Supplementary Materials for

Genetic assignment of large seizures of elephant ivory reveals Africa's major poaching hotspots

S. K. Wasser, L. Brown, C. Mailand, S. Mondol, W. Clark, C. Laurie, B. S. Weir

Correspondence to: wassers@u.washington.edu

This PDF file includes:

Materials and Methods

References

Tables S1 to S4

Figs. S1 to S4

Materials and Methods

We genetically assigned geographic origin to 28 large ivory seizures made between 1996 and 2014 (Table S1). Eight of the seizures made between 1996 and 2009, were acquired concurrently from the Philippines stockpile just prior to their ivory crush in 2013. The original weights of those 8 seizures totaled 6.1 tons, leaving 1.9 tons of ivory unaccounted for. Another key seizure made in the Philippines in 2005 was not available for analysis. That seizure weighed 3.7 tons, was allegedly from Zambia, but was stolen from the warehouse in the Philippines some time before 2007 [21]. The 5 Philippines seizures made between 1996-2005 were initially analyzed separately but then pooled in a single run after originally being assigned to the same area. The last two seizures from 2009 were large enough to warrant running them separately. The first 2009 Philippines seizure (Philppines-6) amplified only 3 tusks, which were assigned to the same locale as the other two 2009 Philippine seizures. Due to its low sample number, that seizure is not included among the 28 seizures described in this paper. The final two Philippines seizures (Philippines 7 and 8) were still intact, having the same number of tusks as when first seized. The remaining 21 seizures analyzed (Table S1) were acquired independently and appeared to be complete at the time of sampling. Of those, one occurred in 2002 (Singapore-2002) and the remainder occurred between 2006 and 2014 (Table S1). The two Togo seizures from 2014 were pooled by authorities prior to being sampled and were thus assigned in a single analysis.

Ivory seizures were sampled to minimize chances of acquiring both tusks from the same individual, and to maximize representativeness of tusks derived from multiple locations in a given seizure, should that occur. To minimize chances of sampling both tusks from

the same individual, tusks were first aligned to maximize chances that one of its adjacent tusks was its pair. Tusks were aligned from smallest to largest based on diameter at the base of the tusk and tusk length. Neighboring tusks were then rearranged, if needed, to assure that the tusks next to each other were also similar in external color and in the distance between the base and the gum-line (where the tusk first protrudes from skin). Every second or third tusk was then retained for analysis. To maximize chances of sampling tusks from multiple locations, we separated the retained tusks into groups according to common external features (e.g., sectioned or whole tusks, orange or white tusk color, weathered appearance as if old, burned above the gum-line as though heated to clean off the tissue, large clumps of tissue left on the tusk, similar writing on the outside of the tusk). We then selected up to 200 samples for analysis across the different groups in proportion to the respective group sizes. However, all selected tusks were generally retained if fewer than 200 tusks remained after separating out the pairs.

Origin assignments were based on genotypes assessed at 16 microsatellite loci [10-13] using DNA isolated from 1350 dung or tissue samples gathered from 71 locations across 29 African countries, with 1-95 samples per location (Table S2). Dung sampling was conducted such that consecutively collected samples were ≥ 1 km apart to reduce chances of multiply sampling the same family unit. The distribution of reference samples across sub-Saharan Africa's five habitat regions was 6.7% northern savanna (NS), 37.9% eastern savanna (ES), 32.0% southern savanna (SS), 3.4% western forest (WF) and 23.3% central forest (CF) (Table S2). Samples were identified as pure savanna or pure forest elephants using the programs SCAT [13] and STRUCTURE [22]. Any identified hybrids were excluded from the reference samples [13].

DNA was extracted and amplified from the dung samples following the methods of Wasser et al. [13], from tissue samples using methods in Roca et al. [9] and Comstock et al [11] and from ivory using the methods in Mailand and Wasser [23]. All samples were extracted twice and each extract amplified 2-3 times per locus, using a multiple tubes approach [13] to minimize allelic drop-out. Depending on the platform, two extraction blank controls per 22 or 48 samples were processed and then amplified with its set. One PCR blank control was added per plate and locus. Two known positive controls were amplified on every PCR plate and loci to monitor PCR and electrophoresis conditions. Genotyping followed strict rules by highly trained individuals. For heterozygotes, each allele had to be observed in at least two separate PCR amplifications. For homozygotes, the single allele needed to be observed in at least three separate PCR amplifications. In the absence of those criteria, both alleles at that locus were coded as missing.

Reference or unknown samples were only included in our analyses if they amplified both alleles at 10 or more loci. In previous simulation studies [13], we observed a 5% difference in error rate between analyses using 10 versus 16 loci. The 10 loci cut-off was based on this small error rate, combined with the benefit it provided in terms of increased number of usable samples. The error model in our assignment method also takes account of deviations from Hardy Weinberg Equilibrium, allowing for null alleles, or non-amplification of one allele, also helping to take account of other excess homozygosity if present [13]. The percentage of samples/seizure that amplified for 10 or more, 15 or more, or all 16 loci is indicated in Table S1.

While rare, duplicate reference or unknown samples from the same individual were identified using Cervus [24], including only one of the matches in any statistical assignment. Two samples were considered a match when their genotypes were identical at all genotyped loci, or if there were a maximum of two non-matching loci, each of which could have been a non-match because of allelic drop-out (i.e., both samples were homozygous for different alleles at the same locus, or a homozygote shared one of the two alleles of a heterozygote at the same locus).

Figure S1 shows the allele frequencies at each of the 16 loci in our reference data set, separated by savanna and forest subspecies. While the allele frequency distributions for loci between forest and savanna elephants occasionally overlap at any single locus, there are several very informative loci. Taken across all 16 loci there is close to zero error in distinguishing forest versus savanna elephants [13], consistent with the $\geq$ 2.5 million year divergence of forest and savanna elephants [9]. Sample assignment thus begins by first identifying samples by African elephant subspecies: forest, savannah, or hybrid [13]. We then proceed to process each species separately, using only forest references to assign forest samples and only savanna references to assign savanna samples. Hybrids are not processed further.

We assign samples to subspecies by grouping reference data into 5 regions: central and western forest, and central, east, and south savannah. Based on the allele frequencies associated with each region, we determine (using maximum likelihoods) the regions of origin of the parents of the sample individual. If both parents originate from forest regions, we classify the sample as forest. If both originate from savannah regions, we

classify the sample as savannah. If one parent originates from forest and the other savannah, we classify the sample as a hybrid.

The second step applies the Smoothed Continuous Assignment Technique (SCAT) [13] to estimate the posterior distribution of the location of origin of each sample, given the reference data. SCAT uses a Bayesian method [25] implemented with Markov Chain Monte Carlo (MCMC) spatial-smoothing [26] to simultaneously estimate allele frequencies at any location in Africa. Allele frequencies are assumed to depend on all reference samples with a spatial correlation that depends on distance between populations. Wasser et al. [13] showed that this approach significantly improves the accuracy of our allele frequency estimates over standard assignment methods (SAM), which estimate allele frequencies in each population separately. The greatest improvements in SCAT over SAM occurred at sampling locations with relatively few reference samples [13].

SAM also has a strong systematic bias against assigning samples to reference locations where sample size is small because small reference sample sizes give inaccurate allele frequencies at that location. However, artificially small likelihoods arising from inaccurate allele frequency estimates can cause gross overconfidence in incorrect conclusions. SCAT alleviates these problems by using all reference samples to estimate allele frequencies at any location, giving more weight to samples from nearby sampling locations. Wasser et al [13] used allele frequency estimates to compute the log-likelihood of each individual's genotype at its actual sampling location. Inaccurate allele frequency estimates tend to give smaller values for these log-likelihoods and these log-likelihoods

were almost always larger for allele frequencies estimated by SCAT versus SAM, again with the greatest improvements for locations with small reference sample sizes.

Each SCAT run begins with 1000 burn-in iterations to estimate allele frequencies across the species' range based on the reference data. Then, we take a sample with known genotype data g and unknown location of origin W and include this in the MCMC algorithm, giving W a uniform prior over the species' range. The MCMC algorithm produces a sampling of the posterior distribution of (g,W) given the reference data. We discard as burn-in the first 1000 iterations after (g,W) is included, and store every 10^{th} of the next 1000 iterations as a sample from the posterior of (g,W), resulting in 100 samples from this posterior. We then remove (g,W) from the MCMC algorithm and repeat the process with the next sample, until all samples have been processed.

To reduce the effects of possible lack of convergence of the MCMC scheme, we perform 9 independent SCAT runs, each starting with a different seed from the pseudo-random number generator, resulting in 900 samples from the posterior of each (g,W). The medians of the latitude and longitude of the posterior samples give an estimate for W, and the spread gives an indication of the precision.

SAMs have the implicit assumption that each sample comes from one of their n sampling locations. Our continuous assignment technique does not make that assumption. It allows for the sample to have arisen from any location in the elephant's range. Our method also estimates how much confidence to place in the assigned sample location by drawing 100 plausible locations from the set of all plausible locations, weighted according to their probability. Tight clustering indicates high confidence, while reduced clustering indicates

lower confidence. Thus, our methods assess uncertainty in each assignment; if few loci are used, or allele frequencies overlap, this will be reflected by increased uncertainty in each assignment [13].

The third and final step post-processes results from SCAT, using a group assignment method that capitalizes on genetic similarities between samples [15]. The group assignment assumes the samples were drawn uniformly from some region *R*, and attempts to simultaneously estimate both *R* and the locations of origin of the samples. The region *R* is assumed to consist of one or more polygons, not necessarily adjacent. We partition the continent into 100 irregular polygons by using a process known as Voronoi tessellation [27]. After specifying a prior distribution for *R*, we use an MCMC scheme to draw an approximate sample from the posterior distribution of *R* given the genotype data (in the form of the SCAT results). We can then compute an estimate of the location of origin of each individual sample by weighting the SCAT results according to the posterior distribution of *R*.

The uniform prior in SCAT presupposes that the region *R* includes the whole of the African (savannah or forest) elephant range, whereas during group assignment *R* is estimated from the data. The latter allows information to be ‘‘borrowed’’ across tusks to improve the precision of individual estimates [15]. The resultant weighted sample is then used to estimate the location of origin of each individual sample (i.e., the estimated posterior mean latitude and longitude is computed from the weighted sample), taking into account the genotype information of all samples simultaneously because all of the genotype information is used to estimate *R*.

Wasser et al [15] showed that simultaneous analysis of multiple samples is able to recognize when samples are originating from a limited geographic region and use this information to produce estimated locations that are more accurate and compact compared to assigning samples one at a time. This improvement in precision is greatest when *R* is a relatively restricted subset of Africa. If, on the other hand, *R* turns out to include samples distributed evenly across the forest or savanna elephants range, all of the weights will be approximately equal, giving virtually identical results to a sample-by-sample analysis using the CAM. Group assignment of large numbers of samples (e.g., large seizures) with a restricted distribution can improve assignment precision even when those samples originated from areas with relatively few reference samples (Figure S2; see also leave-full-location-out validations, Table S3).

Finally, to determine whether run lengths were sufficiently long to obtain reliable results, Wasser et al. [15] applied the group assignment algorithm three times, from three different random starting points, each time using 10,000 iterations and discarding the first 5,000 of these iterations as burn-in. Results from the three different starting points were qualitatively similar, suggesting that these run lengths were sufficient to produce reliable results.

Genetic Distance Among Reference Samples

We used the F-statistic F_{ST} to examine how the genetic distance between two populations changes with the geographic distance between them and how this spatial variation impacts assignment accuracy among forest compared to savanna elephants. Geographic distance between two sites was calculated as a great-circle ``as the crow flies" distance from the latitudes and longitudes of those sites with the R-package geosphere [28]. For

each pair of populations, we calculated the great circle of distance in km and calculated F_{ST} as described in the next paragraph. No parametric values were assumed.

Population-specific allele frequencies were estimated at each locus from the genotypes of individuals sampled from that population, and were summarized as heterozygosities to estimate Fst. Following Bhatia et al. [28] we used an unweighted approach: if p_{ilu} is the observed frequency of allele u at locus l for population i, then the heterozygosity within population i is $H_i=\Sigma_l[n_{il}/(n_{il}-1)](1-\Sigma_u p_{ilu}^2)$ where n_{il} alleles are observed for that population at locus l. The heterozygosity between populations i,j is $H_{ij}=\Sigma_l(1-\Sigma_u p_{ilu}p_{jlu})$. For each pair i,j of populations these heterozygosities are combined to provide estimate F_{ST} as $F_{ST}=1-(H_i+H_j)/(2H_{ij})$ and these quantities serve as genetic distances between the populations [29].

Figure S3 shows a significant positive relation between genetic and geographic distance for both subspecies. Linear regression could be used to estimate the increase in F_{ST} per km. The adjusted r^2 for savanna population was 0.63 with an F-statistic of 793.9 (at 1 and 463 degrees of freedom) and $p < 0.0001$. For forest elephant populations, the adjusted r^2 was 0.31 with an F-statistic of 29.77 (at 1 and 64 degrees of freedom) and $p < 0.0001$. While the slopes of regression lines (Figure S3A and S3B) were not significantly different (savanna=0.0000215; forest=0.0000233), the minimum Fst between any two forest locations is approximately quadruple that for savanna locations (intercept for forest = 0.023, for savanna = 0.006), indicating greater population structure in the forest populations.

Assessing Assignment Accuracy

We established the accuracy of our assignment methods across all 71 separate locations in our dataset for both uniform and Voronoi priors. Assignment accuracy of SCAT with either a uniform or Voronoi prior was assessed by several cross validation analyses on the reference samples, testing cases where there are some or no other reference samples at the location being assigned. We remove either one sample, half of the samples or all of the samples at a reference location, estimate allele frequencies in their absence and use the new frequencies to blindly assign geographic origin to the removed sample(s). Leave-one-sample-out cross validation [30] leaves out each sample, in turn, assigning a location to the removed sample as though its origin was unknown. In that case, all remaining samples for that subspecies, including all remaining references samples at that location are used as a reference set. Single sample assignments can only be examined using a uniform prior because, by definition, they cannot utilize the shared allele information between multiple samples that becomes possible with a Voronoi prior.

In the instance of leave-half-location-out or leave-full-location-out cross validations, respectively, half or all of the samples from a location are removed and each removed sample assigned a location of origin. For the leave-half-location-out and leave-full-location-out cross validation analyses, the samples left out are assigned a location of origin independently using a uniform prior [13] and collectively using a Voronoi prior [15]. The leave half-location-out analysis is done separately for each half. Only locations with two or more samples (21 of the 24 forest locations and 46 of the 47 savanna locations) were included in the leave-half-location-out cross validation analyses while the leave-full-location-out cross validation analyses were conducted on all 71 locations.

Assignment accuracy was greater for the more structured populations in leave-one-sample-out cross validation (median accuracy = 349 km for forest, 491 km for savanna populations, Table S3). These single sample assignments benefit from including all remaining reference samples from the assigned location, which are informative of local allele frequencies. The gain in accuracy when having reference samples at the assigned location is most evident in West African forest populations, suggesting them to be the most structured of all the African elephant populations (median accuracy = 124 km). Among savanna populations, median accuracy was highest for the northern savanna elephant samples (353 km) compared to southern (464 km) and eastern (536 km) savanna regions where gene flow is presumed to be substantially higher.

Consistent with the above, removing all reference samples from the assigned location (leave-full-location-out cross validation) using a uniform prior resulted in a 60% decline in median accuracy over that for leave-one-sample-out cross-validations among forest sample assignments (Table S3). The same comparison among the less structured savanna elephants resulted in only a 33% decline. The greatest decline in median accuracy occurred in western forest sample assignments (527%) while the eastern savanna assignments were the least affected (22% decline in median accuracy) by the removal of all reference samples from an assignment location (Table S3).

By contrast, use of a Voronoi prior in leave-full-location-out cross validations utilizes the maximum amount of shared information for that location in the Voronoi prior group assignment but loses the allele frequency information specific to that location in the reference panel. Applying the Voronoi prior to the leave-full-location-out cross validation improves assignment accuracy over use of a uniform prior by 78% for savanna and 17%

for forest populations. The least improvement from use of the Voronoi prior was in the more structured elephant populations among their respective subspecies, with a 1% increase in assignment accuracy of western forest elephant samples and an 87% improvement in eastern savanna elephant samples over the uniform prior (Table S3). This result indicates that group assignment using a Voronoi prior can still achieve reasonable precision when assigning tusks from large seizures to locations with few or no reference samples because of the improvements resulting from shared information between the large number of samples with restricted origin.

Half-location-out assignments using a Voronoi prior displayed higher accuracy and lower variability than any other combination (Table S3). Assignment accuracy here benefits from both the shared information in a group assignment and from the presence of reference samples at the location being assigned [15]. Overall, 50% of samples were assigned to within 275 km of its actual origin for anywhere in Africa and the majority (75%) of samples to within 430 km of their actual origin in half-location-out assignments using a Voronoi prior. The largest improvement in accuracy of half- compared to full-location-out assignments using a Voronoi prior occurred in the more structured forest versus savanna elephant populations. The greatest improvements occurred in western forest (315%) and the least (9%) in eastern savanna populations (Table S3), again emphasizing the importance of local allele frequencies in the more structured populations.

Figure S4A and S4B provides a more detailed illustration of assignment accuracy using the half-location-out protocol with a Voronoi prior for each of our forest and savanna

locations, respectively, that have sample sizes ≥ 10. In those cases, the spread is generally far less than the distance between most protected areas across Africa.

It is also important to note that all half-location-out assignment accuracy estimates should be considered conservative since that protocol utilizes only half the available number of samples at any given location. Actual assignments of unknowns from those areas would actually have double the reference samples at the location, along with the added benefit of shared information from group assignment of the large number of samples in these seizures (Figure S5, Table S3).

Increasing sample size most improves assignment accuracy when using a Voronoi versus a uniform prior (Figure S5). These sample size improvements are greatest for the less structured savanna populations (Table S3) because a Voronoi prior uses all samples from a common location to diminish the contribution of outliers. The sample size advantage diminishes after approximately 25 samples per location (Figure S5).

Effect of prior on sample size impacts by geographic region

Analysis of covariance (ANCOVA) showed that assignment accuracy increased at larger sample sizes in all savanna populations when using a Voronoi prior over a uniform prior. Southern and eastern savanna Voronoi prior assignments showed particularly steep declines in median distance from true location with increasing sample size at the assigned location compared to the use of a uniform prior ($p=0.0116$ and 0.0091, respectively). The small number of northern savanna locations, which also have relatively high population structure, did not show a strong difference in median assignment distances between priors ($p=0.539$). Central forest assignments also decline in median distance from true location

as sample size increases at that location for the Voronoi versus uniform prior (p=0.0008). However, the western forest region exhibits no effect of sample size with almost identical results for uniform and Voronoi prior assignments (p=0.0574), reflecting their relatively high degree of population structure (see also main text).

Only a few outlying savanna reference locations with large sample sizes displayed high within-location assignment variability, yielding a reduction of median accuracy at those locations. One half of the samples from Queen Elizabeth National Park, Uganda were assigned with a very large spread (Figure S4B). This protected area is proximal to one of the rare areas where forest-savanna hybrids are found [9,11], potentially impacting the purity of sub-specific allele frequency estimates there. Both Tarangire and Amboseli National Parks in Tanzania and Kenya, respectively, also display relatively high assignment error (Figure S4B), likely explained by the tendency of those elephants to move particularly long distances during some times of the year [31-33].

Effect of prior on within and between population variance

An analysis of variance (Table S4) revealed that the use of the Voronoi prior over the uniform prior dramatically reduces within-location variability for savanna populations. After accounting for the strong clustering of assignments that results from Voronoi prior analysis, a larger proportion of the total variability in assignments is explained by differences across the compared locations. Across location variability is higher in savanna versus forest populations. Uniform and Voronoi priors show similar across location variability for forest populations, but within-location variability is reduced greatly with the use of a Voronoi prior over a uniform prior.

Table Legends

Table S1. Details of seizures obtained for origin analysis during 2013.

Table S2. Sample and location information for reference panel sites. Region abbreviations are: CF for Central Forest, WF for West African Forest, NS for Northern Savanna, ES for Eastern Savanna and SS for Southern Savanna.

Table S3. 25^{th}, 50^{th} and 75^{th} percentile of distances (km) from true location for single sample, full location and half location assignments by species, region and assignment method.

Table S4. Standard deviation estimates (km) from ANOVA analyses.

Figure Legends

Figure 1. Allele frequencies of the 16 microsatellite loci used in the DNA assignments, broken down by forest (green) and savanna (orange) subspecies.

Figure S2. Heat maps showing how assignment precision increases with the number of unknowns assigned from a restricted area that has few (n=5) reference samples when using a Voronoi prior. The redness of the grid square represents the fraction of the time that grid square is in the region *R* being computed by the Voronoi program, relative to the most frequently included grid square. This is, up to a constant, the probability that a given grid square belongs to the region R from which the unknowns are assumed to be uniformly drawn.

Figure S3. Estimated *Fst* vs distance for pairs of populations with 10 or more samples for forest (A) and savanna (B) populations.

Figure S4. Assignment accuracy by location for each (A) forest and (B) savanna elephant reference population with ≥ 10 reference samples based on the leave-half-location-out assignment protocol using a Voronoi prior. Each location has two maps, one for each half of the samples removed and subsequently assigned as unknowns. The red cross in each figure is the true location being assigned. The blue dots are the assigned locations of the reference samples treated as unknowns. Green crosses (Figure 3A) represent forest elephant reference sample locations used to make the assignments. Orange crosses (Figure 3B) represent savanna elephant reference sample locations used to make the assignments.

Figure S5. Median distance from true location vs number of total reference samples at location for Voronoi prior half location assignments by region.

Supplemental Tables:

Table S1.

Name of Seizure	Location of Seizure	Known Transit Countries in order of occurrence	Primary Country of Origin	Date of Seizure	Seizure Weight, # of tusks	# of Samples Analyzed	Date Received	Subspecies in Seizure	% DNA Amplification Success at ≥ 10 loci, ≥ 15 loci, all 16 loci
Philippines-1-1996	Philippines	Unknown	Eastern DRC	1996	594 kg 40 tusks	17	9/2013	Forest	47, 35, 35
Philippines-2-1997	Philippines	Unknown	Eastern DRC	1997	61 kg 20 tusks	9	9/2013	Forest	44, 22, 22
Philippines-3-2005	Philippines	Zambia	Eastern DRC	4/2005	158 kg 78 tusks	27	9/2013	Forest	82, 44, 37
Philippines-4-2005	Philippines	Uganda	Eastern DRC	9/2005	203 kg 150 tusks	55	9/2013	Forest	47, 29, 18
Philippines-5-2005	Philippines	Kenya	Eastern DRC	10/2005	168 kg 70 tusks	21	9/2013	Forest	48, 38, 33
Singapore-2002	Singapore	Malawi, South Africa	Zambia, Angola	6/2002	6500 kg 532 tusks	66	8/2004	Savanna	44, 33, 24
Singapore-2007	Singapore	Zambia, Egypt, Bangkok, Hong Kong, Singapore	Zambia	3/2007	470 kg 70 tusks	45	12/2008	Savanna	82, 44, 25
Taiwan1	Taiwan	Tanzania	Tanzania, Mozambique	7/2006	3000 kg 744 tusks	60	1/2008	Savanna	80, 52, 12
Taiwan2	Taiwan	Tanzania	Tanzania, Mozambique	7/2006	1150 kg 350 tusks	89	1/2008	Savanna	62, 49, 35
HongKong	Hong Kong	Tanzania	Tanzania, Mozambique	7/2006	2600 kg 390 tusks	99	11/2006	Savanna	32, 1, 0
HongKong-Cameroon	Hong Kong	Cameroon	TRIDOM, Gabon, Republic of Congo	5/2006	4000 kg 603 tusks	150	2/2008	Forest	76, 65, 51
Philippines-6-2009	Philippines	Kenya	Tanzania, Zambia	3/2009	89 kg 36 tusks	12	9/2013	Savanna	25, 16, 16
Philippines-7-2009	Philippines	Tanzania	Tanzania, Mozambique	3/2009	3378 kg 1086 tusks	192	9/2013	Savanna	42, 31, 27
Philippines-8-2009	Philippines	Tanzania	Tanzania, Mozambique, Zambia	3/2009	1442 kg 615 tusks	89	9/2013	Savanna	29, 19, 15

Avocado	Kenya	Tanzania, Kenya	Tanzania, Mozambique, Kenya	8/2010	1462 kg 317 tusks	141	3/2012	Savanna	81, 67, 57
Thailand	Thailand	Mozambique, Laos	Mozambique	1/2011	336 kg 69 tusks + 4 pieces	73	2/2015	Savanna	62, 32, 19
Pili	Kenya	Tanzania, Kenya	Tanzania, Mozambique, Zambia, Kenya, Uganda, DRC	5/2011	1305 kg 115 tusks	49	2/2012	Savanna	69, 61, 57
Malaysia	Malaysia	Togo	TRIDOM, Gabon, Republic of Congo, Cameroon, DRC, Ghana, Ivory Cost, Togo, Tanzania, Kenya, Uganda, Zambia	12/2012	6000 kg 2300 tusks	266	1/2015	Forest Savanna	79, 57, 46
Hong Kong-Togo	Hong Kong	Togo	TRIDOM, Gabon, Republic of Congo	12/2012	2000 kg 1,148 tusks	200	9/2013	Forest	25, 10, 5
Hong Kong-Kenya	Hong Kong	Kenya, Malaysia	Mozambique, Tanzania	1/2013	1324 kg 779 tusks	106	7/2014	Savanna	48, 26, 23
Singapore 2013	Singapore	Kenya	Tanzania, Kenya	3/2013	1838 kg 1099 tusks	53	7/2014	Savanna	68, 62, 45
Malawi	Malawi	Tanzania	Tanzania, Mozambique	5/2013	2640 kg 781 tusks	200	1/2014	Savanna	31, 24, 22
Sri Lanka	Sri Lanka	Uganda, Kenya	Tanzania	7/2013	1.5 tons 359 tusks	205	7/2013	Savanna	63, 54, 49
Togo-N'Bouke	Togo	Not Applicable. Samples were seized from the warehouse of a dealer selling ivory carvings	Widely dispersed: West Africa to TRIDOM, Gabon, Republic of Congo	8/2013	700 kg; (mostly worked) 69 tusks	69	11/2013	Forest[1]	52, 42, 38
Hong Kong-Nigeria	Hong Kong	Nigeria	TRIDOM, Gabon, Republic of Congo, Cameroon	8/2013	2230 kg 1120 tusks	200	12/2013	Forest[5]	38, 19, 9
Uganda-Oct 2013	Uganda	Uganda	Tanzania	10/2013	2903 kg 823 tusks	187	4/2014	Savanna[3]	62, 54, 48

Uganda-Dec 2013	Uganda	Uganda	Tanzania, Kenya, Uganda	12/2013	1424 kg 440 tusks	188	4/2014	Savanna[4]	56, 46, 41
Togo	Togo	Togo	TRIDOM, Gabon, Republic of Congo	1/2014-1 1/2014-2	Seizure 1: 1689 kg, 1500 raw pieces; Seizure 2: 2166 kg	200	4/2014 g	Forest[2]	40, 33, 27

1. Contained 3 savanna tusks; 2. Contained 7 savanna tusks; 3. Contained 4 forest tusks; 4. Contained 5 forest tusks; 5 Contained 2 savanna tusks

Table S2.

Location Name	Sample Size	Country	Region	Latitude	Longitude
Cabinda	3	Angola	CF	-4.7	12.65
Banyang Mbo	5	Cameroon	CF	5	9.5
Dja	6	Cameroon	CF	3.2	13.7
Congo	28	Congo	CF	0.58	15.61
Dzaga Sanga	94	Congo	CF	2.8	16.35
Garmaba	25	DRC	CF	4	29.46
Kahuzi	1	DRC	CF	-2.2	28.7
Central DRC	15	DRC	CF	4.46	24.66
N Salonga	29	DRC	CF	-1.21	20.97
NW Salonga	9	DRC	CF	-1.44	24.95
S. Salonga	1	DRC	CF	-2.1	21.1
Okapi	10	DRC	CF	1.5	28.5
Moukalaba	14	Gabon	CF	-2	9.9
Minkebe	24	Gabon	CF	1.8	12.7
Lope	17	Gabon	CF	-0.25	11.53
Bamingui	4	NA	CF	7.55	20.18
Virunga	12	Rwanda	CF	-1.5	29.5
Kibale	7	Uganda	CF	0.5	30.4
Tai	1	Cote D'Ivoire	WF	5.8	-7.5
Bia	6	Ghana	WF	6.33	-3
Mole	12	Ghana	WF	9	-2
Guinea	8	Guinea	WF	8.28	-9.25
Sierra Leone	8	Sierra Leone	WF	9.72	-12.13
Togo	10	Togo	WF	8.7	1.05
Benoue	16	Cameroon	NS	8.1	14
Waza	20	Cameroon	NS	11.3	14.8
Chad	29	Chad	NS	11.1	19.5
Logone Oriental	8	Chad	NS	8	16
Mayo	8	Chad	NS	10.45	16.48
Mali	4	Mali	NS	15.5	-1.9
Nigeria	3	Nigeria	NS	9.83	10.58
Virunga	9	DRC	ES	-1.25	29.5
Eritrea	3	Eritrea	ES	15.18	39.78
Mago	4	Ethiopia	ES	5.5	36.3
Aberderes	17	Kenya	ES	-0.42	36.63
Amboseli	68	Kenya	ES	-2.55	37.26
Mount Kenya	3	Kenya	ES	-0.15	37.3
Tsavo	66	Kenya	ES	-3.01	38.51
Akagera	8	Rwanda	ES	-1.6	30.7
Rusizi	2	Rwanda	ES	-2.5	29.3
Nimule	4	Sudan	ES	3.362	32.29

Katavi	18	Tanzania	ES	-6.74	31.05
Mikumi	60	Tanzania	ES	-7.6	37.14
Ngorongoro Crater	15	Tanzania	ES	-3.23	35.57
N Ruaha	19	Tanzania	ES	-5.67	34.58
Ruaha	39	Tanzania	ES	-7.5	34.75
Rukwa	12	Tanzania	ES	-7.2	31.98
Selous	18	Tanzania	ES	-8.8	37.5
Serengeti	16	Tanzania	ES	-1.97	35.03
Tarangire	40	Tanzania	ES	-4.1	36.1
Udzungwa	5	Tanzania	ES	-7.75	36.70
KVCA	17	Uganda	ES	3.84	33.69
Murchison Falls	22	Uganda	ES	2.34	31.54
Queen Elizabeth	29	Uganda	ES	-0.2	30.11
Chobe	81	Botswana	SS	-18.33	24.86
S Malawi	17	Malawi	SS	-14.9	35
C Malawi	12	Malawi	SS	-12.9	33.5
Gorongosa	14	Mozambique	SS	-18.55	34.26
N Mozambique	26	Mozambique	SS	-11.83	37.91
NW Namibia	14	Namibia	SS	-18.75	15.5
Torra Conservancy	2	Namibia	SS	-20.16	13.60
Kruger	95	South Africa	SS	-23.83	31.5
E Zambia	26	Zambia	SS	-12.5	32
Kafue	18	Zambia	SS	-15.2	25.9
Kasanka	4	Zambia	SS	-12.6	30.2
S. Zambia	10	Zambia	SS	-15.5	29.7
Sumbu	7	Zambia	SS	-8.64	30.44
Hwange	54	Zimbabwe	SS	-19.05	26.6
Mashatu	22	Zimbabwe	SS	-22.08	29.11
Sengwa	16	Zimbabwe	SS	-18.22	28.56

Table S3.

	N. Locs Tested	Single Sample Uniform			Full Location Uniform			Full Location Voronoi			Half Location Uniform			Half Location Voronoi		
		25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%
Savanna																
Overall	47	229	491	840	390	654	1026	221	368	557	203	543	909	148	267	437
Northern	7	230	353	648	288	472	967	229	337	456	259	393	750	104	180	396
Eastern	24	230	536	962	394	653	1079	213	349	540	340	585	1021	192	312	492
Southern	16	222	464	750	412	677	993	225	406	528	283	543	849	108	231	380
Forest																
Overall	24	154	349	552	386	557	822	321	475	648	243	421	644	195	301	416
Western	18	82	124	301	378	654	994	395	650	850	95	197	420	116	161	270
Central	6	199	379	565	389	540	795	317	465	610	279	440	661	216	320	429

Table S4.

		Uniform	Voronoi
Savanna			
	Across Locations	151.3	365.6
	Within Locations	571.7	293.7
Forest			
	Across Locations	108.8	125.9
	Within Locations	412.7	224.7

Figure S1.

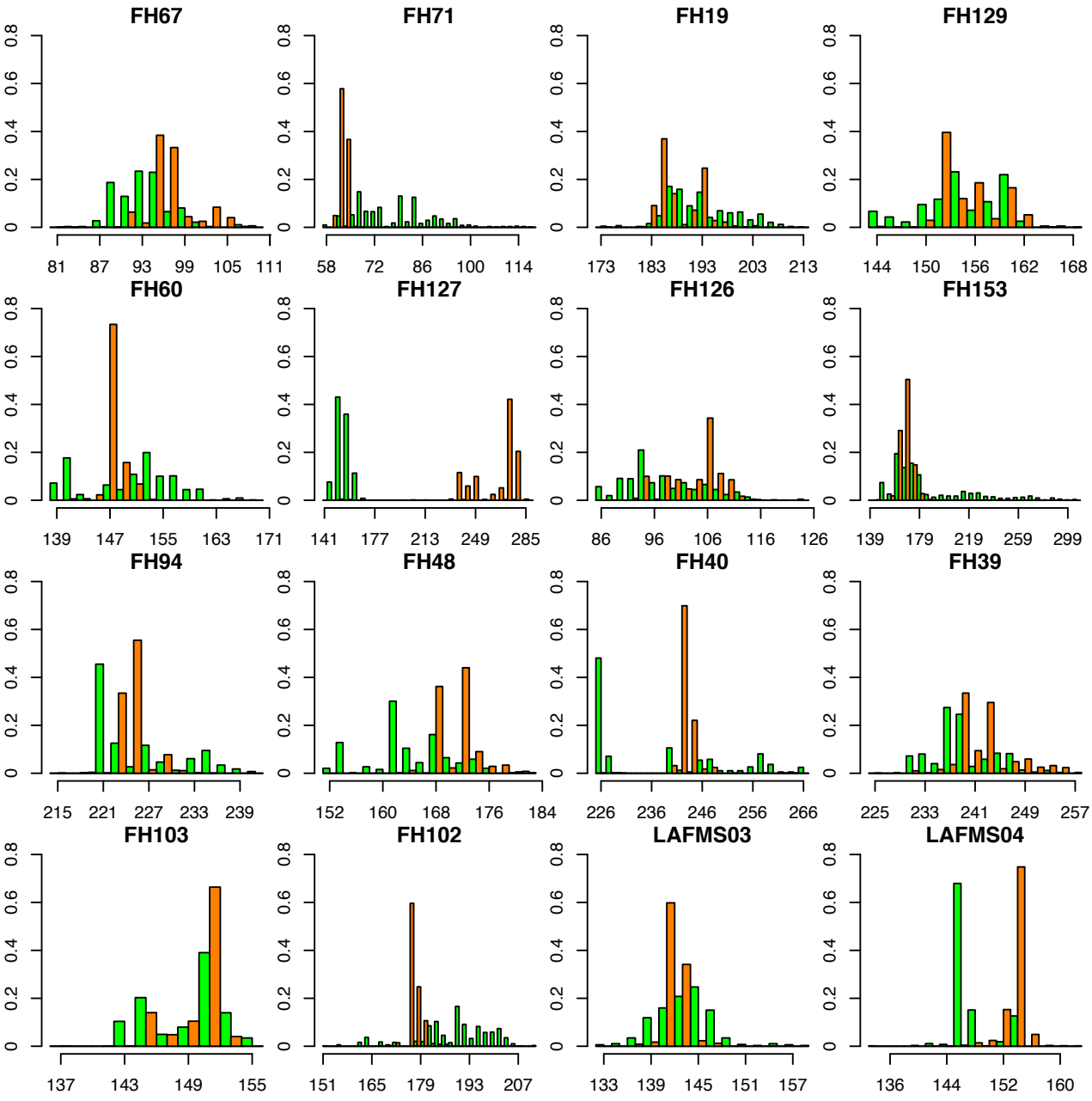
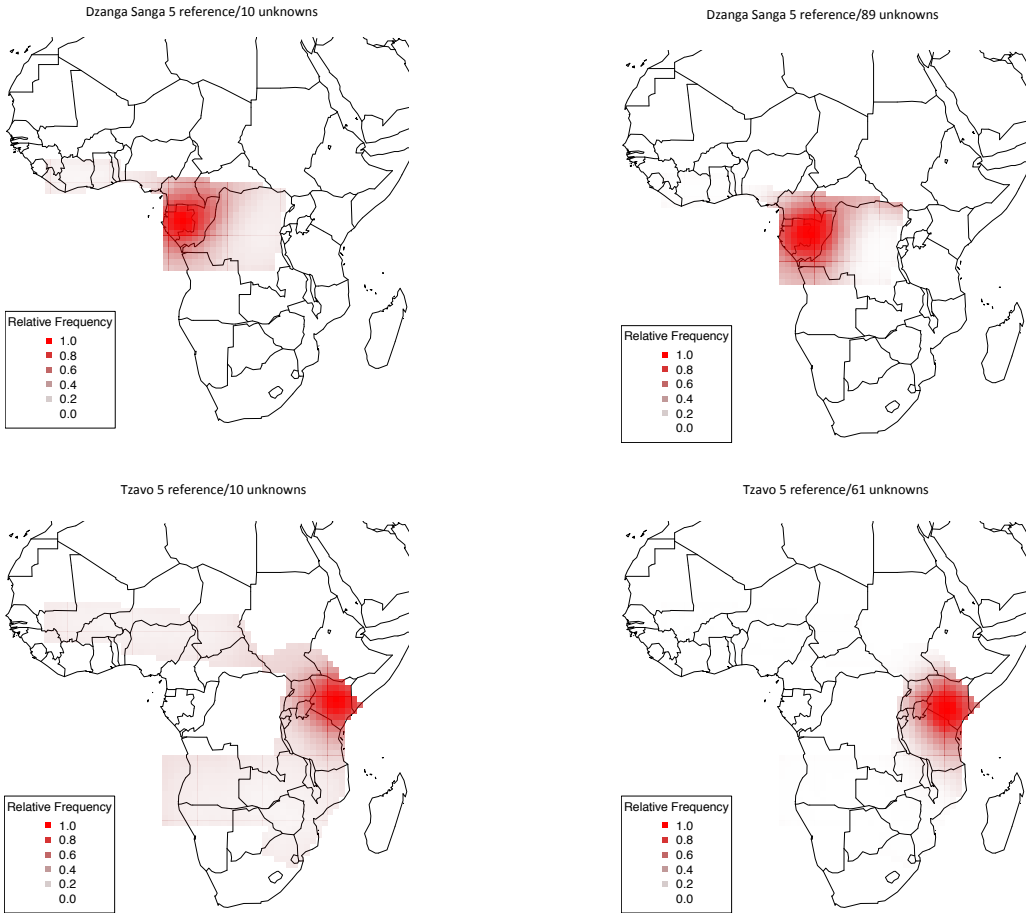
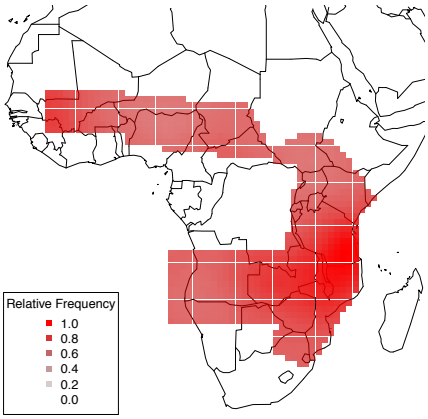

FH67
FH71
FH19
FH129
FH60
FH127
FH126
FH153
FH94
FH48
FH40
FH39
FH103
FH102
LAFMS03
LAFMS04

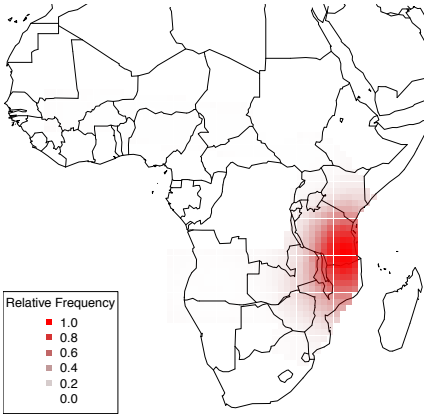
Figure S2.


Dzanga Sanga 5 reference/10 unknowns
Relative Frequency
1.0
0.8
0.6
0.4
0.2
0.0
Dzanga Sanga 5 reference/89 unknowns
Relative Frequency
1.0
0.8
0.6
0.4
0.2
0.0
Tzavo 5 reference/10 unknowns
Relative Frequency
1.0
0.8
0.6
0.4
0.2
0.0
Tzavo 5 reference/61 unknowns
Relative Frequency
1.0
0.8
0.6
0.4
0.2
0.0

Mikumi 5 reference/10 unknowns

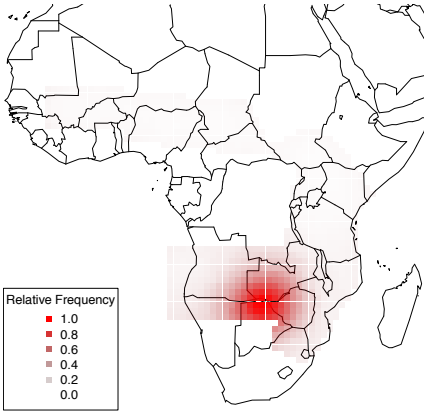


Mikumi 5 reference/76 unknowns



Tzavo 5 reference/61 unknowns

Chobe 5 reference/10 unknowns



Chobe 5 reference/76 unknowns

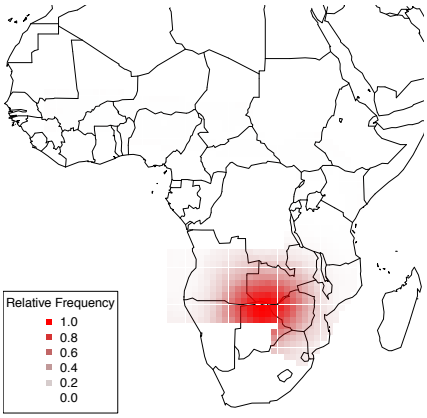


Figure S3

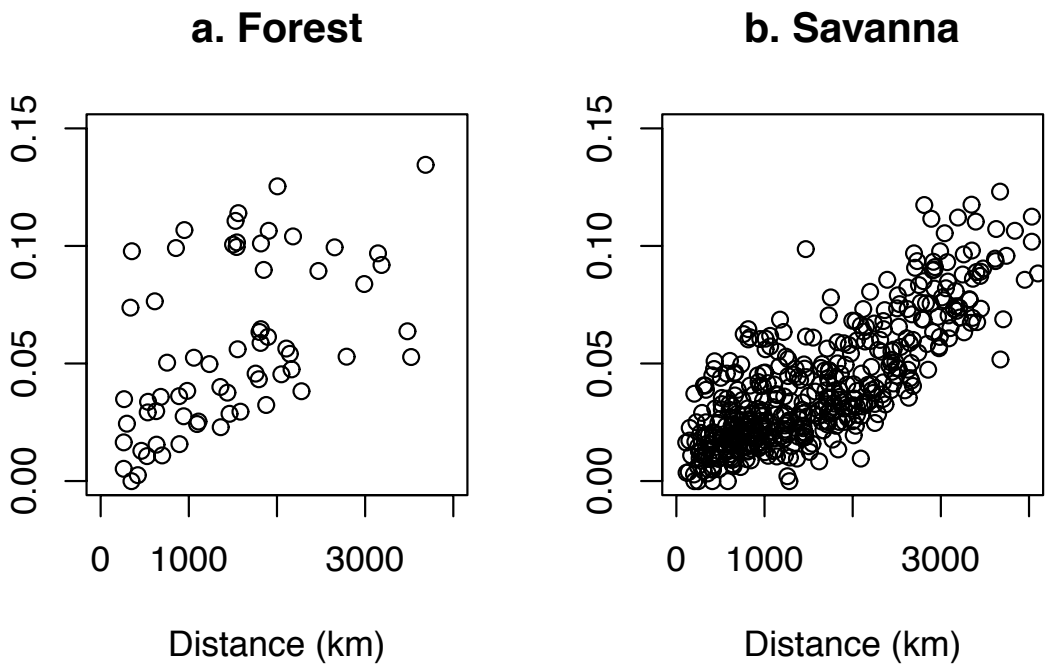
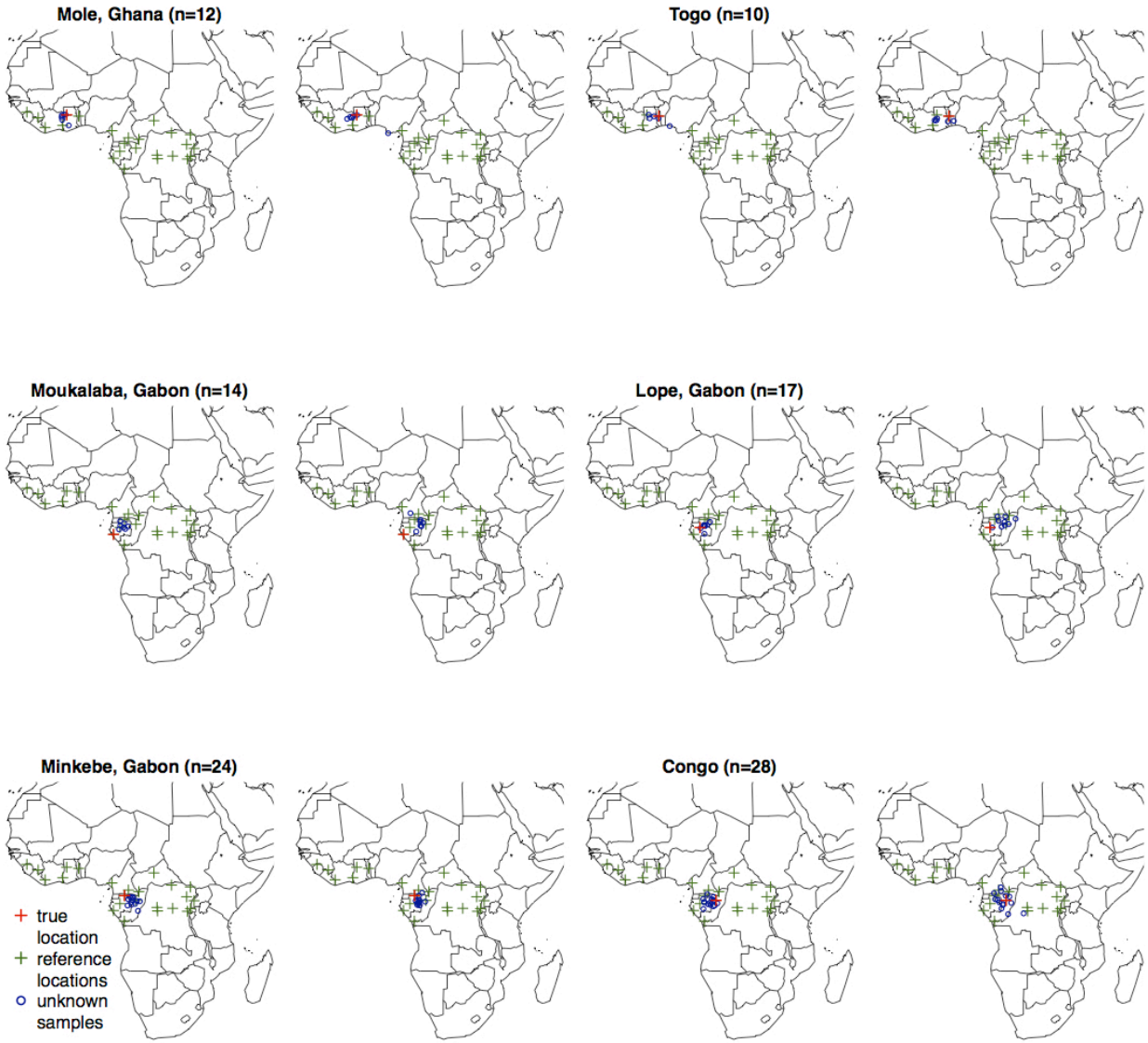


Figure 4A


Mole, Ghana (n=12)
Togo (n=10)
Moukalaba, Gabon (n=14)
Lope, Gabon (n=17)
Minkebe, Gabon (n=24)
Congo (n=28)
true location
reference locations
unknown samples

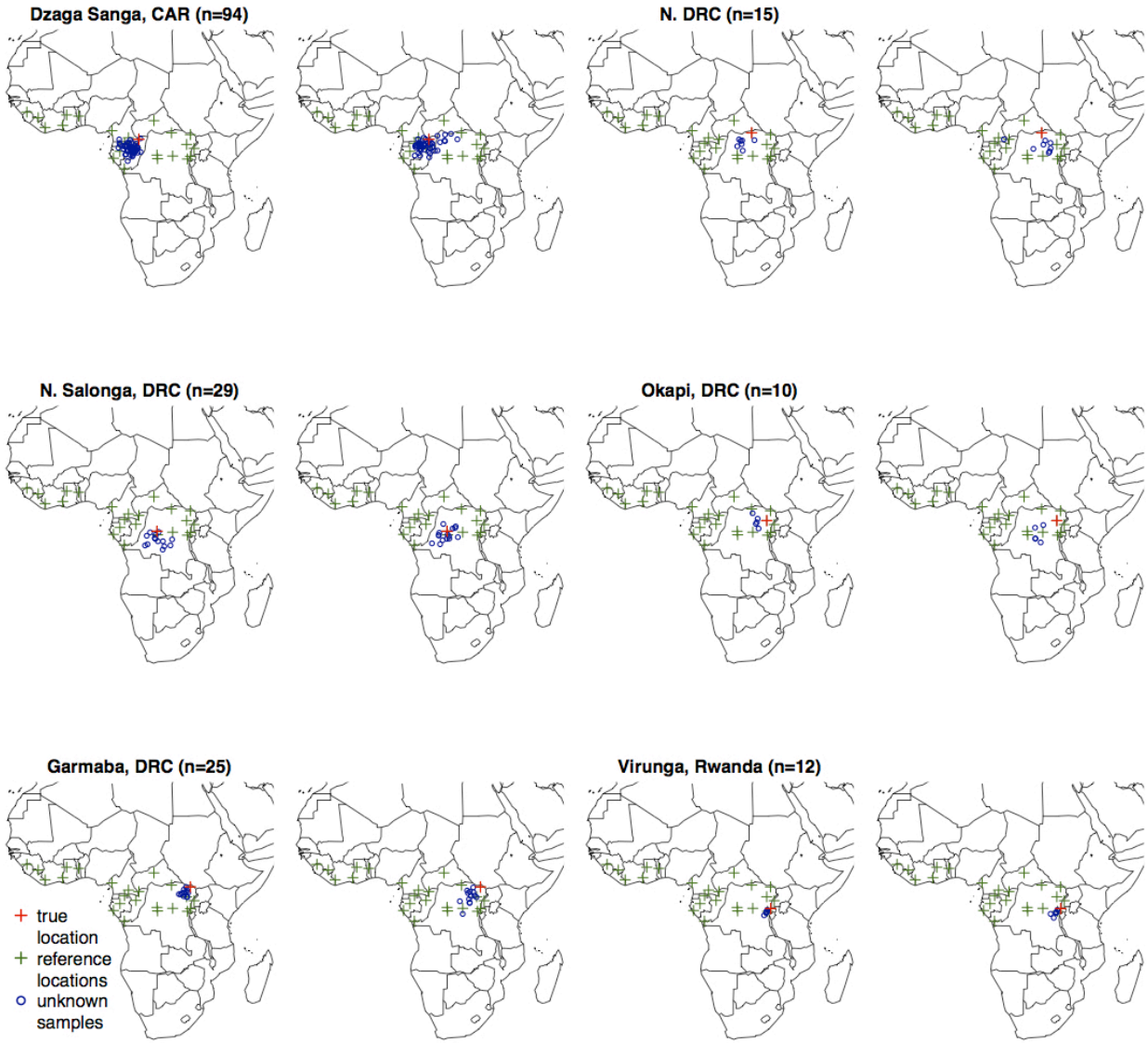
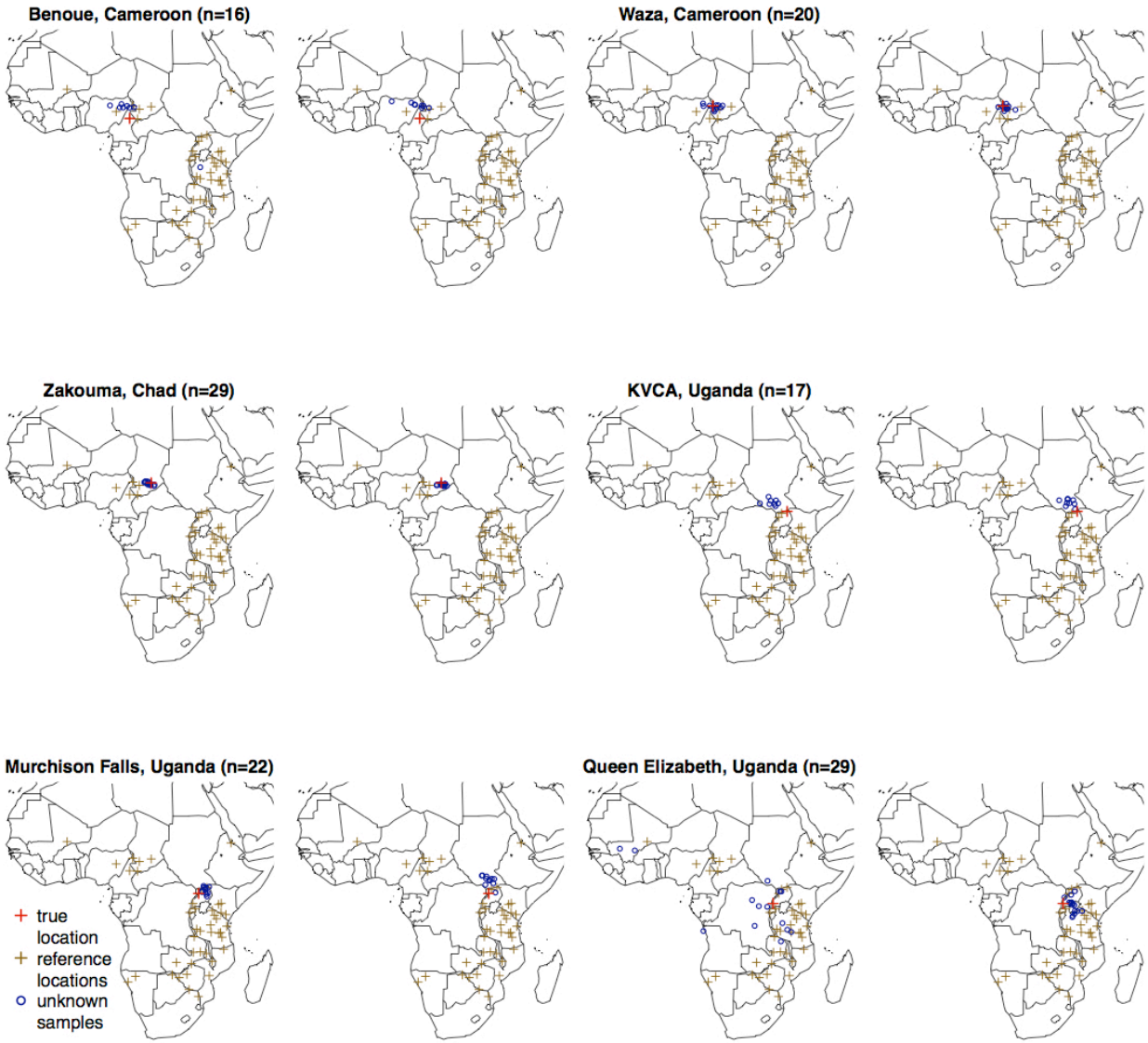
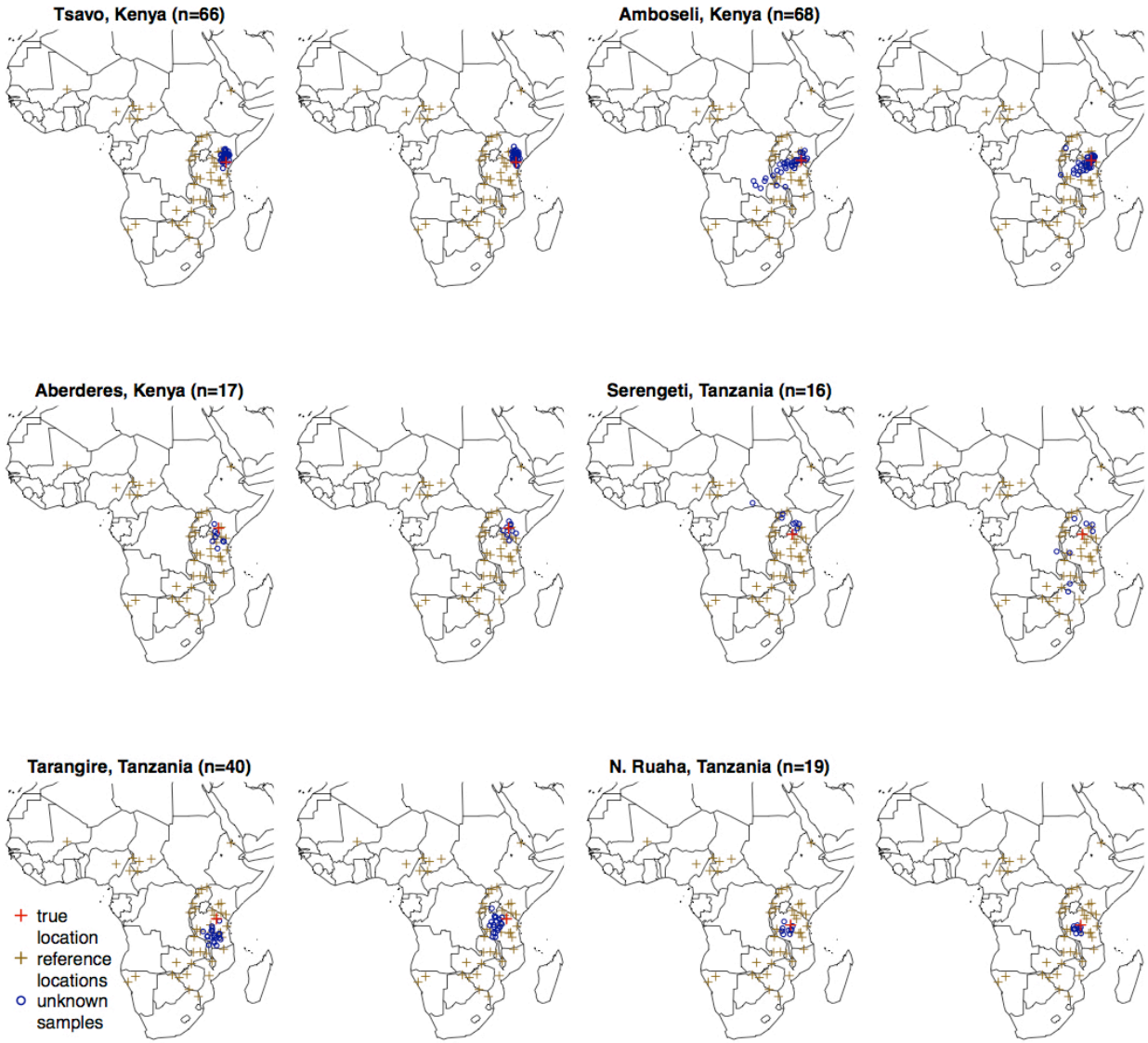
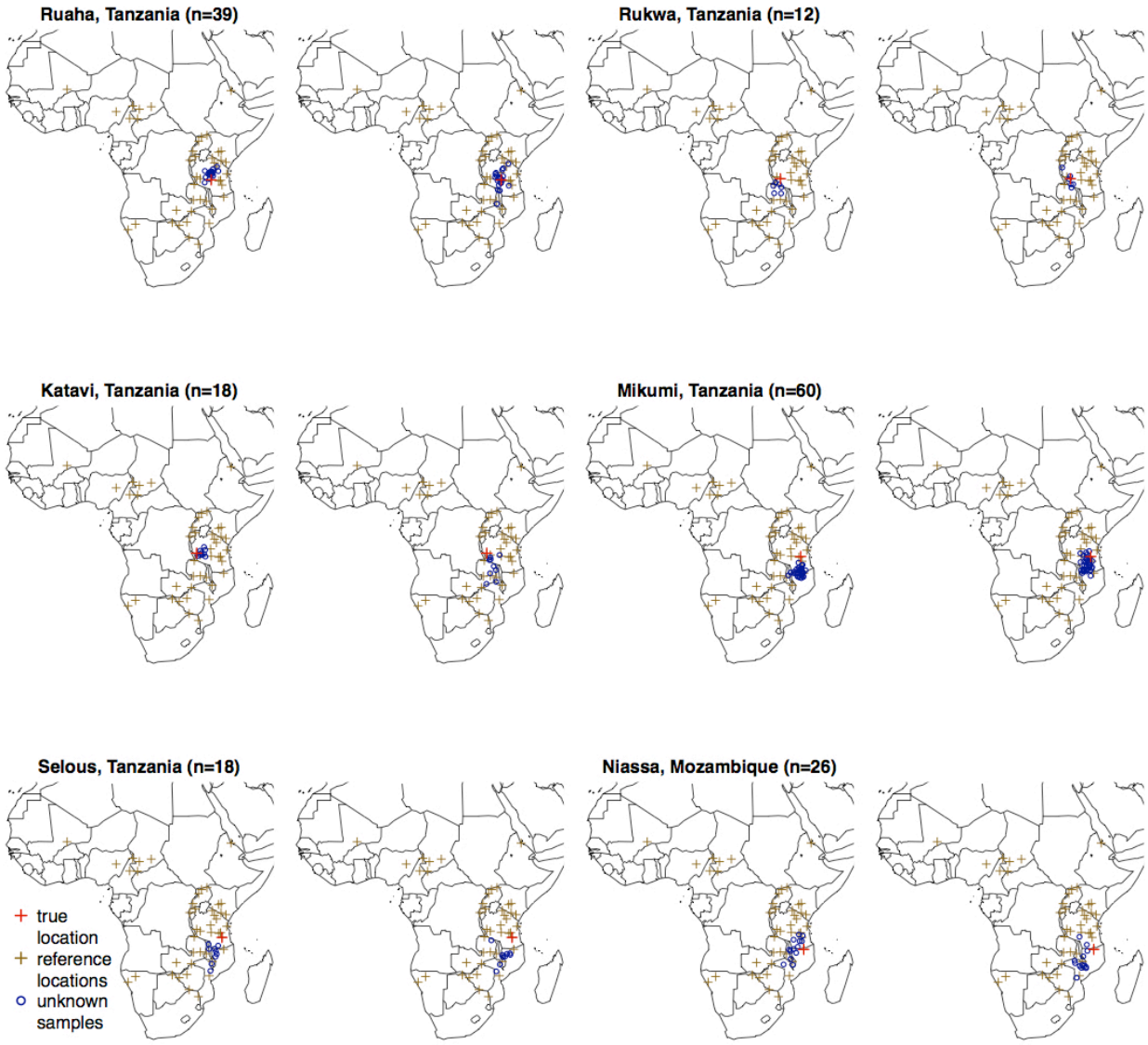
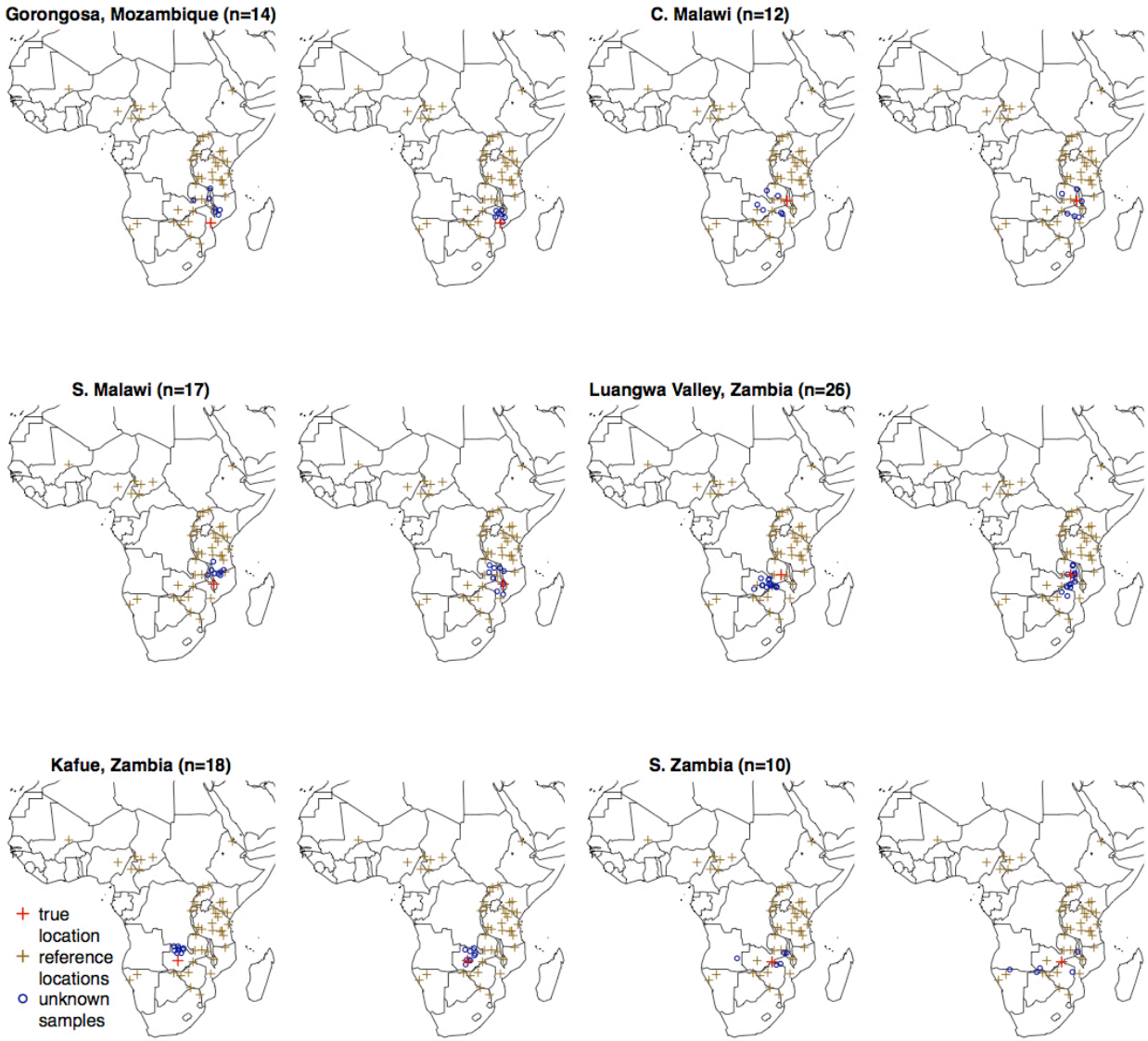

Dzaga Sanga, CAR (n=94)
N. DRC (n=15)
N. Salonga, DRC (n=29)
Okapi, DRC (n=10)
Garmaba, DRC (n=25)
Virunga, Rwanda (n=12)
true location
reference locations
unknown samples

Figure 4B


Benoue, Cameroon (n=16)
Waza, Cameroon (n=20)
Zakouma, Chad (n=29)
KVCA, Uganda (n=17)
Murchison Falls, Uganda (n=22)
Queen Elizabeth, Uganda (n=29)
true location
reference locations
unknown samples


Tsavo, Kenya (n=66)
Amboseli, Kenya (n=68)
Aberderes, Kenya (n=17)
Serengeti, Tanzania (n=16)
Tarangire, Tanzania (n=40)
N. Ruaha, Tanzania (n=19)
true location
reference locations
unknown samples


Ruaha, Tanzania (n=39)
Rukwa, Tanzania (n=12)
Katavi, Tanzania (n=18)
Mikumi, Tanzania (n=60)
Selous, Tanzania (n=18)
Niassa, Mozambique (n=26)
true location
reference locations
unknown samples


Gorongosa, Mozambique (n=14)
C. Malawi (n=12)
S. Malawi (n=17)
Luangwa Valley, Zambia (n=26)
Kafue, Zambia (n=18)
S. Zambia (n=10)
true location
reference locations
unknown samples

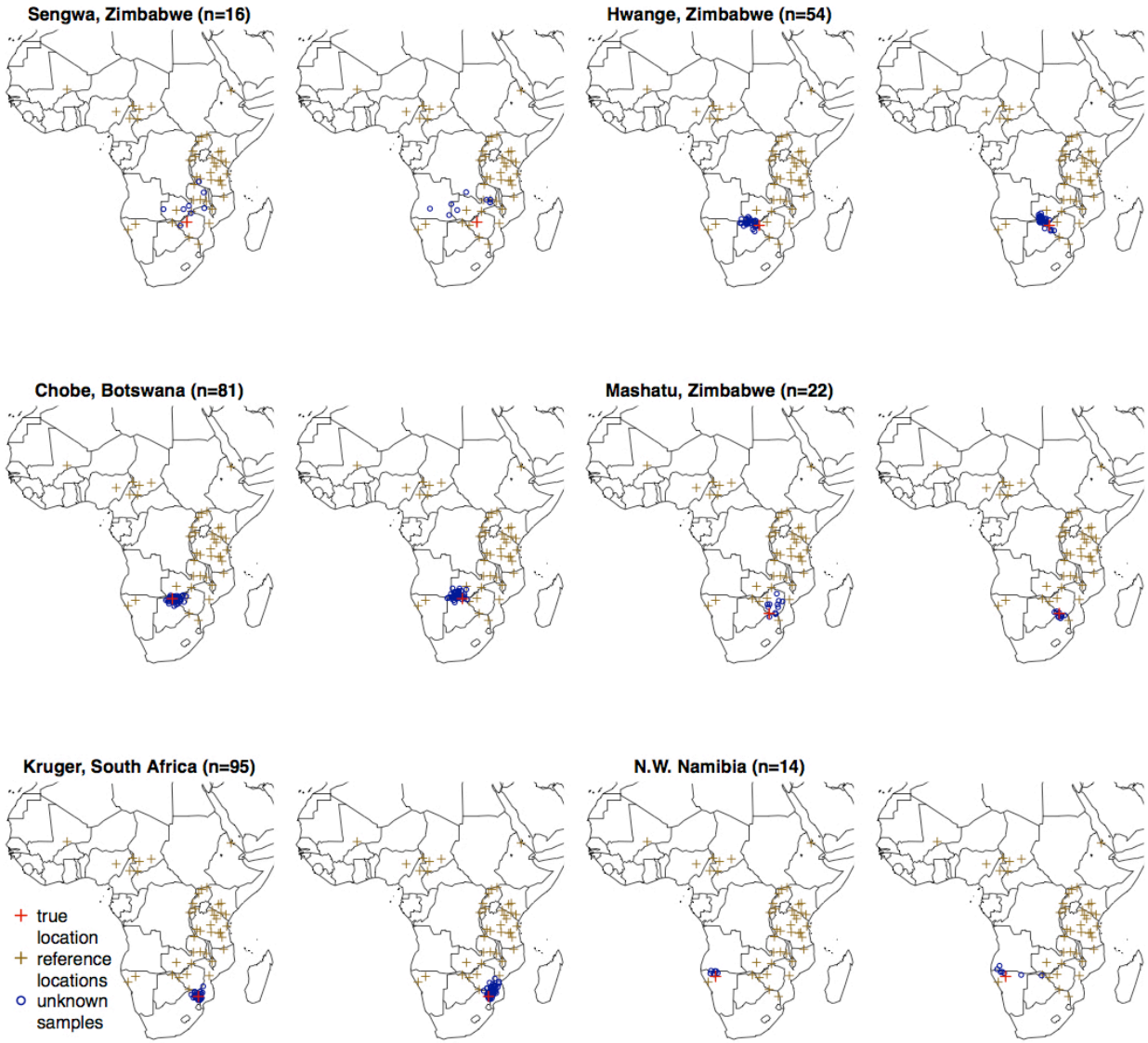
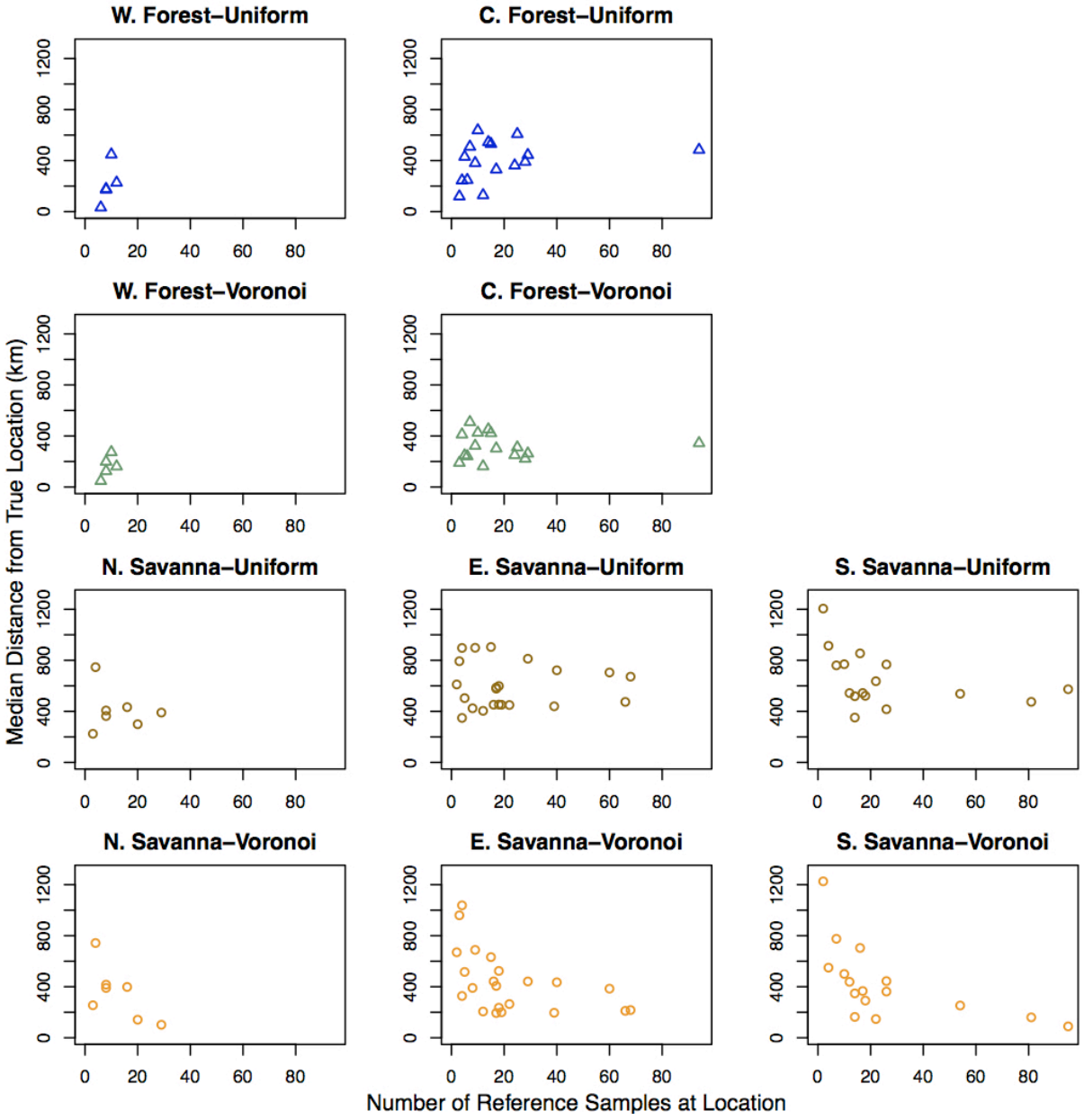

Sengwa, Zimbabwe (n=16)
Hwange, Zimbabwe (n=54)
Chobe, Botswana (n=81)
Mashatu, Zimbabwe (n=22)
Kruger, South Africa (n=95)
N.W. Namibia (n=14)
true location
reference locations
unknown samples

Figure S5.


W. Forest–Uniform
C. Forest–Uniform
W. Forest–Voronoi
C. Forest–Voronoi
N. Savanna–Uniform
E. Savanna–Uniform
S. Savanna–Uniform
N. Savanna–Voronoi
E. Savanna–Voronoi
S. Savanna–Voronoi
Median Distance from True Location (km)
Number of Reference Samples at Location
0
400
800
1200
0
20
40
60
80