Parallel Patterns of Evolution in the Genomes and Transcriptomes of Humans and Chimpanzees

Philipp Khaitovich, Ines Hellmann, Wolfgang Enard, Katja Nowick, Marcus Lehrner, Henriette Franz,Gunter Weiss, Michael Lachmann, Svante Pääbo

The determination of the chimpanzee genome sequence provides a means to study both structural and functional aspects of the evolution of the human genome. Here we compare humans and chimpanzees with respect to differences in expression levels and protein-coding sequences for genes active in brain, heart, liver, kidney, and testis. We find that the patterns of differences in gene expression and gene sequences are markedly similar. In particular, there is a gradation of selective constraints among the tissues so that the brain shows the least differences between the species whereas liver shows the most. Furthermore, expression levels as well as amino acid sequences of genes active in more tissues have diverged less between the species than have genes active in fewer tissues. In general, these patterns are consistent with a model of neutral evolution with negative selection. However, for X-chromosomal genes expressed in testis, patterns suggestive of positive selection on sequence changes as well as expression changes are seen. Furthermore, although genes expressed in the brain have changed less than have genes expressed in other tissues, in agreement with previous work we find that genes active in brain have accumulated more changes on the human than on the chimpanzee lineage.

In some behavioral and cognitive traits, humans have changed dramatically since their evolutionary divergence from a common ancestor shared with chimpanzees (1, 2). It seems reasonable to assume that a number of these changes were driven by positive Darwinian selection. However, although positive selection has been demonstrated for several human genes (3–5), the overall patterns of evolution of chimpanzee and human genes are consistent with selective neutrality (6). It has long been argued that changes in gene expression may provide an additional and crucial perspective on the evolutionary differences between humans and chimpanzees (7), but relevant data to address this issue have only recently started to become available (8). On a more general level, data from yeast, fruit flies, humans, and mice have been used to argue that regulatory evolution and protein evolution act independently of each other and thus that they are “decoupled” (9, 10). However, other results seem to contradict this assertion (11–14). The chimpanzee and human genome sequences now provide the opportunity to address these questions by studying the evolution of gene expression, as well as of the DNA sequences of the genes expressed in various tissues in two closely related mammals. To this end, we have measured gene expression in five different tissues in six humans and five chimpanzees. We find that gene sequences and gene expression evolve in qualitatively similar manners, suggesting that the evolutionary forces that act on them are similar in effect and nature. Through analyses of evolutionary patterns at both levels, it is possible to identify groups of genes that violate neutral expectations and may have been positively selected.

Using probes on Affymetrix U133plus2 arrays that target sequences that are identical between the human and the chimpanzee genomes (15), we analyzed the expression for 51,460 probe sets (∼21,000 genes) in heart, kidney, liver, testis, and prefrontal cortex of the brain from six humans and five chimpanzees (table S1). In each tissue, we measured the extent of differences in gene expression between and within species as an average squared difference in normalized expression across all probe sets with detectable gene expression (table S2). Figure 1 schematically illustrates the results. Two major findings stand out. First, gene expression patterns differ less between humans and chimpanzees in the brain than in the other tissues (bootstrap test, P < 0.0001). Second, the ratio of expression divergence between species to diversity within species is higher in testis than in any other tissue (5.6 versus 1.8 to 2.5, P < 0.0001). Consequently, 32% of the probe sets detected in testis show significant expression differences between humans and chimpanzees, whereas ~8% do so in brain, heart, kidney, and liver (fig. S1). It is conceivable that the patterns of transcriptome divergence and diversity observed among the five tissues are mainly due to differences between tissue-specific genes, i.e., those expressed in one single tissue. Alternatively, the patterns could be due to differences also in genes that are expressed in several tissues. To distinguish between these two alternatives, we analyzed probe sets detected in all five tissues, and probe sets specific to one tissue, separately. We find that both groups of genes show similar patterns of evolution (fig. S2). In particular, brain shows fewer differences than other tissues and testis shows an excess of divergence relative to diversity (table S3). Thus, the different expression patterns observed among tissues are due to effects that a tissue exerts not only on genes expressed in that tissue but also on genes expressed in that as well as in many other tissues. A further noteworthy finding is that ubiquitously expressed genes differ less among individuals within a species as well as between species than do genes expressed in single tissues (table S3; fig. S2).

Next, we analyzed the evolution of protein-coding DNA sequences of genes for which expression was detected in at least one tissue (15). As an estimate of the protein divergence of each gene, we used the number of nonsynonymous nucleotide substitutions per nonsynonymous site (Ka), normalized to the number of substitutions per site in inter-
spersed repeats in a 250-kbp window around the center of each gene (Ki) (6, 15). In general, low Ka/Ki ratios indicate stronger purifying selection acting on amino acid substitutions, whereas higher ratios indicate fewer constraints or possibly an enrichment of amino acid substitutions by positive selection [for a review, see (16)]. In agreement with previous work (17–20), we find that brain-specific genes show Ka/Ki ratios that are significantly lower than those of other tissue-specific genes (Mann-Whitney U-test, \( P < 10^{-6} \)) (Fig. 2A) and that ubiquitously expressed genes show lower Ka/Ki ratios than genes expressed in single tissues (Mann-Whitney U-test, \( P < 10^{-6} \)) (table S4).

When the divergence of gene expression is similarly analyzed with respect to tissues (Fig. 2B), the results show that for both sequence and expression divergence, brain shows the least differences and liver the most, with testis, heart, and kidney at intermediate levels. Consequently, the higher the expression divergence in a tissue, the higher the protein divergence (Pearson’s \( r = 0.94, P < 0.05 \)) (Fig. 2C). Parallel patterns can also be seen with respect to the breadth of expression, i.e., the number of tissues in which a gene is expressed. Genes expressed in only one tissue show the highest expression and sequence divergence, and genes expressed in all five tissues the lowest divergence. This parallelism between expression divergence and protein divergence is also seen when analyzed on a gene-by-gene basis (\( R^2 = 0.0011, P < 10^{-6} \)), implying that similar factors influence protein and expression divergence. Two such factors are the tissues in which a gene is expressed and its expression breadth. Both factors influence expression divergence (multiway analysis of variance \( R^2 = 0.075, P < 10^{-6} \)) and protein divergence (\( R^2 = 0.071, P < 10^{-6} \)). If we correct for the influence of these factors, the relation between expression and protein divergence becomes much weaker but remains significant (\( R^2 = 0.00019, P < 0.05 \)). This is not surprising, given that we do not consider other factors that may affect both expression and sequence divergence, such as protein-protein interactions (21, 22). The weak relation between expression and sequence divergence is likely due to the inherently large measurement errors of expression data. In addition, it may indicate that some evolutionary forces affect gene expression and protein divergence differentially.

We also analyzed the relation between expression divergence and sequence divergence in putative core promoters (Kp), defined as a 1500-bp region upstream and \( \leq 500 \)-bp region downstream of the transcriptional start (15). Kp as well as the ratio Kp/Ki are significantly correlated with expression divergence (\( R^2 = 0.001, P < 10^{-6} \) and \( R^2 = 0.0004, P < 10^{-3} \), respectively) (table S4). Given that genetic differences in promoters are more likely to directly cause differences in expression levels than differences in coding regions, these correlations may seem surprisingly weak. However, many or most sites in these promoter regions are likely not relevant for transcriptional activity (median Kp/Ki = 0.82 versus 0.15 for Ka/Ki) and the relevant transcription start sites might not be identified for all tissues. Much more work is necessary to elucidate the relation between the evolution of promoter sequences and expression levels.

Our analyses show that each tissue is associated with a certain level of evolutionary constraints acting on the genes expressed in it—for instance, brain imposes more constraints than liver. These constraints add up across tissues so that genes expressed in many tissues are subject to more constraints than are genes expressed in few tissues. The signatures of these constraints are seen both at the level of DNA sequence differences and at the level of expression differences. We have recently suggested that the evolution of gene expression patterns largely conforms to the predictions of a neutral model of evolution (23), i.e., that most expression differences observed within and between species are selectively neutral or nearly neutral. Because most evolutionary

---

**Fig. 1.** Schematic illustration of gene expression variation among and between humans and chimpanzees in five tissues. The trees are inferred from the mean of the squared difference of expression intensities of all detected probe sets (15). Brain shows the smallest diversity and diversity. The ratio of divergence to diversity in testis is 5.6, which is significantly different from the ratio in all other tissues (table S2).

**Fig. 2.** Protein sequence and gene expression divergences between humans and chimpanzees and their correlation. (A) Median protein sequence divergence (Ka/Ki), of genes expressed in one tissue (lightest color, left) to five tissues (darkest color, right). (B) Median expression divergence of genes expressed in one tissue (lightest color, left) to five tissues (darkest color, right). (C) Correlation of expression and protein sequence divergences. Tissues [brain (black), heart (red), kidney (green), liver (dark blue), testis (cyan)] with a high amino acid sequence divergence tend to have a high expression divergence (Pearson’s \( r = 0.94, P < 0.05 \)). All error bars in the figure represent 95% confidence intervals of the median values as calculated from 10,000 bootstrap replicates.
changes in nucleotide sequences conform to a neutral theory (24), the parallelism between sequence evolution and expression evolution observed here supports the notion that most evolutionary changes in gene expression are similarly selectively neutral or nearly neutral (23). A consequence of the neutral hypothesis is that the extent of expression differences found between species is largely determined by the time since they shared a common ancestor and the extent of negative selection in a particular tissue [see also (14)].

Our observation that brain, heart, kidney, and liver have similar ratios of expression divergence between species to diversity within species (Fig. 1, table S2) is compatible with a model in which gene expression changes are a function of time. The divergence to diversity ratios are smaller than would be expected if time were the sole factor influencing gene expression. A probable explanation for this is that experimental and environmental variation contributes proportionally more to interindividual differences than to divergence. Because deviations from neutral expectations can indicate the action of positive selection, we next attempted to identify such deviations in gene expression patterns, and—in a subsequent step—to corroborate such indications with observations at the DNA sequence level whenever suitable DNA sequence data were available.

It was recently proposed that a high ratio of gene expression divergence between species to gene expression diversity within species may indicate the action of positive selection (23, 25, 26). This is analogous to tests proposed for quantitative traits (27) and akin to tests that compare between- and within-species differences at functional sites to infer positive selection (28). However, because realistic evolutionary models for neutral expression changes are not yet available and because environmental factors have a considerable influence on gene expression diversity, a high ratio of divergence to diversity represents an indication rather than proof of positive selection. As seen above, testis differs from other organs studied in that the ratio of expression divergence to diversity is higher (Fig. 1). If the cellular composition of testicles differed between humans and chimpanzees more than it does for other tissues, this observation could be explained by only a few genetic differences between the species. However, although human and chimpanzee testicles differ in size, there is no evidence that the cellular composition of this organ differs between the species (29). Another possibility is that the genetic component of the expression diversity in testis is not lower than expected from the expression divergence, but that gene expression patterns in testis have a smaller environmental (i.e., nongenetic) component. In that case, we would expect genes expressed in testis to be subject to as much constraint as genes expressed in tissues such as liver or heart that have a comparable expression divergence. The property of being expressed in testis should then have a similar effect on diversity levels in other tissues as the property of being expressed in, for example, liver. However, we find that among the five tissues, expression in testis is associated with the highest number of significant reductions in diversity in tissues other than testis, whereas expression in liver is associated with the highest number of significant increases of diversity in tissues other than liver (fig. S3) (15). This suggests that strong selective constraints on genes, rather than low environmental influence, account for the low extent of expression diversity in testis. Thus, the higher ratio of gene expression divergence to diversity in testis as compared with the other tissues is indeed indicative of positive selection. Unfortunately, this pattern cannot be corroborated at the DNA sequence level because human DNA sequence diversity data collected in an unbiased way are not yet available. However, we can test predictions about the chromosomal distribution of instances of positive selection in genes active in testis. If a substantial fraction of such positively selected variants are genetically recessive, we would expect differentially expressed genes to be enriched on the X chromosome, where they could exert their full effect in males (30). Therefore, we investigated if genes with expression differences between humans and chimpanzees are unevenly distributed among chromosomes. In testis, genes on the X chromosome show a significant excess of expression differences when compared to the other chromosomes (binomial test corrected for multiple testing, $P < 10^{-5}$), whereas in the other tissues we find no significant differences among chromosomes (Fig. 3). To test if this pattern also exists at the DNA sequence level, we investigated the DNA sequence divergence of genes expressed in different tissues with respect to chromosomal location (fig. S4). For genes expressed in brain, heart, kidney, and liver, neither the autosomes nor the X differ from each other with respect to Ka/Ki. In contrast, among genes expressed in testis, those located on the X have significantly higher Ka/Ki ratios than those located

Fig. 3. The number of expression changes between humans and chimpanzees across chromosomes. Red lines indicate the normalized deviation that would be significant at $P = 0.05$, corrected for 24 tests in five tissues.
on the autosomes (Mann-Whitney U-test, \( P < 0.0005 \)) (table S5). Thus, genes expressed in testis—especially those located on the X—tend to accumulate expression changes as well as sequence changes that may have been positively selected. This is compatible with the observation that genes involved in reproduction tend to evolve under positive selection [examples in apes include (31) and (32)]. At the organismal level, this may correlate with mating strategies in different ape species (33).

Next, we examined whether differences in gene expression are equally distributed along the human and chimpanzee lineages. Because suitable data from outgroup species are lacking for most tissues, we estimated the amount of expression changes along the human and chimpanzee lineages using the observation that up-regulations of gene expression are of bigger amplitude but less numerous than are down-regulations (34). Consequently, if more gene expression changes happened on one of the two lineages, the result would be a skewed distribution of gene expression differences observed between the species [fig. S5 and (15)]. The distributions are positively and significantly skewed for brain, heart, liver, and testis (table S6), suggesting that more gene expression changes occurred on the human evolutionary lineage than on the chimpanzee lineage. In magnitude, this acceleration of gene expression change is largest in brain, and significantly larger than in any of the other tissues (\( P < 0.05 \)) except heart (\( P = 0.10 \)). This is in agreement with previous work that found a larger acceleration of gene expression changes on the human relative to the chimpanzee lineage in brain than in liver when using an orangutan as an outgroup (35–37). Thus, although gene expression is more constrained in brain than in other tissues, it has changed relatively more on the human lineage.

To investigate if such a pattern is seen also at the amino acid sequence level, we inferred how many amino acid changes occurred on the human and chimpanzee lineage, respectively, using alignments of orthologous genes from human, chimpanzee, mouse, and rat (6). For genes expressed specifically in heart, kidney, liver, and testis, the ratios of the numbers of changes on the human and chimpanzee lineages vary between 0.79 and 1.04, whereas for all genes the ratio is 1.12 (Table 1). By contrast, for genes expressed in brain, the ratio of human-specific to chimpanzeespecific amino acid changes is 1.40, higher, though not significantly (\( P = 0.08 \)), than for genes not expressed in brain and higher than for genes expressed in any other single tissue (\( P < 0.05 \)). This finding is in agreement with recent work showing a faster evolution on the human lineage for a set of genes involved in brain function and development (38). Thus, the acceleration seen for gene expression is corroborated on the sequence level for brain but not for other tissues. Such an acceleration on the human lineage could be caused by a relaxation of selective constraints on both the structure and expression of brain proteins during human evolution. A more compelling alternative is that the acceleration is caused by positive selection that changed the functions of genes expressed in the brains of humans more than in the brains of chimpanzees. However, further work elucidating the phenotypic effect of genetic changes on the human lineage is necessary to establish this.

In summary, we find that the patterns of evolutionary change in gene expression are largely compatible with a neutral model, in which different levels of constraints acting in different tissues add up for single genes. These evolutionary constraints act in a similar manner on the coding regions of DNA sequences and thus lead to parallel patterns in expression and sequence evolution. In contrast to the overall picture of selective neutrality, two examples of putative positive selection stand out. First, testis shows an excess of expression differences between species and an enrichment of both expression and amino acid sequence differences on the X chromosome. Second, the brain, although under more constraints than the other tissues, has an excess of gene expression and amino acid changes on the human lineage compared to other tissues. This suggests that evolutionary changes at both the level of gene regulation and the level of protein sequence have played crucial roles in the evolution of certain organs, such as those involved in cognition or male reproduction. Consequently, the modest number of sequence differences in genes between humans and chimpanzees cannot be taken as evidence that regulatory changes would necessarily be more important than structural protein changes during human evolution (7). Rather, both types of changes are likely to have acted in concert.

Table 1. Protein evolution in tissues along lineages. From alignments for human, chimpanzee, mouse, and rat genes (6), we calculated the amino acid changes on the human and the chimpanzee lineages for tissue-specific genes. For each tissue, we used a \( \chi^2 \)-test to determine if the ratio of human- to chimpanzee-specific changes is significantly different from the ratio observed for genes not expressed in that tissue but in one or more of the other tissues.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Genes</th>
<th>Amino acids</th>
<th>Human</th>
<th>Chimpanzee</th>
<th>Ratio</th>
<th>( P )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>201</td>
<td>81,316</td>
<td>113</td>
<td>81</td>
<td>1.40</td>
<td>0.077</td>
</tr>
<tr>
<td>Heart</td>
<td>69</td>
<td>25,604</td>
<td>37</td>
<td>47</td>
<td>0.79</td>
<td>0.173</td>
</tr>
<tr>
<td>Kidney</td>
<td>110</td>
<td>37,772</td>
<td>92</td>
<td>96</td>
<td>0.96</td>
<td>0.549</td>
</tr>
<tr>
<td>Liver</td>
<td>115</td>
<td>42,155</td>
<td>91</td>
<td>94</td>
<td>0.97</td>
<td>0.524</td>
</tr>
<tr>
<td>Testis</td>
<td>504</td>
<td>199,147</td>
<td>487</td>
<td>470</td>
<td>1.04</td>
<td>0.297</td>
</tr>
<tr>
<td>All genes</td>
<td>5268</td>
<td>2,120,869</td>
<td>3417</td>
<td>3051</td>
<td>1.12</td>
<td>–</td>
</tr>
</tbody>
</table>

References and Notes
15. Materials and methods are available as supporting material on Science Online.
39. We thank T. Mikelsen of the Broad Institute and all other members of the Chimpanzee Sequencing Con-
Achieving Stability of Lipopolysaccharide-Induced NF-κB Activation

Markus W. Covert,* Thomas H. Leung,* Jahlionais E. Gaston, David Baltimore†

The activation dynamics of the transcription factor NF-κB exhibit damped oscillatory behavior when cells are stimulated by tumor necrosis factor-α (TNFα) but stable behavior when stimulated by lipopolysaccharide (LPS). LPS binding to Toll-like receptor 4 (TLR4) causes activation of NF-κB that requires two downstream pathways, each of which when isolated exhibits damped oscillatory behavior. Computational modeling of the two TLR4-dependent signaling pathways suggests that one pathway requires a time delay to establish early anti-phase activation of NF-κB by the two pathways. The MyD88-independent pathway required interferon regulatory factor 3–dependent expression of TNFα to activate NF-κB, and the time required for TNFα synthesis established the delay.

The transcription factor NF-κB regulates numerous genes that function in diverse processes, including inflammatory responses, immune system development, apoptosis, learning in the brain, and bone development (1). Aberrant NF-κB activity has been linked to oncogenesis, tumor progression, and resistance to chemotherapy (2). NF-κB has also been identified as a tumor promoter in inflammation-associated cancer (3). Understanding the specificity and temporal mechanisms that govern NF-κB activation may therefore be important in understanding cancer progression, and systems-based and computational approaches are being developed to address this issue (4, 5).

The activity of NF-κB shows damped oscillatory behavior in cells stimulated with TNFα. Using a computational model coordinated to molecular and biochemical techniques, we have demonstrated that the oscillations in NF-κB activity are largely due to negative feedback by the NF-κB inhibitor protein IκBα (6). Another study performed in single cells has provided further evidence for these conclusions (7).

NF-κB mediates cellular responses to a wide variety of stimuli other than TNFα (8), and we wanted to determine whether NF-κB activation dynamics exhibited oscillations under other stimulation conditions. We observed non-oscillatory dynamics of active NF-κB when cells were stimulated with LPS (Fig. 1A). This difference in NF-κB activation could be linked to differences in the TNFα and LPS signaling pathways. Upon TNFα binding to the TNF receptor, the receptors aggregate and bind adaptor proteins, leading to activation of the IκB kinase (IKK) complex. Phosphorylation of IκB by IKK leads to ubiquitination and degradation of IκB and allows free NF-κB to bind target genes. One such target is IκBα, and its production results in a negative feedback loop (9–11).

In contrast, LPS signals through TLR4. TLR4 activates two downstream pathways, each of which is thought to directly activate NF-κB (12–14). The MyD88-dependent pathway recruits the kinases interleukin-1 receptor–associated kinase 1 (IRAK1) and IRAK4, which phosphorylate TNF receptor–associated factor 6 (TRAF6), leading to the activation of the IKK complex. The MyD88-independent pathway leading to NF-κB activation is not fully understood. The pathway is dependent on the TIR domain–containing adaptor inducing interferon-β (Trif) adaptor molecule, and Trif-related adaptor molecule (Tram), receptor-interactor protein 1 (RIP1), and RIP3 have been identified as important factors in the pathway (15–17). However, the end result of these pathways is the same as the end result of the TNFα-activated pathway: degradation of IκBα, which is followed by activation of IκBα gene transcription. We monitored IκBα mRNA transcript and protein levels over a 180-min time course in LPS-stimulated wild-type cells and found that IκBα protein expression decreased and

---

Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: baltimo@caltech.edu

Fig. 1. Dynamics of NF-κB activation. (A) Time course of nuclear NF-κB activation in wild-type (WT) MEFs stimulated with TNFα (10 ng/ml) or LPS (0.5 μg/ml), as indicated. NF-κB–specific mobility shifts were detected by EMSA. (B) Amounts of IκBα protein in wild-type MEFs stimulated with LPS. (C) IκBα gene expression in wild-type, Trif-deficient, and MyD88-deficient MEFs stimulated with LPS, determined by quantitative PCR (qPCR). Error bars show means ± SD. (D) Time course of nuclear NF-κB activation in Trif-deficient and MyD88-deficient MEFs stimulated with LPS. All experiments described here were repeated two or three times with a high degree of reproducibility.
Parallel Patterns of Evolution in the Genomes and Transcriptomes of Humans and Chimpanzees
Philipp Khaitovich et al.
Science 309, 1850 (2005);
DOI: 10.1126/science.1108296

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of March 24, 2016):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
/content/309/5742/1850.full.html

Supporting Online Material can be found at:
/content/suppl/2005/09/15/1108296.DC1.html

A list of selected additional articles on the Science Web sites related to this article can be found at:
/content/309/5742/1850.full.html#related

This article cites 32 articles, 16 of which can be accessed free:
/content/309/5742/1850.full.html#ref-list-1

This article has been cited by 155 article(s) on the ISI Web of Science

This article has been cited by 100 articles hosted by HighWire Press; see:
/content/309/5742/1850.full.html#related-urls

This article appears in the following subject collections:
Evolution
/cgi/collection/evolution