Conservation Genomics of Threatened Animal Species

Cynthia C. Steiner,[1] Andrea S. Putnam,[2] Paquita E.A. Hoeck,[1] and Oliver A. Ryder[1]

[1]Institute for Conservation Research, San Diego Zoo Global, Escondido, California 92027; email: PHoeck@sandiegozoo.org; CSteiner@sandiegozoo.org; ORyder@sandiegozoo.org

[2]Department of Life Sciences, San Diego Zoo Global, San Diego, CA 92112; email: APutnam@sandiegozoo.org

Annu. Rev. Anim. Biosci. 2013. 1:261–281

First published online as a Review in Advance on January 3, 2013

The *Annual Review of Animal Biosciences* is online at animal.annualreviews.org

This article's doi:
10.1146/annurev-animal-031412-103636

Keywords

population demography, adaptive variation, inbreeding depression, hybridization, disease susceptibility

Abstract

The genomics era has opened up exciting possibilities in the field of conservation biology by enabling genomic analyses of threatened species that previously were limited to model organisms. Next-generation sequencing (NGS) and the collection of genome-wide data allow for more robust studies of the demographic history of populations and adaptive variation associated with fitness and local adaptation. Genomic analyses can also advance management efforts for threatened wild and captive populations by identifying loci contributing to inbreeding depression and disease susceptibility, and predicting fitness consequences of introgression. However, the development of genomic tools in wild species still carries multiple challenges, particularly those associated with computational and sampling constraints. This review provides an overview of the most significant applications of NGS and the implications and limitations of genomic studies in conservation.

INTRODUCTION

There is considerable anticipation that next-generation sequencing (NGS) technologies will provide new opportunities to produce genomic data rapidly from a broad swath of species outside the standard biomedical models, especially because NGS costs continue to decline (1). In particular, high-throughput NGS and analyses on genome sequence data and their variation are readily applied to species that are genome enabled, i.e., those that have an assembled reference genome or for which an assembled genome from a closely related species is available. The application of genomic technologies in the field of conservation biology has already resulted in a surge of research and reviews (1–14). Such large-scale sequencing data sets represent an extraordinary source of information that can shed light on aspects of the biology of wildlife species relevant to conservation assessments, monitoring, and management.

Genomic analyses previously limited to model organisms now can be applied to threatened species to estimate recent demographic events, genetic variation, and population structure through population-genomic approaches. The role of natural selection at the genome level and the identification of loci associated with fitness, including local adaptation, inbreeding depression, or disease susceptibility, can also be dissected by implementing quantitative trait loci (QTL) analyses and genome-wide association studies (GWAS). In addition, identifying conservation units for protection, management, and recovery may now be more clearly resolved by using phylogenomics and population genomics (**Table 1**).

All the aforementioned applications rely upon the detection of genome-wide polymorphisms among individuals, populations, or species in the form of single-nucleotide polymorphisms (SNPs), insertion-deletions (indels), or copy number variants (CNVs). This genetic variation can be identified by high-throughput sequencing of whole genomes from multiple individuals or from a reduced representation of genomic fragments as implemented in the restriction-site-associated DNA sequencing (RAD-seq) technique (see sidebar, NGS Approaches). Efforts to detect and map genetic polymorphisms and to conduct genome-wide comparisons of wild organisms also benefit from accurate reference genomes, which must often be assembled de novo. The assembly of de novo genome sequences is progressing rapidly, but taxonomic coverage is still limited, because the efforts to use NGS have been underway for a relatively short period of time. The extent of necessary genomic data, the cost per sample, the DNA quality required, and the equipment needed depend on the NGS platform to be employed and the conservation questions of interest (reviewed in 2).

In this review, we discuss the application of genomic analyses to new and long-standing dilemmas in conservation management, such as demographic analyses, genetic variation associated with local adaptation and fitness, the basis for inbreeding depression, detection and timing of hybridization events, and identification of loci associated with disease susceptibility. We also emphasize the potential pitfalls of genomic technologies and the statistical and computational challenges related to the ever-growing acquisition of genomic data.

Next-generation sequencing (NGS): high parallel DNA sequencing in which hundreds of thousands or millions of reads (sequences) are produced in one run of an automated sequencer; the best known platforms are the Roche 454 FLX Titanium system, Illumina's Genome Analyser (Solexa), and ABI's SOLiD™

Local adaptation: heritable changes in the genotype or phenotype of a population that result in increased fitness within a specific environment

Inbreeding depression: the decrease in fitness owing to inbreeding or random genetic drift

Quantitative trait locus (QTL): a locus that controls a quantitative phenotypic trait, identified by showing statistical association between genetic markers surrounding the locus and phenotypic measurements; QTL are typically mapped by crossing individuals from different populations, which generates an F2 or backcross mapping population

DEMOGRAPHY

Selectively neutral genetic markers (e.g., microsatellites and silent sites) have traditionally been the cornerstone of identifying recent and historic demography, such as population size fluctuations, admixture, gene flow, and geographic population structure. However, to model and analyze complex demographic histories, accurate estimates of summary statistics of genomic variation, such as nucleotide diversity, effective population size (N_e), and recombination, are often necessary because they can vary greatly across a genome. Although traditional molecular approaches have been used successfully to examine demographic

history, greater analytical and statistical power has emerged from the availability of genome-wide data (e.g., SNPs) (3).

In particular, estimating the patterns of recent and historic population size changes has numerous conservation implications. For example, genome-wide SNP analysis of the two extant orangutan species allowed estimation of current and ancestral N_e, recombination rates, and speciation times (15). Demographic modeling found that the two species have complex and very different evolutionary histories. Surprisingly, the authors found greater genetic variation among Sumatran orangutans than Bornean orangutans, despite a much smaller population size. Conservation biologists are often interested in whether a species' population size has been influenced by factors such as climate change or human-mediated events. Genome-wide data will help elucidate these questions by enabling more precise estimates of the timing and extent of population bottlenecks and expansions.

It is well known that population size fluctuations can have a large effect on the ability to detect genetic signatures of selection (16–21). The availability of genomic data increases the statistical power to tease apart changes in genetic variation and allele frequencies due to demography versus selection (22, 23). Maximum likelihood approaches that account for both demography and selection have proven amenable to NGS (reviewed in 24). For example, Nielsen et al. (22) developed a likelihood method that estimates demographic parameters in multiple populations using a joint frequency spectrum and SNPs that are fixed between populations. For species for which SNPs can be mapped to a reference genome, patterns of linkage disequilibrium (LD) may be used to estimate N_e (25, 26). For unmapped SNPs, the site frequency spectrum (SFS) can be used to estimate many demographic parameters (15, 27, 28). Both LD and SFS data across several wild and domesticated canids were used to identify bottleneck severity and timing among species and breeds (29).

Gene Flow and Units of Conservation

Conservation biologists frequently seek to estimate the spatial scale of gene flow to examine historical and contemporary population connectivity. These analyses can guide the designation of conservation units and their boundaries at both the intra- and interspecific level. The use of genomic data to infer population structure and to identify units of conservation has been covered extensively by other recent reviews (2, 30–32) and is covered here only briefly.

In short, gene flow among populations contributes to maintaining genetic diversity, which is fundamental to ensure a species' sustainability and reduce the risk of extinction (33, 34). Very rare events of gene flow can play a disproportionate role in contributing to genetic diversity (35). Genomic data can provide more power to detect these infrequent events, in particular when the Wright's fixation index (F_{ST}) estimator is low (2). For example, Gompert et al. (36) looked at population structure between 12 populations of *Lycaeides* butterflies, including the endangered Karner blue (*Lycaeides melissa samuelis*), using SNPs from across the genome. The authors found that admixture was more widespread than was previously noted using traditional molecular approaches with many fewer markers. Similarly, an extensive SNP survey of wolflike canids found previously unidentified evidence of extensive gene flow (37), which highlights the potential of improved resolution in genome-wide surveys with greatly increased data.

Genome-wide data can also provide greater power in taxonomic classification compared with traditional approaches that use only a handful of genetic markers. A recent study of population history in extant elephants and their extinct relatives examined sequence data from almost 400 loci. Both phylogenetic analyses and demographic models revealed a deep divergence between African savanna and forest elephants, which had been thought by some to be

Genome-wide association studies (GWAS): studies in which a dense array of genetic markers, which capture a substantial proportion of common variation in genome sequence, is typed in a set of DNA samples that are informative for a trait of interest; the aim is to map susceptibility effects through the detection of associations between genotype frequency and trait status

Conservation unit: a population of organisms in the wild that is considered distinct for purposes of conservation; refers to either evolutionarily significant unit or management unit

Phylogenomics: analyses that involve genomic data and phylogenetic reconstructions

Single-nucleotide polymorphism (SNP): a DNA sequence variation that occurs when a single nucleotide in the genome differs between individuals; abundant, codominantly inherited markers that are highly adaptable to large-scale, cost-effective, automated genotyping

Indels: mutation class that includes insertions and deletions as well as the combination of both; unless the length of an indel is a multiple of 3, it produces a frameshift mutation

Table 1

Analyses	Quantitative trait locus mapping	Association studies	Kinship	Pedigree reconstruction	Selection	Functional characterization	Genetic variation	Population genomics	Inbreeding-outbreeding depression	Hybridization	Phylogenetics/ Phylogenomics	Phylogeography	Comparative genomics
Main goal	Identify genetic regions associated with phenotypes of interest	Establish phenotype-genotype interactions	Determine levels of consanguinity or inbreeding	Identify genealogical relationships	Identify genetic regions under selection	Dissect the role of gene regulation versus coding-gene variation in evolutionary processes	Determine the genetic diversity or differentiation of populations and species	Separate locus-specific effects (selection, mutation, recombination) from genome-wide effects (genetic drift, inbreeding, gene flow)	Identify genetic basis of inbreeding depression or predict outbreeding depression in the wild	Detect hybridization events (recent and ancestral), frequency of introgression, behavior of introgressed loci, loci involved in heterosis	Determine the phylogenetic relationship among different taxa	Determine patterns of geographic distribution of genetic diversity (haplotypes)	Identify patterns of genome evolution among species (selection, gene duplication/ deletion)
Tools													
Msats	x	x	x	x	x		x	x	x	x	x	x	
mtDNA							x		x	x	x	x	
Amplified fragment length polymorphism	x	x	x	x	x		x	x	x	x	x	x	
Whole-genome sequencing					x		x	x	x	x	x	x	x
Single-nucleotide polymorphisms	x	x	x	x	x		x	x	x	x	x	x	
Candidate genes	x	x			x		x		x	x	x	x	
Expressed sequence tag	x	x			x	x	x	x	x	x	x	x	x
Microarrays						x							x
Phenotype	x				x	x			x	x	x		
Examples	Deer mouse (63), Cave tetra fish (64), Three-spined stickleback (119)	*Anopheles* (121)	Subdesert mesite (123), Cape Fear shiner (124)	Steelhead trout (127)	Deer mouse (76), Ocellated lizards (130)	Lake cichlid fish (51), Zebra finch (52), Primates (131)	Bison (133), House sparrow (134)	Wolf (37), Chimpanzee (39)	Wolf (91), Bighorn sheep (137), Florida panther (138)	Bison (139), Carrion crow (140)	Ruminants (142), Primates (143)	Pitcher plant mosquitoes (148), Birds (149)	Gorilla (151,152)
Reviewed in	61, 71, 120	122	125, 126	128, 129	23, 73, 77	1, 13, 72, 132	135, 136	9, 61	2, 88	101, 141	144–147	150	153, 154
Potential in conservation	Identify adaptive or fitness-related loci	Determine genetic basis of adaptive traits or disease susceptibility in the absence of pedigree data	Study evolution of cooperative breeding in the wild, minimize average kinship in captive-breeding programs	Characterize pedigree in the wild for studying quantitative trait loci, inbreeding depression	Adaptation to captivity	Characterize functional genes involved in local adaptation and fitness, identify conservation units	Insights into population history, inbreeding, demography, gene flow, effective population size, and disease susceptibility	Determine demographic history of populations (bottlenecks, expansion, admixture events)	Genetic rescue and translocation of individuals into the wild	Assessment of levels of introgression for population management	Identification of species and conservation units, establishment of phylogenetic relationship among taxa of conservation concern	Identify phylogeographic patterns (admixture) and define conservation units	Identify the genetic basis of unique traits, selective pressures in genetic regions

NGS APPROACHES

Numerous NGS approaches have been developed to perform marker discovery and genotyping. Some of the most promising methods that may be applicable to conservation biology studies include:

DNA-Based Methods

All methods involving the use of genomic DNA and NGS proceed under the same key steps, which include random shearing or digestion with one or more restriction enzymes, size selection for an optimal length, sequencing using a NGS platform, and identification of genetic polymorphisms (reviewed in 117).

- **RAD sequencing (RAD-seq):** In restriction-site-associated DNA (RAD) sequencing, genome DNA is cut using one or two restriction enzymes and is indexed (tagged) with a specific oligonucleotide signature (66). These fragments are ligated with adaptors, randomly sheared and amplified, which produces many copies of the DNA adjacent to the restriction enzyme–recognition site. RAD-seq reduces the proportion of the genome sequenced, which enables each marker to be sequenced at high coverage. RAD-seq can be employed in SNP and microsatellite discovery, genotyping, genotype-phenotype association mapping, genome assembly scaffolding through linkage mapping, hybridization and gene flow analysis, phylogeography, and other population genetic analyses.
- **Target DNA sequencing:** In the target DNA sequencing, genomic fragments are captured by hybridization with probes, multiplex-amplified or targeted circularization approaches, and DNA libraries are prepared (118). Hybrid selection methods apply immobilized oligonucleotides on either microarrays or beads for the enrichment of genomic targets from a modified DNA sample. In multiplex polymerase chain reaction (PCR), complex primer sets can be used to selectively amplify targeted regions prior to preparing sequencing libraries. Targeted genomic circularization directly captures one strand of the genomic DNA target by converting the target probe DNA into a circle using in-solution capture oligonucleotides. Targeted DNA-sequencing methods have proven useful for validating variants from whole-genome sequencing and studying disease-relevant gene subsets.

RNA-Based Method

- **RNA sequencing (RNA-seq):** cDNAs from an RNA sample are used to prepare sequencing libraries. In this method, cellular poly-A-tailed RNA is isolated, fragmented, and randomly primed for reverse transcription to generate double-stranded cDNA fragments. Sequencer adaptors are ligated to the cDNA fragments, which are size-selected by gel electrophoresis. A limited-cycle PCR step ensures minimal contamination of RNA. cDNA libraries are sequenced using standard NGS platforms (62). RNA-seq provides information on the diversity and abundance of expressed sequences, thereby elaborating the content and, simultaneously, levels of gene expression of the sample.

a single species based on shared mtDNA haplotypes (38). Similarly, a study involving 700 SNPs (39) uncovered the distinctiveness of a chimpanzee subspecies (*Pan troglodytes ellioti*) (**Figure 1**) when population structure was conflicting using microsatellites (40, 41). Genomic data enable high-resolution analyses, which can reveal a subtle inter- and intraspecific population structure that has far-reaching conservation implications.

It is important to note, however, that population genetic analyses developed for standard genetic techniques may not be directly transferable to analysis of genomic data. For example, SNP

Copy number variant (CNV): a class of DNA sequence variant (including deletions and duplications) in which the result is a departure from the expected diploid representation of the DNA sequence

Hybridization: interbreeding of individuals considered to be different species

Microsatellite: a section of DNA consisting of short nucleotide sequences repeated many times, which vary between individuals

Silent sites: a term used to describe genetic variation that does not result in a change in amino acids (i.e. synonymous sites, introns, and intergenic sites)

Admixture: the production of new genetic combinations in hybrid populations through recombination

Effective population size (N_e): the number of individuals in an idealized population that would show the same amount of genetic drift or inbreeding as the population being studied

Recombination: process during which maternal and paternal loci are mixed into new combinations in the gametes during meiosis

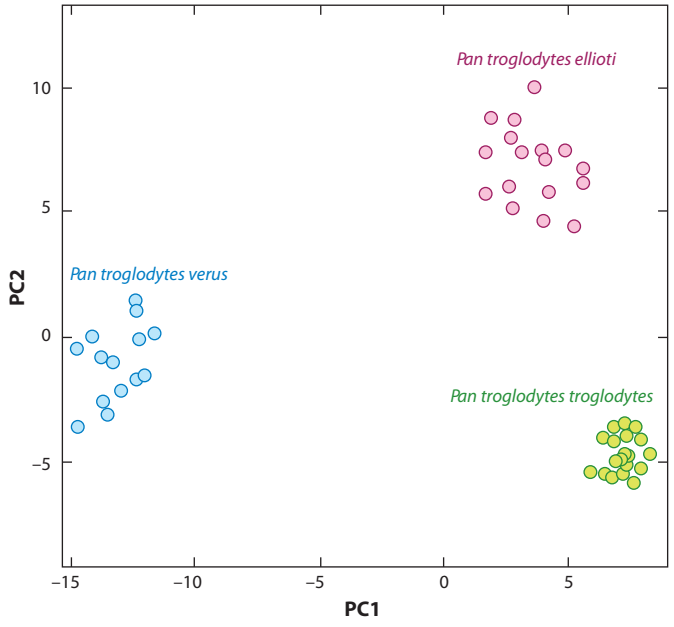

Figure 1

Clustering of chimpanzees based on principal components (PCs) using data from 818 SNPs. Plots of the first two PCs show that chimpanzees in this study form three genetically distinct groups corresponding to different subspecies. In particular, *Pan troglodytes ellioti* is identified as a genetically distinct cluster. This study illustrates the application of high-throughput genomic methods in identifying conservation units. Modified with permission from Bowden et al. (39).

calling must account for the high error rate associated with NGS, and not all commonly used analysis software can support large genomic data sets. New methods developed specifically for NGS data include improving the accuracy of genotype calls through the use of linkage information for genome-enabled species (42), as well as promising new model-based methods that use linkage to uncover subtle population structure quickly and accurately (43). Additionally, flexible models that use likelihood simplifications or allele frequency–spectrum summaries of genomic data will need to be extended for use in wild species (44, 45).

ADAPTIVE GENETIC VARIATION

Unlike neutral genetic variation, adaptive variation is shaped by selective forces. The ability to identify adaptive loci is one of the most exciting uses of genomic approaches in evolutionary studies and conservation (2). Access to genome-wide data and annotated genomes in wild species facilitates identification of the genetic changes that accompany local adaptations and the manner in which these changes influence fitness. This knowledge will assist in defining conservation units in the wild (46, 47) and may help evaluate the potential of populations to respond to changing environments (48). Understanding the geographic distribution of loci influencing local adaptation will also help to assess habitat requirements for population persistence and the ecological exchangeability of divergent populations (49).

In conservation genomics, the use of transcriptomics has gained interest as a valuable starting point for characterizing functional genetic variation in wild organisms. This is especially true in species with large and complex genomes for which complete genome sequencing is unnecessary,

and relatively costly using current sequencing approaches, and for which assembly poses special difficulties (50). An increasing number of transcriptome studies have recently been published (51–55), including studies on species under conservation management (47, 56). Functional SNPs and genome-wide differences in gene-expression levels between populations can be detected from expressed sequence tag (EST) libraries by using the RNA-seq technique (57, 58). This technique has been employed successfully in rainbow trout for identifying SNP markers associated with growth rate (59). Additionally, genome-wide expression data, in conjunction with analysis of neutral variation (microsatellites), have identified conservation units and loci associated with local adaptation of distinct populations of Atlantic salmon (47).

The use of transcriptomics in species of conservation concern does pose challenges given contemporary methodology. Computationally, RNA-seq estimates of expression levels require normalization for transcript length, which has to be inferred from evolutionarily distant species with available genomic sequences (54). From a management perspective, conservation strategies that rely on analyses of a limited number of putatively adaptive loci are problematic. In the wild, managing for a few genes that are under positive selection may fail to consider standing genetic diversity that can be adaptive in the future (2). Additionally, in intensively managed populations, favoring the breeding of some individuals over others based on their genotype will lead to unequal genetic contributions, loss of genetic diversity, and increased relatedness among individuals within the population (60).

Methods to Detect Adaptive Variation

Historically, quantitative genetic methods have been effective in detecting adaptive variation in life-history traits important for species fitness and population viability (61). QTL mapping is an approach frequently used to identify genetic regions associated with phenotypes in populations with available pedigree information. In wild species, QTL mapping has been applied, for example, to the zebra finch (62), deer mouse (63), and cave tetra fish (64) to identify genetic regions involved in adaptive traits.

Although QTL mapping studies are feasible in any organism for which adequate phenotypic or life-history data are available, a handful of wild populations have been mapped using SNP data (reviewed in 65). Cost-effective SNP discovery using techniques such as RAD-seq has facilitated mapping of polymorphisms associated with adaptive traits in three-spined sticklebacks (66) and *Peromyscus leucopus* (67). The RAD-seq methodology was also used to identify the common genetic basis of development-rate variation in two genetically distinct rainbow trout populations (68) (**Figure 2**). These studies, facilitated by NGS, can detect genomic regions and even nucleotide sequence variants that function adaptively in the natal environment of each population, enabling population-specific recovery plans for rainbow trout and other species.

QTL mapping cannot be applied to species in which crosses are infeasible or from which pedigree data are not available. Under these circumstances, GWAS may be an effective approach (69); for instance, it was implemented in wild Soay bighorn sheep, in which genetic variation across 36,000 SNPs was screened to identify a major QTL involved in horn morphology (70). Genetic mapping is limited by the number of individuals that must be sampled (71) to observe a significant signal and by the relatively low map resolution of QTL analyses, which makes it difficult to move from QTL detection to identification of a locus associated with adaptation and fitness (72).

Other methods used to detect adaptive variation are based on genome scans that identify signatures of selection according to levels of genetic polymorphism within populations and levels of divergence among populations. An increase or decrease in population differentiation is often measured with the F_{ST} estimator (73). Regions under positive selection may be identified

Linkage disequilibrium (LD): the non-random association of alleles present at two or more loci, which can cause a genetic correlation

Site frequency spectrum (SFS): the distribution of allele frequencies at a single nucleotide site or across multiple sites

Wright's fixation index (F_{ST}): a measure of population differentiation that is due to genetic structure; it measures the fraction of the total genetic variation that is distributed among subpopulations relative to the total population

Transcriptomics: the sequencing and quantification of expressed genes in specific tissues/cells

Expressed sequence tag (EST): a short DNA fragment (hundreds of base pairs) produced by reverse transcription of mRNA into DNA

Sliding window averaging: the averaging of nucleotide diversity or F_{ST} values within a fixed genome length; when the window is moved across the genome, this method smoothes out variation within regions so that genome-wide patterns can be observed

Selective sweeps: the process by which a new, advantageous mutation eliminates or reduces variation in linked neutral sites as it increases in frequency in the population

through a genomic sliding windows averaging approach because their F_{ST} values lie significantly outside the differentiation observed over the neutral genomic background (2). This method has been used to dissect the genetic basis of adaptive phenotypes and to detect ecological speciation in lake whitefish (74) and three-spined sticklebacks (75). Selective sweeps can also be detected by scanning for regions of reduced genetic variation or an excess of rare alleles, as performed in deer mice (76) and wolves (69) (**Figure 3**) to identify mutations involved in cryptic color variation.

Evaluation of the extent of LD in genomic sliding windows also provides insight into the influence of selection at specific genomic regions among populations. For instance, an extended LD block, combined with a high-frequency haplotype, is expected around a region under positive selection (73). However, large LD blocks might also reflect demographic events associated with inbred or admixed populations (37), for which interpreting genetic signatures of selection may be difficult owing to the confounding influences of complex demographic events as well as varying rates of mutation and recombination (77).

Adaptive Loci for Conservation

The major histocompatibility complex (MHC) region is a major fitness-related locus with a role that extends beyond host defense and pathogen resistance. MHC is involved in important life-history traits such as mate choice, intraspecific territoriality, and kin recognition. Recent advances in sequencing technologies indicate that obtaining measures of MHC genetic diversity is within reach of conservation initiatives (78). Genomic approaches in wildlife species have facilitated the characterization of MHC loci (79, 80) by revealing their complexity in structure and number of alleles as well as the presence of strong balancing selection (81, 82), gene duplication, and neofunctionalization, all of which produce evolutionary novelties for ecological adaptation (83).

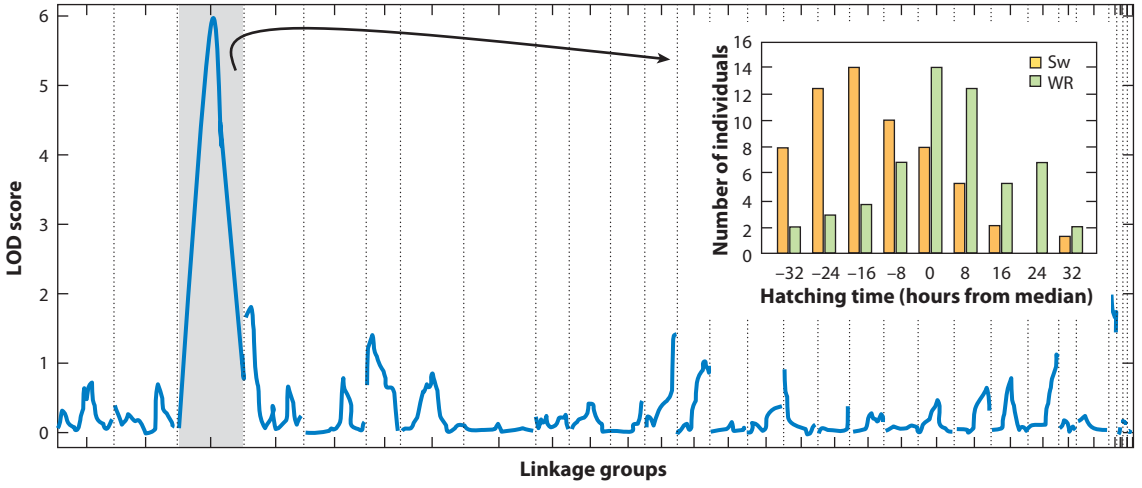

Figure 2

Application of restriction-site-associated DNA sequencing (RAD-seq) to examine the genetic architecture of development rate variation of the rainbow trout (*Oncorhynchus mykiss*). Development rate of the *O. mykiss* clonal lines varies in a manner reflecting the population's natural variation, with one line having an accelerated development rate [referred to as Swanson (Sw)] compared with the other, slower-developing line [referred to as Whale Rock (WR)]. QTL analyses of development rate [logarithm of odds (LOD) profiles] for each of the 29 linkage groups in this study show that the major QTL peak associated with rapid development is located in linkage group 3. The inset histogram represents the development rate phenotypic profile for individuals that inherit either the Sw or WR allele at the major QTL peak. Modified with permission from Miller et al. (68).

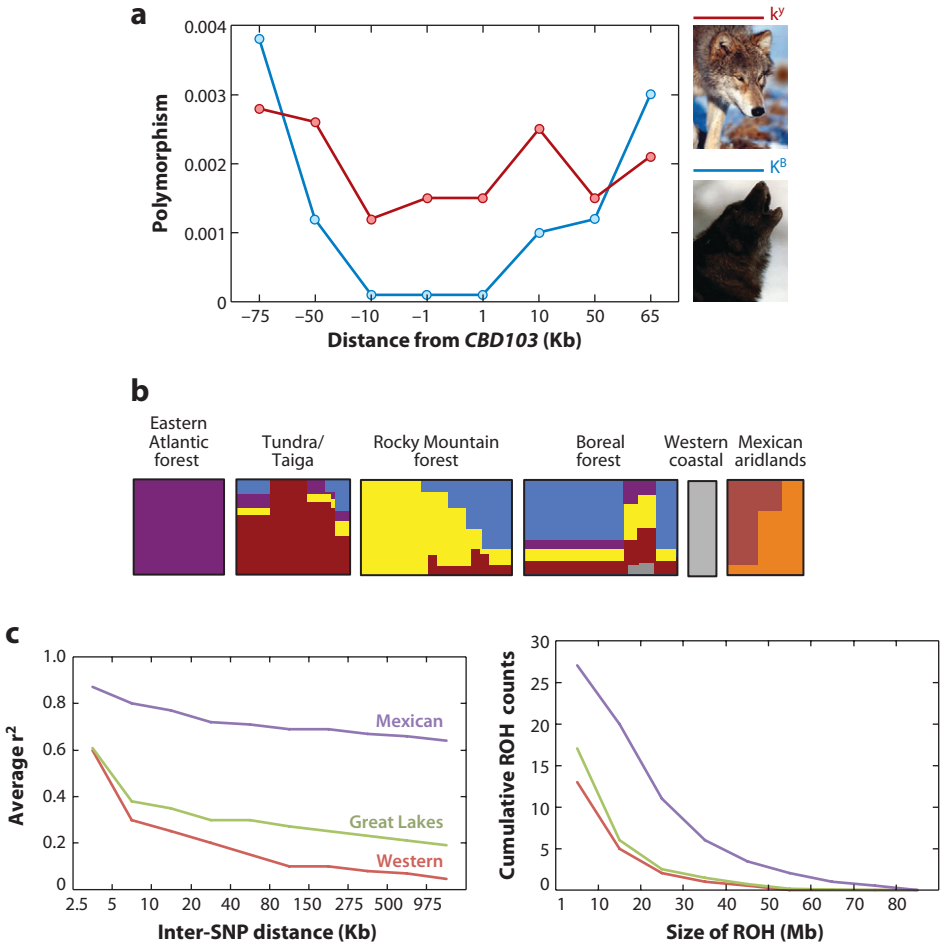

Figure 3

Examples of NGS applications in conservation using the North American wolf as an example. (*a*) Polymorphism around the *CBD103* gene, encoded by the K locus, a component of the melanocortin pathway in North American gray wolves (69). Little genetic variation, combined with extended haplotype homozygosity, demonstrates positive selection around the K locus in the black-colored wolves. (*b*) Genetic clustering analysis using 44,000 SNPs shows population structure and allele composition (represented by different colors) among North American wolves, evidencing high levels of admixture among wolf populations. (*c*) Extent of linkage disequilibrium (LD; genetic association, r^2) and runs of homozygosity. The comparison of LD extent (*left*) and the autozygosity frequency spectrum (*right*) were used to examine inbreeding (Mexican wolf) and admixture (Great Lakes wolf) in comparison with outbred North American wolves (western wolf). Modified with permission from Anderson et al. (69) and vonHoldt et al. (37). Abbreviations: ROH, runs of homozygosity; SNP, single-nucleotide polymorphism.

Levels of polymorphism in MHC loci can provide indirect measures of the immunological fitness of populations as identified in the endangered San Nicolas island fox (81). Captive-breeding programs may also benefit from the characterization of MHC variation for pathogen resistance by monitoring the maintenance of MHC variants, as recommended in the Arabian oryx breeding program (84). Population genomic studies of MHC evolution are a promising area in

Inbreeding coefficient: the probability that at a given locus, both alleles are descended from a common ancestor (i.e., identical-by-descent)

conservation genomics and in the search for genetic variation contributing to fitness. The deeper exploration of MHC function, however, requires more precise assembly of these complex regions. In most wild species, details on order and CNV within the MHC region remain unclear, and MHC organization is known to differ greatly across vertebrates (85). As advances are made in identifying MHC organization and variability in wild organisms, explicit conservation management applications will become increasingly feasible.

FITNESS DECLINES AND INBREEDING

Habitat destruction and fragmentation lead to smaller and ever more isolated populations. This makes populations more susceptible not only to exogenous stochastic events but also to loss of genetic diversity through genetic drift and inbreeding. Inbreeding can result in inbreeding depression, a decrease in evolutionary adaptive potential, which thereby increases extinction risk. Although inbreeding depression is seen across taxa (86), the basis for its stochasticity and the underlying mechanisms that produce inbreeding depression remain poorly understood. Genomics may help to shed light on the genetic architecture of inbreeding depression and to establish a causative link between the phenotypes and the underlying molecular processes. In particular, it may help test hypotheses regarding the number of loci that contribute to inbreeding as well as the underlying genetic mechanisms, such as dominance, overdominance, epistasis, and/or genotype-environment interactions.

Loci contributing to inbreeding depression may be identified by sequencing the whole genomes of parents and offspring (87) or by examining gene-expression profiles (88). Recently, GWAS have been used to successfully identify recessive mutations that cause inherited defects in livestock and dogs (89, 90), in an attempt to characterize genes associated with inbreeding. Another study on inbred Scandinavian wolves, using approximately 250 microsatellite loci, evaluated the effects of inbreeding and selective forces at relatively low resolution and provided evidence of increased LD compared to outbred populations (91). This study set the stage for more detailed analyses using SNP interrogation methods that can facilitate the identification of the loci involved in inbreeding depression. Ultimately, the identification of loci that are associated with inbreeding depression, and their comparison across species, will open new areas of investigation and allow the establishment of new approaches for managing the risks of inbreeding depression in small-population management.

NGS and genomic analyses may also help to better estimate relatedness and inbreeding coefficients among individuals in wild populations. To date, the vast majority of studies measuring inbreeding in wild animal populations that lack pedigree information have been performed by using neutral genetic markers, such as microsatellites. However, analytic tools for relatedness estimates that use microsatellites or small numbers of nuclear markers have very high sampling variances (92, 93), which reduces accuracy. High-density genomic data obtained through NGS, such as SNPs that can provide estimates of inbreeding coefficients, have the potential to decrease this large variance, and new methods are being developed with improved results (94). For example, Li et al. (95) compared methods for estimating individual inbreeding coefficients and pairwise relatedness based on genome-wide SNPs and genealogies separately. They obtained consistent values, in particular for highly inbred animals and pairs of closely related individuals, and concluded that genomic data provide useful information in cases of complex or absent pedigrees. In managed populations, genomic analysis can also assist in estimating the relatedness of the wild founders that make up captive populations. Founders of captive populations are typically assumed to be unrelated to one another, which may be an unrealistic assumption for some species.

HYBRIDIZATION AND INTROGRESSION

Gaining deeper insights into the roles of hybridization and introgressive gene flow in natural populations is critical for understanding how new species are formed and how the genetic distinctiveness and reproductive isolation of species is maintained. Because introgressive gene flow may increase or decrease fitness, a greater ability to detect the occurrence and timing of gene flow between species is relevant to population sustainability and management. Introgression is increasingly being identified in wildlife populations, including African and Asian elephants (96), American bison (97), Darwin's finches (98), and California tiger salamanders (99). Studies of ancient DNA from fossil material have even revealed introgressive gene flow in the recent evolutionary history of humans (100), which broadens recognition of the potential role of hybridization across species boundaries as a common evolutionary process that facilitates adaptation.

Studies incorporating genomics to investigate genetic introgression in wildlife populations are only now beginning to emerge. Undoubtedly, the generation of high-density genomic data will significantly improve our understanding of the genetics of introgression, particularly with regard to long-standing questions about the patterns of introgression in speciation and radiation, detection and timing of ancient hybridization, and understanding of how introgressed loci behave in their new genomic background (reviewed in 101). A recent example of introgression involves polar and brown bears. Deep genome sequencing, in combination with comparative genomic analyses, has identified the timing of divergence of these sister taxa and has brought forth evidence for occasional hybridization through the past 4–5 million years of climatic variation (102). In this study, admixture maps (**Figure 4**) were used to identify the introgressed genomic regions in each bear species. The authors identify one such region that may have also undergone a selective sweep in both polar bears and coastal-living brown bears. This region contains the homolog of *ALDH7A1*, a gene that has been shown in humans to be related to salt tolerance, a potentially adaptive trait for polar and coastal brown bears living in marine environments.

Introgressive gene flow: gene flow from one species into the gene pool of another through repeated interspecific hybridization

Chondrodystrophy: a genetic skeletal disorder that affects the development of cartilage

DISEASE SUSCEPTIBILITY AND RESISTANCE

The application of genome sequencing to advance understanding of disease processes and the genetic basis for risk and resistance has been a hallmark of the Human Genome Project (103). Whole-genome sequencing and genome-wide SNP studies can provide knowledge for management and treatment of diseases in wildlife species (104). For instance, in the Atlantic salmon, comparative genomic studies using analysis of synteny and partial genome sequences have identified a genetic region for resistance to a devastating disease, salmon infectious anemia (105) (**Figure 5**). One gene in particular, *HIV-EP2/MBP-2*, that is implicated in the response to infectious salmon anemia virus is a very strong candidate for resistance.

In wildlife and endangered species, as in humans and domestic animals, diagnostic information, phenotypic data, and epidemiological studies will have to be used in concert with data on genetic variation to understand the genetic basis for disease risk. The limited medical oversight of wild populations limits the amount of phenotypic data related to disease status and risk factors. However, notable examples exist. For example, in California condors a lethal disorder, chondrodystrophy, was identified in the breeding population, with inheritance consistent with an autosomal recessive transmission mode (106). The frequency of mutation can be modeled, but the identification of actual carriers of the trait—in both the reintroduced populations and the captive population—is still based on production of affected chicks (107). The immediate goal for conservation management is to assess carrier status by using genomic technologies to identify a linked

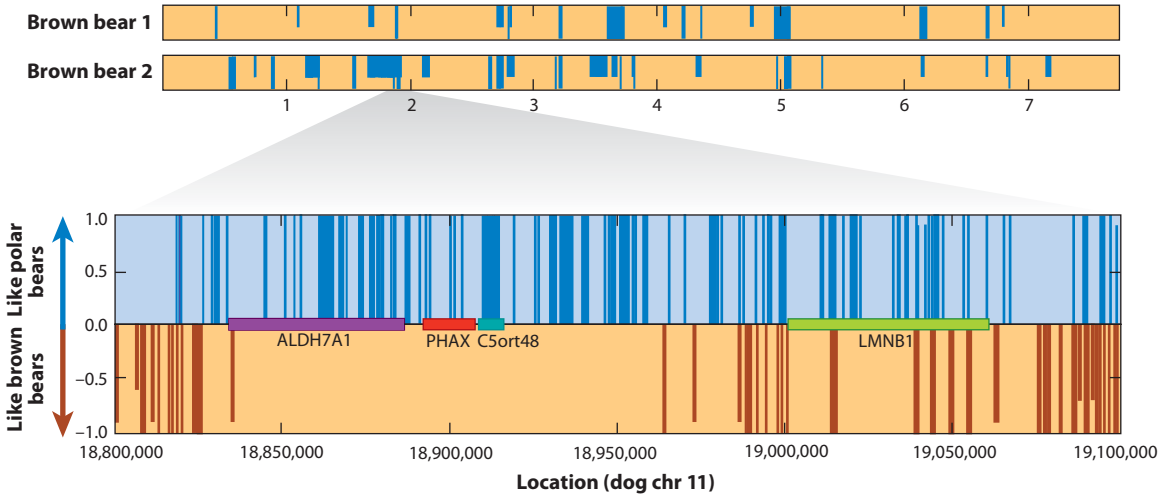

Figure 4

Admixture maps detecting genetic introgression from the polar bear genome into brown bears. The genomic region corresponds to dog chromosome 11 in two brown bears, scaled in units of 10 million bases. Brown areas denote chromosomal regions shared among brown bears, and blue indicates where one or both chromosomes are shared with polar bears. A magnification of that includes a 250-kb interval in which both chromosomes in brown bears are very similar to those of polar bears. The region contains four genes, including the ortholog of ALDH7A1, which may be related to salt tolerance. Modified with permission from Miller et al. (102).

marker to the recessive mutation. A linkage map based on multiple markers, including SNPs, will assist population managers in monitoring and managing the frequency of the deleterious allele both in the captive population, in a manner that assures the conservation of the species gene pool, and in the reintroduced, free-ranging population that forms the basis for species recovery (108).

Worldwide amphibian biodiversity is under a severe threat associated with disease risk to chytrid fungus, *Batrachochytrium dendrobatidis*. Understanding the genetic basis for susceptibility and resistance to the chytrid pathogen is crucial for recovery of species that have not already become extinct. Experimental infection of individuals collected from multiple wild populations of the lowland leopard frog (*Lithobates yavapaiensis*) has shown that, in isolated populations, significant differences in survival of chytrid fungal infection are correlated with MHC heterozygosity and polymorphism (104). MHC alleles that may confer resistance to this pathogen were detected in some but not all populations, a finding that encourages consideration of population structure and diversity in resilience to disease challenges in the wild.

LIMITATIONS OF GENOMICS IN CONSERVATION

Sampling and Application of In Vitro Models for Conservation Genomics

A major impediment to the advancement of genomic studies that involve threatened species is the availability of samples. Disturbance of the population to be sampled and potential impacts on individual threatened animals pose reasonable limitations to sample collection. However, animals that are handled for other purposes, for example, individuals fitted with radio collars in wild populations and those given health evaluations in captive populations, provide appropriate opportunities to obtain samples for genetic analyses that may provide crucial data for population management. To take advantage of genetic sampling opportunities, suggested protocols for collection, shipment, and curation of samples have been established (109–112).

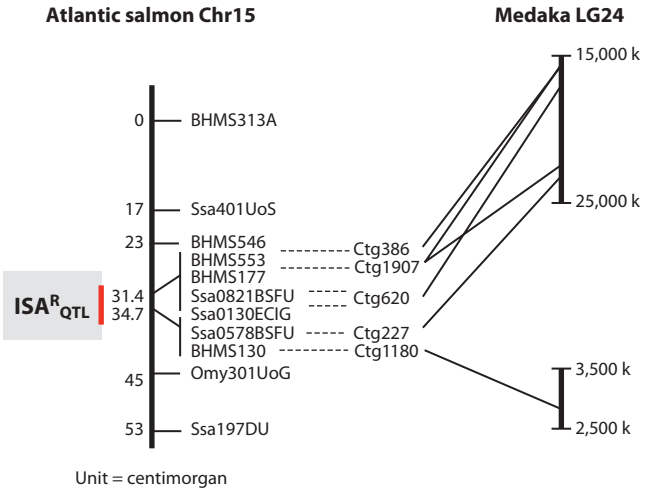

Figure 5

Schematic representation of the genetic map of Atlantic salmon chromosome 15 [linkage group (LG) 8] showing the infectious salmon anemia quantitative trait loci (QTL) region and the corresponding syntenic regions of medaka LG24. One gene in particular within the QTL regions, *HIV-EP2/MBP-2*, is one of the stronger candidates for association with the infectious salmon anemia virus (ISAV) resistance given that it may influence the expression of several other genes that have been implicated in the response to infection by ISAV. Modified with permission from Li et al. (105).

The development of banks that establish and maintain primary cell cultures from biopsies, either freshly collected or processed in the field for later culturing (113), has provided the only available access to high-quality DNA and RNA specimens from numerous endangered species. Experimental studies using cell lines represent a promising in vitro model that can facilitate genomic studies of wild species. Generation of EST libraries from cell cultures provides resources for improving the process of genome annotation, discovering SNPs in coding sequences, and identifying functional genes potentially associated with ecologically relevant traits and adaptation across populations. For example, gene-expression profiles from cultured fibroblast cells were used to identify patterns of species specificity in cellular metabolism in humans and great apes (114), which shows the applicability of in vitro models to evolutionary and conservation-related issues.

Analytical and Practical Limitations

Despite the rapid evolution of genomic technologies, some significant limitations still remain. As mentioned throughout this review, the production of genomic data has become faster and easier, whereas data-analysis techniques frequently lag. In addition, many statistical programs for population genetics need to be adapted to large data sets and require significant advances in bioinformatics and computational biology. Once genomic data are obtained and analyzed, conservation scientists face yet another challenge of a new scale. Application of genetic data may result in defining units of conservation too narrowly, may impede conservation actions, and may stand in the way of endangered species management (115). Conservation scientists and managers face the opposing risks to population survival from depression of fitness owing to the effects of outbreeding and inbreeding. Development of a predictive science that may be applied at the species level to manage genetic risks in populations is not close at hand. However, the empirical data for initiating such approaches is likely to rely heavily on the application of genomic technologies.

The application of genomic approaches in conservation research and management involves consideration of their cost-effectiveness and feasibility. In many cases, traditional genetic markers, such as microsatellites, that already have been developed will be the most economical and efficient solution to a particular conservation genetics question (116). This may be the case for conservation programs in more remote locations, where application of genomic technologies is not currently feasible. Expanding the application of genomic studies to a greater number and broader diversity of species is a pivotal issue for the characterization and conservation of biodiversity and for evaluating genetic aspects of population viability. However, the appropriate application of genomic technologies in the conservation context should be driven by the requisite resolving power and not by the novelty of the techniques (reviewed in 3).

SUMMARY POINTS

1. NGS projects will rapidly expand the number of threatened species for which assembled genomes and detailed information on sequence variation are available. These data will advance investigations relevant to the conservation of biological diversity.
2. The ability to identify adaptive genetic variation will improve the definition of units of conservation management and the evaluation of the potential of populations to respond to changing environments.
3. Knowledge of the genetic mechanisms of inbreeding depression and the loci that contribute to reduction in fitness stands to be advanced through the application of NGS and gene-expression studies.
4. Factors involved in population viability analysis, including kinship evaluation, pedigree reconstruction, migration, introgression, and admixture, will all be facilitated through interrogation of genome-wide variation of SNP markers.
5. The development of biomaterial banks and protocols to collect, maintain, and curate biological samples is necessary to advance genomic studies.
6. Although NGS and genomic analyses stand to fundamentally alter methods of biological inquiry, current obstacles in de novo genome assembly challenge the advancement of comparative genomics and its fruitful application to biodiversity conservation.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review

ACKNOWLEDGMENT

We thank an anonymous reviewer for comments on an early version of this manuscript and Christina Jordan for assistance obtaining permission for use of figures.

LITERATURE CITED

1. Ekblom R, Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107:1–15

2. Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. *Nat. Rev. Genet.* 11:697–709
3. Avise J. 2010. Perspective: conservation genetics enters the genomics era. *Conserv. Genet.* 11:665–69
4. Frankham R. 2010. Challenges and opportunities of genetic approaches to biological conservation. *Biol. Conserv.* 143:1919–27
5. Frankham R. 2010. Where are we in conservation genetics and where do we need to go? *Conserv. Genet.* 11:661–63
6. Hudson ME. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Resour.* 8:3–17
7. Kohn MH, Murphy WJ, Ostrander EA, Wayne RK. 2006. Genomics and conservation genetics. *Trends Ecol. Evol.* 21:629–37
8. Ouborg NJ, Angeloni F, Vergeer P. 2009. An essay on the necessity and feasibility of conservation genomics. *Conserv. Genet.* 11:643–53
9. Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW. 2010. Conservation genetics in transition to conservation genomics. *Trends Genet.* 26:177–87
10. Primmer CR. 2009. From conservation genetics to conservation genomics. *Ann. N. Y. Acad. Sci.* 1162:357–68
11. Schwartz MK, Luikart G, Waples RS. 2007. Genetic monitoring as a promising tool for conservation and management. *Trends Ecol. Evol.* 22:25–33
12. Shokralla S, Spall JL, Gibson JF, Hajibabaei M. 2012. Next-generation sequencing technologies for environmental DNA research. *Mol. Ecol.* 21:1794–805
13. Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, et al. 2010. Adaptation genomics: the next generation. *Trends Ecol. Evol.* 25:705–12
14. Windig J, Engelsma K. 2010. Perspectives of genomics for genetic conservation of livestock. *Conserv. Genet.* 11:635–41
15. Locke DP, Hillier LW, Warren WC, Worley KC, Nazareth LV, et al. 2011. Comparative and demographic analysis of orang-utan genomes. *Nature* 469:529–33
16. Excoffier L, Hofer T, Foll M. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity* 103:285–98
17. Jensen P. 2005. Genomics: the chicken genome sequence. *Heredity* 94:567–68
18. Nei M, Suzuki Y, Nozawa M. 2010. The neutral theory of molecular evolution in the genomic era. *Annu. Rev. Genomics Hum. Genet.* 11:265–89
19. Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–95
20. Lohmueller KE, Bustamante CD, Clark AG. 2011. Detecting directional selection in the presence of recent admixture in African-Americans. *Genetics* 187:823–35
21. Verdu P, Rosenberg NA. 2011. A general mechanistic model for admixture histories of hybrid populations. *Genetics* 189:1413–26
22. Nielsen R, Hubisz MJ, Hellmann I, Torgerson D, Andres AM, et al. 2009. Darwinian and demographic forces affecting human protein coding genes. *Genome Res.* 19:838–49
23. Thornton KR, Jensen JD. 2007. Controlling the false-positive rate in multilocus genome scans for selection. *Genetics* 175:737–50
24. Li J, Li H, Jakobsson M, Li S, Sjödin P, Lascoux M. 2012. Joint analysis of demography and selection in population genetics: Where do we stand and where could we go? *Mol. Ecol.* 21:28–44
25. Tenesa A, Navarro P, Hayes BJ, Duffy DL, Clarke GM, et al. 2007. Recent human effective population size estimated from linkage disequilibrium. *Genome Res.* 17:520–26
26. Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, Eversole KA, et al. 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324:528–32
27. Boyko AR, Williamson SH, Indap AR, Degenhardt JD, Hernandez RD, et al. 2008. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* 4:e1000083
28. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5:e1000695

29. Gray MM, Granka JM, Bustamante CD, Sutter NB, Boyko AR, et al. 2009. Linkage disequilibrium and demographic history of wild and domestic canids. *Genetics* 181:1493–505
30. Engelhardt BE, Stephens M. 2010. Analysis of population structure: a unifying framework and novel methods based on sparse factor analysis. *PLoS Genet.* 6:e1001117
31. Kirk H, Freeland JR. 2011. Applications and implications of neutral versus non-neutral markers in molecular ecology. *Int. J. Mol. Sci.* 12:3966–88
32. Pool JE, Hellmann I, Jensen JD, Nielsen R. 2010. Population genetic inference from genomic sequence variation. *Genome Res.* 20:291–300
33. Frankham R. 2005. Genetics and extinction. *Biol. Conserv.* 126:131–40
34. Reed DH, Frankham R. 2003. Correlation between fitness and genetic diversity. *Conserv. Biol.* 17:230–37
35. Vila C, Sundqvist AK, Flagstad O, Seddon J, Bjornerfeldt S, et al. 2003. Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proc. Biol. Sci.* 270:91–97
36. Gompert Z, Forister ML, Fordyce JA, Nice CC, Williamson RJ, Buerkle CA. 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of *Lycaeides* butterflies. *Mol. Ecol.* 19:2455–73
37. vonHoldt BM, Pollinger JP, Earl DA, Knowles JC, Boyko AR, et al. 2011. A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Res.* 21:1294–305
38. Rohland N, Reich D, Mallick S, Meyer M, Green RE, et al. 2010. Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. *PLoS Biol.* 8:e1000564
39. Bowden R, MacFie TS, Myers S, Hellenthal G, Nerrienet E, et al. 2012. Genomic tools for evolution and conservation in the chimpanzee: *Pan troglodytes ellioti* is a genetically distinct population. *PLoS Genet.* 8:e1002504
40. Becquet C, Patterson N, Stone AC, Przeworski M, Reich D. 2007. Genetic structure of chimpanzee populations. *PLoS Genet.* 3:e66
41. Gonder MK, Locatelli S, Ghobrial L, Mitchell MW, Kujawski JT, et al. 2011. Evidence from Cameroon reveals differences in the genetic structure and histories of chimpanzee populations. *Proc. Natl. Acad. Sci. USA* 108:4766–71
42. Nielsen R, Paul JS, Albrechtsen A, Song YS. 2011. Genotype and SNP calling from next-generation sequencing data. *Nat. Rev. Genet.* 12:443–51
43. Lawson DJ, Hellenthal G, Myers S, Falush D. 2012. Inference of population structure using dense haplotype data. *PLoS Genet.* 8:e1002453
44. Kim SY, Lohmueller KE, Albrechtsen A, Li Y, Korneliussen T, et al. 2011. Estimation of allele frequency and association mapping using next-generation sequencing data. *BMC Bioinformatics* 12:231
45. Pritchard JK. 2011. Whole-genome sequencing data offer insights into human demography. *Nat. Genet.* 43:923–25
46. Manel S, Joost S, Epperson BK, Holderegger R, Storfer A, et al. 2010. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Mol. Ecol.* 19:3760–72
47. Vandersteen Tymchuk W, O'Reilly P, Bittman J, MacDonald D, Schulte P. 2010. Conservation genomics of Atlantic salmon: variation in gene expression between and within regions of the Bay of Fundy. *Mol. Ecol.* 19:1842–59
48. Hoffmann AA, Sgró CM. 2011. Climate change and evolutionary adaptation. *Nature* 470:479–85
49. Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK. 2000. Considering evolutionary processes in conservation biology. *Trends Ecol. Evol.* 15:290–95
50. Parchman T, Geist K, Grahnen J, Benkman C, Buerkle CA. 2010. Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics* 11:180
51. Elmer KR, Kusche H, Lehtonen TK, Meyer A. 2010. Local variation and parallel evolution: morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes. *Philos. Trans. R. Soc. B Biol. Sci.* 365:1763–82
52. Ekblom R, Balakrishnan CN, Burke T, Slate J. 2010. Digital gene expression analysis of the zebra finch genome. *BMC Genomics* 11:219

53. Kunstner A, Wolf JBW, Backstrom N, Whitney O, Balakrishnan CN, et al. 2010. Comparative genomics based on massive parallel transcriptome sequencing reveals patterns of substitution and selection across 10 bird species. *Mol. Ecol.* 19:266–76
54. Tarazona S, Garcia-Alcade F, Dopazo J, Ferrer A, Conesa A. 2011. Differential expression in RNA-seq: a matter of depth. *Genome Res.* 21:2213–23
55. Tzika AC, Helaers R, Schramm G, Milinkovitch MC. 2011. Reptilian-transcriptome v1.0, a glimpse in the brain transcriptome of five divergent Sauropsida lineages and the phylogenetic position of turtles. *EvoDevo* 2:19
56. Milano I, Babbucci M, Panitz F, Ogden R, Nielsen RO, et al. 2011. Novel tools for conservation genomics: comparing two high-throughput approaches for SNP discovery in the transcriptome of the European hake. *PLoS ONE* 6:e28008
57. Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10:57–63
58. Gilad Y, Pritchard JK, Thornton K. 2009. Characterizing natural variation using next-generation sequencing technologies. *Trends Genet.* 25:463–71
59. Salem M, Vallejo RL, Leeds TD, Palti Y, Liu S, et al. 2012. RNA-seq identifies SNP markers for growth traits in rainbow trout. *PLoS ONE* 7:e36264
60. Ballou JD, Lacy RC. 1995. Identifying genetically important individuals for management of genetic diversity in pedigreed populations. In *Population Management for Survival and Recovery*, ed. JD Ballou, M Gilpin, TJ Foose, pp. 76–111. New York: Columbia Univ. Press
61. Stinchcombe JR, Hoekstra HE. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100:158–70
62. Schielzeth H, Forstmeier W, Kempenaers B, Ellegren H. 2012. QTL linkage mapping of wing length in zebra finch using genome-wide single nucleotide polymorphisms markers. *Mol. Ecol.* 21:329–39
63. Steiner CC, Weber JN, Hoekstra HE. 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol.* 5:1880–89
64. Protas M, Tabansky I, Conrad M, Gross JB, Vidal O, et al. 2008. Multi-trait evolution in a cave fish, *Astyanax mexicanus*. *Evol. Dev.* 10:196–209
65. Garvin MR, Saitoh K, Gharrett AJ. 2010. Application of single nucleotide polymorphisms to non-model species: a technical review. *Mol. Ecol. Resour.* 10:915–34
66. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376
67. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7:e37135
68. Miller M, Brunelli J, Wheeler P, Liu S, Rexroad CE, et al. 2012. A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Mol. Ecol.* 21:237–49
69. Anderson TM, vonHoldt BM, Candille SI, Musiani M, Greco C, et al. 2009. Molecular and evolutionary history of melanism in North American gray wolves. *Science* 323:1339–43
70. Johnson HE, Mills LS, Wehausen JD, Stephenson TR, Luikart G. 2011. Translating effects of inbreeding depression on component vital rates to overall population growth in endangered bighorn sheep. *Conserv. Biol.* 25:1240–49
71. Slate J, Gratten J, Beraldi D, Stapley J, Hale M, Pemberton J. 2009. Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica* 136:97–107
72. Barrett RDH, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* 12:767–80
73. Oleksyk TK, Smith MW, O'Brien SJ. 2010. Genome-wide scans for footprints of natural selection. *Philos. Trans. R. Soc. B Biol. Sci.* 365:185–205
74. Renaut S, Bernatchez L. 2011. Transcriptome-wide signature of hybrid breakdown associated with intrinsic reproductive isolation in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Heredity* 106:1003–11
75. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6:e1000862

76. Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE. 2009. On the origin and spread of an adaptive allele in deer mice. *Science* 325:1095–98
77. Nielsen R. 2005. Molecular signatures of natural selection. *Annu. Rev. Genet.* 39:197–218
78. Ujvari B, Belov K. 2011. Major Histocompatibility Complex (MHC) markers in conservation biology. *Int. J. Mol. Sci.* 12:5168–86
79. Siddle HV, Deakin JE, Coggill P, Whilming LG, Harrow J, et al. 2011. The tammar wallaby major histocompatibility complex shows evidence of past genomic instability. *BMC Genomics* 12:421
80. Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrom M, et al. 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477:207–10
81. Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc. Natl. Acad. Sci. USA* 101:3490–94
82. Xu T-j, Sun Y-n, Chen S-l. 2010. Allelic variation, balancing selection and positive selected sites detected from MHC class I alpha gene of olive flounder. *Genetica* 138:1251–59
83. Burri R, Salamin N, Studer RA, Roulin A, Fumagalli L. 2010. Adaptive divergence of ancient gene duplicates in the avian MHC class II β. *Mol. Biol. Evol.* 27:2360–74
84. Hedrick PW, Parker KM, Gutiérrez-Espeleta GA, Rattink A, Lievers K. 2000. Major histocompatibility complex variation in the Arabian oryx. *Evolution* 54:2145–51
85. Kelley J, Walter L, Trowsdale J. 2005. Comparative genomics of major histocompatibility complexes. *Immunogenetics* 56:683–95
86. Keller LF, Waller DM. 2002. Inbreeding effects in wild populations. *Trends Ecol. Evol.* 17:230–41
87. Charlesworth D, Willis JH. 2009. The genetics of inbreeding depression. *Nat. Rev. Genet.* 10:783–96
88. Paige KN. 2010. The functional genomics of inbreeding depression: a new approach to an old problem. *Bioscience* 60:267–77
89. Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, et al. 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nat. Genet.* 40:449–54
90. Karlsson EK, Baranowska I, Wade CM, Salmon Hillbertz NH, Zody MC, et al. 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat. Genet.* 39:1321–28
91. Hagenblad J, Olsson M, Parker HG, Ostrander EA, Ellegren H. 2009. Population genomics of the inbred Scandinavian wolf. *Mol. Ecol.* 18:1341–51
92. Glaubitz JC, Rhodes OE, Dewoody JA. 2003. Prospects for inferring pairwise relationships with single nucleotide polymorphisms. *Mol. Ecol.* 12:1039–47
93. Lynch M, Ritland K. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* 152: 1753–66
94. Gasbarra D, Pirinen M, Sillanpaa MJ, Arjas E. 2007. Estimating genealogies from linked marker data: a Bayesian approach. *BMC Bioinformatics* 8:411
95. Li M-H, Stranden I, Tiirikka T, Sevon-Aimonen M-L, Kantanen J. 2011. A comparison of approaches to estimate the inbreeding coefficient and pairwise relatedness using genomic and pedigree data in a sheep population. *PLoS ONE* 6:e26256
96. Roca AL, Georgiadis N, O'Brien SJ. 2005. Cytonuclear genomic dissociation in African elephant species. *Nat. Genet.* 37:96–100
97. Halbert ND, Derr JN. 2007. A comprehensive evaluation of cattle introgression into US federal bison herds. *J. Hered.* 98:1–12
98. Grant BR, Grant PR. 2010. Conspecific versus heterospecific gene exchange between populations of Darwin's finches. *Philos. Trans. R. Soc. B Biol. Sci.* 365:1065–76
99. Fitzpatrick BM, Johnson JR, Kump DK, Shaffer HB, Smith JJ, Voss SR. 2009. Rapid fixation of non-native alleles revealed by genome-wide SNP analysis of hybrid tiger salamanders. *BMC Evol. Biol.* 9:176
100. Abi-Rached L, Jobin MJ, Kulkarni S, McWhinnie A, Dalva K, et al. 2011. The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science* 333:89–94

101. Twyford AD, Ennos RA. 2012. Next-generation hybridization and introgression. *Heredity* 108:179–89
102. Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-Reina OC, et al. 2012. Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proc. Natl. Acad. Sci. USA* 109:14295–96
103. Collins FS. 1999. Shattuck lecture: medical and societal consequences of the Human Genome Project. *N. Engl. J. Med.* 341:28–37
104. Savage AE, Zamudio KR. 2011. MHC genotypes associate with resistance to a frog-killing fungus. *Proc. Natl. Acad. Sci. USA* 108:16705–10
105. Li J, Boroevich KA, Koop BF, Davidson WS. 2011. Comparative genomics identifies candidate genes for Infectious Salmon Anemia (ISA) resistance in Atlantic salmon (*Salmo salar*). *Mar. Biotechnol.*13:232–41
106. Ralls K, Ballou JD, Rideout BA, Frankham R. 2000. Genetic management of chondrodystrophy in California condors. *Anim. Conserv.* 3:145–53
107. Romanov MN, Tuttle EM, Houck ML, Modi WS, Chemnick LG, et al. 2009. The value of avian genomics to the conservation of wildlife. *BMC Genomics* 10(Suppl. 2):S10
108. Romanov MN, Koriabine M, Nefedov M, de Jong PJ, Ryder OA. 2006. Construction of a California condor BAC library and first-generation chicken-condor comparative physical map as an endangered species conservation genomics resource. *Genomics* 88:711–18
109. Wong PBY, Wiley EO, Johnson WE, Ryder OA, O'Brien SJ, et al. 2012. Tissue sampling and standards for vertebrate genomics. *GigaScience* 1:8
110. Ryder OA, Benirschke K. 1984. Value of frozen tissue collections for zoological parks. In *Collections of Viable Tissues: Value, Management, and Future*, ed. HC Dessauer, MS Hafner, pp. 6–9. Lawrence, KS: Am. Assoc. Syst. Collect.
111. Ryder OA, McLaren A, Brenner S, Zhang Y-P, Benirschke K. 2000. DNA banks for endangered species. *Science* 288:275–77
112. Wildt DE, Rall WF, Critser JK, Monfort SL, Seal US. 1997. Genome resource banks. *Bioscience* 47: 689–98
113. Tovar H, Navarrete F, Rodríguez L, Skewes O, Castro FO. 2008. Cold storage of biopsies from wild endangered native Chilean species in field conditions and subsequent isolation of primary culture cell lines. *In Vitro Cell. Dev. Biol. Anim.* 44:309–20
114. Karaman M, Houck M, Chemnick L, Nagpal S, Chawannakul D, et al. 2003. Comparative analysis of gene-expression patterns in human and African great ape cultured fibroblasts. *Genome Res.* 13:1619–30
115. Frankham R, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, et al. 2012. Implications of different species concepts for conserving biodiversity. *Biol. Conserv.* 153:25–31
116. Kristensen TN, Pedersen KS, Vermeulen CJ, Loeschcke V. 2010. Research on inbreeding in the 'omic' era. *Trends Ecol. Evol.* 25:44–52
117. Davey J, Hohenlohe P, Etter P, Boone J, Catchen J, Blaxter M. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12:499–510
118. Myllykangas S, Natsoulis G, Bell JM, Ji HP. 2011. Targeted sequencing library preparation by genomic DNA circularization. *BMC Biotechnol.* 11:122
119. Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307:1928–33
120. Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10:565–77
121. Weetman D, Wilding CS, Steen K, Morgan JC, Simard F, Donnelly MJ. 2010. Association mapping of insecticide resistance in wild *Anopheles gambiae* populations: major variants identified in a low-linkage disequilibrium genome. *PLoS ONE* 5:e13140
122. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, et al. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.* 9:356–69
123. Seddon N, Amos W, Adcock G, Johnson P, Kraaijeveld K, et al. 2005. Mating system, philopatry and patterns of kinship in the cooperatively breeding subdesert mesite *Monias benschi*. *Mol. Ecol.* 14:3573–83

124. Saillant E, Patton JC, Gold JR. 2005. Genetic variation, kinship, and effective population size in a captive population of the endangered Cape Fear shiner, *Notropis mekistocholas*. *Copeia* 2005:20–28
125. Blouin MS. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends Ecol. Evol.* 18:503–11
126. Ivy JA, Lacy RC. 2012. A comparison of strategies for selecting breeding pairs to maximize genetic diversity retention in managed populations. *J. Hered.* 103:186–96
127. Araki H, Cooper B, Blouin MS. 2007. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science* 318:100–3
128. Pemberton JM. 2008. Wild pedigrees: the way forward. *Proc. R. Soc. B Biol. Sci.* 275:613–21
129. Kirkpatrick B, Li SC, Karp RM, Halperin E. 2011. Pedigree reconstruction using identity by descent. *J. Comput. Biol.* 18:1481–93
130. Nunes VL, Beaumont MA, Butlin RK, Paulo OS. 2011. Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (*Lacerta lepida*) along an environmental gradient. *Mol. Ecol.* 20:193–205
131. Perry GH, Melsted P, Marioni JC, Wang Y, Bainer R, et al. 2012. Comparative RNA sequencing reveals substantial genetic variation in endangered primates. *Genome Res.* 22:602–10
132. Ellegren H, Sheldon BC. 2008. Genetic basis of fitness differences in natural populations. *Nature* 452:169–75
133. Pertoldi C, Wojcik JM, Tokarska M, Kawalko A, Kristensen TN, et al. 2010. Genome variability in European and American bison detected using the BovineSNP50 BeadChip. *Conserv. Genet.* 11:627–34
134. Wetzel DP, Stewart IRK, Westneat DF. 2012. Heterozygosity predicts clutch and egg size but not plasticity in a house sparrow population with no evidence of inbreeding. *Mol. Ecol.* 21:406–20
135. David P. 1998. Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity* 80:531–37
136. Slate J, Pemberton JM. 2002. Comparing molecular measures for detecting inbreeding depression. *J. Evol. Biol.* 15:20–31
137. Miller JM, Poissant J, Hogg JT, Coltman DW. 2012. Genomic consequences of genetic rescue in an insular population of bighorn sheep (*Ovis canadensis*). *Mol. Ecol.* 21:1583–96
138. Johnson WE, Onorato DP, Roelke ME, Land ED, Cunningham M, et al. 2010. Genetic restoration of the Florida panther. *Science* 329:1641–45
139. Halbert ND, Ward TJ, Schnabel RD, Taylor JF, Derr JN. 2005. Conservation genomics: disequilibrium mapping of domestic cattle chromosomal segments in North American bison populations. *Mol. Ecol.* 14:2343–62
140. Wolf JBW, Bayer T, Haubold B, Schilhabel M, Rosenstiel P, Tautz D. 2010. Nucleotide divergence vs. gene expression differentiation: comparative transcriptome sequencing in natural isolates from the carrion crow and its hybrid zone with the hooded crow. *Mol. Ecol.* 19:162–75
141. Roca AL, Schook LB. 2011. Genomics: captive breeding and wildlife conservation. In *Encyclopedia of Biotechnology in Agriculture and Food*, ed. DR Heldman, DG Hoover, MB Wheeler, pp. 320–25. New York: Taylor & Francis
142. Decker JE, Pires JC, Conant GC, McKay SD, Heaton MP, et al. 2009. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proc. Natl. Acad. Sci. USA* 106:18644–49
143. Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, et al. 2011. A molecular phylogeny of living primates. *PLoS Genet.* 7:e1001342
144. Delsuc F, Brinkmann H, Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6:361–75
145. Steele PR, Pires JC. 2011. Biodiversity assessment: state-of-the-art techniques in phylogenomics and species identification. *Am. J. Bot.* 98:415–25
146. Brinkmann H, Philippe H. 2008. Animal phylogeny and large-scale sequencing: progress and pitfalls. *J. Syst. Evol.* 46:274–86
147. Singh ND, Larracuente AM, Sackton TB, Clark AG. 2009. Comparative genomics on the *Drosophila* phylogenetic tree. *Annu. Rev. Ecol. Evol. Syst.* 40:459–80

148. Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, et al. 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proc. Natl. Acad. Sci. USA* 107:16196–200
149. McCormack JE, Maley JM, Hird SM, Derryberry EP, Graves GR, Brumfield RT. 2012. Next-generation sequencing reveals phylogeographic structure and a species tree for recent bird divergences. *Mol. Phylogenet. Evol.* 62:397–406
150. Knowles LL. 2009. Statistical phylogeography. *Annu. Rev. Ecol. Evol. Syst.* 40:593–612
151. Ventura M, Catacchio CR, Alkan C, Marques-Bonet T, Sajjadian S, et al. 2011. Gorilla genome structural variation reveals evolutionary parallelisms with chimpanzee. *Genome Res.* 21:1640–49
152. Scally A, Dutheil JY, Hillier LW, Jordan GE, Goodhead I, et al. 2012. Insights into hominid evolution from the gorilla genome sequence. *Nature* 483:169–75
153. Ureta-Vidal A, Ettwiller L, Birney E. 2003. Comparative genomics: genome-wide analysis in metazoan eukaryotes. *Nat. Rev. Genet.* 4:251–62
154. Miller W, Makova KD, Nekrutenko A, Hardison RC. 2004. Comparative genomics. *Annu. Rev. Genomics Hum. Genet.* 5:15–56