

Speciation and inversions: chimps and humans

Jody Hey

Summary

A new set of models has resurrected a role for chromosomal inversions in the formation of new species.^(1–3) Traditional models, which are generally considered to be unlikely in most cases, had imagined that inversions might aid speciation by directly causing low hybrid fitness. In contrast, the newer models focus on the effect that inversions have on local recombination rates. A test of these models found a strikingly high rate of amino-acid substitution within regions where humans and chimpanzees differ by inversions, suggesting perhaps that our ancestral species underwent a divergence process in which gene flow and inversions played a key role.⁽⁴⁾ However, it remains uncertain whether this interesting finding is actually consistent with the proposed model.

BioEssays 25:825–828, 2003.

© 2003 Wiley Periodicals, Inc.

Introduction

The processes by which new species arise are the subject of much evolutionary research, and it is one in which many biologists hold an abiding interest. Indeed, there is one particularly interesting case of speciation that is close to home for all of us, and that is the split that occurred in the common ancestral species of humans and chimpanzees. Molecular evidence suggests that this event occurred sometime between five and seven million years ago.^(5,6) Hominin fossils from that time are known only from Africa and present-day chimpanzees are known only from Africa (there are no fossils for chimpanzees, Ref. 7), so it seems pretty likely that Africa is where the speciation occurred. However we know little else about how it might have happened.

Recently some new population genetic theory and some DNA sequence comparisons have been brought together in support of a model of how our ancestral species initially separated from each other. The findings are somewhat surprising in that the theory relies upon chromosomal inversions that distinguish the genomes of humans and chimpanzees. Geneticists have long known that such inversions could in principal be the basis for some cases of reproductive isolation

between newly arisen species.⁽⁸⁾ However, the simplest models, in which inversion heterozygotes have low fertility, are not very plausible because such inversions are not expected to rise high in frequency except by chance in very small populations. Indeed the new inversion-based theory, which was first sketched out by Rieseberg,⁽¹⁾ does not rely upon the fitness cost of inversion heterozygosity, but rather upon the strong suppressing effect that inversions have on recombination within and near inverted regions in heterozygotes.

Inversion theory

The theory begins by considering two closely related populations that have some opportunity for gene exchange, such as might happen if they share a common geographic or environmental boundary (parapatry). Next, suppose that each population has, at high frequency, alleles at one or more genes that are 'incompatible' with genes in the other population. Such incompatibilities might be manifested directly as inviable or partly sterile hybrids; however, incompatibility might also be mediated through the environment, if an allele from one population is disadvantageous in the environment of the other population. These incompatible alleles are not expected to spread from one population to the other because of their selective disadvantage. Thus hybrids between the populations will usually be heterozygous for such incompatible alleles. If the effects of incompatibility are so strong as to cause a complete loss of hybrid fitness, then speciation is complete. Suppose, however, that the populations are only slightly divergent and that hybrid fitness is only reduced a small amount. Then it will still be the case that the incompatible alleles do not move between the populations, even while genes that are not linked to the incompatible alleles may move between the species. The presence of the incompatible alleles creates what is called a 'genetic barrier' to gene exchange for sites that are tightly linked to those alleles.⁽⁹⁾ This distinction, between the more restricted gene flow for genes linked to incompatible alleles and the less restricted flow of unlinked alleles, is fundamental to understanding the possible effects of chromosomal inversions.

Now consider a chromosomal inversion that is common in one population and rare in the other. If each form of the chromosome, inverted and non-inverted, carries an incompatible allele at one of the genes within the span of the inversion, then the genetic barrier caused by these alleles extends to include the entire region of tight linkage to the alleles, which is

Department of Genetics, Rutgers, the State University of New Jersey,
604 Allison Rd. Piscataway, NJ 08854-8082.

E-mail: hey@biology.rutgers.edu

DOI 10.1002/bies.10336

Published online in Wiley InterScience (www.interscience.wiley.com).

the region in which the inversion suppresses recombination in heterozygotes. Furthermore, this genetic barrier to gene flow is stable, since neither form of the chromosome can replace the other because of the incompatible alleles. So long as the inversion difference persists, that region of the genome can be the site of accumulation of additional alleles that are favorable in one population but not the other. This is essentially the model of Noor et al., who realized that an inversion greatly increases the chance that pairs of incompatible alleles, one in each population, may be caught in a stable configuration when they are both spanned by an inversion.⁽²⁾ The prediction of this model is that species that arose in the presence of gene flow should differ from their sister species by inversions, more so than should species pairs that diverged without gene flow. Put another way, the model says that allopatric species pairs do not need inversions to diverge, and that sympatric or parapatric populations without inversion differences will have been less likely than those that do differ by inversions to actually diverge into separate species. Indeed *Drosophila* species that are sympatric with their sister species do differ from them by inversions more so than do allopatric pairs of sister taxa.⁽²⁾

Navarro and Barton generalize this model by considering the dynamics that arise when two sorts of alleles arise within two parapatric populations.⁽³⁾ Alleles of one class are simply advantageous throughout both populations whereas alleles of a second class are favored only in one population (and are incompatible and disfavored in the other population). When the former pass between populations, they become fixed in both and cause the populations to be more similar. However, if such a favored mutation arises near the site where there is an incompatible allele in the other population, the selective sweep will be retarded because of the reduced gene flow in the presence of the genetic barrier. Navarro and Barton ask us to envision a race between mutations that sweep through both populations and remove genetic barriers as they do, and a different class of mutations that become fixed in just one population and cause incompatibilities and genetic barriers to gene exchange. Unless gene flow is very low, the latter type, which cause incompatibilities, will lose this race and the populations will tend to merge. But suppose that one or more incompatible alleles occur in a region that is spanned by a chromosomal inversion that distinguishes the two populations. Now the genetic barrier caused by the incompatible alleles includes many genes (i.e., all those in the region that experienced reduced crossing over in inversion heterozygotes). Now there is a much larger region with sharply reduced gene flow and thus a much larger region in which new incompatible alleles can arise and persist and not be removed by favorable mutations that sweep through both populations.

Regardless of inversions, we expect that two populations that have zero or extremely low gene flow will diverge from one another, both by genetic drift and by the action of mutations that are beneficial in only the population in which they occur. In

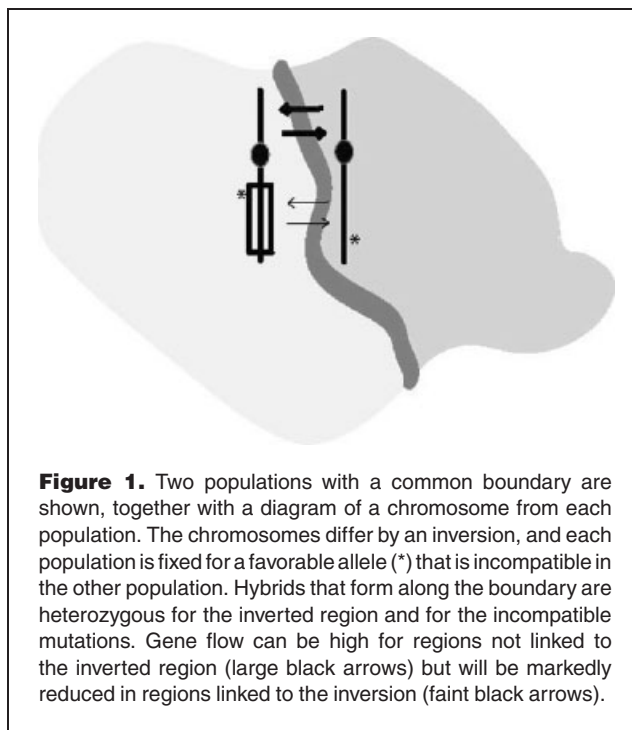


Figure 1. Two populations with a common boundary are shown, together with a diagram of a chromosome from each population. The chromosomes differ by an inversion, and each population is fixed for a favorable allele (*) that is incompatible in the other population. Hybrids that form along the boundary are heterozygous for the inverted region and for the incompatible mutations. Gene flow can be high for regions not linked to the inverted region (large black arrows) but will be markedly reduced in regions linked to the inversion (faint black arrows).

theory, the effect of genetic barriers to gene exchange is to shift the situation for populations that have too much gene flow to a situation that is effectively one of much lower gene flow in which additional incompatibilities can arise. If we imagine that there is a level of gene flow below which speciation occurs and above which populations merge, then the effect of inversions is to shift the effective level of gene flow near genetic barriers to a lower level and make speciation more likely (Fig. 1).

Testing the theory

This model has the same prediction as that of Noor et al.⁽²⁾ that sympatric sister species should tend to have more inversion differences than allopatric sister species. But there is also another interesting prediction which is that genes within inversions should preferentially have been the site of population-specific selective sweeps associated with selection of alleles that caused incompatibilities and contributed to the genetic barrier associated with the inversion. Thus, during the period when populations were exchanging genes, genes near the genetic barriers should be accumulating selective differences and diverging faster than genes away from the genetic barriers. It is this prediction that inspired Navarro and Barton's second paper on the human/chimpanzee speciation event.⁽⁴⁾ Ten of our chromosomes differ from those of chimpanzees by either an inversion or, in one case, a chromosomal fusion. Thus the question was whether genes on these chromosomes that might have been the site of genetic barriers to gene exchange have experienced higher levels of divergence for selectively favored substitutions than have genes on other chromosomes.

The authors used a conventional index of relative substitution rates, K_A/K_S , where K_A is an estimate of the rate at which amino acid changes accumulate at positions where mutations can cause an amino acid change, and K_S is the corresponding estimate for non-amino acid (i.e., synonymous) changes. Under a classic neutral model in which synonymous mutations are neutral and amino-acid changes are nearly all disadvantageous, the ratio is expected to be considerably less than 1. Values greater than 1 are traditionally taken as evidence that a number of selectively favored amino acid substitutions have elevated the K_A above what one would expect, even if amino acid changes were completely neutral.⁽¹⁰⁾

There were 115 genes in the study, and they were roughly split between those on rearranged chromosomes and those on collinear chromosomes. The former showed a mean K_A/K_S ratio of 0.84 whereas the latter had a mean ratio of 0.37. The distributions of the ratios for the two groups of genes were significantly different. Furthermore, those genes on rearranged chromosomes that actually lie within the span of the rearrangement had higher K_A/K_S values (though not significantly so) than those that flanked it, as expected if these regions had lower net recombination in inversion heterozygotes. Under the model, the elevated K_A/K_S value within rearranged regions is not because these regions have more selectively favored mutations, but rather because it is within these regions that more of the favored mutations have been restricted to one species or the other and it is these that show up in an assessment of divergence and substitution rate.

This seems to be quite a remarkable observation, as it implies a greatly elevated amino acid substitution rate for a large fraction of the human genome. Navarro and Barton do have an elegant model, but this particular test of it seems to raise more questions than it answers. Perhaps the first question concerns a simpler prediction than that associated with the K_A/K_S ratio. Under the model, incompatible alleles accumulate in the inverted regions (and not elsewhere prior to complete speciation) because the diverging populations are exchanging genes. Thus one seemingly straightforward expectation under the model is that there should be greater divergence between humans and chimpanzees in regions where their chromosomes differ by inversions. Furthermore, this difference in divergence levels should apply to all sequences (genes and intergenic regions, amino acid sites and silent sites, introns etc). When K_S was compared between inverted and non-inverted regions, the latter actually had slightly higher values. Though not statistically significant, the observation is in the opposite direction of what is expected under the model. However other data do support the contention that rearranged regions have higher levels of divergence between humans and chimps.^(11,12) It is also noteworthy that the prediction regarding divergence levels and inversions has been partly confirmed in some other species pairs.^(13–15)

Additional questions arise when considering the original articulation of the model, in which two kinds of mutations are considered, those that are favored and sweep between populations unless they are near a genetic barrier and those that arise near a genetic barrier and cause an incompatibility.⁽³⁾ Both kinds of mutations are advantageous, but the incompatible ones are only advantageous in one population. Following speciation, there is no gene flow and both kinds of mutations would be expected to accumulate in all regions of the genome. But prior to speciation, during a period of gene flow, regions characterized by genetic barriers will also be diverging by accumulating incompatible alleles. Thus, under the model, the elevation of K_A/K_S values in rearranged regions is caused by the accumulation of these incompatible alleles during the period when the populations exchanged genes. But why would the fixation of these incompatible alleles be associated with a markedly different K_A/K_S ratio prior to speciation than would the fixation of both types of mutations following speciation? This question gets to the heart of the prediction and test conducted by the authors.

A final question that arises even if the model is correct, and about which the authors speculate, is how to get such a large difference in K_A/K_S values. The observed difference implies very high rates of amino acid substitutions over a large fraction of the genome for a long period of time (i.e., millions of years, Ref. 16). Yet under the model it is expected that once incompatible alleles start to accumulate, a cascade develops in which an increasingly large fraction of substitutions in one species would be incompatible in the other species.⁽¹⁷⁾ This may lead fairly rapidly to the evolution of much stronger isolation and finally, complete speciation. In other words, under the model there should not be a very long time period in which incompatible alleles accumulate preferentially in regions linked to genetic barriers.

Conclusions

Navarro and Barton have developed a novel model that may well come to play a significant role in future speciation studies. They have also made a striking observation about relative rates of amino acid and silent substitutions in the divergence of humans and chimps. This observation, though cast as befitting the speciation model, actually raises more question than it answers.

Acknowledgments

I am grateful to Arcadi Navarro and Nick Barton for comments on the manuscript.

References

1. Rieseberg LH. Chromosomal rearrangements and speciation. *Trends Ecol Evol* 2001;16:351–358.
2. Noor MA, Grams KL, Bertucci LA, Reiland J. Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci USA* 2001; 98:12084–12088.

3. Navarro A, Barton NH. Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* 2003;57:447–459.
4. Navarro A, Barton NH. Chromosomal speciation and molecular divergence—accelerated evolution in rearranged chromosomes. *Science* 2003;300:321–324.
5. Chen F-C, Li W-H. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am J Human Genetics* 2001;68:444–456.
6. Takahata N, Satta Y. Evolution of the primate lineage leading to modern humans: phylogenetic and demographic inferences from DNA sequences. *Proc Natl Acad Sci USA* 1997;94:4811–4815.
7. Gee H. Palaeontology. Return to the planet of the apes. *Nature* 2001;412:131–132.
8. White MJD. *Modes of Speciation*. San Francisco: W.H. Freeman, 1978.
9. Barton N, Bengtsson BO. The barrier to genetic exchange between hybridising populations. *Heredity* 1986;56:357–376.
10. Kimura M. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press, 1983.
11. Ebersberger I, Metzler D, Schwarz C, Paabo S. Genomewide Comparison of DNA Sequences between Humans and Chimpanzees. *Am J Hum Genet* 2002;70:1490–1497.
12. Yi S, Ellsworth DL, Li WH. Slow molecular clocks in Old World monkeys, apes, and humans. *Mol Biol Evol* 2002;19:2191–2198.
13. Rieseberg LH, Whitton J, Gardner K. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* 1999;152:713–727.
14. Wang RL, Wakeley J, Hey J. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* 1997;147:1091–1106.
15. Machado C, Kliman RM, Markert JM, Hey J. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and its close relatives. *Molecular Biology and Evolution* 2002;19:472–488.
16. Rieseberg LH, Livingstone K. Chromosomal Speciation in Primates. *Science* 2003;300:267–268.
17. Orr HA, Turelli M. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution Int J Org Evolution* 2001;55:1085–1094.