Distinct genomic signatures of adaptation in pre- and postnatal environments during human evolution

Monica Uddin[†], Morris Goodman[†‡§], Offer Erez[¶], Roberto Romero[¶], Guozhen Liu[†], Munirul Islam[‖], Juan C. Opazo[†], Chet C. Sherwood[††], Lawrence I. Grossman[†], and Derek E. Wildman[†§¶‡‡]

[†]Center for Molecular Medicine and Genetics and [‡]Department of Anatomy and Cell Biology, Wayne State University School of Medicine, Detroit, MI 48201; [¶]Perinatology Research Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892; [‖]Department of Computer Science, Wayne State University, Detroit, MI 48201; [††]Department of Anthropology, The George Washington University, Washington, DC 20052; and [‡‡]Department of Obstetrics and Gynecology, Wayne State University School of Medicine and Hutzel Women's Hospital, Detroit, MI 48201

Contributed by Morris Goodman, December 31, 2007 (sent for review December 3, 2007)

The human genome evolution project seeks to reveal the genetic underpinnings of key phenotypic features that are distinctive of humans, such as a greatly enlarged cerebral cortex, slow development, and long life spans. This project has focused predominantly on genotypic changes during the 6-million-year descent from the last common ancestor (LCA) of humans and chimpanzees. Here, we argue that adaptive genotypic changes during earlier periods of evolutionary history also helped shape the distinctive human phenotype. Using comparative genome sequence data from 10 vertebrate species, we find a signature of human ancestry-specific adaptive evolution in 1,240 genes during their descent from the LCA with rodents. We also find that the signature of adaptive evolution is significantly different for highly expressed genes in human fetal and adult-stage tissues. Functional annotation clustering shows that on the ape stem lineage, an especially evident adaptively evolved biological pathway contains genes that function in mitochondria, are crucially involved in aerobic energy production, and are highly expressed in two energy-demanding tissues, heart and brain. Also, on this ape stem lineage, there was adaptive evolution among genes associated with human autoimmune and aging-related diseases. During more recent human descent, the adaptively evolving, highly expressed genes in fetal brain are involved in mediating neuronal connectivity. Comparing adaptively evolving genes from pre- and postnatal-stage tissues suggests that different selective pressures act on the development vs. the maintenance of the human phenotype.

fetal | human disease | mitochondria | placenta | thyroid

Even before the sequencing of the human genome (1, 2) it was realized that a human genome evolution project was needed to gain a better understanding of the genetic underpinnings of the distinctive human phenotype (3). With the current availability of human, chimpanzee, rhesus monkey, and other vertebrate genome sequences, the proposed human genome evolution project is now well underway. So far in pursuing this project, most efforts have focused on identifying adaptive genotypic changes that occurred in relatively late human ancestors (4–6). However, some of humankind's most striking features arose from changes that occurred more anciently in our evolutionary heritage (7). For example, bipedalism would not have been effective without the development of upright posture in the common ancestor of all living apes with the advent of the mobile shoulder joint (8). Similarly, it has been suggested that the relatively extended intrauterine fetal life characteristic of humans (9) benefited from the development of a villous type of maternal–fetal placental interdigitation in an ancestor of Primates (10).

An additional important dimension to consider is the environment in which selection occurs. Placental mammals exist in two primary environments during their lifetimes that may select for different adaptive changes: the prenatal and postnatal. Thus the prenatal environment, with its maternal fostering of fetal development, could favor a different course of adaptive evolution than postnatal environments, when individuals face a much broader range of external challenges. Furthermore, to shape complex morphologies or physiologies, the targets of natural selection may have been not just individual genes but, rather, networks of genes comprising specific biological pathways.

In light of these considerations, the present study was directed at identifying the positively selected genetic changes that may have helped to shape key, distinctive (and in some cases unique) phenotypic features of humans. We first identified genes that have been the targets of positive selection during human evolutionary history, broadly defined to encompass the lineage that descended from the last common ancestor (LCA) of primates and rodents to present day humans (Fig. 1*A*). After exclusion of the genes that were also targets of positive selection in rodent descent, the remaining genes were referred to as human ancestry-specific (HAS). We next identified highly expressed HAS genes from human tissues representing pre- and postnatal stages of development (11): fetal brain, adult brain, fetal liver, adult liver, fetal lung, adult lung, fetal thyroid, adult thyroid, placenta, adult uterus, and adult heart (Fig. 1*B*). For the purpose of this study, postnatal is defined as the age of the individual tissue samples, which were exclusively adult.

For each HAS gene, we identified period(s) exhibiting the signature of positive selection. Also, we determined whether the chosen tissues show functionally distinct sets of highly expressed HAS genes. Functionally distinct sets are here defined as those genes with annotations (*i*) that are overrepresented when compared with the annotations assigned to all of the genes in the human genome and (*ii*) that are functionally clustered together into particular biological pathways by using the DAVID algorithm (12). We find that HAS genes group into functionally distinct sets of overrepresented annotation clusters. Among the functional annotation clusters (FACs), those consisting of highly expressed HAS genes in specific tissues provide insights into the adaptive molecular changes underlying the morphologies or physiologies of these tissues.

Results and Discussion

HAS Genes in Recent vs. Ancient Periods of Human Ancestry. A pattern of HAS adaptive evolution was identified for 1,692 RefSeq mRNA IDs for which 1,240 unique genes could be assigned. This number, although larger than some previous reports of other whole-genome

Author contributions: M.U., M.G., and D.E.W. designed research; M.U., M.I., J.C.O., and D.E.W. performed research; G.L. contributed new reagents/analytic tools; M.U., M.G., O.E., R.R., J.C.O., C.C.S., L.I.G., and D.E.W. analyzed data; and M.U., M.G., O.E., C.C.S., L.I.G., and D.E.W. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[§]To whom correspondence may be addressed. E-mail: mgoodwayne@aol.com or dwildman@med.wayne.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0712400105/DC1.

© 2008 by The National Academy of Sciences of the USA

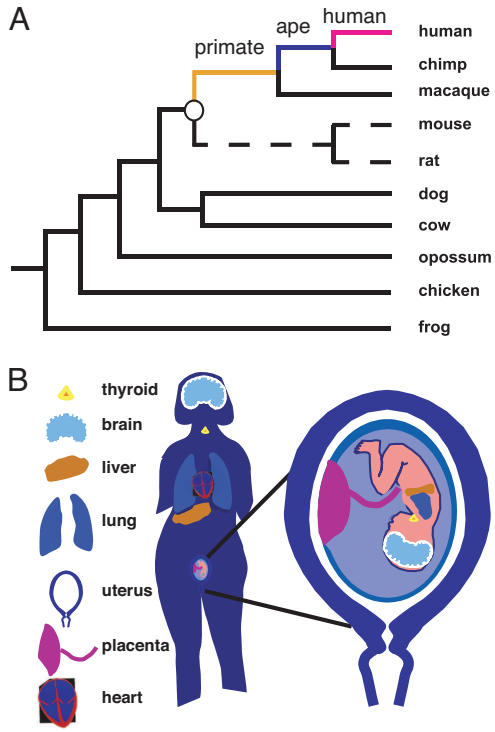

Fig. 1. Experimental design. (*A*) Phylogenetic tree showing the taxa sampled in the present study. Open circle indicates the LCA of rodents and primates. HAS lineages are indicated in color: yellow, primate stem lineage; blue, ape stem lineage; magenta, human terminal lineage. Dashed lines indicate rodent lineages. A gene was retained on the HAS list if it showed dN/dS >1 on any HAS lineage(s) and dN/dS <1 on all rodent-specific lineages. (*B*) Tissues analyzed in the present study. (*Left*) Postnatal tissues. (*Right*) Prenatal tissues.

scans of positive selection (4, 5, 13–17), reflects our broader definition of human evolutionary history, defined as the descent of humans since the time of their LCA with rodents. FACs determined from the full HAS list identified glycoprotein, glycosylation site: N-linked (GlcNAc. . .) and disulfide bond as the annotations in the most enriched cluster (enrichment score = 26.32, Table 1). This finding demonstrates that adaptive changes in the posttranslational modification of membrane and secreted proteins are among the major evolutionary transformations during human ancestry.

FACs were also identified for the 585 HAS genes on the human terminal, the 546 HAS genes on the ape stem, and the 188 HAS genes on the primate stem. Eighty-two HAS genes also showed a signature of selection on the chimpanzee terminal, but only 25 showed their signature of selection on both the human and chimpanzee terminals and on no other, earlier period of shared human and chimpanzee ancestry. A complete listing of FACs for all HAS genes, and for each branch of human ancestry, is provided in supporting information (SI) Tables 3–6.

The top FACs comprise distinctive sets of annotations and indicate that different biological pathways became the targets of positive selection during successive periods of humankind's evolutionary history. For example, aerobic energy production pathways were particularly strong targets of selection on the ape stem, as evidenced by that period's annotations of mitochondrion and oxidative phosphorylation for HAS genes. Similarly, immune system functions appear to have been strong targets of selection on the primate stem, as evidenced by annotations of cytokine activity and immune response for HAS genes. Although some annotations do appear more than once during different periods of human evolutionary history (e.g., transmembrane, glycoprotein) only a small number of genes underlying these annotations are shared between the different periods (SI Table 7). This suggests that the adaptive changes that occurred during earlier periods set the stage for selection of different sets of adaptive changes during the later evolutionary periods, even within the same biological pathway. Additional analyses in which all HAS genes are clustered according to only disease-related annotations support the finding that earlier periods of human evolutionary history also play a role in "maladaptive" human phenotypes, particularly with respect to autoimmune and longevity-related disease (*SI Results*).

Highly Expressed HAS Genes. The top FACs for all highly expressed genes from each of the 11 tissues, and the subset of those that are also adaptively evolving, are presented in Table 2. Among the 1,240 HAS genes, 287 appeared at least once on the lists of genes highly expressed in the 11 studied tissues. In all four adult-stage tissues with a fetal counterpart, and in the uterus/placenta comparison, the annotations comprising the top clusters of highly expressed HAS genes are overrepresented to a greater extent (as indicated by enrichment scores) than in the prenatal-stage tissue. Furthermore, when considered in aggregate, enrichment scores from the top three clusters of highly expressed HAS genes from all five adult tissues are significantly different from the enrichment scores determined for the top three clusters in the five predominantly fetal tissues (Mann–Whitney U test = 45; P = 0.006, two-tailed test). When heart is added to the adult-stage tissues, the significant difference from fetal tissues remains (Mann–Whitney U test = 52; P = 0.003). These greater adult enrichment scores are in contrast to the more variable pattern of enrichment scores obtained for all highly expressed genes for which certain tissues (e.g., thyroid, lung) show higher enrichment in the fetal stage and others (e.g., whole

Table 1. Functional annotation clustering of HAS genes in total and by period of descent

Period of descent	Total no. of genes	Top three functional annotation clusters			
		Cluster	Enrichment score	No. of annotations	No. of genes
Total HAS list	1,240	Glycoprotein	26.32	3	289
		Olfaction	18.68	21	273
		Transmembrane	18.33	6	346
Human terminal lineage	585	Olfaction	19.07	25	207
		Transmembrane, glycoprotein	13.62	10	201
		Interferon regulatory factor	1.77	3	3
Ape stem lineage	546	Transmembrane	6.92	6	143
		Mitochondrion	4.9	4	32
		Oxidative phosphorylation	4.11	8	34
Primate stem lineage	188	Signal	7.31	3	50
		Cytokine activity	4.36	6	16
		Immune response	3.31	10	38

Table 2. Functional annotation cluster analyses of highly expressed and highly expressed HAS genes by tissue

Tissue	Total no. of 3X genes	Top three clusters determined from highly expressed 3X genes	Enrichment score	Total no. of 3X∩HAS genes	Top three clusters determined from HAS highly expressed 3X∩HAS genes	Enrichment score
Uterus	1,385	Extracellular matrix	33.36	53	Transmembrane, glycoprotein	3.3
		Signal, glycoprotein	28.11		Cellular morphogenesis	2.35
		Negative regulation of cellular processes	18.35		Cell motility	1.64
Placenta	1,675	Signal, glycoprotein	49.74	70	Hormone activity	2.38
		Extracellular matrix	15.76		Plasma membrane	1.52
		Transmembrane	15.04		Blood coagulation	1.48
Whole brain	1,836	Nervous system development	33.15	63	Mitochondrion	5.04
		Cadherin-like	17.81		Electron carrier activity	3.57
		Membrane	13.07		Oxidative phosphorylation	2.48
Fetal brain	1,469	Cadherin-like	23.7	35	Cadherin, cell-adhesion	1.65
		Repeat: PXXP	13.74		Membrane, glycoprotein	0.98
		Microtubule	8.55		Plasma membrane	0.74
Liver	1,029	Signal, glycoprotein	27.43	67	Microsome	3.93
		Amino acid and derivative metabolism	24.37		Immune response	2.75
		Microsome	21.7		Signal, glycoprotein	2.52
Fetal liver	742	Signal	16.18	47	Signal	3.58
		Blood coagulation	12.18		Localization	2.24
		Protease inhibitor activity	11.57		Plasma membrane	2
Lung	908	Signal, glycoprotein	28.99	48	Signal	7.4
		Response to pest, pathogen, or parasite	13.07		Response to pest, pathogen, or parasite	3.21
		Vacuole	9.88		Transmembrane protein	2.75
Fetal lung	1,126	Signal, glycoprotein	66.37	68	Response to pest, pathogen, or parasite	2.92
		Response to pest, pathogen, or parasite	23.09		Surface antigen	2.73
		Extracellular matrix	20.38		Blood coagulation	2.31
Thyroid	2,430	Actin-binding	13.22	110	Membrane	2.85
		Signal, glycoprotein	12.48		Oxidoreductase	1.62
		Golgi stack	9.48		Plasma membrane	1.35
Fetal thyroid	646	Extracellular matrix	19	22	Response to external stimulus	0.62
		Signal, glycoprotein	15.29		Nonmembrane bound organelle	0.34
		Contractile fiber	10.58		Transport	0.28
Heart	896	Mitochondrion	31.45	52	Hydrogen ion transporter activity	6
		Hydrogen ion transporter activity	13.94		NADH dehydrogenase activity	4.54
		NADH dehydrogenase activity	12.73		Oxidative phosphorylation	1.82

Unless otherwise indicated, the listed tissues were obtained from adult individuals.

brain, liver) show higher enrichment in the adult stage. Notably, enrichment scores determined for the top three clusters of all highly expressed genes in the adult vs. fetal tissues are not significantly different (Mann–Whitney U test = 100; $P = 0.619$, two-tailed test).

Tissue Specificity in Adaptively Evolving Biological Pathways. Examples of HAS-determined tissue specificity include the 5 genes involved in hormone activity in the placenta ($P < 0.001$), 6 genes involved in cellular morphogenesis in the uterus ($P < 0.001$), 6 genes involved in cell adhesion in the fetal lung ($P < 0.01$), and 9 genes involved in oxidoreductase activity in the thyroid ($P < 0.05$). The top FACs for the list of all highly expressed genes (not just HAS genes) generally identified quite different biological pathways. A full listing of FACs by tissue for all highly expressed genes, and the subset of those that are also adaptively evolving, is provided in SI Tables 8–29. These results suggest that the FACs composed of HAS genes are not just random subsets of the FACs composed of all highly expressed genes; rather, they represent tissue-specific biological pathways that have become the targets of selection during human descent.

Pre- vs. Postnatal Environments. It has been suggested that genes involved in determining species-specific morphologies are more likely to be expressed during developmental rather than adult stages (18) and that the species-specific morphologies are more evident later rather than earlier during development (19, 20). Although results presented here do not include a developmental time series of expression data, the pattern of fetal vs. adult-expressed HAS genes may be interpreted to conform to this general pattern in which the positively selected genes that shape species-specific morphologies are more likely to be highly expressed in fetal rather than in adult tissues: Only the prenatal and gestation-related tissues show 50% or more of their adaptive signatures on the human terminal lineage (Fig. 2). Indeed, among the 127 genes that are highly expressed in fetal tissues, only two show a signature of selection on both the

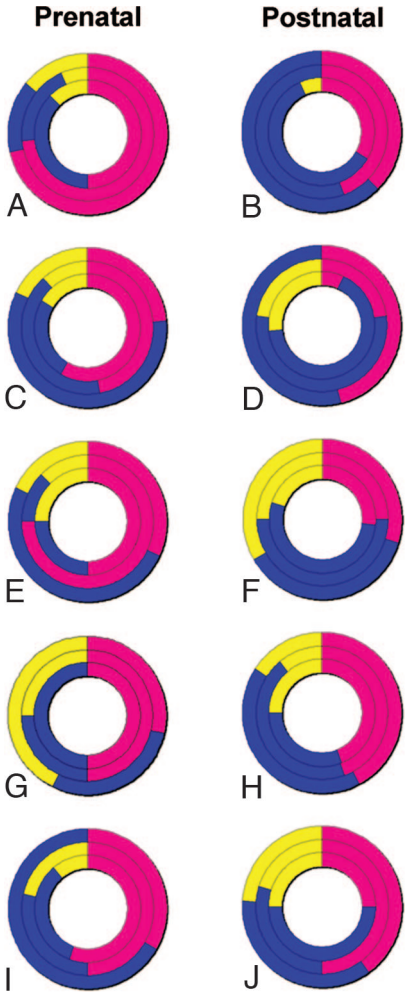

Fig. 2. Proportion of adaptive signatures on each HAS branch, by tissue and cluster. The top three FACs are represented, with cluster one appearing as the outermost ring. Yellow, proportion of adaptive signatures occurring on the primate stem lineage; blue, proportion of adaptive signatures occurring on the ape stem lineage; magenta, proportion of adaptive signatures occurring on the human terminal lineage. (*A*) Fetal brain. (*B*) Adult whole brain. (*C*) Fetal liver. (*D*) Adult liver. (*E*) Fetal lung. (*F*) Adult lung. (*G*) Fetal thyroid. (*H*) Adult thyroid. (*I*) Placenta. (*J*) Adult uterus. Only among fetal tissues (*A*, *C*, *E*, *G*, *I*) and the adult uterus do 50% or more adaptive events occur on the human terminal in at least one of the top three FACs.

human and chimpanzee terminals (and on no other, earlier period of human ancestry): smoothelin and v-crk sarcoma virus ct10 oncogene homolog (avian). Furthermore, only one of these—smoothelin—appears in one of the top three clusters identified among fetally expressed HAS genes (fetal lung). This observation is particularly noteworthy in that, in terms of mean dN calculated from the three human ancestry periods for each tissue with fetal and adult stages, the fetal tissues show lower values in all cases but the lung, which had similar mean dN values in its adult and fetal stages (0.058 vs. 0.059; other data not shown). Despite this overall lower level of nonsynonymous changes, HAS genes highly expressed in fetal tissues show a more pronounced pattern of human terminal-specific change than do their adult counterparts. When considering the total number of adaptive signatures that occur among the top three FACs by HAS branch, this difference in fetal vs. adult tissues is statistically significant ($\chi^2 = 7.52$, df = 2, $P = 0.02$). As apparent from examples discussed below, the top FACs for adult tissues are usually implicated in physiological pathways, whereas the top FACs for fetal tissues can be implicated in developmental pathways. This interpretation of our results is in line with the proposal that investigations into the evolution of physiological and morphological characters are best treated separately (18).

Adaptive Evolution in Brain-Expressed HAS Genes. Results obtained from highly expressed HAS genes in fetal brain suggest that in human descent, there have been important modifications in the development of neuronal connectivity. The most enriched cluster determined from fetal brain tissue includes annotations of cadherin, cell adhesion, and calcium ion binding (SI Table 15) and comprises seven genes, five of which show the signature of positive selection on the human terminal lineage. The importance of protocadherins, represented in this cluster by *PCDHGA8* and *PCDHGB1*, and other cell adhesion molecules to the formation of neuronal networks during development is well established (21). In addition, the involvement of voltage-sensitive calcium channels, such as that represented by *CACNA1G*, may be involved in neuronal firing patterns important to information processing and cell growth processes (22). Furthermore, *KIAA0319*, a gene involved in neuronal migration and axon guidance in the neocortex, has been linked to developmental dyslexia (23). Given the species differences between humans and other apes in the details of neocortical wiring suggested by magnetic resonance imaging and histological studies (24–26), these FAC results provide congruent evidence that unique aspects of human brain wiring have been molded by selection to generate specific patterns of connectivity.

In contrast, the most enriched cluster determined from adult whole-brain tissue includes annotations of mitochondrion, oxidative phosphorylation, and ion transporter activity and comprises 16 genes (SI Table 13), 10 of which show the signature of positive selection on the ape stem lineage. Given the documented accelerated evolution in electron transport chain genes in anthropoid primates (27) and the relatively higher expression of ETC-related genes in both chimpanzee and human brains (28), these results suggest that energy metabolism pathways supporting adult neuronal function have been modified during portions of the ancestral lineage shared with chimpanzees. In support of this, the proportion of metabolic-supporting glial cells relative to neurons in the neocortex increases in anthropoid primates as overall brain size enlarges, with humans and other apes displaying the highest glia–neuron ratios (29).

Because the heart exhibits a mass-specific metabolic rate that is more than twice that of the brain (30), we also examined HAS gene expression in adult heart tissue. Although aerobic energy metabolism-related clusters also appear in these analyses, the FACs so obtained appear to be largely subsets of the most enriched FACs obtained when considering all highly expressed genes in that tissue (Table 2). Furthermore, whereas brain and body sizes vary widely among primates (31), the heart is an extremely conserved organ whose four-chambered anatomical structure is present in all mammals and whose function is essential to life. Nevertheless, the heart could have benefited from adaptive changes in this pathway if, for example, the adaptive evolution favored more efficient energy production accompanied by less free radical production, which would, in turn, favor longer life spans.

Adaptive Evolution in Thyroid-Expressed HAS Genes. The thyroid produces hormones that are known to play a major role in regulating metabolic rate and the growth rate and function of many bodily systems. In particular, thyroid hormones play a significant role in early brain development (32–35). Despite low levels of triiodothyronine (T_3) in the fetus, from weeks 13 to 20—when the mother is the only source of thyroxine (from which T_3 is converted)—levels of this hormone increase in fetal cerebral cortical tissue to levels higher than in adults (34). The provision of this important hormone by the mother to the fetus may thus

help account for the much greater number of HAS genes identified in adult vs. fetal thyroid, providing a particularly striking example of the differing signatures of selection depending on the developmental stage of the tissue. Consistent with this finding, the topmost cluster from adult thyroid predominates in membrane- and transmembrane-related annotations, suggesting that the genes so annotated could play a role in trafficking thyroid hormones and their related products.

Although genes directly related to thyroid hormone generation did not meet the criterion of our HAS list, oxioreductase and other aerobic energy metabolism annotations comprised the second-most-enriched FAC of HAS genes in the adult thyroid tissue. This finding is consistent with the view (36–38) that thyroid hormones are regulators of mitochondrial function, because they are known to increase oxygen consumption (39, 40) and stimulate cytochrome *c* oxidase activity (37). In addition, a direct mitochondrial T3 pathway has been implicated in cellular differentiation (41). This dual involvement in energy production and development has led to the proposal that thyroid hormone's regulation of mitochondrial activity may represent a major link between metabolism and development (38), a proposal consistent with the findings presented here.

Adaptive Evolution in Placentally Expressed HAS Genes. Although the placenta shows >50% of its adaptive signatures on the human terminal lineage in two of its top three FACs, results from the topmost FAC indicate that hormonal pathways that function in pregnancy have been particularly strong targets of adaptive evolution during the 19 or so million years before the LCA of humans and chimpanzees. Five of the six genes in this topmost cluster have annotations of hormone and/or hormone activity. Four of these genes show adaptive evolution on the ape stem lineage: *MUC1*, *CCK*, *GH2*, and *CGA*. *MUC1* is expressed in both villous and extravillous trophoblasts (42, 43) and has been implicated in implantation (44). This gene is also highly expressed in the amnion (45). Humans have interstitial implantation compared with the superficial form with fewer invasive trophoblasts seen in Old World monkeys (46). *CCK* has been shown to regulate leptin levels (47), and leptin expression levels have been shown to change in the obstetrical syndromes of preeclampsia and intrauterine growth retardation (48). *GH2* also increases its expression in patients with preeclampsia (49). Interestingly, preeclamptic-like symptoms have not been described in species other than apes.

CGA encodes the α polypeptide in heterodimeric glycoprotein hormones including chorionic gonadotropin (CG), luteinizing hormone, follicle stimulating hormone, and thyroid stimulating hormone. Although these hormones share an identical α polypeptide, their β polypeptides are unique and confer biological specificity. CG is produced by placental trophoblast cells and stimulates production of ovarian steroids essential for pregnancy maintenance. Additional members of the CGβ gene family, predicted to function in implantation and placental development, have recently been identified as originating in the common ancestor of humans and African great apes (50). Taken together, the identification of HAS genes involved in hormonal pathways active during gestation supports the proposal that the evolution of the endocrine system in recent and more distant periods of human evolutionary history is related to adaptations in pregnancy and development.

Concluding Remarks. This study presents a framework for understanding the genetic underpinnings of humankind. FACs of HAS genes by period of descent confirm the deep genetic roots of human distinctiveness and emphasize the importance of considering how the adaptive changes in earlier evolutionary periods set the stage for adaptive changes in later evolutionary periods. In addition, analyses of highly expressed HAS genes from different developmental stages demonstrate that the prenatal environment selects for different HAS genes than the postnatal environment. Indeed, the prenatal environment appears to have favored the adaptive evolution of genes involved in human-specific developmental and morphological changes during the last 6 million years, whereas the postnatal environment has favored the adaptive evolution of genes involved in key physiological processes, the adaptations having been established during earlier periods of human ancestry. A potential caveat when considering our results is that, particularly among fetal tissues, expression data were obtained from samples that included heterogeneous mixtures of ages, gestational stages, and sex, and may thus represent gene expression patterns from a variety of developmental stages. However, that we were able to detect strong signals of tissue-specific adaptations despite the heterogeneous nature of the sampled tissues' origins speaks to the advantage of focusing on functional analyses of only highly expressed HAS genes.

Many clues regarding additional important components in the development of human phenotypic characters are hinted at by our results. For example, the identification of glycoprotein and glycosylation sites among the annotations in the topmost FACs in both the total HAS and human terminal lists suggests a genetic basis for changes in carbohydrate biochemistry and is consistent with the well documented human- and other species-specific differences in glycans and glycan-binding proteins (51, 52). Furthermore, in our tissue-based lists of positively selected genes, the annotations glycoprotein, glycosaminoglycan binding, carbohydrate binding, and glycosylation site appear in the FACs of 10 of the 11 studied tissues.

An additional example comes from annotations comprising genes of the protocadherin (*PCDH*) gene family. The identification of two *PCDH* genes in the most enriched cluster obtained from highly expressed HAS genes in fetal brain tissue is noteworthy in light of recent reports that there are primate-specific *cis*-antisense transcripts for *PCDH* genes that are expressed in fetal and adult human brain (53). Together, these findings point to a key biological pathway that has undergone both protein coding and regulatory evolution in human ancestry. Our study thus offers insight into the links between genotype and phenotype and underscores the importance of protein-encoding loci in the process of phenotypic change.

Materials and Methods

Determination of Positive Selection During Human Ancestry. Genes positively selected during human ancestry were determined with the assistance of OCPAT (54). Subsequent analyses were conducted using codeml (55), under the free-ratio model. The use of this model alone to identify positively selected lineages may, in some cases, lead to inaccurate results. However, testing the model to several other potential alternative models for every gene is not computationally feasible and could result in overfitting the data. Genes with <50 codons were eliminated from further analyses. The remaining list of 23,945 RefSeq IDs, and their accompanying summary data, is available in SI Table 30. Analyses were conducted with data current as of November 2006. HAS genes were identified if dN/dS > 1 on one or more HAS branches depicted in Fig. 1*A*. A value of dN/dS >1 is very rare across the entire length of aligned protein-coding sequences (56); therefore, values >1 provide robust support for positive selection. RefSeq IDs where S*dS = 0 on one or more HAS branch were retained if they met the criterion of dN > mean dS on the dN/dS-undefined branch. HAS candidate genes showing a similar pattern on one or more branches of rodent descent were eliminated from further consideration.

Identification of Highly Expressed Genes. Genes expressed at three times or greater than the median of the 79 analyzed tissues were obtained from Human GNF1H, gcRMA gene expression data indexed by SymAtlas (11) for: placenta, adult uterus (uterus or uterus corpus), whole adult brain, fetal brain, adult liver, fetal liver, adult thyroid, fetal thyroid, adult lung, fetal lung, and adult heart. Each tissue-based list was intersected with the HAS list to identify the set of RefSeq IDs that was both highly expressed and positively selected in each of the 11 tissues.

Identification of Functional Annotation Clusters. To identify major biological themes underlying this large assortment of HAS genes, functional analyses were conducted by using DAVID (Database for Annotation, Visualization and Integrated Discovery) (12). DAVID is a publicly available, National Institute of Allergy

and Infectious Diseases-administered bioinformatics resource that provides a comprehensive set of functional annotation tools designed to assist investigators in understanding the biological meaning behind large sets of genes. Several DAVID tools identify enriched or overrepresented biological annotations within a given set of genes, usually determined in relation to the proportion of similar annotations for an entire genome. For the purpose of this study, results were obtained with the FAC tool, which clusters similar annotations based on the cooccurrence of particular gene sets. The tool also calculates an associated enrichment score for each cluster based on the geometric mean of the *P* values determined for each of its component annotations. The κ similarity threshold was set to 0.7 to identify annotation sets that shared strong agreement among their underlying genes. Other options were set to their default values.

FAC results of the HAS list, and the subset of HAS genes that were highly expressed in each tissue, were obtained in April and May 2007. All text, tables, and figures report results in terms of DAVID genes. *P* values from DAVID analyses referred to in the text are reported in terms of the modified Fisher-exact *P* value implemented by the DAVID tool (12). Cluster titles indicated in Tables 1 and 2 were obtained by selecting the overrepresented annotation that conveyed the broadest biological meaning from among the top three overrepresented annotations found in the cluster.

ACKNOWLEDGMENTS. We thank Dr. Michael Kruger for statistical advice and Dr. Jamal Uddin for useful discussions. Pat Schoff provided assistance with artwork. This work was supported in part by the Intramural Research division of the National Institute of Child Health and Human Development, National Institutes of Health (NIH), Department of Health and Human Services; National Science Foundation (NSF) Grant BCS-0550209; and the Thomas E. Starzl Transplantation Institute at the University of Pittsburgh School of Medicine. Additional support was provided by NSF Grants BCS-0515484 and BCS-0549117, NIH Grant NS-42867, and James S. McDonnell Foundation Grant 22002078.

1. Lander ES, *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921.
2. Venter JC, *et al.* (2001) The sequence of the human genome. *Science* 291:1304–1351.
3. McConkey EH, Goodman M (1997) A Human Genome Evolution Project is needed. *Trends Genet* 13:350–351.
4. Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437:69–87.
5. Bakewell MA, Shi P, Zhang J (2007) More genes underwent positive selection in chimpanzee evolution than in human evolution. *Proc Natl Acad Sci USA* 104:7489–7494.
6. Clark AG, *et al.* (2003) Inferring nonneutral evolution from human–chimp–mouse orthologous gene trios. *Science* 302:1960–1963.
7. Goodman M (1999) The genomic record of humankind's evolutionary roots. *Am J Hum Genet* 64:31–39.
8. Thorpe SK, Holder RL, Crompton RH (2007) Origin of human bipedalism as an adaptation for locomotion on flexible branches. *Science* 316:1328–1331.
9. Martin RD (2003) Human reproduction: A comparative background for medical hypotheses. *J Reprod Immunol* 59:111–135.
10. Wildman DE, *et al.* (2006) Evolution of the mammalian placenta revealed by phylogenetic analysis. *Proc Natl Acad Sci USA* 103:3203–3208.
11. Su AI, *et al.* (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci USA* 101:6062–6067.
12. Dennis G, Jr, *et al.* (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4:P3.
13. Arbiza L, Dopazo J, Dopazo H (2006) Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome. *PLoS Comput Biol* 2:e38.
14. Bustamante CD, *et al.* (2005) Natural selection on protein-coding genes in the human genome. *Nature* 437:1153–1157.
15. Gibbs RA, *et al.* (2007) Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 316:222–234.
16. Nielsen R, *et al.* (2005) A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol* 3:e170.
17. Sabeti PC, *et al.* (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* 449:913–918.
18. Nei M (2007) The new mutation theory of phenotypic evolution. *Proc Natl Acad Sci USA* 104:12235–12242.
19. Davidson EH, Erwin DH (2006) Gene regulatory networks and the evolution of animal body plans. *Science* 311:796–800.
20. Goodman M (1961) The role of immunochemical differences in the phyletic development of human behavior. *Hum Biol* 33:131–162.
21. Frank M, Kemler R (2002) Protocadherins. *Curr Opin Cell Biol* 14:557–562.
22. Huguenard JR (1996) Low-threshold calcium currents in central nervous system neurons. *Annu Rev Physiol* 58:329–348.
23. Paracchini S, *et al.* (2006) The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Hum Mol Genet* 15:1659–1666.
24. Rilling JK, Seligman RA (2002) A quantitative morphometric comparative analysis of the primate temporal lobe. *J Hum Evol* 42:505–533.
25. Schenker NM, Desgouttes AM, Semendeferi K (2005) Neural connectivity and cortical substrates of cognition in hominoids. *J Hum Evol* 49:547–569.
26. Sherwood CC, Hof PR (2007) The evolution of neuron types and cortical histology in apes and humans. *The Evolution of Primate Nervous Systems*, ed Preuss TM, Kaas JH (Academic, Oxford), pp 355–378.
27. Grossman LI, Wildman DE, Schmidt TR, Goodman M (2004) Accelerated evolution of the electron transport chain in anthropoid primates. *Trends Genet* 20:578–585.
28. Uddin M, *et al.* (2004) Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles. *Proc Natl Acad Sci USA* 101:2957–2962.
29. Sherwood CC, *et al.* (2006) Evolution of increased glia-neuron ratios in the human frontal cortex. *Proc Natl Acad Sci USA* 103:13606–13611.
30. Aiello LC, Wheeler P (1995) The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution. *Curr Anthropol* 36:199–221.
31. Fleagle JG (1999) *Primate Adaptation and Evolution* (Academic, San Diego), Second Ed.
32. Bernal J (2005) Thyroid hormones and brain development. *Vitam Horm* 71:95–122.
33. de Escobar GM, Obregon MJ, del Rey FE (2004) Maternal thyroid hormones early in pregnancy and fetal brain development. *Best Pract Res Clin Endocrinol Metab* 18:225–248.
34. Kester MH, *et al.* (2004) Iodothyronine levels in the human developing brain: Major regulatory roles of iodothyronine deiodinases in different areas. *J Clin Endocrinol Metab* 89:3117–3128.
35. Morreale de Escobar G, Obregon MJ, Escobar del Rey F (2004) Role of thyroid hormone during early brain development. *Eur J Endocrinol* 151 Suppl 3:U25.
36. Arnold S, Goglia F, Kadenbach B (1998) 3,5-Diiodothyronine binds to subunit Va of cytochrome-*c* oxidase and abolishes the allosteric inhibition of respiration by ATP. *Eur J Biochem* 252:325–330.
37. Sheehan TE, Kumar PA, Hood DA (2004) Tissue-specific regulation of cytochrome *c* oxidase subunit expression by thyroid hormone. *Am J Physiol* 286:E968–E974.
38. Wrutniak-Cabello C, Casas F, Cabello G (2001) Thyroid hormone action in mitochondria. *J Mol Endocrinol* 26:67–77.
39. Fernandez V, *et al.* (2006) Thyroid hormone-induced oxidative stress in rodents and humans: A comparative view and relation to redox regulation of gene expression. *Comp Biochem Physiol C Toxicol Pharmacol* 142:231–239.
40. Rose RW, Kuswanti N (2004) Thyroid function and the development of endothermy in a marsupial, the Tasmanian bettong, *Bettongia gaimardi* (Demarest 1822). *Gen Comp Endocrinol* 136:17–22.
41. Rochard P, *et al.* (2000) Mitochondrial activity is involved in the regulation of myoblast differentiation through myogenin expression and activity of myogenic factors. *J Biol Chem* 275:2733–2744.
42. Jeschke U, *et al.* (2002) Expression of the Thomsen–Friedenreich antigen and of its putative carrier protein mucin 1 in the human placenta and in trophoblast cells *in vitro*. *Histochem Cell Biol* 117:219–226.
43. Shalom-Barak T, *et al.* (2004) Peroxisome proliferator-activated receptor gamma controls Muc1 transcription in trophoblasts. *Mol Cell Biol* 24:10661–10669.
44. Thirkill TL, *et al.* (2007) MUC1 is involved in trophoblast transendothelial migration. *Biochim Biophys Acta* 1773:1007–1014.
45. Sood R, Zehnder JL, Druzin ML, Brown PO (2006) Gene expression patterns in human placenta. *Proc Natl Acad Sci USA* 103:5478–5483.
46. Carter AM (2007) Animal models of human placentation—a review. *Placenta* 28 Suppl A:S41–S47.
47. Bado A, *et al.* (1998) The stomach is a source of leptin. *Nature* 394:790–793.
48. Mise H, *et al.* (2007) The relationship between maternal plasma leptin levels and fetal growth restriction. *Endocr J* 10.1507/endocrj. K06-225.
49. Mittal P, *et al.* (2007) Placental growth hormone is increased in the maternal and fetal serum of patients with preeclampsia. *J Matern Fetal Neonatal Med* 20:651–659.
50. Hallast P, Rull K, Laan M (2007) The evolution and genomic landscape of CGB1 and CGB2 genes. *Mol Cell Endocrinol* 260–262:2–11.
51. Varki A (2007) Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature* 446:1023–1029.
52. Koike C, *et al.* (2007) Functionally important glycosyltransferase gain and loss during catarrhine primate emergence. *Proc Natl Acad Sci USA* 104:559–564.
53. Lipovich L, *et al.* (2006) Primate-specific endogenous *cis*-antisense transcription in the human 5q31 protocadherin gene cluster. *J Mol Evol* 62:73–88.
54. Liu G, *et al.* (2007) OCPAT: An online codon-preserved alignment tool for evolutionary genomic analysis of protein coding sequences. *Source Code Biol Med* 2:5.
55. Yang Z (1997) PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13:555–556.
56. Yang Z (2006) *Computational Molecular Evolution.* (Oxford Univ Press, Oxford).