



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Human-Specific Gain of Function in a Developmental Enhancer

Citation for published version:

Prabhakar, S, Visel, A, Akiyama, JA, Shoukry, M, Lewis, KD, Holt, A, Plajzer-Frick, I, Morrison, H, Fitzpatrick, DR, Afzal, V, Pennacchio, LA, Rubin, EM & Noonan, JP 2008, 'Human-Specific Gain of Function in a Developmental Enhancer' *Science*, vol 321, no. 5894, pp. 1346-1350. DOI: 10.1126/science.1159974

Digital Object Identifier (DOI):

[10.1126/science.1159974](https://doi.org/10.1126/science.1159974)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Science

Publisher Rights Statement:

NIH public access Author manuscript

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Published in final edited form as:

Science. 2008 September 5; 321(5894): 1346–1350. doi:10.1126/science.1159974.

Human-specific gain of function in a developmental enhancer

Shyam Prabhakar^{1,3}, Axel Visel¹, Jennifer A. Akiyama¹, Malak Shoukry¹, Keith D. Lewis^{1,4}, Amy Holt¹, Ingrid Plajzer-Frick¹, Harris Morrison⁵, David R. FitzPatrick⁵, Veena Afzal¹, Len A. Pennacchio^{1,2}, Edward M. Rubin^{1,2}, and James P. Noonan^{1,6}

¹ Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

² United States Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA.

⁵ MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK.

Abstract

Changes in gene regulation are thought to have contributed to the evolution of human development. However, *in vivo* evidence for uniquely human developmental regulatory function has remained elusive. In transgenic mice, a conserved noncoding sequence (*HACNS1*) that evolved extremely rapidly in humans acted as an enhancer of gene expression that has gained a strong limb expression domain relative to the orthologous elements from chimpanzee and rhesus macaque. This gain of function was consistent across two developmental stages in the mouse and included the presumptive anterior wrist and proximal thumb. *In vivo* analyses with synthetic enhancers, in which human-specific substitutions were introduced into the chimpanzee enhancer sequence or reverted in the human enhancer to the ancestral state, indicated that 13 substitutions clustered in an 81-basepair module otherwise highly constrained among terrestrial vertebrates were sufficient to confer the human-specific limb expression domain.

Genome sequence changes that altered the molecular machinery of development likely facilitated the evolution of uniquely human morphological traits (1,2). Although these genetic modifications remain largely unidentified, it has long been thought that changes in gene expression, due to positive selection for nucleotide substitutions that modified the activity of *cis*-regulatory elements, played a prominent role (3). Several cases of putatively adaptive sequence change, including polymorphisms among human populations and apparently fixed differences between humans and other primates, have been shown to affect *in vitro* promoter or enhancer function in cell line reporter assays (4-7). However, the impact of human-specific nucleotide substitutions on the *in vivo* activity of developmental regulatory elements remains obscure.

In vivo analyses of evolutionarily conserved noncoding sequences have revealed them to be enriched in *cis*-regulatory transcriptional enhancers that confer specific expression patterns during development (8-11). Recent efforts have identified conserved noncoding sequences that evolved rapidly on the human lineage, but it is not known if these sequences include regulatory

Correspondence and requests for materials should be addressed to E.M.R. (EMRubin@lbl.gov) or J.P.N. (james.noonan@yale.edu).

³Present address: Computational and Mathematical Biology, Genome Institute of Singapore 138672, Singapore.

⁴Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.

⁶Present address: Department of Genetics, Yale University School of Medicine, New Haven, CT 06520, USA.

Supporting Online Material

Materials and Methods

Figs. S1 to S6

Table S1

References

elements with altered activities in humans (12-15). Here we focus on functionally characterizing the most rapidly evolving human noncoding element yet identified, which we termed human-accelerated conserved noncoding sequence 1 (*HACNS1*) (12). Although this 546 basepair element is highly constrained in all sequenced terrestrial vertebrate genomes, it has accumulated 16 human-specific sequence changes in the ~6 million years since the human-chimpanzee split (Fig. 1A). We evaluated the significance of this evolutionary acceleration using a test statistic that represents the log-likelihood, or information theoretic “surprisal,” of observing the human sequence given the orthologous sequences from multiple terrestrial vertebrates. Assuming *HACNS1* is under functional constraint in humans, its rapid divergence is highly unexpected based on its strong conservation in these other species (surprisal test P value = $9.2e-12$; 16). This divergence also significantly exceeds the ~4 substitutions expected if *HACNS1* were evolving at the neutral substitution rate in humans (surprisal test P value = $1.3e-6$; 16). One explanation for this dramatic acceleration is that *HACNS1* has undergone several instances of positive selection during human evolution that may have altered its function.

To test this hypothesis, we evaluated the ability of *HACNS1* and its orthologs from chimpanzee and rhesus macaque to function as transcriptional enhancers during development using a transgenic mouse enhancer assay in which the activity of each sequence is assessed through a β -galactosidase (*lacZ*) reporter gene coupled to a minimal *Hsp68* promoter (17). We initially examined the potential enhancer activity of *HACNS1* at embryonic day 11.5 (E11.5). We tested a 1.2 kb DNA fragment encompassing *HACNS1* that also contained nonconserved sequences flanking the element, in order to include possible functional sequences near *HACNS1* not detected by conservation (Table S1). At E11.5, the human element drove strong and reproducible reporter gene expression in the anterior limb bud, pharyngeal arches, developing ear and eye, suggesting that *HACNS1* acts as a robust enhancer during development (Figs. 1B, C and S1). In striking contrast to the highly reproducible staining driven by the human enhancer, which extended into the most distal region of the anterior limb bud in 5 out of 6 *lacZ*-positive embryos (Figs. 1C and S1; *HACNS1* embryos 1-5), the chimpanzee and rhesus orthologs failed to drive reproducible reporter gene expression in the distal limb bud, though they did drive moderately reproducible expression at the base of the limb (Figs. 1B, C and chimpanzee enhancer embryos 1, 2, 6-8 and rhesus enhancer embryos 1-4, 6-8 and 10 in Fig. S1). Two of the embryos that were transgenic for the chimpanzee ortholog and showed this pattern also exhibited diffuse, low-level staining that extended into the anterior limb, suggesting that the chimpanzee enhancer may possess a weak capacity to drive expression in this structure (embryos 6 and 7 in Fig. S1). However, this infrequent pattern was in stark contrast to the strong and highly reproducible pattern of the human enhancer. Furthermore, pharyngeal arch, eye and ear expression was less reproducible and, where present, generally weaker in multiple positive embryos for both nonhuman orthologs, suggesting additional sites of reduced overall enhancer activity compared to human. To assess the *HACNS1* limb expression pattern at higher resolution, we sectioned *HACNS1* transgenic embryos and found that staining in the forelimb was restricted to the mesenchyme, forming a continuous expression domain that extended deep into the limb bud along the antero-posterior axis at the handplate and shoulder while remaining more anterior in between (Fig. S3). These results provide evidence that the human-specific sequence changes in *HACNS1* have resulted in a gain of function in this otherwise highly conserved enhancer, increasing its overall robustness and producing a strong human-specific expression domain in the anterior limb bud mesenchyme at E11.5. Since the chimpanzee and rhesus orthologs yield similar patterns to each other and show consistent differences compared to human, a parsimonious conclusion would be that the chimpanzee and rhesus patterns reflect the ancestral primate state from which the human-specific pattern has evolved.

To explore the activity of *HACNS1* at a more advanced stage of limb development, we compared the expression patterns of the human, chimpanzee and rhesus enhancers in E13.5

transgenic mouse embryos. At this stage, the human element continued to drive reproducible reporter gene expression in the anterior developing forelimb, particularly in the shoulder and the anterior junction of the forearm and handplate in 11 of 12 positive embryos (Figs. 2A, B, C and embryos 1-10 and 12 in Fig. S2). In 4 of these embryos the reporter gene activity extended into the future anterior-most digit of the forelimb (Fig. 2B). Similar expression, although with weaker staining, was also observed in the corresponding structures in the hindlimb. Imaging of lacZ staining in a representative *HACNS1* transgenic embryo using optical projection tomography (OPT; 18) revealed that the anterior expression evident in the whole mount extended deep inside the limb at the forearm-handplate junction (Fig. S4). The orthologous chimpanzee and rhesus elements failed to drive reproducible expression in the distal limbs at this time point, though a subset of positive embryos in each case showed reporter gene expression in the shoulder region of the limb bud, thus recapitulating the proximal tip of the expression domain of the human enhancer (4/10 for chimpanzee, Fig. 2A and chimpanzee enhancer embryos 2-5 in Fig. S2, and 3/12 for rhesus, Fig. 2A and rhesus enhancer embryos 1, 5 and 6 in Fig. S2). OPT imaging confirmed the absence of reproducible lacZ staining inside the distal limb in representative embryos transgenic for the rhesus and chimpanzee enhancers (Fig. S4). These results indicate that the human-specific enhancer activity persists across multiple developmental stages, and suggest that the robust anterior limb expression pattern of *HACNS1* evolved from a weaker ancestral pattern that is largely confined to the base of the limb bud, as evident in the activities of the chimpanzee and rhesus enhancers at both time points.

We next sought to identify human-specific sequence changes responsible for the functional change in the human enhancer. While the 16 human-specific substitutions within the 546 bp conserved region corresponding to *HACNS1* are the most striking feature of the 1.2 kb orthologous segments we initially tested for enhancer function, these segments also included ~650 bp of nonconserved DNA containing additional human-chimpanzee sequence differences. To isolate the effect of the substitutions within *HACNS1* on enhancer function, we synthesized a chimeric 1.2 kb enhancer in which we transferred all 16 substitutions into the chimpanzee sequence background (16). This “humanized” chimpanzee enhancer produced an E11.5 expression pattern nearly identical to that of the native human enhancer, suggesting that the human-specific sequence changes within *HACNS1* are responsible for the gain of function we observe (8/8 embryos; Figs. 3D and S1). Strikingly, these human-specific substitutions are significantly clustered: 13 of 16 substitutions occur within an 81 basepair region of the 546-bp conserved element (permutation test P value = $1.7e-7$; 16), suggesting that this region may be particularly relevant to the human-specific function of *HACNS1* (Figs. 1A and 3A). To test this hypothesis, we synthesized a chimeric 1.2 kb enhancer in which the 13 clustered human substitutions were introduced into the chimpanzee sequence background (Table S2). At E11.5, this element produced an anterior limb bud pattern highly similar to *HACNS1* (6/6 positive embryos; Figs. 3 and S1). We also performed the reciprocal experiment, synthesizing a complementary chimeric enhancer where we replaced the 13 human-specific nucleotides in the human enhancer sequence with their putatively ancestral orthologs from chimpanzee. This “reverted” enhancer yielded a pattern very similar to the chimpanzee and rhesus enhancer patterns shown in Figure 1B, with expression in the anterior limb bud greatly reduced or absent (Figs. 3 and S1). These results confirm the robustness of the functional differences we observed between *HACNS1* and its chimpanzee and rhesus orthologs and indicate that the *HACNS1* anterior limb bud pattern is largely attributable to one or more of the 13 clustered human-specific substitutions we identified. To further dissect the functional contribution of these substitutions, we introduced independent groups of 6 substitutions and 3 substitutions into the chimpanzee enhancer sequence (Fig. S5). These enhancers drove variable expression in the anterior limb bud, suggesting that at least two human-specific substitutions are required for the gain of function in *HACNS1*.

The precise molecular mechanism by which the substitutions in *HACNS1* confer the human-specific expression pattern remains to be determined. Computational analysis of predicted transcription factor binding sites in *HACNS1* and its nonhuman orthologs suggested that multiple sites have been gained and lost in this enhancer during human evolution (Fig. S6). Predicted human-specific binding sites for the developmental transcription factors PAX9 and ZNF423 may contribute to *HACNS1* enhancer activity, as the known expression pattern of PAX9 in the mouse limb overlaps the human-specific limb domain of *HACNS1* at E11.5 and E13.5, and ZNF423 is expressed in the mouse handplate mesenchyme from E10.5 through E12.5 (19-21).

Several lines of evidence suggest that the functional changes in *HACNS1* are due to adaptive evolution. The rate of human-specific accelerated evolution in *HACNS1* is more than four times the local neutral rate. Moreover, this rapid evolution is not consistent with biased gene conversion (BGC), a neutral mechanism postulated to cause hotspots of accelerated evolution in the genome by increasing the local fixation rate of AT to GC substitutions (22,23). Under the neutral BGC hypothesis, one would expect an increase in the overall substitution rate across the entire region of increased AT to GC substitution (23). An excess of AT to GC substitutions is indeed present in *HACNS1* (binomial test P value = 1.1×10^{-4} ; 16), and the element lies in a ~5-kb genomic region enriched in such substitutions (Fig. 4). However, the human-specific substitution rate is elevated only in the narrow 81-bp region in *HACNS1* described above and is close to the local average outside of this window (Fig. 4). These data, coupled with the human-specific functional changes in *HACNS1*, argue against a neutral explanation for the rapid evolution of this element in humans.

Our results evoke the hypothesis that human-specific adaptive evolution in *HACNS1* has contributed to uniquely human aspects of digit and limb patterning. The increased dexterity of the human hand compared to other primates is due to morphological differences that include rotation of the thumb toward the palm and an increase in the length of the thumb compared to the other digits (1). Human-specific changes in hindlimb morphology, such as the characteristic inflexibility and shortened digits of the human foot, facilitated habitual bipedalism. The gain of function in *HACNS1* may have influenced the evolution of these or other human limb features by altering the expression of nearby genes during limb development. *HACNS1* is located within an intron of *CENTG2*, which encodes a GTPase activating protein involved in the regulation of endosome function, and ~300 kb downstream of *GBX2*, an essential developmental transcription factor (24,25). The potential role of *CENTG2* in limb development has not been evaluated. Although mouse *Gbx2* is expressed in the developing limb, *Gbx2* null mice have not been described as showing abnormal limbs (25). However, we note that the *HACNS1* expression pattern in transgenic mice may not entirely recapitulate the precise *HACNS1* expression pattern in the human embryo. The accelerated evolution of *HACNS1* might therefore reflect selection for changes in structures other than, or in addition to, the limb. Elucidating the role of *HACNS1* in human morphological evolution requires further lines of evidence, including the analysis of *GBX2* and *CENTG2* expression during human development and the generation of *HACNS1* targeted replacement mice. Independent of these considerations, our study suggests that adaptive nucleotide substitution altered the function of a developmental enhancer in humans, and illustrates a strategy that could be employed across the genome to understand at a molecular level how human development evolved through *cis*-regulatory change.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank members of the Pennacchio and Rubin laboratories for critical comments on the manuscript. Research was done under Department of Energy Contract DE-AC02-05CH11231, University of California, E.O. Lawrence Berkeley National Laboratory. L.A.P. was supported by grant HL066681, Berkeley-Program for Genomic Applications, under the Programs for Genomic Applications, funded by the National Heart, Lung and Blood Institute, and HG003988 funded by the National Human Genome Research Institute. S.P. was additionally supported by the Agency for Science, Technology, and Research of Singapore. A.V. was supported by an American Heart Association postdoctoral fellowship. J.P.N. was supported by NIH National Research Service Award fellowship 1-F32-GM074367 and funds from the Department of Genetics, Yale University School of Medicine.

References

1. Trinkaus, E. Evolution of human manipulation. In: Jones, S.; Martin, R.; Pilbeam, D., editors. The Cambridge Encyclopedia of Human Evolution. Cambridge Univ. Press; New York: 1993. p. 346-349.
2. Carroll SB. Nature 2003;422:849. [PubMed: 12712196]
3. King MC, Wilson AC. Science 1975;188:107. [PubMed: 1090005]
4. Tournamille C, et al. Nat. Genet 1995;10:224. [PubMed: 7663520]
5. Rockman MV, et al. PLoS Biol 2005;3:e387. [PubMed: 16274263]
6. Tishkoff SA, et al. Nat. Genet 2007;39:31. [PubMed: 17159977]
7. Wray GA. Nat. Rev. Genet 2007;8:206. [PubMed: 17304246]
8. Boffelli D, Nobrega MA, Rubin EM. Nat. Rev. Genet 2004;5:456. [PubMed: 15153998]
9. Nobrega MA, Ovcharenko I, Afzal V, Rubin EM. Science 2003;302:417.
10. Pennacchio LA, et al. Nature 2006;444:499. [PubMed: 17086198]
11. Visel A, et al. Nat. Genet 2008;40:158. [PubMed: 18176564]
12. Prabhakar S, Noonan JP, Pääbo S, Rubin EM. Science 2006;314:786. [PubMed: 17082449]
13. Pollard KS, et al. Nature 2006;443:167. [PubMed: 16915236]
14. Bird CP, et al. Genome Biol 2007;8:R118. [PubMed: 17578567]
15. Bush EC, Lahn BT. BMC Evol. Biol 2008;8:17. [PubMed: 18215302]
16. Materials and methods are available as supporting material on *Science* Online.
17. Kothary R, et al. Development 1989;105:707. [PubMed: 2557196]
18. Sharpe J, et al. Science 2002;296:541. [PubMed: 11964482]
19. Neubüser A, Koseki H, Balling R. Dev. Biol 1995;170:701. [PubMed: 7649395]
20. Peters H, Neubüser A, Kratochwil K, Balling R. Genes Dev 1998;12:2735. [PubMed: 9732271]
21. Warming S, Suzuki T, Yamaguchi TP, Jenkins NA, Copeland NG. Oncogene 2004;23:2727. [PubMed: 15048087]
22. Pollard KS, et al. PLoS Genet 2006;2:e168. [PubMed: 17040131]
23. Galtier N, Duret L. Trends Genet 2007;23:273–277. [PubMed: 17418442]
24. Nie Z, et al. J Biol. Chem 2002;50:48965. [PubMed: 12388557]
25. Wassarman KM, et al. Development 1997;124:2923. [PubMed: 9247335]
26. Blanchette M, et al. Genome Res 2004;14:708. [PubMed: 15060014]
27. Siepel A, et al. Genome Res 2005;15:1034. [PubMed: 16024819]
28. Kent WJ, et al. Genome Res 2002;12:996–1006. [PubMed: 12045153]

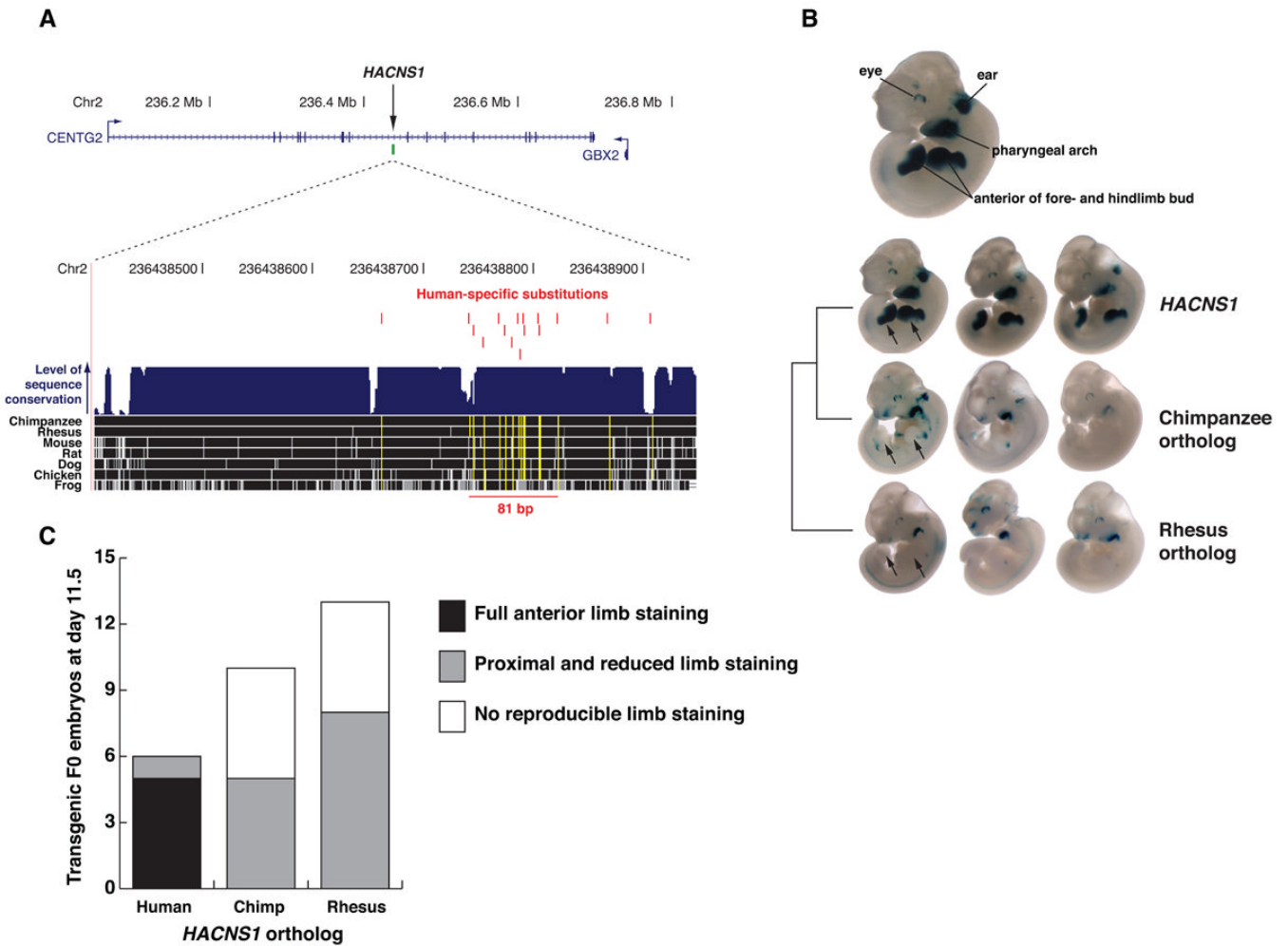


Figure 1. Human-specific gain of function in *HACNS1*

A. Top: Location of *HACNS1* in NCBI build 36.1 of the human genome assembly. **Bottom:** Sequence alignment of *HACNS1* with orthologs from other vertebrate genomes; positions identical to human are shown in black. A quantitative plot of sequence conservation is shown in blue above the alignment (26-28). The location of each human-specific substitution is indicated by a vertical red line, and the depth of nonhuman evolutionary conservation at human-substituted positions is shown by a vertical yellow line that indicates whether each sequence is identical to chimpanzee and rhesus at that position. The cluster of 13 human-specific substitutions in 81 basepairs is also indicated. **B.** Expression patterns obtained from the *HACNS1* enhancer and orthologous sequences from chimpanzee and rhesus driving expression of a *lacZ* reporter gene in E11.5 mouse embryos. Arrows indicate positions in the anterior limb bud where reproducible reporter gene expression is present or absent. A representative *HACNS1* embryo is shown at top to illustrate the relevant anatomical structures. Three embryos resulting from independent transgene integration events are shown for each orthologous sequence tested. **C.** Number of embryos transgenic for each sequence displaying the limb expression patterns described in the text.

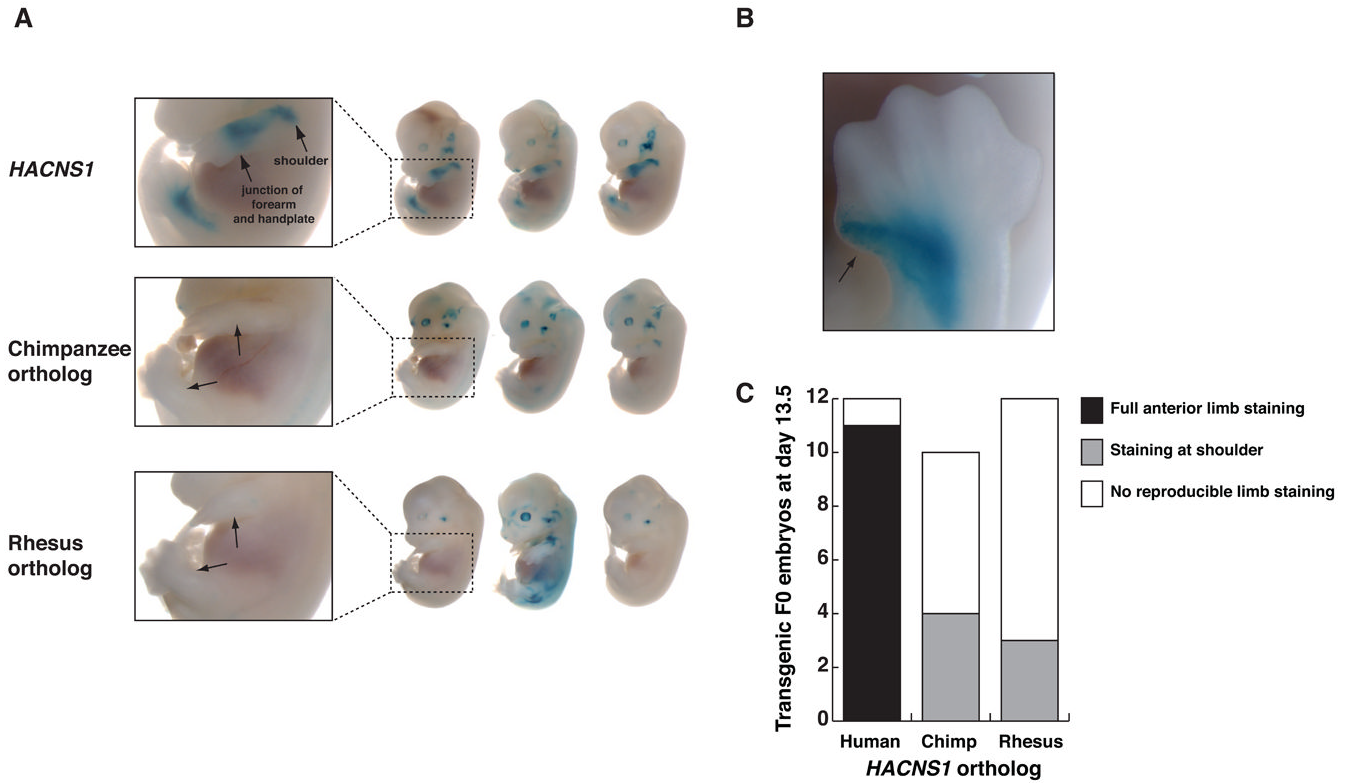


Figure 2. Gain of function in *HACNS1* persists at E13.5

A. Expression patterns obtained from *HACNS1* and its chimpanzee ortholog in E13.5 mouse embryos. Three embryos resulting from independent transgene integration events are shown for each construct. Close-up views of forelimb and hindlimb expression in a representative embryo for each construct are shown at left, and arrows indicate positions where limb expression is present or absent. **B.** Dorsal view of reporter gene expression in the distal anterior forelimb of a *HACNS1* E13.5 transgenic embryo. Arrows indicate the most anterior digit. **C.** Number of embryos transgenic for each construct that display the limb expression patterns described in the text.

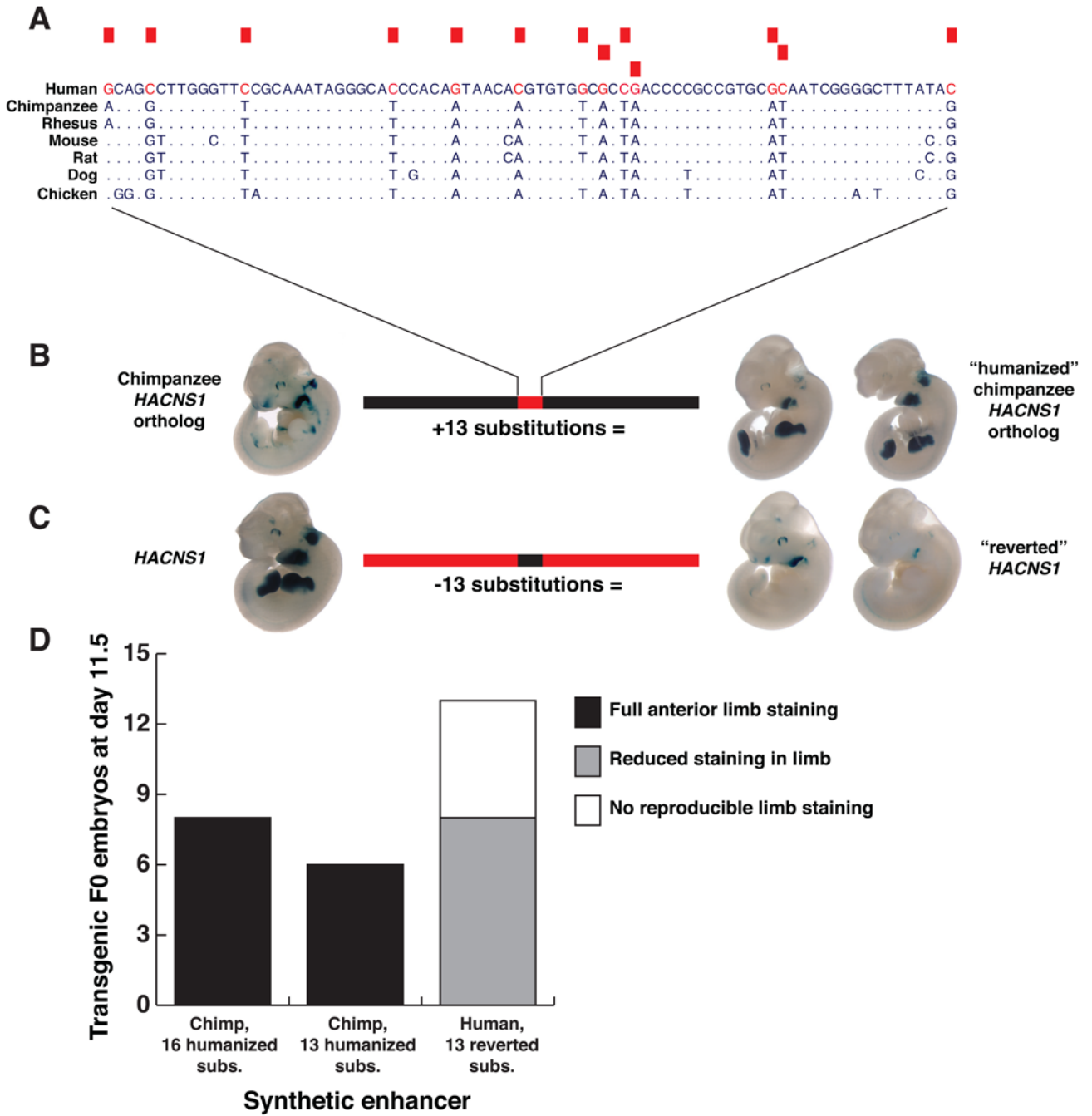


Figure 3. Identification of human-specific substitutions contributing to the gain of function in *HACNS1*

A. Alignment of *HACNS1* with orthologous sequences from other vertebrate genomes, focused on an 81-bp region in the element that contains 13 human-specific substitutions. The position of each substitution is indicated by a red box above the alignment and each human-specific nucleotide is highlighted in red. Positions in the nonhuman genomes that are identical to the human sequence are displayed as dots. **B.** Expression pattern of a synthetic enhancer in which the 13 human-specific substitutions (red box) are introduced into the orthologous 1.2 kb chimpanzee sequence background (black bar). **C.** Expression pattern of a synthetic enhancer obtained by reversion of these substitutions (black box) in the human sequence (red bar) to the

nucleotide states in chimpanzee and rhesus. **D.** Number of embryos transgenic for each synthetic enhancer that show full, partial or no expression in the limb at E11.5.

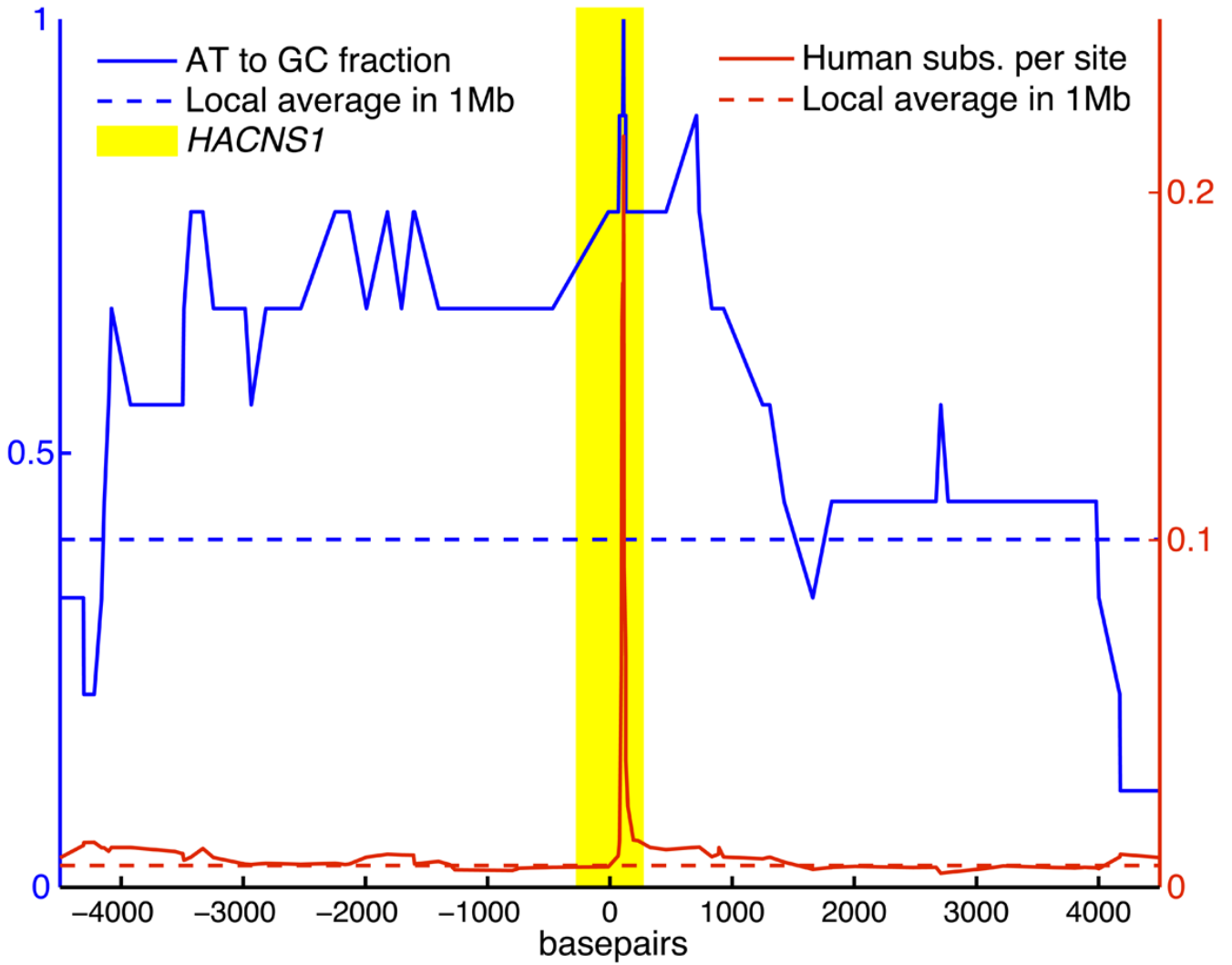


Figure 4. Human-specific substitution rate and proportion of AT to GC substitutions in *HACNS1* and flanking genomic region

The rate of human-specific substitutions (solid red line) and the fraction of human substitutions that are AT to GC events (solid blue line) were estimated in sliding windows across a 9 kb interval around *HACNS1*. The average values of each metric for 1 Mb of genomic sequence centered on *HACNS1* are shown for reference (dashed lines). The 546 bp interval corresponding to *HACNS1* is highlighted in yellow.