

**Comment on "Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa"**

Michael Cysouw, *et al.*  
*Science* **335**, 657 (2012);  
DOI: 10.1126/science.1208841

*This copy is for your personal, non-commercial use only.*

**If you wish to distribute this article to others**, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

**Permission to republish or repurpose articles or portions of articles** can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of February 13, 2012):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/335/6069/657.2.full.html>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/content/suppl/2012/02/09/335.6069.657-b.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/335/6069/657.2.full.html#related>

This article **cites 9 articles**, 4 of which can be accessed free:

<http://www.sciencemag.org/content/335/6069/657.2.full.html#ref-list-1>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/335/6069/657.2.full.html#related-urls>

This article appears in the following **subject collections**:

Evolution

<http://www.sciencemag.org/cgi/collection/evolution>

Technical Comments

[http://www.sciencemag.org/cgi/collection/tech\\_comment](http://www.sciencemag.org/cgi/collection/tech_comment)



minimizing the Bayesian information criterion (BIC) (7) of the regression between phonemic inventory size and geographic distance, including further control variables. Atkinson selects those locations at most four BIC units away from this optimum as having considerable support in being the origin of the expansion. A quick computation shows that this implies accepting models that are at most  $e^2 \sim 7.4$  times less likely than the optimal one (2), which strikes us as rather arbitrary. Further, this BIC optimization method necessarily “spreads” any origin across a contiguous geographic region, even in the case of totally random data (fig. S12).

Notwithstanding this criticism, we replicated Atkinson’s method using the UPSID data, but instead of a single origin in western Africa, we found two separate “origins,” one in eastern Africa and one in the Caucasus (fig. S10). The BIC+4 range of possible origins covers a large area, including also the Middle East and southern Africa. Although this finding does not necessarily contradict an expansion from Africa, it does not provide clear support in its favor, either. Further, adding a quadratic distance factor to the model substantially improves the fit and suggests an alternative origin located in New Guinea with a small phoneme inventory (fig. S10). Even more problematic, when we apply the original method to other inventory-like linguistic characteristics from WALS (Fig. 1), we find origins of global clines all over the world, not just in Africa, and not always corresponding to the highest structural “complexity” (fig. S11). Therefore, the observation of an Africa-based phoneme inventory cline does not generalize to other linguistic characteristics of a similar kind.

Third, Atkinson’s explanation crucially depends on a positive correlation between phonemic inventory size and speaker community size,

which, unfortunately and contrary to his own claim [see figure S1 in (1)], does not hold for small populations when using UPSID data ( $r = 0.04$ ,  $P = 0.64$ ) (fig. S6). This correlation reaches significance at the 5% level only when languages with speaker populations above  $10^5$  are included, but such large speaker community sizes only arose in the context of agriculture long after the peopling of most of the globe (8).

Fourth, the geographic patterning of tone might be influenced by a genetic bias postdating the out-of-Africa migration by tens of thousands of years (9). Moreover, consonant inventories (and to a lesser extent, vowel inventories) do not seem to be phylogenetically stable enough (10) to conserve the kind of deep signal necessary for the proposed scenario, whereas other, more stable, features show non-African “origins” (fig. S11).

Finally, we believe that Atkinson’s interpretation of the reported worldwide cline in terms of a linguistic serial founder effect is problematic because of the extraordinary large amount of horizontal processes affecting language (11, 12) and because the underlying mechanism proposed by Atkinson is linguistically not plausible (13). Further, global clines in linguistics, like in genetics, do not necessarily equate with a serial founder effect and can have other causes (2, 14).

Summarizing, the reported linguistic evidence for an expansion from Africa is unfortunately an artifact of various methodological decisions and biased interpretations. We consider this to be unfortunate, because we would very much welcome any new insights into human prehistory based on geographic patterns of linguistic diversity. In this respect, we applaud Atkinson for further developing this approach (15) and renewing the methodological discussion, because only explicit testing and refutation opens the way for the formulation of more specific hypotheses con-

cerning the identification of possible linguistic signatures of ancient demographic events.

#### References and Notes

1. Q. D. Atkinson, *Science* **332**, 346 (2011).
2. Materials and methods are available as supporting material on *Science* Online.
3. I. Maddieson, K. Precoda, The UCLA Phonological Segment Inventory Database (1990); [http://web.phonetik.uni-frankfurt.de/upsid\\_info.html](http://web.phonetik.uni-frankfurt.de/upsid_info.html).
4. M. Haspelmath, M. S. Dryer, D. Gil, B. Comrie, Eds., *The World Atlas of Language Structures* (Oxford Univ. Press, Oxford, 2005).
5. A. Manica, W. Amos, F. Balloux, T. Hanihara, *Nature* **448**, 346 (2007).
6. L. Betti, F. Balloux, W. Amos, T. Hanihara, A. Manica, *Proc. Biol. Sci.* **276**, 809 (2009).
7. G. E. Schwarz, *Ann. Stat.* **6**, 461 (1978).
8. S. Mithen, *After the Ice* (Orion Books, London, 2003).
9. D. Dediu, D. R. Ladd, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10944 (2007).
10. D. Dediu, *Proc. Biol. Sci.* **278**, 474 (2011).
11. S. G. Thomason, T. Kaufman, *Language Contact, Creolization, and Genetic Linguistics* (Univ. of California Press, Berkeley, CA, 1988).
12. T. F. Jaeger, P. Graff, W. Croft, D. Pontillo, *Linguist. Typol.* **15**, 281 (2011).
13. C. Bownen, *Linguist. Typol.* **15**, 207 (2011).
14. I. Maddieson, T. Bhattacharya, D. E. Smith, W. Croft, *Linguist. Typol.* **15**, 267 (2011).
15. J. Nichols, *Linguistic Diversity in Space and Time* (Univ. of Chicago Press, Chicago, 1992).

**Acknowledgments:** The authors declare no conflict of interests. M.C., D.D., and S.M. designed the research; M.C. and S.M. organized the data; M.C. and D.D. performed the analysis and wrote the paper. M.C. and S.M. are funded by European Research Council starting grant 240816 QuantHistLing.

#### Supporting Online Material

[www.sciencemag.org/cgi/content/full/335/6069/657-b/DC1](http://www.sciencemag.org/cgi/content/full/335/6069/657-b/DC1)  
Materials and Methods  
SOM Text  
Figs. S1 to S12  
Table S1  
References

24 May 2011; accepted 3 January 2012  
10.1126/science.1208841