BMC
Evolutionary Biology

**RESEARCH ARTICLE**                                                                 **Open Access**

# Environmental variation *and* rivers govern the structure of chimpanzee genetic diversity in a biodiversity hotspot

Matthew W Mitchell[1,2*†], Sabrina Locatelli[2,3†], Paul R Sesink Clee[1,2], Henri A Thomassen[4] and Mary Katherine Gonder[1,2*]

## Abstract

**Background:** The mechanisms that underlie the diversification of tropical animals remain poorly understood, but new approaches that combine geo-spatial modeling with spatially explicit genetic data are providing fresh insights on this topic. Data about the diversification of tropical mammals remain particularly sparse, and vanishingly few opportunities exist to study endangered large mammals that increasingly exist only in isolated pockets. The chimpanzees of Cameroon represent a unique opportunity to examine the mechanisms that promote genetic differentiation in tropical mammals because the region is home to two chimpanzee subspecies: *Pan troglodytes ellioti* and *P. t. trogolodytes*. Their ranges converge in central Cameroon, which is a geographically, climatically and environmentally complex region that presents an unparalleled opportunity to examine the roles of rivers and/or environmental variation in influencing the evolution of chimpanzee populations.

**Results:** We analyzed microsatellite genotypes and mtDNA HVRI sequencing data from wild chimpanzees sampled at a fine geographic scale across Cameroon and eastern Nigeria using a spatially explicit approach based upon Generalized Dissimilarity Modeling. Both the Sanaga River and environmental variation were found to contribute to driving separation of the subspecies. The importance of environmental variation differed among subspecies. Gene-environment associations were weak in *P. t. troglodytes,* whereas environmental variation was found to play a much larger role in shaping patterns of genetic differentiation in *P. t. ellioti.*

**Conclusions:** We found that both the Sanaga River and environmental variation likely play a role in shaping patterns of chimpanzee genetic diversity. Future studies using single nucleotide polymorphism (SNP) data are necessary to further understand how rivers and environmental variation contribute to shaping patterns of genetic variation in chimpanzees.

## Background

The Gulf of Guinea and the Congo River Basin biomes (Figure 1) collectively house ~20% of all plants and animals species, with a very high number of endemic taxa [1]. However, very little is known about the mechanisms that have created this region's rich biodiversity. Understanding how local factors contributed to generating patterns of genetic differentiation is important to better understand the evolutionary history of tropical taxa. These patterns of genetic differentiation in tropical taxa, and particularly in forest-dwelling primates, may have been shaped by forest history during the Pleistocene, as well as by geographic barriers to dispersal, including rivers and pronounced environmental gradients [2-4].

Tropical Africa is also noted for exhibiting a complex forest history that has undergone considerable change along with the Earth's climate, and the behavioral ecology of African primates often reflects a shared history with their habitats [5-7]. This relationship has been proposed to have influenced the diversification of rainforest taxa due to the fact that previously continuous populations were isolated from one another in forest "refuges" that persisted during periods of maximum glaciation
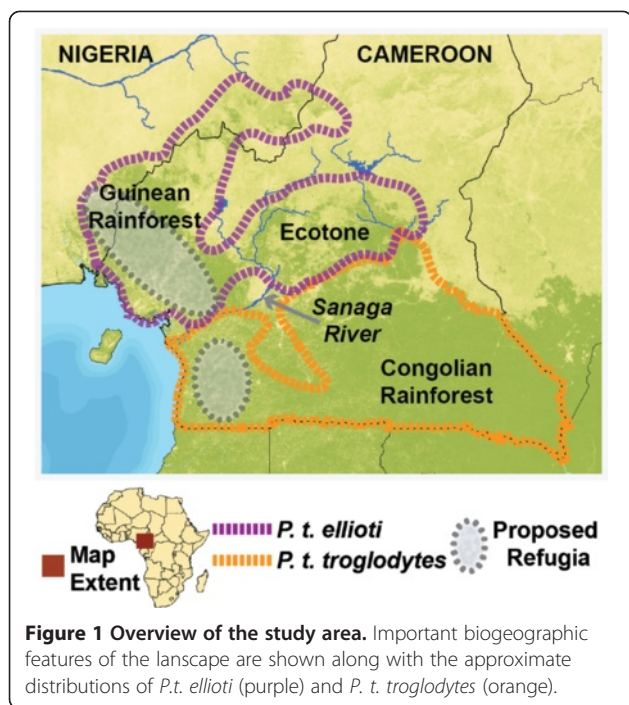
* Correspondence: mwmitchell@drexel.edu; gonder@drexel.edu
†Equal contributors
[1]Department of Biology, Drexel University, Philadelphia, Pennsylvania 19104, USA
[2]Department of Biological Sciences, University at Albany – State University of New York, Albany, New York 12222, USA
Full list of author information is available at the end of the article

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 2 of 13



**Figure 1 Overview of the study area.** Important biogeographic features of the lanscape are shown along with the approximate distributions of *P.t. ellioti* (purple) and *P. t. troglodytes* (orange).

[5-8]. One of the most widely cited examples to support the 'Pleistocene Refuge Hypothesis' are the forests of western Africa, which are divided into two regions located in Upper Guinea and the Gulf of Guinea. Each forest region has its own species assemblages, which are often used to point to the existence of forest refuges in Africa [9]. The Dahomey Gap currently separates these two forest regions and their associated taxa. The Gap is a large, dry, open area that presently extends for about 700 km through Benin, Togo and eastern Ghana [10], but it may have been up to 1,400 km wide during the cool, arid phases of the Pleistocene [3]. Consequently, the Gap has been proposed to have been a barrier to dispersal for several species which resulted in the formation of new species. This process of separation and speciation across the Gap has been proposed to account for the unique faunal assemblages of the Upper and Gulf of Guinea forest blocks, which are both two of the world's original Global Biodiversity Hotspots [2,4,11,12].

Rivers have also been proposed to play an important role in delimiting the distribution of many species [2,3,10]. For example, the Sanaga River in central Cameroon (Figure 1) has been proposed to delimit the distributions of several pairs of primate species and subspecies: *Mandrillus leucophaeus/M. sphinx, Cercopithecus erythrotis/C. cephus, C. nictitans martini/C. n. nictitans, C. pogonias pogonias/C. p. grayi, P. t. ellioti/P. t. troglodytes* and possibly, *Gorilla gorilla diehli/G. g. gorilla* [1,13-18]. The Congo separates chimpanzees (*Pan troglodytes*) from bonobos (*P. paniscus*) [15], and the Ubangi, Niger and Sanaga Rivers in central Africa may be important in delimiting chimpanzee

subspecies from one another. However, the role of riverine barriers in shaping patterns of diversity across the landscape are poorly understood [4,19]. Changes in river size and course over time may dramatically affect the significance of a given river to act as a barrier to dispersal, and habitat changes in the vicinities of rivers may confound distiguishing between the role of the river and changes in the environment [4]. The Sanaga River, for instance, has also been proposed to act as an historical boundary that separates *P. t. ellioti* and *P. t. troglodytes* [20-23], but its significance is questionable [23]. Little is known about the history of the Sanaga, but it is entirely possible that its course and size have changed over time, especially given this region's sensitivity to climatic oscillations [9,24]. And although chimpanzee genetic populations are seemingly partitioned along the location of the Sanaga, there is evidence of gene flow occurring between these populations [22,23,25].

Until very recently, the role of ecological gradients in driving the diversification of tropical taxa has been under-appreciated. Ecological gradients are zones of transition between habitat types that display marked differences in ecological variables (i.e. precipitation, temperature, vegetation density, etc.) across their range. These gradients have been linked with driving adaptive variation in several taxa in different parts of the world, for example, suture zones in Australia [26] and more recently in Cameroon [27]. In this ecological gradient model, speciation across habitat gradients is driven by local adaptation, whereas the genetic differentiation of allopatric populations is driven by genetic drift resulting from historical isolation in refugia or separation by geographic barriers, including rivers and other features of the landscape, while the habitats remain generally the same [28]. A prominent gradient is present in central Cameroon, which transitions from Guinean forest in the west to Congolian rainforest in the south and to Sahelian habitats in the north and east [29] (Figure 1). This forest-savanna mosaic was termed by Smith *et al.* [30] as an 'ecotone'. A growing body of evidence suggests that this ecotone has been important in promoting the evolutionary diversification of insects [31], reptiles [27] and birds [30,32]. Complementary genetic datasets for mammals occupying this region remain sparse, which makes it difficult to distinguish between the relative influence of environmental and topographic factors in governing diversification and population structuring of mammals, particularly primates where distributional data are generally the only information available.

Recent analyses suggest that the two subspecies of chimpanzees present in Cameroon may be divided into genetically distinct populations: *P. t. troglodytes* occurs south of the Sanaga River, while *P. t. ellioti* occurs north.

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 3 of 13

*P. t. ellioti* may be further subdivided into two additional populations: a population that occurs in forested regions of western Cameroon, *P. t. ellioti* (Rainforest), and a second population that occupies central Cameroon, *P. t. ellioti* (Ecotone) [23]. While it appears that the Sanaga River has played an important role in separating these subspecies, the estimated ranges of these populations coincide with the transition of the two rain forests types and the ecotone [23], and all three populations occupy significantly different habitat types [33]. This suggests that a relationship exists between the environmental variation and the partitioning of genetic variation in chimpanzees found across the study area. Taken together, these factors make it difficult to distinguish between the relative importance of the Sanaga River or habitat variation in shaping the partitioning of chimpanzee population genetic variation given their close proximity to one another (Figure 1).

The available genetic data for wild chimpanzees sampled across this region consist of 21 autosomal microsatellite loci and the HVRI region of the mtDNA [23]. Since these loci meet expectations of neutrality, the genetic data set cannot be used to directly examine the role of adaptation in shaping patterns of variation in chimpanzees across the study area [27,34]. However, it is still possible to use these neutrally-evolving genetic makers to infer whether environmental factors other than, or in addition to, the Sanaga have contributed to shaping the patterning of genetic variation found in chimpanzees from the region, indicating that chimpanzees in Cameroon and Nigeria may follow a pattern of isolation-by-environment, a relationship driven by selective evolutionary processes [35].

Table 1 lists expected patterns of genetic variation depending upon whether isolation across the Sanaga River or environmental variation occupies a dominant role in shaping patterns of genetic variation in chimpanzees across the study region. Collectively, these predictions are related to: (*i*) the diversity and distribution of alleles; (*ii*) the dates of divergence between populations; (*iii*) the location of barriers to gene flow; and (*iv*) the demographic characteristics of populations (e.g. population size and population growth). This study is an extension of a complimentary paper [23] which looked at the population genetic structure and demographic history (*ii* and *iv*) of chimpanzees in Cameroon and Nigeria. That study found evidence that positive selection likely plays a role in shaping patterns of chimpanzee genetic diversity, and that this may be the results of landscape variation. This study is the next step, and focuses on using a spatially explicit modeling technique, Generalized Dissimilarity Modeling (GDM) [36], to quantify and visualize associations between landscape variation and genetic variation in chimpanzees. The results of the analyses presented here were used to support or reject predictions regarding spatial patterns of genetic diversity (*i*) and the location of barriers to gene flow (*iii*).

## Results and discussion

### Regional patterns of isolation-by-environment

The data set used for analysis was used in a previous study on the population genetic structure of chimpanzees [23]. The genetic data consisted of autosomal microsatellite genotype profiles of 187 unrelated individuals sampled from 28 locations across eastern Nigeria and Cameroon and 604 mtDNA sequences sampled from 35 locations across the study region (Figure 2). These data were subjected to rigorous quality control procedures when generated, including separate DNA extractions and several independent calculations of allele sizes [23]. All loci met expectations of neutral evolution, as measured by the results of an outlier test [37] and meeting expectations of Hardy-Weinberg equilibrium [23]. We ran GDM's using microsatellite genotypes and mtDNA sequencing data in order to assess the relative contribution of spatial variables to patterns of genetic differentiation. This method fits genetic distance matrices from both microsatellite ($F_{ST}$) and mtDNA data against straight-line geographic distance, topographic, climatic and vegetation variables assumed to contribute to chimpanzee habitat ecology and biogeographic boundaries. We ran models using: (*i*) geographic distance between sample locations only; (*ii*) environmental variables only (topography, climate and vegetation); (*iii*) environmental

**Table 1 Models to explain the partitioning of chimpanzee genetic variability across the study region, and associated predictions**

| Predictions | Riverine barriers | Ecological variation |
|---|---|---|
| **Diversity of neutral alleles** | Highest on opposite banks of Sanaga and Mbam Rivers | Significant correlations between the distribution of allelic diversity and variation in one or more environmental variables. |
| **Population divergence (T_{MRCA})** | Broad spectrum of time intervals | Broad spectrum of time intervals |
| **Barriers to gene flow are located** | At the Sanaga, with highest resistance at the Sanaga Delta and decreasing towards headwaters | At or near ecotone boundaries, but not at adjacent Guinean-Congolian rainforest boundary in western Cameroon |
| **Population history includes** | Demographic stability, other scenarios possible | Demographic stability, other scenarios possible |

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 4 of 13

variables and geographic distance between sample locations, and (*iv*) environmental variables, distance and riverine barrier layers ('complete barrier' layers and resistance surfaces). The percent of genetic variation accounted for by the microsatellite genotype data and the mtDNA data for three different classes of predictor variables are shown in Table 2.

Pairwise differences ($F_{ST}$) between sample locations at autosomal microsatellites are shown in Additional file 1. The correlation between $F_{ST}$ and predictor variables was then interpolated across the study area using the GDM method, which revealed that geographic distance alone accounted for 4% and environmental variation alone accounted for 10% of the observed genetic variation (Table 2). A combination of geographic distance and environmental variation only increased the explanatory power of the model to 11% (Table 2), even though distance was a significant contributor to the model. These findings are consistent with the results of from a previous study [23], that revealed that genetic population structure in chimpanzees in the region can only be weakly explained by a pattern of isolation-by-distance. In the model including geographic distance and environmental variables jointly, genetic differentiation was predicted along a general west–east cline from Guinean rainforest to ecotone to Congolian rainforest (Figure 3a). This model was mostly explained by minimum NDVI during the yearly period with the least amount of new vegetation (NDBR) and climatic variables relating to precipitation, and to a lesser extent by slope and geographic distance (Figure 3c).

When we included the Sanaga and Mbam River layers, both rivers significantly contributed to the partitioning of genetic variation but the total variation explained increased to only 13% in that particular model (Table 2). The cost-resistance surface layer of riverine disperal barriers did not significantly contribute to the population structure of chimpanzees across the study region. The pattern of predicted genetic differentiation changed and featured a sharp divide at the Sanaga, a less pronounced, but still obvious divide across the Mbam, and a distinct gradient across elevations in the Cameroon Highlands (Figure 3b). The relative contribution of the environment variables was similar to the previous model, with the Sanaga and Mbam rivers contributing only moderately (Figure 3d).

Pairwise differences between sample locations at the mtDNA HVRI locus are shown in Additional file 2 and Tamura and Nei indices [38] are shown in Additional file 3. The response between pairwise distance and the predictor variables was interpolated across the study area using the GDM method, which revealed that geographic distance alone accounted for 6% and environmental variation alone accounted for 22% of the observed genetic variation (Table 2). The response between Tamura and Nei distance and the predictor variables was also calculated, which revealed that geographic distance alone accounted for 11% and environmental variation alone accounted for 32% of the observed genetic variation (Table 2). Geographic distance did not significantly contribute to the partitioning of genetic variation when either model was run combined with environmental variation. When this analysis was run with only environmental variables, genetic differentiation was predicted along two pronounced clines, one from coastal rainforest to montane forests and ecotone, and another from coastal rainforest and ecotone southward to Congolian rainforest (Figure 4a).

When the Sanaga and Mbam River layers were included, the amount of genetic variation accounted for by the pairwise distance model increased to 56%, while the Tamura and Nei model increased to 72%, the highest percentage for all models across the entire study area included in this study (Table 2). Again, as with the microsatellite data, the cost-resistance surface layer of riverine disperal barriers did not significantly contribute to the model. Genetic differentiation was predicted across a sharp divide at the Sanaga River, with further differentiation occurring along elevational gradients in western Cameroon and eastern Nigeria (Figure 4b). The Sanaga River was by far the most important contributor, with variability in several environmental factors (e.g. surface moisture, temperature, precipitation and diurnal range) moderately contributing to the overall partitioning of genetic diversity (Figure 4d).

The Sanaga River is the most important contributor to the partitioning of mtDNA genetic diversity, and the
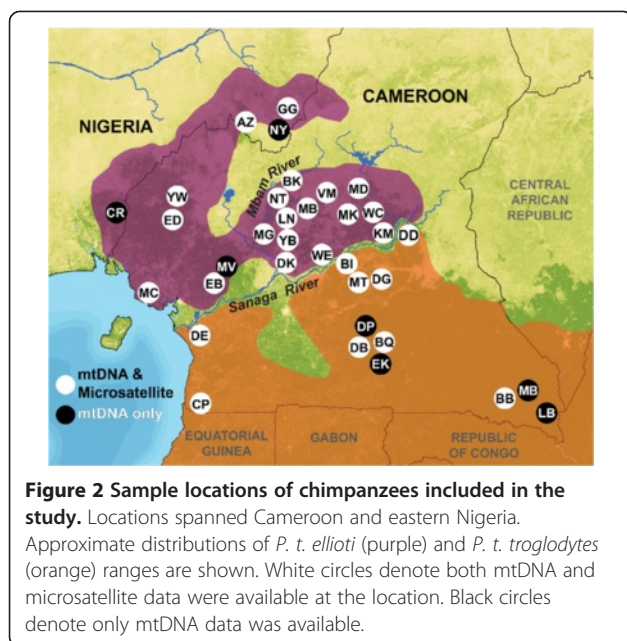
**Table 2 Region wide gene-environment relations**

| Genetic dataset | GDM model | | | |
|---|---|---|---|---|
| | Distance only[a] | Environment only[b] | Environment + Distance[c] | Environment + Distance + Rivers |
| Autosomal Microsatellites ($F_{ST}$) | 4 | 10 | 11 | 13 |
| mtDNA – Pairwise Differences | 6 | 22 | - | 56 |
| mtDNA – Tamura and Nei | 11 | 32 | - | 72 |

[a]Model includes only geographic distance between sample locations as a predictor variable.
[b]Model includes environmental data layers as well as resistance matrices, which incorporate river barriers and habitat suitability, as predictor variables.
For all models, resistance matrices were not included, as they had no significant contribution.
[c]Entries are blank because geographic distance was not a significant contributor to the final model.

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 5 of 13



**Figure 2 Sample locations of chimpanzees included in the study.** Locations spanned Cameroon and eastern Nigeria. Approximate distributions of *P. t. ellioti* (purple) and *P. t. troglodytes* (orange) ranges are shown. White circles denote both mtDNA and microsatellite data were available at the location. Black circles denote only mtDNA data was available.

model using both river layers accounted for the greatest amount of the genetic variation of all the region-wide models (Table 2, Figure 4). This result is unsurprising, especially given the results of a previous study, showing a deep break of mtDNA haplotypes across the Sanaga [23]. The GDMs for microsatellites suggest that rivers play an important, but incomplete, role in separating *P. t. ellioti* and *P. t. troglodytes* from one another. While the Sanaga and Mbam Rivers contributed significantly to the analysis from microsatellite genotypes, these rivers were never the top contributors, and they only slightly increased the overall explanatory power of the models. These GDMs showed the strongest associations with different habitat types (Figure 3). These results, taken together, show that the Sanaga River is indeed a contributor to the partitioning of genetic diversity of chimpanzees in the region. However, given the importance of environmental variables in the microsatellite models, it is apparent that environmental variation also contributes to shaping patterns of genetic diversity in chimpanzees in the region.

### Intra-population patterns of isolation-by-environment

We completed additional GDMs to assess the partitioning of genetic variation among populations within subspecies, by including a two-population grouping and a three-population grouping [33] based on cluster analysis from a previous study [23]. We ran these models only using microsatellite genetic differentiation as estimated by $F_{ST}$. We completed the analysis for the *P. t. ellioti* group using environmental variables, geographic distance, and the Mbam River layer. This model accounted for 20% of the observed genetic variation (Table 3). Neither geographic distance nor the Mbam River significantly contributed to

this model, although a model run with geographic distance alone could explain 6% of genetic variation (Table 3). Genetic differentiation was predicted along a gradual west–east cline from the Guinean Rainforest to the ecotone of central Cameroon (Figure 5a). A variety of environmental variables were important contributors, including NDVI of the least green season, precipitation variables, temperature seasonality and leaf area index (Figure 5c).

We completed the analysis for the *P. t. troglodytes* group using environmental variables and geographic distance as response variables. Geographic distance alone accounted for 12% and environmental variation accounted for 42% of observed genetic variation. Combining geographic distance and environmental variation as predictors increased the explanatory power of the model to 48% (Table 3). Geographic distance was a significant contributor to the overall partitioning of genetic variation found in *P. t. troglodytes*, although a pocket of differentiation occurred in southwest Cameroon that did not appear to follow the clinal model found across the rest of this subspecies' range in southern Cameroon (Figure 5b). In addition to geographic distance, several environmental variables relating to precipitation, temperature and surface moisture (annual mean and variability), were also associated with genetic variation in *P. t. troglodytes* across the study area (Figure 5d).

Environmental variation is strongly correlated with the partitioning of genetic diversity in *P. t. ellioti*. The analysis of the rainforest population of *P. t. ellioti* revealed that environmental variables accounted for 91% of the observed genetic variation (Table 3), particularly differences in vegetation, precipitation and slope (Figure 6d). Geographic distance was not a significant contributor to the model when combined with the environmental variables, and could only account for 2% of the genetic variation when included as the sole predictor. Genetic differentiation was predicted along a gradient that spans coastal to montane forests (Figure 6a), consistent with other analyses that included all of *P. t. ellioti*.

Environmental variables accounted for 37% of the partitioning of genetic variation of the *P. t. ellioti* (Ecotone) population (Table 3). Geographic distance was not a significant contributor when combined with the environmental variables and could only account for 1% of the genetic variation when included as the sole predictor. In addition, genetic differentiation followed a west–east gradient, across a wide breadth of habitats that includes a forest-savanna mosaic that spans central Cameroon (Figure 6b). The most important contributors to influencing the partitioning of genetic variation among chimpanzees located in central Cameroon include precipitation variables, surface moisture and vegetation density (Figure 6e), which are all important variables that define forest and savanna habitats.
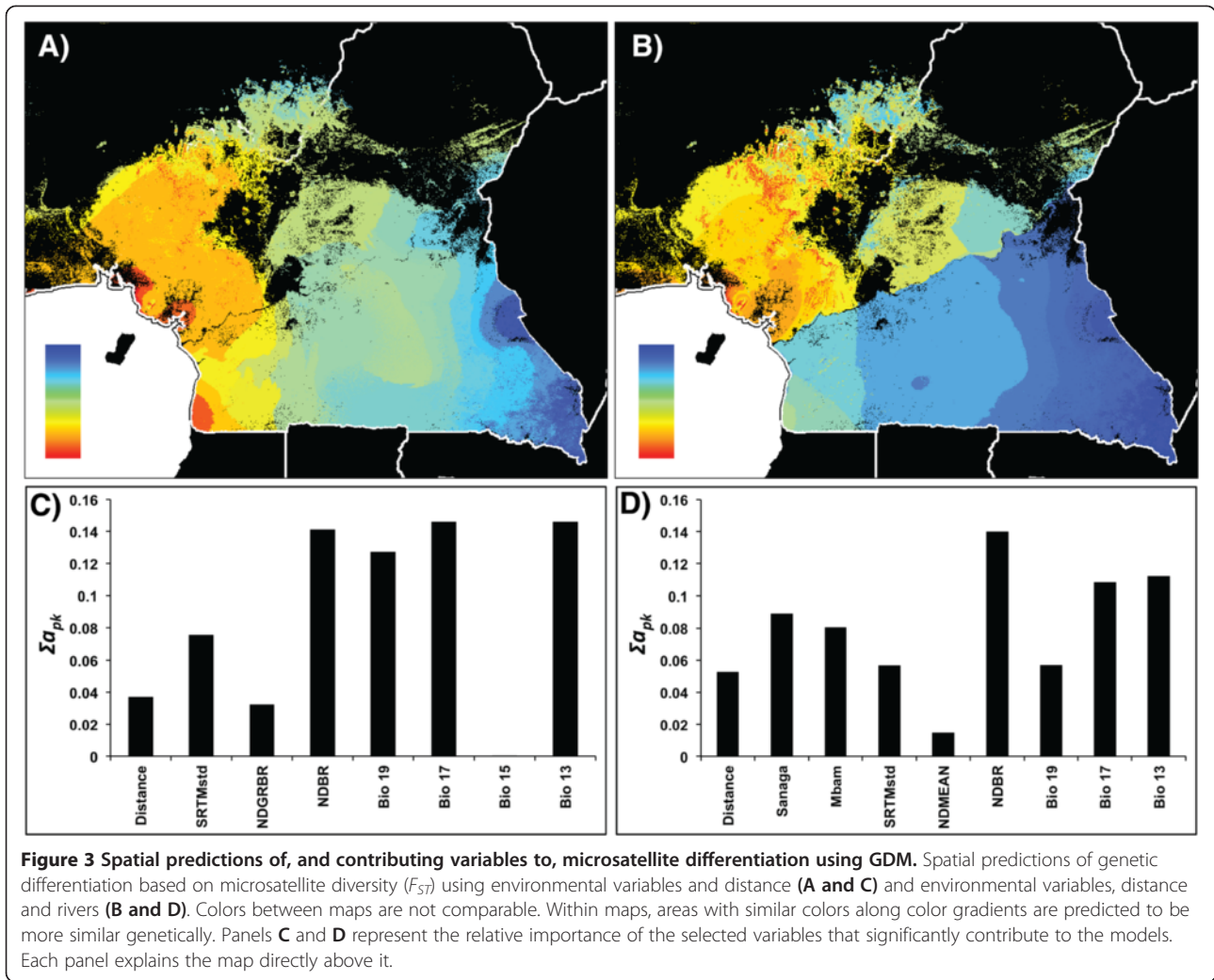
Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 6 of 13



**Figure 3 Spatial predictions of, and contributing variables to, microsatellite differentiation using GDM.** Spatial predictions of genetic differentiation based on microsatellite diversity ($F_{ST}$) using environmental variables and distance **(A and C)** and environmental variables, distance and rivers **(B and D)**. Colors between maps are not comparable. Within maps, areas with similar colors along color gradients are predicted to be more similar genetically. Panels **C** and **D** represent the relative importance of the selected variables that significantly contribute to the models. Each panel explains the map directly above it.

### Table 3 Percent of genetic variation within chimpanzee populations explained by GDM[a]

| Chimpanzee Population[b] | Two-population model | | |
|---|---|---|---|
| | Distance only[c] | Environment only[d] | Environment + Distance[e] |
| *P. t. elliot*[f] | 6 | 20 | - |
| *P. t. troglodytes* | 12 | 42 | 48 |
| | Three-population model | | |
| | Distance only | Environment only | Environment + Distance |
| *P. t. elliot* (Rainforest) | 2 | 91 | - |
| *P. t. elliot* (Ecotone) | 1 | 37 | - |
| *P. t. troglodytes* | 12 | 42 | 48 |

[a]In order to examine genetic diversity within populations, only microsatellite data ($F_{ST}$) was used.
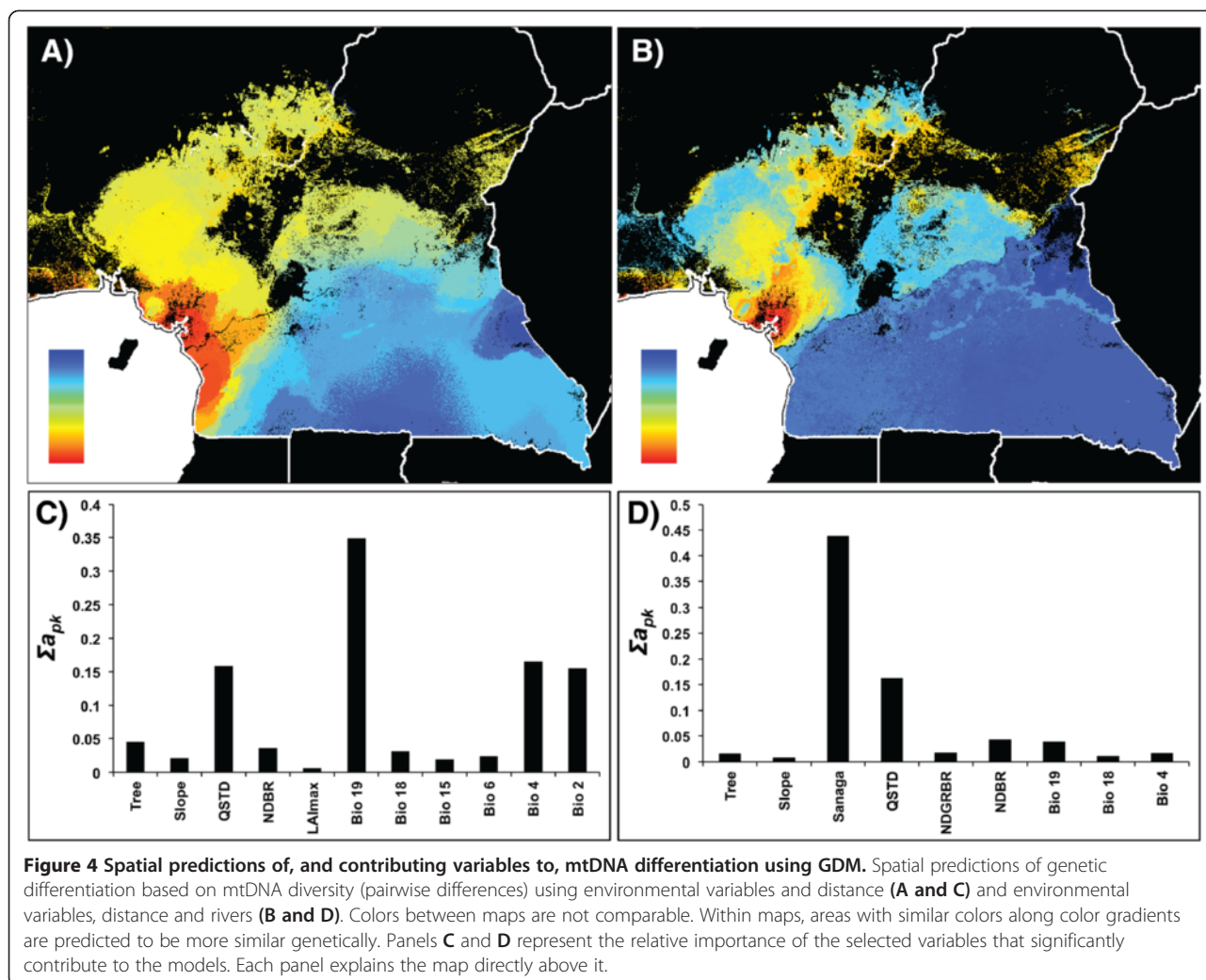[b]Sub-groupings of chimpanzee populations correspond to distinct genetic populations as determined by Mitchell *et al.* [23]. The *P. t. troglodytes* group was included in both the two- and three-population model, as only the *P. t. elliot* group is sub-divided in the three-population model.
[c]Model includes only geographic distance between sample locations as a predictor variable.
[d]Model only includes environmental data layers as predictor variables.
[e]Entries are blank because geographic distance was not a significant contributor to the final model.
[f]A least cost path layer for the Mbam River was included when running the *P. t. elliot* two-population model, but as it was not a significant contributor, was not included in this table.

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 7 of 13



**Figure 4 Spatial predictions of, and contributing variables to, mtDNA differentiation using GDM.** Spatial predictions of genetic differentiation based on mtDNA diversity (pairwise differences) using environmental variables and distance **(A and C)** and environmental variables, distance and rivers **(B and D)**. Colors between maps are not comparable. Within maps, areas with similar colors along color gradients are predicted to be more similar genetically. Panels **C** and **D** represent the relative importance of the selected variables that significantly contribute to the models. Each panel explains the map directly above it.
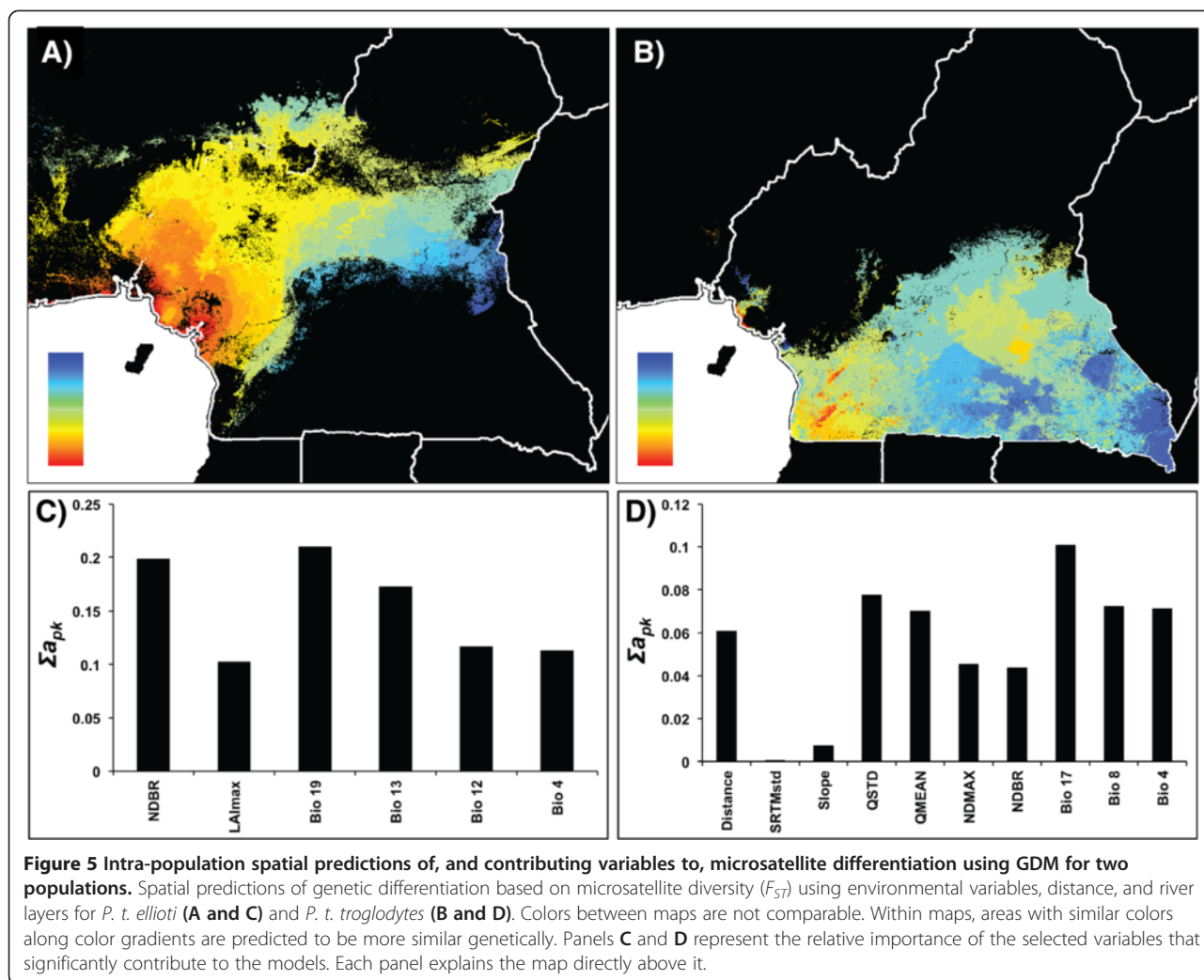
Significant environmentally associated gradients were observed within all tested populations north of the Sanaga River (*P. t. ellioti* and its two sub-groupings). In western Cameroon and eastern Nigeria elevational gradients were once again observed to account for the highest proportion of variation (Figures 5a, and c, 6a and d). In central Cameroon, genetic differentiation was predicted across habitats with variability in vegetation density, moisture and precipitation (Figure 6b and 6e). In addition, *P. t. ellioti* habitats in the forested regions of western Cameroon are characterized by steep, densely forested areas with pronounced precipitation seasonality in northwest Cameroon, and a wider breadth of habitats that include both forest and savanna that experience even more seasonal variability in temperature and precipitation [32,33]. In southern Cameroon, geographic distance was a major predictor, and there was no apparent spatial association between genetic diversity and environmental clines (Figure 5b and 6e). These patterns are unsurprising, especially given that ecological niche models show that *P. t.*

*troglodytes* habitat are relatively homogeneous in terms of elevation, temperature and precipitation [33]. In contrast, habitat variability increases dramatically north of the Sanaga River [27,29,32,33,39].
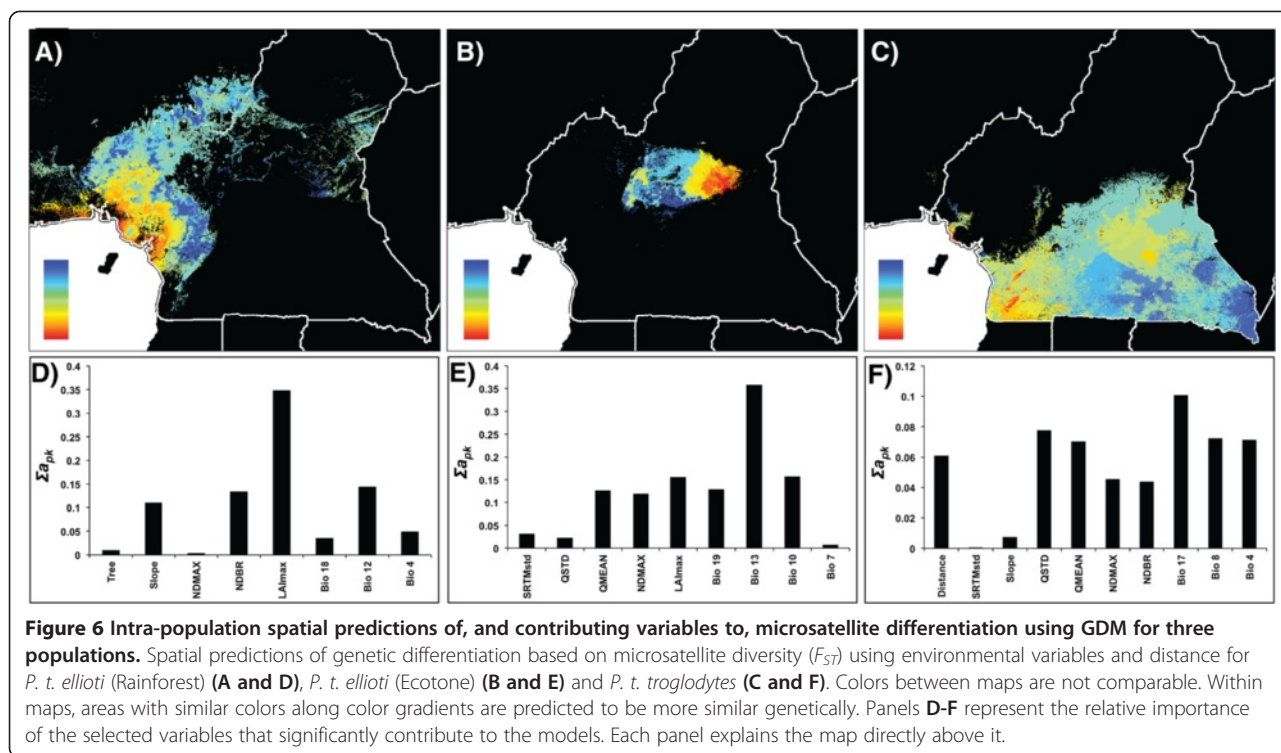
### Hypothesis testing

Using the geospatial modeling approaches presented here, we can provide preliminary assessments of the likely contributors that underlie the partitioning of chimpanzee population genetic diversity: Riverine barriers and environmental variation. Overall, predicted patterns of neutral allele diversity across the study region were found to be relatively consistent with each other in both of the region-wide models (Figures 3 and 4). Without the river layers, the models both showed a pronounced association between allele diversity and habitat type, with a general west–east gradient that might be consistent with differentiation between Guinean Rainforest – ecotone – Congolian Rainforest habitat types.

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 8 of 13



**Figure 5 Intra-population spatial predictions of, and contributing variables to, microsatellite differentiation using GDM for two populations.** Spatial predictions of genetic differentiation based on microsatellite diversity ($F_{ST}$) using environmental variables, distance, and river layers for *P. t. ellioti* (**A and C**) and *P. t. troglodytes* (**B and D**). Colors between maps are not comparable. Within maps, areas with similar colors along color gradients are predicted to be more similar genetically. Panels **C** and **D** represent the relative importance of the selected variables that significantly contribute to the models. Each panel explains the map directly above it.

How do rivers contribute to generating the pattern found in chimpanzees across the study area? Results from previous work show a distinct division of population clusters precisely at the location of the Sanaga River. However, the Sanaga is an incomplete boundary, as there is evidence of migration across the river [12,21-23,25]. Population genetic analysis shows that chimpanzees across the Sanaga River exchange migrants at a rate of approximately 1 migrant per generation, lending even more support to the face that the Sanaga River is not a complete boundary to gene flow [23]. Including the river layers increased the explanatory power of both region-wide models presented in this study (Table 2). The Sanaga River is always an important contributor to patterns of region-wide genetic differentiation, but never the only one (Figures 3 and 4), and in some cases not even the most important (Figure 3). Additionally, the Mbam River was not found to be a significant contributor to genetic differentiation north of the Sanaga River (Figure 6a and c).

How does environmental variation contribute to chimpanzee genetic diversity? Although variation in forest cover, preciptation, and slope are generally more important across the study area than previously appreciated, these variables impact *P. t. troglodytes* and *P. t. ellioti* differently. Amongst *P. t. troglodytes* south of the Sanaga, the distribution of allelic diversity does not appear to be associated with any environmental clines, and instead, appears to follow a pattern of isolation-by-distance (Figures 3 and 4). In contrast, the partitioning of genetic variation in *P. t. ellioti* appears to be substantially influenced by changes in the environment (Figures 3 and 4). Finally, these associations between habitat and the partitioning of genetic variation are especially noticeable when the data for *P. t. elloti* are split into two subpopulations in the western rainforest and central ecotone regions (Figures 5 and 6), in which environmental varation accounts for 91% and 37%, respectively, of the variation explained by the GDM for this subspecies. These observations provide support for the hypothesis

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 9 of 13



**Figure 6 Intra-population spatial predictions of, and contributing variables to, microsatellite differentiation using GDM for three populations.** Spatial predictions of genetic differentiation based on microsatellite diversity ($F_{ST}$) using environmental variables and distance for *P. t. ellioti* (Rainforest) **(A and D)**, *P. t. ellioti* (Ecotone) **(B and E)** and *P. t. troglodytes* **(C and F)**. Colors between maps are not comparable. Within maps, areas with similar colors along color gradients are predicted to be more similar genetically. Panels **D-F** represent the relative importance of the selected variables that significantly contribute to the models. Each panel explains the map directly above it.

that habitats play an important role in structuring chimpanzee populations.

The analyses presented here show, for the first time, that an important relationship exists between the partitioning of genetic variation in chimpanzees and environmental variation in Cameroon and eastern Nigeria, particularly regarding changes in slope, climate and vegetation. Characterizing these relationships is important because it might explain why the separation of chimpanzee subspecies across the Sanaga only partially explains the distribution of genetic variation and that the role of adaptation to local environmental conditions may be substantially underappreciated in the evolution of chimpanzee subspecies. Finally, these findings suggest that the role of environmental variation may be under-appreciated in other primates whose distributions may have been influenced by the Sanaga River [13,14]. Future studies that examine patterns of isolation-by-environment [35] in primates and other mammals that occur sympatrically with chimpanzees may reveal that the differentiation across the ecotone is a more common feature of these species' evolutionary history than previously believed.

## Conclusions

Although this study focused on quantifying a pattern of isolation-by-environment in chimpanzees using genetic data comprised of a relatively small number of neutral loci, several important conclusions have emerged from this analysis. The Sanaga River is an important contributor

to patterns of genetic diversity in chimpanzees in Cameroon (Figure 4). However, it is not the only contributing factor (Figures 3 and 4). Habitat and elevational gradients play a major role in partitioning genetic differentiation, especially in *P. t. ellioti*. This was especially evident when GDMs were run for sub-groupings of chimpanzee populations (Figures 5 and 6), where both forest type across elevational and precipitation/moisture gradients played a major role, which is consistent with the ecological niches inhabited by each subpopulation.

Future studies that include data from loci that might be subject to selection are needed to better understand these complex associations, as this study is limited in its ability to detect environmentally diven natural selection in chimpanzees. Genome-wide single nucleotide polymorpism (SNP) data, for example, will allow for quantifying these associations, and provide improved spatial and temporal resolution to disentangle the relative role of Pleistocene refugia in generating genetic diversity. The evolutionary impacts of the environmental gradient in Cameroon have been examined in a small number of taxa [27,31,32], but for the most part, there are not enough studies examining the relative roles of ecotones versus other biogeographic barriers (i.e. the Sanaga River).

In addition to being an important region for chimpanzees, there are also a number of other primate species and subspecies that are presumably influenced by the Sanaga River, including: *Mandrillus leucophaeus/*

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 10 of 13

*M. sphinx, Cercopithecus erythrotis*/*C. cephus, C. nictitans martini*/*C. n. nictitans*, and *C. pogonias pogonias*/*C. p. grayi* [1,13-15,21]. Because these pairs of primate taxa occupy vastly different niches [13,14], it is important to also investigate the role that environmental variation might play in shaping their own patterns of genetic diversity. This study provides an important first step in this process by showing that chimpanzee population structure in eastern Nigeria and Cameroon is not solely driven by separation across riverine barriers, as previously thought.

## Methods

### Overview of modeling patterns of isolation-by-environment

There are many spatially explicit genetic analysis methods that can interpolate a population's genetic structure across a study region [34,40-45]. Recent advances in these methods and the increasing number of publicly-available remote sensing data sets have improved evaluating how ecological, geological and environmental variables influence the genetic structure of populations [34]. Not only can these new models quantify statistical associations of genetic differentiation of sampled populations of study taxa and their habitats, but they can also be used to project inferred patterns of diversity across unsampled areas of the study taxa's projected habitat [34]. These continuous projections represent expected genetic or phenotypic variation, given environmental and/or topographic variation [34]. These models have been used to answer questions from a wide range of topics, including conservation prioritization [46,47], disease ecology [48-51], and biological diversification and speciation [27,32,52,53].

There are a variety of spatial modeling techniques, which include simple regression methods, such as spatially auto-correlated principal components analysis [54], and random forests [55,56] that are used to evaluate environmental and biodiversity associations. GDM is a versatile technique that was originally developed as a matrix regression technique used to study species beta diversity [36]. GDM has also been used to study the relationship between environmental matrices and matrices of morphological and genetic turnover [27,32,34]. In brief, GDM evaluates dissimilarities between environmental and topographic 'predictor' variables and 'response' variables, which can include pairwise genetic distances (e.g. $F_{ST}$) or morphological diversity among populations. The relationships between these types of variables are often non-linear, such as pairwise genetic distances, which are scaled between 0–1, while environmental variables may increase or decrease beyond this scale. GDM's use of non-linear regression algorithms is especially appropriate for understanding the complex influences that landscape features have on shaping patterns of genetic diversity [34].

### Data curation and generation

We used DNA from non-invasively collected chimpanzee hair and fecal samples from a previous study [23]. All samples from this study were transported from Cameroon to the United States in full compliance with Convention of International Trade in Endangered Species of Wild Fauna and Flora (CITES) and Center for Disease Control (CDC) export and import regulations. This research was carried out with IACUC approval from the University at Albany – State University of New York. The genetic data included in the analysis consisted of 21 autosomal microsatellite loci, all determined to be selectively neutral, as well as a 501-base pair (bp) fragment spanning the hypervariable D-loop of mitochondrial (mt) DNA [23]. One hundred and eighty seven individuals from 28 sampling locations found across the study region of Cameroon and eastern Nigeria were genotyped at 21 microsatellite loci (Figure 2). We also included 354 mtDNA control region sequences from 35 sampling locations found across the study region (Figure 2).

For the microsatellite data set, we calculated pairwise $F_{ST}$ values between sample locations using Arlequin version 3.5 [57]. Each pairwise genetic distance was determined by 100,000 replications. In addition, we calculated pairwise difference and Tamura and Nei [38] indices using mtDNA sequence data, from previous studies [20,22,23,58,59], between sample locations from eastern Nigeria and Cameroon using Arlequin version 3.5 [57]. We used a 5:1 transition to transversion weight [22,60], and each estimate was determined by 100,000 replications. All microsatellite loci and mtDNA haplotypes met expectations of Hardy-Weinberg equilibrium and were found to be selectively neutral by an outlier analysis [23].

We obtained environmental data layers from several remote sensing platforms to characterize the habitats available to chimpanzees across the study region. These layers are described in Additional file 4 and can be grouped into three broad categories of factors that describe variation in topography, climate, and vegetation. The topography of the region was described by several variables, including elevation and slope, each sampled at 1 km resolution from the Shuttle Radar Topography Mission dataset [33,61]. We also mapped and included major rivers across the study area in a hydrography layer [62]. Climatic factors describe annual variation in seasonality, temperature and precipitation. Eighteen layers were obtained from WorldClim 1.4 [63] and sampled at 1 km resolution. The category of variables shown in Additional file 4 describes a series of vegetation factors. These factors describe variation in ground cover across the study region. We calculated percent Tree Cover and Leaf Area Index, two measures of vegetation density, from Moderate Resolution Imaging Spectroradiometer (MODIS) imagery [64,65]. Two layers included in this

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 11 of 13

category describe surface moisture, leaf water content and deciduousness of vegetation (QMEAN and QSTD) as measured from the Quick Scatterometer satellite data set [66]. These layers represent annual mean and variability (standard deviation) of surface moisture. Finally, we described variation in vegetation density and deciduousness of vegetation across the study area using a series of Normalized Difference Vegetation Indices, or NDVI, layers [27]. NDVI measurements include mean annual NDVI (NDMEAN), maximum annual NDVI (NDMAX), maximum NDVI during the yearly period of the commencement of the rainy season when new vegetation occurs (NDGR), minimum NDVI during the yearly period with the least amount of new vegetation (NDBR), and NDVI seasonality (NDGRBR).

In order to run the model as accurately as possible, we restricted all spatial analysis to areas within Cameroon and eastern Nigeria where chimpanzees are likely to occur at present time. We used an ecological niche model (ENM) of chimpanzees specifically created for chimpanzees in the study region [33] that was created using a maximum entropy method applied in the program MAXENT [67] to approximate chimpanzee occurrence. This particular model is the most extensive to date for chimpanzees in our region of study, drawing upon occurrence points from the ranges of *P. t. ellioti* (656 total) and *P. t. troglodytes* (98 total). We used a logistic distribution of values from the ENM to establish a threshold in which all cells with a value of less than 10% suitability were excluded from the training mask used as a background for all GDMs, based on a visual comparison of the maps. We excluded these cells to accurately estimate the known range of chimpanzees in the region [68]. By excluding all cells with a value of less than 10% suilability, we eliminated ~77% of the area of Cameroon and Nigeria, area where chimpanzees do not occur [33]. The spatial extent of the mask had no influence on the relative contributions of each of the predictor variables to the models, and was used only as an extent for testing the model and displaying the results.

We identified confounding spatial variables by Pearson correlation tests that were performed using the R software package (http://www.r-project.org) using the topography, vegetation and climate data layers. We clipped each data layer to the extent of the estimated chimpanzee distribution mask. This revealed that several of these variables were highly cross-correlated (Additional file 5). Variables were considered highly correlated for $r^2 > 0.8$. Only BioClim variables [63] exhibited significant cross-correlation with one another. This resulted in Bio 9, 11, 14 and 16 being excluded from the analysis. None of the topographic variables or vegetation indices exhibited significant cross-correlations and were retained for final model construction (Additional file 5).

Incorporating both environmental variables and landscape features into the GDM approach is necessary to estimate how they affect the partitioning of genetic variation and gene flow. River data layers were generated using the hydrography layer [62], to simulate the Sanaga and Mbam Rivers as complete dispersal barriers. In order to imitate a complete dispersal barrier, we coded these variables as 0 on one side, and 1 on the other side. These layers were generated for the Sanaga and Mbam Rivers using ArcMap version 10 (ESRI Corp., Redlands, CA).

We also used resistance surfaces to understand how various landscape features affect gene flow between populations by using an approach that is rooted in circuit theory and implemented in the program CIRCUITSCAPE [69]. CIRCUITSCAPE-based resistance surfaces predict dispersal routes based on predicted cost of travel. These resistance surfaces predict and quantify connectivity between pathways along multiple pathways by using cost weighted distance and suitability of the study area to determine (*i*) a cost-effective route between locations and (*ii*) a matrix of pairwise values representing the incurred cost of travel between locations [27,70]. CIRCUITSCAPE simultaneously integrates all possible pathways that connect populations [69], and has been shown to accurately predict patterns of genetic diversity among animals [27,69,71]. The resistance matrix generated for this study incorporates several variables, including habitat suitability, as determined by the ENM [33], and the size and strength of riverine barriers, based on the Strahler stream order of the major rivers in Cameron and Nigeria [62,72]. In the resistance surface used in this study, low levels of habitat suitability and larger sized rivers between sample locations generated higher levels of resistance.

### Spatial modeling

We used GDM [36] to quantify how much of the variation in the genetic data set could be explained by variation in the environmental data sets, as well as to create explicit spatial predictions of these observed patterns. Using this approach involves incorporating predicted species distributions, environmental data layers, resistance surfaces, and straight-line geographic distance as different predictors. We also used GDM to complete matrix regressions that fit nonlinear relationships between these variables. We then used the model to predict spatial patterns of genetic variation, which facilitated evaluating how environmental data may contribute to generating patterns of genetic diversity [36,73]. This resulted in a continuous spatial prediction of genetic variation across the study area. This spatial prediction was generated using metric multidimensional scaling using 5000 random sample points across the study area, with scores at neighboring pixels achieved by a *k*-nearest neighbors interpolation [36]. We

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 12 of 13

ran GDM's using the entire data set for both microsatellite and mtDNA markers, and spatial interpolation maps were generated for the entire training mask of suitable habitat for chimpanzees across Cameroon and eastern Nigeria, using the one population MAXENT model [33]. For each genetic data set, we ran various models using different groups of predictor variables, including: (1) environmental variables only, (2) environmental variables and geographic distance, and (3) environmental variables, geographic distance and riverine barriers.

We ran additional GDM's for individual population subsets for two- and three-population models as identified by a previous study [23]. In the two-population grouping, we grouped sample locations according to their classification as originating north (*P. t. ellioti*) or south of the Sanaga River (*P. t. troglodytes*). The three-population grouping included sample locations south of the Sanaga River (*P. t. troglodytes*) and we further subdivided presence points north of the Sanaga River into two groups, *P. t. ellioti* (Rainforest) and *P. t. ellioti* (Ecotone). We only ran the GDM for the *P. t. troglodytes* group once, as this group was composed of identical sampling locations. The training masks for all GDM's using the two- and three-population groups were obtained from the MAXENT models of corresponding groups [33].

## Additional files

**Additional file 1: Matrix of microsatellite pairwise differences ($F_{ST}$).** All values in bold were determined as significant by 10,000 permutations of the data in Arlequin [57].

**Additional file 2: Matrix of mtDNA pairwise differences.** All values in bold were determined as significant by 10,000 permutations of the data in Arlequin [57].

**Additional file 3: Matrix of mtDNA Tamura and Nei differences.**

**Additional file 4: List of environmental predicting variables.**

**Additional file 5: Results from Pearson Correlation test comparing environmental variables.** Values shaded grey show all values above 0.8. Values shaded grey and bold show all values above 0.9 and had a p < 0.001.

## Authors' contributions
MWM, SL and MKG designed the research. MWM, PRSC and HAT performed research. MWM and HAT analyzed data. MWM, SL, PRSC and MKG wrote the paper. All authors read and approved the final manuscript.

## Author details
[1]Department of Biology, Drexel University, Philadelphia, Pennsylvania 19104, USA. [2]Department of Biological Sciences, University at Albany – State University of New York, Albany, New York 12222, USA. [3]Institut de Recherche pour le Développement (IRD) and Université Montpellier 1 (UM1), 34394 Montpellier, France. [4]Institute of Evolution and Ecology, Eberhard Karls Universität Tübingen, D-72076 Tübingen, Germany.

## References

1. Oates JF, Bergl R, Linder J. Africa's gulf of guinea forests: biodiversity patterns and conservation priorities. Advances in Applied Biodiversity Science. 2004;6:1–95.
2. Booth AH. The Niger, the Volta and the Dahomey Gap as geographic barriers. Evolution. 1958;12:48–62.
3. Moreau RE. Climatic changes and the distribution of forest vertebrates in West Africa. J Zool. 1969;158(1):39–61.
4. Oates JF. The distribution of Cercopithecus monkeys in West African forests. In: Gautier-Hion A, Bourlière F, Gautier JP, Kingdon J, editors. A Primate Radiation: Evolutionary Biology of the African Guenons. Cambridge: Cambridge University Press; 1988. p. 79–103.
5. Haffer J. Speciation in Amazon forest birds. Science. 1969;165:131–7.
6. Endler JA. Pleistocene forest refuges: Fact or fancy? In: Prance GT, editor. Biological Diversification in the Tropics. New York: Columbia Univeristy Press; 1982. p. 641–57.
7. Mayr E, O'Hara RJ. The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. Evolution. 1986;40(1):55–67.
8. Haffer J. Hypotheses to explain the origin of species in Amazonia. Braz J Biol. 2008;68(4):917–47.
9. Maley J. The African rain forest - Main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. Proc Roy Soc Edinb B Biol Sci. 1996;104:31–73.
10. Robbins CB. The Dahomey gap—a reevaluation of its significance as a faunal barrier to West African high forest mammals. Bull Carnegie Mus Nat Hist. 1978;6:168–74.
11. Myers N, Mittermeler RA, Mittermeler CG, Da Fonseca GAB, Kent J. Biodiversity hotspots for conservation priorities. Nature. 2000;403(6772):853–8.
12. Gonder MK, Locatelli S, Ghobrial L, Mitchell MW, Kujawski JT, Lankester FJ, Stewart CB, Tishkoff SA. Evidence from Cameroon reveals differences in the genetic structure and histories of chimpanzee populations. Proc Natl Acad Sci. 2011;108(12):4766–71.
13. Groves CP. Primate Taxonomy. Smithsonian Institution: Washington, D.C; 2001.
14. Grubb P, Butynski TM, Oates JF, Bearder SK, Disotell TR, Groves CP, Struhsaker TT. Assessment of the diversity of african primates. Int J Primatol. 2003;24(6):1301–57.
15. Kingdon J. The Kingdon Field Guide to African Mammals. Princeton University Press: Princeton; 2003.
16. Clifford SL, Anthony NM, Bawe-Johnson M, Abernethy KA, Tutin CE, White LJ, et al. Mitochondrial DNA phylogeography of western lowland gorillas (Gorilla gorilla gorilla). Mol Ecol. 2004;13(6):1551–65.
17. Anthony NM, Johnson-Bawe M, Jeffery K, Clifford SL, Abernethy KA, Tutin CE, et al. The role of Pleistocene refugia and rivers in shaping gorilla genetic diversity in central Africa. Proc Natl Acad Sci. 2007;104(51):20432–6.
18. Harcourt AH, Wood MA. Rivers as barriers to primate distributions in Africa. Int J Primatol. 2012;33(1):168–83.
19. Chapman CA. Speciation of tropical rainforest primates of Africa: insular biogeography. Afr J Ecol. 1983;21(4):297–308.
20. Gonder MK, Oates JF, Disotell TR, Forstner MR, Morales JC, Melnick DJ. A new west African chimpanzee subspecies? Nature. 1997;388(6640):337.
21. Gagneux P, Gonder MK, Goldberg TL, Morin PA. Gene flow in wild chimpanzee populations: what genetic data tell us about chimpanzee movement over space and time. Phil Trans Biol S. 2001;356(1410):889–97.
22. Gonder MK, Disotell TR, Oates JF. New genetic evidence on the evolution of chimpanzee populations, and implications for taxonomy. Int J Primatol. 2006;27(4):1103–27.
23. Mitchell MW, Locatelli S, Ghobrial L, Pokempner AA, Sesink Clee PR, Abwe EE, et al. The population genetics of wild chimpanzees in Cameroon and Nigeria suggests a positive role for selection in the evolution of chimpanzee subspecies. BMC Evol Biol. In review.
24. Nichol JE. Geomorphological evidence and Pleistocene refugia in Africa. Geogr J. 1999;165(1):79–89.

Mitchell *et al. BMC Evolutionary Biology* (2015) 15:1

Page 13 of 13

25. Ghobrial L, Lankester F, Kiyang JA, Akih AE, de Vries S, Fotso R, et al. Tracing the origins of rescued chimpanzees reveals widespread chimpanzee hunting in Cameroon. BMC Ecology. 2010;10(1):2.

26. Schneider CJ, Smith TB, Larison B, Moritz C. A test of alternative models of diversification in tropical rainforests: Ecological gradients vs. rainforest refugia. Proc Natl Acad Sci. 1999;96(24):13869–73.

27. Freedman AH, Thomassen HA, Buermann W, Smith TB. Genomic signals of diversification along ecological gradients in a tropical lizard. Mol Ecol. 2010;19(17):3773–88.

28. Moritz C, Patton JL, Schneider CJ, Smith TB. Diversification of rainforest faunas: An integrated molecular approach. Annu Rev Ecol Syst. 2000;31(1):533–63.

29. White F. The Vegetation of Africa. Paris: UNESCO; 1983.

30. Smith TB, Wayne RK, Girman DJ, Bruford MW. A role for ecotones in generating rainforest biodiversity. Science. 1997;276(5320):1855–7.

31. Simard F, Ayala D, Kamdem G, Pombi M, Etouna J, Ose K, Fotsing JM, Fontenille D, Besansky N, Costantini C. Ecological niche partitioning between *Anopheles gambiae* molecular forms in Cameroon: the ecological side of speciation. BMC Ecol. 2009;9(1):17.

32. Smith TB, Thomassen HA, Freedman AH, Sehgal RNM, Buermann W, Saatchi S, et al. Patterns of divergence in the olive sunbird *Cyanomitra olivacea* (Aves: Nectariniidae) across the African rainforest-savanna ecotone. Biol J Linn Soc. 2011;103(4):821–35.

33. Sesink Clee PR, Abwe EE, Ambahe RD, Anthony NM, Fotso R, Locatelli S, et al. Chimpanzee population genetic structure in Cameroon and Nigeria is associated with habitat variation that may be lost under climate change. BMC Evol Biol. In review.

34. Thomassen HA, Cheviron ZA, Freedman AH, Harrigan RJ, Wayne RK, Smith TB. Spatial modelling and landscape-level approaches for visualizing intra-specific variation. Mol Ecol. 2010;19(17):3532–48.

35. Wang IJ, Bradburd GS: Isolation by environment. Mol Ecol. 2014;23(23):5649-5662.

36. Ferrier S, Manion G, Elith J, Richardson K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Divers Distrib. 2007;13(3):252–64.

37. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G: LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics. 2008;9(1):323.

38. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993;10(3):512–26.

39. Maisels F. Mbam Djerem National Park, Cameroon: at the forest's edge. Canopee. 2005;27:2–6.

40. Guillot G, Mortier F, Estoup A. GENELAND: a computer package for landscape genetics. Mol Ecol Notes. 2005;5(3):712–5.

41. Chen C, Durand E, Forbes F, François O. Bayesian clustering algorithms ascertaining spatial population structure: A new computer program and a comparison study. Mol Ecol Notes. 2007;7(5):747–56.

42. Corander J, Marttinen P, Sirén J, Tang J. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics. 2008;9:539.

43. Durand E, Jay F, Gaggiotti OE, François O. Spatial inference of admixture proportions and secondary contact zones. Mol Biol Evol. 2009;26(9):1963–73.

44. Jay F. POPS: Prediction of population genetic structure - Program documentation and tutorial. Grenoble, France: University Joseph Fourier; 2011.

45. Mitchell MW, Rowe B, Sesink Clee PR, Gonder MK. TESS Ad-Mixer: A novel program for visualizing TESS *Q* matrices. Conserv Genet Resour. 2013;5 (4):1075–8.

46. Vandergast AG, Bohonak AJ, Hathaway SA, Boys J, Fisher RN. Are hotspots of evolutionary potential adequately protected in southern California? Biol Conserv. 2008;141(6):1648–64.

47. Thomassen HA, Buermann W, Milá B, Graham CH, Cameron SE, Schneider CJ, Pollinger JP, Saatchi S, Wayne RK, Smith TB. Modeling environmentally associated morphological and genetic variation in a rainforest bird, and its application to conservation prioritization. Evol Appl. 2010;3(1):1–16.

48. Gilbert M, Xiao X, Pfeiffer DU, Epprecht M, Boles S, Czarnecki C, et al. Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. Proc Natl Acad Sci. 2008;105(12):4769–74.

49. Harrigan RJ, Thomassen HA, Buermann W, Cummings RF, Kahn ME, Smith TB. Economic conditions predict prevalence of west nile virus. PLoS ONE. 2010;5(11):e15437.

50. Loiseau C, Harrigan RJ, Cornel AJ, Guers SL, Dodge M, Marzec T, Carlson JS, Seppi B, Sehgal RNM. First evidence and predictions of plasmodium transmission in alaskan bird populations. PLoS ONE. 2012;7(9):e44729.

51. Cornuault J, Khimoun A, Harrigan RJ, Bourgeois YXC, Milá B, Thébaud C, Heeb P. The role of ecology in the geographical separation of blood parasites infecting an insular bird. J Biogeogr. 2013;40(7):1313–23.

52. Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C. Stability predicts genetic diversity in the Brazilian Atlantic Forest Hotspot. Science. 2009;323(5915):785–9.

53. Pease KM, Freedman AH, Pollinger JP, McCormack JE, Buermann W, Rodzen J, et al. Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. Mol Ecol. 2009;18(9):1848–62.

54. Novembre J, Stephens M. Interpreting principal component analyses of spatial population genetic variation. Nat Genet. 2008;40(5):646–9.

55. Breiman L. Random Forests. Mach Learn. 2001;45(1):5–32.

56. Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ. Random forests for classification in ecology. Ecology. 2007;88(11):2783–92.

57. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010;10(3):564–7.

58. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, et al. Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1. Science. 2006;313(5786):523–6.

59. Neel C, Etienne L, Li Y, Takehisa J, Rudicell RS, Bass IN, et al. Molecular epidemiology of simian immunodeficiency virus infection in wild-living gorillas. J Virol. 2010;84(3):1464–76.

60. Brown WM, Prager EM, Wang A, Wilson AC. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J Mol Evol. 1982;18(4):225–39.

61. Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, et al. The shuttle radar topography mission. Rev Geophys. 2007;45(2): RG2004.

62. Lehner B, Verdin K, Jarvis A. New global hydrography derived from spaceborne elevation data. Eos, Trans Am Geophys Union. 2008;89(10):93–4.

63. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005;25 (15):1965–78.

64. Hansen MC, Defries RS, Townshend JRG, Sohlberg R. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int J Remote Sens. 2000;21(6–7):1331–64.

65. Myneni RB, Hoffman S, Knyazikhin Y, Privette JL, Glassy J, Tian Y, et al. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sens Environ. 2002;83(1–2):214–31.

66. Long DG, Drinkwater MR, Holt B, Saatchi S, Bertoia C. Global ice and land climate studies using scatterometer image data. Eos, Trans Am Geophys Union. 2001;82(43):503.

67. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190(3–4):231–59.

68. Morgan B, Adeleke A, Bassey T, Bergl R, Dunn A, Fotso R, et al. Regional Action Plan for the Conservation of the Nigeria-Cameroon Chimpanzee (*Pan troglodytes ellioti*). Gland, Switzerland and San Diego, CA, USA: IUCN/SSC Primate Specialist Group and Zoological Society of San Diego, CA, USA; 2011.

69. McRae BH, Beier P. Circuit theory predicts gene flow in plant and animal populations. Proc Natl Acad Sci. 2007;104(50):19885–90.

70. Spear SF, Balkenhol N, Fortin MJ, McRae BH, Scribner K. Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. Mol Ecol. 2010;19(17):3576–91.

71. Schwartz MK, Copeland JP, Anderson NJ, Squires JR, Inman RM, McKelvey KS, Pilgrim KL, Waits LP, Cushman SA. Wolverine gene flow across a narrow climatic niche. Ecology. 2009;90(11):3222–32.

72. Strahler AN. Quantitative analysis of watershed geomorphology. Trans Am Geophys Union. 1957;38(6):913–20.

73. Thomassen HA, Fuller T, Buermann W, Milá B, Kieswetter CM, Jarrín-V P, et al. Mapping evolutionary process: A multi-taxa approach to conservation prioritization. Evol Appl. 2011;4(2):397–413.