Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban

Samuel K. Wasser, Celia Mailand, Rebecca Booth, Benezeth Mutayoba, Emily Kisamo, Bill Clark, and Matthew Stephens

PNAS published online Feb 26, 2007;
doi:10.1073/pnas.0609714104

Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban

Samuel K. Wasser*†, Celia Mailand*, Rebecca Booth*, Benezeth Mutayoba‡, Emily Kisamo§, Bill Clark¶, and Matthew Stephens‖**

*Center for Conservation Biology, Department of Biology, and ‖Department of Statistics, University of Washington, Seattle, WA 98195; ‡Faculty of Veterinary Medicine, Department of Veterinary Physiology, Biochemistry, Pharmacology, and Toxicology, Sokoine University of Agriculture, Morogoro, Tanzania; §Lusaka Agreement Task Force, Nairobi, Kenya; and ¶Interpol Working Group on Wildlife Crime and Department of Law Enforcement, Israel Nature and Parks Authority, Jerusalem 95463, Israel

Edited by John C. Avise, University of California, Irvine, CA, and approved December 26, 2006 (received for review November 2, 2006)

The illegal ivory trade recently intensified to the highest levels ever reported. Policing this trafficking has been hampered by the inability to reliably determine geographic origin of contraband ivory. Ivory can be smuggled across multiple international borders and along numerous trade routes, making poaching hotspots and potential trade routes difficult to identify. This fluidity also makes it difficult to refute a country's denial of poaching problems. We extend an innovative DNA assignment method to determine the geographic origin(s) of large elephant ivory seizures. A Voronoi tessellation method is used that utilizes genetic similarities across tusks to simultaneously infer the origin of multiple samples that could have one or more common origin(s). We show that this joint analysis performs better than sample-by-sample methods in assigning sample clusters of known origin. The joint method is then used to infer the geographic origin of the largest ivory seizure since the 1989 ivory trade ban. Wildlife authorities initially suspected that this ivory came from multiple locations across forest and savanna Africa. However, we show that the ivory was entirely from savanna elephants, most probably originating from a narrow east-to-west band of southern Africa, centered on Zambia. These findings enabled law enforcement to focus their investigation to a smaller area and fewer trade routes and led to changes within the Zambian government to improve antipoaching efforts. Such outcomes demonstrate the potential of genetic analyses to help combat the expanding wildlife trade by identifying origin(s) of large seizures of contraband ivory. Broader applications to wildlife trade are discussed.

African elephant | forensics | *Loxodonta africana* | DNA assignments | poaching

The illegal trade in elephant ivory has once again escalated to the devastating levels that occurred before the 1989 Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) ivory trade ban (1–5). Between August 2005 and August 2006, there have been 12 major seizures of African elephant ivory being shipped to the Far East, totaling 23,461 kg, plus 91 unweighed tusks. Most of this ivory was deemed to be from freshly killed elephants (B.C., unpublished observation). It is commonly assumed that customs intercepts 10% of all contraband (e.g., drugs, weapons, pirated compact discs). We conservatively assume that this percentage is also the case for ivory; most enforcement agencies do not "target" ivory as they do drugs or weapons, and technological advances (such as drug scanners and detection dogs) do not help with interception of contraband ivory. Thus, the above 23,461 kg should correspond to 234,610 kg of smuggled ivory from ≈23,000 elephants killed this past year. Knowing the origin of ivory in such large seizures enhances understanding of where elephants are being slaughtered and routes by which the contraband ivory is smuggled. Law-enforcement efforts could be fruitfully focused with such information. It also creates accountability that compels nations to be more responsive to poaching in their country. We previously described a method to infer the geographic origin of individual samples of African elephant DNA (6). Here, we extend the approach to multiple samples and apply this method to infer the origin of the largest seizure of contraband ivory since the 1989 ivory trade ban (the second largest seizure in the entire history of the trade).

In late June 2002, an investigative team consisting of officers from the Zambia Wildlife Authority, the Lusaka Agreement Task Force, and the Anti-Corruption Bureau of Malawi uncovered vital information concerning the shipment of a 20-ft container packed with >6.5 tons of contraband elephant ivory in Malawi, destined for the Far East. (Based on the above assumptions, this would have resulted from poaching of between 3,000 and 6,500 elephants.) The container had been shipped via South Africa to Singapore, where it was seized later that month. The seizure contained 532 tusks of widely diverse sizes and weights. The average weight of tusks was >11 kg, substantially larger than the average tusk in the current ivory trade. The seizure also contained 42,120 "hankos," believed to have been manufactured in Malawi. Hankos are round ivory cylinders, ≈6.5 cm in length and 1.5–2 cm in diameter, cut from the solid portion of the tusk. Some Asian communities carve their personal seal on the end of these cylinders to be used as a prestigious stamp (7). The hankos alone in this shipment were worth an estimated $8.4 million (U.S.), and represented >20% of Japan's annual hanko trade (B.C., unpublished observation). The enormous size of this consignment indicates the existence of an elaborate network in the Far East that is capable, with a single delivery, to receive and launder tens of thousands of hankos and hundreds of tusks into existing legal markets.

Investigative work revealed that the ivory had been carried from Zambia into Malawi in small lots, before shipping, but it was unknown whether the ivory came from Zambian elephants. Our analysis tested two broad competing hypotheses for the origin of the seized ivory:

Hypothesis 1. The ivory originated from within, or in close proximity to, Zambia and/or Malawi, the original shipping locale. This hypothesis would require minimal preshipment transport (smuggling), but the size of the seizure would suggest

Author contributions: S.K.W. and M.S. contributed equally to this work; S.K.W., E.K., and B.C. designed research; S.K.W. and B.M. performed research; C.M., R.B., and M.S. analyzed data; and S.K.W., B.C., and M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Freely available online through the PNAS open access option.

Abbreviations: CITES, Convention on International Trade in Endangered Species of Wild Fauna and Flora; MCMC, Markov Chain Monte Carlo.

†To whom correspondence should be addressed. E-mail: wassers@u.washington.edu.

**Present address: Departments of Human Genetics and Statistics, University of Chicago, Chicago, IL 60637.

that poaching intensity in this region was substantially greater than previously believed or acknowledged.

Hypothesis 2. The ivory originated from numerous locations across forest and savanna Africa, with stockpiles smuggled into Malawi before shipping. This hypothesis, which suggests the existence of a relatively sophisticated and widespread organizational network, was supported by several factors, including the large volume of the shipment, the considerable mean and variation in tusk size, and extensive poaching in the nearby Democratic Republic of Congo[††] and Selous Game Reserve in Tanzania.

Results

We selected 67 of the 532 tusks for DNA analysis, using a stratified sampling scheme aimed at maximizing the chances of acquiring tusks from multiple locations (see *Materials and Methods*). Amplification success varied greatly across samples; a total of 13 tusks had all 16 loci amplify successfully, and 23 tusks had at least 14 loci amplify successfully, whereas 18 tusks had no loci amplify successfully. In total, 37 tusks (55%) amplified at seven or more loci and were included in the subsequent assignment analysis [the cutoff of seven loci being chosen for consistency with the way the reference database was assembled (6)]. Among these 37 samples, the average number of successful loci was 13.5. Hankos were excluded from these analyses because initial attempts to amplify DNA from hankos were unsuccessful. Hanko samples are derived from the core of the tusk and were subsequently found to require a decalcification step before their extraction; analyses of the hankos are ongoing.

DNA obtained from the tusks was compared with a reference database of DNA samples of known geographic origin. The reference data were from Wasser *et al.* (6), augmented with 165 samples from Zambia, Malawi, and Southern Tanzania. The combined samples provided an updated reference database of 525 samples (see *Materials and Methods*). Initial comparison of alleles obtained from each tusk against reference allele frequency distributions for forest vs. savanna elephants suggested that all of the tusks were most likely derived from savanna elephants [likelihood ratios in favor of savanna origin, computed as in Wasser *et al.* (6), ranged from 2.5×10^4 to 9.1×10^{10}].

We developed a statistical assignment method to infer the most likely savanna locations of the sampled tusks. Existing assignment methods estimate the likely source of each tusk independently, assuming the tusks were independently and uniformly sampled from some set of possible sources. This assumption is problematic here because it implies that the tusks likely originated from a wide range of locations, essentially ignoring the possibility that they came from a restricted region (hypothesis 1). Our approach (see *Materials and Methods*) extends the smoothed continuous assignment method for individual DNA samples from Wasser *et al.* (6) to analyze multiple tusks simultaneously, allowing that they may have arisen either from a wide range of locations or from one (or a few) narrow geographic region(s).

Fig. 1 *A–D* illustrates the improved performance that can be achieved by analyzing multiple samples simultaneously rather than one sample at a time. Specifically, the figure compares results from our approach, which jointly analyzes multiple samples, with results from sample-by-sample analysis using the method described in ref. 6. We applied both methods to groups of samples known to originate from Malawi ($n = 18$) (Fig. 1*A*), Zambia ($n = 29$) (Fig. 1*B*), and the Selous Game Reserve in Southern Tanzania ($n = 12$) (Fig. 1*C*), with each analysis constituting a random sample of half of the samples available from its respective origin, and to a group of samples from

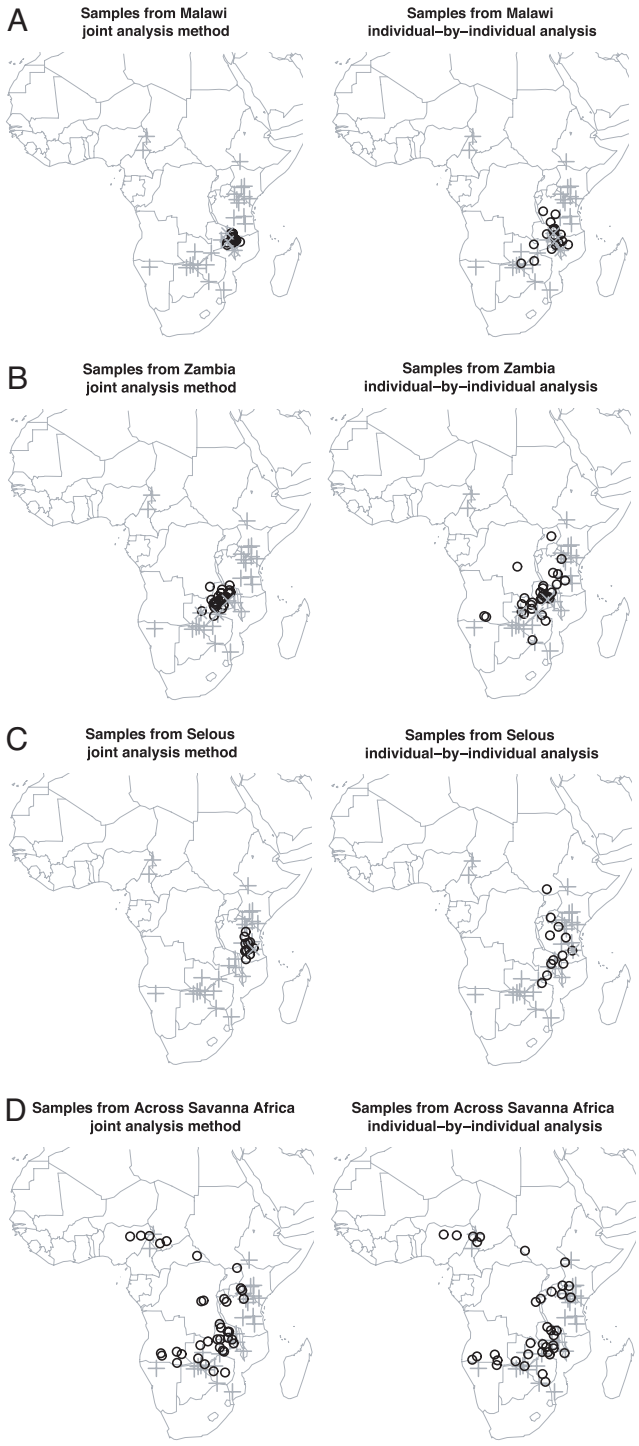

Fig. 1. Comparison of results from the new assignment method for jointly analyzing multiple samples (*Left*) with those obtained by independently analyzing each sample by using the assignment method from Wasser *et al.* (6) (*Right*). Results obtained for a batch of samples of known origin from Malawi (*A*), Zambia (*B*), and Selous Game Reserve in Tanzania (*C*) and for dung and tissue samples originating from across savanna Africa (*D*) are shown. Circles show the estimated location of origin of each sample, whereas crosses indicate locations of reference samples from savanna habitats used to make the assignments. In *A–C*, ×s are used to indicate the actual locations of the samples of known origin.

[††]Mubalama, L. (2005) Rapport sur L'Enquete du Marche D'Ivoire la ville de Kinshasa, March 9–19, 2005, Wildlife Conservation Society and Monitoring of Illegal Killing of Elephants, Kinshasa, Democratic Republic of Congo.

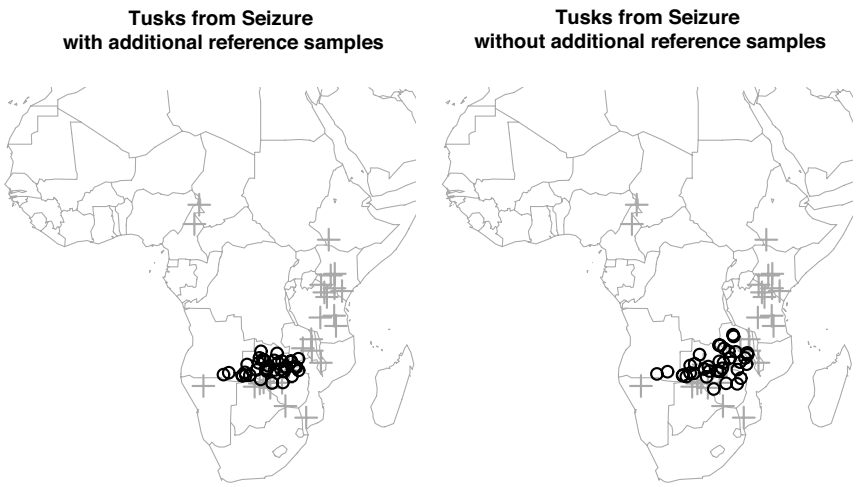

Fig. 2. Assignment results for 37 tusks from the Singapore seizure. The estimated locations of origin (circles) of the 37 tusks analyzed are shown. (*Left*) Results using the additional reference samples from Zambia, Malawi, and Selous. (*Right*) Results without these additional reference samples. Crosses are the same as in Fig. 1.

numerous locations scattered throughout savanna Africa (n = 37, chosen to match the number of tusks analyzed from the Singapore seizure) (Fig. 1*D*). The remaining halves of samples from Zambia, Malawi, and Selous contributed to the reference samples for these analyses. When applied to a batch of samples originating from a limited geographical region (Fig. 1 *A–C*), our joint analysis method is able to recognize this fact and produces estimated locations of origin that are both accurate and compact. In contrast, estimates from independent analysis of each sample, although centered on approximately the correct location, are considerably more diffuse and tend to (wrongly) suggest that the samples came from a relatively wide geographic region. Conversely, when applied to samples that actually originated from a wide geographic region (Fig. 1*D*), our joint analysis method is also able to deduce this from the data and produces estimated locations of origin that are very similar to those from the sample-by-sample analysis.

We applied the new joint analysis method to 37 tusks acquired from the Singapore seizure. The results (Fig. 2 *Left*) suggest that the tusks originated from a relatively restricted part of southern Africa, concentrated near Zambia, lending support to hypothesis 1. The tusks were genotyped on the same platform and at the same time as the reference samples from Malawi, Selous, and Zambia and at a different time and platform than the majority of the other reference samples from East and Savanna Africa. We therefore checked to determine whether our results were not unduly affected by unidentified systematic differences between the way different reference samples were treated, by reanalyzing the tusk DNA without the additional reference samples from Zambia, Malawi, and Selous. The results (Fig. 2 *Right*) were similar to those obtained with the additional reference samples, with estimated tusk origins being slightly more diffuse and centered slightly farther south.

Discussion

Using DNA, it is possible to determine, with near 100% accuracy, whether an individual sample originated from a savanna or forest elephant (6). The DNA from all of the tusks that we examined from the Singapore seizure pointed to a savanna origin for these samples. This simple inference alone immediately rules out many countries that are habitat for forest elephants (*Loxodonta cyclotis*); it also lends some support to the hypothesis that the tusks may have originated from a restricted geographic region rather than from a pool of many stockpiles from across the continent. More sophisticated analytic methods, able to accurately determine the likely geographic origin of DNA samples on a finer scale, point to a relatively narrow band of Southern Africa, centered on Zambia, as the likely source of tusks in this seizure. The estimated locations of origin for the tusks spread east and west from Zambia and may include regions of Mozambique and savanna Angola from which no reference samples are yet available. Reference samples from these locations could increase the precision of these estimates and help to confirm or rule out these countries as possible contributors to the seizure.

The 37 tusks analyzed here represent a subset of tusks that produced the most complete genotype data. Although visual inspection revealed no obvious systematic differences between these tusks and others that failed to yield such complete data (see *Materials and Methods*), it is difficult to entirely rule out the possibility that the seizure could contain some tusks of different origin that failed to produce good genotype data.

These caveats notwithstanding, the analysis of available DNA data from these samples has greatly facilitated law-enforcement efforts. As described in hypotheses 1 and 2, authorities strongly suspected that this ivory had multiple origins, including forest habitat. Our results caused law enforcement to substantially narrow the area of origin and the trade routes being investigated.

These results also had a number of consequences for Zambia. The seizure immediately followed Zambia's application to CITES for a one-off sale of their ivory stockpiles at COP12 (Conference of the Parties). That application maintained that only 135 elephants were known to have been illegally killed in Zambia during the previous 10 years, woefully shy of the 3,000–6,500 elephants we estimate to have been killed in Zambia surrounding the seizure, let alone during that entire 10-year period. Subsequent to being informed of our findings, the Zambian government replaced its director of wildlife and began imposing significantly harsher sentences for convicted ivory traffickers in its courts. However, one still has to wonder whether this will be enough.

Virtually no one has been prosecuted for this case. Moreover, just 3 years after the Singapore seizure, when we were in the thick of our DNA analyses, another 6 tons of ivory was seized in the Philippines en route from Zambia. (That ivory was subsequently stolen from the warehouse that Philippine customs had contracted to hold the contraband.) This begs the question: How can a poor country like Zambia, with only token international assistance, have the physical capacity to act effectively against

criminals exploiting the dynamic market demands of the financially robust Far East?

Wildlife trade represents a serious and growing area of organized crime that can irreparably damage a country's ecosystems and economy and has demonstrable links to other serious crime. The illegal ivory market exemplifies this. Elephants are a keystone species, whose loss significantly alters natural habitat. The ivory trade has corresponded with massive declines in elephant numbers (>50% continent-wide and up to 90% in some areas), including areas where habitat loss (the other most likely cause of decline) has remained unchanged (8–11). Moreover, the illegal trade this year has seen its largest increases ever, based on marked increases in seizures without any commensurate increases in the capacities of the seizing agencies (B.C., unpublished observation). Much of this increase in trade is being driven by wholesale prices of high-quality ivory in China and Japan (12), which have risen from $100 per kilogram in the late 1990s to $200 per kilogram by 2004 to a now staggering $750 per kilogram (B.C., unpublished observation). This disproportionately large 3.5-fold rise in the past 2 years has raised concern that commodity speculators may be buying up much of the ivory. Certainly, these trends suggest that the market is being heavily stimulated, adding to current fears that China's growing demand for illegal ivory could jeopardize elephants throughout Africa and Asia (5, ††).

Given that syndicated ivory crime has reached such international scale, we suggest that the most effective way to combat this trade is to prevent the ivory from ever entering the international market. Genetically tracking the origin of large ivory seizures can help by identifying poaching hotspots, focusing urgently needed policing of elephant poaching and associated trafficking in contraband ivory. This approach places emphasis on saving elephants before they are killed. By identifying common patterns among large seizures, such as homogeneity of origin and proximity to original shipping locale, our methods could also highlight likely smuggling routes (e.g., major roads, train routes, or nearby ports) and suggest how illegal ivory is being moved to global markets outside Africa. These effects should also increase tusk seizure rates, further helping to stop the trade before it leaves Africa. Strategic changes in these smuggling patterns over time could also be detected, as could changes in the quantity and distribution of ivory from specific locales in the world's major ivory markets. Monitoring such changes, coincident with CITES trade decisions, could provide critically needed tools to determine whether sanctioned sales influence poaching rates across the continent.

Although our methods can enhance the effectiveness of law enforcement in wildlife trade, what is really needed is to combine this with a major reinfusion of law-enforcement aid at the scale that coincided with the 1989 ivory ban. For this reinfusion to occur, industrialized nations need to be reeducated about the seriousness of the poaching problem to encourage their governments to once again provide this needed law-enforcement support. The United Nations has declared many of Africa's natural resources to be "World Heritage," and the rest of the world needs to help protect this shared heritage. To ensure that such aid is not endless, law-enforcement aid needs to be coupled with education aimed at reducing demand in the Far East and at engendering respect for natural resources in Africa. Improved management is also needed in Africa to restore the historical abilities of elephants to self-regulate their population sizes and reduce elephant/human conflict. Ironically, stopping poaching may help reduce such conflict, if elephants can once again be made to feel safe enough to remain in protected areas (13). Stopping poaching will also prevent loss of tourism in wildlife-rich countries, along with the disproportionately large amounts of foreign currency it generates.

The international community virtually stopped ivory poaching once (14), and it can stop it again. The enhanced law-enforcement effort that coincided with the 1989 ban dramatically suppressed the illegal ivory trade. However, believing that the problem was solved, western aid was largely withdrawn by 1993. Law enforcement rapidly declined in poor African countries, and poaching began to steadily increase all over again (14). A more comprehensive approach is needed this time, one that combines law enforcement with DNA analyses, education, and improved management. We have to act now, before it is too late. We hope that the results of this study will encourage such timely conservation efforts, thereby helping to curb a criminal trade that is once again imperiling elephants.

We also believe that these techniques can prove useful for other species that are substantially represented in the wildlife trade. The ability to acquire DNA from feces, coupled with new methods that markedly enhance fecal sampling rates over large remote areas (15), makes this approach highly feasible for a diverse array of at-risk species.

Materials and Methods

Additional Reference Samples. We augmented the database of 399 DNA samples of known origin from Wasser *et al.* (6) with 165 dung samples for DNA from Malawi ($n = 40$), Zambia ($n = 58$), and the Selous Game Reserve in southern Tanzania ($n = 67$). To minimize chances of sampling elephants from the same family group, no two samples were collected within 1 km of one another. This method proved highly effective: all protected areas in Zambia and Malawi were sampled in just 2 weeks, and the Selous was sampled in <1 month. At the time of collection, dung samples were mixed with a gloved hand, and then ≈25 g of the mixed sample was placed in 35 ml of 20% DMSO in TNE (Tris base, NaCl, Na_2–EDTA) buffer for storage and shipped to our laboratory in the United States (with the appropriate United States Department of Agriculture import permits).

DNA Extraction and Amplification. DNA was extracted and amplified from the dung samples as described by Wasser *et al.* (6). Amplification success across all 16 loci was 92% for Zambia and Malawi and 51% for Selous. Sample freshness was more varied in the Selous, resulting in bimodal amplification success, with one-third having amplification success >80%. For consistency with the way the existing reference database was assembled (6) and to reduce potential problems caused by low-quality DNA, we ignored samples where fewer than seven loci amplified successfully. Data from the 126 samples where at least seven loci amplified were added to the database of 399 reference samples from Wasser *et al.* (6) to create an updated database of 525 reference samples.

Ivory Sampling. In selecting tusks for analysis, we used a stratified sampling scheme aimed at maximizing chances of acquiring tusks from multiple locations. We attempted to match each tusk to its pair to avoid duplicate sampling of a single individual. Paired tusks were then grouped by external markings. Many tusks were similarly colored (as although buried in the same soil). Some tusks also had similar writing on them (e.g., YOKOHOMA or ALA), suggesting multiple contributors to the consignment. We randomly selected equal numbers from each group, taking only one sample from any matched pair. An ≈10-cm^2 piece was cut from the base of each of the 127 tusks and shipped to our lab for DNA analysis, with all appropriate CITES permits. Sixty-seven of these tusks were genotyped, of which 37 yielded sufficient genotypes (i.e., both alleles confirmed for at least seven loci) to be included in the analysis. These 37 tusks included multiple representatives from each of the above groupings, including a group specifically selected for its morphological similarities to forest elephant ivory.

DNA Extraction and Amplification from Ivory. DNA was extracted from the ivory as described by Comstock *et al.* (16) with the following modifications:

1. Several small pieces from each 10-cm^2 sample were jointly pulverized in a freezer mill (16) to increase the homogeneity of DNA in the extracted subsample.
2. Digestions used 100–160 mg of pulverized ivory in 200 μl of ATL buffer (Qiagen, Valencia, CA). The mixture was vortexed (10 sec) until all ivory powder was in solution before being placed in an incubator shaker for 18–20 h at 56°C. Tubes were vortexed at medium setting until homogenous and spun in a centrifuge (10 min at 13,000 rpm), and the supernatant was transferred to a new tube. The new tube was centrifuged (10 min at 13,000 rpm; Biofuge Pico, PP3-96 #3324; Heraeus Instruments, Osterode, Germany) to pellet any remaining ivory pieces, the supernatant was transferred to a new tube, and its final volume (n) was recorded. The equivalent amounts (n) of AL buffer (Qiagen) and ethanol were added (1:1:1 = 3n). Samples were eluted twice with 100 μl of AE buffer for 15 min per elution.
3. PCR amplification was increased from 30 to 40 cycles, and the PCR product was analyzed on an ABI 3100 genetic analyzer (Applied Biosystems, Foster City, CA).

Allele Calling. All allele sizes were called independently by two trained individuals for both dung and ivory. Intercaller discrepancies averaged ≈1% of all calls and were resolved by a third caller. Only confirmed alleles were included in the analyses. For heterozygous loci, each allele was observed in a minimum of two separate PCRs of the same sample. For homozygous loci, the allele was seen alone in a minimum of three separate PCRs. This method guarded against allelic dropout associated with low template DNA. If only one of the two alleles was confirmed, only that allele was included, and the other allele in the genotype was treated as "missing." (Confirmation of only one allele occurred in only 7% of loci. Such loci were not counted as "successful" for the purposes of requiring seven successful loci to be included in our analysis, or in the other results reported regarding numbers of successful loci.) In our statistical analyses, we assume that missing alleles are missing at random. Thus, for example, at any given location, the probability of observing a genotype with one allele equal to A and the other allele missing is taken to be proportional to the frequency of the allele A at that location.

Determination of Savanna vs. Forest Origin. As in Wasser *et al.* (6), we initially determined whether the DNA sample from each tusk was likely of savanna or forest origin by computing a likelihood ratio for savanna vs. forest origin. The forest and savanna allele frequency estimates necessary to compute these likelihood ratios were obtained from the updated reference database of 525 samples.

Assignment Method for Multiple Samples. Wasser *et al.* (6) described a method that aims to estimate the location of origin of a single DNA sample of unknown origin by comparing alleles obtained from that sample with an allele frequency map obtained from applying a spatial smoothing scheme to observations on reference samples of known origin. Here, we used similar ideas to develop a method aimed at simultaneously estimating the likely origins of multiple samples. In this approach, rather than attempting to independently estimate the origin of each individual sample, we assume that the samples of unknown origin were sampled uniformly from some region R, and attempt to simultaneously estimate both R and the locations of origin of the samples.

We adopt a flexible approach to specifying the region R, assuming it to consist of one or more polygons, which may or may not be adjacent to one another. In particular, this approach allows both for the possibility that the samples could have been drawn from a single contiguous region containing a relatively restricted part of the continent and for the possibility that the samples may have been drawn from several distinct widespread sources across the continent.

Formally, we specify a prior distribution for R and then use a Markov Chain Monte Carlo (MCMC) scheme to draw an approximate sample from the posterior distribution of R given the genotype data and to compute estimates of the location of origin of each individual tusk. These estimates are similar in spirit to those from Wasser *et al.* (6), except that in that study, the region R was presupposed to contain the whole of the African (savanna or forest) elephant range, whereas here, R is estimated from the data. By estimating R, the new method allows information to be "borrowed" across tusks to improve the precision of individual estimates, the improvement in precision being greatest when R is a relatively restricted subset of Africa.

The Prior Distribution for *R*. The prior distribution for R in our MCMC scheme can be described in two steps:

1. Randomly partition the continent into n irregular polygons (we used $n = 100$ for the results shown here) by using a process known as Voronoi tessellation [see Guillot *et al.* (17) for an example]. A Voronoi tessellation containing n polygons is created by randomly placing n points $x_1,\ldots,x_n$ uniformly on a space containing the whole of Africa; the set of points that are closer to x_i than to any other x_j form a polygon P_i, and the polygons $P_1,\ldots,P_n$ partition the space.
2. For each polygon P_i, decide whether to include P_i in R as follows. Independently for each i, include P_i in R with probability p (where p is a hyperparameter to be estimated); otherwise do not include P_i in R. We find it helpful here to introduce variables $\gamma_1,\ldots,\gamma_n$, where $\gamma_i = 1$ if P_i is included in R and where $\gamma_i = 0$ otherwise.

Note that knowing the vectors x and γ introduced above is enough to determine R, and the above two steps can be thought of as specifying the prior distribution for R by specifying distributions for x (uniform on a space containing Africa) and for γ [$\Pr(\gamma_i = 1) = p$]. We also assume a uniform prior on [0,1) for the hyperparameter p. Note that, under this prior distribution, each location has a prior probability of 0.5 of being included in R. The quantities (x, γ, p) are treated as unknown parameters and estimated by using an MCMC scheme (see below). In particular, the locations, shapes, and sizes of the polygons, determined by x, are allowed to vary in each iteration of the MCMC scheme, rather than being fixed.

The above description ignores one detail: in practice, for this analysis, we allowed R to contain only parts of Africa that fall in the range of the savanna elephants (after determining all tusks to be of savanna origin; see *Results*). Thus, we formed R as the intersection of a collection of polygons (as described above) within the savanna range.

MCMC Estimation Procedure. We developed an MCMC scheme to sample from the posterior distribution of R given the available genotype data and to estimate the location of origin of each sample. Our approach involves approximating the posterior distribution of R by discretizing the continent into a regular grid and postprocessing results from the sample-by-sample analysis described by Wasser *et al.* (6)

1. Define a 67 by 67° grid, each cell of which is 1° latitude by 1° longitude, covering the area from 17N to 50S and from 36W to 31E. We will index the cells of this grid using i and j.
2. For each sample $s = 1,\ldots S$, in the batch of samples of unknown origin, apply the method from Wasser *et al.* (6),

implemented in the software SCAT (for Smoothed and Continuous AssignmenTs; http://www.stat.washington.edu/stephens/software.html). This method provides a sample of M points, $w_{s1}. . . w_{sM}$, from the posterior distribution of the location of origin of each sample, s, assuming a uniform prior on this location across the entire savanna range. (We used $M = 900$ by applying the algorithm nine times for each sample, using nine different values for the seed of the random number generator, using the default software options that create 100 sample points per application). Let c_{sij} be the count of the number of w_{sm} that fall in grid cell (i,j). Then c_{sij}/M is an estimate of the posterior probability that sample s originated from grid cell (i,j) for a uniform prior on its location of origin, independently of the other samples in the batch. Because of the uniform prior, c_{sij}/M is also, up to an unknown constant of proportionality, an estimate of the probability of the genotype data for sample s, given that the sample arose from grid cell (i,j). [Although most cells will contain no reference samples, the continuous assignment method from Wasser *et al.* (6) uses spatial smoothing to estimate allele frequencies even in locations containing no reference samples. Thus, unlike traditional assignment approaches, the continuous assignment method allows samples to be assigned to cells for which no reference samples are available.]

3. Now implement a new MCMC scheme to sample from the posterior distribution of R given the genotype data of samples $s = 1,. . . ,S$. As described above, R is defined by the Voronoi cell points x and the indicator variables γ, whose distribution depends on a hyperparameter p. The MCMC scheme aims to sample from the posterior distribution of R by sampling from the posterior distribution of x,γ,p given the genotype data G. This posterior distribution is determined by two quantities: (*i*) the prior distribution on (x,γ,p), and (*ii*) the likelihood, $\Pr(G|x,\gamma,p)$. As described above, for the prior distributions, we assumed the elements of the vector x to be independent and identically distributed (iid), with a uniform distribution on the area covered by the 67° by 67° grid; the elements of the vector γ to be iid Bernoulli(p); and p to have a uniform prior distribution on the interval [0,1). For the likelihood, we used an approximation based on discretizing R. Specifically, for a given value of R, let $z_{ij} = 1$ if R contains the center of the grid cell (i,j) and this grid cell lies within the range of savanna elephants; otherwise, let $z_{ij} = 0$. Then $\Pr(G|R)$ is approximately proportional to $\Pi_s\ [\Sigma_{i,j} z_{ij} c_{sij}/\Sigma_{i,j} z_{ij}]$. (Here, we are assuming that all grid cells have the same area, which will not be true because of the curvature of the earth, but will be a somewhat reasonable approximation because all grid cells are close to the equator.)

With the prior and likelihood thus specified, we implement a simple random-walk Metropolis–Hastings MCMC scheme, with the following updates: (*i*) update each component of γ in turn (the proposed new value of γ_k being one minus the current value); (*ii*) update the vector x [the proposed new value being obtained by adding independent N (mean = 0, SD = 5) random variables to the latitude and longitude coordinates of each component of x, with reflecting boundaries at the limits of the grid]; (*iii*) update p [the proposed new value being obtained by adding an N (mean = 0, SD = 0.02) random variable to the current value].

4. Let $R_1,. . . ,R_N$ denote the posterior sample of R, obtained in step 3 above. This sample can then be used to obtain a sample from the posterior distribution of the location of each sample s by appropriately weighting the original sampled location points $w_{s1}. . . w_{sM}$. Specifically, we first compute iteration-specific weights for these sampled location points: at each iteration I, for each sample s, those w_{sm} that lie in R are given weight $1/N_I$, where N_I is the number of such points that lie in R; the w_{sm} that lie outside R_I are given weight 0 (so that the weights for each individual sum to 1 each iteration). Finally, each sampled location point is given an overall weight, which is the average, across all iterations, of its iteration-specific weights.

The end result is that for each individual s, we have the original sampled location points, $w_{s1}. . . w_{sM}$, and corresponding weights, $q_{s1}. . . q_{sM}$, that sum to 1. This weighted sample can be used to estimate the location of origin for each individual (e.g., we used the estimated posterior mean latitude and longitude computed from the weighted sample). This estimated location of origin takes account of the genotype information on all individuals simultaneously because all of the genotype information is used to estimate R. If, as in Fig. 1*D*, R turns out to include most of the savanna range, then all of the weights will be approximately equal (to $1/M$), and the method will produce results almost identical to a sample-by-sample analysis.

For the analyses described here, we applied the MCMC algorithm (step 3 above) 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 sufficiently long to obtain reliable results.

This work is part of a cooperative inquiry involving the Lusaka Agreement Task Force, the Zambia Wildlife Authority, the Anti-Corruption Bureau of Malawi, and the Singapore Agri-Food and Veterinary Authority, all of whom contributed samples to this study, along with B.C. (representing the Interpol Working Group on Wildlife Crime and the Israel Nature and Parks Authority). We also thank all of the people who have contributed reference samples to our elephant database in the past (6) and the World Wildlife Fund-Tanzania for contributing samples from the Selous Game Reserve to these analyses. This work was financially supported by the United States Fish and Wildlife Service African Elephant Conservation Fund and the International Fund for Animal Welfare.

1. Martin E, Stiles D (2000) *The Ivory Markets of Africa* (Save the Elephants, Nairobi).
2. Stiles D, Martin E (2001) *Pachyderm* 30:24–36.
3. Martin E, Stiles D (2003) *The Ivory Markets of East Asia* (Save the Elephants, Nairobi).
4. Martin E (2005) *Pachyderm* 39:67–76.
5. Clark B (2006) *Between the Conference of the Parties* 6:6–7.
6. Wasser S-K, Shedlock A, Comstock K-E, Ostrander E-A, Mutayoba B, Stephens M (2004) *Proc Natl Acad Sci USA* 101:14847–14852.
7. Scigliano E (2002) *Love, War, and Circuses: The Age-Old Relationship Between Elephants and Humans* (Houghton Mifflin, Boston).
8. Convention on International Trade in Endangered Species of Wild Fauna and Flora (2000) *Eleventh Meeting of the Conference of the Parties, Proposal 11.24* (Conv Int Trade Endangered Species Wild Fauna Flora, Nairobi).
9. Douglas-Hamilton I (1987) *Pachyderm* 8:1–10.
10. Douglas-Hamilton I (1988) *African Elephant Database Project: Phase Two, December* (United Nations Environment Program, Nairobi).
11. Said M-Y, Change R-N, Craig G-C, Thouless C-R, Barnes R-F-W, Dublin H-T (1995) *African Elephant Database* (International Union for Conservation of Nature and Natural Resources, Gland, Switzerland).
12. Courable M, Hurse F, Milliken T (2003) *More Ivory Than Elephants: Domestic Ivory Markets in Three West African Countries* (Traffic International, Cambridge, UK).
13. Bradshaw G-A, Schore A-N, Brown J-L, Poole J-H, Moss C-J (2005) *Nature* 433:807.
14. Dublin H-T, Milliken T, Barnes R-F-W (1995) *Four Years After the CITES Ban: Illegal Killings of Elephants, Ivory Trade and Stockpiles* (International Union for Conservation of Nature and Natural Resources, Gland, Switzerland).
15. Wasser S-K, Davenport B, Ramage E-R, Hunt K-E, Parker M, Clarke C, Stenhouse G (2004) *Can J Zool* 82:475–492.
16. Comstock K-E, Ostrander E-A, Wasser S-K (2003) *Conserv Biol* 17:1–4.
17. Guillot G, Estoup A, Mortier F, Cosson J-F (2005) *Genetics* 170:1261–1280.