Research News

The genus *Pan*: population genetics of an endangered outgroup

Pascal Gagneux

A new study of Y-chromosome variation has shed fresh light on the population history of the genus *Pan*, which includes our closest living relatives, bonobos and chimpanzees. The study confirms a great diversity in both species and suggests these species have substantially larger effective population sizes than humans. Y-chromosome lineages appear distinct between bonobos and chimpanzees, and also between the different chimpanzee subspecies. During the Pleistocene, forest fragmentation led to recurrent, transient subdivisions within regional *Pan* subspecies; did these contribute to their apparently higher effective population sizes?

In human genetics, chimpanzees are used regularly as the preferred OUTGROUP (see Glossary). To this end, genetic data from one or very few individuals usually suffice. It was not until the late 1980s that researchers started to consider larger numbers of chimpanzees in genetic studies, with the aim of understanding chimpanzee phylogeography. In 1994, pioneering work by Morin *et al.* [1] used non-invasive techniques, with DNA being extracted from the animals' shed hair, to document genetic variation in chimpanzees from known locations across Africa (Box 1). After the initial euphoria about the feasibility of genetic studies, it quickly became apparent that the quantity and quality of DNA from non-invasive samples such as shed hair, chewed fruit pulp, feces or urine was very limited. This has prevented studies of nuclear DNA other than short microsatellite loci. Recent improvements in sample collection and DNA amplification will hopefully make studies of non-repetitive nuclear DNA possible [2].

The contrast between the low genetic diversity and lack of subdivision in humans on the one hand, and the high diversity with geographical subdivision in chimpanzees on the other, was first pointed out by Ferris *et al.* [3], using mitochondrial DNA (mtDNA) sequences. The same pattern has since been confirmed at several additional genetic loci, both nuclear and mitochondrial [1,4–6].

Recent data on human Y-chromosome variation form an obvious exception to this rule, showing substantial subdivision between continents [7]. This higher subdivision of paternally inherited markers could reflect the history of repeated FOUNDER EVENTS outside Africa and low levels of male dispersal (Box 2). The pattern of male PHILOPATRY and female dispersal is rare among mammals, but it is one of the many traits shared between humans and chimpanzees.

Estimation of divergence time and effective population size in *Pan*

Recently, Stone and colleagues [8] published the first comparative study of Y-chromosome variation among humans, chimpanzees and bonobos, in an effort to compare the population genetics and history of our closest relatives with our own. The study is based on the characterization of 23 polymorphisms at ten sequence-tagged sites over 2787 bp of the non-recombining region of the Y chromosome (NRY) from humans ($N = 42$) and captive apes ($N = 109$). To investigate male and female EFFECTIVE POPULATION SIZE (N_e), the authors also obtained the sequence of the gene encoding NADH dehydrogenase subunit 2 from the mitochondrial genome of a subset of the individuals.

As with most other loci studied to date, including nuclear autosomal, X-linked, and mitochondrial, the NRY region reveals substantially higher levels of variation in chimpanzees than in humans. This is true for both *Pan troglodytes* (chimpanzees) and *Pan paniscus* (bonobos). Extrapolating from the degree of polymorphism detected over this short region of the NRY and comparing it with published data on world-wide human polymorphism across a 63-kb stretch of NRY [9], chimpanzees and bonobos appear to have six to seven times more variation than humans. Based on a coalescence analysis (Box 3), the time elapsed since the most recent common ancestor estimated for this segment of the Y chromosome in chimpanzees and bonobos is >500 000 years, compared with a mere 50 000 years for humans. None of the unique NRY HAPLOGROUPS were shared

Glossary

Effective population size (N_e): The theoretical notion of the genetic composition of a population. A number reflecting the size of an idealized population (i.e. large, with random mating, even sex ratio, and non-overlapping generation times) that is affected by genetic drift to the same extent as the population under consideration.
Founder event: The establishment of a new population by a small number of individuals from the original population. The small number of individuals leads to lower levels of genetic variation in subsequent generations.
Haplotype: A unique combination of polymorphism on a segment of DNA.
Lineage sorting: Random loss of gene lineages in sibling species after speciation, eventually causing the absence of any shared lineages between species.
Monophyletic group: A group of sequences that share a distinct common ancestor much more recently than other sequences in the sample.
NRY haplogroup: A combination of point mutations observed across ten sequence tagged sites (STS) on the non-recombining part of the Y chromosome. Each haplogroup can include different haplotypes, if studied in more detail.
Operational sex ratio: The ratio between males and females actually contributing alleles to next generations.
Outgroup: A closely related species used to define the root of a phylogeny.
Philopatry: The individual remains in their native social group for life.
Population bottleneck: A strong reduction in effective population size, causing loss of genetic diversity, which will still be evident after the population reaches a larger size.

http://tig.trends.com 0168-9525/02/$ – see front matter PII: S0168-9525(02)02695-1

Box 1. From outgroup to principal focus of interest

Historically, the taxonomy of chimpanzees has swung from inflated numbers of taxa (up to 12 genera) during early exploration, to the treatment of all chimpanzees as a single undifferentiated species in biomedical research. Such confusion is not surprising given the vast overlap of phenotypic traits between regional populations, and the substantial variability within each subspecies. Further studies have confirmed that different chimpanzee subspecies do not share mitochondrial DNA haplotypes [a–c].

Thus, in 1997, Gonder and colleagues suggested the existence of a new subspecies *Pan troglodytes vellerosus* in Eastern Nigeria and Northern Cameroun [c].

There remain large gaps in the sampling of the wild chimpanzee population, especially around three rivers that probably formed subspecific barriers: the Niger, Sanaga and Ubangui Rivers (Fig. I). The only *P.t. vellerosus* sample included in the recent Y-chromosome study [d] had a haplogroup identical to *Pan troglodytes verus* chimpanzees.

The two divergent NRY haplotypes from animals in European Zoos [d] could indicate the existence of additional subdivision among *Pan troglodytes.*

In 1969, Hill had described four subspecies of chimpanzees: *P. t. verus*, *Pan troglodytes troglodytes*, *Pan troglodytes schweinfurthii*, and the mysterious *Pan troglodytes koolakamba* [e], and a study of transferrin protein polymorphism seemed to confirm this classification [f]. Could it be that the NRY lineage Pt?1 (Fig. 2) is actually from Hill's Koolakamba lineage?

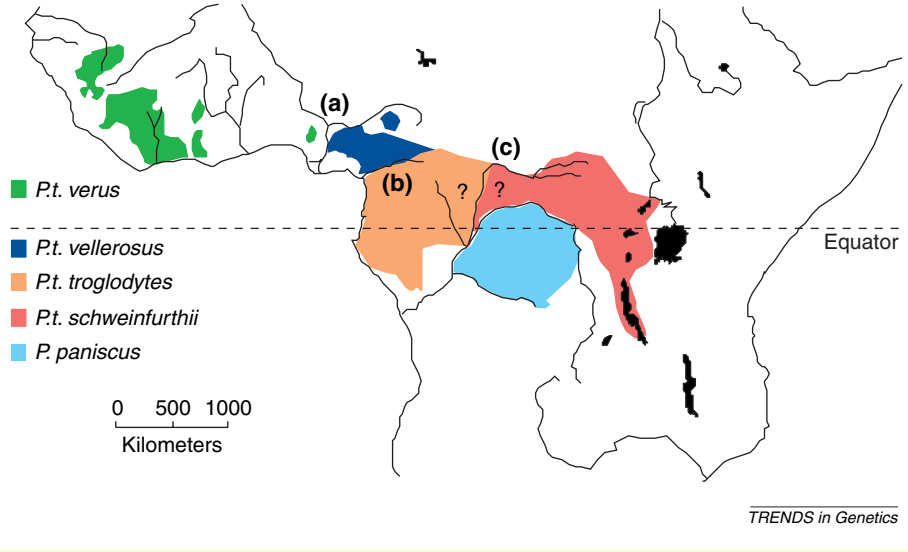

Fig. I. Map of Central Africa with ranges of chimpanzee subspecies and bonobos, and showing major rivers (potential barriers between species and subspecies): (a) the Niger River; (b) the Sanaga River and (c) the Ubangui River. Question marks indicate unsampled areas.

References

a Goldberg, T.L. and Ruvolo, M. (1997) The geographic apportionment of mitochondrial genetic diversity in east African chimpanzees, *Pan troglodytes schweinfurthii*. *Mol. Biol. Evol.* 14, 976–984

b Gonder, M.K. *et al.* (1997) A new west African chimpanzee subspecies. *Nature* 388, 337

c Gagneux, P. *et al.* (1999) Mitochondrial sequences show diverse evolutionary histories of African hominoids. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5077–5082

d Stone, A.C. *et al.* (2002) High levels of Y chromosome nucleotide diversity in the genus *Pan*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 43–48

e Hill, O.W.C. (1969) The nomenclature, taxonomy and distribution of chimpanzees. In *The Chimpanze* (Bourne, G.H., ed.), pp. 22–49, Karger

f Goodman, M. and Tashian, R.E. (1969). A geographic variation in the serum transferrin and red cell phosphoglucomutase polymorphisms of chimpanzees. *Hum. Biol.* 41, 237–249

between species or subspecies. This is in contrast to estimates from a 10-kb segment of noncoding X chromosome studied by Kaessmann *et al.* [5] and three autosomal loci studied by Deinard and Kidd [6]. Here, the time elapsed since they underwent speciation appears to have been too short for complete LINEAGE SORTING of autosomal and X-linked loci to occur between chimpanzees and bonobos.

Stone and colleagues also used coalescent analysis to estimate the divergence time between bonobos and chimpanzees at 1.8 (±0.3) million years (Fig. 1), which is intermediate between previous estimates based on other parts of the genome (930 000 years for the X-chromosome [5], and 2.5 million years

Box 2. Population genetics

Low genetic variation at neutral loci can result from inbreeding and/or small numbers of breeding individuals (e.g. POPULATION BOTTLENECK, founder event). Population size effects are usually reflected by most loci in the genome, in contrast to changes in levels of variation due to selection, which acts only on specific loci.

Genetic diversity can be measured as nucleotide diversity (π), which is based on the average nucleotide differences per site between two sequences randomly selected from a sample. Alternatively, genetic diversity can be measured as θ, which is based on the proportion of segregating sites corrected for the sample size. Under equilibrium conditions with respect to mutation and drift, both π and θ measure the neutral parameter $4N_e\mu$ for autosomal loci, $3N_e\mu$ for X-linked loci and $2N_e\mu$ for Y-chromosome and mitochondrial DNA (mtDNA) loci, where N_e is the effective population size and μ is the mutation rate.

Genetic diversity in the context of the Y chromosome refers to the number of distinct Y chromosomes and their respective frequencies in the sample. Y chromosomes are not subject to recombination (except for a small 'pseudoautosomal region') and are inherited from male to male. Each unique Y chromosome can be defined by characterizing a combination of mutations on it (a haplogroup, possibly containing several haplotypes, depending on the detail of analysis). mtDNA is only inherited from female to female (with rare exceptions) and can be used to study female population history of a species. Differences in the estimate for effective population sizes based on maternally and paternally inherited loci (mtDNA and Y chromosome, respectively) can indicate different operational sex ratios, or differences in dispersal patterns between the sexes.

Population subdivision refers to the absence or strong limitation of genetic exchange (migration) between populations of the same species. Conspecific populations of mammals that appear to have been subdivided for significant amounts of time (>100 000 of years) are often called 'subspecies'. However, there are no clear criteria for the definition of subspecies.

Box 3. Coalescence

The theory of the coalescent is an approach to the analysis of genetic variation that starts with current samples and aims to trace back in time events that took place since the most recent common ancestor of the sample. Each DNA sequence must have had an ancestor, but not all DNA sequences in previous generations left any descendants, and some left more than one descendant. Hence, all extant DNA sequences must invariably share a common ancestor. Unlike classical population genetics theory, which describes properties of entire populations, coalescent theory describes properties of samples. It develops algorithms for simulating population samples under a variety of population genetic models.

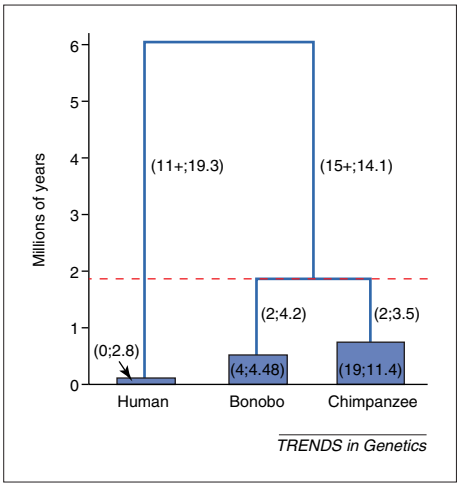

Fig. 1. Coalescence tree for NRY haplogroups of humans, bonobos and chimpanzees. Numbers represent number of mutations within species and along branches leading to hypothetical ancestors, in the form (observed; expected). The tree was calculated using the program SPLIT_TIME. Plus symbols refer to mutations that could not be assigned to one particular branch. The dotted line indicates the chimpanzee bonobo divergence time. Figure reproduced from Ref. [7], with permission of the National Academy of Sciences USA.

for β-globin and mtDNA [10,11]). Stone *et al.* conclude that, apart from the East African subspecies *Pan troglodytes schweinfurthii*, all other taxa (species and subspecies) of the genus *Pan* can be characterized by large N_e of relatively constant size, more-ancient origins and a high degree of subdivision. Significantly larger N_e for common ancestors of humans and apes have also been estimated on the basis of interspecies polymorphisms at more than 50 loci [12].

Differences in operational sex ratio

Stone *et al.* found pronounced differences between the estimates of chimpanzee male and female effective population sizes ($N_{em} = 21\,000$, $N_{ef} = 48\,000$). This was also true in bonobos, although to a lesser degree ($N_{em} = 24\,000$, $N_{ef} = 28\,000$). The adult OPERATIONAL SEX RATIO is biased towards females in most research populations of both chimpanzees and bonobos, although to varying degrees, ranging from 0.3 to almost 1 [13]. Factors reducing the operational sex ratio in chimpanzees are the variance in male reproductive success and higher mortality in males [14]. Data on reproductive success remain scant owing to the technical difficulties of obtaining quality DNA from non-invasive samples. However, there is good evidence for male–male competition in chimpanzees from behavior studies and comparative anatomy of male reproductive organs [15,16]. Confirmation of reproductive skew in males will have to await future non-invasive studies.

More subdivision than meets the eye?

All calculations of coalescence in the study by Stone *et al.* were performed under the assumption of equilibrium conditions; that is, constant large population size and random mating. In contrast to humans who exploited novel ecosystems inside and outside Africa, African great apes have remained wed to their forests. Recurrent contractions of tropical forests in Africa due to glaciation episodes during the Plio/Pleistocene (the past 5 million years) have frequently fragmented tropical forests into small refugia [17]. The question arises then of the extent to which such fragmentation caused subdivision, even within the different chimpanzee subspecies. The separation could have been transient, followed by re-unification of separated populations after re-expansion of the forests. The 21 quaternary glaciations each lasted 20 000–100 000 years, which is within the range of the coalescence time for the human NRY. Recurrent subdivision within each subspecies would have led to an inflation of the apparent N_e in the re-united population, as subdivision accumulates total variation due to local drift of various sub-populations [18]. Detecting current subdivision within regional populations of chimpanzees is only possible by analyzing samples from known geographical locations. With one exception, the few existing studies on gene flow in chimpanzees have not detected such subdivision [1,19,20].

Box 4. Estimated chimpanzee populations

In 1900 [a]
Throughout Africa: >2 million

Present day

Wild populations [a]	
Throughout Africa:	100 000–200 000
West Africa, *Pan troglodytes verus*:	12 000
Central Africa, *Pan troglodytes troglodytes*:	80 000
Nigeria, Cameroun, *Pan troglodytes vellerosus*:	No estimate
East Africa, *Pan troglodytes schweinfurthii*:	14 000
Democratic Republic of Congo, *Pan paniscus*:	15 000
Captive chimpanzees	
Sanctuaries and orphanages in Africa [a]: >	100
Outside Africa [b–d]:	3200
North America:	2500
Europe:	700
Asia (Japan and Taiwan):	500

References

a Oates, J.F. (1996) *Status Survey and Conservation Action Plan* (revised edn), The World Conservation Union (IUCN)
b National Research Council Staff (1997) *Chimpanzees in Research*, National Academic Press (http:www.isis.org/abstracts/mammalia.asp)
c Wilde, J. *et al.* (1994) A census for captive primates in Europe and North Africa. *Primate Report* 38, 1–81
d Matsuzawa, T. (1998) Chimpanzees in Japan need help. *Pan Africa News* 5, 5

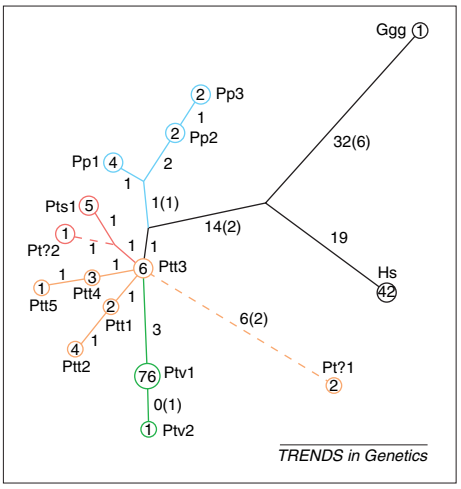

Fig. 2. Maximum parsimony network for NRY haplogroups. Ggg, *Gorilla gorilla*; Hs, *Homo sapiens*; Pp, *Pan paniscus*; Ptv, *Pan troglodytes verus* (green); Ptt, *Pan troglodytes troglodytes* (yellow); Pts, *Pan troglodytes schweinfurthii* (red); Pt?, *Pan troglodytes* unknown subspecies. Numbers next to lines represent point mutation events followed by the number of indels in parentheses. Dotted lines indicate uncertainty of subspecies origin. Figure reproduced from Ref. [7], with permission of the National Academy of Sciences USA.

Ex situ populations of captive apes

Before the enactment of the Convention on International Trade in Endangered Species (CITES), chimpanzees were caught in the wild, and shipped to research and entertainment facilities in the USA, Japan and Europe. In the process, animals from different geographical regions were mixed in captive populations. In Africa, several rescue centers and sanctuaries are caring for chimpanzees (Box 4). Many captive animals sampled for this recent study are of unknown origin, but earlier studies have indicated that mitochondrial HAPLOTYPES are not shared between different chimpanzee subspecies [1]. This allowed the authors to assign subspecies status based on the mitochondrial haplotype for all wild-born animals. Intriguingly, the most divergent NRY haplogroup (Pt?1) was from two captive-born animals living in European zoos; their origin is undetermined and their wild-born fathers are dead (Fig. 2). A second unique haplogroup from an animal of unknown origin was assumed to be *P.t. schweinfurthii* by Stone and colleagues (Pt?2). This is not necessarily justifiable. If *P.t. schweinfurthii* diverged recently from *Pan troglodytes troglodytes* [19], they could still share NRY inherited from the ancestral *P.t. troglodytes* population. In fact, even the more-rapidly evolving mtDNA hypervariable region 1 (HV1) haplotypes from *P.t. schweinfurthii* do not form a MONOPHYLETIC clade in phylogenetic reconstructions [20].

The future of Great Ape populations

Coordinated efforts by geneticists and primatologists could allow the collection of many valuable samples from captive populations, while promoting ethical treatment and providing financial support for many of the African facilities. Blood could be collected during regular health examinations, and could be used to establish cell lines. However, because of the unknown lineage of many captive Apes, studies of wild populations are preferable, and these studies will become increasingly feasible following recent improvement in non-invasive genotyping techniques.

However, time is rapidly running out for the study of natural Ape populations. A group of prominent field primatologists has launched GRASP, the Great Ape Survival project, through the United Nations Environmental Program. Geneticists would have much to contribute to this effort (http://www.unep.org/grasp/strategies.asp). It is a sad irony that thousands of wild Apes are being butchered yearly to feed the growing bush meat market, just as an increasing number of scientists are trying to obtain samples. Many more Great Apes are examined for the quality of their smoked tissues in African markets than are studied by anthropologists, geneticists or virologists combined. As geneticists, we have an important part to play in helping wild Ape populations persist.

Acknowledgements

I thank Jim Moore and Olivier Gagneux for their helpful comments.

References

1 Morin, P.A. *et al.* (1994) Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* 265, 1193–1201
2 Morin, P.A. *et al.* (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Mol. Ecol.* 10, 1835–1844
3 Ferris, S.D. *et al.* (1981) Extensive polymorphism in the mitochondrial DNA of apes. *Proc. Natl. Acad. Sci. U. S. A.* 78, 6319–6323
4 Jorde, L.B. *et al.* (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *Am. J. Hum. Genet.* 66, 979–988
5 Kaessmann, H. *et al.* (1999) Extensive nuclear DNA sequence diversity among chimpanzees. *Science* 286, 1159–1162
6 Deinard, A.S. and Kidd, K.K. (2000) Identifying conservation units within captive chimpanzee populations. *Am. J. Phys. Anthropol.* 111, 25–44
7 Hammer, M.F. *et al.* (2001) Hierarchical patterns of global human Y-chromosome diversity. *Mol. Biol. Evol.* 18, 1189–1203
8 Stone, A.C. *et al.* (2002) High levels of Y-chromosome nucleotide diversity in the genus *Pan*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 43–48
9 Shen, P. *et al.* (2000) Population genetic implications from sequence variation in four Y chromosome genes. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6927–6929
10 Horai, S. *et al.* (1992) Man's place in Hominoidea revealed by mitochondrial DNA genealogy. *J. Mol. Evol.* 35, 32–43
11 Bailey, W.J. *et al.* (1992) Reexamination of the African hominoid trichotomy with additional sequences from the primate beta-globin gene cluster. *Mol. Phylog. Evol.* 1, 97–135
12 Chen, F.C. and Li, W.H. (2001) Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am. J. Hum. Genet.* 68, 444–456
13 Boesch, C. (1996) Social grouping of Taï chimpanzees. In *Great Ape Societies* (McGrew, W. *et al.*, eds), pp. 101–113, Cambridge University Press
14 Hill, K. *et al.* (2001) Mortality rates among wild chimpanzees. *J. Hum. Evol.* 40, 437–450
15 Tutin, C.E.G. (1979) Mating patterns and reproductive strategies in a community of wild chimpanzees. *Behav. Ecol. Sociobiol.* 6, 29–38
16 Dixson, A.F. (1998) Sexual selection and evolution of the seminal vesicles in primates. *Folia Primatol.* 69, 300–306
17 Maley, J. (1991) The African rain forest vegetation and paleoenvironments during late Quaternary. *Climatic Change* 19, 79–98
18 Nei, M. and Takahata, N. (1993) Effective population size, genetic diversity, and coalescence time in subdivided populations. *J. Mol. Evol.* 37, 240–244
19 Goldberg, T.L. and Ruvolo, M. (1997) The geographic apportionment of mitochondrial genetic diversity in east African chimpanzees, *Pan troglodytes schweinfurthii*. *Mol. Biol. Evol.* 14, 976–984
20 Gagneux, P. *et al.* (2001) Gene flow in wild chimpanzees. What genetic data tell us about chimpanzee movements over space and time. *Philos. Trans. R. Soc. London Ser. B.* 356, 889–897

Pascal Gagneux
University of California San Diego, 9500 Gilman Drive, Glycobiology Research and Training Center, 0687, Cellular and Molecular Medicine East, La Jolla, CA 92093-0687, USA.
e-mail: gagneux@biomail.ucsd.edu