Response to Comment on "Human-Specific Gain of Function in a Developmental Enhancer"
Shyam Prabhakar, *et al.*
Science **323**, 714d (2009);
DOI: 10.1126/science.1166571

Response to Comment on "Human-Specific Gain of Function in a Developmental Enhancer"

Shyam Prabhakar,[1]* Axel Visel,[1] Jennifer A. Akiyama,[1] Malak Shoukry,[1] Keith D. Lewis,[1]† Amy Holt,[1] Ingrid Plajzer-Frick,[1] Harris Morrison,[2] David R. FitzPatrick,[2] Veena Afzal,[1] Len A. Pennacchio,[1,3] Edward M. Rubin,[1,3]‡ James P. Noonan[1]‡§

Duret and Galtier argue that human-specific sequence divergence and gain of function in the *HACNS1* enhancer result from deleterious biased gene conversion (BGC) with no contribution from positive selection. We reinforce our previous conclusion by analyzing hypothesized BGC events genomewide and assessing the effect of recombination rates on human-accelerated conserved noncoding sequence ascertainment. We also provide evidence that AT → GC substitution bias can coexist with positive selection.

Duret and Galtier (*1*) suggest that biased gene conversion (BGC) alone drove the rapid human-specific evolution of human-accelerated conserved noncoding sequence 1 (*HACNS1*) (*2*, *3*) and that positive selection played no role. BGC is a nonadaptive, recombination-driven mechanism hypothesized to increase the rate of AT → GC substitutions in a given interval (*4*, *5*). In support, it has been shown that the AT → GC substitution rate is more strongly correlated with male recombination rate than with GC content (*6*). Clusters [100 to 1000 base pairs (bp)] of human-specific substitutions occurring in regions of elevated male recombination rate show an AT → GC bias, which suggests that BGC may have contributed to some substitution hotspots in the genome (*5*). These correlations are stronger at the 1 Mb scale than at finer scales, likely because fine-scale recombination rates are less stable over time than coarse-scale rates (*5*, *7*, *8*). In addition, human-accelerated regions (HARs) identified in a whole-genome screen of conserved genomic elements have an excess of AT → GC substitutions and are enriched in broad regions of high average recombination (*4*, *9*, *10*).

In light of these findings, Duret and Galtier's observation that some HARs and HACNSs could be due solely to BGC is reasonable (*4*). However, their present attempt to extrapolate this argument to *HACNS1* requires several unsupported assumptions regarding the ascertainment of HACNSs compared with HARs and the known distribution of recombination hotspots in the genome. We examine the authors' claims below.

The robust gain of limb expression in *HACNS1* is not in itself evidence of adaptation. However, the most likely effect of deleterious substitution in the enhancer would be to randomly distort the existing function rather than to strengthen the ancestral expression pattern while introducing a new, robust expression domain. The precise, highly spatially restricted gain of function in *HACNS1* is exactly what one would expect from an adaptive process that altered a specific functional module in the enhancer, as suggested by the pattern of substitutions we observed, while leaving the ancestral expression domains intact. Therefore, we find the adaptive hypothesis qualitatively more parsimonious than the authors' alternative (*1*).

Overrepresentation of HARs in subtelomeric regions, which display high recombination rates, has been cited as evidence that some HARs are due to BGC rather than to positive selection (*4*, *10*). However, in assuming that HACNSs suffer from a similar systematic bias, Duret and Galtier misunderstand how HACNSs were identified. In contrast to HARs, which were ascertained relative to whole-genome averaged substitution rates, HACNSs were defined based on their acceleration relative to the local neutral substitution rate (*2*, *9*). Thus, the method used to identify HACNSs mitigates the local average effect of recombination-driven BGC on substitution rates. Although some HACNSs may be due to BGC, a comparison of HACNS distribution relative to male-specific recombination rates demonstrates that HACNSs as a class show no enrichment in regions of high recombination (Table 1) (*11*). Moreover, the fraction of conserved noncoding sequences within a 10-Mb window centered on *HACNS1* that are HACNSs (0.8%) is lower than the whole-genome average (0.9%). The elevated male-specific recombination rate around *HACNS1* thus has no bearing on whether *HACNS1* is affected by BGC.

Because HACNSs show no correlation with coarse-scale recombination rates, Duret and Galtier's claim that BGC explains human-specific acceleration in *HACNS1* rests on a single point: the excess of AT → GC substitutions in the enhancer. As we discussed in (*3*), this AT → GC excess is not confined to the 546-bp *HACNS1* element; instead, it extends into the nonconserved flanking regions, covering ~5 kb in all (*3*). However, the overall human-specific substitution rate is close to the local (1 Mb) average rate across nearly this entire region, with one critical exception: the sharp spike of 13 substitutions in a highly constrained 81-bp subregion within *HACNS1*.

Duret and Galtier (*1*) propose a nonparsimonious "BGC only" scenario for the spatial profile of substitutions in and around *HACNS1*: a sharp (~80 bp), extremely strong recombination hotspot embedded in a larger (~5 kb) weak hotspot. In support, they claim that such a spatial profile is "in agreement with current knowledge about the spatiotemporal distribution of recombination in humans" and reference Myers *et al.* (*12*). However, this is a misinterpretation, because that study analyzed recombination rates at a low resolution of ~2 kb and therefore cannot provide insight on the likelihood of ~80-bp hotspots within larger "warmspots." Moreover, the authors' proposed BGC-driven cluster of 13 substitutions in 81 bp is implausibly narrow and dense, even when compared with the 200 most extreme putative BGC-driven substitution hotspots in the genome, which average ~370 bp and have an average substitution density equivalent to <5 substitutions in 81 bp (*5*).

We reject the authors' assumption that BGC and adaptation are mutually exclusive. The highly localized spike in the substitution rate in *HACNS1* is more consistent with positive selection than with BGC. However, the AT → GC substitution

Table 1. Lack of HACNS enrichment in regions of high male-specific recombination (*11*).

	All CNSs	HACNSs	Fisher two-sided *P*
Total in genome	110,549	992	
No. in regions with male recomb ≥2.0 cM	13,542 (12.3%)	135 (13.6%)	0.25
No. in regions with male recomb ≥2.5 cM	9,330 (8.44%)	87 (8.77%)	0.73

[1]Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [2]MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK. [3]United States Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA.

*Present address: Computational and Mathematical Biology, Genome Institute of Singapore 138672, Singapore.
†Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.
‡To whom correspondence should be addressed. E-mail: EMRubin@lbl.gov (E.M.R.); james.noonan@yale.edu (J.P.N.)
§Present address: Department of Genetics, Yale University School of Medicine, New Haven, CT 06520, USA.

bias in the rapidly evolving segment of *HACNS1* and in the ~5-kb flanking sequence where the overall substitution rate is not elevated suggests that BGC may have influenced the evolution of this enhancer. The substitution pattern in *HACNS1* may thus be due to positive selection synergizing with a moderate BGC domain. A recombination hotspot of mild strength, that extended over ~5 kb and elevated the fraction of AT → GC substitutions without increasing the overall substitution rate, would explain the observed substitution pattern in the neutral sequence surrounding *HACNS1*. Within *HACNS1*, positive selection for new GC-rich transcription factor binding sites or loss of AT-rich sites likely amplified a modest AT → GC fixation bias caused by BGC, producing a sharp increase in the substitution rate within an ~80 bp cluster of binding sites that generated the human-specific gain of limb expression.

Evidence that an extreme AT → GC substitution bias can coexist with positive selection comes from a 382-bp putative BGC event (chr18:897,475-897,857 in hg18) identified by Dreszer *et al*. (*5*), which contains 19 substitutions and overlaps exon 2 of *ADCYAP1*. In parallel to *HACNS1*, the human-specific substitutions comprising this BGC event are significantly clustered in functional sequence: The 132-bp exon contains 13/19 substitutions (Fisher's exact test, $P = 0.002$), despite being conserved across terrestrial vertebrate genomes. All 13 substitutions are AT → GC and produce 11 amino acid changes. The human-specific nonsynonymous substitution rate in the exon is well above the synonymous rate (dN/dS = 3.2), a classic signature of adaptive evolution (*13*). It is difficult to imagine how neutral or deleterious BGC could selectively favor nonsynonymous substitutions. We hypothesize that accelerated evolution in *ADCYAP1* and *HACNS1* are due to a common mechanism: synergy between BGC and positive selection producing a cluster of AT → GC substitutions at functional sites.

In summary, we find no basis for the authors' claim that the gain of function in *HACNS1* is solely due to BGC. Our analysis reinforces our initial conclusion, that positive selection likely played a role in *HACNS1* evolution.

References and Notes

1. L. Duret, N. Galtier, *Science* **323**, 714 (2009); www.sciencemag.org/cgi/content/full/323/5915/714c.
2. S. Prabhakar, J. P. Noonan, S. Pääbo, E. M. Rubin, *Science* **314**, 786 (2006).
3. S. Prabhakar *et al*., *Science* **321**, 1346 (2008).
4. N. Galtier, L. Duret, *Trends Genet.* **23**, 273 (2007).
5. T. R. Dreszer, G. D. Wall, D. Haussler, K. S. Pollard, *Genome Res.* **17**, 1420 (2007).
6. L. Duret, P. F. Arndt, *PLoS Genet.* **4**, e1000071 (2008).
7. C. C. Spencer *et al*., *PLoS Genet.* **2**, e148 (2006).
8. W. Winckler *et al*., *Science* **308**, 107 (2005).
9. K. S. Pollard *et al*., *Nature* **443**, 167 (2006).
10. K. S. Pollard *et al*., *PLoS Genet.* **2**, e168 (2006).
11. A. Kong *et al*., *Nat. Genet.* **31**, 241 (2002).
12. S. Myers, L. Bottolo, C. Freeman, G. McVean, P. Donnelly, *Science* **310**, 321 (2005).
13. Z. Yang, *Mol. Biol. Evol.* **24**, 1586 (2007).
14. Research was done under Department of Energy contract DE-AC02-05CH11231, University of California, E. O. Lawrence Berkeley National Laboratory, and supported by National Heart, Lung and Blood Institute grant HL066681 and National Human Genome Research Institute grant HG003988 (L.A.P.); the Agency for Science, Technology, and Research of Singapore (S.P.); an American Heart Association postdoctoral fellowship (A.V.); and NIH National Research Service Award fellowship 1-F32-GM074367 and the Department of Genetics, Yale University School of Medicine (J.P.N.).

28 October 2008; accepted 12 January 2009
10.1126/science.1166571