Comment on "Ongoing Adaptive Evolution of *ASPM*, a Brain Size Determinant in *Homo sapiens*" and "*Microcephalin*, a Gene Regulating Brain Size, Continues to Evolve Adaptively in Humans"

Mathias Currat,[2] Laurent Excoffier,[2] Wayne Maddison,[1] Sarah P. Otto,[1]* Nicolas Ray,[2] Michael C. Whitlock,[1] Sam Yeaman[1]†

Mekel-Bobrov *et al*. and Evans *et al*. (Reports, 9 Sept. 2005, p. 1720 and p. 1717, respectively) examined sequence data from modern humans within two gene regions associated with brain development, *ASPM* and *microcephalin*, and concluded that selection of these genes must be ongoing. We show that models of human history that include both population growth and spatial structure can generate the observed patterns without selection.

Evolutionary processes, including selection, migration, and population size expansion, alter the probability that mutations persist within a species. Thus, DNA sequence comparisons within and among species can provide insight into evolutionary history. Unfortunately, many evolutionary processes leave similar signals in DNA sequences. To conclude that selection has shaped genetic sequence data, one must first reject reasonable alternative explanations based on demographic models alone.

Recent papers by Mekel-Bobrov *et al*. (*1*) and Evans *et al*. (*2*) examined sequence data from an ethnically diverse group of humans within two gene regions: *ASPM* and *microcephalin*, respectively. These genes were previously known to affect brain size based on clinical features of individuals carrying loss-of-function mutations (primary autosomal recessive microcephaly). The authors observed a single haplotype at high frequency in each of these genes (haplotype 63 at 21% in *ASPM*; haplotype 49 at 33% in *microcephalin*). Given the length of DNA sequenced (*ASPM*, 62114 bp; *microcephalin*, 29027 bp) and the substantial number of polymorphic sites (*ASPM*, 166; *microcephalin*, 220), observing single haplotypes at high frequency is notable. Indeed, the authors used coalescent simulations of nine different demographic models describing the growth and movement of human populations, none of which generated the observed levels of homozygosity or single haplotypes at high frequency for the estimated rates of recombination and conversion. Consequently, the authors argued that selection must have acted to raise the frequency of certain haplotypes within the human population.

Unfortunately, the demographic histories that were examined were only a small subset of the larger number consistent with what is known about human history. Thus, rejecting a subset (even a large subset) may not be relevant. Indeed, a straightforward demographic explanation of the data is provided by a mixture of the models considered by the authors.

Mekel-Bobrov *et al*. (*1*) and Evans *et al*. (*2*) examined models of human population growth [models 2 to 5, Supporting Online Material for (*1*)] and models of structured populations over space (models 6 to 9), but they did not consider a population that is both structured and growing. The data are consistent, however, with a demographic history where the population is initially structured, following, for example, a founder event, and subsequently undergoes population growth. If an ancestral subpopulation makes a large contribution to the present-day population, drift in that subpopulation can result in high frequencies of a particular haplotype, while other subpopulations account for the allelic diversity observed among the remaining haplotypes. Simulations of simplified demographic models of this nature are provided in Table 1. Although we focused on subpopula-

Table 1. Coalescent simulations can generate the observed haplotype data without selection. A thousand replicate coalescent simulations traced the ancestry of genes back in time, conditioned upon the observed number of polymorphic sites, local recombination and gene conversion rates, and sample size as in (*1*, *2*). Percentages give the fraction of simulations in which the overall level of homozygosity for the most common haplotype (columns a and c) or the frequency of the most common haplotype (columns b and d) equaled or exceeded the observations in (*1*, *2*). Even though selection was absent, the bolded cases were often consistent with the observed levels of homozygosity and high-frequency haplotypes at *ASPM* and *microcephalin*. In models (i) to (iii), a subpopulation split off from a core population, 1000 generations before the present, and grew exponentially from an initial effective population size of n_s to 10^7 diploid individuals in the present. The core population size was (i) constant at 10^5 (stable core), (ii) grew from an historical size of 10^4 individuals 1000 generations ago to 10^7 at present (growing core), or (iii) grew from 10^4 individuals 5000 generations ago to 10^7 at present (extended growth core). Model (iv) was equivalent to model (iii) except that the fission event occurred earlier (at 5000 generations), with the subpopulation remaining at size n_s until 1000 generations ago, after which it grew to 10^7 at present (early fission). The effect of migration was also explored using model (iv); the results were unchanged for low migration rates but fell toward zero when at least 0.025% of the core and subpopulation were composed of migrants every generation (requiring substantial migration between Africa and Eurasia). Samples were drawn equally from the core and subpopulation.

Simulation	*ASPM*			*Microcephalin**	
	n_s	a	b	c	d
(i) Stable core	10^1	**97.2%**†	**98.0%**	**91.9%**	**92.0%**
	10^2	**12.6%**	**12.7%**	3.5%	3.7%
(ii) Growing core	10^1	**90.4%**	**91.2%**	**17.0%**‡	**14.5%**
	10^2	4.2%	3.9%	0.1%	0.1%
(iii) Extended growth core	10^1	**96.5%**§	**96.7%**	**76.6%**	**75.2%**
	10^2	**8.8%**	**8.4%**	0.9%	0.7%
(iv) Early fission	10^1	**99.3%**	**99.4%**	**79.6%**	**79.6%**
	10^2	**98.0%**	**98.6%**	**62.2%**	**61.6%**
	10^3	**40.3%**	**40.7%**	3.9%	3.1%
Observed data		7/90	37/180	18/89	59/178

*We used the local recombination rate of 1.9 cM/MB rather than the genome-wide rate of 1 cM/MB used in the code of (*2*). Using a higher recombination rate makes it less likely to observe the data. †Code using the software package ms (*6*): ./ms 180 1 -s 166 -r 472.3324 62149 -c 1.1 290 -I 2 90 90 -n 2 100.00 -G 0 -g 2 5526.20422 -eG 0.0025 0 -ej 0.0025 2 1. ‡Code using the software package ms (*6*): ./ms 178 1 -s 220 -r 22060.52 29027 -c 1 100 -I 2 89 89 -n 2 1.00 -G 276310.2112 -g 2 552620.422 -eG 0.000025 0 -ej 0.000025 2 1. §Code using the software package ms (*6*): ./ms 180 1 -s 166 -r 47233.24 62149 -c 1.1 290 -I 2 90 90 -n 2 1.00 -G 55262.04223 -g 2 552620.422 -eG 0.000125 0 -ej 0.000025 2 1.

[1]Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. [2]Computational and Molecular Population Genetics (CMPG), Zoological Institute, University of Bern, 6 Baltzerstrasse, CH-3012 Bern, Switzerland.

*To whom correspondence should be addressed. E-mail: otto@zoology.ubc.ca

†Authors are listed alphabetically.

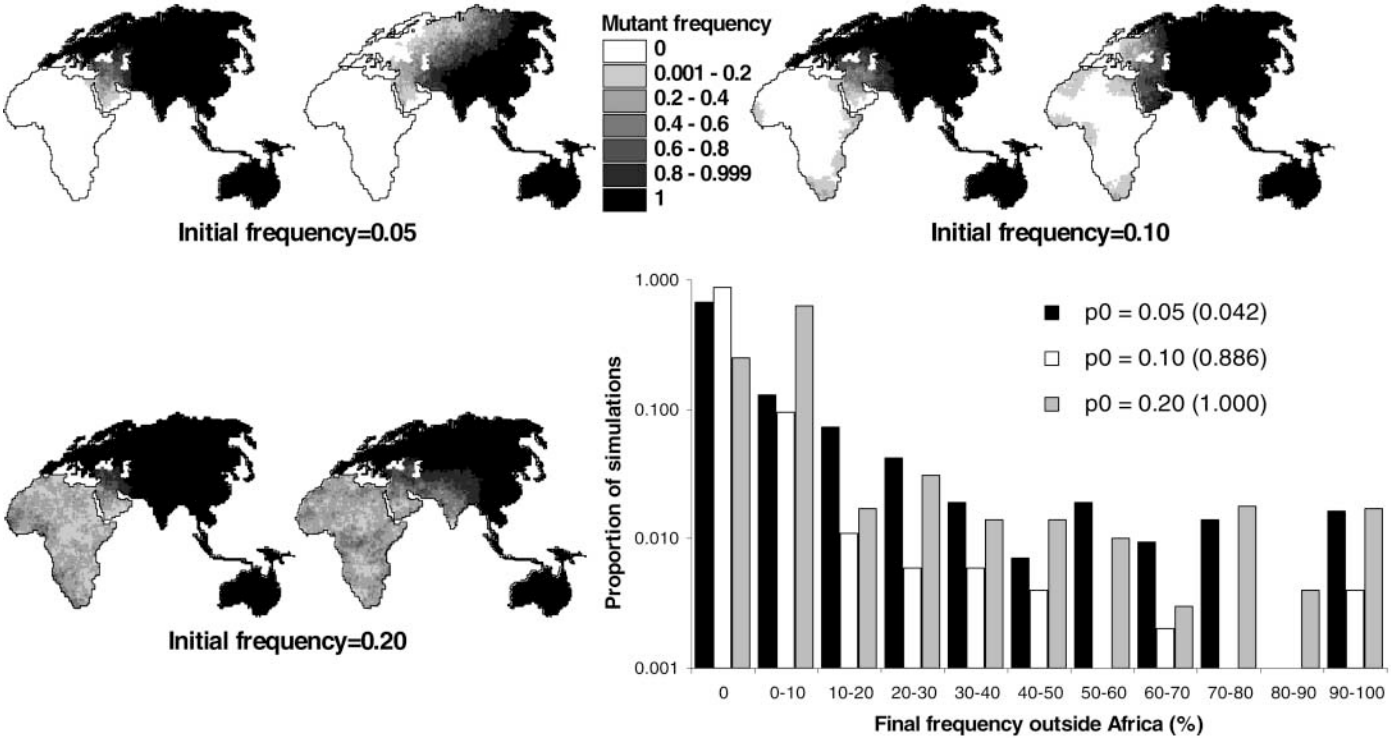

Fig. 1. Spatial simulations are consistent with observed clines in allele frequency. The Old World was modeled as a two-dimensional stepping-stone divided into 9226 demes of 100 by 100 km, each with a carrying capacity of 50 diploid individuals (100 genes). Each generation, each gene had a 25% probability of migrating to an adjacent deme. We modeled a range expansion out-of-Africa through the Sinai Peninsula that started 1400 generations ago (~40,000 years), as in (*7*). At that time, we assumed that Africa was fully occupied, with all demes being at carrying capacity and harboring an allele with a uniform initial frequency p_0 of 0.05, 0.1, or 0.2. We simulated the stochastic evolution of this allele in Africa and in Eurasia until today, keeping only simulations where the allele has persisted until today somewhere in the Old World. (The proportion of simulations kept was equal to 0.042, 0.886, and 1.000 for p_0 values of 0.05, 0.1, or 0.2, respectively, as reported in the legend of the histogram.) The histogram (**lower right**) reports the frequency distribution of the allele in non-African populations, demonstrating that the allele can reach very high frequencies by drift and surf (*4*) outside of Africa. The distributions were obtained from 1000 simulations, except for $p_0 = 0.05$, where only 424 simulations out of 10,000 were successful. For each parameter value, the maps show two spatial frequency distributions where alleles reached a minimum average frequency of 50% outside Africa, illustrating how alleles with low initial frequencies in Africa can reach high frequency by colonization and drift in non-African populations, while remaining at low frequencies within Africa.

tions that were small when initially formed (10 to 1000 individuals), reductions in effective population size due to disease, repeated founder effects, and/or variability in reproductive success could generate similar patterns even in larger subpopulations. We conclude that human demographic models with structure followed by population growth can explain the haplotype frequency data at *ASPM* and *microcephalin* without invoking selection.

A second demographic model that can explain the data without requiring selection involves population growth across space, as occurred during the range expansion of humans (*3*, *4*). Population growth over space can be described by a wave of advance. The few individuals on the wave of advance will have, by luck, disproportionate numbers of descendants. Haplotypes that happen to find themselves in the wave front can rise to high frequency by chance alone and surf on the wave of advance (*4*). Indeed, an explicit spatial model of human demography, with expansion out of Africa starting around 40,000 years ago can also generate a high frequency of a single haplotype in non-African populations (Fig. 1).

These models do not predict that every gene should exhibit a high frequency haplotype. It is a matter of chance whether one haplotype will drift up in frequency during the growth of a subpopulation or the spread of a wave front. An empirically important question is how often this pattern is observed in putatively neutral regions of the human genome. If few neutral regions exhibit high-frequency haplotypes, then there would be an empirical basis for arguing against the demographic processes explored in this comment.

In summary, the high haplotype frequency, high levels of homozygosity, and spatial patterns observed by Mekel-Bobrov *et al*. (*1*) and Evans *et al*. (*2*) can be generated by demographic models of human history involving a founder effect out-of-Africa and a subsequent demographic or spatial population expansion, a very plausible scenario (*5*). Thus, there is insufficient evidence for ongoing selection acting on *ASPM* and *microcephalin* within humans.

References and Notes

1. N. Mekel-Bobrov *et al.*, *Science* **309**, 1720 (2005).
2. P. D. Evans *et al.*, *Science* **309**, 1717 (2005).
3. C. A. Edmonds, A. S. Lillie, L. L. Cavalli-Sforza, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 975 (2004).
4. S. Klopfstein, M. Currat, L. Excoffier, *Mol. Biol. Evol.* **23**, 482 (2006).
5. S. Ramachandran *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15942 (2005).
6. R. R. Hudson, *Bioinformatics* **18**, 337 (2002).
7. N. Ray, M. Currat, P. Berthier, L. Excoffier, *Genome Res.* **15**, 1161 (2005).
8. This work was partly supported by Swiss National Science Foundation grant 3100A0-100800 to L.E. and by National Science and Engineering Research Council of Canada Discovery grants to W.M., S.P.O., and M.C.W.

16 November 2005; accepted 19 April 2006
10.1126/science.1122712