REPLY TO DISCUSSANTS: RECENT COMMON ANCESTORS OF ALL PRESENT-DAY INDIVIDUALS

I would like to thank the editors for suggesting a discussion for this paper and for inviting such a distinguished panel of discussants, making this a genuine treat for me. My gratitude to the discussants for their comments may interfere with maintaining a suitably contrary attitude, but I will try.

Much of the discussion concerned one-parent models. Several (including myself in the paper under discussion) pointed out how one-parent models are relevant in genetics and how different genetic loci can have very different genealogies. Slatkin and Donnelly imagine, for each position in the DNA, an ancestral history that changes along the genome; this picture was also discussed in the introduction of the paper. I like Slatkin's suggestion that a theory somehow interpolating between this genetic picture and the two-parent results of the present paper would be interesting, and in fact Wiuf and Hein make a start in this direction.

Although the relevance of one-parent models in genetics is clear, I still tend to feel uneasy about assertions of 'robustness' of the coalescent, Ewens sampling formula and Poisson–Dirichlet distribution. How far does this robustness extend? Kingman summarized, 'Everything turns on the independence of the genealogical structure of the family tree on the one hand, and the process of mutation on the other. If this is true, everything else follows.' But what about geographic considerations, which the paper emphasized as a concern about the applicability of many models? Questions of robustness under various types of population subdivision and geographic structure remain largely open, as far as I know. I would imagine that geography can play havoc with purportedly robust relationships while still satisfying Kingman's independence condition, and reasonable forms of geographic barriers and population subdivision are quite capable of making the formulas fit very poorly.

I also suspect that certain robustness statements in two-parent models might be established under conditions rather similar to those giving robustness results in one-parent models. Such statements would further highlight a gap between current theoretical formulations and realistic applications. If two-parent models enjoy robustness results, and if we find certain two-parent predictions implausible, then should we be comforted by the one-parent robustness results?

The effects of geography seem important but difficult to analyse, and most coalescent type analyses ignore them. Donnelly sees an 'urgent need' for the development of even partially realistic models that incorporate geographic considerations, whereas Ewens hints that the situation here is more satisfactory, referring to 'more modern' calculations that would incorporate such effects. I tend more toward Donnelly's viewpoint here. One interesting modern example is the *Y*-chromosome Adam paper of Dorit *et al.* (1995) and the subsequent technical comments on that paper that appeared in the May 31, 1996, issue of *Science*. While these comments contained detailed criticisms and refinements of the original analysis, the contribution of Donnelly and Tavaré was the only one that even mentioned geography, and then just briefly and qualitatively.

I appreciate Donnelly's thoughts about species trees. Donnelly feels it is better to look at the set of topologies from different loci than to try to find a single, 'correct' species tree. I must admit I feel much the same way myself, and I still struggle with the question of what a species tree is really supposed to be. But many scientists appear to want a clear separation between the concepts of gene trees and species tree, endowing the species tree with a reality

that is not tied to or derived from gene trees. Such a conceptual framework often seems to be presupposed in, for example, arguing over the 'true' species tree for a given set of species and analysing possible disagreements between gene trees and species trees. The idea mentioned in the paper came from asking myself what a reasonable definition of species tree in this spirit might be. I am unsure about whether it may be useful.

I was puzzled to see Ewens arguing for the obvious proposition that a one-parent calculation is more appropriate for dating 'mitochondrial Eve' than a two-parent calculation; this is true, of course, and I hope I did not give the impression that I felt otherwise. I would be eager to hear how Ewens views a question closer to the subject of the paper: When did the most recent common ancestors of mankind live? For example, is the answer something like 100 000 years ago, or perhaps more like 600 years ago, or something very different? Why?

The answer might be that the question is of no interest. According to Kingman, 'The concept of being a common ancestor of a whole subsequent population is a very weak one, and it is not clear that it has any real significance.' Such doubts may arise from a genetic perspective, although, as warned, genetics is not really what this paper was about. The descendants of a common ancestor need not share any particular DNA from that ancestor, and it is even possible that none of the descendants has inherited any DNA from the ancestor. If you and I were investigating our common ancestry, we might conceive of an extreme case in which your mother's father is the same as my mother's father, but our common grandfather passed along no genes to either of us. Our ability to detect this common ancestor may be affected by these genetic circumstances, but the fact that we have a common grandfather would remain. I suspect a certain interest in our common ancestors comes naturally to most of us. For example, consider again Svante Pääbo's statement, 'What the data seem to show, however, is that we have a much more recent common ancestor, one who lived some 100 000 or 200 000 years ago.' Who are 'we' here? I doubt 'we strands of mitochondrial DNA' is the intended interpretation, but rather, Pääbo is talking about us people and the overlapping and joining of our family histories. The anthropological significance of this type of genetic evidence seems to have little to do with whether it comes from mitochondrial DNA, the Y chromosome, or elsewhere, and I imagine if an anthropologist were somehow to divine or discover when and where our most recent common ancestors lived, the information would be valued.

Wiuf and Hein touch on connections with genetics, with one question being whether a most recent common ancestor of two individuals, say, is a genetic ancestor of some part of the genomes of the two individuals. If the part of the genome is not prespecified, so that searching over the genome is allowed, the answer should depend on aspects of the model of reproduction, the size of the genome, and the size of the population, as also observed by Donnelly. On the question of the steady state fraction of a past population that are ancestral to present-day individuals, the calculation 'about 80% of the population is biologically related to individual i' of Wiuf and Hein agrees with the statement 'a randomly chosen individual would be a common ancestor with probability about 0.8' of the final section of the paper; here 0.8 refers to $1 - \rho \approx 0.7968$, where ρ is the extinction probability for a branching process with offspring distribution Poisson(2).

Other questions about how we might use available data to investigate common ancestors and about connections between the results of this paper and genetics remain to be worked out. I hope the expectations of Donnelly, Slatkin, and Wiuf and Hein that these connections could be interesting will be realized. I started on this subject because the questions seemed natural and the analysis and results seemed interesting and useful, conceptually at least. I share the feeling Wiuf and Hein express about being motivated by 'basic curiosity' here, and I particularly

enjoyed the way they entered into the spirit of the investigation. Their generalization of the two-parent model is a lot of fun, in addition to modeling more diverse organisms, including 'a hypothetical extraterrestrial life form.' Somewhere in the universe, this work may find practical application after all.

References

DORIT, R. L., AKASHI, H. AND GILBERT, W. (1995). Absence of polymorphism at the ZFY locus on the human Y chromosome. *Science* **268,** 1183–1185.

JOSEPH T. CHANG
Statistics Department,
Yale University,
Box 208290,
New Haven,
CT 06520, USA.
Email address: joseph.chang@yale.edu