymorphisms using nucleotide arrays. *Genome Biology*, **11**, R44.

Poelstra JW, Vijay N, Bossu CM *et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, **344**, 1410–1414.

Pujolar JM, Jacobsen MW, Als TD *et al.* (2014) Genome-wide single-generation signatures of local selection in the panmictic European eel. *Molecular Ecology*, **23**, 2514–2528.

Reich DE, Cargill M, Bolk S *et al.* (2001) Linkage disequilibrium in the human genome. *Nature*, **411**, 199–204.

Richards PM, Liu MM, Lowe N, Davey JW, Blaxter ML, Davison A (2013) RAD-Seq derived markers flank the shell colour and banding loci of the *Cepaea nemoralis* supergene. *Molecular Ecology*, **22**, 3077–3089.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution*, **16**, 351–358.

Rieseberg L (2011) Adaptive introgression: the seeds of resistance. *Current Biology*, **21**, R581–R583.

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, **152**, 713–727.

Rosenblum EB, Hoekstra HE, Nachman MW (2004) Adaptive reptile color variation and the evolution of the *Mc1r* gene. *Evolution*, **58**, 1794–1808.

Rosenblum EB, Rompler H, Schoneberg T, Hoekstra HE (2010) Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 2113–2117.

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807–820.

Schwander T, Libbrecht R, Keller L (2014) Supergenes and complex phenotypes. *Current Biology*, **24**, R288–R294.

Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, **19**, 198–207.

Seehausen O, Butlin RK, Keller I *et al.* (2014) Genomics and the origin of species. *Nature Reviews Genetics*, **15**, 176–192.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW (2013) Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics*, **14**, 483–495.

Song Y, Endepols S, Klemann N *et al.* (2011) Adaptive introgression of anticoagulant rodent poison resistance by hybridization between Old World mice. *Current Biology*, **21**, 1296–1301.

Spichtig M, Kawecki TJ (2004) The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *American Naturalist*, **164**, 70–84.

Stapley J, Reger J, Feulner PGD *et al.* (2010) Adaptation genomics: the next generation. *Trends in Ecology & Evolution*, **25**, 705–712.

Thompson MJ, Jiggins CD (2014) Supergenes and their role in evolution. *Heredity*, **113**, 1–8.

Tripet F, Wright J, Cornel A *et al.* (2007) Longitudinal survey of knockdown resistance to pyrethroid (KDR) in Mali, West Africa, and evidence of its emergence in the Bamako form of *Anopheles gambiae* s.s. *American Journal of Tropical Medicine and Hygiene*, **76**, 81–87.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, **3**, e285.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics*, **42**, 260–263.

Twyford AD, Friedman J (2015) Adaptive divergence in the monkeyflower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution*, **69**, 1476–1486.

Wasserman M (1968) Recombination-induced chromosomal heterosis. *Genetics*, **58**, 125–139.

White MJD (1978) Chain processes in chromosomal speciation. *Systematic Biology*, **27**, 285–298.

Whitney KD, Randell RA, Rieseberg LH (2006) Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *American Naturalist*, **167**, 794–807.

Whitney KD, Broman KW, Kane NC, Hovick SM, Randell RA, Rieseberg LH (2015) Quantitative trait locus mapping identifies candidate alleles involved in adaptive introgression and range expansion in a wild sunflower. *Molecular Ecology*, **24**, 2194–2211.

Wolf JBW, Lindell J, Backström N (2010) Speciation genetics: current status and evolving approaches. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 1717–1733.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.

Wright S, Dobzhansky T (1946) Genetics of natural populations. XII. Experimental reproduction of some of the changes caused by natural selection in certain populations of *Drosophila pseudoobscura*. *Genetics*, **31**, 125.

Yang J, Benyamin B, McEvoy BP *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, **42**, 565–569.

Yeaman S (2013) Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, E1743–E1751.

Yeaman S (2015) Local adaptation by small-effect alleles. *American Naturalist*, **186**, S74–S89.

Yeaman S, Jarvis A (2006) Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proceedings of the Royal Society of London B: Biological Sciences*, **273**, 1587–1593.

Yeaman S, Otto SP (2011) Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution*, **65**, 2123–2129.

Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-selection balance. *Evolution*, **65**, 1897–1911.

Yelina NE, Choi K, Chelysheva L *et al.* (2012) Epigenetic remodeling of meiotic crossover frequency in *Arabidopsis thaliana* DNA methyltransferase mutants. *PLoS Genetics*, **8**, e1002844.

A.T. and V.L.F. outlined the review. A.T. wrote the manuscript with the help of V.L.F.. A.T. prepared all figures and tables.

Data accessibility

No data are associated with the manuscript.