Complex speciation of humans and chimpanzees

Arising from: N. Patterson, D. J. Richter, S. Gnerre, E. Lander & D. Reich *Nature* **441**, 1103–1108 (2006)

Genetic data from two or more species provide information about the process of speciation. In their analysis of DNA from humans, chimpanzees, gorillas, orangutans and macaques (HCGOM), Patterson *et al.*[1] suggest that the apparently short divergence time between humans and chimpanzees on the X chromosome is explained by a massive interspecific hybridization event in the ancestry of these two species. However, Patterson *et al.*[1] do not statistically test their own null model of simple speciation before concluding that speciation was complex, and—even if the null model could be rejected—they do not consider other explanations of a short divergence time on the X chromosome. These include natural selection on the X chromosome in the common ancestor of humans and chimpanzees, changes in the ratio of male-to-female mutation rates over time, and less extreme versions of divergence with gene flow (see ref. 2, for example). I therefore believe that their claim of hybridization is unwarranted.

Patterson *et al.*[1] estimate the divergence time between humans and chimpanzees on the X chromosome to be 0.835 times the average autosomal divergence time; they note that this is less than the value of 0.94 predicted by their model of simple speciation (see Methods). They also computed a 'genome minimum' divergence time of 0.86 by measuring divergence close to sites where humans and chimpanzees share a derived base ('HC' sites[1]), and cited the similarity between this and the X-chromosome divergence time of 0.835 as support of hybridization. However, the HC speciation time estimated using the data and methods of Patterson *et al.*[1] is only 0.76 (see Methods): the X-chromosome divergence and the genome minimum divergence therefore occurred in the common ancestor of humans and chimpanzees, rather than after an initial speciation event (Fig. 1). The apparently low estimated divergence time on the X chromosome thus provides only indirect evidence for complex speciation.

It has been suggested that the autosomal data may be consistent with the null model of simple speciation and that a test is lacking[3]. With reference just to the apparent reduction in divergence time on the X chromosome, a proper statistical test would compute the probability under the null model of observing a difference between the X chromosome and the autosomes as large as, or larger than, the difference between 0.94 and 0.835. This test would require all relevant sources of variation to be accounted for, including the stochastic variation in coalescence times under the null model and the variation in α, the ratio of male-to-female mutation rates. The jack-knife procedure described in the Supplementary Information of ref. 1 does not do this.

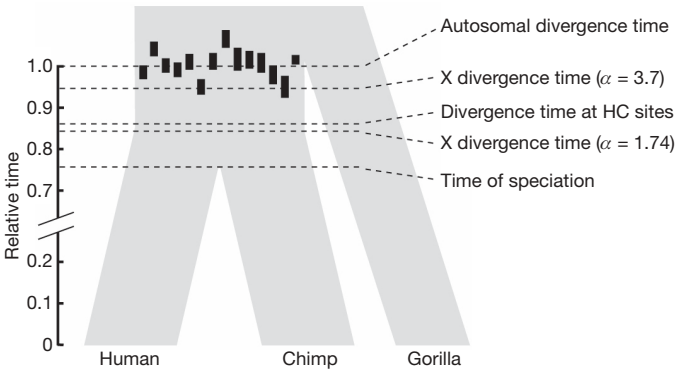

Figure 1 | Simple Speciation of humans and chimpanzees. Black bars show the ranges of autosomal human–chimpanzee divergence by chromosome, redrawn (in the same order) from Fig. 3a of Patterson *et al.*[1]. Dashed lines display quantities discussed in the text. The drawing is to scale, with all times given relative to the average autosomal divergence time between humans and chimpanzees.

Assuming that the null model can be rejected, the suggestion of hybridization needs to be supported by rejecting other kinds of complex speciation, which Patterson *et al.*[1] do not. Their Supplementary Note 11 considers whether modifications of the null model could produce a large reduction of HC divergence on the X chromosome, and they conclude that natural selection must explain the reduction. Note that, because fossil dates were included in the analysis, Supplementary Note 11 seeks to explain a more extreme reduction ($R < 0.29$[1]) than implied by the genetic data alone. Different scenarios that include natural selection and models of complex speciation other than hybridization are not considered.

An alternative to hybridization is that α has changed during the course of primate evolution. Another study[4] finds $\alpha = 3.88$ (95% CI = 2.90–6.07) in primates and $\alpha = 3.79$ (95% CI = 2.71–5.99) in perisodactyls. In contrast, Patterson *et al.*[1] use $\alpha < 2$, based on the observed divergences to macaque in their data. Relative genetic divergences on the X chromosome versus the autosomes are converted into the relative divergence times above by using the factor $3(1+\alpha)/(4+2\alpha)$, which accounts for the different fractions of time the X and the autosomes spend in males versus females. For example, summing the counts of patterns in which human and chimpanzee differ and dividing by the total number of bases in Table 1 of ref. 1, gives genetic divergences of $\pi^{(X)}_{HC} = 0.0053$ and $\pi^{(A)}_{HC} = 0.0070$. The relative genetic divergence is $\pi^{(X)}_{HC}/\pi^{(A)}_{HC} \approx 0.76$. A similar value ($0.0094/0.0123 \approx 0.764$) is obtained from the whole genomes of humans and chimpanzees[5]. If $\alpha = 1.74$, then $0.76 \times 3(1+\alpha)/(4+2\alpha) = 0.835$ (ref. 1). Instead, if $\alpha = 3.7$, then $0.76 \times 3(1+\alpha)/(4+2\alpha) = 0.94$, as predicted by the fitted null model[1]. A similar value ($\alpha = 3.55$) makes the X-autosome relative genetic divergence 0.61 within humans[6], consistent with the simple prediction of 3/4 for the relative divergence time.

METHODS

The values 0.76 and 0.94 are obtained from the HCGOM autosomal data in Supplementary Table 4 of ref. 1 by using the model and method in Supplementary Note 2 of ref. 1, and assuming that the effective population sizes of HC and HCG ancestors are the same. The three model parameters are in units of expected numbers of mutations in the sampled portions of the autosomal genome: γ is the time back to the HC speciation event, β is the time between the HCG and HC speciation events, and $\theta/2$ is the pairwise coalescence time within a species. The expected values in Supplementary Note 2 of ref. 1 become $E[n_H + n_{HG}] = \gamma + \theta/2$; $E[n_H + n_{HC}] = \gamma + \beta + \theta/2$; $E[n_{HG}] = \theta e^{-2\beta/\theta}/6$, where n_H, n_{HC} and n_{HG} are the numbers of mutations unique to humans, shared uniquely by humans and chimpanzees, and shared uniquely by humans and gorillas. After correcting for multiple mutations at single sites[1], the observed values become $n_H = 28{,}173$, $n_{HC} = 7{,}990$ and $n_{HG} = 866$. By equating the expected and observed values, one finds $\gamma \sim 21{,}946$, $\beta \sim 7{,}124$ and $\theta \sim 14{,}186$. Thus, relative to the average autosomal divergence, the estimated HC speciation time is $2\gamma/(2\gamma+\theta) \approx 0.76$. With the standard assumption of a 3/4 ratio of effective population sizes on the X chromosome versus the autosomes[1], the predicted value for the relative HC divergence time on the X chromosome is $(2\gamma + 0.75\theta)/(2\gamma+\theta) \approx 0.94$.

John Wakeley[1]

[1]Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA.
e-mail: wakeley@fas.harvard.edu

Received 22 February 2007; accepted 17 January 2008.

1. Patterson, N., Richter, D. J., Gnerre, S., Lander, E. & Reich, D. Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441**, 1103–1108 (2006).
2. Innan, H. & Watanabe, H. The effect of gene flow on the coalescent time in the human–chimpanzee ancestral population. *Mol. Biol. Evol.* **23**, 1040–1047 (2006).
3. Barton, N. H. Evolutionary biology: how did the human species form? *Curr. Biol.* **16**, R647–R650 (2006).

4. Goetting-Minetsky, M. P. & Makova, K. D. Mammalian male mutational bias: impacts of generation time and regional variation in substitution rates. *J. Mol. Evol.* **63,** 537–544 (2006).
5. The Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437,** 69–87 (2005).
6. The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409,** 928–933 (2001).

doi:10.1038/nature06805

Patterson et al. *reply*

Replying to: J. Wakeley *Nature* 452, doi:10.1038/nature06805 (2008)

In his communication[1], Wakeley does not find any flaw in our argument for complex speciation of humans and chimpanzees[2], nor has he identified a simple demographic model that can explain the notable differences in genetic divergence that we observe between humans and chimpanzees when comparing chromosome X and the autosomes.

Our argument[2] for complex speciation rests on the difference in genetic divergence time that we observe between chromosome X and the autosomes, and not on the wide range of genetic divergence times observed within the autosomes (which can indeed be explained by a large ancestral population size[3,4]). To reiterate our argument for complex speciation, we began with a null model of simple speciation in which the ancestral populations of humans and chimpanzees were separated by a barrier with no subsequent gene flow. Fitting this model to our data, we obtain an expectation for the genetic divergence on chromosome X. However, the observed chromosome X data differ from this, and we could not explain the difference even by using more elaborate demographic histories. Wakeley's model of demographic history can explain only the autosomal data, and does not reconcile the autosomal and chromosome X data, even though comparing these two parts of the genome provides the key signal for complex speciation. Table 1 shows three statistics that are each sensitive to human–chimpanzee genetic divergence on chromosome X: all are significantly reduced (4.4–8.3 s.d.) compared with that predicted from Wakeley's simple model fitted to our autosomal data.

Wakeley[1] suggests that the ratio of male-to-female mutation rate in primates (α) might be higher than we estimated. This would not explain the data, especially in light of our human–gorilla comparison. Using Wakeley's method[1] for calculating α with raw genetic divergence data (rather than by comparison with an outgroup[2]), we estimate $\alpha = 3.19$ from the human–chimpanzee comparison, similar to the value Wakeley reports (the slight difference is owing to our correction for recurrent mutation). But this disagrees with our $\alpha = 1.57$ obtained from the same calculation using the human–gorilla comparison.

If mutation-rate differences alone could explain the observed data, we would expect a consistent value for α from the human–chimpanzee and human–gorilla divergence data, but estimates of α are significantly different ($P = 0.001$). A high value of α also cannot explain other important features in Table 1: the near-absence of sites on chromosome X that cluster humans and gorillas or chimpanzees and gorillas; or why human–gorilla divergence should not be reduced on chromosome X (such a reduction would be expected if high male mutation rate were responsible for low human–chimpanzee genetic divergence on chromosome X).

What could explain the evidence for reduced chromosome-X time divergence? We suggested hybridization[2]. In hybrids, genetic barriers to gene flow often map to chromosome X (ref. 5). A corollary is that if a hybrid population overcomes these barriers, it may experience intense selection to eliminate most or all the chromosome X contribution from one of the ancestral populations—the population cannot tolerate the sequence of both X chromosomes and one is selected away. When hybridization between two populations establishes a third population, the divergence on chromosome X will be large relative to one ancestral population and small relative to the other, depending on which ancestral chromosome X is selected away. Depending on which populations survive to the present day, the divergence on chromosome X will be very high or very low relative to that on the autosomes (but not intermediate), which would explain the low observed chromosome X divergence.

We therefore reject the simple model of speciation that Wakeley proposes, having also investigated other simple speciation models and found none to explain the data[2]. In contrast, hybridization followed by natural selection across the entire chromosome X to eliminate hybrid sterility or inviability loci does explain the data[2]. To argue against the evidence for complex speciation, an alternative simple model is needed that explains the reduced chromosome X divergence in humans and chimpanzees with no similar reduction for humans and gorillas. No one has yet succeeded in identifying such a model.

Nick Patterson[1], Daniel J. Richter[1], Sante Gnerre[1], Eric S. Lander[1,2] & David Reich[1,3]
[1]Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA.
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
[3]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.
e-mail: reich@genetics.med.harvard.edu

1. Wakeley, J. Complex speciation of humans and chimpanzees. *Nature* **452,** doi:10.1038/nature06805 (2008).
2. Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S. & Reich, D. Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441,** 1103–1108 (2006).
3. Barton, N. H. Evolutionary biology: how did the human species form? *Curr. Biol.* **16,** 647–650 (2006).
4. Innan, H. & Watanabe, H. The effect of gene flow on the coalescent time in the human–chimpanzee ancestral population. *Mol. Biol. Evol.* **23,** 1040–1047 (2006).
5. Coyne, J. A. & Orr, H. A. in *Speciation and its consequences* (eds Otte, D. & Endler, J. A.) 180–207 (Sinauer Associates, Sunderland, Massachusetts, 1989).

doi:10.1038/nature06806

Table 1 | Statistics showing extreme reduction in genetic divergence time between humans and chimps on chromosome X

Statistic of interest*	Expectation of X-to-autosome ratio from Wakeley's model†	Observation of X-to-autosome ratio in our actual data	Significance of difference between observed and expected (s.d.)
Human–chimpanzee genetic divergence averaged across the genome and normalized by human–macaque divergence	0.937 ± 0.002	0.839 ± 0.022	−4.4
Sum of divergent sites clustering humans and gorillas (HG) or chimpanzees and gorillas (CG) and normalized by human–macaque divergence	0.548 ± 0.023	0.022 ± 0.059	−8.3
Ratio of human–chimpanzee to human–gorilla divergence averaged across the genome and calculated on the human side of the tree	0.987 ± 0.001	0.862 ± 0.022	−5.8

* Corrected for recurrent mutation.
† Parameters fitted to autosomal data, then extrapolated to chromosome X. The extrapolation to chromosome X assumes a ratio of chromosome X to autosomal population size of three-quarters throughout history.