



Published in final edited form as:

Sci Signal. ; 2(89): pe59. doi:10.1126/scisignal.289pe59.

Human-Specific Genes May Offer a Unique Window into Human Cell Signaling

Philip D. Stahl* and Marisa J. Wainszelbaum

Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO 63110, USA.

Abstract

The identification and characterization of human-specific genes and the cellular processes that the encoded proteins control have the potential to help us understand at the molecular level what makes humans different from other species. The sequencing of the human genome and the genomes of closely related primates has revealed the presence of a small number of human- or human-lineage-specific genes that have no orthologs in lower species. Human-specific and human-lineage-specific genes are likely to function as regulators of cell signaling events, and by fine-tuning pathways, the encoded proteins may contribute to human-specific characteristics and behaviors. In addition, human-specific genes may represent biomarkers for examining human-specific characteristics of various diseases. Investigation of the gene encoding TBC1D3 is one example of a search that may lead to understanding the evolution and the function of human-specific genes, because it is absent in lower species and present in high copy number in the human genome.

Lack of knowledge about the function of human-specific genes and the proteins or RNAs that they encode represents a major gap in our understanding of signaling in human cells and, by extension, in human physiology. We posit that human-specific genes offer an opportunity to understand the evolution of signaling pathways in humans and to identify human-specific biomarkers that may open windows into human disease.

From sequencing a range of mammalian genomes, it has been posited that most human genes have orthologs in the genomes of nonprimates (1). Indeed, this assumption has served as the rationale for the use of rodent and other model organisms to simulate human cell signaling and disease. However, the human genome also contains a small number of genes that are either specific to humans or found only among the hominoids, which include the great ape lineages and the human lineages. It has been estimated that the human genome has as many as 168 human-specific genes (2).

Evidence points to segmental duplication and mutation as major mechanisms by which new genes are created (3, 4). Segmental duplication is often accompanied by increased gene copy number [that is, copy number variation (CNV)], which may influence gene transcription and function (5). Up to 12% of human euchromatic sequence is characterized by CNVs (3), which suggests that there could be scores of uncharacterized genes. The evolution of cell signaling pathways has taken place over eons, so these genes are unlikely to encode proteins that create new signaling “hubs” (proteins that interact with multiple partners) (6, 7); rather, they are more likely to modulate or regulate evolutionarily conserved pathways in a human-specific manner. Indeed, human-specific genes may include genes that account, at least in

part, for the signaling events that define or underlie the human condition—for example, the specific ways that humans think, communicate, and reproduce.

Hominoid-specific genes and the signaling pathways that they influence may serve as biomarkers to track human physiology and disease. This concept is predicated on the assumption that among the hundreds of hominoid-specific genes in our genome, some, perhaps most, will play a role in cellular processes specific to humans or their most closely related species, the chimps, gorillas, gibbons, and orangutans. The four hominoid-specific genes for which the encoded proteins are known are definitely or likely involved in cell signaling and transport. Because this is uncharted territory, human-specific biomarkers may potentially reveal a number of attributes and variations of human physiology and disease that have not been detected by medical research or practice—that is to say, we are ignorant of these variations because there is no human-specific or even hominoid-specific database that codifies them. The potential medical applications of human-specific biomarkers may catalyze the study of human-specific genes. Biomarkers (8) have multiple functions in evaluating and diagnosing disease and predicting outcomes. Human-specific biomarkers may also permit the investigation of normal human-specific processes.

Terra Incognita—Human-Specific Genes and the Human-Specific Interactome

One approach to understanding human-specific genes is to employ a top-down systems strategy by using computational methods similar to those used in yeast. A goal of this strategy would be to identify all expressed human genes and to create a human “interactome” map representing physical interactions among the encoded proteins, both those that are human-specific and those that are evolutionarily conserved (9). Hubs or proteins that interact with many partners (6, 7) are key components of interactome maps. Although it is conceivable that human-specific hubs will emerge from this analysis, it seems more likely that these human-specific proteins will serve modulatory roles in otherwise conserved processes.

Although a top-down approach provides an initial glimpse into the potential roles of the human-specific genes and their protein products, a bottom-up reductionist approach that examines one human-specific gene at a time would reveal the regulation and function of the gene and its product in detail. The bottom-up approach would provide information about the regulation of the expression of the gene and production of the encoded protein, such as in what tissues and during what stages of development or in response to what environmental cues it is expressed. Cellular localization studies would provide a context for how the protein operates in living cells. Analysis of the interacting proteins would provide clues about which signaling or transport pathways the human-specific protein affects. Finally, if the gene is implicated in a human disease, this would provide, a priori, a rationale for in-depth cellular and molecular biology studies. The top-down and bottom-up approaches are complementary and may converge in the development of a comprehensive human interactome map that would consist of all human genes and their interacting partners. Once a human-specific gene has been identified, validated, and characterized, its product can be placed within the interactome (Fig. 1, red lines).

Aside from genes that encode RNAs, such as HAR1 (10), only a small handful of the encoded proteins of human- or hominoid-specific genes have been studied. Several of these are proteins with as-yet-uncharacterized functions, such as C1ORF37-DUP, a transmembrane protein (11). The neuroblastoma breakpoint family (*NBPF1*) of genes was initially identified as a translocation breakpoint (12), and CNV in one member (*NBPF23*) is associated with neuroblastoma in humans (13). Another member of this family, the

transcript for DUF1220 is highly expressed in human brain and thought to be involved in cognitive function (14). Nuclear pore complex–interacting protein (NPIP) is encoded by a member of the rapidly evolving *Morpheus* gene family (15), and the mRNA encoding NPIP has been reported to be overexpressed in the retina of patients with macular degeneration (16). *TBC1D3* (17–22), an oncogene overexpressed in breast and prostate cancers, encodes a protein with a TBC (Tre-2, Bub2, and CDC16) domain, which alters growth factor receptor signaling.

TBC1D3: A Hominoid-Specific Gene Implicated in Growth Factor Receptor Signaling

TBC1D3 was identified two decades ago, in a study in which genomic DNA from human tumor cells was transfected into NIH3T3 cells that were injected into mice where they formed tumors (20). Human DNA was carried in these tumors as they were passaged from animal to animal, and *TBC1D3* was among the human genes retained. Pei *et al.* (22) identified PRC17 (which is another name for *TBC1D3*, based on the fact that the gene is amplified on chromosome 17 in cells derived from patients with prostate and breast cancer) and reported that it encodes a protein with a “TBC domain” that operates as a Rab5 guanosine triphosphatase–activating protein (GAP). Paulding *et al.* (21) showed that *TBC1D3* evolved ~25 million years ago by segmental duplication and that multiple copies are present in the human genome. Sequencing of human chromosome 17 further confirmed that *TBC1D3* underwent intrachromosomal duplication (23), and eight homologous, yet unique, paralogs of *TBC1D3* are transcribed from two clusters at 17q12 (19) and are present in multiple human tissues. A comparative genomics study with the chimp (*Pan troglodytes*) and human genomes (24) revealed that *TBC1D3* is present in the chimp genome as a single-copy gene. Because eight copies are found in humans (19), positive evolutionary pressure may be enhancing *TBC1D3* copy number in humans (24).

Functional studies of *TBC1D3* showed that *TBC1D3* expression substantially enhances the cellular response to epidermal growth factor (EGF) (17, 18). EGF has two well-studied effects on cells—activation of signaling pathways that regulate gene expression and activation of pathways that alter cell motility and macropinocytosis. The presence of *TBC1D3* enhanced the activity of the kinases Akt and extracellular signal–regulated kinase (ERK) (both in magnitude and duration) in response to EGF stimulation in (18). EGF receptor endocytosis and degradation were delayed by *TBC1D3*; delayed receptor trafficking correlated with enhanced signaling. Although the TBC domain of *TBC1D3* bound Rab5, the protein failed to show Rab5 GAP activity (17). The presence of *TBC1D3* enhanced the effect of EGF on macropinocytosis—an effect that required both Rab5 and Arf6, which are two proteins involved in endocytosis. In sum, *TBC1D3* interdicts and regulates a highly conserved growth factor pathway. How does this hominoid-specific gene influence the phenotypes associated with EGF signaling? Moreover, are all of the major signaling pathways in humans subject to similar regulation by newly evolved genes?

Human-specific genes represent uncharted territory with enormous potential to illuminate what makes humans human. Newly created genes are likely to have evolved by segmental duplication (4), and evidence suggests that rapidly evolving genes are often associated with cell proliferation, immunity, and inflammation (24). A new field of inquiry focusing on the role of human-specific genes in cell signaling may evolve as the functions of these recently evolved genes and the encoded proteins are revealed.

Acknowledgments

The authors thank M. Brent, X. Su, T. Brett, and H. Piwnica-Worms for helpful comments and C. Kong and P. Srikanth for invaluable discussions. We also thank D. Owyong and J. Francis for help with editing. This work was supported by NIH grant GM42259 and support from the Washington University School of Medicine.

References and Notes

1. Varki A, Geschwind DH, Eichler EE. Explaining human uniqueness: Genome interactions with environment, behaviour and culture. *Nat. Rev. Genet.* 2008; 9:749–763. [PubMed: 18802414]
2. Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, Kellis M, Lindblad-Toh K, Lander ES. Distinguishing protein-coding and noncoding genes in the human genome. *Proc. Natl. Acad. Sci. U.S.A.* 2007; 104:19428–19433. [PubMed: 18040051]
3. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad EF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. *Nature.* 2006; 444:444–454. [PubMed: 17122850]
4. Bailey JA, Eichler EE. Primate segmental duplications: Crucibles of evolution, diversity and disease. *Nat. Rev. Genet.* 2006; 7:552–564. [PubMed: 16770338]
5. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavare S, Deloukas P, Hurles ME, Dermitzakis ET. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science.* 2007; 315:848–853. [PubMed: 17289997]
6. Friedman A, Perrimon N. Genetic screening for signal transduction in the era of network biology. *Cell.* 2007; 128:225–231. [PubMed: 17254958]
7. Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJ, Cusick ME, Roth FP, Vidal M. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature.* 2004; 430:88–93. [PubMed: 15190252]
8. Gerszten RE, Wang TJ. The search for new cardiovascular biomarkers. *Nature.* 2008; 451:949–952. [PubMed: 18288185]
9. Figeys D. Mapping the human protein interactome. *Cell Res.* 2008; 18:716–724. [PubMed: 18574500]
10. Pollard KS, Salama SR, Lambert N, Lambot MA, Coppens S, Pedersen JS, Katzman S, King B, Onodera C, Siepel A, Kern AD, Dehay C, Igel H, Ares M Jr, Vanderhaeghen P, Haussler D. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature.* 2006; 443:167–172. [PubMed: 16915236]
11. Yu H, Jiang H, Zhou Q, Yang J, Cun Y, Su B, Xiao C, Wang W. Origination and evolution of a human-specific transmembrane protein gene, *c1orf37-dup*. *Hum. Mol. Genet.* 2006; 15:1870–1875. [PubMed: 16644869]
12. Vandepoele K, Van Roy N, Staes K, Speleman F, van Roy F. A novel gene family NBPF: Intricate structure generated by gene duplications during primate evolution. *Mol. Biol. Evol.* 2005; 22:2265–2274. [PubMed: 16079250]
13. Diskin SJ, Hou C, Glessner JT, Attiyeh EF, Laudenslager M, Bosse K, Cole K, Mosse YP, Wood A, Lynch JE, Pecor K, Diamond M, Winter C, Wang K, Kim C, Geiger EA, Mc-Grady PW, Blakemore AI, London WB, Shaikh TH, Bradfield J, Grant SF, Li H, Devoto M, Rappaport ER, Hakonarson H, Maris JM. Copy number variation at 1q21.1 associated with neuroblastoma. *Nature.* 2009; 459:987–991. [PubMed: 19536264]
14. Popesco MC, Maclaren EJ, Hopkins J, Dumas L, Cox M, Meltesen L, McGavran L, Wyckoff GJ, Sikela JM. Human lineage-specific amplification, selection, and neuronal expression of DUF1220 domains. *Science.* 2006; 313:1304–1307. [PubMed: 16946073]

15. Johnson ME, Viggiano L, Bailey JA, Abdul-Rauf M, Goodwin G, Rocchi M, Eichler EE. Positive selection of a gene family during the emergence of humans and African apes. *Nature*. 2001; 413:514–519. [PubMed: 11586358]
16. Hornan DM, Peirson SN, Hardcastle AJ, Molday RS, Cheetham ME, Webster AR. Novel retinal and cone photoreceptor transcripts revealed by human macular expression profiling. *Invest. Ophthalmol. Vis. Sci*. 2007; 48:5388–5396. [PubMed: 18055785]
17. Frittoli E, Palamidessi A, Pizzigoni A, Lanzetti L, Garre M, Troglio F, Troilo A, Fukuda M, Di Fiore PP, Scita G, Confalonieri S. The primate-specific protein TBC1D3 is required for optimal macropinocytosis in a novel ARF6-dependent pathway. *Mol. Biol. Cell*. 2008; 19:1304–1316. [PubMed: 18199687]
18. Wainszelbaum MJ, Charron AJ, Kong C, Kirkpatrick DS, Srikanth P, Barbieri MA, Gygi SP, Stahl PD. The hominoid-specific oncogene TBC1D3 activates Ras and modulates epidermal growth factor receptor signaling and trafficking. *J. Biol. Chem*. 2008; 283:13233–13242. [PubMed: 18319245]
19. Hodzic D, Kong C, Wainszelbaum MJ, Charron AJ, Su X, Stahl PD. TBC1D3, a hominoid oncoprotein, is encoded by a cluster of paralogues located on chromosome 17q12. *Genomics*. 2006; 88:731–736. [PubMed: 16863688]
20. Huebner K, Cannizzaro LA, Nakamura T, Hillova J, Mariage-Samson R, Hecht F, Hill M, Croce CM. A rearranged transforming gene, *tre*, is made up of human sequences derived from chromosome regions 5q, 17q and 18q. *Oncogene*. 1988; 3:449–455. [PubMed: 3274085]
21. Paulding CA, Ruvolo M, Haber DA. The *Tre2* (*USP6*) oncogene is a hominoid-specific gene. *Proc. Natl. Acad. Sci. U.S.A.* 2003; 100:2507–2511. [PubMed: 12604796]
22. Pei L, Peng Y, Yang Y, Ling XB, Van Eindhoven WG, Nguyen KC, Rubin M, Hoey T, Powers S, Li J. *PRC17*, a novel oncogene encoding a Rab GTPase-activating protein, is amplified in prostate cancer. *Cancer Res*. 2002; 62:5420–5424. [PubMed: 12359748]
23. Zody MC, Garber M, Adams DJ, Sharpe T, Harrow J, Lupski JR, Nicholson C, Searle SM, Wilming L, Young SK, Abouelleil A, Allen NR, Bi W, Bloom T, Borowsky ML, Bugalter BE, Butler J, Chang JL, Chen CK, Cook A, Corum B, Cuomo CA, de Jong PJ, DeCaprio D, Dewar K, FitzGerald M, Gilbert J, Gibson R, Gnerre S, Goldstein S, Grafham DV, Grocock R, Hafez N, Hagopian DS, Hart E, Norman CH, Humphray S, Jaffe DB, Jones M, Kamal M, Khodiyar VK, LaButti K, Laird G, Lehoczyk J, Liu X, Lokyitsang T, Loveland J, Lui A, Macdonald P, Major JE, Matthews L, Mauceli E, McCarroll SA, Mihalev AH, Mudge J, Nguyen C, Nicol R, O’Leary SB, Osoegawa K, Schwartz DC, Shaw-Smith C, Stankiewicz P, Steward C, Swarbreck D, Venkataraman V, Whittaker CA, Yang X, Zimmer AR, Bradley A, Hubbard T, Birren BW, Rogers J, Lander ES, Nusbaum C. DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. *Nature*. 2006; 440:1045–1049. [PubMed: 16625196]
24. Perry GH, Yang F, Marques-Bonet T, Murphy C, Fitzgerald T, Lee AS, Hyland C, Stone AC, Hurles ME, Tyler-Smith C, Eichler EE, Carter NP, Lee C, Redon R. Copy number variation and evolution in humans and chimpanzees. *Genome Res*. 2008; 18:1698–1710. [PubMed: 18775914]

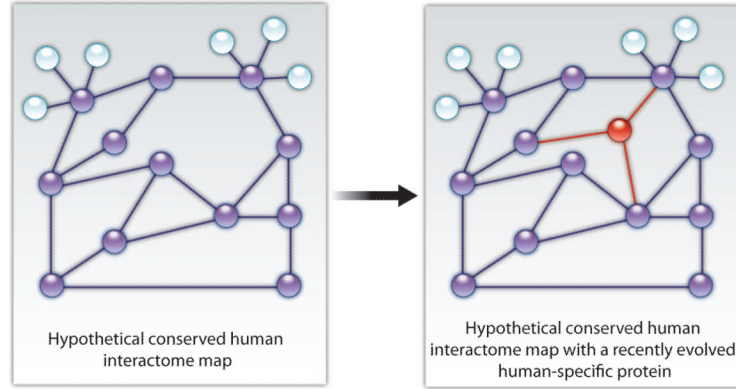


Fig. 1. Hypothetical role of a human-specific gene in signaling. Human-specific genes may complement the human interactome map by producing proteins with interactions not present in other species that alter signaling responsiveness through tighter regulation. **(Left)** A hypothetical human interactome map of evolutionarily conserved hubs (purple) represents how a complex human network might operate with multiple protein-protein interactions. Inputs such as growth factor receptors are shown in light blue. **(Right)** An interactome superimposed with a recently evolved human-specific gene (red) might exploit novel interactions that enhance or attenuate signal quality and amplitude.