Introduction

Speciation genetics: current status and evolving approaches

Jochen B. W. Wolf1,*, Johan Lindell1 and Niclas Backström1,2

1Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden
2Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA

The view of species as entities subjected to natural selection and amenable to change put forth by Charles Darwin and Alfred Wallace laid the conceptual foundation for understanding speciation. Initially marred by a rudimental understanding of hereditary principles, evolutionists gained appreciation of the mechanistic underpinnings of speciation following the merger of Mendelian genetic principles with Darwinian evolution. Only recently have we entered an era where deciphering the molecular basis of speciation is within reach. Much focus has been devoted to the genetic basis of intrinsic postzygotic isolation in model organisms and several hybrid incompatibility genes have been successfully identified. However, concomitant with the recent technological advancements in genome analysis and a newfound interest in the role of ecology in the differentiation process, speciation genetic research is becoming increasingly open to non-model organisms. This development will expand speciation research beyond the traditional boundaries and unveil the genetic basis of speciation from manifold perspectives and at various stages of the splitting process. This review aims at providing an extensive overview of speciation genetics. Starting from key historical developments and core concepts of speciation genetics, we focus much of our attention on evolving approaches and introduce promising methodological approaches for future research venues.

Keywords: selection; reproductive isolation; next generation sequencing; gene expression; hybrid; speciation research in the post-genomic era

1. INTRODUCTION

The formation of new species lies at the very heart of evolutionary biology. Indeed, the vast diversity of life on Earth can only be explained by speciation, a process that continuously generates independently evolving lineages. One and a half centuries ago, this ‘mystery of mysteries’ was subject to bold speculation, as the philosopher John Herschel communicates in a letter to Charles Lyell (Herschel 1836). Several years later Charles Robert Darwin and Alfred Russell Wallace made a considerable contribution to demystify the origin of new species and laid the foundation for evolutionary biology by suggesting natural selection and common ancestry as cornerstones of organismic evolution (Darwin & Wallace 1858). Yet, despite its title, Darwin’s opus ‘On the Origin of Species by Means of Natural Selection’ (Darwin 1859) did not focus on the rise of new species, but instead emphasized natural selection as a mechanism for the adaptive change of populations in response to the prevailing conditions. Furthermore, Darwin highlighted the transition from populations to species as a gradual continuum (Mallet 2008) without formally treating the isolation factors that reduce gene flow among populations (Mayr 1942; see also Barraclough 2010; Mallet 2010). However, Darwin lacked an understanding of the genetic basis of heredity. This eluded evolutionary biology until four decades later, following the rediscovery of Mendelian principles of inheritance in 1900. Yet, it was not until the theoretical framework of population genetics was amalgamated with Darwinian evolution and gave rise to the Modern Synthesis during the 1930s that the species problem was seriously considered (Dobzhansky 1937; Mayr 1942). This fusion put a premium on a population genetic viewpoint and hence allowed examining the speciation process from a genic perspective. By explicit modelling, the Modern Synthesis and influential derivatives such as the Neutral and Nearly Neutral Theory (Kimura 1983; Ohta 1992) conceptually reduced the evolutionary process to several tractable parameters like mutation, drift, selection and recombination, which can be estimated with empirical data. From the excitement about modes and mechanisms of speciation that characterized the Modern Synthesis, a consensus had emerged that speciation represented complete reproductive isolation of biological species, which could most likely be acquired through

* Author for correspondence (jochen.wolf@ebc.uu.se).

One contribution of 11 to a Theme Issue ‘Genomics of speciation’.
geographic isolation (Mayr 1942). Speciation research fell into a state of dormancy and largely revolved around the relative importance of different geographic speciation scenarios for several decades. However, one and a half centuries after ‘On the Origin of Species’ there is once again much excitement about speciation. Over the last two decades, the spectrum of researchers with an interest in speciation has expanded considerably. Given the breadth of scientific disciplines that contribute to contemporary speciation research, this review is naturally limited in focus and will capitalize on research concerned with the genetics of speciation.

Genetic approaches have always been central to speciation research, but despite significant progress over the last years in speciation genetic research, many fundamental questions about the molecular basis of the splitting process await to be answered. Which genetic elements are of particular relevance to speciation? How many loci are involved, how large is the effect of a specific locus and how important is epistasis or pleiotropy? Where in the genome are the determinants located and what is the importance of the genomic landscape? What is the role of recombination, mutation, chromosomal rearrangements, gene conversion and other molecular forces? How does divergence in gene expression compare with structural changes? How crucial is sex-linkage? Are different functional classes of genes relevant at different stages of the speciation process? What is the role of natural selection and how can we best detect its genetic footprints?

Finding full answers to even a subset of these questions over a broad taxonomic range will probably be wishing for too much. Still, focusing on several well-chosen speciation models, we may come close to an educated guess. Being empiricists, we will focus on the empirical side by highlighting where recent advancements have been and are expected to be made and only mention the relevant theoretical work in passing. We provide the conceptual background on general key concepts where deemed necessary. As much of recent research has focused on the role of natural selection in speciation, the review reflects this bias. We start by addressing research concerned with the genetic basis of intrinsic postzygotic isolation, which has been the traditional stronghold of speciation genetics. We then expand the framework of speciation genetics into an ecological context and try to infer how the field will be transformed, as novel genomic tools allow for detailed analysis of organisms, where previously no genomic resources have been available.

2. THE ROOTS OF SPECIATION GENETICS

Speciation involves the build-up of reproductive isolating barriers between diverging populations which are most palpable in malfunctional heterospecific hybrids. The evolution of postzygotic isolation giving rise to hybrid problems posed an important challenge to Darwinism: how can natural selection allow the production of maladaptive phenotypes and unfit hybrids? Most speciation theories have subsequently focused on resolving this dilemma. The theories can be divided into two groups corresponding to distinct forms of postzygotic isolation. In extrinsic postzygotic isolation, hybrid phenotypes fare poorly in their interaction with the environment, falling between the niches of the parental phenotypes (Schluter & Conte 2009). In intrinsic postzygotic isolation, hybrids are unfit because they suffer inherent developmental defects, resulting in partial or complete sterility or inviability (Orr & Turelli 2001). There are numerous examples from nature where both extrinsic and intrinsic postzygotic isolation appear to be at work (e.g. Rogers & Bernatchez 2006; Rieseberg & Willis 2007; Fuller 2008). Some recent work suggests that extrinsic postzygotic isolation may be more common and more important than intrinsic postzygotic isolation, specifically in the early stages of divergence (Schluter 2009; Schluter & Conte 2009; Johannesson et al. 2010). However, given the relative ease with which the genetic basis of reproductive isolation can be evaluated in laboratory model organisms, along with a great amount of theoretical work on the topic, most of what we know about the genetics of speciation deals with intrinsic postzygotic isolation.

Four kinds of genetic problems have been identified as the likely causes of intrinsic hybrid difficulties: ploidy levels, chromosomal rearrangements, genic incompatibilities and interaction between nuclear genomes and endosymbionts, which can arguably be regarded as a special case of the latter (Rieseberg 2001; Coyne & Orr 2004; Hoffmann & Rieseberg 2008). These mechanisms vary in importance depending on the system. Ploidy levels, for example, are of major importance in plant speciation (Rieseberg & Willis 2007), where chromosomal rearrangements has also been extensively discussed (White 1969; Hoffmann & Rieseberg 2008). In research on genetic model organisms such as Drosophila (Kulathinal et al. 2009), there has also been much interest in chromosomal rearrangements (Noor et al. 2001; Noor & Feder 2006). Nevertheless, it appears that genic incompatibilities may be the most important cause of intrinsic postzygotic isolation; they play a common role in both hybrid sterility and inviability, and affect both animals and plants (Orr & Turelli 2001; Coyne & Orr 2004). Genic incompatibilities in hybrids most commonly involve between-locus interactions; an allele at one locus from one of the parental species does not interact well with an allele at another locus from the other parental species (Turelli & Orr 2000; Coyne & Orr 2004). This in line with early suggestions of the Modern Synthesis that negative epistatic interactions among genes constitute a plausible mechanism that can cause hybrid sterility and inviability. The ‘Dobzhansky–Muller’ model, initially discussed by Bateson (1909) and later developed by Dobzhansky (1937) and Muller (1942) (we will refer to it as the Bateson–Dobzhansky–Muller (BDM) model throughout), was proposed as a solution to the problem of how hybrid sterility can evolve without selection opposing any intermediate step. In short, allopatric populations that evolve independently each accumulate different mutations that contribute to genetic differences between the populations. Subjected to evolutionary forces including genetic drift and natural selection, specific mutations may function well in the genetic make-up of their particular population.

Phil. Trans. R. Soc. B (2010)
However, because alleles from different populations have not been tested together, such mutations are on average less likely to function with alleles of a different ancestral background in hybrid individuals. Hybrid sterility or inviability may therefore simply evolve as a by-product of genomic differentiation after extended periods of geographic separation. Accordingly, the evolution of BDM incompatibilities provides an elegant solution to the production of unfit hybrids, because natural selection need not oppose any step in this process (Orr & Turelli 2001). BDM incompatibilities are expected to accumulate with the square of the number of substitutions separating two species (Orr & Turelli 2001). The emergence of reproductive isolation through BDM is thus expected to be a slow process that gets ever more efficient as time progresses (Coyne & Orr 1997; Price & Bouvier 2002). If the evolution of epistatic BDM incompatibilities were commonplace, this ‘snowball effect’ of accelerating decline in reproductive compatibility should generally be visible. A recent exploratory meta-analysis by Gourbière & Mallet (2010), however, suggests that for most of the investigated taxa, the decay of reproductive compatibility is better predicted by linear or slowdown models. This finding calls the general importance of BDM compatibilities into question and much rather suggests that incompatibilities accumulate linearly without BDM effects and provides novel evidence for a role of reinforcement. Further meta-analyses of this kind are needed to better judge the relative contribution of these processes in generating reproductive isolation.

Much of the research on intrinsic postzygotic isolation and genic incompatibilities has focused on Haldane’s rule, the preferential effect of sterility or inviability on hybrids of the heterogametic sex (Haldane 1922). Following decades of relative stasis in speciation genetics, the field was remarkably reinvigorated in the mid-1980s by newfound interest in this phenomenon (Coyne 1985). Four main ideas have been suggested as general causes of Haldane’s rule: the dominance theory, the faster-male theory, the faster-X theory and meiotic drive (Coyne & Orr 2004), with the two former recognized as main factors in causing Haldane’s rule. New genomic data have underscored the importance of sex chromosomes in speciation (Mank et al. 2007; Presgraves 2008; Ellegren 2008a). While Haldane’s rule is commonly viewed as important in the initial stages of speciation (Kulathinal & Singh 2008), it may, however, be argued that sex chromosomes are comparatively more important in later stages of speciation, completing the process following initial differentiation (Qvarnström & Bailey 2008).

There has also been considerable debate regarding how many and what type of genes are important in causing reproductive isolation. While the population genetic approach of the Modern Synthesis held that adaptation and population differentiation was the cumulative effect of numerous genes, each with small effect (Fisher 1930), recent research on intrinsic postzygotic isolation has focused on the effect of a small number of genes each with large effect (Orr 2001). This view is taken to the extreme in research on bona fide speciation genes, of which only a handful of examples are known (Wittbrodt et al. 1989; Ting et al. 1998; Barbash et al. 2000; Presgraves et al. 2003; Ortiz-Barrientos & Noor 2005; Brideau et al. 2006; Mihiola et al. 2009; Phadnis & Orr 2009; Tang & Presgraves 2009). Only recently, the first case was documented in which both genes of a pair of epistatically interacting loci causing hybrid incompatibility (Brideau et al. 2006; Presgraves 2007). While further research is needed to resolve the function of factors causing intrinsic postzygotic isolation, exciting new evidence points to a role both for epigenetic interactions and genetic conflict (Orr et al. 2007; Mihiola et al. 2009; Phadnis & Orr 2009; Presgraves 2010).

A somewhat special case of gene–gene interactions is given by mitonuclear interactions and may deserve some extra attention. Until lately, the effect of endosymbionts on speciation has received comparatively little attention and has largely focused on incompatibilities caused by cytoplasmically inherited parasites like Wolbachia (Bordenstein et al. 2001). While mitochondrial DNA was long regarded as a neutral marker invaluable for tracing evolutionary history (Avise 2000), accumulating evidence questions the assumption of neutrality, with implications for evolutionary biology including speciation (Meiklejohn et al. 2007; Dowling et al. 2008). Specifically, as mitochondrial function is closely tied to energy production through oxidative phosphorylation and organismal fitness (Rand et al. 2004), maladaptive combinations of mitochondrial and nuclear genes in hybrids may act to reduce gene flow and drive population differentiation (Dowling et al. 2008). For example, hybrid breakdown owing to mitonuclear incompatibilities (leading to reduced energy production) has been observed in population crosses of marine copepods (Ellison & Burton 2008) and Nasonia parasitoid wasps (Ellison et al. 2008). Interestingly, large-scale analysis of the Nasonia nuclear genome implies strong effects of natural selection on nuclear genes of relevance for mitochondrial function, in line with strong selection for mitonuclear coadaptation (TNGWG 2010). Incompatibilities between nuclear and mitochondrial genes have also been shown to cause hybrid sterility in yeast (Lee et al. 2008). Clearly, the role of mitochondrial DNA in speciation deserves further attention (Levin 2003; Gershoni et al. 2009) and may explain intriguing patterns such as asymmetric introgression between incipient species (Turelli & Moyle 2007).

The nature of hybrid sterility and infertility has not only been elucidated by studies concerned with intrinsic postzygotic isolation. Investigations of postmating prezygotic isolation in externally fertilized organisms such as sea urchins and mussels have importantly contributed (Palumbi 2009). A plethora of approaches over the last two decades have revealed that proteins on gamete surfaces (e.g. bindin and lysin) are of major importance in reducing gene flow by disrupting fertilization (Lee et al. 1995; Metz & Palumbi 1996; McCartney & Lessios 2004). Egg and sperm proteins seem to have engaged in an arms race driven by sexual conflict, which may accelerate the formation of reproductive isolation (Gavrilets 2000; Swanson & Vacquier 2002). Considered in isolation this would make for a
simple story. A combination of ecological, genetic and physiological approaches, however, suggests that speciation in these systems include manifold processes that may be simultaneously necessary for the emergence of discrete clusters. These include sexual conflict, sperm competition, cryptic female choice, frequency-dependent selection and reinforcement, as well as ecological aspects such as individual density distributions (Birkhead & Pizzari 2002; Swanson & Vacquier 2002).

Both the examples of intrinsic postzygotic isolation and postmating prezygotic isolation in external fertilizers highlight a role for genetic interactions in the build-up of reproductive isolation (referring in the former case to epistatic interactions within one hybrid individual, in the latter between proteins relevant for communication between gametes). They also strengthen the idea that only few genes of major importance may suffice in driving two populations apart (Orr 2001).

3. EXTENDING THE FRAMEWORK OF SPECIATION GENETICS

The considerable amount of data that has been collected on the basis of genetically encoded hybrid incompatibilities over the last decades has biased our view of speciation towards the genetics of postzygotic isolation between rather divergent lineages. Conceptually, however, the genetics of speciation has a much broader definition, with speciation genes being functional genomic elements that convey some degree of ecological, sexual, pre- or postmating, pre- or postzygotic isolation. Furthermore, as different genes will act during different stages of the speciation process, speciation research should encompass nascent species as well as species that have accomplished a certain degree of reproductive isolation (Via 2009). For example, while understanding Haldane’s rule and the action of BDM incompatibilities is undoubtedly highly relevant, it remains unclear to what extent observed incompatibilities have contributed to the initial branching of lineages or if they merely reflect a subsequent accumulation of incompatibility factors completing the speciation process. Research into the genetics of intrinsic postzygotic isolation therefore represents a retrospective look at the speciation process (Via 2009), and an exclusive focus on hybrid sterility and inviability will impede a deeper understanding of the molecular basis of all aspects and stages of the speciation process. Phylogenetic approaches including taxa which already have diverged significantly can give important insight into the tempo and mode of speciation (Price 2010). However, a strict retrospective inference precludes a role of ecology a priori and can thereby only speculate about the conditions under which the speciation process was initiated. The evidence that prezygotic isolation seems to evolve faster than postzygotic isolation (Coyne & Orr 1997) and that postzygotic isolation can be achieved much more readily if driven by extrinsic factors (Schluter & Conte 2009) suggests that other approaches to understanding the speciation process are needed. Fortunately, current speciation research is also embracing approaches that focus on the causes of initial divergence in populations that are only partly isolated (Via 2009). From such work on evolutionary young lineages that are at incipient states of divergence, it has increasingly been recognized that barriers to gene flow can evolve as a result of ecologically based divergent or disruptive selection. This perception is bolstered both from theoretical work evaluating the role of natural selection in a growing number of empirical systems (Gavrilets 2004; Dieckmann et al. 2004a; Gavrilets & Losos 2009; Barton 2010). Clearly, a central limitation of this forward-looking approach is that one cannot foresee whether the speciation process will be driven to completion. Still, we predict that an approach combining both ecology and genetics in young systems will be fruitful, particularly so if the recently emerging genomic tools are applied to the well-established ecological model systems with long study histories (Kruuk & Hill 2008). Merging both worlds will eventually paint a broader picture of the relevant mechanisms involved in speciation. In the following sections, we highlight some areas where recent progress has been made towards understanding the role of ecology and selection in speciation.

(a) Ecological speciation in a tube

One way to address the importance of ecologically imposed divergent selection in speciation is given by experimental evolution studies on micro-organisms. This approach dates back to one of Darwin’s contemporaries, William Dallinger, but had not gathered weight until the early 1990s. A number of in vitro experiments have demonstrated that fitness trade-offs between heterogeneous environments are easy to achieve and can be stably maintained (Rainey & Travisano 1998; Buckling et al. 2009). The genes involved in adaptations can potentially be mapped and in recent years it has become possible to monitor the evolution of whole viral and prokaryotic genomes and this will likely be increasingly feasible in eukaryotes (Bomblies & Weigel 2010). From these studies it emerges that ecological adaptations often seem to entail lowered hybrid fitness between divergent lineages by negative epistatic interactions sensu BDM incompatibilities (Dettman et al. 2007; Duffy et al. 2007; Barrick et al. 2009). This establishes the link between ecological adaptation and postzygotic isolation and suggests that BDM incompatibilities can arise as incidental by-products of positive natural selection (Dettman et al. 2007; Bomblies & Weigel 2010). The observation that most of the hybrid incompatibility genes identified so far show signatures of adaptive evolution further supports this idea (Orr et al. 2007; but see Presgraves 2010).

(b) Ecology and the concept of adaptive speciation

Darwin was a clear proponent of the idea that environmental differences can generate divergent selection pressures that eventually drive two populations apart. He foreshadowed the idea that the splitting process
itself may be adaptive and not only a by-product of geographical isolation (Darwin 1859). However, in contrast to the experimental evolution studies described above (which are essentially allopatric in their setup), natural populations might be connected by gene flow for some period of time and experience gradually changing environments. The only realistic scenario for the splitting process itself to be adaptive then occurs by frequency-dependent intraspecific interactions that can result in disruptive selection (Dieckmann & Doebeli 1999). The most radical representation of such ecologically mediated speciation is that of ‘adaptive speciation’, which refers to speciation processes in which the ‘splitting is an adaptive response to disruptive selection caused by frequency-dependent biological interactions’ (Dieckmann et al. 2004b). While this concept can, under special circumstances, also work in allopatry it is essentially related to speciation under conditions of gene flow (‘divergence-with-gene-flow’).

While much of the literature on ‘ecological speciation’ revolves around similar ideas as in ‘adaptive speciation’, ecological speciation is broader in its definition (Rundle & Nosil 2005) and encompasses all instances whereby reproductive isolation can evolve as a by-product of adaptation to different environments (Schluter 2001, 2009; Rundle & Nosil 2005). Nonetheless, the two concepts have several features in common that make them explicitly different from neutral speciation models in allopatry caused by the random accumulation of negative epistatic mutations. A central theme of both is the importance of natural selection acting on a set of few key traits associated with resource use, mate choice or, in plants, pollination. Consequently, it is predicted that in the early phase of the divergence process, taxa are reproducibly isolated only at a small number of locally confined areas in the genome (‘genomic islands of speciation’), while remaining indistinguishable throughout the parts of the genome that are unaffected by selection (Turner et al. 2005; Harr 2006). During the course of genomic differentiation, this divergence mix slowly attains a higher degree of phylogenetic concordance through independent responses to genetic drift and selection within the new species (Nosil et al. 2009). Ecologically motivated speciation scenarios are thus genic in their view and put a premium on the early stages of speciation, where speciation boundaries are still porous and branching patterns are established by a few, but crucial changes.

(c) Establishing empirical evidence for divergence-with-gene-flow
Quantification of divergent or disruptive selection in the wild is not trivial. Therefore, indirect means are usually sought to evaluate the role of adaptation in speciation. For example, evidence for prezygotic isolation between subpopulations from different environments is indicative of adaptive differentiation. Examples where habitat-based prezygotic isolation has been documented include cichlid fish (Kocher 2004; Barluenga et al. 2006; Elmet et al. 2010), Galapagos finches (Grant & Grant 2008), guppies (Reznick et al. 2008), pea aphids (Hawthorne & Via 2001), butterflies (Jiggins et al. 2001), monkeyflowers (Bradshaw & Schemske 2003) and other flowering plant species (Lowry et al. 2008). Further evidence for ecologically driven divergence comes from studies examining the level of differentiation between subpopulations that have adapted to different environments as opposed to subpopulations that reside in similar environments (Funk et al. 2006). With an increasing availability of analytical tools (Foll & Gaggiotti 2006), we can expect that the environmental factors determining genetic structure in populations will be identified in new systems.

An important observation that relies on a model of speciation that invokes adaptation is the phenomenon of parallel evolution, where similar ecotypes have evolved repeatedly upon recurrent colonization of new habitats (Schluter & Nagel 1995). Hard evidence for parallel evolution is difficult to collect as it needs to be demonstrated that parallel divergence is the result of independent colonization events and not of subsequent gene flow between similar ecotypes (Schluter 2009). Still, there are examples from natural populations indicating that parallel speciation may occur. A prominent case is given by the divergence of threespined stickleback ecotypes (limnetic versus benthic and marine versus fresh water) (Rundle et al. 2000; McKinnon et al. 2004). Similar adaptations to limnetic and benthic niches are observed in lake whitefish (Rogers & Bernatchez 2007; Whiteley et al. 2008; Bernatchez et al. 2010) and arctic char (Skuåsen et al. 1996; Ott & Smith 1998). Other cases where parallel speciation seems to be in progress include adaptations to environments with and without predators in mosquitofish (Langerhans et al. 2007), adaptations to areas of different wave exposure in snails (Sadedin et al. 2009; Johannesson et al. 2010; Butlin in press), and adaptations to different host plants in walking sticks (Nosil et al. 2008) and pea aphids (Peccoud et al. 2009).

(d) Adaptations from novel mutations or standing genetic variation?
Common to the examples listed above is that parallel divergence occurred within thousands rather than millions of years. Given the speed of the process, it is most probable that genetic variants underlying a given adaptive phenotype are independently recruited from standing genetic variation, since adaptation restricted to selection of novel mutations would be far slower (Barrett & Schluter 2008). There are indeed some well-established cases of adaptive divergence occurring as a result of recent selection for particular alleles recruited from ancestral polymorphism (Schluter & Conte 2009). For example, phylogenetic studies in sticklebacks have shown that stream living, less-armoured ecotypes have arisen independently several times at different locations, but that the alleles contributing to body armouring (Ectodysplasin, Eda) were already present in their common ancestor (Colosimo et al. 2005). In the comparable case of lake whitefish (Rogers & Bernatchez 2007; Whiteley et al. 2008; Jeukens et al. 2009;
Bernatchez et al. (2010), it is likely that standing genetic variation has contributed since the ecotypes evolved repeatedly, in parallel, in very recent times. A fascinating case has been reported for the apple maggot, *Rhagoletis pomonella*. In this species, a particular eco-type started using domestic apple instead of hawthorn and in less than two centuries this change in behaviour has caused almost complete reproductive isolation between these ecotypes (Feder et al. 2003a, b). While many of the examples of parallel speciation might be indicative of selection on standing genetic variation, caution needs to be taken. In sticklebacks, the *Pituitary homeobox transcription factor 1* (*Pitx1*) locus has been shown to be involved in the development of the pelvic apparatus. Comparable to the situation of armour plating described above, most sticklebacks develop a normal pelvic apparatus. In over a dozen widely distributed populations, however, phenotypes with reduced spines seem to have evolved in parallel and seem to be under selection as a response to habitat-related factors such as predator pressure or calcium availability (see Shapiro et al. 2006 and references therein). Chan et al. (2010) have recently demonstrated that recurrent deletions in a highly mutable enhancer region of the *Pitx1* gene, rather than recruitment from ancestral variation, are responsible for a phenotype with reduced pelvic spines.

(e) The problem of recombination

In the early stages of adaptive divergence, reproductive isolation is expected to be concentrated around a small number of locally adapted genes. It remains an important challenge to understand how reproductive isolation progresses from a genetic mosaic pattern to genome-wide divergence. Particularly, under conditions with homogenizing gene flow, the association between genes involved in local adaptation and those influencing premating isolation are generally considered to be vital (Rundle & Nosil 2005; Bolnick & Fitzpatrick 2007). A key problem in speciation-with-gene-flow models is therefore to understand how natural selection can maintain adaptive gene combinations when faced with the deteriorating force of recombination. Several possibilities to overcome the disruptive influence of recombination have been suggested, including close physical linkage (Butlin 2005), reduction in effective recombination rate in the regions under divergently selection (Via & West 2008) and pleiotropy (Kirkpatrick & Barton 1997; Kirkpatrick & Ragné 2002). These scenarios are hard to disentangle as long as the causative variants for both isolation and adaptation have not been elucidated (Rundle & Nosil 2005). Still, to understand the strength of selection needed to result in the build-up of reproductive isolation, an effort to discriminate them is essential. At present, the genetic elements governing the traits involved in ecological speciation remain largely unknown. Chromosomal speciation models postulate an important role for rearrangements in the build-up of reproductive isolation among incipient species (Rieseberg 2001; Hoffmann & Rieseberg 2008). Indeed, rearrangements have been found to be associated with hybrid inviability for example in *Helianthus* sunflowers (Rieseberg et al. 1999) and *Drosophila* (Noor et al. 2001; Brown et al. 2004). Chromosome rearrangements definitely provide a means to resolve the problem of linkage, but of course require that major causative variants for both isolation and adaptation are located within the inverted region.

Pleiotropic genes are particularly attractive candidates, as they entirely bypass the problems of linkage and recombination. By definition, pleiotropy occurs when a single gene influences several phenotypic traits. Translated into the speciation context, they convey habitat-specific selective advantage and at the same time ensure assortative mating with reference to the trait under selection. Examples from several studies point towards their existence. Beak size in Darwin’s finches is both relevant to ecologically mediated fitness and species recognition (Grant & Grant 2008); the genetic background to beak shape seems to be largely confined to one locus involved in the calmodulin pathway (Abzhanov et al. 2006). Empirical data also indicate that pleiotropy might govern copper tolerance and pollinator shifts in monkeyflowers (Macnair & Christie 1983; Bradshaw & Schemske 2003), and the coupling of reproductive isolation and host switch in pea aphids (Hawthorne & Via 2001). In addition, wing colour and mate preference actually map to the same gene (*wingless*) in butterflies (Kronforst et al. 2006) and many floral traits that affect pollinator shift in columbines seem to be restricted to a small genomic region (Hodges et al. 2002). However, causative variants are not fine-mapped and verified and the results could also be explained if there is tight linkage between the locus governing local adaptation and the locus governing reproductive isolation (Rundle & Nosil 2005). Another way of establishing evidence for pleiotropic genes could potentially come from candidate gene approaches. While rapidly developing genomic tools will soon allow examining classes of candidate genes (e.g. pigmentation genes, early development genes) or entire gene families (Mamanova et al. 2010), candidate gene approaches are at present still limited to a handful of genes. For example, the *major histocompatibility* locus (*MHC*) has been extensively studied in a behavioural ecological framework. Pleiotropy with regard to parasite resistance and signal for mate choice make it a good candidate for studies on ecological speciation under conditions of gene flow (Eizaguirre et al. 2009). Another extensively studied class of genes that may be relevant to speciation are pigmentation genes like the melanocortin-1-receptor (*MC1R*) or the *Agouti signalling protein* (*ASIP*) that have been shown to influence coat and plumage colour in several organisms (Mundy et al. 2004; Hoekstra et al. 2006; Limen et al. 2009). It is intuitively clear that the match between body coloration and substrate is relevant to predator-mediated selection, as has been recently shown for *Peromyscus* mice (Mullen et al. 2009). Another convincing non-genic mechanism that generates an immediate association between ecological adaptation and mate choice is habitat learning (Beltman & Haccou 2005), which has been suggested to be relevant in several vertebrate systems.
(Musiani et al. 2007; Wolf et al. 2008). One has to be careful, however, to a priori attribute the link between ecological adaptation and assortative mating to learning, as simple non-genic mechanism of assortative mating, in which the mating trait arises as a pleiotropic effect of genes responsible for ecological adaptation, is also credible in viral evolution (Duffy et al. 2007).

Pleiotropy certainly is an appealing idea, but to date it remains unclear to what extent the different mechanisms of coupling the trait under selection and assortative mating are involved. Regardless, the studies mentioned above show that the coupling seems to be possible and rather widespread.

(f) A role for sexual selection
While sexual selection in itself need not be linked to ecological speciation (Schluter 2001, 2009), it has been discussed as a means to enhance divergence in an ecological context (Grant & Grant 1997; Edwards et al. 2005; Ritchie 2007; van Doorn et al. 2009). Under certain circumstances, the effect of a sexually selected trait depends on the environment in which it is displayed, so that the divergence in mating traits will eventually be governed by adaptation to the environment (‘sensory drive’; Boughman 2002). In a recent study, Seehausen et al. (2008) established the link between colour variants in cichlid fishes and water turbidity and provided compelling evidence for speciation through sensory drive in sympathy. Although sensory drive may promote speciation in some systems, it is conceivable that other modes of sexual selection are the driving forces of speciation (e.g. good genes (Andersson 1994) or Fisherian runaway selection (Fisher 1930; Kirkpatrick & Hall 2004)). Specific examples where sexual selection has been argued to promote divergence include the rapid diversification of cichlid fish (Seehausen et al. 1999; Kocher 2004; Elmer et al. 2010) and cricket species (Shaw & Parsons 2002; Mendelson & Shaw 2005).

Theory predicts that sexual selection is expected to be more powerful in organisms with female heterogamety (Reeve & Pfennig 2003), such as birds and lepidopterans. There is indeed some empirical indication that traits of importance for species recognition are sexually selected in Heliconius butterflies (suggesting a role for the wingless locus; Kronforst et al. 2006) and Ficedula flycatchers (Seetzer et al. 2007; Qvarnström et al. 2010). Interestingly, in the case of flycatchers there is character displacement, presumably driven by reinforcement (Sætre et al. 1997), and in crickets, butterflies and flycatchers there is also evidence for physical linkage of trait and preference loci, which could partly resolve the problem of recombination. One important point is that in most cases where sexual selection has been the suggested force of speciation, it is still to be resolved if the differentiation has evolved as a by-product of diversifying selection driven by environmental factors (ecological speciation) or through fixation by sexual selection of different mutations in populations with similar selection regimes (Ritchie 2007).

A fresh perspective on how sexual selection could facilitate speciation under sympatric conditions has been put forth by van Doorn et al. (2009). In a simple model, the authors explore how sexual selection and disruptive ecological selection can join forces to curtail gene flow, promote local adaptation and eventually lead to speciation. Key to their model is the incorporation of condition-dependent mate choice, which only involves a pre-existing mate choice machinery instead of having to rely on concomitant divergence of ecologically and arbitrarily sexually selected traits. By the introduction of this genotype-by-environment interaction they entirely circumvent the problem of earlier models to link ecological performance and assortative mating without having to invoke the presence of fortuitous pleiotropy between ecological and mating traits.

4. GETTING TO THE GENES UNDER SELECTION
Much progress has been made over the last years in identifying the genes responsible for BDM incompatibilities in model organisms (see above), but the quest for genes underlying adaptive divergence in organisms of ecological interest where few genetic resources are available has only begun. Bringing wild strains into the laboratory will not yield the same clear-cut insights as in the study of postzygotic hybrid breakdown even if the same genetic tools existed as for model species like Drosophila. The fitness effect of a given trait (and its underlying genetic basis) should preferably be investigated under the full set of environmental conditions in the wild and will be difficult to study under laboratory conditions (Calisi & Bentley 2009). The quest is further exacerbated by the fact that traits important for local adaptation are likely to be quantitative and are hence thought to have a complex genetic background (Weedon & Frayling 2008; Hendry 2009). In the years to come, many efforts will nonetheless be devoted to deciphering the genetic basis of speciation driven by adaptive divergence. Although getting to specific genes will understandably be difficult in most cases, much progress is expected in finding candidate regions of interest and in answering more general questions about the underlying genetic mechanisms. Is extensive adaptive divergence based on few loci with major effect (Gavrilets & Losos 2009) or by many loci of small effect (Fisher 1930)? Do gene interactions between traits undergoing adaptive divergence also lead to intrinsic postzygotic isolation? Are candidate genes special with regard to the level of pleiotropy or their position in protein networks? How is the homogenizing effect of gene flow and recombination overcome?

In the following, we will shortly mention some promising avenues for addressing these questions.

(a) Phenotype–genotype association
Most investigations conducted in natural populations so far have used a limited number of genetic markers, typically tens to thousands of microsatellites or amplified fragment-length polymorphisms (AFLPs), which are suboptimal with regard to large-scale genome
scans (Butlin in press). The advent of massively parallel sequencing technologies holds much promise for speeding up the progress in understanding the genetic basis of (ecological) speciation (Noor & Feder 2006; Ellegren 2008b; Vera et al. 2008; Gilad et al. 2009). The scope of these recent fast developing methods is nicely illustrated by a collation of 21 articles on next generation molecular ecology (Tautz et al. 2010). Once statistical challenges of how to appropriately deal with short-read shotgun sequences in a population genetic context are overcome, population genomic analysis will be directly based on the sequencing data itself and replace the marker-based approaches. At present still, the most straightforward application consists in scanning large proportions of the genome for polymorphisms that may be used as genetic markers for subsequent genotyping using array-based high-throughput genotyping techniques (Svyànen 2005). The increase of genetic markers by orders of magnitudes is expected to boost genetic mapping studies that have been so far often limited by the number of available markers. The first step in establishing the link between phenotype and genotype usually involves obtaining a detailed linkage map. In a few natural populations, this has already been achieved (e.g. Wang & Porter 2004; Stemshorn et al. 2005; Gharbi et al. 2006; Àkesson et al. 2007; Rogers et al. 2007; Backström et al. 2008). Nonetheless, such pedigree-based approaches require access to multigeneration samples of related individuals; that can be extremely challenging in natural populations, and again stresses the importance of long-term ecological model systems for the study of speciation. With the availability of larger marker sets one could anticipate that mapping efforts will be focused on other methods, such as association scans (linkage disequilibrium mapping) using population samples, an approach most well developed in model species with the available genome sequences (Nordborg & Weigel 2008; Goddard & Hayes 2009; Bombliès & Weigel 2010). In divergent natural populations or hybrid zones, it may be of particular interest to make use of the extended linkage disequilibrium resulting from the admixture of differentiated populations (Rieseberg & Buerkle 2002; Smith & O’Brien 2005). Recent progress in analytical approaches (Gompert & Buerkle 2009) further increases the applicability of this method making a strict geographical sampling regime dispensable. A striking example demonstrating the potential of this approach comes from a study on the genetics of introgression across Cottus hybrid zones, which basically suggests that different forms of selection affect much of the genome and provides numerous candidate regions for future studies (Nolte et al. 2009). While extended linkage after admixture is useful for identifying the genomic regions of interest, short-range linkage disequilibrium is needed to be able to resolve selection at the level of the gene. However, adaptive divergence may put a lower boundary on the resolution with which genes can be mapped, as selection can reduce the effective (interspecific/interpopulation) recombination rates in regions harbouring genetic determinants of local adaptation (‘divergence hitchhiking’; Via & West 2008). This results in large regions spanning significant portions of the genome increasing the degree of differentiation at marker loci located far from the target of selection (but see Yatabe et al. 2007; Wood et al. 2008). Nevertheless, as exemplified in disease-mapping studies in dog, there may still be potential for designing mapping studies so that within-breed long-range linkage disequilibrium (admixture or divergence hitchhiking in the case of incipient species) is used to find candidate regions and then between-breed short-range linkage disequilibrium (within species) allows for more detailed searches (Sutter et al. 2004; Lindblad-Toh et al. 2005).

Additional possibilities to characterize the genes involved in early speciation spring from the ever increasing characterization of gene function in model species from where it will be possible to extract candidate loci for investigation in the focal species (Hoekstra et al. 2004; Mundy 2005). This approach certainly takes the risk that there might be different genetic backgrounds to similar phenotypes also between closely related species or between populations within species (Hoekstra & Nachman 2003).

(b) Phenotype uninformed methods: evolutionary genomics
All the above-mentioned approaches require some previous knowledge about the phenotype involved in the adaptive process (top down; mapping). An alternative approach lies in the application of population genetic approaches to detect selection directly from DNA sequence data without a priori knowledge of the phenotypic effect (bottom-up; evolutionary genomics). With an ever increasing availability of genome-wide polymorphism and divergence data, it will be possible to scan genomes of diverging populations for regions with higher than expected differentiation indicative of ongoing or recent diversifying selection (Akkey et al. 2004; Beaumont 2005; Excoffier et al. 2009). Additionally, high-density marker data can be used to trace regions indicative of recent directional selection (selective sweeps) within populations (Nielsen 2005; Nachman 2006). An example of where this approach has been successfully applied comes from wild mice populations (Harr 2006; Teschke et al. 2008). However, it is well recognized that population structure and other demographic scenarios can severely affect the expected distribution of parameter values (Thornton & Andolfatto 2006; Pool & Nielsen 2007; Excoffier et al. 2009; Hermisson 2009). Encouragingly, numerous methods have been developed to estimate the demographic scenario under which to search for the footprints of selection (Hey & Nielsen 2004; Hey 2006; Becquet & Przeworski 2007). Clearly, studies on rich datasets that infer selection under detailed demographic scenarios will set the future standards (Nielsen et al. 2009).

Valuable insights into the genetic basis of adaptation may further come from comparative genomic studies that set out to find signatures of selection by comparing sequences of orthologous genes from two or more organisms (Ellegren 2008a). Genes affected by diversifying selection on protein structure are

Phil. Trans. R. Soc. B (2010)
expected to have a higher ratio of non-synonymous to synonymous differences among taxa than genes that evolve under purifying selection. For this approach to be meaningful, however, the number of fixed mutations must clearly exceed the number of polymorphisms, which restricts it to comparisons of lineages that have speciated millions of years ago. Still, it may be informative if we consider which genes or gene ontology classes are repeatedly identified to be under positive selection, as these genes may also be involved in earlier stages of the splitting process. A complementary approach that can identify the spread of beneficial mutations in single lineages consists of contrasting polymorphism to divergence data between species with McDonald–Kreitman (McDonald & Kreitman 1991) and Hudson–Kreitman–Aguadetypetype (Hudson et al. 1987) approaches (Begun et al. 2007).

So far, past comparative genomic approaches have been limited to a small number of organisms where whole genome sequences have been available (Kosiol et al. 2008). Having entered the era of massively parallel sequencing, this will rapidly change and the first large-scale examples of comparative genomic analyses on non-model organisms are being published (Künstner et al. 2010).

5. A ROLE FOR GENE EXPRESSION IN SPECIATION
(a) Structural variation versus variation in expression
Thirty-five years ago, King & Wilson (1975) expressed their amazement that homologous protein and DNA sequences appeared to be almost identical between humans and chimpanzees. This influential paper touched upon an important concept whose basic postulate is still valid. Functional polymorphism in genes relevant to evolutionary change is not restricted to coding variation, which ultimately alters amino acid composition and protein structure; it also includes regulatory variation modulating the expression of a gene. Several lines of research have made clear that changes in gene expression are indeed relevant in speciation (Tautz 2000; Wittkop et al. 2008). This applies to a broad variety of taxa and ranges from colour patches in the wings of flies (Gompel et al. 2005) to beak size in Galapagos finches (Abzhanov et al. 2006). While the evolutionary implications of structural variation have been extensively explored in an evolutionary framework both in theory and practice, scrutiny of the evolution of gene expression remains a big challenge.

Analysis of the role of gene expression in speciation faces many obstacles. Part of this relates to technological restrictions. For studying structural variation, the one-dimensional DNA sequence can nowadays be read with great ease and the amino acid composition can directly be derived from the genetic code. However, it is technically more demanding to work with RNA and quantify gene expression. In contrast to DNA sequencing, where clear quality standards have been established that enable comparative results across laboratories, quantification of transcript abundance differs between technologies such as qRT-PCR and microarray studies. Furthermore, gene expression studies often fail to reflect the (major) quantity of interest: protein abundance (Schrmpf et al. 2009). Nonetheless, the development of promising new approaches (Wang et al. 2009) will facilitate investigations of the role of gene expression in speciation. Instead, biological complexity may pose a greater challenge. Numerous complex and interacting processes like transcription, transcript stability, splicing, regulatory RNAs and translational efficiency eventually determine protein abundance in a cell and it is difficult to make allowances for all. Another complicating factor is the sometimes widely different expression profiles among tissues and developmental stages, which makes it hard to pick the right time and place to study the evolution of expression differences between lineages. Similarly, expression profiles are notoriously plastic (Cheviron et al. 2008), which limits expression studies to species that can be bred under common garden conditions with relative ease.

Despite the analytical difficulties associated with studying gene expression, a body of literature has accumulated over the last years showing that regulatory variants are a primary substrate for the evolution of species (Wray 2007). In the past, it has been customary to focus on structural sequence variation and consider each gene as a separate unit of evolution in both population genetic theory and empirical practice. However, phenotypic traits are controlled by a large number of different genes and changes in their temporal and spatial coordination have far-reaching consequences. Stern & Orgogozo (2009) posit that once we start considering the interactions of genes and transcripts, we may understand that not all genes are equal in the eyes of evolution, and that evolutionarily relevant changes may accumulate in certain hotspot genes located at specific positions in regulatory networks. Results of experimental evolution studies corroborate this claim (Cooper et al. 2003). Tapping the full potential of such a perspective can shed new light on evolutionary phenomena like parallel evolution in divergent lineages that are difficult to explain otherwise. For example, does parallel speciation commonly involve parallel changes in expression patterns, following strong selection for particular beneficial alleles (Unckless & Orr 2009)? The integration of network thinking into evolutionary genetics may constitute a similar quantum leap as the integration of Mendelian genetics into Darwin’s evolutionary framework (Koonin 2009). Without doubt, the burgeoning interest in the contribution of gene expression to species divergence is part of this transition.

(b) A role for selection on gene expression?
Under the assumption that most characters are controlled by a large number of interacting genes, we can expect that the underlying genetic network may be largely resilient to slightly deleterious changes in one of its elements. Likewise, advantageous mutations may not be manifested significantly in a genetic pathway with many developmental ramifications (Wagner 2000). Many features of transcriptional networks may indeed be described by non-adaptive processes and a non-
negligible number of regulatory changes may thus be expected to evolve approximately in a neutral fashion (Lynch 2007). Khaitovich et al. (2005) have laid out the theoretical basis for a neutral evolutionary model of gene expression which predicts an approximately linear accumulation of expression differences with time as well as a correlation of expression variance within a species and expression differences between species. Several studies explicitly addressing this question are in line with a predominantly neutral scenario (Khaitovich et al. 2004; Staubach et al. 2010). If divergence in gene expression progresses simply as a function of time in a largely neutral fashion, we may expect regulatory incompatibilities to arise analogous to the BDM model for structural variation. Indeed, experimental evidence suggests that divergence in gene regulation is a major contributor to BDM incompatibilities between several species of Drosophila (Haerty & Singh 2006), which may be a more widespread phenomenon as hybrid mis-expression between taxa is not restricted to drosophilids (Cowles et al. 2002; Tirosch et al. 2009). Still, it may be premature to argue that regulatory incompatibilities generally arise analogous to genetic incompatibilities sensu the BDM model. Alternative scenarios such as compensatory changes among interacting gene products or gene products and regulatory elements need to be taken into consideration (Landry et al. 2007).

(c) Cis or trans?
The study of hybrid mis-expression has also proven to yield valuable insights on the relative role of *trans* and *cis* factors in the evolution of novel phenotypes. Changes in *cis*-acting sites occur on the regulatory sequences of the gene itself. This way, their effect is restricted to the sequences of their own DNA (or RNA) molecule. *Trans* factors (such as transcription factors) on the other hand are separate molecules (proteins or RNAs) that can influence the activity of a broad variety of targets. Studies in several model species suggest that expression divergence is predominantly owing to changes in *cis* factors (Wray 2007; Wittkopp et al. 2008). A broad study in yeast (Tirosch et al. 2009) further suggests that upstream components involved in transduction of environmental or internal signals to direct regulatory elements (sensory *trans* factors) seem to be more often involved than direct transcription and chromatin regulators (regulatory *trans* factors). Taken together, these results suggest that not all divergence processes are strictly neutral. Reviewing the empirical evidence to date, Fay & Wittkopp (2008) conclude that adaptation often occurs by changes in gene regulation and that *cis*-regulatory sequences appear to play a special role in adaptive divergence.

What is the evolutionary relevance of *cis*-regulatory changes and how can they affect the speciation process? It is often proposed that natural selection can operate more efficiently on *cis*-regulatory mutations. First, alleles in diploid organisms are largely transcribed independently suggesting that—in contrast to structural mutations that are mostly recessive—mutations in *cis*-regulatory sequences are often co-dominant and thereby directly accessible to selection (see Wray 2007 and references therein). Second, modular organization and tissue-specific expression governed by enhancer elements reduce the degree of negative pleiotropy. Wray (2007) and Prud’homme et al. (2007) provide a broad array of empirical examples that nicely illustrate the vast creative potential of *cis*-regulatory evolution. In sticklebacks, for instance, *cis*-regulatory changes in the *Ptx1* gene seem to be associated with differences in pelvic skeleton structure between marine and freshwater forms (Shapiro et al. 2004; Chan et al. 2010). The repeated evolution of pelvic reduction in freshwater populations is not limited to populations of the three-spined sticklebacks, but can also be observed in the distantly related lineage of nine-spined sticklebacks (*Pungitius pungitius*; Shapiro et al. 2006). This observation of repeated habitat-associated changes in the regulatory region of *Ptx1* is not only suggestive of a role in *cis*-elements in parallel evolution, but it also underlines the importance of studying the genetic basis of speciation in combination with ecological research. It will be of crucial importance to our understanding of regulatory divergence processes to apply genomic tools to non-model species that have been extensively studied on morphological, ecological and behavioural grounds.

(d) Expression studies in non-model organisms
A few years ago genomic studies of gene expression in non-model organisms were out of reach. Today, they certainly still constitute a technological challenge, as genomic resources are, with the exception of a few speciation models like *Anopheles* (Cassone et al. 2003), usually unavailable. Interspecific microarrays have been successfully applied (Cheviron et al. 2008; Renaut et al. 2009), but will always remain a compromise. We can expect, however, that digital measures of expression on the basis of next generation sequencing (RNAseq) will be a major breakthrough (Gilad et al. 2009; Wang et al. 2009). As sequence and expression data are simultaneously generated, this approach has the advantage that structural and expression divergence can be directly compared. It further enables a much more detailed view on expression, e.g. by considering allele-specific expression patterns (Fontanillas et al. 2010) or by characterizing splicing variants (Harr & Turner 2010), which are not tractable by interspecific microarrays. First studies exploring the potential of RNAseq in non-model organisms which rely on distant genomic resources are promising (Buggs et al. 2010; Goetz et al. 2010; Wolf et al. 2010) and document the dawning of an era where high-resolution transcript-profiling in non-model organisms will become commonplace (Gilad et al. 2009).

6. CONCLUSION
Since the conception of evolutionary biology, interest in speciation has gone through periods of intense discussion and times of relative stasis. Over the last two decades speciation research has gained enough momentum to address the genetics of the splitting process in earnest. This advancement has primarily been driven by an interest in the genetics of intrinsic postzygotic isolation with particular reference to genic

Phil. Trans. R. Soc. B (2010)
Introduction. Speciation genetics J. B. W. Wolf et al. 1727

incompatibilities. Despite considerable achievements, this field undoubtedly awaits further insights. Experimental studies will continue to contribute to the understanding of divergent selection in the accumulation of genic incompatibilities and the fast development in massive parallel sequencing technologies now makes it affordable to monitor the evolutionary process across entire genomes.

Technological and analytical developments and a newfound interest in the role of ecology in promoting differentiation have recently encouraged speciation genetic research to broaden its perspective. From being confined to the study of genic incompatibilities in hybrid crosses, speciation genetics now extends to research on the forces involved in the initial phase of the speciation process and the role played by diversifying selection. This development enhances the chance to establish the link between phenotypes and genotypes and to bring the study of speciation into the wild. Population genomic analyses and studies on the role of expression and copy number variation will soon be common practice also in non-model organisms. A burgeoning field of great potential herein is the genomics of hybrid zones where generations of hybrid and backcross individuals lay the foundation for genome scans, taking advantage of differences in within- and between-population levels of linkage disequilibrium. Similarly, long studied ecological models with well-established pedigree information and knowledge on traits involved in reproductive isolation will remain highly valuable resources. In particular, wild species that can be kept under controlled conditions bear the potential to reveal the underlying genetic causes of speciation. Finally and unquestionably, unravelling the mechanistic foundations underlying the ‘mystery of mysteries’ will ever more benefit from combining the expertise from many fields.

This review results from a symposium on speciation entitled ‘Origin of Species—150 Years Later’, held at the Sven Lovén Center for Marine Sciences in Fiskebäckskil, Sweden. We would like to thank the organizers Hans Ellegren, Staffan Ulfstrand and Michael Thorndyke along with the Center for Marine Sciences in Fiskebäckskil, Sweden. We further acknowledge post-doctoral research funding (grant: 2008-1840 to J.L.) and the Swedish Research Council FORMAS (grant: 2008-1840 to J.L.) and the Swedish Research Council (grant: 2009-693 to N.B.), respectively.

REFERENCES
Bomblies, K. & Weigel, D. 2010 Arabidopsis and relatives as models for the study of genetic and genomic


Introd. Speciation genetics  J. B. W. Wolf et al. 1729


Haldane, J. B. S. 1922 Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12, 279–281.


Kirkpatrick, M. & Barton, N. H. 1997 The strength of indir-


---

*Phil. Trans. R. Soc. B* (2010)
Introducotton. Speciation genetics J. B. W. Wolf et al. 1731


Sæther, S. A. 2010 J. B. W. Wolf


---

*Phil. Trans. R. Soc. B* (2010)
Introduction. Speciation genetics  J. B. W. Wolf et al. 1733


