INVITED REVIEWS AND META-ANALYSES

A road map for molecular ecology

ROSE L. ANDREW,1 LOUIS BERNATCHEZ,2 AURÉLIE BONIN,3 C. ALEX. BUERKLE,4 BRYAN C. CARSTENS,3 BREN C. EMERSON,6 DANY GARANT,7 TATIANA GIRAUD,8 NOLAN C. KANE,1 SEAN M. ROGERS,9 JON SLATE,10 HARRY SMITH,11 VICTORIA L. SORK,12 GRAHAM N. STONE,13 TIMOTHY H. VINES,14 LISETTE WAITS,15 ALEX WIDMER16 and LOREN H. RIESEBERG1,17

1Department of Botany, University of British Columbia, 3529-6270 University Blvd, Vancouver, BC V6T 1Z4, Canada, 2DInstitut de Biologie Intégrative et des Systèmes, Département de Biologie, 1030, Avenue de la Médecine, Université Laval, Québec, QC G1V 0A6, Canada, 3Laboratoire d’Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France, 4Department of Botany, University of Wyoming, 1000 E. University Ave., Laramie, WY 82071, USA, 5Department of Evolution, Ecology and Organismal Biology, 318 W. 12th Ave., The Ohio State University, Columbus, OH 43210, USA, 6Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), C/ Astrofísico Francisco Sánchez 3, La Laguna, Tenerife, Canary Islands 38206, Spain, 7Département de Biologie, Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada, 8Laboratoire Ecologie, Systématique et Evolution, UMR 8079 CNRS-LIPS-AgroParisTech, Bâtiment 360, Univ. Paris Sud, 91405 Orsay cedex, France, 9Department of Biological Sciences, University of Calgary, 2500 University Drive N.W., Calgary, AB T2N 1N4, Canada, 10Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK, 1179 Melton Road, Burton-on-the-Wolds, Loughborough LE12 5TQ, UK, 12Department of Ecology and Evolutionary Biology, University of California Los Angeles, 4139 Terasaki Life Sciences Building, 610 Charles E. Young Drive East, Los Angeles, CA 90095, USA, 13Institute of Evolutionary Biology, University of Edinburgh, The King’s Buildings, West Mains Road, Edinburgh EH9 3JT, UK, 14Molecular Ecology Editorial Office, 6270 University Blvd, Vancouver, BC V6T 1Z4, Canada, 15Department of Fish and Wildlife Sciences, University of Idaho, 875 Perimeter Drive MS 1136 Moscow, ID 83844, USA, 16ETH Zurich, Institute of Integrative Biology, Universitätsstrasse 16, Zurich 8092, Switzerland, 17Department of Biology, Indiana University, 1001 E. 3rd St., Bloomington, IN 47405, USA

Abstract

The discipline of molecular ecology has undergone enormous changes since the journal bearing its name was launched approximately two decades ago. The field has seen great strides in analytical methods development, made groundbreaking discoveries and experienced a revolution in genotyping technology. Here, we provide brief perspectives on the main subdisciplines of molecular ecology, describe key questions and goals, discuss common challenges, predict future research directions and suggest research priorities for the next 20 years.

Keywords: behaviour, community phylogeography, ecological genomics, hybridization, landscape genomics, microbial diversity, molecular adaptation, phylogeography, speciation, trophic ecology

Received 15 March 2013; revision received xxxx; accepted 16 March 2013

Introduction

Molecular ecology refers to a diversity of approaches that use molecular genetic techniques to address ecological questions. This nascent discipline has expanded into a field encompassing a broad range of ecological and evolutionary questions, largely shaped since 1991 by the journal Molecular Ecology. In the past decade, the field of molecular ecology has been revolutionized, and this revolution is ongoing. Almost daily, new technologies and analytical approaches open up novel ways to address classic questions and enable us to test hypotheses that were previously unanswerable. Today, a typical issue of Molecular Ecology may include papers using molecular genetic approaches to investigate the interactions among species (including difficult to culture microbes), the genetics and evolution of ecologically
important traits, the relatedness among individuals and (based on this information) their dispersal and behaviour, the movement of individuals across landscapes, the formation of new species and the consequences of hybridization between divergent lineages. Here, we examine the current state of the key subdisciplines within the field of molecular ecology and predict possible future directions and ongoing changes.

What will the field of molecular ecology look like in years to come? Much of the current revolution is enabled by new technology, with advances in genomics, satellite imagery, computer hardware and software, and microfluidics leading to major changes in how we can best investigate molecular ecological questions. Molecular ecologists must therefore be masters of integrating these new technological approaches and applying them in innovative ways to key biological questions. And this brings up what is most exciting about the current state and direction of the field—the biological questions remain at the fore.

A classic metaphor relating to many scientific endeavours is ‘the drunk under the streetlight’, who searches for his lost keys in the area illuminated by a streetlight even though he knows his keys are far away in a darkened area where he cannot see. While many scientific questions are indeed difficult to investigate due to our limited observational abilities, these new technologies are dramatically broadening our horizons. Much of early molecular ecology research was limited by the difficulties and expense in obtaining genetic information, but today we can assess thousands of loci in numerous individuals routinely and relatively cheaply (Davey & Blaxter 2010; Davey et al. 2011). We are now at the point where the whole street is illuminated, and many ‘key’ questions that we have wanted to ask all along can finally be addressed.

This ‘road map’ paper builds on a symposium about the future of molecular ecology that was held on 6 July 2012 at the First Joint Congress on Evolutionary Biology in Ottawa. The symposium showcased some of the leading research in molecular ecology, mapping out future research priorities with panel discussions. On 24 October 2012, we held an Online Forum to obtain additional feedback from the molecular ecology community about challenges and priorities. These talks and discussions are distilled below into brief perspectives on many of the key research areas in molecular ecology, including discussions of important questions, challenges and priorities.

**DNA sequence–based trophic ecology**—by Brent C. Emerson

Trophic interactions represent the primary data for investigations ranging from single species conservation through to the resolution of community food webs. However, our capacity to directly acquire data on trophic interactions is limited by our ability to record interactions between consumer and resource species. Monitoring these interactions is challenging when consumers are difficult to observe and becomes more challenging still if a consumer is a resource-use generalist, as opposed to specialist.

In response to these difficulties, efforts have been directed towards the quantification of diet by indirect means such as the analysis of faecal material (morphological remains, plant alkanes, near infrared reflectance spectroscopy), analysis of gut contents (as for faecal analysis, but also including prey-specific antibodies and protein electrophoresis), laboratory feeding experiments and stable isotope analysis of consumer tissues (reviewed in Valentini et al. 2009; Pompanon et al. 2012).

The expansion of publicly available genomic resources, technical advances in PCR amplification, the development of next-generation DNA sequencing technology and increased understanding of the degradation properties of DNA sequence have been capitalized upon to refine and improve the general approach of developing species-specific primers (Asahida et al. 1997). The DNA of resource species can potentially be sampled from the moment the organism is consumed (upper digestive tract) through to the moment it is expelled as faeces, with DNA quality expected to progressively decay with sampling time.

Recent work has shown that regions of chloroplast DNA of more than 500 bp can be amplified from DNA extracted from herbivorous beetles, facilitating the identification of resource species consumed (Jurado-Rivera et al. 2009). Similar work on rare weevil species from Mauritius reveals that the upper digestive tract frequently contains plant material from a single resource species, sufficient for co-extraction, cpDNA amplification and sequencing (J. N. J. Kitson, B. H. Warren, V. Florens, C. Baider, D. Strasberg & B. C. Emerson, unpublished data). In such cases where reasonably long DNA regions can be amplified from a single resource species, traditional DNA sequencing approaches can be employed, but degraded samples will most profitably gain from parallel sequencing technology (for reviews see Valentini et al. 2009; Pompanon et al. 2012). In both cases, the ability to select plant-specific genes (cpDNA) eliminates unwanted co-amplification of consumer DNA, which poses more of a problem for the characterization of predator–prey relationships (for a review see King et al. 2008a).

In the case of predators and prey, if there is sufficient prior genomic knowledge and phylogenetic distance between consumer and resource species, resource-
specific primers may be possible, as has been done to compare insectivory among sympatric New World primates (Pickett et al. 2012). However, as the genomic distance between consumer and resource species narrows, the challenge of selectively amplifying the resource species increases, but primer blocking does hold some promise for minimizing this problem (Vestheim & Jarman 2008; Shehzad et al. 2012).

Promises and challenges

The main goals of DNA sequence–based trophic ecology are to (i) characterize resource species utilization by a focal species and (ii) estimate the proportional representation of each resource species to the focal species’ diet. Unfortunately, there is not a one size fits all solution to these goals, because of the inherent variance and complexity of consumer resource systems (Fig. 1). Several of the main challenges and priorities for DNA sequence–based trophic ecology over the next two decades are listed below:

1 Development of methods that provide accurate and unbiased identification of all resource species. Common to all systems is the desire to amplify a DNA sequence region, or regions, across the full range of resource species, in the absence of amplification bias, with each resource species being uniquely identifiable within a fully inventoried DNA sequence reference library of resource species. While this may be considered the gold standard, it may often be difficult to achieve in full.

2 Development of methods that provide accurate and unbiased quantitative estimates of the proportional representation of resource species (see Pompanon et al. 2012 for a discussion). Current estimates rely upon a quantifiable relationship between the copy number of an amplified DNA sequence region and resource biomass; these estimates also require equal amplification efficiency of the target DNA region across different resource species.

3 Elimination of DNA amplification step (i.e. PCR). Problems arising as a result of PCR amplification bias can potentially be overcome if technology eliminates PCR altogether. This becomes theoretically possible as next-generation sequencing capacity increases and logistically possible as costs come down. If the genomes of each resource species become available as a reference tool, PCR-free sequencing of tissue sampled from the alimentary system of a consumer may permit the estimation of presence, absence and abundance of resource species DNA.

4 Connection of spatial variation in abundance of consumer and resource species with dietary patterns. Combining ecological surveying and sampling with a DNA sequence–based approach to trophic ecology could connect spatial variation in resource species utilization by consumer species to spatial variation in abundance of either group. Such data would contribute to understanding the evolution of diet within groups of evolutionary interest such as plant-feeding insects.

5 Evaluation of diet preferences of consumers. Molecular characterization of resource utilization may be a useful tool for conservation biologists, providing a means to evaluate which resource species are favoured by a consumer species, or whether rarity among consumer species is related to resource species limitations, such as specialization to rare resource species, or competitive exclusion by other consumers.

Even in the absence of quantitative estimates, DNA sequence–based characterization of trophic ecology has already taken us far beyond where we were, and we can expect more to come. The key to success, as always, will be clever questions, careful experimental design and cautious interpretation of data.

Influential passengers: microbial diversity within multicellular organisms—by Graham N. Stone

Cohabiting microorganisms (bacteria, fungi, protists) play diverse roles in the biology of multicellular hosts. Improved methods of molecular detection—and particularly high-throughput sequencing—are driving an explosion of studies detecting bacterial and fungal contributions to the genomic and transcriptomic diversity present within other organisms. It is increasingly clear...
that many aspects of organismal phenotypes reflect contributions from a diverse associated microbiome (Zilber-Rosenberg & Rosenberg 2008; Gibson & Hunter 2010).

Some of the newly revealed phenotypic diversity stems from discovering new roles for otherwise well-known symbionts. For example, Wolbachia bacteria, primarily known for their influence on the reproductive biology of their hosts (Hilgenboecker et al. 2008), are now also known to allow the larvae of leaf-mining moths to maintain ‘green islands’ of living plant food within fallen leaves (Kaiser et al. 2010), and to suppress populations of human pathogens within mosquito vectors (Pan et al. 2012). Other new discoveries are of familiar microbial groups in unfamiliar places. Examples include recently demonstrated associations between multiple lineages of nitrogen-fixing bacteria and ants, whether inhabiting the nests of leaf-cutters (Pinto-Tomás et al. 2009) or specialized organs within the bodies of honeydew feeders (van Borm et al. 2002). Growing numbers of studies, however, reveal unexpected and major roles for microbes in mediating interactions within (Sharon et al. 2010) and between other species (Cafaro et al. 2011; van der Heide et al. 2012; McFrederick et al. 2012; Oliver et al. 2012; Zhang et al. 2012). A new paradigm may be that some of the variation in most host traits can be attributed to such ‘influential passengers’ (O’Neill et al. 1997). However, the microbiomes of the vast majority of organisms remain unsurveyed.

Future promise and challenges

Three general approaches should allow rapid advances in this field in the near future:

1. First, the falling cost of high-throughput sequencing allows large-scale DNA-barcode-based surveys of host-associated microbial diversity. It is now possible to ask how microbial floras vary both across host species (Oliver et al. 2010; Anderson et al. 2012; Sullam et al. 2012) and within them (e.g. Qi et al. 2009; Blaalid et al. 2012), including our own (Yatsunenko et al. 2012).

2. Second, genome and transcriptome libraries for a given focal species inevitably contain contributions from associated microorganisms. Informatics tools used to filter out nonhost contributions during host genome/transcriptome assembly can also be used to focus on host–symbiont associations (Kumar & Blaxter 2011), and differences in base composition and coverage between host and bacterial genomes allow contributions from these sources to be visualized (Fig. 2).

3. Third, sequencing experiments can be designed to distinguish environmental contaminants from microorganisms that are genuinely facultative or obligate symbionts. Given what is known of the genomic relationships between hosts and their symbionts (Dunning-Hotopp et al. 2007; The International Aphid Genomics Consortium 2010; Suen et al. 2011), it seems

![Fig. 2 Visualization of host and associated microbial genomes in two systems. (a) the host nematode Dirofilaria immitis and its known Wolbachia symbiont and (b) the host nematode Caenorhabditis sp. 5 and its newly revealed microbiome. On each of the two plots, points map individual contigs reconstructed from whole-organism genomic libraries based on variation in GC content and sequence coverage (for full details, see Kumar & Blaxter 2011). In each figure, 10 000 randomly selected contigs were annotated by BLAST comparison with the NCBI nt database and coloured by taxon for the best match identified (non-BLASTed contigs are shown in grey in both figures). In (a), we see clear separation between contigs attributable to the host worm and its symbiont. In (b), the host nematode is clearly separated from microbial contigs attributed to ten distinct clusters from seven taxonomic orders. While some bacterial sequences were expected in this library because the worms had been fed Escherichia coli, the roles and origins of the other bacteria are unknown. Modified with permission from Kumar & Blaxter (2011).](2608.png)
likely that some aspects of gene diversity and expression in any host will only make sense when considered at the scale of the combined 'holobiont' (Zilber-Rosenberg & Rosenberg 2008).

Phylogeography—by Bryan C. Carstens

Phylogeographic investigations seek to identify the forces that influence the geographical distribution of genetic variation. As befits a discipline that was developed to conduct phylogenetic investigations within species, studies are commonly concerned with identifying cryptic diversity (e.g. King et al. 2008b), estimating lineage divergence (e.g. Evans et al. 2011) or understanding species boundaries (e.g. Barrett & Freudenstein 2011; Niemiller et al. 2011). However, phylogeography in 2012 bears little resemblance to studies published 20 or even 10 years ago. Changes in the types of data collected by phylogeographers and the analytical methods employed have accelerated in recent years, and the net effect of these changes has been to dramatically improve the quality of phylogeographic inference.

Originally conceived as a mitochondrial bridge between phylogenetics and population genetics (Avise et al. 1987), phylogeographic investigations nowadays commonly include multiple nuclear loci (e.g. Amaral et al. 2012) or microsatellites (e.g. Zakharov & Hellman 2012). In addition, next-generation sequencing (NGS) has already produced a number of compelling phylogeographic data sets (e.g. Emerson et al. 2010; Gompert et al. 2010; Zellmer et al. 2012). The acquisition of these data is motivated by concerns about selection on the mitochondrial genome (e.g. Kivisild et al. 2006), the realization that stochastic forces can lead to incongruence between the history of a population and the history of any single locus (e.g. Maddison 1997), as well as the desire for improved parameter estimates (e.g. Beerli 2006).

The methods used for data analysis have also changed, becoming more reliant on models (e.g. Knowles 2009; Beaumont et al. 2010). One exciting development is model comparison, where the probability of multiple models is calculated given the data and subsequently ranked by calculating the relative posterior probability of models (e.g. Fagundes et al. 2007; Peter et al. 2010) or by using an information theory approach (e.g. Carstens et al. 2009; Provan & Maggs 2012). Rather than making qualitative inferences derived from patterns in the data, phylogeographers can now model specific evolutionary scenarios and evaluate the probability of each given the data. This approach is less prone to over-interpretation (Knowles & Maddison 2002) and confirmation bias (Nickerson 1998) and also less likely to be misled by inaccurate parameter estimates.

Future promise and challenges

1 One persistent challenge facing phylogeography is the discovery of cryptic species-level variation. While methods for species delimitation are improving, we lack methods that can accurately discover cryptic diversity from genetic data across a broad range of relevant parameter space. Clustering approaches such as Structure are broadly applicable (e.g. Rittmeyer & Austin 2012), but methods that utilize species trees are mostly limited to validating proposed species boundaries rather than jointly estimating these boundaries and the species phylogeny (but see O’Meara 2010).

2 A second challenge is identification of optimal sampling design, where the axes of variation are the number of samples and the number of loci. Historically, the discipline has relied largely on sampling that maximized the former because phylogeographic breaks and cryptic diversity are difficult to discover across an undersampled landscape. However, this question needs to be re-examined in the light of our expanded capacity to collect data along the second axis, in order to optimize the sampling design for particular questions.

3 A third challenge is developing methods for comparative phylogeography. To date, most comparative work has proceeded by inferring phylogeographic history on a species-by-species basis and secondarily comparing these results. While integrative community-level approaches to data analysis are generally lacking (but see Hickerson & Meyer 2008), such an approach could dramatically improve our ability to estimate the evolutionary history of ecological communities, particularly those that are coevolved (e.g. Smith et al. 2011; and see following section).

Community phylogeography—by Graham N. Stone

As described above, a range of new phylogeographic techniques now allow formal comparison of the support in observed data for alternative scenarios of population history (Bertorelle et al. 2010; Hickerson et al. 2010; Huang et al. 2011). When applied to sets of species, these approaches allow identification of shared routes of range expansion or barriers to gene flow (e.g. Hickerson & Meyer 2008). Data sets for species in the same guild (Bell et al. 2011; Dolman & Joseph 2012) or in interacting trophic levels (Smith et al. 2011; Stone et al. 2012) allow testing of alternative models of community assembly, and hence link population ecology with macroecology (Byrne et al. 2011; Ricklefs & Jenkins 2011). It is possible to ask, for example, whether species that
now form important components of interaction networks (such as food webs or pollination webs) have a long history of co-occurrence, tracking each other through space and time from a shared origin, or instead represent recent associations with discordant phylogeographic histories (Fig. 3; Stone et al. 2012).

While discrimination between complex scenarios can be challenging because of the number of parameters that must be estimated, comparative phylogeography research programmes are becoming increasingly accessible. Generation of the data sets required to infer population history in a coalescent framework is becoming increasingly affordable with high-throughput sequencing, making it ever easier to select target species on the basis of their biological interest rather than their tractability with existing markers (Emerson et al. 2010). Further, the changing emphasis from a few loci in many individuals to many loci in a few individuals (e.g. Lohse et al. 2012) makes it much easier to incorporate species that are either rare, difficult to sample or exist only in museum collections. As an example, Lohse et al. (2012) used multilocus data in only a single haploid male individual from each of three populations to infer contrasting times of range expansion across the Western Palearctic in a guild of parasitoid wasp species.

Future promise and challenges

1 The major challenge in community phylogeography is accurate estimation of the topology of population relationships, and the timing of population splits and dispersal events, for multiple species (such as guilds or trophic levels) (Fig. 3). Even for simple models, this is a very data-hungry problem.

2 Next-generation and third-generation sequencing technologies offer enormous promise for this field, providing increasing power to estimate parameters from population genomic data at low cost. The challenge has been to develop analytical approaches that make best use of many hundreds or thousands of sequences. A major attraction of such population genomic data is that support for alternative simple population models can be estimated in a likelihood framework using very small numbers of individuals per population. Although simple, these models can reasonably be applied to real-world scenarios (Lohse et al. 2012; Smith et al. 2012).

3 Additional promise is provided by ABC approaches that incorporate data for many loci in many species (Huang et al. 2011). Because this approach is simulation based, an ongoing aim for this field is to overcome the computational challenge of extending multispecies models to population genomic-scale data sets.

4 Once phylogeographic relationships for sets of species are determined, it will become possible to use these to control for statistical nonindependence in population-level analyses of interactions between species (Stone et al. 2011). This importance of this issue
is increasingly recognized, but remains challenging, in analyses of local adaptation and community genetics (Marko 2005; Evans et al. 2008).

**Landscape genomics—by Rose L. Andrew and Victoria L. Sork**

Landscape shape gene flow by influencing the ease with which organisms, propagules or gametes move, the production of propagules and the density of receiving populations. The subject of many review articles (e.g. Manel et al. 2003; Storfer et al. 2007, 2010; Holderegger & Wagner 2008; Sork & Waits 2010), methods papers (Balkenhol et al. 2009; Cushman & Landguth 2010a,b; Spear et al. 2010), a special edition of *Molecular Ecology* (Waits & Sork 2010) and landscape genetics is an archetypical molecular ecology discipline through its integration of molecular approaches, ecology, population genetics, spatial statistics and geographical tools to address landscape-scale research questions and hypotheses.

Connectivity and the movement of organisms through the landscape matrix are central to landscape genetics (Beier & Noss 1998; Dyer & Nason 2004; McRae & Beier 2007; Braunisch et al. 2010). Both individual- and population-level approaches have been developed for modelling how organisms (and genes) actually move across a landscape, rather than as simply isolation by linear distance. These models have valuable applications in conservation and management (Segelbacher et al. 2010), allowing practical analysis of whether landscape changes have interfered with gene movement (Braunisch et al. 2010) or the presence of corridors facilitated it (Epps et al. 2007; Beier et al. 2011). In the context of conservation and management, methods such as least cost paths, circuit theory, ocean simulations and population networks have provided tools to assess barriers, corridors and overall patterns of connectivity (Beier & Noss 1998; Dyer & Nason 2004; McRae & Beier 2007; Braunisch et al. 2010; Galindo et al. 2010).

Other landscape genetic studies have investigated the impact of the local environment on patterns of genetic variation on the landscape, especially with climate variables (Manel et al. 2010b; Sork et al. 2010; Poelchau & Hamrick 2012). In a study of valley oak (*Quercus lobata*), multivariate genotypes of nuclear microsatellites were significantly associated with climate variables, even after the confounding effects of spatial location were taken into account (Grivet et al. 2008; Sork et al. 2010). One explanation is that climate shapes migration patterns creating similar gradients. Alternatively, immigrants from localities with different climates may be maladapted to the new location, limiting gene flow in both neutral and adaptive portions of the genome. For example, in the context of ecological speciation, landscape genetics can be used to generate more realistic null hypotheses when testing whether adaptation reduces gene flow between habitats by eliminating poorly adapted immigrants (e.g. sunflowers, Andrew et al. 2012).

As the discipline of landscape genetics extended its focus to adaptive genetic variation (Holderegger et al. 2006; Lowry 2010; Manel et al. 2010a; Schoville et al. 2012), it became clear that patterns of adaptive genetic variation could be distinguished from those created by background demographic processes (Beaumont & Balding 2004; Joost et al. 2007). The initial use of AFLPs provided a means of scanning numerous loci across the genome to identify candidate loci, and associations of loci with habitat or climate were a harbinger of the landscape genomics studies we see today. The most compelling of the AFLP-based studies were conducted in concert with other disciplines, such as ecological niche modelling and historical demography (e.g. Freedman et al. 2010 and Manel et al. 2012). Their spatially explicit models provided excellent opportunity to separate the impacts of gene flow, demographic history and selection on the geographical structure of genetic variation.

The availability of NGS tools (Helyar et al. 2011) has facilitated the transition from landscape genetics to landscape genomics. Even for nonmodel systems, it is often possible to link markers to functional genes based on rapidly growing databases of transcriptome sequences (see section on Ecological Genomics and Molecular Adaptation). These tools enable surveys of thousands of genetic variants [single nuclear polymorphisms (SNPs)] found across the genome, greatly facilitating the simultaneous analysis of background genetic structure created by neutral processes, such as population expansion or contraction and gene flow, and identification of candidate genes under natural selection. Similar to the genome-wide association studies (GWAS) that are commonly used for finding genes underlying specific traits in model systems, such as *Arabidopsis thaliana* (Bergelson & Roux 2010; Kover & Mott 2012), several statistical models are available to identify loci that are correlated with environmental gradients while controlling for spatial autocorrelation and demographic effects (Hancock & Di Rienzo 2008; Coop et al. 2010; Kang et al. 2010). One limitation of these models is that they test a large number of SNPs and climate variables one at a time for phenotypic traits that are often polygenic and for which epistatic interactions may be more important than simple additive effects (Holliday et al. 2012; Le Corre & Kremer 2012). Consequently, we are calling for increased use of multivariate statistical approaches (Sork et al. 2013).
A spatially explicit perspective is complementary to evolutionary and ecological genomics. Understanding the landscape and habitat factors shaping the distribution of adaptive genetic variation in nature is important to a comprehensive picture of the evolution of locally adapted genes (Lowry 2010), for instance, genes identified from common-garden experiments testing for local adaptation or associations with phenotypes (Anderson et al. 2010). Conversely, to understand the relationship between patterns of putatively adaptive genetic variation on the landscape and associated phenotypes under selection, it will be necessary to couple landscape genomics with experimental approaches such as ecological genomics and gene expression studies (Stinchcombe & Hoekstra 2008). Under controlled conditions, it is possible to assess which genes are expressed during drought stress, photoperiod changes or other manipulations. Experimental tests of whether environment-associated genes also affect phenotypes and fitness in reciprocal transplants will provide convincing evidence for their adaptive importance and offer insight into the evolutionary mechanisms maintaining geographical structure (Anderson et al. 2013). Increasing cross talk between landscape genomics and evolutionary/ecological genomics is a promising way forward for both fields.

Major challenges

Several challenges have accompanied landscape genomics research from its inception and are now emerging for landscape genomics. Much progress has been made in inferring historical divergence (Gugger et al. 2013), gene flow (e.g. Andreasen et al. 2012) and molecular demography (e.g. Schoville et al. 2012), but incorporating them into landscape genetics and genomics studies remains a challenge. Another big issue pertains to the scale and intensity of sampling, not only for statistical power but also for accurately detecting topographic and environmental effects. Preferably, studies are designed such that an a priori assumption of the important spatial scales is not implicit but can instead be identified (Galpern et al. 2012), and downloaded climate variables are at the same scale as the local samples. Historical environmental data are also highly desirable, especially when considering the evolution and spread of locally adapted alleles, and the availability of such data at suitable scales imposes a limit on the scope of landscape genetics and genomics.

The increasing ease of using SNPs as genetic markers has created opportunities, but challenges remain as recently summarized in the study by Helyar et al. (2011). A few issues are worth highlighting here. The impact of physical linkage on tests of selection is well known in population genetics through the hitchhiking effect (Barton 2000; Schlötterer 2003). With the increasing marker density created by NGS techniques, statistical techniques that account for linkage are desirable for future landscape genomics. Detecting common variants strongly associated with a given environmental variable is relatively straightforward (Coop et al. 2010; Hancock et al. 2011); however, rare and population-specific variants are difficult to identify based on global analyses, just as in genome-wide association studies (Buckler et al. 2009).

Potential future directions

1 Going beyond identifying loci associated with environments is critical for the ongoing development of landscape genomics as a field. For example, by quantifying connectivity across the ranges of species, landscape genomics can provide a novel perspective on the question of how gene flow promotes or constrains adaptation to new habitats. Essential to such studies will be the coupling of landscape genomics with other approaches, especially experiments, demography and niche modelling.

2 Localized introgression can shape genome-wide genetic structure and is amenable to landscape genomics (Fitzpatrick & Shaffer 2007; Kane et al. 2009). In addition to being a potential confounding factor in environmental association analysis, the geographical extent of gene flow between species and the factors driving genome-wide patterns of introgression are important questions (see Hybridization and Speciation).

3 Spatially explicit simulations have made important contributions to the development of landscape genetics (Balkenhol et al. 2009; Cushman & Landguth 2010a; Epperson et al. 2010) but are not routinely used in empirical studies, especially those concerning adaptation. As real populations represent only a single iteration of the evolutionary process (Buerkle et al. 2011), simulations may be essential for gauging the uncertainty around inferences in landscape genomics.

4 The role of epigenetics in plant response to the environment is receiving increased attention (Bosssdorf et al. 2008; Jablonka & Raz 2009; Becker & Weigel 2012); however, we know little about its prevalence in natural populations. It is now possible to survey DNA sequence variation and epigenetic marks, such as DNA methylation, simultaneously (Feng et al. 2011). A landscape genomic analysis could provide first-level evidence for the association of both genetic and epigenetic variation with environmental gradients.
Ecological genomics and molecular adaptation: back to the future—by Sean M. Rogers, Louis Bernatchez, Aurelie Bonin and Jon Slate

From the time when E.B. Ford ‘invented’ the field of ecological genetics (Ford 1964), followed by the advent of electrophoretic surveys of genomic variation in 1966, molecular ecologists have been challenged to explain the large amounts of standing genetic variation in populations and the degree to which this variation can be explained by adaptive evolution (Lewontin 1991). The field of ecological and evolutionary genomics (EEG) emerged from efforts to understand the genomic mechanisms underlying organismal responses to abiotic and biotic environments (Feder & Mitchell-Olds 2003; Ungerer et al. 2008). This framework proposed experimental approaches towards elucidating the genomic architecture of ecologically important traits, how these traits affect fitness and the evolutionary processes by which these traits may arise and persist—the overarching objective of linking genotype to phenotype and ultimately fitness (Dalziel et al. 2009).

The merging of genomics with ecology includes more than just the incorporation of a new genomic toolbox. Emerging technologies are providing unparalleled insight into the genomes of species, leading to new questions that need to be tested, while existing questions can be addressed in ways that were not previously possible (Barrett & Hoekstra 2011). In addition, Feder & Mitchell-Olds (2003) predicted that the promises of large-scale genomic data would not change the fact that ecological and physiological knowledge would remain crucial for the interpretation of genomic and postgenomic data. Molecular ecologists have indeed risen to this challenge by (i) demonstrating the significance of standing genetic variation to adaptive evolution (e.g. Colosimo et al. 2005); (ii) revealing that even small changes in the sequences of genes (including regulatory regions) may result in striking adaptive evolution (e.g. Hoekstra et al. 2006; Chan et al. 2010; Rosenblum et al. 2010); (iii) measuring selection and validating candidate genes (e.g. Barrett et al. 2008, 2011; Bonin et al. 2009; Gratten et al. 2012); (iv) elucidating the genetic bases of microevolutionary changes in natural populations (e.g. Gratten et al. 2008, 2012; Johnstone et al. 2011); (v) determining the genomic architecture of adaptive evolution and ecological speciation (e.g. Kane & Rieseberg 2007; Rogers & Bernatchez 2007; Nosil et al. 2012); (vi) establishing the importance of plasticity in adaptive evolution (e.g. Ghalambor et al. 2007; McCAirns & Bernatchez 2010; McCAirns et al. 2012); and (vii) estimating the role of life history trade-offs in shaping patterns of genome-wide gene expression and linking these trade-offs with adaptive divergence (e.g. Derôme et al. 2006; St-Cyr et al. 2008; Colbourne et al. 2011).

Nonetheless, these studies and increasingly others have revealed additional questions and highlighted significant challenges for the future, including broadening the scope to include a wider range of organismal diversity, especially keystone species. EEG studies should also pay greater attention to relatively undisturbed habitats in the native range of species, unique ecology and behaviours, and long-term synthetic and natural ecological experiments (Feder & Mitchell-Olds 2003; Gratten et al. 2008, 2012; Grant & Grant 2011).

Regardless of the organism being studied, the progress over the last 10 years of EEG research highlights at least six priorities that should be considered over the next two decades.

**Priority 1: Extended evolutionary theory**

1. What does evolutionary theory predict for the consequences of standing genetic variation, dominance, molecular quantitative genetics and nongenetic inheritance (e.g. epigenetic inheritance, parental effects) during adaptive evolution?
2. What does theory predict for the outcome of phenotypic plasticity during adaptive evolution?
3. Beyond single locus traits, what are the predicted consequences of polygenic inheritance during adaptive evolution?

**Priority 2: Ecological annotation of genes**

1. Annotation of genes is the main limiting factor when making functional inferences for genomic variation, especially in nonmodel species (Pavey et al. 2012).
2. There is an urgent need for ecological gene annotation, which will require better data integration and functional analyses.

**Priority 3: Phenomics**

1. Phenotypes are the variation that selection can see, so greater attention should be paid to the measurement and reporting of phenotype-environment associations.
2. How do different levels of biological organization (from the gene to the different steps of regulation, transcription, signal transduction to networks and pathways) give rise to this variation?
3. Acquiring detailed phenotypic data will be as crucial as molecular data in building genotype-phenotype maps (Houle et al. 2010).
Priority 4: Become predictive about organismal response to environmental variation

1. There is a need for further analytical methods development that allows robust discrimination between the consequences of drift and selection.
2. What are the consequences of variation in genetic architecture?
3. Will genome scans of temporal changes in allele frequency reveal targets of selection?

Priority 5: Become predictive about organismal solutions to environmental heterogeneity

1. Under what ecological and evolutionary conditions do organisms respond to environmental heterogeneity by local adaptation (which generates population structure) vs. the maintenance of balanced polymorphisms or a plastic response of phenotypes?
2. How does genetic variation affect population demography and vice versa?
3. Does the stability of communities depend upon ongoing eco-evolutionary feedbacks—or is evolution ecologically trivial?

Priority 6: Experiments

1. Ultimately, we need to move towards a holistic approach and aim to fully integrate multidimensional high-throughput ‘omics’ measurements (i.e. ecological systems biology). This should include dynamic, sequential common-garden experiments in the laboratory and in nature.
2. Also needed is validation of the adaptive significance of candidate genes: how repeatable will the findings be?

Given the rise of molecular ecology since the early 1990s, we are optimistic about the future. These priorities should nonetheless serve as a reminder to students of ecological genomics that there is much work to be carried out. Future advances will continue to require re-engineering of scientific attitudes, training and a focus on multidisciplinarity (Feder & Mitchell-Olds 2003).

Speciation and hybridization—by C. Alex Buerkle, Tatiana Giraud, and Alex Widmer

Interest in the evolutionary and genetic processes that lead to the appearance and maintenance of new species has been a driving force in the development of evolutionary theory and genetics. Much of current speciation research involves understanding how isolation between populations might arise and increase as a result of evolutionary processes, in different spatial and ecological settings, and how isolation might be maintained when it is tested by potential hybridization. Presently, there is great interest in ‘speciation with gene flow’, both in terms of the origin and maintenance of diversity. This includes ecological speciation, that is, the possibility of reproductive isolation arising directly from adaptation to ecological conditions. This work provides a link between ecology and evolution (Egan & Funk 2009; Schluter 2009; Giraud et al. 2010; Gladieux et al. 2011) and draws attention to the potentially short timescale over which ecological determinants might give rise to evolutionarily relevant isolation.

Until recently (as is true for much of molecular ecology), empirical studies were severely limited by our ability to assay genomic variation in natural populations. This situation has changed dramatically, and studies for any organism can now be based on orders of magnitude more individuals and loci than before. However, access to population genomic data comes with new challenges, in particular for handling and analysing huge data sets. At this stage, we have only just begun to apply these to key questions in speciation and hybridization and to recognize and accommodate the new complexities that arise from sampling the genome at high resolution.

In the last few years, we have begun to document population genomic variation among species and their hybrids, but we have made less progress towards tying patterns of variation to underlying evolutionary processes. Previously, this was a substantial challenge even with single or small sets of loci, because the inferred evolutionary parameters are only applicable to the models that we specify (Excoffier & Heckel 2006; Wegmann et al. 2010). This challenge is now compounded when sampling a larger fraction of the genome, as it involves analysis of many loci with potential differences in their evolutionary histories (mutation, recombination, drift and background selection, positive selection, etc.). In the sections that follow, we expand on this overview with more specific examples of how the study of speciation and hybridization might develop further.

Theory

1. Dobzhansky-Muller (DM) incompatibilities have been identified and now dominate discussions of the genetics of reproductive isolation (e.g. Lee et al. 2008; Burton & Barreto 2012). However, theory shows that DM incompatibilities can only serve as an effective
isolation mechanism if F1 hybrids have near zero fitness (Gavrile 1997). Progress towards understanding their role in isolation will come from additional modelling, including using existing theoretical models with high levels of abstraction (e.g. Barton & Rodríguez de Cara 2009) for generating predictions for parameters that can be estimated in experimental and natural populations.

We also need to develop models that better predict the evolution of reproductive isolation with genetic distance, depending on the underlying causes and their effects on fitness, for example a snowball effect due to the accumulation of DM incompatibilities (Matute et al. 2010; Moyle & Nakazato 2010) vs. a linear increase in reproductive isolation due to ecological isolation (Gourbière & Mallet 2010; Giraud & Gourbière 2012).

Experiments and study of natural cases

Estimates of the fitness effect of DM incompatibilities in a larger number and diversity of organisms are needed, as Drosophila may not be a good model for all eukaryotes (see data in fungi: Gourbière & Mallet 2010; Giraud & Gourbière 2012). In addition, the generality of the model of accumulation of DM incompatibilities vs. that of reproductive isolation evolution by adaptation to different environments remains to be assessed (Egan & Funk 2009; Schluter 2009). Indeed, while the accumulation of DM incompatibilities with genetic distance has been documented in a variety of organisms (Dettman et al. 2007; Anderson et al. 2010; Matute et al. 2010; Moyle & Nakazato 2010), a whole-genome screen has indicated their absence in yeasts (Kao et al. 2010). It would also be interesting to further study cases of ecological speciation and assess their prevalence, and in particular those where adaptation generates reproductive isolation through pleiotropy (Giraud et al. 2010; Gladieux et al. 2011; Servedio et al. 2011).

Population genomics of experimental and natural populations may reveal whether and how genomic regions involved in adaptation to a particular environmental factor contribute to isolation. It is, for instance, an open question whether adaptive divergence between species leads to dysfunction and isolation in hybrids or merely to selection against maladapted immigrants (Nosil et al. 2009; Giraud & Gourbière 2012).

Inferences of the evolutionary history of loci have the potential to indicate the relative times at which isolation was obtained for loci associated with different phenotypic components of isolation (e.g., loci associated with flowering time vs. hybrid sterility: Hey & Nielsen 2004; Gladieux et al. 2011; Cornille et al. 2012) and to thereby truly reconstruct the history and initial causes of speciation (but see Gaggiotti 2011; Sousa et al. 2011; Strasburg & Rieseberg 2011 for a discussion of the challenges associated with such an approach).

NGS may finally allow tests of the genic view of speciation (Wu 2001). Many studies find heterogeneous genomic divergence across the genome (Fig. 4). It remains to be assessed whether this is typical for early stages of divergence and speciation, for example because adaptation has a complex genetic basis that leads to divergence in multiple genomic regions from the beginning. Growing knowledge of genomic variation in recombination rates will play a large role in tying the patterns of heterogeneous population genomic divergence to the underlying evolutionary processes (Via & West 2008; Nosil et al. 2009; Nachman & Payseur 2012; Roesti et al. 2012; Via 2012). Suppressed recombination can allow divergence to accumulate at a higher than expected rate and to prevent the breakdown of adaptive allelic combinations in the face of gene flow (Rieseberg 2001; Turner et al. 2005; Kirkpatrick & Barton 2006; Noor & Bennett 2009; Turner & Hahn 2010; Joron et al. 2011).

Empirical study of hybridization and isolating barriers

In the coming years, we will have the opportunity to determine the extent to which the fitness of hybrids in natural populations is predicted based on genetic mapping results. Beyond the assessment of the contribution of DM incompatibilities to isolation, taking results from laboratory and controlled crosses into natural populations will probably lead to important advances in the genetics of speciation. Expectations for the transferability of any trait mapping result to a new population must be informed by our knowledge of genetic polymorphism for trait loci, epistatic interactions among genotypes and the effects of environment on organismal trait expression. For example, we already know from mapping studies that polymorphism for isolation traits segregates within populations (Scopece et al. 2010; Rieseberg & Blackman 2010; Lindtke et al. 2012). Consequently, isolating barriers could be polymorphic among populations with different genetic compositions and ecological settings. A greater understanding of this potential polymorphism, particularly in the context of the timing of origin for different components of isolation, is likely to affect our conception of species’ origin and maintenance.

Similarly, future population genomic studies of hybrid zones will teach us more about how selection
and introgression in hybrids is (or is not) tied to population genomic differentiation between species. Furthermore, advances in analytical methods will provide estimates of the recency of introgression and about the strength and direction of selection.

3 An additional dimension of the genomics of hybridization involves changes in transposable element frequency and activity. At the sequence level, we know that historical, homoploid hybridization has led to TE proliferation and genome expansion (Rieseberg 1991; Baack et al. 2005; Ungerer et al. 2006; although see Kawakami et al. 2011). Hybridization can also directly lead to TE activation and expression, and gene expression misregulation (Parisod et al. 2010; Hegarty et al. 2011; Buggs 2012; Debes et al. 2012). A question little studied so far is whether the mechanisms controlling TE are involved in hybrid inviability and sterility. For instance in fungi, repeat-induced point mutation (RIP) specifically changes repeated sequences, and therefore, aneuploid hybrids may be inviable or sterile because RIP inactivates genes in duplicated chromosome arms (Galagan & Selher 2004; Giraud et al. 2008). RNA interference (RNAi) may also have such effects of inactivation of abnormally duplicated sequences in hybrids.

Kinship, parentage and behaviour—by Dany Garant and Lisette Waits

The last 20 years

The field of kinship, parentage and behaviour (henceforth KPB) has been an important component of molecular ecology studies over recent decades. This field has contributed much to our understanding of mating systems, behavioural ecology, sexual selection and the impacts of inbreeding on individual fitness. Yet, the last 20 years have seen major changes in the techniques being applied and the questions being targeted. For instance, studies published two decades ago rarely focused on more than one population and often on only a few families when assessing patterns of relatedness, mate choice and/or reproductive success (see Achmann et al. 1992; Patton & Smith 1993; Signer et al. 1994 for examples). Development and reliability of techniques were also important in these early studies as allozymes, mtDNA and minisatellites were commonly used (Burke & Bruford 1987; Wetton et al. 1987; Chakraborty et al. 1988; Lehman et al. 1992). Nonetheless, at that time, results provided by such studies were regarded as important breakthroughs in the field of KPB as they improved our understanding of behavioural and ecological processes that were, up until then, based solely on observations.

Since then, the number of KPB studies published in Molecular Ecology increased between 1990 and 2005 and has remained fairly constant over recent years with c. 10% of papers (between 25 and 35 papers per year) on average over the last 5 years. Microsatellite loci have become the marker of choice for KPB studies (Tautz 1989; Jones & Ardren 2003), and researchers have shown that KPB research can be conducted using low-quality DNA obtained from faecal and hair samples (Bradley et al. 2004; Walker et al. 2008; DeBarba et al. 2010; Stenglein et al. 2011), which greatly increases our ability to apply this research to rare and hard-to-capture species. Recent studies have moved on from a focus on tech-
niques to a more informative emphasis on processes and now typically involve the analyses of several populations and thousands of individuals (see references below).

Current state of the field

The field of KPB currently includes studies defining mating systems (Wright et al. 2012) and characterizing processes underlying sexual selection (While et al. 2011), such as mate choice (Wang & Lu 2011). KPB research also focuses on estimating determinants of reproductive success (Thöß et al. 2011) and describing potential inbreeding effects (Nielsen et al. 2012) or avoidance (Waser et al. 2012), as well as the social structure of populations through assessments of relatedness, dispersal patterns and networking approaches (Rollins et al. 2012). Furthermore, the number of studies reconstructing extensive wild pedigrees over multiple generations continues to increase providing a rich resource for KPB studies.

An overview of the studies published in 2011 in *Molecular Ecology* revealed important differences in the methods and approaches between KPB studies and research conducted in other subfields in molecular ecology. Like most molecular ecology studies, microsatellites are widely used, but KPB studies in *Molecular Ecology* on average use a lower number of markers (c. 10 loci in 2011) than studies conducted in other subfields (c. 20 microsatellite loci in 2011; see Rieseberg et al. 2012). Why? Part of the answer may lie in the kind of questions being targeted in this field, which could often be satisfactorily answered using a more limited number of markers. For example, studies of parentage can sometimes reach a good assignment success by using only a handful of highly polymorphic markers (e.g. Oddou-Muratorio et al. 2011). However, a low number of markers could be problematic and limit power and resolution of methods, especially when one is interested in both conducting parentage assignments and quantifying genetic diversity of possible parents (Wetzel & Westneat 2009). Most studies should now aim at increasing the number of markers and making good use of the recent developments in NGS techniques that allow rapid and inexpensive development of panels of markers for multiple species (Glenn 2011; Guichoux et al. 2011). However, as much of the research in KPB involves matching genotypes, increasing the number of loci will require that more attention be paid to genotyping errors.

The mean number of individuals being analysed in KPB studies (on average >1300 individuals per study in 2011) is higher than in other subfields (on average <500 individuals per study in 2011; see details in Rieseberg et al. 2012). This observation is likely related to questions of interest in KPB, which are now typically addressed across several populations with as many individuals per population as possible. Finally, the statistical methods employed and software packages developed over the years are suitable to most biological systems (reviewed in Blouin 2003; Jones & Ardren 2003) and are also generally accessible (Kalinowski et al. 2007; Wang 2007, 2011; Jones & Wang 2010). However, we still lack a comparison of the performance of these different approaches/software in estimating variables of interest in a context-specific manner. Such a comparison should be a priority for KPB, because it would provide useful guidance for assessing the most productive approach/software in a given species/population/environmental context.

Future promise and challenges

Even though many improvements have occurred over the last decades, several critical key elements remain less well defined in studies of KPB:

1 Measuring lifetime fitness in the wild should be a major focus of KPB studies in the coming decades. While central to evolutionary biology, lifetime reproductive success and longevity as proxies of fitness are still often difficult to assess empirically (Clutton-Brock & Sheldon 2010), despite the advances in molecular techniques and the increasing number of analytical software programs available.

2 Future studies should quantify relevant indicators of human-induced variation on populations and aim at generalizing findings across a broad range of natural environmental conditions. Both human-induced and natural variation in environment should have tremendous impacts on the processes targeted by KPB studies. For example, recent studies have suggested that the extent of social organization could be impacted by changes in population density and ecological conditions (Messier et al. 2012; Schradin et al. 2012) and hunting pressure (Jedrzejewski et al. 2005; Rutledge et al. 2010; Onorato et al. 2011). Others have shown theoretically (Blyton et al. 2012) and empirically (Bergon et al. 2011) that mating system may be variable depending on resource availability. Elucidating the interaction between change in ecological conditions and evolutionary mechanisms will allow researchers to assess the importance of eco-evolutionary dynamics in a broader range of systems (Pelletier et al. 2009).

3 More studies should conduct research across multiple habitats to quantify spatial variation but also to develop and maintain long-term ‘individual-based’ studies to accurately describe temporal variation (see...
Clutton-Brock & Sheldon 2010). Developing such long-term studies will also allow more researchers to reconstruct pedigrees in different wild populations. Pedigrees are valuable tools for obtaining precise inbreeding coefficients and estimating important quantitative genetics parameters (see Dunn et al. 2011; Nielsen et al. 2012; Richards-Zawacki et al. 2012 for recent effective reconstructions). Having several pedigrees available across different biological systems will also help validate/refute the patterns found with neutral molecular markers, as well as document the efficiency of different markers for reconstructing pedigree relationships (Garant & Kruuk 2005).

4 Studies of KPB over the next years will benefit from employing a multitool and multitrait approach (include several traits and different markers) but also from advances in NGS (for instance through the development of SNP markers in nonmodel species—see Van Bers et al. 2010 for example), in mapping and gene expression. For example, Laine et al. (2012) recently studied nine-spined stickleback (Pungitius pungitius) and used available mapping information to separate markers in functional categories. Significant heterozygosities–behaviour correlations were detected with functional markers but not when all markers were combined.

Overall, the field of KPB has provided important advances over the last decades to our understanding of evolutionary, ecological and biological processes and will face stimulating challenges and prospects in the near future.

Conclusions

Molecular ecology represents a spectacularly successful example of cross-disciplinary science, in which the tools and methods of molecular biology, genomics and bioinformatics have been merged with the theory, concepts and approaches of organismal biology, including ecology, evolution, conservation and behaviour. As can be seen from the subdiscipline perspectives outlined above, the questions addressed by molecular ecologists include longstanding discipline-specific problems that can now be investigated with new tools and approaches (e.g. Lodge et al. 2012; Malek et al. 2012; Orozco-terWengel et al. 2012; Parchman et al. 2012; Pompanon et al. 2012; Tedersoo et al. 2012), as well as new questions that have resulted from merging formally disparate disciplines (e.g. Kraaijeveld et al. 2012; Nosil & Feder 2012; Ozawa et al. 2012; Simms & Porter 2012).

Over the past two decades, molecular ecology has experienced huge advances in genotyping, from allozymes to RFLPs to minisatellites to AFLPs to microsatellites to genotyping arrays to NGS (Mobley 2012; Rowe et al. 2012). Likewise, there have been impressive advances in the analytical approaches employed in molecular ecology, with coalescent and landscape genetic approaches providing considerably more robust inferences about the demographic and geographical history of populations than were previously possible (Beaumont et al. 2010; Sork & Waits 2010; Storfer et al. 2010; Andrew et al. 2012; Holderegger & Gugerli 2012; Li et al. 2012; Lohse et al. 2012; Salzburger et al. 2011).

While most molecular ecology studies over this period have addressed questions about the biology of organisms and communities, this is slowly changing, as ecological and evolutionary information are increasingly being employed to identify and functionally characterize ecologically important genes and their products (Blackman et al. 2011; Beuler-Martinez et al. 2011; Johnstone et al. 2011; Kent et al. 2011; Bloomer et al. 2012). This general approach includes various types of population genomic scans for ‘outlier loci’ (Bonin et al. 2009; Paris et al. 2010; Buckley et al. 2012; Collin & Fumagalli 2012; Laurent et al. 2012; Midamegbe et al. 2011; Prunier et al. 2012). In addition, information about geographical location, habitat, phenotype, ecological community and so forth are being used for candidate gene discovery and to make inferences about allelic function (Cox et al. 2011; Fischer et al. 2011; Johnstone et al. 2011; Gratten et al. 2012; Orsini et al. 2012; Manel et al. 2012; Paris & Despres 2012).

Despite the diversity of the questions and problems being addressed by molecular ecologists, many of the challenges are similar. In particular, two themes stand out. One of these concerns the difficulties associated with managing, analysing and integrating the very large data sets that are becoming increasingly commonplace in molecular ecology studies. As noted by Paterson and Piertney (2011):

‘The challenge will be ensuring that the onslaught of data that accompanies approaches such as NGS, genome scans [and array- and sequence-based analyses of the transcriptome and epigenome] can be coupled to appropriate ecological and phenotypic metadata to allow meaningful analysis to be undertaken.’

A second general challenge concerns the inadequacy of our analytical methods toolbox (despite advances over the past 20 years) for making inferences about the ecology and evolution of organisms, as well as the ecological effects of molecular variation. Frequently discussed needs include better analytical tools for (i) distinguishing the genomic consequences of different ecological and evolutionary processes; (ii) estimating
the timing of gene flow during population divergence; (iii) inferring phylogeography using many genes, populations and species; (iv) distinguishing between contemporary and historical effects of the landscape on patterns of genetic variation; (v) establishing the ecological functions of genes and alleles in natural populations; and (vi) estimating lifetime reproductive success in natural populations from molecular marker data. Because natural selection increasingly appears to play an important role in shaping patterns of molecular variation within and among species (Sella et al. 2009), it is important that these analytical tools be robust to non-neutral variation.

Despite these challenges, the future of molecular ecology is bright. New genotyping and analytical tools are allowing us to address key questions and problems with a rigour that was not possible even a decade ago. Of greater importance, however, has been the training of a new generation of molecular ecologists with diverse skills—from fieldwork to computational biology to molecular functional studies. We are confident that this next generation of molecular ecologists has the conceptual and analytical skill sets to successfully respond to the challenges faced by our discipline.

Acknowledgements

We thank the participants in the Molecular Ecology Symposium and Online Forum for many of the ideas put forward in this article. We also thank everyone at Wiley-Blackwell, and Sarah Burrows in particular, for making the symposium and road map paper possible.

References


Amaral AR, Beheragaray LB, Bildmann K et al. (2012) Influences of past climate changes on historical population structure and demography of a cosmopolitan marine predator, the common dolphin (genus Delphinus). Molecular Ecology, 21, 4854–4871.


Oliver KM, Degnan PH, Burke GR et al. (2010) Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annual Review of Entomology*, 55, 247–266.


All authors contributed to the writing of the paper.

© 2013 Blackwell Publishing Ltd