### **REVIEW ARTICLE**

# **Evolutionary Anthropology** WILEY

# Primate landscape genetics: A review and practical guide

Revised: 12 June 2020

Darice Westphal<sup>1,2</sup> | Amanda N. Mancini<sup>1,2</sup> | Andrea L. Baden<sup>1,2,3</sup>

#### <sup>1</sup>Department of Anthropology, The Graduate Center, City University of New York. New York, New York

<sup>2</sup>The New York Consortium in Evolutionary Primatology (NYCEP), New York, New York

<sup>3</sup>Department of Anthropology, Hunter College, New York, New York

#### Correspondence

Andrea L. Baden, Department of Anthropology, Hunter College, City University of New York, New York, NY 10065. Email: andrea.baden@hunter.cuny.edu

#### Abstract

Landscape genetics is an emerging field that integrates population genetics, landscape ecology, and spatial statistics to investigate how geographical and environmental features and evolutionary processes such as gene flow, genetic drift, and selection structure genetic variation at both the population and individual levels, with implications for ecology, evolution, and conservation biology. Despite being particularly well suited for primatologists, this method is currently underutilized. Here, we synthesize the current state of research on landscape genetics in primates. We begin by outlining how landscape genetics has been used to disentangle the drivers of diversity, followed by a review of how landscape genetic methods have been applied to primates. This is followed by a section highlighting special considerations when applying the methods to primates, and a practical guide to facilitate further landscape genetics studies using both existing and de novo datasets. We conclude by exploring future avenues of inquiry that could be facilitated by recent developments as well as underdeveloped applications of landscape genetics to primates.

#### KEYWORDS

gene flow, genetic connectivity, least-cost path, remote sensing, resistance distance

#### INTRODUCTION 1

Population genetics, a subfield of genetics that examines and models changes in genetic variation within and among populations, has allowed scientists to investigate the consequences of migration, gene flow, selection, and demography.<sup>1</sup> Such studies have helped elucidate past influences on current population structure,<sup>2</sup> genetic patterns due to introgression,<sup>3</sup> and best practices for species management.<sup>4</sup> However, as with all molecular techniques, the field is rapidly advancing, continuously incorporating new methodologies, statistical analyses, and sequencing technologies that improve our abilities to ask evermore refined and focused questions.<sup>5-7</sup> Such advances in population genetics, when combined with modern advances in molecular phylogenetics and more recently genomics, has allowed molecular ecologists to better understand nuances of species behavior (sex-biased dispersal<sup>8</sup>) and evolution (adaptation to climate change<sup>9</sup>) that are otherwise difficult to detect from observations alone. Characterizing and understanding the mechanisms driving species diversity are especially important as we enter the Anthropocene.<sup>10</sup> As human activities drive species toward extinction, the knowledge derived from molecular genetic techniques will be essential for informing effective conservation management.<sup>11</sup> Toward this end, landscape genetics<sup>12</sup> has emerged. Landscape genetics combines methods from traditional population genetics with georeferenced spatial data to investigate how features in the landscape facilitate and/or impede gene flow.<sup>13-16</sup> Most simply, landscape genetics compares genetic distances among individuals (or, when appropriate, populations) to heterogeneity in the landscape, including geographic (rivers, mountains) or topographic barriers (ridges, slopes, or valleys), habitat type (rainforest, dry forest, savanna), and measures of anthropogenic disturbance (habitat quality, distance to forest edge, distance to village, presence of farmland, roads, railways, or dams).

Landscape datasets are then transformed to represent least-cost paths (isolation-by-barrier) or resistance distances (isolation-byresistance),14,17 and correlated with genetic distances, the results of which are compared to a null hypothesis of isolation-by-distance.<sup>18</sup> Where landscape genetics differs from more traditional population genetics is in its use of spatially explicit model testing; whereas

#### BOX 1 GLOSSARY

Population genetics: Subfield of molecular genetics that examines and models changes in genetic variation within and among populations.

- Landscape genetics: Combines methods from traditional population genetics with georeferenced spatial data to investigate how features in the landscape facilitate and/or impede gene flow by comparing genetic distances among populations (or more ideally individuals) to heterogeneity in the landscape in terms of composition and configuration.
- Spatial extent: Size or boundaries of the study area. This is an important consideration in landscape genetic studies as spatial extent needs to incorporate dispersal distance of the species as well as encompass landscape features of interest.
- Spatial resolution: Grain or pixel size of classified landscape features. This will often depend on the source of the landscape data and the limitations on sensitivity of the equipment used to gather the data.
- Sampling scheme: Systematic way in which sampling locations are selected. This will vary depending on species distribution and field site conditions.
- Sampling intensity: How many genetic samples are collected at each sampling location. More samples are not necessarily always best, and low samples per site may be overcome with careful choice of genetic markers with greater polymorphism.
- Genetic distance: Representation of genetic variation derived from collected genetic data, often in the form of a matrix between all individuals sampled. More distantly related individuals or populations will have higher genetic distances.
- Resistance distance: Representation of the distance that reflects the route(s) of gene flow around barriers that connect two individuals or populations. Derived from circuit theory, a resistance distance allows for more than one path between connected individuals weighted by genetic connectivity.
- Isolation-by-distance: A common null hypothesis for population and landscape genetic studies. Individuals that are more geographically distant are also more genetically distant.
- Isolation-by-barrier: Geographically close individuals are less related than expected due to a physical barrier to gene flow occurring between them. Barriers commonly include rivers, valleys, and mountains.
- Isolation-by-resistance: Genetic distance and impediments to gene flow are positively correlated. This is an expanded idea of isolationby-barrier and includes several barriers (usually in the form of landscape cover, terrain, and climate) that together may form more complicated impediments to gene flow than a single barrier.
- Least-cost-distance/path: A single distance or path that reflects the most parsimonious route of gene flow around barriers that connect two individuals or populations.
- Remotely sensed imagery: Imagery obtained without visiting field locations, often via satellites and unmanned vehicles. Types of data collected vary by sensor, but common ones include collection of many wavelengths of electromagnetic radiation that can be compiled into landcover data.
- Parameterizing: Used to determine how to weight landscape features to reflect ability of a species in moving across said feature when working with resistance surfaces. For example, multiple parameterization schemes are often tried in order to determine how much more difficult it is for a species to cross a river versus a road.

population genetic assessments such as Discriminate Analysis of Principal Components and Bayesian clustering assess patterns of genetic structuring without a priori assumptions of sampling provenience or explicit interest in relating the observed structure to landscape variables, landscape genetic methods allow researchers to statistically evaluate landscape features hypothesized to affect the process of gene flow across a continuous landscape (Figure 1).

One common area of confusion surrounding landscape genetics is its distinction from phylogeography. Phylogeography and landscape genetics differ in terms of scale (see<sup>20</sup> for dissenting view). Whereas phylogeography focuses on historical processes that have generated observed genetic patterns through deep time, such as historic climate change/glaciation cycles and expansion/contraction of habitat in response to these changes,<sup>21</sup> landscape genetics investigates more recent processes, such as deforestation due to agriculture or expanding urbanization. Studies can combine phylogenetic and landscape genetic techniques in ways that complement each other; however, time scales need to be clearly defined.

In this review, we synthesize the current state of landscape genetics, specifically as it relates to wild non-human primates, with a goal of encouraging primatologists and conservationists to incorporate landscape genetic methods into their existing research toolkit. Our goals in this review are fourfold: First, we review the literature as it stands on the application of landscape genetics to primates, highlighting the major drivers of and impediments to primate gene flow identified by such studies. Second, we highlight several difficulties associated with primates as a study system, including their sociality, dispersal biases, and life histories, that make the application of landscape genetic methods to primates particularly complex. Third, we provide recommendations for overcoming common landscape genetic problems such as scale,



**FIGURE 1** Examples of georeferenced landscape data that can be used in downstream landscape genetic analyses, including (from left to right) habitat type, rivers, and roads; distance to nearest village; and topographic position index (TPI). Once parameterized, landscape resistance surfaces are correlated to genetic distances, and paths of least resistance (i.e., highest gene flow) among sampling localities can be illustrated via current density maps using programs like Circuitscape (far right). Warm colors indicate areas of high conductive value (i.e., low resistance to gene flow, high dispersal ability); cool colors indicate areas of low conductive value (i.e., high resistance to gene flow, low dispersal ability). Figure modified from Baden et al<sup>19</sup>

time lags, and sampling, coupled with guidelines and suggestions for conducting a landscape genetic analysis with consideration given to issues applying this analysis to primates. And finally, we touch on future directions for landscape genetic studies in primates and describe scenarios in which landscape genetics can provide valuable insight into longstanding evolutionary and ecological questions.

### 2 | LANDSCAPE GENETICS IN PRIMATES

Although nearly 20 years has passed since the inception of landscape genetics, its application to primates is still in its infancy.<sup>13</sup> To date, a total of 17 published primate studies have used landscape genetic methods in some form (Tables 1 and 2). We do not include studies that did not explicitly test landscape features but rather only assessed isolation-by-distance. Half of these focused on cercopithecoids, including one African (*Procolobus gordonorum*<sup>34</sup>) and two Asian colobines (*Rhinopithecus bieti*<sup>30-32</sup> and *Trachypithecus leucocephalus*<sup>35</sup>), vervets (*Chlorocebus pygeryhtrus*<sup>29</sup>), and long-tailed macaques (*Macaca fascicularis*<sup>28</sup>). The remaining eight studies span diverse taxa including lemurs, <sup>19,22,23</sup> platyrrhines,<sup>24-27</sup> and hominoids.<sup>36,37</sup> Of these, the vast

majority have been published in the last 6 years as methods have become more accessible via freely available software such as R and QGIS. This recent uptick in studies suggests that landscape genetic methods are becoming more commonplace, an encouraging trend that we hope continues.

To date, primate landscape genetics studies have been motivated by conservation, with their primary aims focused on identifying the natural and anthropogenic impediments to gene flow across human modified landscapes. In some cases, studies were direct extensions of earlier work, allowing researchers to empirically test hypothesized drivers of genetic structuring as inferred from initial population genetic investigations.<sup>19,31,32,35</sup> However, a few studies were new, applying landscape genetic analyses in combination with traditional population genetic tests for the first time in that species or population,<sup>26,34,36</sup> or taking an adaptive focus from previous neutral population genetics analyses.<sup>29</sup>

While nearly all studies found evidence of restricted gene flow, the landscape barriers varied by taxon, geographic region, and the intensity of anthropogenic threat. In many cases, results supported traditional assumptions of population genetic theory: areas with intact habitat were almost unanimously identified as features that facilitated

o genetic analyses (excludes studies using geographi	
Only includes studies with specific landscape features in addition t	
Comparison of primate landscape genetic studies.	ly)
<b>TABLE 1</b>	distance on

distance only)										
Таха	Habitat type	Landscape features	Main application	Size of study area (km²)	Dispersal distance [home range size]	Generation length (years)	Landscape features evaluated	Source of genetic samples	Collection of samples	Sample size
Lemurs										
Microcebus ravelobensis <sup>22</sup>	CDF	A,N	U	955	2 km [85 m]	1	AI	μ	ď	187
Propithecus tattersalli <sup>23</sup>	FrDF	A,N	C, E	2,450	800 m [200 m]	Us	Ri, Ro, Ve	ш	ď	230
Varecia variegata <sup>19</sup>	C&FrTF	A,N	C, E	$\sim 107,000^{a}$	200 km [15-20 ha]	7-8	Ri, Ro, To, Ve, vi	В, F	Ч	200
Cebids										
Saimiri oerstedii <sup>24</sup>	FrTF	A,N	C, E	1,800	Unk [200 ha]	3-6	Ag, P, Ri, Ve, Vi	ш	C	244
Callitrichids										
Saguinus (simulated) <sup>25</sup>	SimF	z	ш	16	Unk [27 ha]	CJ	Ve	S	U	288
Leontopithecus rosalia <sup>26</sup>	FrTF	A,N	C, E	~4,000ª	6 km [47–67 ha]	Us	Ag, P, Ri, Ro, Ve	т	Ъ	201
Atelids										
Alouatta caraya <sup>27</sup>	FrTF	A,N	C, E	$\sim$ 500,000 <sup>a</sup>	1,000 km [Us]	5	Ag, P, Ri, Ro	ш	U	138
Cercopithecines										
Macaca fascicularis <sup>28</sup>	DisF	A,N	ш	5,780	Us [100 ha]	Us	Al, Ag, Tu, Ve, Vi	B, F	C + D	345
Chlorocebus pygerythrus <sup>29</sup>	S	U	۲	$\sim$ 1,220,000 <sup>a</sup>	Us [Us]	Us	U	F	¢.	81
Colobines										
Rhinopithecus bieti <sup>30–33</sup>	FrTeF	z	C, E	17,000	Unk [Us]	J.	Al, Ro, Ve, Vi, P	В, F, Т	с.	135-157
Procolobus gordonorum <sup>34</sup>	FrTF	A,N	U	19,000	Us [Us]	6-8	F, Ra, To, Ve, Vi	ш	U	121
Trachypithecus leucocephalus <sup>35</sup>	FrTF	A,N	U	256	14 km [Us]	Us	Ro, Ve	щ	ď	214
Hominoids										
Pongo pygmaeus <sup>36</sup>	FrTF	z	C, E	270	Unk [Us]	Us	Ri	F,H	Ч	200
Pan troglodytes <sup>37</sup>	CTF	C, N	ш	$\sim 500,000^{a}$	Us [Us]	Us	C, Ri, Ve	F,H	Ъ	187-345
Note: Habitat Type: C = C = climatic; Main Appli unknown at time study = topography, Tu = tou <sup>a</sup> Size of Study Area: stuc	continuous, cation: $A = I$ was conduc rism, Ve = v( ly area size v	, Dis = disturbed Adaptation, C = ( Adaptation, C = ( Adaptation, Vi = v estettion, Vi = v was estimated fr	<ul> <li>I, Fr = fragmentei</li> <li>conservation, E =</li> <li><i>Length</i>: Us = uns</li> <li>illages; Source Ge</li> <li>om figures and v</li> </ul>	<ul> <li>d, D = deciduous, 1</li> <li>ecology: Dispersal</li> <li>tated: Landscape Fi</li> <li>snetic Samples: T =</li> <li>vas not explicitly st</li> </ul>	e = temperate, T = tropical, Distance [Home Range]: Us atures Evaluated: Ag = agri tissue, F = fecal, B = blood, ated within articles.	F = forest, S = savan = unstated, no menti culture, Al = altitude, ' S = simulated, H = ha	na, Sim = simulated; <i>Lan</i> on of feature in article, <sup>I</sup> C = climate, F = fire, P = iir; <i>Collection of Samples</i> :	<i>ddscape Features</i> : N = Unk = unknown, featu pastures, Ri = rivers, : P = previous study, (	natural, A = anthre ure was mentionec Ra = railroad, Ro <sup>=</sup> C = current study.	pogenic, I but : Roads, To

4 WILEY Evolutionary Anthropology

TABLE 2 Compariso	n of genetic ma	ırkers, landscape data,	and statistical methods	used in primate land	scape genetics stu	udies organized b	y taxon		
Таха	Genetic markers	Genetic distance metric	Genetic computer program	Landscape data source	Landscape programs	Spatial resolution	Parameterizing method	Correlation method	Correlation program
Lemurs									
Microcebus ravelobensis <sup>22</sup>	8 8	Ľ	S	F	AV	1 km	EX	Σ	S
Propithecus tattersalli <sup>23</sup>	13 m	Я	S	L	EN	30 m	υ	M, PM	Us
Varecia variegata <sup>19</sup>	10 m	F <sub>ST</sub>	щ	D, G, Mad	AG	1 km	G, R	_	lme4, rGA
Cebids									
Saimiri oerstedii <sup>24</sup>	16 m	Я	A	A, L	AG	20 m	R, S	M, PM	ZT
Callitrichids									
Saguinus (simulated) <sup>25</sup>	20 m (sim)	F <sub>ST</sub>	F, h	Sim	U	8 m	EX	M, PM	FSTAT
Leontopithecus rosalia <sup>26</sup>	14 m	F <sub>ij</sub>	S	ĻТ	ra, rg	30 m	Ex	ט	Я
Atelids									
Alouatta caraya <sup>27</sup>	7-10 m, mt	$D_{ps}$	٩	Ш	Us	Us	EX	M, PM	vegan
Cercopithecines									
Macaca fascicularis <sup>28</sup>	11 m	Я	60	GS, Mi	Us	1 km	Us	ξ	ade3, gstat
Chlorocebus pygerythrus <sup>29</sup>	с С	F <sub>ST</sub> , Hd	A, D	>	Us	Us	Us	GI, M	A, PAST
Colobines									
Rhinopithecus bieti <sup>30–33</sup>	10 m, mt	Я	S	S	AG, AV, ER	60 km-50 m	Ex, P	L., M, MR, PM	A, E, G, R
Procolobus gordonorum <sup>34</sup>	10 m	Ж	S		AG	500 m	G, R		lme4, rGA
Trachypithecus Ieucocephalus <sup>35</sup>	10-15 m	ж	S	_	Us	Us	Ĕ	M, PM	A
Hominoids									
Pongo pygmaeus <sup>36</sup>	14 m	F <sub>ST</sub>	G, R	Us	ц	Us	Ex	Σ	ц
Pan troglodytes <sup>37</sup>	21 m, mt	F <sub>ST</sub> , TN	A	H, M, N, Q, W	AG	1 km	Ec, R	Gd	gdm
Note: Genetic Markers: g = index, Hd = haplotype dive packages start with lowerr L = LANDSAT, M = MODI Biome data (Mucina & Rut with lowercase letters), Us P = permutation tests, R = t-tests, L = linear mixed eff lowercase letters), rGA = R	genes, m = micr ersity, R = Rouss case letters), S = S, Mad = CEPF h therford 2006), V = unstated; <i>Spa</i> resistance analy fects models, M ResistanceGA, S	saatellites, mt = mtDNA set's a, TN = Tamura & N SPAGeDi; <i>Landscape Dc</i> Madagascar Vegetation V = WORLDCLIM; <i>Lana</i> <i>V</i> = WORLDCLIM; <i>Lana</i> <i>V</i> = wORLDCLIM; <i>Lana</i> <i>V</i> = wORLDCLIM; <i>Lana</i> <i>V</i> = morel, <i>V</i> = unstitiple = Mantel, MR = multiple = SPAGeDi, Us = unstat	, sim = simulated; Genetic lei's index; Computer Prog ta Source: A = aerial phot Mapping Project, Mi = mi lscape Programs: AG = Arc lscape Programs: AG = Arc tated; Parameterizing Met tated; Parameterizing det ed.	Distance Metric: D <sub>ps</sub> = ram: A = Arlequin, D = os, D = DIVA-GIS, E = litary maps, N = NAS/ isdIS, AV = ArcVIEW, I hod: C = classification hod: C = classification ation Method: Gd = ge Mantel, Correlation P	<ul> <li>1 - proportion of</li> <li>E DnaSP, F = fstat,</li> <li>E European space a</li> <li>A SRTM, S = SPOT,</li> <li>EN = ENVI, ER = Ef</li> <li>Ibased on opennes</li> <li>Ibased dissimilation</li> <li>rogram: A = Arelqui</li> </ul>	shared alleles bet G = Genetix, g = g gency portal, G = 1 Sim = simulated, RDAS, G = GENES s, Ec = Ecological rity models, GI = g in, E = ECODIST, (	ween populations, $F_{ij}$ = stat, h = hierfstat, p = GeoNET, GS = ground T = topographic map, I YS, ra = raster, rg = rg, riche model, Ex = exp niche model, Ex = exp eneralized linear effec	<ul> <li>kinship coefficient</li> <li>PopGenReport, R =</li> <li>Poryev, H = Hyrdo</li> <li>Js = unstated, V = 'V</li> <li>Jal, R = R (specific fill</li> <li>fat opinion, G = ger</li> <li>ts models, K = krigit</li> <li>specific R packages</li> </ul>	t, F <sub>51</sub> = fixation : R (specific R SHEDS, Vegetation & R packages start netic algorithms, ng with pairwise start with

5

gene flow (14 of 17 studies),<sup>28,29,34</sup> and well-known biogeographic dispersal barriers such as rivers and mountain ranges were characteristically identified as being difficult to traverse (that is, animals encountering these landscape features experienced high resistance to movement).<sup>22,23,36,37</sup> Moreover, deforestation,<sup>19,24,26,27,30–33,35</sup> urbanization,<sup>19,28,32,34</sup> and high human population densities<sup>33</sup> and/or activity<sup>26,27,30–32,35</sup> were typically identified as posing higher resistance to primate gene flow, than areas experiencing less anthropogenic disturbance. These patterns are not unique to primates and have been found to affect taxa at a global scale.<sup>14,38</sup>

Landscape genetic studies in primates that integrated both natural and anthropogenically-driven landscape features often identified anthropogenic features as being better predictors of current genetic structure than naturally occurring barriers to gene flow. For instance, Ruiz-Lopez et al.<sup>34</sup> expected altitude to be the most likely barrier to gene flow in the Udzungwa red colobus monkey (Procolobus gordonorum) due to the ecological tolerances of the species. However, analyses instead identified distance to nearest village and fire density to be better predictors, with increased fire density resulting in reduced forest coverage leading to lower genetic connectivity. Similarly, proximity to nearby human settlements and habitat degradation best explained species-wide population genetic structure in black-and-white ruffed lemurs (Varecia variegata),19 despite earlier evidence suggesting that the Mangoro River was the most likely barrier to gene flow across the species range.<sup>39</sup> Both V. variegata and P. gordonorum have longer generation times (6-8 years, Table 1) for their body size and thus results were unexpected (see Primate Considerations below), leading authors to surmise that rapid and intensifying anthropogenic pressures are effectively swamping signatures of historic gene flow in these species.<sup>19,34</sup> It will be interesting to see whether these same patterns hold up in other taxa as studies increase.

Equally important are landscape genetic studies that disentangle the anthropogenic features impacting species. For instance, in a study of Central American squirrel monkeys (Saimiri oerstedii), oil palm plantations, the predominant landscape feature in the region, were found to pose greater barriers to movement than smaller, more isolated cattle pastures or residential areas,<sup>24</sup> suggesting that dispersal costs can become magnified over greater distances. Studies such as this allow researchers to at once identify both the scale and configuration of relevant landscape features, allowing conservation practitioners to develop targeted conservation action plans.<sup>24</sup> There are also rare occasions where anthropogenic features are found to facilitate rather than impede gene flow, as in long-tailed macaques (Macaca fascicularis).<sup>28</sup> Large urban areas with high human population density had no impact on genetic connectivity in M. fascicularis, demonstrating that macaques can move freely across the landscape and that provisioning sites encourage rather than restrict gene flow, even in seemingly inhospitable areas.<sup>28</sup> Results such as these are equally as valuable in their ability to identify particularly flexible species. In fact, they may help to highlight other more pressing threats, such as an increased potential for human-wildlife conflict or disease transmission, thereby facilitating more targeted conservation efforts.

Given the vast diversity of primates, it is perhaps unsurprising that landscape features impact species differently. For instance, roads have been found to limit dispersal in golden-brown mouse lemurs (*Microcebus ravelobensis*),<sup>22</sup> but not the golden-crowned sifaka (*Propithecus tattersalli*),<sup>23</sup> outcomes that are likely tied to differences in body size, locomotion, and dispersal abilities.

Finally, when considering landscape permeability, timing matters. In their study of sifakas (*Propithecus tattersalli*), Quéméré et al.<sup>23</sup> found that seasonality can influence barrier detection in relation to species behavior. Despite being dry for most of the year, rivers acted as major barriers to golden-crowned sifaka movement because the rainy season, when rivers form, coincides with major dispersal events, thereby reducing dispersal ability during this critical period. Taken together, these points highlight the unique benefits of landscape genetic methods that simultaneously allow us to consider multiple features in the landscape, as well as the importance of combining multiple lines of evidence including genetics, ecology, and behavior to best understand animal behavior and evolution.

Although the vast majority of landscape genetic studies in primates have been conservation-minded, landscape genetics has great potential to inform our understanding of basic ecological influences for species. For instance, Mitchell et al.<sup>37</sup> used landscape genetic methods to investigate the role of rivers as barriers to gene flow between two chimpanzee subspecies as well as the role of habitat and elevation in genetic differentiation within subspecies. Additionally, landscape genetic studies can identify the role of adaptation to varying environments. For example, Coetzer et al.<sup>29</sup> identified genes associated with immune adaptation to different disease environments to determine appropriate locations for reintroductions for rehabilitated vervets. Unfortunately, far fewer studies focus on these questions and we believe this area of inquiry presents a valuable avenue for future research.

### 3 | CONSIDERATIONS FOR PRIMATES

Primates, as an order, are known for their broad diversity in ecological tolerances, social organizations, and life history variables. In this section, we highlight aspects of primate life history and behavior that warrant special consideration when designing and implementing primate landscape genetic (hereafter PLG) studies. While these traits and issues are not unique to primates, and can apply to many long-lived highly social mammals, they require careful consideration because poor application of landscape genetic techniques can lead to erroneous inferences.

#### 3.1 | Social organization

Primates are a particularly difficult group of animals to study using landscape genetics due to their gregariousness. Despite exhibiting extensive variation in social organization, most primates form groups, which de facto creates a clumped distribution of animals across the landscape. By extension, this also leads to gaps (discontinuities) between sampling locations. When sampling occurs in too close proximity or from too dispersed locations this can create artificial groups and may lead to erroneous support for isolation-by-barriers.<sup>40</sup> Moreover, as with many endangered taxa, collecting sufficient samples from primates is often further complicated by the external challenges of fieldwork, including obtaining permits for sample collection and export, financial, logistical, and/or ethical sampling constraints, and logistical difficulties encountered on-the-ground at the study site(s).

One way to address these challenges is by choosing an appropriate sampling scheme (see discussion on sampling scheme and intensity in project design and data collection below). Most PLG studies reviewed herein did not mention their choice of sampling scheme when designing their study (although see Lane-deGraaf et al<sup>28,34</sup>). Of those that did, very few studies employed a systematic scheme that sampled from almost all social groups across the entire range of the species (but see Liu et al<sup>30-32</sup>). Nevertheless, even when sampling scheme was not explicitly reported, all studies reported that sampling occurred on either side of suspected landscape barriers. Moreover, it is important to highlight that many PLG studies were not designed a priori and were either extensions of previous work or opportunistic use of samples that had been collected for other purposes (Table 1).

#### 3.2 | Dispersal biases

In addition to exhibiting extensive variation in social organization, primates are among the few mammals to exhibit near complete dispersal bias.<sup>41</sup> Nevertheless, they show great variability in their patterns of philopatry, dispersal distance, and home range size (Table 1). Sex-biases in dispersal and/or sex-specific preferences for dispersal routes can lead to differential landscape effects on connectivity and evolutionary processes between sexes of the same species. When dispersal biases are present, it is therefore advisable to conduct separate analyses to determine sex-specific responses to the landscape features in question.<sup>42</sup>

In many of the published PLG studies to date, dispersal distance and/or home range size was unknown for the study species and thus approximating dispersal distance was among the primary objectives. Home range is often used as a proxy for minimum dispersal distance when actual dispersal distance is unknown. However, many studies made no statement about dispersal distance, dispersal biases, or home range size. Without this information, it is difficult to determine whether the landscape features identified as influencing genetic structure were actually the cause of the observed pattern or rather if their results were a mismatch between sampling regimen and dispersal distance.

It is therefore important to take dispersal distance into consideration when designing a landscape genetic study, as this will influence the *spatial extent*, that is, the size of the study area, that is appropriate for analysis. Study areas varied enormously across PLG studies (Table 1), from all of South Africa  $(\sim 1.22 \text{ million km}^2)^{29}$  to a small simulated landscape (16 km<sup>2</sup>).<sup>25</sup> In some cases, study areas include the entire known range of the species, as in the cases of *Rhinopithecus bieti*,<sup>30-32</sup> *Procolobus gordonorum*,<sup>34</sup> and *Varecia variegata*,<sup>19</sup> or in areas where dispersal is naturally constrained as in *Macaca fascicularis* on the island of Bali, Indonesia.<sup>28</sup>

### 3.3 | Life history and demography

As with dispersal distance, generation length can also impact whether a given landscape feature's influence on gene flow will be detectable, as time lags exist.<sup>43</sup> Primates have a wide range of generation lengths, ranging from approximately one year (*Microcebus*<sup>22</sup>) to 20 years (*Pan*<sup>37</sup>), and although a landscape feature, such as a road, might influence gene flow in both a mouse lemur and a chimpanzee, its effects may only be visible in the species where sufficient generations have passed to allow for its detection via population genetic methods. Furthermore, short dispersal distances can compound with generation length to increase the effect of time lags in a species, such that those species with short dispersal distances will need more time before showing a genetic response to landscape features.<sup>20,43</sup>

Despite this, only about half of PLG studies report the generation length of their species (Table 1). Many studies evaluated features that have existed without much change for conceivably longer than even the longest generation times (i.e., rivers and altitude). However, some studies also included anthropogenic features, such as villages, roads, agriculture, and even tourism, which can appear rapidly and change over short periods of time. Such features should only be used with species of generation lengths appropriate for evaluating the feature(s) in question (Table 1).

#### 3.4 | Ethics of sampling nonhuman primates

Many primate taxa are threatened with extinction (42.3%), with 128 species currently listed as Endangered and 68 listed as Critically Endangered.<sup>44</sup> This creates numerous ethical challenges and legal limitations in regard to sample collection from these taxa,<sup>45</sup> as often only non-invasive, low-quality samples (feces) can be obtained at the sufficient extent and numbers (>200 individuals) required for a landscape genetic assessment.<sup>46</sup> The quality of samples that can be collected will influence the choice of genetic markers selected, along with the conclusions that can be drawn from results.

### 4 | IMPLEMENTING A PRIMATE LANDSCAPE GENETICS STUDY

When designing and implementing a landscape genetic study, there are several key considerations (Figure 2). The following section aims to introduce interested primatologists to the major steps necessary to implement a landscape genetics study, as well as highlighting current best practices in the field. For additional guidance see Hall and Beissinger<sup>47</sup> and refer to references cited for a more detailed

# **WILEY\_Evolutionary Anthropology\_**





#### BOX 2 Working with existing data

While a landscape genetics study should ideally be designed from the ground up with regard to sampling strategy and marker choice, researchers can incorporate existing datasets to make use of landscape genetic techniques, so long as several considerations are addressed. Already collected samples may be used if enough samples exist with even coverage over the landscape or additional samples can be collected or combined across collaborators. Samples should all be from similar time periods with respect to the species' generation length, such that species with longer generation times may use samples collected from several years. If older samples are used, be careful that landscape data match the time frame of the genetic samples in question. For example, do not use landscape data collected from 2016 to understand genetic connectivity in samples collected in 2000.

Historical datasets or data collected from long-term field sites present a unique opportunity to look at changes in landscape connectivity.<sup>48, 49</sup> In the first ever application of time series data to landscape genetics, Draheim et al.<sup>50</sup> found that the landscape features varied in impact on the influence of genetic connectivity in black bears. In some years, only landcover correlated with genetic connectivity and in other years rivers did. This pattern was revealed even over relatively short times spans (4 years) and was surprising given the longer generation times of bears (6 years).<sup>50</sup> Though not yet evaluated in primates, such time series data are undoubtedly available and would provide myriad opportunities for such investigations.

It is important to remember, however, that when working with existing data, the original sampling scheme may limit the types of questions that can be asked. For example, if the sampling scheme is clumped and does not include samples from both sides of a suspected barrier, the research should not focus on determining if the suspected barrier reduces gene flow. That said, many behavioral based studies have unique data sets that could allow them to examine questions that are not otherwise possible with less well-known species. For example, inclusion of favored fruiting trees, salt licks,<sup>51</sup> and sleeping trees<sup>52</sup> can represent unique landscape features that could be further tied to movement studies. If these types of data are not available, researchers can explore what types of questions could be answered if these data were collected and consider adding it to their data collection regimen.

discussion of the topics addressed herein. While the initial considerations and design may seem daunting, this should not discourage research groups from applying landscape genetic methods to pre-existing datasets (see Box 2). However, understanding the possible biases inherent in co-opted datasets is a necessary step to ensure that interpretations do not overreach study limitations.

9

#### 4.1 | Project design and data collection

The spatial extent of a study should be chosen to encompass landscape features that are thought to influence dispersal in a species. When testing landscape variables it is necessary to consider when they first emerged, as time lags are common between species reacting to changes in the habitat and the corresponding change in their genetic patterns.<sup>43</sup> Time lags are primarily the result of species generation length and dispersal distance, as noted above, although population size, genetic structure, and the choice of genetic marker and its mutation rate also influence the time between the emergence of a landscape feature and its resulting effects in the genetic structure of a species. In general, at least five generations are needed in order for the genetic signatures of a landscape feature to be detectable in the population of interest.<sup>53</sup>

In addition to study extent, researchers must ensure the distances between sampling points remain within the dispersal distance of the species, otherwise spurious genetic structuring is likely to occur.<sup>40,46</sup> Given the variation in primate dispersal abilities (see Section 4.2 above), it is clear that certain features may highly influence one species while having no effect on another within the same extent. If dispersal distance is unknown, home range size can be used to estimate this metric.<sup>54</sup> Estimates based on home range size can be corroborated by paternity analyses<sup>55</sup> or migrant detection when both datasets are available.<sup>56,57</sup> When using existing datasets where large gaps between sampling locations exceed known or estimated dispersal capabilities, it is best to reduce the study area to a smaller extent or to subdivide the study area to remove large sections where samples could not be collected.

Simulations have demonstrated that random sampling outperforms other more directed sampling schemes to correctly identify landscape drivers of observed genetic structure.<sup>40,46</sup> Thus, when sampling de novo, an ideal scheme should include a random collection of samples distributed evenly across the study extent and on either side of suspected dispersal barriers. This can often be logistically infeasible due to several factors, including a species distribution, social organization, and/or mating system, or due to the topography, terrain, and/or accessibility of the study region. Given the highly social nature of most primates, it may therefore be more feasible to employ a systematic (i.e., collecting samples at regularly space intervals), clustered, or hybrid sampling scheme.<sup>34,40</sup> These sampling schemes also perform well in simulation studies and are the most likely to recover individuals in clumped space while minimizing bias from overand under-sampling. Whether designing a new study or capitalizing on previously collected data, it is important to consider the spatial distribution of the samples, as these can influence genetic signatures in downstream analyses. Additional sampling schemes that may be appropriate are reviewed and evaluated in Oyler-McCance et al.<sup>46</sup> and Schwartz and McKelvey.<sup>40</sup>

Sampling from primate taxa often occurs in extremely remote and challenging environments and trade-offs exist when making decisions about *sampling intensity*. These trade-offs include sampling additional individuals from one group or from additional populations, as well as whether to use many genetic loci or few loci with higher variability. Studies using simulated data suggest large samples sizes (>200 individuals) coupled with a high number (>20 microsatellites) of highly variable (>10 alleles) genetic markers are necessary to detect influences of landscape on genetic structure.<sup>46,58</sup> When deciding between sample size (i.e., number of individuals) versus marker number and variability, previous work suggests that the latter should be prioritized.<sup>58</sup> Benchmarks for sample size using single nucleotide polymorphism (SNP) data are not currently available, as simulation studies assessing sample size in landscape genetics have thus far focused on microsatellites. However, it has been established that population genetic inferences are more robust to small sample size when using SNP data compared to microsatellites. Furthermore, as few as 1,000 SNP loci are sufficient to reliably discern patterns in genetic structure.<sup>59</sup> Many SNP ascertainment methods produce a large number of alleles (between 100-100,000 bi-allelic SNPs<sup>60-62</sup>), suggesting that researchers can focus more on obtaining large sample sizes when using these markers. At present, SNP data are best achieved from high quality samples, often blood or tissue (although newer methods are quickly evolving; see Supporting Information). Thus, to achieve these aims primate researchers must consider the feasibility of accessing populations and their ability to acquire the desired number and quality of samples. Obtaining a sufficient sample size will likely entail collecting low-quality samples (feces) from across the landscape due to the ethical and legal limitations in acquiring high-quality samples (blood and tissue) from most primate taxa.

#### 4.2 | Analyzing genetic data

When designing a new study, SNP data are recommended, as technological advances now allow large numbers of SNPs to be generated at a similar cost to microsatellite data without the need for marker development<sup>61</sup> (but see Helyar et al<sup>63</sup>). SNPs are now being employed in landscape genomics approaches (reviewed in Li et al<sup>64</sup>), with the benefit being that one can investigate the landscape drivers of both neutral and adaptive genetic variation.<sup>65</sup> That said, SNPs often require higher quality samples (blood and tissue) for development of the SNP array or availability of a reference genome<sup>66</sup> for bioinformatic processing, with reference-based processing outperforming de novo approaches in accuracy of downstream genetic inferences.<sup>67</sup> Numerous primate reference genomes are now freely available,<sup>68</sup> and more are currently in development (www.hgsc.bcm.edu/non-human-primates), facilitating the use of genomic data in future investigations. Furthermore, recent methodological advancements now allow researchers to increase the quality of DNA extracted from non-invasive samples like feces, 60,69 albeit with varying success.<sup>70,71</sup> See Supporting Information for further detail on SNP development in PLG studies.

Despite these recent advancements toward using genomic data, most landscape genetic studies to date have relied primarily on microsatellite markers.<sup>38,72</sup> Microsatellites can capture a high degree of variability with relatively few loci and are feasible to employ with low-quality samples. When used with an appropriate dataset, microsatellite markers can reliably uncover landscape drivers of population connectivity and are an option for primatologists with existing datasets. Studies being developed de novo, however, should focus on the use of SNPs in place of microsatellites due to their many advantages (see Supporting Information for full discussion). Markers with biased inheritance patterns, such as mitochondrial DNA (mtDNA), are generally not ideal for landscape genetic analyses, unless specifically interested in questions concerning historical landscape features, and are often better suited to phylogeographic studies.<sup>20</sup>

Regardless of which marker is ultimately used, landscape genetic methods are best suited for use with individual-based distance estimates, as these do not a priori or artificially assign individuals to populations.<sup>73,74</sup> Distances based on principle component analysis (PCA) with several axes (more than 16) and Rousset's *a* are the most robust metrics for evaluating genetic distances when sample size and/or genetic structure is low; all others perform equally well with a large sample size and high degree of genetic clustering.<sup>75</sup> Population clustering analyses, although not explicitly landscape genetics, can be used to evaluate the level of genetic clustering present in the data. Many other genetic distances can and have been used in landscape genetic studies and a full review of those metrics including comparisons and utility for landscape genetics is provided by Shirk and coauthors.<sup>76</sup> Unmodified  $F_{ST}$  values should be avoided, as they are sensitive to time lags.<sup>43</sup>

Relatedness should also be accounted for early on in sample processing to reduce bias in downstream analyses. Because primates are social animals, this often means that samples collected in the same area or sample location belong to related individuals. When there is high relatedness within a genetic sample or several related individuals cluster geographically, estimates of population structure can be biased.<sup>77</sup> When left in the sample, the resulting genetic pattern is swamped by the signal from these highly related individuals. Relatives can be identified by parentage and sibling analyses and removed prior to later analyses; for a review of available parentage analysis programs see Jones et al.<sup>78</sup> Moreover, individuals sampled from the same group will have the same geographic location, which can further bias results. To correct for this, researchers can use the harmonic mean of the genetic distance measure for individuals sampled from the same location, leaving a single genetic measure for the reported spatial location.<sup>34</sup>

Most of the PLG studies reviewed here used microsatellites combined with one or more mtDNA regions (Table 2); the number of microsatellites used ranged from 7 to 21 (Table 2). Additionally, most studies used Rousset's *a*, a derivation of  $F_{ST}$  ( $F_{ST}/[1 - F_{ST}]$ ), while some used unmodified  $F_{ST}$  (Table 2). All the PLG studies reviewed performed a population structure analysis either in previous studies using the same samples or prior to landscape genetic analysis in the same study. Some even included structure analyses incorporating spatial information.<sup>28</sup>

### 4.3 | Working with landscape data

In addition to genetic data, a PLG analysis requires that the researcher have corresponding georeferenced landscape data. These are typically sourced from *remotely sensed imagery*, which can be obtained from private companies or more commonly from government or open source databases.<sup>79,80</sup> Landscape data should encompass both the

composition (classification) of landscape features as well as their arrangement (that is, their configuration). The availability and accuracy of these data can ideally be found on local government sources or openly available platforms like DIVA-GIS (www.diva-gis.org), EARTH-DATA (earthdata.nasa.gov), OpenStreetMap (www.openstreetmap.org), or the Hansen Global Forest Change dataset.<sup>81</sup> Researchers should always check the *spatial resolution* (or grain; Figure 3) of their landscape data to be sure that the spatial resolution of genetic and landscape data are comparable prior to analysis. Spatial resolution can be modified as necessary, but this must be determined a priori. Once acquired, landscape data can be manipulated in propriety programs like ArcGIS (www.esri.com) as well as in free, open source packages like QGIS, many of which are based in the R statistical computing environment (Table 2); see details in Supporting Information for additional resources.

Landscape data can also be obtained from species distribution or ecological niche models,36 light detection and ranging (LiDAR),82 and landscape features noted from cartographic maps or encountered during field work. Primatologists have access to a rich depth of information on their species that other researchers may lack, as primates are often the focus of countless long-term behavioral projects. Therefore, beyond the usual landscape and environmental variables such as forest coverage, elevation, and climatic variables, researchers can add in the locations of sleeping trees, fruiting trees, mineral licks, or other features of importance to their study species.<sup>51,52</sup> If landscape data are not readily available, researchers can generate layers themselves. It should be noted, however, that the learning curve is steep, and the skills needed to create such surfaces can require extensive training. The handbook edited by Wegmann et al.<sup>80</sup> provides a great starting place, as do intensive workshops offered at various universities or through private institutions. Alternatively, researchers would benefit from collaboration with remote sensing experts who can generate these data on a much shorter time scale and with less frustration than doing so oneself.

Before relating the two datasets, landscape data must first be parameterized (that is, weighted). This involves assigning resistance or cost values to the landscape features in question; these values are meant to reflect the difficulty individuals of a given species have dispersing across each feature.<sup>18</sup> The best *parameterization* methods are currently debated<sup>83</sup>; however, most studies still use expert opinion.<sup>84</sup> In so doing, this method draws on the knowledge of experienced researchers to determine what types of habitat a species tends to avoid and which they prefer. Inevitably, this method suffers from a lack of objectivity and may result in mismatches between the inferred resistance values and the actual resistance experienced by the animal.<sup>83</sup> At present, the most objective and feasible methods include using ecological niche models or genetic algorithms to parameterize resistance surfaces. Ecological niche models (ENMs) predict species distributions based on climatic variables and other landscape data.<sup>83</sup> Although ENMs typically reflect day-to-day movement rather than a species dispersal capacity (but, see Kamilar et al<sup>85</sup>), it has been found that species tend to disperse across landcover most similar to their preferred habitat, making these models a potentially valid method for calculating resistances.<sup>86</sup> More recently, genetic algorithms have been used to parameterize landscape surfaces.87 This method, implemented in the R

## Evolutionary Anthropology\_WILEY

Resulting landscape

spatial resolution

**FIGURE 3** Influence of pixel size on contained landscape information. A 144 m<sup>2</sup> study area could be represented by 144 pixels each 1 m  $\times$  1 m. All the landscape features contained in the 1 m  $\times$  1 m area would then be averaged and represented by one pixel. With larger pixel sizes, the landscape data represented by each pixel becomes coarser, losing information. Each pixel within our hypothetical study area of 144 m<sup>2</sup> could instead have a grain size of 4 m  $\times$  4 m, resulting in just 9 pixels, and a very poor spatial resolution Landscape features with spatial grain overlay



package *ResistanceGA*,<sup>87</sup> creates a workflow that optimizes resistance surfaces based on testing different parameterizations of an input surface using a maximum likelihood method within the framework of genetic algorithms. Using both methods and testing the robusticity of each may be the best way to proceed in a landscape genetics analysis. Other empirical methods for parameterization are reviewed in Zeller et al.<sup>84</sup> and discussed in the Supporting Information.

Once genetic distances and parameterized resistance surfaces have been created, there are numerous correlation methods that can then be used to evaluate their relationships (reviewed in Balkenhol et al<sup>88</sup>; evaluated in Shirk et al<sup>75,76</sup>). Performance of these correlation methods varies based on which landscape variables are included (continuous, categorical, linear), their parameterizations, and the amount of structure within the genetic dataset.<sup>75,89</sup> Based on a recent simulation study, linear mixed effects models are considered the most robust method for correlating genetic and landscape datasets under a variety of scenarios.<sup>75</sup> Researchers have historically used Mantel and partial Mantel tests to correlate datasets, however Balkenhol et al.<sup>88</sup> found that Mantel tests cannot reliably distinguish between isolation-by-distance and true signatures of landscape resistance and should therefore be avoided. See the Supporting Information for a discussion of various software packages used for such analyses.

Most PLG studies to date have incorporated a variety of landscape features (Table 1), though common features included vegetation, roads, and rivers. These types of landscape features are often classified from imagery-based remote sensors like Landsat images from NASA (Table 2). For those PLG studies that classified their own imagery, some included accuracy checks through comparison to reference points<sup>24,28,31,32</sup> or through computational checks.<sup>23</sup> Many PLG studies also incorporated other sources of remotely sensed data, including climatic variables, tourism rates, and fire density (Table 1).

Most PLG studies parameterized data using expert opinion (Table 2), though some employed the genetic algorithms approach described above.<sup>19,34</sup> Some PLG studies used ecological niche modeling (ENM) for parameterization. For example, Mitchell et al. developed an ENM for chimpanzees in Maxent<sup>90</sup> to determine the correct weighting for different habitat types, with low ENM predictions signaling a high cost to traversing that habitat type.<sup>37</sup> Moreover, while various statistical methods were employed (for example, generalized linear models, multiple regression, linear mixed effects models, and generalized dissimilarity modeling; Table 2), the most common statistical method used was Mantel tests and partial Mantel tests (Table 2). As stated above the use of these tests has been highly criticized and their use is not recommended for future studies.<sup>14,91</sup>

12

# <sup>2</sup> WILEY **Evolutionary Anthropology**

### 5 | FUTURE DIRECTIONS

As landscape genetics incorporates methods from both population genetics and landscape ecology, the field has the ability to answer several open questions about the role of ecology in the process of evolution. While the vast majority of primate studies have centered around conservation issues, there are a wide range of questions that can be answered using landscape genetics. These questions range from basic evolutionary processes, such as the role of gene flow in the evolution of populations,<sup>15</sup> to more specific questions relating to the role of the environment in the spread of infectious disease.92 Besides identifying barriers to gene flow, landscape genetics can also help to understand the role of spatial and temporal scale in ecological processes.<sup>16</sup> For example, the impact of historical demography on current patterns of genetic diversity and the role of climatic refugia can be elucidated.<sup>20</sup> This topic ties into analyses of animal movement and migration as well as source-sink dynamics, a function of population stability and viability.<sup>16</sup> While these types of topics have been addressed in landscape genetics in general, these topics have not been fully developed in primate landscape genetic studies and could be a fruitful avenue of future research.

Furthermore, many unresolved questions relate to conservation issues, such as assessing the impact of anthropogenic activities and features on genetic connectivity, the impact of invasive species on local ecology,<sup>92</sup> and the impacts of climate change on both current and projected gene flow patterns for the future.<sup>14</sup> Primates are of particular conservation concern as many species are Endangered, Threat-ened, or susceptible to climate change.

With evermore sophisticated methods becoming increasingly available, landscape genetics has also begun including genomic data built on inclusion of hundreds of thousands of single nucleotide polymorphisms derived from next-generation sequencing and focusing on questions related to adaptation and selection.<sup>64</sup> Storfer et al.<sup>65</sup> provide an excellent overview of how driving questions, sampling design, and analytical techniques differ when scaling up from landscape genetics to genomics. One major difference between landscape genetics and genomics revolves around the inclusion of paired populations used in landscape genomics to detect signatures of selection. This of course requires changes to sampling design and other project decisions that need to be made before data collection begins. These paired comparisons are not only beneficial for identifying signatures of selection but can also help to elucidate different patterns that populations of the same species may have in response to the same landscape feature(s).

Landscape genetics approaches have also started to incorporate graph theory which borrows from network analyses (such as internet and server hubs) to understand relatedness in terms of connections and clustering.<sup>93</sup> Insights drawn from these types of analyses can guide researchers toward understanding which groups or populations are most influenced by restricted genetic connectivity, and can help researchers determine which groups to prioritize for conservation management.

While complex, landscape genetics is a tool for understanding the influence of landscape features on genetic patterning and has the ability to open up interesting avenues of research and evaluate longstanding questions in new ways. Landscape genetics can help primatologists understand how landscape features influence their study species, whether their questions are geared toward understanding ecological impacts, local adaptations, or conservation management.

### DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

#### ORCID

Andrea L. Baden D https://orcid.org/0000-0002-4722-0532

### REFERENCES

- Excoffier L, Heckel G. 2006. Computer programs for population genetics data analysis: a survival guide. Nat Rev Genet 7: 745-758.
- [2] Liu Z, Ren BP, Wei FW, et al. 2007. Phylogeography and population structure of the Yunnan snub-nosed monkey (*Rhinopithecus bieti*) inferred from mitochondrial control region DNA sequence analysis. Mol Ecol 16:3334–3349.
- [3] Zinner D, Groenveld LF, Keller C, et al. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.)–Indication for introgressive hybridization? BMC Evol Biol 9:83.
- [4] Ruiz-García M, Sánchez-Castillo S, Castillo MI, et al. 2018. How many species, taxa, or lineages of *Cebus albifrons* (Platyrrhini, primates) inhabit Ecuador? Insights from Mitogenomics. Int J Primatol 39: 1068–1104.
- [5] Alvarado-Serrano DF, Knowles LL. 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. Mol Ecol Resour 14:233–248.
- [6] Csilléry K, Blum MGB, Gaggiotti OE, et al. 2010. Approximate Bayesian computation (ABC) in practice. Trends Ecol Evol 25:410–418.
- [7] Metzker ML. 2010. Sequencing technologies—the next generation. Nat Rev Genet 11:31–46.
- [8] Prugnolle F, De Meeus T. 2002. Inferring sex-biased dispersal from population genetic tools: a review. Heredity 88:161–165.
- [9] Prunier J, Laroche J, Beaulieu J, et al. 2011. Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. Mol Ecol 20:1702–1716.
- [10] Dirzo R, Young HS, Galetti M, et al. 2014. Defaunation in the Anthropocene. Science 345:401–406.
- [11] Caballero A, Rodríguez-Ramilo ST, Ávila V, et al. 2010. Management of genetic diversity of subdivided populations in conservation programmes. Conserv Genet 11:409–419.
- [12] Richardson JL, Brady SP, Wang IJ, et al. 2016. Navigating the pitfalls and promise of landscape genetics. Mol Ecol 25:849–863.
- [13] Manel S, Schwartz MK, Luikart G, et al. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol 18:189–197.
- [14] Manel S, Holderegger R. 2013. Ten years of landscape genetics. Trends Ecol Evol 28:614–621.
- [15] Holderegger R, Wagner HH. 2008. Landscape genetics. Bioscience 58:199–207.
- [16] Storfer A, Murphy MA, Evans JS, et al. 2007. Putting the "landscape" in landscape genetics. Heredity 98:128–142.
- [17] Marrotte RR, Bowman J. 2017. The relationship between least-cost and resistance distance. PLoS ONE 12:e0174212.
- [18] Shah VB, McRae BH. 2008. Circuitscape: a tool for landscape ecology. In: Varoquaux G et al., editors. Proc. 7th Python Sci. Conf. p 62–66. Pasadena, CA: SciPy.

# Evolutionary Anthropology\_WILEY

- [19] Baden AL, Mancini AN, Federman S, et al. 2019. Anthropogenic pressures drive population genetic structuring across a critically endangered lemur species range. Sci Rep 9:16276.
- [20] Epps CW, Keyghobadi N. 2015. Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. Mol Ecol 24:6021–6040.
- [21] Wang IJ. 2010. Recognizing the temporal distinctions between landscape genetics and phylogeography. Mol Ecol 19:2605–2608.
- [22] Radespiel U, Rakotondravony R, Chikhi L. 2008. Natural and anthropogenic determinants of genetic structure in the largest remaining population of the endangered golden-brown mouse lemur, *Microcebus ravelobensis*. Am J Primatol 70:860–870.
- [23] Crouau-Roy B, Rabarivola C, et al. 2010. Landscape genetics of an endangered lemur (*Propithecus tattersalli*) within its entire fragmented range. Mol Ecol 19:1606–1621.
- [24] Blair ME, Melnick DJ. 2012. Scale-dependent effects of a heterogeneous landscape on genetic differentiation in the central American squirrel monkey (*Saimiri oerstedii*). PLoS ONE 7:e43027.
- [25] Di Fiore A, Valencia LM. 2014. The interplay of landscape features and social system on the genetic structure of a primate population: an agentbased simulation study using "Tamarins.". Int J Primatol 35:226–257.
- [26] Moraes AM, Ruiz-Miranda CR, Galetti PM Jr, et al. 2018. Landscape resistance influences effective dispersal of endangered golden lion tamarins within the Atlantic Forest. Biol Conserv 224:178–187.
- [27] Oklander LI, Miño CI, Fernández G, et al. 2017. Genetic structure in the southernmost populations of black-and-gold howler monkeys (Alouatta caraya) and its conservation implications. PLoS One 12:e0185867.
- [28] Lane-deGraaf KE, Fuentes A, Hollocher H, et al. 2014. Landscape genetics reveal fine-scale boundaries in Island populations of Indonesian long-tailed macaques. Landsc Ecol 29:1505–1519.
- [29] Coetzer WG, Turner TR, Schmitt CA, et al. 2018. Adaptive genetic variation at three loci in south African vervet monkeys (*Chlorocebus pygerythrus*) and the role of selection within primates. PeerJ 6:e4953.
- [30] Liu Z, Ren B, Wu R, et al. 2009. The effect of landscape features on population genetic structure in Yunnan snub-nosed monkeys (*Rhinopithecus bieti*) implies an anthropogenic genetic discontinuity. Mol Ecol 18:3831–3846.
- [31] Li W, Clauzel C, Dai Y, et al. 2017. Improving landscape connectivity for the Yunnan snub-nosed monkey through cropland reforestation using graph theory. J Nat Conserv 38:46–55.
- [32] Li L, Xue Y, Wu G, et al. 2015. Potential habitat corridors and restoration areas for the black-and-white snub-nosed monkey *Rhinopithecus bieti* in Yunnan, China. ORYX 49:719–726.
- [33] Zhao X, Ren B, Li D, et al. 2019. Effects of habitat fragmentation and human disturbance on the population dynamics of the Yunnan snubnosed monkey from 1994 to 2016. PeerJ 7:e6633.
- [34] Ruiz-Lopez MJ, Barelli C, Rovero F, et al. 2016. A novel landscape genetic approach demonstrates the effects of human disturbance on the Udzungwa red colobus monkey (*Procolobus gordonorum*). Heredity 116:167–176.
- [35] Wang W, Qiao Y, Li S, et al. 2017. Low genetic diversity and strong population structure shaped by anthropogenic habitat fragmentation in a critically endangered primate, *Trachypithecus leucocephalus*. Heredity 118:542–553.
- [36] Goossens B, Chikhi L, Jalil MF, et al. 2005. Patterns of genetic diversity and migration in increasingly fragmented and declining orangutan (*Pongo pygmaeus*) populations from Sabah, Malaysia. Mol Ecol 14:441–456.
- [37] Mitchell MW, Locatelli S, Sesink Clee PR, et al. 2015. Environmental variation and rivers govern the structure of chimpanzee genetic diversity in a biodiversity hotspot. BMC Evol Biol 15:1.
- [38] Storfer A, Murphy MA, Spear SF, et al. 2010. Landscape genetics: where are we now? Mol Ecol 19:3496–3514.
- [39] Baden AL, Holmes SM, Johnson SE, et al. 2014. Species-level view of population structure and gene flow for a critically endangered primate (*Varecia variegata*). Ecol Evol 4:2675–2692.

- [40] Schwartz MK, McKelvey KS. 2009. Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conserv Genet 10:441–452.
- [41] Lawson Handley LJ, Perrin N. 2007. Advances in our understanding of mammalian sex-biased dispersal. Mol Ecol 16:1559–1578.
- [42] Graves T, Chandler RB, Royle JA, et al. 2014. Estimating landscape resistance to dispersal. Landsc Ecol 29:1201–1211.
- [43] Landguth EL, Cushman SA, Schwartz MK, et al. 2010. Quantifying the lag time to detect barriers in landscape genetics. Mol Ecol 19: 4179-4191.
- [44] IUCN. 2019. IUCN 2019. IUCN Red List Threat Species 2019-3.
- [45] Fernandez-Duque M, Chapman CA, Glander KE, et al. 2018. Darting primates: steps toward procedural and reporting standards. Int J Primatol 39:1009–1016.
- [46] Oyler-McCance SJ, Fedy BC, Landguth EL, et al. 2013. Sample design effects in landscape genetics. Conserv Genet 14:275–285.
- [47] Hall LA, Beissinger SR. 2014. A practical toolbox for design and analysis of landscape genetics studies. Landsc Ecol 29:1487–1504.
- [48] Lindenmayer DB, Likens GE, Andersen A, et al. 2012. Value of longterm ecological studies. Austral Ecol 37:745–757.
- [49] Kappeler PM, Watts DP. 2012. Kappeler PM, Watts DP, editors. Long-term field studies of primates, Berlin: Springer-Verlag Berlin Heidelberg.
- [50] Draheim HM, Moore JA, Fortin MJ, et al. 2018. Beyond the snapshot: landscape genetic analysis of time series data reveal responses of American black bears to landscape change. Evol Appl 11: 1219–1230.
- [51] Ampeng A, Shukor MN, Sahibin AR, et al. 2016. Patterns of mineral lick use by Northwest Bornean orangutans (*Pongo pygmaeus pygmaeus*) in the Lanjak Entimau Wildlife Sanctuary, Sarawak, Malaysia. Eur J Wildl Res 62:147–150.
- [52] Alexander C, Korstjens AH, Hankinson E, et al. 2018. Locating emergent trees in a tropical rainforest using data from an unmanned aerial vehicle (UAV). Int J Appl Earth Obs Geoinf 72:86–90.
- [53] Murphy MA, Evans JS, Cushman SA, et al. 2008. Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. Ecography 31: 685–697.
- [54] Bowman J, Jaeger JAG, Fahrig L. 2002. Dispersal distance of mammals is proportional to home range size. Ecology 83:2049–2055.
- [55] Kamm U, Rotach P, Gugerli F, et al. 2009. Frequent long-distance gene flow in a rare temperate forest tree (*Sorbus domestica*) at the landscape scale. Heredity 103:476–482.
- [56] Paetkau D, Slade R, Burden M, et al. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol 13: 55–65.
- [57] Piry S, Alapetite A, Cornuet JM, et al. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. J Hered 95:536–539.
- [58] Landguth EL, Fedy BC, Oyler-McCance SJ, et al. 2012. Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. Mol Ecol Resour 12:276–284.
- [59] Aguirre-Liguori JA, Luna-Sánchez JA, Gasca-Pineda J, et al. 2020. Evaluation of the minimum sampling design for population genomic and microsatellite studies. An analysis based on wild maize. Frontiers in genetics 11:870. https://doi.org/10.3389/fgene.2020.00870.
- [60] Chiou KL, Bergey CM. 2018. Methylation-based enrichment facilitates low-cost, noninvasive genomic scale sequencing of populations from feces. Sci Rep 8:1975.
- [61] Puritz JB, Matz MV, Toonen RJ, et al. 2014. Demystifying the RAD fad. Mol Ecol 23:5937–5942.
- [62] Stetz JB, Smith S, Sawaya MA, et al. 2016. Discovery of 20,000 RAD-SNPs and development of a 52-SNP array for monitoring river otters. Conserv Genet Resour 8:299–302.

# 14 WILEY Evolutionary Anthropology

- [63] Helyar SJ, Hemmer-Hansen J, Bekkevold D, et al. 2011. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Mol Ecol Resour 11:123–136.
- [64] Li Y, Zhang XX, Mao RL, et al. 2017. Ten years of landscape genomics: challenges and opportunities. Front Plant Sci 8:2136.
- [65] Storfer A, Patton A, Fraik AK. 2018. Navigating the Interface between landscape genetics and landscape genomics. Front Genet 9:68.
- [66] Morin PA, Foote AD, Hill CM, et al. 2018. SNP discovery from single and multiplex genome assemblies of non-model organisms. Methods Mol. Biol. 1712:113–144.
- [67] Shafer ABA, Peart CR, Tusso S, et al. 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. Gilbert M, editor. Methods Ecol Evol 8:907–917.
- [68] Koepfli K-P, Paten B, the Genome 10K Community of Scientists, et al. 2015. The genome 10K project: a way forward. Annu Rev Anim Biosci 3:57–111.
- [69] Orkin JD, de Manuel M, Krawetz R, et al. 2018. Unbiased whole genomes from mammalian feces using fluorescence-activated cell sorting. bioRxiv 366112. https://doi.org/10.1101/366112.
- [70] Escoda L, Fernández-González Á, Castresana J. 2019. Quantitative analysis of connectivity in populations of a semi-aquatic mammal using kinship categories and network assortativity. Mol Ecol Resour 19:310–326.
- [71] Chiou KL, Bergey CM, Burrell AS, et al. 2019. Genomic signatures of extreme body size divergence in baboons. bioRxiv . https://doi.org/ 10.1101/578740.
- [72] Cushman SA, Shirk AJ, Howe GT, et al. 2018. The least cost path from landscape genetics to landscape genomics: challenges and opportunities to explore NGS data in a spatially explicit context. Front Genet 9:215.
- [73] Manel S, Gaggiotti OE, Waples RS, et al. 2005. Assignment methods: matching biological questions with appropriate techniques. Trends Ecol Evol 20:136–142.
- [74] Segelbacher G, Cushman SA, Epperson BK, et al. 2010. Applications of landscape genetics in conservation biology: concepts and challenges. Conserv Genet 11:375–385.
- [75] Shirk AJ, Landguth EL, Cushman SA. 2017. A comparison of individual-based genetic distance metrics for landscape genetics. Mol Ecol Resour 17:1308–1317.
- [76] Shirk AJ, Landguth EL, Cushman SA. 2018. A comparison of regression methods for model selection in individual-based landscape genetic analysis. Mol Ecol Resour 18:55–67.
- [77] Puechmaille SJ. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Mol Ecol Resour 16:608–627.
- [78] Jones AG, Small CM, Paczolt KA, et al. 2010. A practical guide to methods of parentage analysis. Mol Ecol Resour 10:6–30.
- [79] Horning N, Robinson JA, Sterling EJ, et al. 2010. Remote sensing for ecology and conservation: a handbook of techniques, Oxford: Oxford University Press.
- [80] Wegmann M, Leutner B, Dech S, et al. 2016. Remote sensing and GIS for ecologists using open source software. Exeter: Pelagic Publishing.
- [81] Hansen MC, Potapov PV, Moore R, et al. 2013. High-resolution global maps of 21st-century forest cover change. Science 342:850–853.
- [82] Clark ML, Clark DB, Roberts DA. 2004. Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. Remote Sens Environ 91:68–89.
- [83] Milanesi P, Holderegger R, Caniglia R, et al. 2017. Expert-based versus habitat-suitability models to develop resistance surfaces in landscape genetics. Oecologia 183:67–79.
- [84] Zeller KA, McGarigal K, Whiteley AR. 2012. Estimating landscape resistance to movement: a review. Landsc Ecol 27:777–797.

- [85] Kamilar JM, Blanco M, Muldoon KM. 2016. Lehman SM, Radespiel U, Zimmermann E, editors. Ecological niche modeling of mouse lemurs and its implications for their species diversity and biogeography. Dwarf and mouse lemurs of Madagascar: biology, behavior and conservation biogeography of the Cheirogaleidae, Cambridge: Cambridge University Press. p 449-461.
- [86] Eycott AE, Stewart GB, Buyung-Ali LM, et al. 2012. A meta-analysis on the impact of different matrix structures on species movement rates. Landsc Ecol 27:1263–1278.
- [87] Peterman WE. 2018. ResistanceGA: an R package for the optimization of resistance surfaces using genetic algorithms. Methods Ecol Evol 9:1–10.
- [88] Balkenhol N, Waits LP, Dezzani RJ. 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32:818–830.
- [89] Kierepka EM, Latch EK. 2015. Performance of partial statistics in individual-based landscape genetics. Mol Ecol Resour 15: 512–525.
- [90] Phillips SJ, et al. 2004. A maximum entropy approach to species distribution modeling. In: Proc. Twenty-First Int. Conf. Mach. Learn. p 655–662.
- [91] Epperson BK, McRae BH, Scribner KIM, et al. 2010. Utility of computer simulations in landscape genetics. Mol Ecol 19:3549–3564.
- [92] Montgelard C, Zenboudji S, Ferchaud AL, et al. 2014. Landscape genetics in mammals. Mammalia 78:139–157.
- [93] Locher A, Scribner KT, Moore JA, et al. 2015. Influence of landscape features on spatial genetic structure of white-tailed deer in humanaltered landscapes. J Wildl Manage 79:180–194.

#### **AUTHOR BIOGRAPHIES**

**Darice Westphal** is a PhD candidate at The Graduate Center (CUNY). She studies landscape genetics in mouse lemurs and deforestation trends in Madagascar.

**Amanda N. Mancini** is a PhD candidate at The Graduate Center (CUNY). Her interests are in spatial ecology, movement ecology, and landscape genetics, with a focus on Madagascar's lemurs.

Andrea L. Baden is an Associate Professor of Anthropology at Hunter College (CUNY) and Director of the Hunter Primate Molecular Ecology Lab (www.HunterPMEL.org). At present, her work focuses on two main avenues of inquiry: How landscape features—past, present, and future—influence distribution, migration, and genetic diversity in wild primate populations; and how the resulting patterns of genetic relatedness within populations and groups can help to understand the evolution and maintenance of complex social behaviors.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Westphal D, Mancini AN, Baden AL. Primate landscape genetics: A review and practical guide. *Evolutionary Anthropology*. 2021;1–14. <u>https://doi.org/10.</u> 1002/evan.21891