Great ape DNA sequences reveal a reduced diversity and an expansion in humans

The extent of DNA sequence variation of chimpanzees is several-fold greater than that of humans[1]. It is unclear, however, if humans or chimpanzees are exceptional among primates in having low and high amounts of DNA sequence diversity, respectively. To address this, we have determined approximately 10,000 bp of noncoding DNA sequences at Xq13.3 (which has been extensively studied in both humans and chimpanzees[1,2]) from 10 western lowland gorillas (*Gorilla gorilla gorilla*) and 1 mountain gorilla (*Gorilla gorilla beringei*; that is, from 2 of the 3 currently recognized gorilla subspecies[3]), as well as 8 Bornean (*Pongo pygmaeus pygmaeus*) and 6 Sumatran (*Pongo pygmaeus abelii*) orang-utans, representing both currently recognized orang-utan subspecies[4]. We show that humans differ from the great apes in having a low level of genetic variation and a signal of population expansion.

To compare the extent of diversity among humans and the great apes, we calculated Watterson´s estimate of the parameter theta[5], θ_w, which is based on the number of variable positions and takes the sample size analyzed in each species into account. The results (Table 1) show that θ_w in gorillas is about twice that in humans, whereas in chimpanzees and orang-utans it is about 3.1 and 3.5 times greater, respectively. One may argue that the more extensive diversity in the great apes is the result of more extensive population subdivision in these species. Although mtDNA data indicate genetically distinct subspecies[6,7], nuclear DNA shows the subspecies of all the great apes to be intermixed[1,8,9] (data not shown), indicating a rather recent separation of the subspecies. Thus, it appears to be justified to compare overall levels of diversity within the great ape species with those of humans. It is notable, however, that θ_w is 1.6 to 4.0 times greater compared with the human estimate even when it is calculated for each of the ape subspecies separately (Table 1). The only exception is western African chimpanzees, for which only zoo samples, which may not represent the entire diversity of this subspecies, were available.

The more extensive diversity in the great apes is reflected in the longer times to the most recent common ancestors (MRCA) in the ape species. The age of the MRCA for humans is about 540,000 years; for chimpanzees, 1,900,000 years; for gorillas, 1,160,000 years; and for orang-utans, 2,120,000 years (Table 1). A maximum likelihood[10] tree (Fig. 1) relating the great ape and human DNA sequences illustrates the occurrence of deep branches within all great ape species, whereas humans constitute only a small cluster with short branches.

Studies of mitochondrial DNA (mtDNA) have shown a three- to fourfold higher nucleotide diversity in chimpanzees compared with humans[1], as well as a greater extent of mtDNA sequence variation in gorillas and orang-utans[11]. Thus, both mtDNA and Xq13.3, the only two loci for which comparative intra-specific DNA sequence information is available from humans as well as the great apes, yield a congruent picture in which humans are unique compared with the great apes in having little genetic variation. Although it is possible that this is due to selection, this is unlikely because the putative selective factors would have had to affect both loci in a similar way in all these species. Rather, these data indicate that the population history of humans differs from that of our closest evolutionary relatives. For example, one may speculate that archaic human forms such as Neandertals, who disappeared about 30,000 years ago[12], carried additional Xq13.3 lineages not present in modern humans. Thus, had the Neandertals or other archaic humans survived until today, contemporary humans would perhaps have been more like the great apes in terms of genetic diversity.

To elucidate whether signals of past expansions in population size can be seen in the Xq13.3 sequences, we used Fu and Li´s D* test[13], which compares the number of singletons and the total number of mutations. Assuming that the noncoding Xq13.3 sequences evolve selectively neutral, this test can be used to detect past population growth. Although the hypothesis of constant population size cannot be rejected for any of the great ape species, it is rejected in humans ($P<0.02$; Table 1). Thus, humans differ from the great apes in having both a low level of genetic variation and a signal of expansion at Xq13.3. To estimate the date of the beginning of the expansion, we used

Table 1 • Sequence diversity at Xq13.3

	Xq13.3							
	diversity					TMRCA		
	n	s	θ_w	Fu and Li´s D*	*P*	mode	mean	95% confidence interval
humans	70	33	6.8	–3.35	<0.02	541,000	645,000	319,000–1,150,000
chimpanzees	30	84	21.2	–2.03	>0.05	1,910,000	2,100,000	1,160,000–3,350,000
central	12	64	21.2	n.d.	n.d.	1,810,000	2,000,000	1,170,000–3,030,000
western	17	23	6.8	n.d.	n.d.	534,000	636,000	325,000–1,090,000
eastern	1	0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
gorillas	11	41	14.0	0.78	>0.10	1,160,000	1,280,000	723,000–2,020,000
western	10	39	13.8	n.d.	n.d.	1,140,000	1,250,000	709,000–1,970,000
mountain	1	0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
orang-utans	14	77	24.2	0.44	>0.10	2,120,000	2,340,000	1,360,000–3,530,000
Sumatran	6	62	27.1	n.d.	n.d.	2,130,000	2,260,000	1,390,000–3,290,000
Bornean	8	28	10.8	n.d.	n.d.	893,000	988,000	552,000–1,580,000

"n", sample size; "s", number of variable positions in the respective primate. θ_w values are given per sequence. The significance of the Fu and Li D* test statistic for each species is indicated as "*P*". The mode, mean and 95% confidence interval of the time to the most recent common ancestor (TMRCA) are given. Some estimates were not determined (n.d.) due to small sample sizes. We estimated the time to the MRCA based on the number of variable positions by a Bayesian procedure as described[21], assuming a generation time of 20 years. We used a γ-distribution with mode 1.878×10^{-6} as prior for the mutation rate. Uncertainty in the population size estimates was modeled by a lognormal prior with mean 32,700 and standard deviation 10,000. We chose these priors to cover all plausible parameter values. MRCA estimates based on the maximum distances between two individuals in each species fall within the range obtained with the Bayesian procedure (humans, approximately 480,000 years; chimpanzees, 1,500,000 years; gorillas, 1,500,000 years; orang-utans, 2,700,000 years).

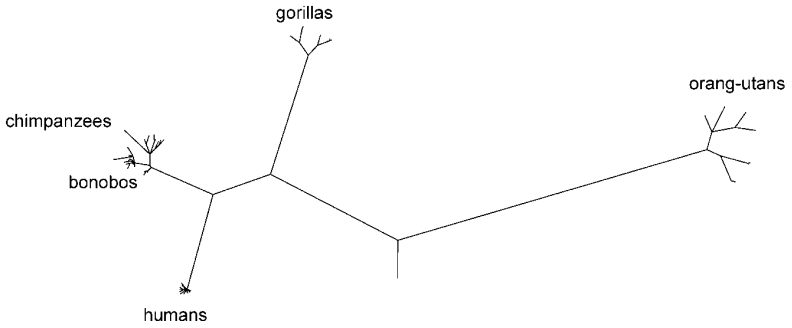

Fig. 1 Phylogenetic tree[10] of human (n=70; ref. 2), chimpanzee (n=30), bonobo (n=5; ref. 1), gorilla (n=11) and orang-utan (n=14) Xq13.3 sequences. Gorilla DNA samples (n=11) were obtained from zoos and primate research institutes. Orang-utan samples were from skin fibroblast cell lines (n=9; collected using remote biopsy darts) of wild orang-utans from both Borneo and Sumatra[4] as well as from zoos and primate research institutes (n=5). We performed PCR amplification and sequencing as described[1]. We used a gibbon sequence as an outgroup. The maximum likelihood tree reconstruction was performed with PUZZLE 4.0 (ref. 10) assuming a Tamura-Nei model with γ-distributed rates.

two different approaches, one based on the coalescent[14] and the other based on mismatch distributions[15]. The maximum likelihood value from the coalescent analysis indicates that the human population started to expand approximately 190,000 years ago from an initial effective population size of about 3,700, whereas the mismatch approach indicates 160,000 years ago as the start of the expansion.

Mitochondrial DNA sequences indicate an expansion of modern humans 40,000 to 50,000 years ago[16,17], a date that is associated with a change in human behaviour as indicated by a transition to more advanced and varied tool industries and the appearance of art. Due to its recent coalescence to one common ancestor, mtDNA may have 'captured' only this more recent population expansion, whereas Xq13.3, which has a MRCA approximately 540,000 years ago, may reveal an earlier expansion in the history of modern humans starting 160,000 to 190,000 years ago. Because it is not possible to obtain true confidence intervals around these estimates with current methods, we cannot rigorously exclude that mtDNA and Xq13.3 reflect the same human population expansion.

We note that three other studies of nuclear DNA sequence variation in humans[18] have failed to detect a population expansion. These DNA sequences are from transcribed genes that carry alleles implicated in disease and are therefore likely to be influenced not only by demographic phenomena but also by selection. It is possible that Xq13.3, which is non-coding, may be more suitable for elucidating historical demography. Two other recent studies of non-coding loci on chromosome 1 and 22, for which multiple human sequences of similar length were determined, also indicate a substitutional pattern consistent with a population expansion in humans[19,20]. It will be extremely important to study long DNA sequences from additional nuclear loci in humans as well as the great apes to elucidate whether a reduced diversity and a tendency to expansion relative to the great apes is typical for the human genome.

Acknowledgments

We thank C. Allen, S. O'Brien, P. Morin, W. Rietschel, C. Roos and O. Takenaka for DNA samples. This work was supported by grants from the Bundesministerium für Bildung und Forschung and the Max-Planck-Gesellschaft.

Henrik Kaessmann, Victor Wiebe, Gunter Weiss & Svante Pääbo

Max-Planck-Institute for Evolutionary Anthropology, Leipzig, Germany. Correspondence should be addressed to S.P. (e-mail: paabo@eva.mpg.de).

Received 29 September; accepted 20 December 2000.

1. Kaessmann, H., Wiebe, V. & Pääbo, S. *Science* **286**, 1159–1162 (1999).
2. Kaessmann, H., Heissig, F., von Haeseler, A. & Pääbo, S. *Nature Genet.* **22**, 78–81 (1999).
3. Groves, C.P. *J. Zool. Lond.* **161**, 287–300 (1970).
4. Zhi, L. *et al. Curr. Biol.* **6**, 1326–1336 (1996).
5. Watterson, G.A. *Theor. Popul. Biol.* **7**, 256–276 (1975).
6. Morin, P.A. *et al. Science* **265**, 1193–1201 (1994).
7. Gagneux, P. *et al. Proc. Natl. Acad. Sci. USA* **96**, 5077–5082 (1999).
8. Deinard, A.S. & Kidd, K. *Am. J. Phys. Anthropol.* **111**, 25–44 (2000).
9. Deinard, A. & Kidd, K. *J. Hum. Evol.* **36**, 687–703 (1999).
10. Strimmer, K.S. & von Haeseler, A. *Mol. Biol. Evol.* **13**, 964–969 (1996).
11. Ferris, S.D., Brown, W.M., Davidson, W.S. & Wilson, A.C. *Proc. Natl. Acad. Sci. USA* **78**, 6319–6323 (1981).
12. Stringer, C. in *The Human Inheritance* (ed. Sykes, B.) 33–44 (Oxford University Press, Oxford, 1999).
13. Fu, Y.X. & Li, W.H. *Genetics* **133**, 693–709 (1993).
14. Weiss, G. & von Haeseler, A. *Genetics* **149**, 1539–1546 (1998).
15. Rogers, A.R. & Harpending, H. *Mol. Biol. Evol.* **9**, 552–569 (1992).
16. Stoneking, M. *J. Bioenerg. Biomemb.* **3**, 251–259 (1994).
17. Ingman, M., Kaessmann, H., Pääbo, S. & Gyllensten, U. *Nature* **408**, 708–713 (2000).
18. Fay, J.C. & Wu, C.I. *Mol. Biol. Evol.* **16**, 1003–1005 (1999).
19. Yu, N. *et al. Mol. Biol. Evol.* (in press).
20. Zhao, Z. *et al. Proc. Natl. Acad. Sci. USA* **97**, 11354–11358 (2000).
21. Tavare, S., Balding, D.J., Griffiths, R.C. & Donnelly, P. *Genetics* **145**, 505–518 (1997).

Transgenic rescue of defective Cd36 ameliorates insulin resistance in spontaneously hypertensive rats

Spontaneously hypertensive rats (SHR) display several features of the human insulin-resistance syndromes. Cd36 deficiency is genetically linked to insulin resistance in SHR. We show that transgenic expression of Cd36 in SHR ameliorates insulin resistance and lowers serum fatty acids. Our results provide direct evidence that Cd36 deficiency can promote defective insulin action and disordered fatty-acid metabolism in spontaneous hypertension.

The spontaneously hypertensive rat (SHR) is the most widely studied animal model of hypertension. In the SHR, as in many humans with essential hypertension, increased blood pressure clusters with other risk factors for cardiovascular disease, including insulin resistance and dyslipidemia[1]. Despite intense effort, however, little progress has been made in the molecular identification of quantitative trait loci (QTL) that regulate the complex phenotypes of insulin resistance and dyslipidemia that cluster in animals and humans with hypertension. In SHR descended from the colony established at the National Institutes of Health (NIH), we found that a spontaneous deletion in *Cd36* (encoding a fatty acid transporter[2]) was linked to the transmission of insulin resistance, defective fatty-acid metabolism and hypertension[3,4]. Although linkage studies of complex traits cannot establish proof of QTL identity at the molecular level, the genetic studies of defective *Cd36* in SHR suggest that the hypertension