Opinion TRENDS in Genetics Vol.22 No.11

Did brain-specific genes evolve faster in humans than in chimpanzees?

Peng Shi*, Margaret A. Bakewell* and Jianzhi Zhang

Department of Ecology and Evolutionary Biology, University of Michigan, 1075 Natural Science Building, 830 North University Avenue, Ann Arbor, MI 48109, USA

One of the most distinctive characteristics of humans among primates is the size, organization and function of the brain. A recent study has proposed that there was widespread accelerated sequence evolution of genes functioning in the nervous system during human origins. Here we test this hypothesis by a genome-wide analysis of genes that are expressed predominantly or specifically in brain tissues and genes that have important roles in the brain, identified on the basis of five different definitions of brain specificity. Although there is little overlap among the five sets of brain-specific genes, none of them supports human acceleration. On the contrary, some datasets show significantly fewer nonsynonymous substitutions in humans than in chimpanzees for brain-specific genes relative to other genes in the genome. Our results suggest that the unique features of the human brain did not arise by a large number of adaptive amino acid changes in many proteins.

Introduction

The human brain differs substantially from those of other primates in size, organization and function. For instance, in comparison to that of chimpanzees, the brain weight of humans is over 300% greater but the body is only 35% heavier [1] (Figure 1). The structural asymmetry between the left and right hemispheres is especially pronounced in human brains [2]. Humans also have speech or language and other high-order cognitive functions that are absent in non-human primates. The genetic changes that have been responsible for the emergence of these human-unique brain features are a topic of enduring interest. Generally speaking, the marked evolution of the human brain could be due to modifications of either a small or a large number of genes, where the modifications might be in gene expression or protein function.

If widespread changes in many genes were the cause of human brain evolution, the signatures of such events might be identifiable from a genome-wide analysis. Recently, Dorus *et al.* [3] analyzed a set of nervous system genes at the protein sequence level and found that these genes evolved significantly faster in primates than in rodents, in hominoids than in Old World monkeys, and in humans than in chimpanzees. They further suggested that the accelerated evolution was due to positive Darwinian selection for advantageous amino acid changes. Their analysis, however, suffered from four shortcomings. First, they compared only 24 nervous system genes between human and chimpanzee – the most relevant species pair for studying evolution of the human brain. Second, their list of nervous system genes was manually compiled and might thus be incomplete or biased (see later). Third, they used house-keeping genes as controls in some of the analyses, which seems inappropriate because tissue-specific genes and house-keeping genes are expected to have different evolutionary patterns [4,5]. Fourth, a recent comparison between the dog and mouse genomes found that 18 nervous system genes that evolved faster in primates than in rodents also evolved faster in carnivores than in rodents [6], suggesting that the findings of Dorus *et al.* [3] might partially be due to rodent deceleration rather than primate acceleration. A more recent analysis of 5268 genes has also found more amino acid substitutions in humans than in chimpanzees for brain-specific genes; however, the statistical significance of the difference is uncertain ($P = 0.03$–0.08, depending on which genes are used as controls) and the results are inconclusive [7].

Here we conduct a comparison of sequence evolution of brain-specific genes between the human and chimpanzee lineages, using genome sequences of human, chimpanzee and macaque monkey, and human transcriptome data.

Compilation of the primate gene dataset

From Ensembl (http://www.ensembl.org), we obtained the DNA and amino acid sequences of all of the proteins predicted from the genome sequences of human (*Homo sapiens*), chimpanzee (*Pan troglodytes*) and macaque (*Macaca mulatta*). To identify orthologous genes, we used human proteins as queries to search chimpanzee proteins with BLASTP (see Supplementary Methods). Reciprocal best hits are considered as orthologs. Similarly, we used human sequences to search the macaque proteins with BLASTP. A total of 19 422 proteins with reciprocal best hits in both the human–chimpanzee and the human–macaque searches were found, and alignments of the human–chimpanzee–macaque orthologous proteins were obtained.

We discarded alignments containing fewer than 100 amino acids because most of these were caused by gaps in draft genome sequences. DNA sequence alignments were obtained from the protein alignments. We further removed 161 alignments that showed exceptionally

Corresponding author: Zhang, J. (jianzhi@umich.edu).
* Authors contributed equally to this work.
Available online 15 September 2006.

 0168-9525/$ – see front matter doi:10.1016/j.tig.2006.09.001

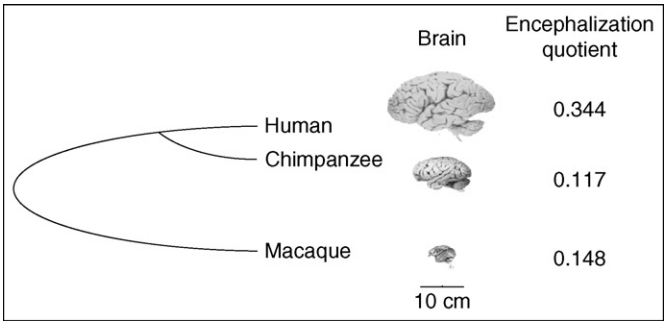

Figure 1. Evolutionary tree of human, chimpanzee and macaque monkey. Also shown are the brains of the three species drawn to scale and the encephalization quotients (EQs). The EQ measures the brain mass relative to the total body mass and is computed by E/P^a, where E is the brain mass, P is the body mass, and a is the exponent. The EQ values are taken from Ref. [1]; $a = 0.75$ on the basis of previous analyses of primates [28] or catarrhine primates (i.e. humans, apes and Old World monkeys) [29]. The brain images are adapted from those in the Comparative Mammalian Brain Collections (http://brainmuseum.org).

high divergences among the species and were probably the results of misalignment or non-orthology (see Supplementary Methods). The proportion of brain-specific genes was lower in the removed alignments than in the remaining alignments. Finally, each protein was assigned to a gene on the basis of its Ensembl annotation, resulting in 13 955 distinct genes for further analysis. After the removal of alignment gaps, these genes contain 18 287 982 nucleotide sites or 6 095 994 codons, covering >50% of all protein-coding regions in a primate genome.

We consider that a nucleotide position has a human-specific substitution if the sequence is identical between the chimpanzee and macaque but different in human at this position. We similarly define chimpanzee-specific substitutions. A nucleotide substitution is then classified as either synonymous or nonsynonymous depending on whether it alters the amino acid encoded. We observed 57 545 chimpanzee-specific nucleotide substitutions and 50 254 human-specific substitutions. Thus, the nucleotide substitution rate seems to be 1.15 times (57 545/50 254) higher in chimpanzees than in humans. This rate difference is probably due to the relatively low quality of the 4-coverage chimpanzee draft genome sequence [8], as compared with that of the finished human genome sequence [9].

A recent study estimated that the error rate in the chimpanzee genome sequence has an upper limit of 0.07% [10], ~70 times higher than the error rate in the human sequence [9]. The observed chimpanzee to human divergence is 0.59% in our dataset of coding sequences. If we assume that the actual substitution rates in humans and chimpanzees are identical, then the chimpanzee to human substitution rate ratio ($R_{C/H}$) might appear as high as 1.27, simply because of the 0.07% sequencing errors in the chimpanzee genome (Supplementary Methods). If we also consider that the mutation rate per year is slightly (3%) lower in humans than in chimpanzees [11], $R_{C/H}$ might appear as high as 1.30 (Supplementary Methods). Our $R_{C/H}$ value of 1.15 is within these limits. Our result is also comparable to a recent estimate of 1.11–1.18 for the $R_{C/H}$ for large numbers of intergenic sequences and introns obtained from a comparison of the draft chimpanzee genome sequence and the finished human sequence [11].

Several measures of the rate of protein sequence evolution have been well established by molecular evolutionists [12]. For example, let n be the number of nonsynonymous substitutions for a group of genes in a particular lineage and s be the corresponding number of synonymous substitutions; and let N and S be the numbers of nonsynonymous and synonymous sites, respectively, for the group of genes [12]. For any large group of genes in our dataset, $N/S = 2.45$ (Supplementary Methods). Thus, the nonsynonymous-to-synonymous rate ratio (w), which is commonly used to measure the rate of protein evolution controlled by the mutation rate, becomes $(n/N)/(s/S) = (n/s)/(N/S) = (n/s)/2.45 = 0.408n/s$. Because most genes in a genome have a w value of <1, whereas sequencing errors are expected to have a w value of 1, the errors cause overestimation of w. Thus, we would see a higher w for chimpanzees than for humans owing to chimpanzee sequencing errors. Furthermore, the bias is more serious for genes with low w than for genes with high w (when $w < 1$). Because brain-specific genes tend to have lower w values than other genes in the genome (Table 1), the former are affected by sequencing errors to a greater extent than the latter. Thus, we expect to observe a higher w in chimpanzees than in humans for brain-specific genes, even when benchmarked by other genes in the genome (Supplementary Table 1).

To rectify this problem, we add the same number of random 'sequencing errors' to the human genome sequence as the number that occurred in the chimpanzee sequence. Although sequencing errors will still affect the w of brain-specific genes more than that of other genes, the human and chimpanzee lineages can now be compared. Assuming that the total numbers of substitutions in our 13 955 genes are equal between the human and chimpanzee lineages, we estimate that the error rate in the chimpanzee sequence is 0.04%, which is equal to 7315 errors (Supplementary Methods). We thus randomly add this number of errors to the human sequence and then compare the human and chimpanzee sequences. Although the 4.6-coverage macaque genome sequence might also contain numerous sequencing errors, these errors are not expected to bias our comparison between human and chimpanzee because the macaque is used as an outgroup.

Analysis of brain-specific genes

It is not an easy task to define those genes that function specifically in the brain. We therefore use five different definitions to examine whether they provide consistent results.

Analysis based on microarray data

Our first definition is based on a human microarray gene expression dataset [13], which includes the expression signals of almost all human genes in 73 normal tissues. Because many of the 73 tissues are from the same organs, we group the tissues into 40 tissue groups (Supplementary Table 2). For example, the brain tissue group includes 17 tissues that represent different developmental stages or parts of the brain. Brain-specific genes are defined as those genes for which the highest expression is found in one of the brain tissues and this highest expression is at least twice the expression level in any non-brain tissues. As a

Table 1. Evolutionary rates of brain-specific genes and other genes in humans and chimpanzees

Tissue-specificity definitions	Genes	No. of genes	Human lineage			Chimpanzee lineage				
			n_H[a]	s_H[b]	w_H[c]	n_C[a]	s_C[b]	w_C[c]	w_H/w_C[d]	n_H/n_C[e]
Microarray (2×)	Brain-specific genes	249	286	571	0.205	318	655	0.198	1.03	0.90
	Other tissue-specific genes	1544	2897	3621	0.327	2833	3765	0.307	1.06	1.02
	Non-tissue-specific genes	12 162	21 710	28 241	0.314	20 394	29 620	0.281	1.12****	1.06
	All genes	13 955	24 893	32 432	0.313	23 546	34 040	0.282	1.11****	1.06
	Ratio of brain to other tissue-specific				0.626			0.646	0.97	0.88
	Ratio of brain to non-tissue-specific				0.652			0.706	0.92	0.84*
Microarray (4×)	Brain-specific genes	72	91	182	0.205	104	192	0.221	0.93	0.88
	Other tissue-specific genes	502	973	1178	0.337	962	1176	0.334	1.01	1.01
	Non-tissue-specific genes	13 381	23 829	31 071	0.313	22 480	32 671	0.281	1.11****	1.06
	Ratio of brain to other tissue-specific				0.607			0.662	0.92	0.87
	Ratio of brain to non-tissue-specific				0.653			0.786	0.83	0.83
EST	Brain-specific genes	165	294	430	0.279	324	493	0.268	1.04	0.91
	Other tissue-specific genes	819	1963	2083	0.385	1891	2313	0.334	1.15**	1.04
	Non-tissue-specific genes	12 971	22 637	29 920	0.309	21 331	31 234	0.279	1.11****	1.06
	Ratio of brain to other tissue-specific				0.725			0.804	0.90	0.87
	Ratio of brain to non-tissue-specific				0.903			0.962	0.94	0.85
SAGE	Brain-specific genes	209	356	494	0.295	368	550	0.273	1.08	0.97
	Other tissue-specific genes	632	1214	1485	0.334	1238	1580	0.320	1.04	0.98
	Non-tissue-specific genes	13 114	23 323	30 454	0.313	21 939	31 911	0.281	1.11****	1.06
	Ratio of brain to other tissue-specific				0.883			0.854	1.03	0.99
	Ratio of brain to non-tissue-specific				0.942			0.974	0.97	0.91
Nervous system genes[f]	Nervous system genes	146	196	341	0.235	193	407	0.193	1.22	1.02
	Developmental	37	53	91	0.237	47	113	0.169	1.40	1.13
	Physiological	61	59	135	0.178	66	170	0.159	1.12	0.89
	Unclassified	48	85	115	0.300	80	124	0.263	1.14	1.06
	Other genes	13 809	24 697	32 091	0.314	23 353	33 633	0.283	1.11**	1.06
	Ratio of nervous system to other genes				0.748			0.682	1.10	0.96
	Ratio of developmental to other genes				0.754			0.596	1.26	1.07
Overlapping sets[g]	Brain-specific genes	74	86	176	0.199	117	216	0.221	0.90	0.73
	Other genes	13 881	24 808	32 256	0.314	23 429	33 824	0.283	1.11****	1.06
	Ratio of brain to other genes				0.632			0.781	0.81	0.69**

[a]Number of nonsynonymous substitutions in the lineage indicated.
[b]Number of synonymous substitutions in the lineage indicated.
[c]Nonsynonymous/synonymous substitution rate ratio, computed by $0.408n/s$.
[d]Statistically significant deviation from 1 is indicated by asterisks: Significance level: $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$. Simulation tests are used for comparing ratios between groups of genes, whereas χ^2-tests are used within groups of genes.
[e]Statistically significant deviation from 1 (between groups of genes) is indicated by asterisks. χ^2-tests are used.
[f]From Dorus *et al.* [3].
[g]Genes identified to be brain-specific in at least two of the four definitions ('Microarray [2×]', 'EST', 'SAGE' and 'Nervous system genes').

result, 249 brain-specific genes are identified. Similarly, we identified tissue-specific genes for the other 39 tissue groups, and the total number of these other tissue-specific genes is 1544. The remaining 12 162 genes are referred to as non-tissue-specific genes.

For our second definition, we used the same human microarray gene expression dataset but with more stringent criteria, requiring that the highest expression level in a brain tissue is at least four times that in any non-brain tissue for a gene to be called brain-specific. Because the results based on these two definitions are almost identical, below we describe in detail only those from the first definition (Table 1, 'Microarray [2×]'; see 'Microarray [4×]' for the results from the second definition).

We find that, for brain-specific genes, the w value in the human lineage (w_H) is 0.205 and that in the chimpanzee lineage (w_C) is 0.198. Their ratio ($w_H/w_C = 1.03$) is not significantly different from 1 ($P > 0.5$, χ^2-test; Table 1). As a comparison, w_H/w_C equals 1.06 ($P > 0.05$) for other tissue-specific genes, 1.12 ($P < 10^{-4}$) for non-tissue-specific genes, and 1.11 ($P < 10^{-4}$) for all of the genes considered together. The observation of $w_H/w_C > 1$ for all genes together is consistent with previous findings and is explainable by a smaller effective population size and thus weaker purifying selection and a higher nonsynonymous substitution rate in the human lineage than in the chimpanzee lineage [7,8,14]. We find that the w_H/w_C ratio of brain-specific genes is slightly lower than that of other tissue-specific genes, but the difference is not statistically significant ($P > 0.5$, simulation test; Table 1).

Similar results are obtained when brain-specific genes are compared with non-tissue-specific genes (Table 1). Because the same genes are compared between human and chimpanzee, we can compute the ratio of the number of nonsynonymous substitutions in the human lineage (n_H) to that in the chimpanzee lineage (n_C) and compare this ratio (n_H/n_C) between different groups of genes. Interestingly, we find that n_H/n_C is significantly lower for brain-specific genes than for non-tissue-specific genes ($P = 0.04$, χ^2-test; Table 1), suggesting a possible human slowdown (or chimpanzee acceleration) of the evolution of brain-specific genes, when benchmarked by non-tissue-specific genes.

The n_H/n_C values are not, however, significantly different between brain-specific genes and other tissue-specific genes, or between other tissue-specific genes and non-tissue-specific genes.

Analysis based on EST data

Because the microarray data might be inaccurate [15], we repeated the above analysis using a third definition of brain-specific genes based on expression sequence tags (ESTs). Here, tissue-specific genes are those for which ESTs are found in only one tissue. We used a recently compiled human EST dataset that includes 4.9 million ESTs from 44 tissues [16] and classified the 13 955 primate genes into 165 brain-specific genes (i.e. ESTs are found only in the brain), 819 other tissue-specific genes, and 12 971 non-tissue-specific genes. The results from the EST data (Table 1) are similar to those from the microarray data. Although w_H/w_C is significantly greater than 1 for other tissue-specific genes and non-tissue-specific genes, it is not significantly greater than 1 for brain-specific genes. Consequently, the w_H/w_C ratio is slightly lower for brain-specific genes than for other genes, although the difference is not statistically significant (Table 1). Similarly, the n_H/n_C ratio appears lower, although not significantly, in brain-specific genes than in other genes (Table 1).

Analysis based on SAGE data

We also repeated the above analysis using a fourth definition of tissue specificity based on serial analysis of gene expression (SAGE). Brain-specific genes are defined as those for which SAGE tags are detected only in the brain. On the basis of a recently compiled SAGE dataset [16], the 13 955 primate genes include 209 brain-specific genes and 632 other tissue-specific genes. The remaining genes are considered to be non-tissue-specific. The results obtained from the SAGE data (Table 1) are similar to those from the microarray and EST data. That is, there is no significant difference between w_H and w_C for brain-specific genes, regardless of whether other genes are used as controls or not. There is also no significant difference between the n_H/n_C ratios of brain-specific genes and other genes.

Analysis based on a list of nervous system genes

Dorus *et al.* [3] compiled a list of 214 nervous system genes on the basis of (i) literature suggesting important gene functions in the nervous system, (ii) SAGE and EST data showing gene expression exclusively or predominantly in the brain, and (iii) information on genes implicated in nervous system diseases [3]. We found 146 of these 214 genes in our list of 13 955 primate genes. Because Dorus

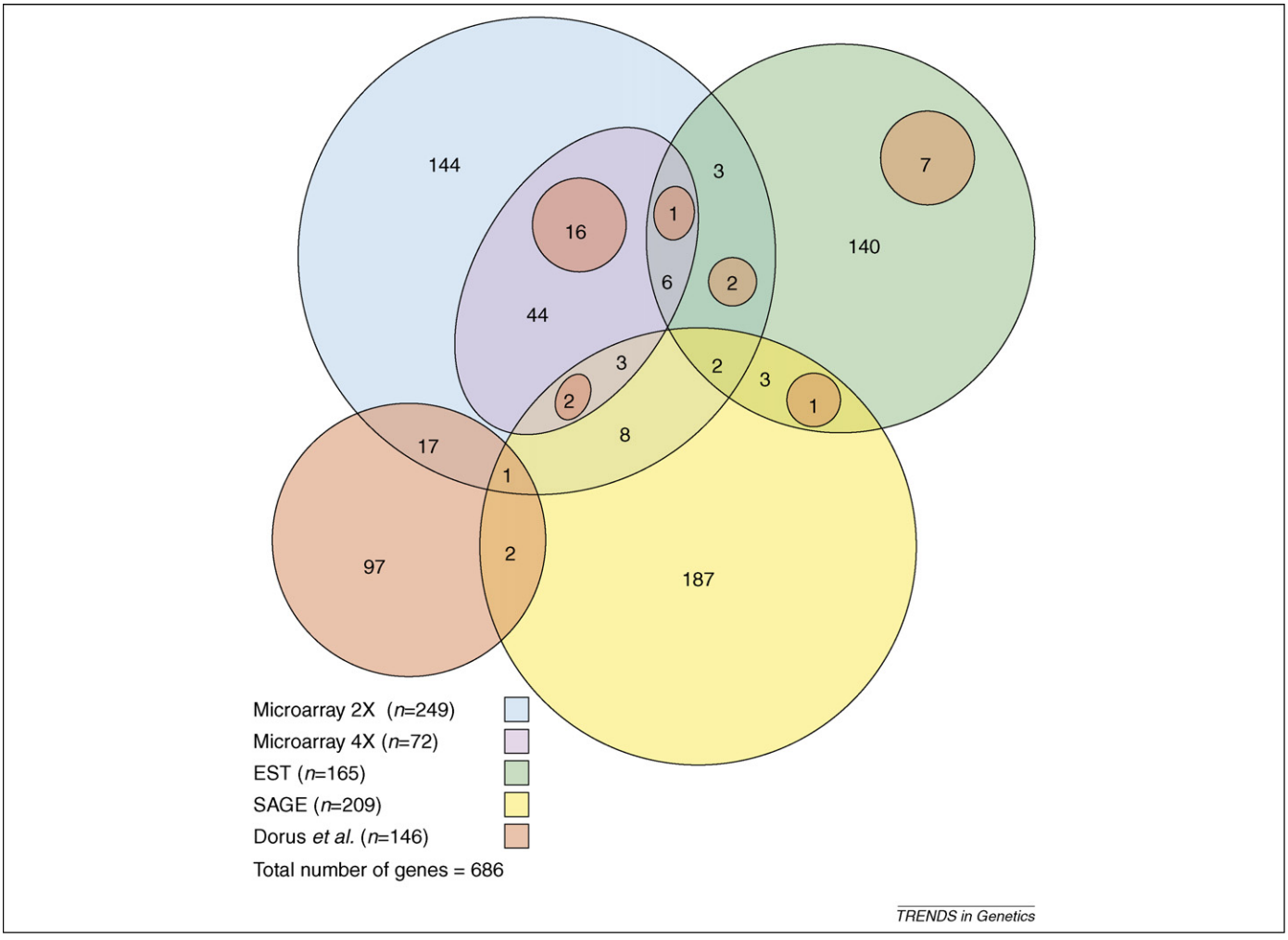

Figure 2. Venn diagram of brain-specific genes identified on the basis of five different definitions. The number of genes is given in each circle. The overlapping sets of the nervous system genes (from Dorus *et al.* [3]; colored red) are shown in separate circles because of the difficulty in connecting all of the circles.

et al. [3] did not define other tissue-specific genes, we analyzed these nervous system genes by using the remaining 13 809 genes in our dataset as a control. We find no significant difference between w_H and w_C for nervous system genes, with or without comparison to other genes (Table 1). Dorus *et al.* [3] suggested that the human lineage acceleration is particularly pronounced for a subset of genes that control nervous system development, but is absent for genes with physiological roles and minimal for the remaining (i.e. unclassified) nervous system genes. Our data, however, provide no statistical evidence for these claims (Table 1). We also failed to detect a difference in n_H/n_C between nervous system genes (or developmental nervous system genes) and other genes (Table 1).

The main reason why we cannot repeat the result of the faster evolution of humans than chimpanzees even when we use the list of nervous system genes that Dorus *et al.* [3] compiled seems to be because Dorus *et al.* did not compare all of the 214 nervous system genes between human and chimpanzee. Instead, between humans and chimpanzees they compared only 24 genes that were known to evolve faster in the human lineage than in the macaque lineage when the squirrel monkey was used as an outgroup. In other words, they used a small and biased gene set in their human–chimpanzee comparison.

Caveats

Although our results from the five analyses are congruent in showing that there has been no accelerated evolution of human brain-specific genes, this congruence would be expected if there were large overlaps among the five groups of brain-specific genes identified under the five different definitions. Interestingly, however, except for those identified by the two microarray-based definitions, only a few genes overlap from any two of the five groups of brain-specific genes and no genes overlap among all five groups (Figure 2).

Although this finding suggests that the five analyses are largely independent, it also raises the issue of how to identify brain-specific genes accurately. The level of gene expression in a tissue is a continuous variable. For the EST (or SAGE) data, we identified brain-specific genes as those that lack ESTs (or SAGE tags) in non-brain tissues, which actually means genes that have a lower expression level in non-brain tissues than in the brain. This definition is qualitatively the same as that used for the microarray data, where brain-specific genes are required to show expression at least twice as high in the brain as in any other tissue. Although it might be argued that a gene that exclusively functions in the brain could have a lower expression in this organ than in other tissues, such a situation is unlikely, particularly when expression in the brain is defined by the highest expression level among all temporal and spatial brain samples.

All five definitions that we used consider gene expression patterns, although the fifth definition also includes genes with known brain functions and genes implicated in brain diseases. On the one hand, considering gene function provides additional information that might help to reduce the reliance on gene expression, which is sometimes a poor indicator of function. On the other hand, gene function information is usually incomplete and it is difficult to know whether a gene functions exclusively in the brain. Our results suggest that it is still a challenging task to define genes that function specifically in a tissue. A potential way of increasing the accuracy of identifying brain-specific genes is to use more than one criterion. We therefore analyzed a subset of 74 genes that are brain-specific by at least two of our definitions 1, 3, 4 and 5; we excluded definition 2 because it is a subset of definition 1 (Figure 2). The difference between w_H and w_C of brain-specific genes, with or without comparison to other genes in the genome, is still not significant (Table 1). Interestingly, however, the n_H/n_C ratio is significantly lower for brain-specific genes than for other genes in the genome (Table 1).

Our analysis also highlights the intricacy of genome-wide comparisons between humans and chimpanzees in the presence of sequencing errors. As eloquently articulated by Taudien *et al.* [10], a small leak can sink a great ship. In our analysis, the chimpanzee sequencing errors, when not appropriately controlled, generate a significantly higher w_C than w_H for brain-specific genes, even when compared with other genes in the genome (Supplementary Table 1). This difference disappears when we add the same number of 'sequencing errors' to the human sequence. In our addition of sequencing errors to the human sequence, we assumed that the substitution rate for the whole set of 13 955 genes is identical between the human and chimpanzee lineages. If the mutation rate is slightly lower in humans than in chimpanzees [11] and the total substitution rate is also lower in humans than in chimpanzees, we might have added more 'sequencing errors' than needed, which would have raised w_H/w_C and favored the human acceleration hypothesis. In other words, our result of no human acceleration is conservative (see also Supplementary Methods).

To verify the results obtained from our approach of error addition, we also used the approach of error removal. We removed errors from the chimpanzee sequence by using only nucleotide sites with quality scores $\geq$20 (or accuracy $>$ 99%) [8]. The new dataset contained 13 888 genes. Again, none of the analyses shows a significantly higher evolutionary rate of brain-specific genes in humans than in chimpanzees (Supplementary Table 3).

Concluding remarks

We have analyzed almost 14 000 human, chimpanzee and macaque genes to test the hypothesis that human brain-specific genes have undergone widespread accelerated protein-sequence evolution since the human lineage separated from the chimpanzee lineage. Our results, based on five different definitions of brain-specificity, show no evidence that supports this hypothesis. Because our data include over 50% of all human genes, it is appropriate to conclude that our results reject the hypothesis of widespread accelerated sequence evolution of human brain-specific genes.

In fact, in several but not all of our analyses, the n_H/n_C ratio is significantly lower for brain-specific genes than for other genes in the genome, suggesting that – relative to other genes – brain-specific genes evolved more slowly in

humans than in chimpanzees. This phenomenon might reflect higher importance of brain-specific genes and therefore stronger purifying selection on them in human evolution than in chimpanzee evolution. Our findings imply that the unique features of the human brain did not arise by a large number of adaptive amino acid substitutions in many proteins. This conclusion, however, does not preclude the possibility that substantial accelerations occurred in the evolution of a few nervous system genes during human origins. Indeed, several such examples are known, including genes that control brain size and speech development [17–23]. It also remains possible that the origin of the human-unique brain features was due to expression changes (rather than coding sequence changes) of many genes, as has been suggested from some microarray data [7,24] (but see also Refs [25–27]).

Acknowledgements

We thank Laurent Duret for providing the EST and SAGE datasets, the macaque genome sequencing team for making the draft sequence available, and Soochin Cho, Wendy Grus and Ben-Yang Liao for valuable comments. This work was supported by research grants from the University of Michigan and the National Institutes of Health to J.Z.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tig.2006.09.001.

References

1 Williams, M.F. (2002) Primate encephalization and intelligence. *Med. Hypotheses* 58, 284–290
2 Deacon, T.W. (1994) The human brain. In *The Cambridge Encyclopedia of Human Evolution* (Jones, S. *et al.*, eds), pp. 115–123, Cambridge University Press
3 Dorus, S. *et al.* (2004) Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. *Cell* 119, 1027–1040
4 Duret, L. and Mouchiroud, D. (2000) Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol. Biol. Evol.* 17, 68–74
5 Zhang, L. and Li, W.H. (2004) Mammalian housekeeping genes evolve more slowly than tissue-specific genes. *Mol. Biol. Evol.* 21, 236–239
6 Lindblad-Toh, K. *et al.* (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438, 803–819
7 Khaitovich, P. *et al.* (2005) Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* 309, 1850–1854
8 Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437, 69–87
9 International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945
10 Taudien, S. *et al.* (2006) Should the draft chimpanzee sequence be finished? *Trends Genet.* 22, 122–125
11 Elango, N. *et al.* (2006) Variable molecular clocks in hominoids. *Proc. Natl. Acad. Sci. U. S. A.* 103, 1370–1375
12 Nei, M. and Kumar, S. (2000) *Molecular Evolution and Phylogenetics.* Oxford University Press
13 Su, A.I. *et al.* (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6062–6067
14 Eyre-Walker, A. and Keightley, P.D. (1999) High genomic deleterious mutation rates in hominids. *Nature* 397, 344–347
15 Liao, B.Y. and Zhang, J. (2006) Evolutionary conservation of expression profiles between human and mouse orthologous genes. *Mol. Biol. Evol.* 23, 530–540
16 Semon, M. *et al.* (2006) No evidence for tissue-specific adaptation of synonymous codon usage in humans. *Mol. Biol. Evol.* 23, 523–529
17 Zhang, J. (2003) Evolution of the human ASPM gene, a major determinant of brain size. *Genetics* 165, 2063–2070
18 Evans, P.D. *et al.* (2004) Adaptive evolution of ASPM, a major determinant of cerebral cortical size in humans. *Hum. Mol. Genet.* 13, 489–494
19 Evans, P.D. *et al.* (2004) Reconstructing the evolutionary history of microcephalin, a gene controlling human brain size. *Hum. Mol. Genet.* 13, 1139–1145
20 Wang, Y.Q. and Su, B. (2004) Molecular evolution of microcephalin, a gene determining human brain size. *Hum. Mol. Genet.* 13, 1131–1137
21 Enard, W. *et al.* (2002) Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature* 418, 869–872
22 Zhang, J. *et al.* (2002) Accelerated protein evolution and origins of human-specific features: Foxp2 as an example. *Genetics* 162, 1825–1835
23 Kouprina, N. *et al.* (2004) Accelerated evolution of the ASPM gene controlling brain size begins prior to human brain expansion. *PLoS Biol.* 2, E126
24 Enard, W. *et al.* (2002) Intra- and interspecific variation in primate gene expression patterns. *Science* 296, 340–343
25 Hsieh, W.P. *et al.* (2003) Mixed-model reanalysis of primate data suggests tissue and species biases in oligonucleotide-based gene expression profiles. *Genetics* 165, 747–757
26 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. U. S. A.* 101, 2957–2962
27 Lemos, B. *et al.* (2005) Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. *Evolution Int. J. Org. Evolution* 59, 126–137
28 Martin, R.D. (1981) Relative brain size and basal metabolic rate in terrestrial vertebrates. *Nature* 293, 57–60
29 Pagel, M.D. and Harvey, P.H. (1989) Taxonomic differences in the scaling of brain on body weight among mammals. *Science* 244, 1589–1593