7 Applications of Genomic Methods to Studies of Wild Primate Populations

Mary A. Kelaita

7.1 Introduction: Molecular Tools for Field Primatology

Nonhuman primates have been instrumental subjects of research intended to investigate the genetic basis of human health and diseases, given that many of their similarities with humans were inherited from a recently shared ancestor. Recent genomic technological advances have facilitated biomedical research conducted on captive nonhuman primates. Therefore, genomic efforts have focused on key nonhuman primate taxa considered to be the model species for biomedical research, such as macaques (Gibbs et al. 2007) and baboons (Rogers et al. 2009). For example, the use of the baboon as a model to determine how genetic variation influences a complex disease is discussed in this volume (e.g., Chap. 16; Comuzzie's chapter). In addition, special attention has also been given to apes owing to their status as our closest relatives (Stone and Verrelli 2006), which contributes to the understanding of what makes humans unique.

This chapter, however, is dedicated to studies of wild primate populations. Field studies of wild primates can be equally important, offering answers to ecological and evolutionary questions that cannot be addressed with data from captive populations. For example, the extent of genotypic and phenotypic variation both within and among populations can often only be observed in the wild. In addition, an evolutionary framework often requires the relevant ecological and environmental factors that shape primate lineage diversity (Tung et al. 2010). Wild population studies allow researchers to determine the effects of different environmental factors on an individual's phenotype. Some phenotypic variation may only be observed in the presence of specific genotype-by-environment interactions, and could suggest the need for the investigation of gene regulating mechanisms in that developmental pathway (Tung et al. 2011). It is also possible to test hypotheses about how some genotypes influence survival and reproduction (and therefore fitness) in wild populations given that they are under natural selective pressures (Bradley and Lawler 2011).

Apart from providing an ecological and evolutionary context, the diversity of wild primate populations suggests that phylogenetic comparisons within the primate order (including model and nonmodel species and their subpopulations) can shed light on when and how unique human adaptations evolved and what processes resulted in the observed current human-wide genomic variation. Phylogenetic relationships can be more accurately ascertained when samples are obtained with special considerations for the geographic distribution of and the extent of variation within wild populations (Luikart et al.

M.A. Kelaita (✉)
Department of Anthropology, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA
e-mail: mary.kelaita@utsa.edu

R. Duggirala et al. (eds.), *Genome Mapping and Genomics in Human and Non-Human Primates*,
Genome Mapping and Genomics in Animals 5, DOI 10.1007/978-3-662-46306-2_7,
© Springer-Verlag Berlin Heidelberg 2015

2003; Thalmann et al. 2007). In these cases, genomic data can inform debates regarding the taxonomic placement of certain populations. Primate phylogenomics is a field that is already well underway in taking advantage of genomic tools (Moulin et al. 2008; Siepel 2009; Ting and Sterner 2013).

Bradley and Lawler (2011) present a comprehensive review on how field primatologists can take advantage of genomic tools to uncover genetic variation underlying primate adaptations, including candidate gene approaches, genome-wide association studies, and expression analyses. Additionally, Ting and Sterner (2013) recently reviewed the status of primate molecular phylogenetics after the introduction of genomic tools. This chapter will build on this existing body of knowledge by providing a brief background on genomic methods but will focus on other tools and applications of relevance to field primatologists that center on detecting variation in wild populations. Population genetics methods can inform studies of primate conservation (Chaves et al. 2011), hybridization (Kelaita and Cortés-Ortiz 2013), behavior and social organization (Di Fiore 2003), and demographic history (Lawler 2011). DNA can now be obtained through noninvansive sampling methods, which are preferred by many primatologists knowing the potential harm that could result during capture. Many primates cannot be habituated to human presence, leaving DNA as the only viable method for identifying individuals and measuring real and effective population sizes (Vigilant 2009). To that effect, molecular biology has already revolutionized the field of primatology, providing tools such as gel electrophoresis, restriction enzyme mapping, polymerase chain reaction (PCR), and finally DNA sequencing (Charlesworth 2010; Di Fiore 2003).

Thus far, the majority of the methods utilized by field primatologists have for the most part relied on inferences made from a few loci discovered through inefficient methods (Raveendran et al. 2006). This could result in inaccurate measures of variation, inability to discern relationships in parentage analysis, or unreliable estimates of divergence, given that various parts of the genome may have been under different evolutionary pressures. The most significant promise of the genomic revolution is the potential to acquire massive amounts of genetic data. Now, with the ability to study thousands to millions of genetic markers, field primatologists will be able to answer questions that they have been unable to with only a limited number of loci. Indeed, the decreasing costs of new technologies and the discovery of novel methods have generated a great deal of interest in determining how genomics can benefit wildlife biology and ecology studies (Thomas and Klaper 2004; Ryder 2005; Primmer 2009; Allendorf et al. 2010; Avise 2010; Ouborg et al. 2010; Steiner et al. 2013). While primates have been at the forefront of genomic sequencing efforts relative to other organisms (Tung et al. 2010), wild primate studies have been slow to incorporate many of the methods reviewed in the ecological genomics literature.

Neutral markers are used to generate estimates of parameters such as effective population size (Ne) and migration rate (m) (Allendorf et al. 2010) as well as nucleotide diversity and recombination rates (Steiner et al. 2013); therefore, the inclusion of a large number of markers from across the entire genome is necessary for accurate parameter estimation. For example, sequencing of the Sumatran and Bornean orangutan genomes revealed a much larger effective population size and greater genetic diversity in the Sumatran species and a divergence time that is more recent than those proposed by previous studies (Locke et al. 2011). Larger data sets enable researchers to test for outlier loci before estimating population parameters, thereby testing assumptions of neutrality (Luikart et al. 2003). Larger data sets also have the potential to uncover historical events such as population bottlenecks and expansions (Ryder 2005), especially given the “mosaic” nature of the genome, different regions of which may have undergone recombination and been subject to different selective pressures (Degnan and Rosenberg 2009). A greater number of markers would reveal linked loci and can improve haplotype inference in order to detect the extent and directionality of migration (Allendorf et al.

2010). With whole genome data, comparisons of the entire genome can be made across taxa which can shed light on the processes generating diversity in primate lineages (Hudson 2008).

7.2 Making the Transition from Genetics to Genomics

Primatologists who plan on taking advantage of the genomic revolution may find it difficult to make the transition, considering that overall few eukaryote species have received attention for genomic resource development (Hudson et al. 2008). When the first genome of a species is assembled and published, it serves as a reference map for assembling genomes of other individuals from the same species (Baker 2012). In addition, it can be scanned for the identification of polymorphic markers, as has been done for rhesus macaques (Raveendran et al. 2006). Many nonmodel primate species lack a fully assembled reference genome. Obtaining a fully sequenced genome in the absence of a reference genome requires a great investment in time and resources for *de novo* genome assembly. This is the case even despite recent advances in assembling genomes on a massively parallel scale (Wheeler 2008). Primatologists interested in using genomic tools currently have two options: either work with model organisms that already have significant genomic resources available or use the resources available from a closely related species for which a reference genome exists and apply them to a species of interest (Thomas and Klaper 2004).

Recently, after the sequencing of the first complete human genome, efforts have been in full force to sequence whole genomes of nonhuman primates, beginning with some species identified as sequencing targets for various reasons. Some were assigned the highest priority, owing to their taxonomic placement as index species in the primate phylogeny, their use in biomedical research (Marques-Bonet et al. 2009), or their conservation status (Ryder 2005). Currently, there are 32 ongoing primate genome projects (reviewed in Bradley and Lawler 2011, and listed on http://www.genome.gov/10002154). Field primatologists can begin to take advantage of published data by accessing a number of available online databases with built-in alignment search tools. Some researchers are conducting partial genome sequencing projects in an effort to provide more sequence data resources for nonmodel primate species for which no whole genome sequencing is currently planned. For example, Jameson et al. (2012) developed and annotated sequence reads from three platyrrhine species from genomic shotgun libraries of 3,000 individual sequences. These data can provide a resource for marker discovery in other related New World taxa.

Once a genome project is completed, the assembled and annotated genomes can be used as reference sequences in what is termed “massively parallel” or “next generation sequencing (NGS) technology”, allowing for millions of simultaneous reads in each run. For some nonmodel species, an assembled genome of a closely related species can serve as a scaffold. These “genome-enabled” species studies can benefit from many of the currently available resources (Thompson et al. 2010), but must factor in genome assembly errors that result from low coverage and actual variation between the two species (Bradley and Lawler 2011). There remains a number of nonhuman primate species which have been ecologically well characterized but have not received much attention in sequencing projects (e.g., howler monkeys), possibly due to the perceived lack of their research’s direct implications for understanding human health and evolution as well as their conservation status. Given the predicted reductions in costs and effort needed to assemble new genomes, this may change in the near future. Until then, primatologists can take steps toward making the transition from the genetic to the genomic era.

The first step for many primatologists is recognizing the different types of newly developed genomic technologies. This can be daunting given the accelerated rate at which new technologies are being introduced and utilized. The traditional Sanger technology provided sequence

data of up to 2 kilobases through the detection of labeled nucleotides as they are incorporated during DNA synthesis (Zhang et al. 2011). Given the sequence length limitations, "shotgun" sequencing was introduced, so-called because DNA was sheared and inserted into cloning vectors, which were randomly fragmented and sequenced to produce short reads. Whole genomes were originally acquired in this manner, through the assembly of these reads into larger fragments, thereby generating sequence data for the entire genome of the individual. The challenge with assembling the first genome for any species is therefore the correct spatial mapping of reads in the absence of a fully mapped genome that can serve as a comparative reference. This is by no means a simple task; the assembly of a draft genome requires considerable bioinformatics know-how and computing resources. The task is further complicated by the presence of structural variation in the genome, including gene duplication (Davey et al. 2011).

NGS technology similarly accomplishes sequencing of the entire genome through the random fragmentation of DNA followed by their sequencing. The use of cloning is eliminated, and sequences are instead bound to adapters (Zhang et al. 2011). However, NGS technology actually comprises several types including Roche 454 pyrosequencing Illumina sequencing by synthesis, ABI SOLiD sequencing by ligation, and Helicos tSMS single-molecule sequencing, whose advantages and disadvantages have been compared (Hudson 2008; Ekblom and Galindo 2011). These technologies have a number of different applications which will be discussed below, but all come with their own set of challenges (Pool et al. 2010). When a reference genome is available, sequencing other individuals of the same species to uncover variation in the population is referred to as "resequencing" (Bentley 2008). This is most preferable given that complete genomic information for each individual is obtained, including coding and noncoding regions, allowing for inferences to be made about the evolutionary pressures that shaped genomes of extant species and uncovering sequence as well as structural variation. For some nonhuman primate species, *de novo* whole genome assembly remains impractical considering the amount of time, funding, expertise, and infrastructure necessary. An additional challenge is that the NGS instruments' data analysis software is usually designed to assemble and annotate human, rat, and mouse sequences. Working with other species requires further development of sequence assembly and annotation pipelines even when a fully assembled reference genome is available. Finally, analyzing a large number of individuals is essential for addressing population genetics questions, but obtaining whole genome sequences for each individual in a sample remains an unfeasible and costly endeavor.

A useful tool for nonmodel species research is expressed sequence tags (ESTs), which are short sequences produced by translating mRNA transcripts into complementary DNA, and represents only protein coding regions (Rudd 2003). ESTs are relatively inexpensive to produce and have been used extensively by molecular ecologists (Bouck and Vision 2007). Therefore, an alternative to genomics involves an analysis of the transcriptome, the mRNA obtained from different tissues at different life stages (Vera et al. 2008). Assembly of a species' transcriptome can be more feasible than that of the genome, given that it only involves mapping of coding sequences. This approach is often recommended for ecologists who plan to begin genomics projects for species that lack a reference genome (Cahais et al. 2012). Transcriptome characterization can be carried out on model organisms with available reference genomes or EST data, but can also involve *de novo* assembly (Cahais et al. 2012; Vera et al. 2008). In fact, Perry et al. (2012) developed a method for *de novo* transcriptome assembly and assembled thousands of sequences for 16 mammalian species, including 11 primate species. Interestingly, RNA comparisons revealed that endangered lemur populations exhibit considerable genetic variation, likely since factors that have impacted lemur populations occurred too recently to be reflected in observed genetic diversity measures. Such comparisons can now be made by accessing publicly available data. For example, Pipes et al.

(2012) developed a nonhuman primate reference transcriptome resource (http://nhprtr.org) presently hosting RNA sequence data for 13 primate species.

Random-primed cDNA libraries can be created and used to analyze nucleotide variation or they can provide information on whether and to what degree genes are expressed. In addition to the potential for massive, parallel investigations of gene expression, NGS can be used to produce the actual mRNA sequences for later assembly (Hudson 2008). Once a transcriptome is assembled, it can be used as a template for further resequencing or the development of markers and constructions of microarrays for expression profiling (Ekblom and Galindo 2011). The transcriptome, therefore, can be a viable method for generating genetic markers for wild population studies.

NGS technologies can be used to generate large amounts of sequence data even without assembling them into a full genome, and these data can be further interrogated for marker discovery. Also, given the difficulty in obtaining whole genome data for many individuals, there are a number of methods utilizing NGS technologies that sample some of the overall variation present in a population, sometimes referred to as genome complexity reduction (GCR) methods (Davey et al. 2011; Dou et al. 2012). For example, a number of known loci can be targeted through the selective capture of DNA prior to sequencing but high coverage sequencing of these regions provides intraspecific variation information that can be useful for population genetics analyses (Ekblom and Galindo 2011). Bi et al. (2012) performed an exon capture in chipmunks relying on a low-coverage draft genome of the ground squirrel that is 30 mya divergent from the chipmunk. They developed transcripts from different tissues and identified ~12,000 exons for capture from these transcripts. Unfortunately, this approach is limited to functional regions, although "exon-primed intron-crossing" (EPIC) markers were developed which can also span intron regions. EPIC markers have the unique property of being variable but also generally conserved across a broad range of species (Thompson et al. 2010). Finally, targeted sequencing of variable parts of the genome can be used as a barcoding approach as well (Ekblom and Galindo 2011), a method that can be of use for identifying plant and bacterial species from fecal samples.

Yet another GCR method ideal for population genetics analyses is called restriction site-associated DNA sequencing (RADSeq, Davey and Blaxter 2010). After genomic DNA is sheared with restriction enzymes, adapters with unique molecular identifiers for each individual are ligated to the fragments, allowing them to bind to the Ilumina flow cell. These fragments are then pooled, randomly sheared, and ligated to a second adapter with a divergent end that can only be amplified upon the amplification of the first adapter containing the molecular identifier. The resulting library is sequenced, generating sequence data of the adapters and the DNA flanking the restriction site, where polymorphisms can be found (Davey and Blaxter 2010). A similar method involves RNA sequencing (RNASeq) where cDNA libraries are used instead of genomic DNA (Wang et al. 2009).

Single nucleotide polymorphisms (SNPs) are especially suited for measuring genetic diversity, a large number of which can be discovered through resequencing (Hudson 2008). SNPs can be utilized as neutral markers for measuring genetic diversity but can also occur in coding or regulatory regions. SNPs can be employed in genome-wide association studies in pedigreed populations which are designed to discover statistically significant correlations between particular regions of the genome and the phenotype in question (Slate et al. 2009). The most feasible high-throughput method for SNP discovery is likely to be through transcriptome sequencing and resequencing (Hudson 2008) or through capture of sequences using EPIC markers, so that SNPs can be identified in a number of species related to the focal organism even without existing sequence data (Slate et al. 2009). Central to many population genetics analyses are measures of linkage disequilibrium (LD), which provides information about historical and demographic events, and can be determined from SNP

data through the construction of linkage maps, which incidentally also aid in locating genes under selection (Thompson et al. 2011).

Recently, Bergey et al. (2013) applied the RADSeq technique to five primate species, including humans, representing major lineages within the primate order. They were able to detect a large number of SNPs that can be compared across closely related species at a relatively low cost. Therefore, the method can be adopted to search for SNPs that exhibit intraspecific variation, but also SNPs that can be used in phylogenetic analyses of relatively shallow trees. The RADSeq method requires high-quality DNA, preferably obtained from tissue or blood samples. However, there are promising methods for extracting DNA from fecal samples for genomic analyses (Perry et al. 2010), and together these studies show real promise for the ability of primatologists to work with large-scale genomic data when resources are scarce.

It is important to note that while using a subset of the genome through GCR methods for marker discovery is more feasible, whole genome sequences could still be more advantageous for demographic analyses given the presence of rare variants and could provide a more complete picture of allele frequencies (Pool et al. 2010).

7.3 Further Applications for Wild Primate Populations

7.3.1 Pedigree Reconstruction

To date, a large number of wild primate population studies lack pedigree information. Long-term studies of wild primate populations tracking several generations are rare. Knowing relatedness among individuals is important for identifying quantitative trait loci and measuring heritability (Pemberton 2008), as well as for measuring reproductive skew and for studying kin-directed behaviors (Di Fiore 2009). Many wild population studies have relied on microsatellite markers, which are highly variable, to infer relationships among individuals (Di Fiore 2009). However, the power to accurately determine pedigree relationships not only depends on how polymorphic a marker is but also the number of markers employed (Blouin 2003). SNP markers, while having lower power than microsatellites for resolving relationships, can be identified using high-throughput methods, providing ample numbers of markers for parentage analysis, and are less prone to genotyping errors (Hauser et al. 2011). For example, large numbers of SNPs have helped to determine relatedness among individuals in a zebra fish population (Santure et al. 2010). SNPs can potentially provide power for determining different categories of kinship beyond those of parent–offspring pairs or full sibs (Avise 2010). Microsatellites have so far remained the marker of choice for wild primate relatedness inference but with the availability of SNP discovery methods, primatologists can begin to construct accurate and specific relationships in natural populations.

7.3.2 Metagenomics

The field of metagenomics has allowed comparisons of microbial ecosystems across primate taxa, encompassing gastrointestinal and vaginal microbiomes. Microbial ecosystems reflect different species' phylogenetic history, dietary quality and availability, and even health outcomes in response to their respective environments (Amato et al. 2013). Gut microbes are thought to influence the evolution of their host, given their role in metabolizing certain nutritional components. Metagenomics studies have thus far provided evidence that microbial community composition is often not only species-specific but can also reflect habitat differences. Given that gut bacteria are largely parentally inherited, gut microbiota evolutionary history should coincide with that of their hosts (Ochman et al. 2010). Yildirim et al. (2010) utilized pyrosequencing technology of the small subunit rRNA (a region of the 16S rRNA gene) of different nonhuman primate species. They found greater similarity in microbial community composition within species than between species, and that gastrointenstinal microbiomes are highly

associated with their host taxa. Overall, gut microbiota among great ape species was found to be phylogentically conserved (Ochman et al. 2010). However, a number of factors including ecological differences among the hosts' environments may shape gut microbial composition. The role of habitat differences (and further, dietary differences) was further confirmed by Amato and colleagues (2013), who assessed microbial community composition from howler monkey fecal samples by sequencing the same region of the rRNA gene. They found habitat specific microbial taxa composition, diversity, and richness, which is predicted by habitat type and shaped by the availability of plants in the diet.

7.3.3 Hybridization

Hybridization in primates has been garnering a great deal of attention recently as molecular tools have made it possible to detect more instances of gene flow across established taxonomically distinct primate taxa (Cortés et al. 2007). Debate regarding the importance of the role of hybridization in primate evolution continues (Zinner et al. 2011), and is receiving renewed interest given the finding that a number of genes have introgressed from Neanderthals into modern humans (Green et al. 2010). So far, researchers have been able to detect hybrid primate individuals using relatively few diagnostic microsatellite loci (Cortés-Ortiz et al. 2007; Tung et al. 2008; Kelaita and Cortés-Ortiz 2013). Yet, initial identification of these loci and subsequent testing is time consuming and cumbersome. Not only must loci successfully amplify and be highly variable, they must also possess fixed allelic differences between the parental species. SNPs, which instead can be identified with high-throughput methods, can also serve as diagnostic loci in hybridization studies (Finger et al. 2009; Hohenlohe et al. 2011).

Further, while few microsatellite loci can aid in the detection of hybrids, understanding the dynamics of gene flow and introgression across the hybrid zone is important for determining mechanisms of reproductive isolation and barriers to gene flow. Such an endeavor requires the use of a much greater number of loci (Allendorf et al. 2010). Teeter et al. (2009) discovered selection against hybrid genotypes and for some introgressed genotypes in a mouse hybrid zone using 41 SNPs. Whole genome data could potentially address the role of the number of loci and the size of their effects, dominance, epistasis, or chromosomal rearrangements in causing outbreeding depression. Hybridization has been shown to produce highly variable morphological characteristics in nonhuman primates (Ackermann et al. 2010; Kelaita and Cortés-Ortiz 2013) and it remains unclear what genetic interactions are the cause of this variability. Genomic approaches could also produce more accurate estimates of each hybrid's proportion of admixture (Allendorf et al. 2010; Steiner et al. 2013). With this information, morphology, behavior, and fitness can be compared across individuals of varying genomic background. Finally, genomic data promises to uncover past hybridization events that could have led to the formation of new species and the emergence of novel adaptations (Keller et al. 2012).

7.4 Concluding Remarks

It is likely that the number of genome-enabled nonhuman primate species will increase in the near future. This chapter has outlined a number of approaches that are feasible for wild population studies, some of which are relatively inexpensive and require little effort. These methods enable making evolutionary and functional inferences for a broader range of species, including nonmodel primate species that have generally received less attention in genomic resource development. However, field primatologists are likely to still face a number of obstacles to fully engaging in this type of research. A consistent concern in wild primate population studies is access to high-quality DNA, which is harder to obtain from noninvasive sampling methods. In addition, as Tung et al. (2010) recommend, considerable statistical and programming skill is required to undertake genome-scale

analyses. Successful genomic endeavors often involve collaborations with researchers who have access to the infrastructure (both laboratory and computing) necessary or who possess expertise in these areas, but building on these skills as more resources become available is necessary given that technological discoveries are enabling investigators to conduct genomic studies with the budget and equipment of a small laboratory. Finally, while primatologists may be eager to acquire massive amounts of genetic data for a seemingly unlimited potential to answer important evolutionary and ecological questions, a well-designed project can help identify the minimum number of loci necessary for the analysis, the ideal sequencing technologies with the least amount of error produced, and the most time- and cost-efficient approaches for achieving one's goals.

References

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697–709

Amato KR, Yeoman CJ, Kent A, Righini N, Carbonero F, Estrada A, Gaskins HR, Stumpf RM, Yildirim S, Torralba M, Gillis M, Wilson BA, Nelson KE, White BA, Leigh SR (2013) Habitat degradation impacts black howler monkey (*Alouatta pigra*) gastrointestinal microbiomes. Int Soc Microb Ecol 7:1344–1353

Avise J (2010) Perspective: conservation genetics enters the genomics era. Conserv Genet 11:665–669

Baker M (2012) *De novo* genome assembly: what every biologist should know. Nat Methods 9:333–337

Bentley DR (2008) Whole-genome re-sequencing. Curr Opin Genet Dev 16:545–552

Bergey CM, Pozzi L, Disotell TR, Burrell AS (2013) A new method for genome-wide marker development and genotyping holds great promise for molecular primatology. Int J Primatol 34:303–314

Bi K, Vanderpool D, Singhal S, Linderoth T, Moritz C, Good JM (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genom 13:403–417

Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends Ecol Evol 18:503–511

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. Mol Ecol 16:907–924

Bradley BJ, Lawler RR (2011) Linking genotypes, phenotypes, and fitness in wild primate populations. Evol Anthropol 20:104–119

Cahais V, Gayral P, Tsagkogeorga G, Melo-Ferreira J, Ballenghien M, Weinert L, Chiari Y, Belkhir K, Ranwez V, Galtier N (2012) Reference-free transcriptome assembly in non-model animals from next-generation sequencing data. Mol Eco Res 12:834–845

Charlesworth B (2010) Molecular population genomics: a short history. Genet Res 92:397–411

Chaves PB, Alvarenga CS, Possamai CB, Dias LG, Boubli JP, Strier KB, Mendes SL, Fagundes V (2011) Genetic diversity and population history of a critically endangered primate, the Northern muriqui (*Brachyteles hypoxanthus*). PLoS ONE 6:e20722

Cortés-Ortiz L, Duda TF, Canales-Espinosa D, García-Orduña F, Rodríguez-Luna E, Bermingham E (2007) Hybridization in large-bodied New World primates. Genetics 176:2421–2425

Davey JW, Blaxter ML (2010) RADSeq: next generation population genetics. Brief Funct Genomics 9:416–423

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499–510

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol Evol 24(6):332–340

Di Fiore A (2003) Molecular genetic approaches to the study of primate behavior, social organization, and reproduction. Yearb Phys Anthropol 46:62–99

Di Fiore A (2009) Genetic approaches to the study of dispersal and kinship in New World primates. In: Garber PA, Estrada A, Bicca-Marques JC, Heymann EW, Strier KB (eds) South American primates, comparative perspectives in the study of behavior, ecology, and conservation, series: developments in primatology: progress and prospects. Springer, Heidelberg, pp 211–250

Dou J, Zhao X, Fu X, Jiao W, Wang N, Zhang L, Hu X, Wang S, Bao Z (2012) Reference-free SNP calling: improved accuracy by preventing incorrect calls from repetitive genomic regions. Biol Direct 7:17–26

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

Finger AJ, Stephens MR, Clipperton NW, May B (2009) Six diagnostic single nucleotide polymorphism markers for detecting introgression between cutthroat and rainbow trouts. Mol Ecol Resour 9:759–763

Green R, Krause J, Briggs A, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz M (2010) A draft sequence of the Neandertal genome. Science 328:710–722

Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC (2007) Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222–234

Hauser L, Baird M, Hildborn R, Seeb LW, Seeb JS (2011) An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. Mol Ecol Resour 11(S1):150–161

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Mol Ecol Resour 11(S1):117–122

Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Mol Ecol Resour 8:3–17

Jameson NM, Xu K, Yi SV, Wildman DE (2012) Development and annotation of shotgun sequence libraries from New World monkeys. Mol Ecol Resour 12:950–955

Kelaita MA, Cortés-Ortiz L (2013) Morphological Variation of Genetically Confirmed *Alouatta pigra* x *A. palliata* hybrids from a natural hybrid zone in Tabasco Mexico. Am J Phys Anth 150:223–234

Keller I, Wagner CE, Greuter L, Mwaiko S, Selz OM, Sivasundar A, Wittwer S, Seehausen O (2012) Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. Mol Ecol 22:2848–2863

Lawler RR (2011) Demographic concepts and research pertaining to the study of wild primate populations. Yearb Phys Anthropol 54:63–85

Locke DP, Hillier LW, Warren WC, Worley KC, Nazareth LV et al (2011) Comparative and demographic analysis of orangutan genomes. Nature 469:529–533

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev 4:981–994

Marques-Bonet T, Ryder OA, Eichler EE (2009) Sequencing primate genomes: what have we learned? Annu Rev Genomics Hum Genet 10:355–386

Moulin S, Gerbault-Seureau M, Dutrillaux B, Richard FA (2008) Phylogenomics of African guenon. Chrom Res 16:783–799

Ochman H, Worobey M, Kuo C, Ndjango JN, Peeters M, Hahn BH, Hugenholtz P (2010) Evolutionary relationships of wild hominids recapitulated by gut microbial communities. PLOS Biol 8:e1000546

Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. Trends Genet 26:177–187

Pemberton JM (2008) Wild pedigrees: the way forward. Proc R Soc B 275:613–621

Perry GH, Marioni J, Pall M, Gilad Y (2010) Genomic-scale capture and sequencing of endogenous DNA from feces. Mol Ecol 19:5328–5331

Perry GH, Melsted P, Marioni J, Wang Y, Bainer R, Pickrell JK, Michelini K, Zehr S, Yoder AD, Stephens M, Pritchard JK, Gilad Y (2012) Comparative RNA sequencing reveals substantial genetic variation in endangered primates. Genome Res 22:602–610

Pipes L, Li S, Bozinoski M, Palermo R, Peng X, Blood P, Kelly S, Weiss JM, Thierry-Mieg J, Thierry-Mieg D, Zumbo P, Chen R, Schroth GP, Mason CE, Katze MG (2012) The non-human primate reference transcriptome resource (NHPRTR) for comparative functional genomics. Nucleic Acids Res 41: D906–D914

Pool JE, Hellmann I, Jensen JD, Nielsen R (2010) Population genetic inference from genomic sequence variation. Genome Res 20:291–300

Primmer CR (2009) From conservation genetics to conservation genomics. Ann N Y Acad Sci 1162:357–368

Raveendran M, Harris RA, Milosavljevic A, Johnson Z, Shelledy W, Cameron J, Rogers J (2006) Designing new microsatellite markers for linkage and population genetic analyses in rhesus macaques and other non-human primates. Genomics 88:706–710

Rogers J, Mahaney MC, Cox LA (2009) The development and status of the baboon genetic linkage map. In: Barrett L (ed) The baboon in biomedical research. Developments.in primatology: progress and prospects. Springer, New York

Rudd S (2003) Expressed sequence tags: alternative or complement to whole genome sequences? Trends Plant Sci 8:321–329

Ryder OA (2005) Conservation genomics: applying whole genome studies to species conservation efforts. Gytogenet Genome Res 108:6–15

Santure AW, Stapley J, Ball AD, Birkhead TR, Burke T, Slate J (2010) On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. Mol Ecol 19:1439–1451

Siepel A (2009) Phylogenomics of primates and their ancestral populations. Genome Res 19:1929–1941

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

Steiner CC, Putnam AS, Hoeck PEA, Ryder OA (2013) Conservation genomics of threatened animal species. Annu Rev Anim Biosci 1:261–281

Stone A, Verrelli B (2006) Focusing on comparative ape population genetics in the postgenomic age. Curr Opin Genet Dev 16:586–591

Thalmann O, Fischer A, Lankester F, Pääbo S, Vigilant L (2007) The complex evolutionary history of gorillas: Insights from genomic data. Mol Biol Evol 24:146–158

Thomas MA, Klaper R (2004) Genomics for the ecological toolbox. Trends Ecol Evol 19:441–445

Thompson RC, Wang IJ, Johnson JR (2010) Genome-enabled development of DNA markers for ecology, evolution and conservation. Mol Ecol 19:2184–2195

Ting N, Sterner KN (2012) Primate molecular phylogenetics in a genomic era. Mol Phylogenet Evol 66:565–568

Tung J, Charpentier MJ, Garfield DA, Altmann J, Alberts SC (2008) Genetic evidence reveals temporal change in hybridization patterns in a wild baboon population. Mol Ecol 17:1998–2011

Tung J, Alberts S, Wray G (2010) Evolutionary genetics in wild primates: combining genetic approaches with field studies of natural populations. Trends Genet 26:353–362

Tung J, Akinyi MY, Mutura S, Altmann J, Wray GA, Alberts SC (2011) Allele-specific gene expression in a wild nonhuman primate population. Mol Ecol 20:725–739

Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanski I, Marden JH (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol Ecol 17: 1636–1647

Vigilant L, Guschanski K (2009) Using genetics to understand the dynamics of wild primate populations. Primates 50:105–120

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63

Wheeler DA, Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM (2008) The complete genome of an individual by massively parallel DNA sequencing. Nature 452:872–876

Yildirim S, Yeoman CJ, Sipos M, Torralba M, Wilson BA, Goldberg TL, Stumpf RM, Leigh SR, White BA, Nelson KE (2010) Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. PLoS ONE 5: e13963

Zhang J, Chiodini R, Badr A, Zhang G (2011) The impact of next-generation sequencing on genomics. J Gene Genomics 38:95–109

Zinner D, Arnold ML, Roos C (2011) The strange blood: natural hybridization in primates. Evol Anthropol 20:96–103