DOI: 10.1534/genetics.109.101428

Minimal Effect of Ectopic Gene Conversion Among Recent Duplicates in Four Mammalian Genomes

Casey L. McGrath,* Claudio Casola*,† and Matthew W. Hahn*,†,[1]

*Department of Biology and †School of Informatics, Indiana University, Bloomington, Indiana 47405

Manuscript received February 3, 2009
Accepted for publication March 19, 2009

ABSTRACT

Gene conversion between duplicated genes has been implicated in homogenization of gene families and reassortment of variation among paralogs. If conversion is common, this process could lead to errors in gene tree inference and subsequent overestimation of rates of gene duplication. After performing simulations to assess our power to detect gene conversion events, we determined rates of conversion among young, lineage-specific gene duplicates in four mammal species: human, rhesus macaque, mouse, and rat. Gene conversion rates (number of conversion events/number of gene pairs) among young duplicates range from 8.3% in macaque to 18.96% in rat, including a 5% false-positive rate. For all lineages, only 1–3% of the total amount of sequence examined was converted. There is no increase in GC content in conversion tracts compared to flanking regions of the same genes nor in conversion tracts compared to the same region in nonconverted gene-family members, suggesting that ectopic gene conversion does not significantly alter nucleotide composition in these duplicates. While the majority of gene duplicate pairs reside on different chromosomes in mammalian genomes, the majority of gene conversion events occur between duplicates on the same chromosome, even after controlling for divergence between duplicates. Among intrachromosomal duplicates, however, there is no correlation between the probability of conversion and physical distance between duplicates after controlling for divergence. Finally, we use a novel method to show that at most 5–10% of all gene trees involving young duplicates are likely to be incorrect due to gene conversion. We conclude that gene conversion has had only a small effect on mammalian genomes and gene duplicate evolution in general.

THE evolutionary processes affecting duplicated genes have been of great interest since Ohno (1970) suggested that duplicates play a major role in the evolution of new traits. Genome sequencing has revealed that gene duplication is widespread in eukaryotic genomes (Zhang 2003), and functional studies of many gene duplicates have supported Ohno's claims about its importance in evolution (reviewed in Hahn 2009). Elucidating how gene duplicates evolve over time is therefore fundamental to our understanding of organismal evolution and adaptation.

Several studies have recently begun assessing the role gene conversion plays in the evolution of duplicated genes. Gene conversion, the nonreciprocal transfer of genetic information between homologous sequences, is a type of concerted evolution thought to be responsible for the homogenization of small segments of DNA, generally smaller than several hundred base pairs (Chen *et al.* 2007). This is in contrast to unequal crossing over, which is usually implicated in homogenizing larger tracts of DNA. Gene conversion is often categorized on the basis of the location of donor and recipient sequences and can generally be classified as either allelic (conversion between alleles on sister chromatids or homologous chromosomes) or nonallelic (conversion between paralogous sequences either on the same chromosome or between chromosomes). In this article, we discuss only the effects of conversion events that occur between duplicated loci (nonallelic or "ectopic" gene conversion).

If widespread, gene conversion between paralogs could greatly impact the evolution of gene families by homogenizing variation among duplicates, thus slowing evolutionary divergence. This pattern has been demonstrated, for example, in the rDNA gene family (Arnheim *et al.* 1980) and visual pigment genes in Old World monkeys (*e.g.*, Winderickx *et al.* 1993; Zhou and Li 1996). Conversely, it has been suggested that gene conversion may generate diversity among paralogs through reassortment of genetic variation in the major histocompatibility complex gene family (*e.g.*, Weiss *et al.* 1983; Ohta 1997; Martinsohn *et al.* 1999). In addition, gene conversion between allelic sequences has been found to be biased such that G or C alleles preferentially convert A or T alleles (Galtier *et al.* 2001), resulting in

Supporting information is available online at http://www.genetics.org/cgi/content/full/genetics.109.101428/DC1.

[1]*Corresponding author:* 1001 E. 3rd St. Indiana University, Bloomington, IN 47405. E-mail: mwh@indiana.edu

more substitutions of G and C over time. This bias is a consequence of GC-biased repair of mismatches in heteroduplex intermediates during recombination, though there have so far been few studies showing that this mechanism affects ectopic gene conversion (Galtier 2003; Kudla *et al.* 2004; Benovoy *et al.* 2005).

Recent studies have begun to assess genomewide rates of conversion between duplicates, attempting to address whether gene family evolution is influenced largely by conversion or by other processes (*e.g.*, Nei and Rooney 2005). Most of these studies indicate that gene conversion may not be so extensive as to have significant effects on gene family evolution. Using a statistical method for inferring conversion events based on the distribution of differences between duplicates (*i.e.*, the software package GENECONV; Sawyer 1989), Drouin (2002) found a genomewide rate of gene conversion (number of conversion events/number of gene pairs) of 7.8% among gene families with more than two members in the yeast *Saccharomyces cerevisiae*. The same method found a conversion rate of 0.88% in humans (Benovoy and Drouin 2009). The rate of gene conversion detected in the *Caenorhabditis elegans* genome using a similar method was only 2% (Semple and Wolfe 1999). Using more limited "quartet" methods—which require two related paralogs in each of two species—Wang *et al.* (2007b) found that ~8% of *Oryza sativa japonica* paralogs on chromosomes 11 and 12 have been affected by gene conversion since the split with *O. sativa indica*. A similar study in humans (Jackson *et al.* 2005) also using a quartet method estimated a conversion rate of 5% among a subset of gene families. Finally, using a composite method that includes both quartet-based and GENECONV analyses, Ezawa *et al.* (2006) detected evidence for conversion in 18% of mouse and rat gene families (quartets) and 13% of mouse gene pairs.

The only study to find extremely high rates of gene conversion compared 68 pairs of duplicates in *S. cerevisiae*, using yet another method intended to detect conversion indirectly (Gao and Innan 2004). This study found that 81% of paralogs in yeast have been recently converted. If gene conversion rates are this high, methods that estimate the rate of gene duplication based on the number of highly similar pairs of paralogs in a genome (*e.g.*, Lynch and Conery 2000) will badly overestimate this rate (Lynch and Conery 2000; Gao and Innan 2004). This is because even ancient paralogs will appear to be recently duplicated if conversion has homogenized their sequences.

Experimental evidence has demonstrated that even slight increases in divergence between homologous sequences can greatly reduce the frequency of conversion (Lukacsovich and Waldman 1999). In this study, therefore, we focus on patterns of gene conversion among young, lineage-specific duplicates. While studies of gene conversion in human (Jackson *et al.* 2005; Benovoy and Drouin 2009) or mouse and rat (Ezawa *et al.* 2006) have been performed previously, these studies used either more limited methods or data sets that included much older, more divergent paralogs. As young, lineage-specific paralogs are the most likely to undergo conversion, focusing our study on these duplicates will not only provide an upper bound on rates and effects of conversion genomewide but will also give us more power to detect patterns of conversion. Here, we estimate independent rates of ectopic gene conversion among young duplicates in four mammalian lineages (human, macaque, mouse, and rat) using the method implemented in GENECONV (Sawyer 1989). This method does not require multiple coparalogs in multiple species and can therefore be used to study gene conversion genomewide. To ensure the accuracy of our results, we also use simulations to determine the power of GENECONV to detect gene conversion events within our data and to determine the false-positive rate. Finally, we use a novel method to show that at most 5–10% of all gene trees involving young duplicates are likely to be incorrect due to gene conversion among paralogs and, therefore, that estimates of gene duplication are not greatly affected by conversion.

METHODS

Simulation of gene conversion: While the power and false-positive rate of GENECONV has been tested in other studies (Posada and Crandall 2001; Posada 2002), the simulated and empirical data sets used were significantly different from those used in this study (*e.g.*, no alignments of only two sequences were included). We therefore determined GENECONV's power and false-positive rate among simulated sequences that more accurately resemble our data. Simulated sequences were generated in PAML using the program Evolver (Yang 2007). Each data set consisted of 1000 duplicates of two sequences representing a coding region of 1500 nucleotides. Duplicates were built under a pattern of two site classes: $d_N/d_S = 0$ and $d_N/d_S = 1$, 0.5 frequency each (d_N is the number of nonsynonymous substitutions per nonsynonymous site; d_S is the number of synonymous substitutions per synonymous site). Divergence (d_S) was fixed at 0.01, 0.02, 0.05, 0.075, 0.1, and 0.18 in different data sets. These divergences were chosen to be representative of those found in our data set. Note that while likelihood estimates of d_S correct for multiple substitutions, the correction is negligible at such low divergences, making these values of d_S approximately equal to the true proportion of synonymous substitutions. When a third sequence was added to the alignment, its divergence from each of the duplicates was twice the divergence between the duplicates. Converted tracts of 45, 90, 150, 252, 402, and 501 bp were then transferred from donor to recipient sequences at random. The conversion tract lengths for the simulations were chosen on the basis of the tract lengths observed in

our data and in other studies (*e.g.*, Semple and Wolfe 1999; Drouin 2002). Codon frequency was uniform (1/61) and transition/transversion rate ratio was fixed at $\kappa = 2$. No rate variation among sites was used, though the effect of such variation—if it affects both synonymous and nonsynonymous mutation rates—will be to inflate the false-positive rate.

Detecting conversion using GENECONV: GENECONV v.1.81 (http://www.math.wustl.edu/~sawyer/geneconv) (Sawyer 1989) was used to identify gene conversion events. Significance is determined based on 10,000 permuted data sets. GENECONV determines both global and pairwise *P*-values, the former corrected for the number of sequences in the alignment. Because we sought to compare gene families of various sizes, we used pairwise *P*-values to determine significance comparably across families. Calculating conversion rates with pairwise *P*-values (number of pairs with significant pairwise *P*-values/number of total pairs analyzed) indicates the percentage of all gene pairs with evidence for conversion. GENECONV was run using all default settings except for the addition of the option to display pairwise *P*-values (--ListPair) and the option to include monomorphic sites in the calculation when there were only two sequences in an alignment (--Includemonosites). This last option removes controls for constant sites but is necessary for analyzing an alignment with only two sequences. Significant "Pairwise Inner" fragments were considered gene conversion events. No mismatches were allowed in conversion tracts. Only duplicate pairs with at least three differences between the two sequences were considered for analysis. Analysis of average conversion tract lengths and the distribution of tract lengths included only conversion events that do not cross intron/exon boundaries or either end of the gene coding sequence, as our study does not determine to what extent the conversion tracts extend into introns or flanking sequences. All conversion events, however, were used for calculation of the total proportion of sequence converted. Subsequent analyses (position of tract in gene; GC content of converted *vs.* nonconverted pairs and conversion tracts *vs.* flanking regions; divergence of flanking regions of converted pairs *vs.* nonconverted pairs; correlation between probability of conversion and meiotic recombination rate) were performed using in-house perl scripts.

Alignment and analysis of mammalian gene duplicates: We used Ensembl v41 gene models for human, macaque, mouse, and rat. Construction of the gene trees for each gene family and inference of duplications from gene trees are described in Hahn *et al.* (2007). Briefly, 9920 gene trees were constructed from protein alignments (including homologs from an outgroup, the dog genome), followed by gene-tree/species-tree reconciliation conducted using NOTUNG (Chen *et al.* 2000). Duplication events specific to each lineage (*i.e.*, in mouse after the split with rat, in rat after the split with mouse, in human after the split with macaque, and in macaque after the split with human) were identified for each tree. Following identification of duplication events, cDNA sequences of lineage-specific paralogs were aligned by first aligning the protein sequences with ClustalW and then threading the nucleotide sequences through the protein alignments. Families containing transposable elements mistakenly annotated as genes were filtered out.

Since duplication events can incorrectly appear to be lineage specific when a copy is lost in an outgroup, we further filtered the duplicates on the basis of branch lengths for our analysis of conversion. We required the distance (d_S) between any two paralogs to be less than twice the distance since the speciation event separating sister lineages (*i.e.*, human–macaque and mouse–rat). This requirement simply identified and removed those duplicates that only appeared to be lineage specific artifactually and that are, in reality, more ancient duplicates. The average d_S values for each of the four lineages were taken from the genomic average of 9448 one-to-one orthologs (Wang *et al.* 2007a): human, $d_S = 0.032$; macaque, $d_S = 0.038$; mouse, $d_S = 0.095$; rat, $d_S = 0.095$. For example, this requirement means that for two paralogs to be considered lineage specific along the human branch, their divergence must be less than $(2) \times (0.032) = 0.064$. Nucleotides present in only one gene in an alignment and the corresponding gaps in all other genes were removed before analysis with GENECONV. Gaps aligned with sequence present in at least two genes, however, were maintained.

Gene tree *vs.* CAFE analysis: To compare the number of lineage-specific duplications inferred by gene tree analyses and copy number analyses, we considered the 9920 gene families used above. For each of these families we counted the number of lineage-specific duplicates inferred from the gene tree along the branch leading to each of human, macaque, mouse, and rat using NOTUNG (Chen *et al.* 2000). We compared these counts for each family to the number of lineage-specific duplicates inferred from the number of copies in each lineage using CAFE (Hahn *et al.* 2005; De Bie *et al.* 2006). The number of families along each lineage with a greater number of duplicates inferred by the gene tree method was divided by the total number of families with two or more genes in that lineage, resulting in the proportion of trees possibly affected by gene conversion (see results).

RESULTS

Assessing false-positive and false-negative rates: Among simulated sequences representative of our data set, we determined that GENECONV has higher statistical power to detect recent gene conversion when the divergence between the duplicates is higher and when the conversion tract is longer (supporting information, Table S1). At the highest tested divergence, 0.18

substitutions per site, GENECONV detected only 21.6% of conversions when the tract was 45 bp but detected all conversions when the tract was at least 90 bp (in the 1500-bp sequence). At the lowest divergence, 0.01 substitutions per site, however, GENECONV only detected 37.1% of conversion events at even the largest tract length, 501 bp. These simulations indicate that GENECONV is able to detect almost all conversion events that are $> \sim$200 bp when duplicates are at least as divergent as $\sim$0.05 substitutions per site. Addition of a third sequence to the alignment (with no additional conversion event simulated) had no effect on the power of GENECONV to detect conversion between the original two sequences.

We also performed simulations to determine GENECONV's false-positive rate under the default conditions (three or more sequences) and when including "monomorphic" sites (two sequences). It has been suggested previously that the false-positive rate may be particularly high when only two sequences are present in an alignment (Drouin 2002; Mondragon-Palomino and Gaut 2005). In our simulations of alignments with only two sequences, the false-positive rates (number of conversion events detected/number of gene pairs) for the divergences of 0.016, 0.05, and 0.1 were 5.7%, 4.9%, and 4.4%, respectively. The average conversion tract length detected was negatively correlated with the divergence of the duplicates. The overall proportion of total sequence implicated in a (false) conversion event was therefore highest (0.45%) for the lowest divergence (0.016). When a third sequence was added to the alignment and GENECONV was run under default conditions, the false-positive rate was still <5%: at a divergence of 0.05, the fraction of false positives per pairwise comparison was 2.7% with three sequences, compared to 4.9% with two sequences.

These simulations indicate that GENECONV has reasonable power to detect true conversion events in our data, though comparison of very young duplicates is undoubtedly underpowered. In addition, we find no evidence that the false-positive rate is aberrantly high when only two sequences are present in an alignment. The rate of false positives of GENECONV appears to be what is expected when a significance threshold of $P < 0.05$ is used.

Conversion rates and patterns in mammalian genomes: To obtain independent estimates of gene conversion in each of the four species, we compared only lineage-specific paralogs within each lineage (methods). Higher divergence between paralogs leads to less frequent gene conversion as well as shorter conversion tracts (Lukacsovich and Waldman 1999); we are therefore confident that an analysis focused on less divergent paralogs captures the majority of gene conversion events occurring in these genomes. It also provides an upper bound on the rate and effects of gene conversion genomewide. Our final data consisted of 261 alignments of lineage-specific duplicates (549 pairwise comparisons) in humans, 206 alignments (363 pairs) in macaque, 629 alignments (1913 pairs) in mouse, and 603 alignments (1171 pairs) in rat.

Among all lineage-specific gene pairs analyzed, we found the rate of gene conversion (number of conversion events/number of gene pairs) to be 12.57% in human, 8.26% in macaque, 14.58% in mouse, and 18.96% in rat at $P < 0.05$ (see Table S2 for a list of predicted conversion events between gene pairs). The actual rates, however, are likely even lower as these values include a false-positive rate of 5% at this *P*-value. The distribution of conversion tract lengths illustrates that most conversion events extend $< \sim$500 bp (Figure 1); it also reflects the poor power of GENECONV to detect conversions $< \sim$100 bp in length. The average length of the conversion tracts is 210 bp in human, 229 bp in macaque, 190 bp in mouse, and 172 bp in rat. Because the method used to detect conversion looks for long stretches of identity that must be bounded on either side by a difference between the paralogs, the conversion tract lengths detected by GENECONV are maximum estimates of the size of the tract. The positions of conversion tracts within genes were uniformly distributed, with the start of most tracts in the first 25% of the gene sequence. The overall proportion of total sequence that has been converted is 2.16% in human, 1.76% in macaque, 2.57% in mouse, and 2.15% in rat. This indicates that gene conversion among duplicates is likely to affect a mere 1–3% of total sequence within recently duplicated mammalian genes (and even smaller amounts among older duplicates).

Biased gene conversion between allelic sequences has been shown to lead to an increase in the GC content of conversion tracts (Galtier and Duret 2007). There have been few studies, however, to investigate the effects of nonallelic gene conversion on GC content (Galtier 2003; Kudla *et al.* 2004; Benovoy *et al.* 2005). Among alignments in our analysis with only two sequences, the average GC content within conversion tracts was not significantly greater than the average GC content of nonconverted flanking sequence and was actually slightly lower in some lineages: 52.0% *vs.* 53.8% in human, 50.6% *vs.* 49.6% in macaque, 46.8% *vs.* 47.3% in mouse, and 47.5% *vs.* 47.3% in rat (paired *t*-test, $P > 0.05$ for all). This comparison could potentially miss an increase in GC content in converted tracts, however, as it compares different regions of genes (conversion tracts *vs.* flanking sequences). We therefore also compared the GC content of a conversion tract with the same gene segment from nonconverted paralogs when there were more than two sequences in an alignment. Again, there was no significant trend toward higher GC content in converted sequences *vs.* nonconverted sequences: 54.0% *vs.* 52.5% in human, 49.4% *vs.* 49.2% in macaque, 45.9% *vs.* 45.7% in mouse, and 47.4% *vs.* 47.4% in rat (paired *t*-test, $P > 0.05$ for all).

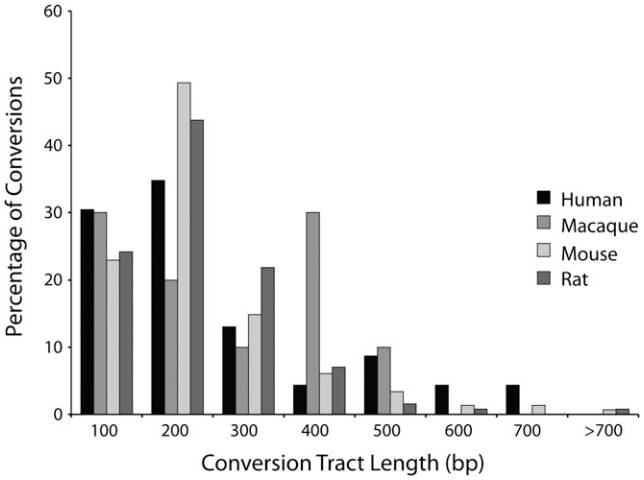

FIGURE 1.—Lengths of conversion events detected in the human, macaque, mouse, and rat genomes. Bin values on the *x*-axis represent the maximum tract length included in each bin. Only conversion events that do not cross intron/exon boundaries are included. The underrepresentation of conversion events <100 bp likely reflects the low power of GENECONV to detect short conversions.

While gene conversion is known to occur more frequently between more similar duplicates (LUKACSOVICH and WALDMAN 1999), the distribution of the divergences of nonconverted gene pairs compared to those of converted pairs (excluding the conversion tract) does not clearly demonstrate such a pattern (Figure S1). One reason for the apparent lack of the expected pattern is GENECONV's poor power to detect conversion events when divergence between duplicates is very low. It is also of course true that conversion between highly similar genes will often have no homogenizing effect, as there may be no nucleotide differences in the conversion tracts to begin with.

Many recent gene duplication events result in paralogs that reside on different chromosomes (Figure 2), though there is evidence for an expansion in intrachromosomal duplications along the human lineage (SHE *et al.* 2006). The majority of duplicated genes that have undergone gene conversion are located on the same chromosome in all four species (Figure 2). The excess of intrachromosomal conversion relative to the number of intrachromosomal duplicates is statistically significant in every genome (Fisher's exact test, all $P < 0.05$). In addition, intrachromosomal conversion occurs at a disproportionately higher frequency between duplicates that are close together (<50 kb apart), and there is a significantly negative correlation ($P < 0.05$) between rates of conversion and intrachromosomal distance in human, mouse, and rat (Figure S2). However, neighboring paralogs are more likely to be recently duplicated and thus less divergent (KATJU and LYNCH 2003), and it is possible that interchromosomal duplicates may on average be more divergent, confounding the factors of chromosomal location, physical distance, and divergence. Linear regressions demonstrate that while chromosomal location (intrachromosomal *vs.* interchromosomal) is still a significant predictor of conversion after correcting for divergence ($P < 0.01$ for all genomes), physical distance between intrachromosomal duplicates is not a significant predictor of conversion once divergence is accounted for ($P > 0.1$ for all genomes).

For intrachromosomal duplicates, we also hypothesized that gene conversion might be influenced by the orientation of duplicates relative to each other. We therefore classified each pair of intrachromosomal duplicates as head to tail, head to head, or tail to tail. If duplicates are arranged randomly, we expect 50% in a head-to-tail orientation and 25% in each of head-to-head and tail-to-tail orientations. Among all mammalian duplicates we found a significant excess of head-to-tail arrangements for intrachromosomal paralogs within 50 kb of each other (Fisher's exact test, all $P < 0.05$; Figure S3), though there was only an excess for all intrachromosomal paralogs in rat and mouse. Contrary to our expectations, however, there was no excess of gene conversion associated with any specific orientation of paralogs in any of the four genomes (Figure S4). These patterns of correlation between conversion and chromosomal location, distance between paralogs, and gene orientation largely agree with those found previously for conversion events between older paralogs in mouse (EZAWA *et al.* 2006) and human (BENOVOY and DROUIN 2009), though these studies did not consider the confounding effects of divergence and physical distance.

While meiotic recombination is responsible for both allelic gene conversion and crossovers, the relationship between meiotic recombination rate and ectopic gene conversion is unclear. We therefore looked for a relationship between human recombination rates based on the deCODE map (KONG *et al.* 2002) and the frequency of gene conversion among human paralogs.

The proportion of converted *vs.* nonconverted duplicated pairs shows no correlation with recombination rates for pairs <1 or <5 Mb apart ($r = -0.007$ and $r = 0.039$). Similar results are obtained using all duplicated pairs and averaging the recombination rates of the two genes ($r = 0.024$). This is contrary to the results of BENOVOY and DROUIN (2009), who found a significant positive correlation between meiotic recombination rate and frequency of gene conversion in humans. This difference in results could be due to a difference in methods or recombination rates used.

Effect of conversion on gene trees and estimates of duplication rates: LYNCH and CONERY (2000) proposed a method to estimate rates of gene duplication by counting the number of very young duplicates (*i.e.*, $d_S < 0.01$) in a genome and dividing by the total number of genes. This method therefore assumes that low divergence between duplicates reflects recent duplication events and is not due to gene conversion

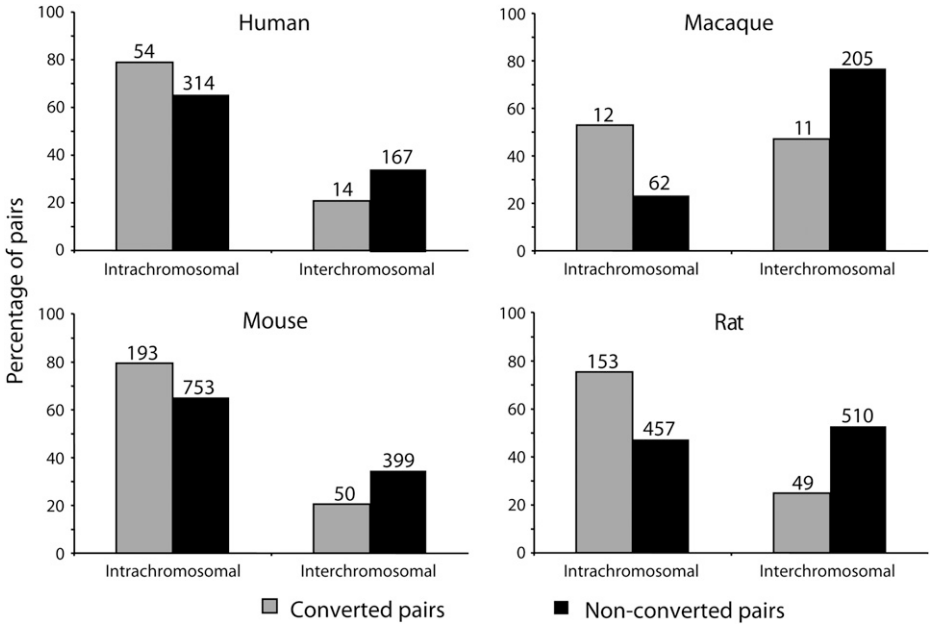

FIGURE 2.—Overrepresentation of intrachromosomal gene pairs among pairs showing evidence of gene conversion. The percentage of converted pairs that are present on the same chromosome (intrachromosomal) and the percentage present on different chromosomes (interchromosomal) are shaded; percentages of nonconverted pairs are solid. The number of pairs in each category is given above the bars. In all four genomes, conversion appears to occur preferentially between intrachromosomal duplicates.

among paralogs (LYNCH and CONERY 2000). A study of gene conversion in yeast has cast doubt on the results of this method by showing extremely high rates of conversion in this species, implying that actual rates of gene duplication are much lower than previously thought (GAO and INNAN 2004). However, the yeast study only indirectly inferred gene conversion and was limited to 68 pairs of duplicates; its results were also in conflict with previous studies of the rate of gene conversion in yeast that used GENECONV (DROUIN 2002).

We have recently introduced a method for estimating rates of gene duplication and loss that only relies on changes in the number of paralogous genes among species and not on sequence identity (HAHN *et al.* 2005). This method will not overestimate rates of gene duplication due to gene conversion, as the number of duplicates in a genome does not change because of conversion (HAHN *et al.* 2007). For example, if human and macaque each had two duplicate copies of a gene and other mammals had only one copy, this method (as implemented in the program CAFE) (DE BIE *et al.* 2006) would infer a single duplication in the human–macaque ancestor, regardless of the similarity between the human paralogs. We can therefore use this method to confirm that gene conversion among young duplicates in mammalian genomes is not leading to widespread error in gene trees and duplication estimates. To do this we compared the number of lineage-specific duplications inferred from gene trees—constructed from the protein sequences of the genes—to the number inferred by CAFE in all gene families with a size of at least two (METHODS). If gene conversion has recently homogenized pairs of duplicates, gene tree-based methods will overestimate the number of duplication events. This is because conversion will cause the intraspecific duplicates to be more similar to each other, leading to an estimation of two recent duplication events, one in each lineage, rather than one duplication event that preceded speciation (Figure 3).

In all four lineages, the percentage of gene families where the number of duplications inferred by gene trees was greater than the number inferred by CAFE (*i.e.*, families where gene conversion *may* be affecting the tree) was very low: 227/3378 (6.7%) in human, 276/3560 (7.8%) in macaque, 301/3505 (8.6%) in mouse, and 328/3388 (9.7%) in rat. We should not assume, however, that all of the cases where the gene tree has inferred more duplications are due to gene conversion (*i.e.*, the CAFE estimate is correct while the gene tree estimate is incorrect), as some are undoubtedly due to true parallel duplications or multiple duplications coupled with gene loss (*i.e.*, the gene tree estimate is correct while the CAFE estimate is incorrect). To provide a rough estimate of the rate of parallel duplication *vs.* gene conversion, we examined the seven cases where a gene family had exactly two gene copies in both human and macaque, independent duplications had been implied by the gene tree, and where all four genes have been assigned to a chromosomal location. Of the seven cases, only three show both duplicates maintained on homologous chromosomes between species. The remaining four families have one ortholog on homologous chromosomes between human and macaque (likely the single gene present in the most recent common ancestor) while the additional copies are on nonhomologous chromosomes between species. While gene conversion followed (or preceded) by translocation cannot be ruled out in these four cases, we believe it is more likely that they represent parallel duplications in the two lineages. It is therefore likely that the percentage of families where gene conversion might lead to an overestimation of duplications is even <5–10%, perhaps less than half this value.

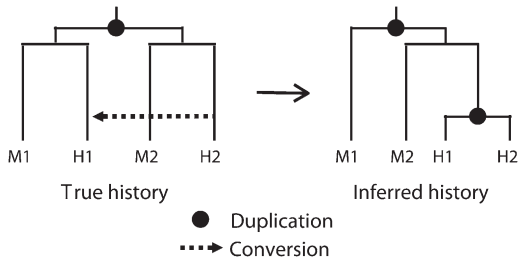

Figure 3.—Duplication followed by gene conversion can lead to an overestimation of the number of duplications in a gene family. The true history of a gene family is shown on the left, with a single duplication event in the human–macaque ancestor followed by speciation giving rise to two macaque paralogs (M1 and M2) and two human paralogs (H1 and H2). If the human paralogs subsequently undergo gene conversion such that H2 converts H1, phylogenetic analysis of the gene family will yield the inferred tree illustrated on the right, where H1 and H2 are highly similar sequences. This tree would imply two duplication events as shown.

DISCUSSION

This study shows that the overall impact of conversion among young gene duplicates in mammalian genomes is likely to be minimal. This conclusion is consistent with that of Nei and Rooney (2005), who suggested that the contribution of gene conversion to gene family evolution is minor in the long term. We found rates of conversion between recently duplicated genes in human, macaque, mouse, and rat to be low: <5–15% of duplicate pairs showed evidence of conversion (when the 5% false-positive rate is considered). We also found no increase in GC content of converted sequences, indicating that biased gene conversion is not a significant driver of nucleotide content evolution in gene duplicates in these genomes. On the whole, only 3–6% of the total sequence analyzed was involved in a conversion event, meaning only 1–3% of sequence was actually converted (a recipient of gene conversion). These numbers are comparable to the 2–13% conversion frequencies observed previously for the yeast, nematode, mouse, and rice lineages (Semple and Wolfe 1999; Drouin 2002; Ezawa *et al.* 2006; Wang *et al.* 2007b), indicating that gene conversion is likely to be far from ubiquitous in most genomes. In particular, our estimate for the percentage of gene pairs undergoing conversion in mouse, 14.56%, is highly consistent with the percentage estimated by Ezawa *et al.* (2006), 13%. This is striking when we consider the different methodologies and data sets used—our study was limited to lineage-specific duplicates while the Ezawa *et al.* data excluded lineage-specific duplicates and focused on duplicates that arose in the mouse–rat ancestor.

Our estimate for conversion rate among young duplicates in human (12.57%), on the other hand, is much larger than the 0.88% frequency recently estimated by Benovoy and Drouin (2009). This is to be expected, however, as Benovoy and Drouin included duplicate pairs with at least 60% protein identity over at least 50% of the sequence. Inclusion of more divergent duplicates should lower the observed conversion rate, as the young duplicates in our study likely undergo the highest rates of conversion of any duplicates in the genome. In addition, Benovoy and Drouin utilized GENECONV's global *P*-values in their calculation of conversion rate, which makes direct comparison with our values difficult but which is also likely to decrease the observed conversion rate.

While we believe our study provides an important estimate of the upper bound of the frequency and effects of conversion among duplicates in these four mammalian genomes, there are some limitations to our analysis. Our method is underpowered for detecting conversion events between duplicates $< \sim 5\%$ divergent, though such conversion events are likely to have the smallest impact on the genome as they will lead to few substitutions in the converted copies. However, this lack of power at very low divergences is potentially responsible for the slightly lower conversion rates in human and macaque compared to mouse and rat, as there are more lineage-specific duplicates with higher divergence in the rodent lineages (methods). In addition, our estimates of tract length (and therefore total sequence involved in conversions) are likely to be somewhat overestimated, as conversion tracts identified by GENECONV must necessarily be bounded by differences between duplicates; in actuality, however, the conversion tract may have been shorter. Because we did not allow mismatches within gene conversion tracts detected by GENECONV, our analysis may miss older events where one or more mutations have occurred after conversion. This would cause our numbers to be underestimates of the actual conversion rates in these genomes. However, because our study is focused on conversion events between recent duplicates, we believe this is not likely to be a significant source of error. Finally, GENECONV does not take into account purifying selection that may be acting differentially on different gene segments. If selection were maintaining identical sequences between duplicates in one part of the gene but relaxed selection were allowing mutations in another region, this could lead to the appearance of gene conversion. However, we believe this type of false positive is unlikely in our data, as our analysis included not only nonsynonymous sites but synonymous sites as well. Because the large majority of synonymous mutations are believed to be silent, purifying selection should generally not affect mutations at synonymous sites. Situations where an identical stretch of coding sequence between duplicates has been maintained by purifying selection at both nonsynonymous and synonymous sites must therefore be very rare, if they occur at all, in these data.

Perhaps most importantly, our comparison of the number of duplications inferred by gene trees compared to the number inferred by copy number demonstrates

that gene conversion does not lead to widespread gene tree inconsistencies and large overestimates of the gene duplication rate. Even if we have missed conversion events between young duplicates using GENECONV, or conversion has occurred across the full length of two paralogs, the comparison of gene trees and copy number indicates that the overall effects of gene conversion must be minimal. Simply the fact that copy numbers do change at such high rates—even in yeast (Hahn *et al.* 2005)—supports the original contention of Lynch and Conery (2000) that rates of gene duplication are high.

While our results emphasize the minor impact of gene conversion genomewide, other studies have highlighted the important role gene conversion can play in duplicate gene evolution in certain gene families (*e.g.*, Hoffmann *et al.* 2008). Those studies, in the context of our results, imply that variation in the frequency and selective advantage of conversion among gene families may be high. Despite these rare cases, however, when all gene families with young duplicate genes are considered, gene conversion clearly does not play a major role across the genome.

We thank Mira Han for sharing her data. This work was supported by a grant from the National Science Foundation (DBI-0543586) to M.W.H.

LITERATURE CITED

Arnheim, N., M. Krystal, R. Schmickel, G. Wilson, O. Ryder *et al.*, 1980 Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc. Natl. Acad. Sci. USA **77:** 7323–7327.

Benovoy, D., and G. Drouin, 2009 Ectopic gene conversions in the human genome. Genomics **93:** 27–32.

Benovoy, D., R. T. Morris, A. Morin and G. Drouin, 2005 Ectopic gene conversions increase the G+C content of duplicated yeast and *Arabidopsis* genes. Mol. Biol. Evol. **22:** 1865–1868.

Chen, J.-M., D. N. Cooper, N. Chuzhanova, C. Ferec and G. P. Patrinos, 2007 Gene conversion: mechanisms, evolution and human disease. Nat. Rev. Genet. **8:** 762–775.

Chen, K., D. Durand and M. Farach-Colton, 2000 NOTUNG: a program for dating gene duplications and optimizing gene family trees. J. Comput. Biol. **7:** 429–447.

De Bie, T., N. Cristianini, J. P. Demuth and M. W. Hahn, 2006 CAFE: a computational tool for the study of gene family evolution. Bioinformatics **22:** 1269–1271.

Drouin, G., 2002 Characterization of the gene conversions between the multigene family members of the yeast genome. J. Mol. Evol. **55:** 14–23.

Ezawa, K., S. Oota and N. Saitou, 2006 Genome-wide search of gene conversions in duplicated genes of mouse and rat. Mol. Biol. Evol. **23:** 927–940.

Galtier, N., 2003 Gene conversion drives GC content evolution in mammalian histones. Trends Genet. **19:** 65–68.

Galtier, N., and L. Duret, 2007 Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution. Trends Genet. **23:** 273–277.

Galtier, N., G. Piganeau, D. Mouchiroud and L. Duret, 2001 GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. Genetics **159:** 907–911.

Gao, L., and H. Innan, 2004 Very low gene duplication rate in the yeast genome. Science **306:** 1367–1370.

Hahn, M. W., 2009 Distinguishing among evolutionary models for the maintenance of gene duplicates. J. Hered. (in press).

Hahn, M. W., T. De Bie, J. E. Stajich, C. Nguyen and N. Cristianini, 2005 Estimating the tempo and mode of gene family evolution from comparative genomic data. Genome Res. **15:** 1153–1160.

Hahn, M. W., J. P. Demuth and S. G. Han, 2007 Accelerated rate of gene gain and loss in primates. Genetics **177:** 1941–1949.

Hoffmann, F. G., J. C. Opazo and J. F. Storz, 2008 Rapid rates of lineage-specific gene duplication and deletion in the alpha-globin gene family. Mol. Biol. Evol. **25:** 591–602.

Jackson, M. S., K. Oliver, J. Loveland, S. Humphray, I. Dunham *et al.*, 2005 Evidence for widespread reticulate evolution within human duplicons. Am. J. Hum. Genet. **77:** 824–840.

Katju, V., and M. Lynch, 2003 The structure and early evolution of recently arisen gene duplicates in the *Caenorhabditis elegans* genome. Genetics **165:** 1793–1803.

Kong, A., D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson *et al.*, 2002 A high-resolution recombination map of the human genome. Nat. Genet. **31:** 241–247.

Kudla, G., A. Helwak and L. Lipinski, 2004 Gene conversion and GC-content evolution in mammalian Hsp70. Mol. Biol. Evol. **21:** 1438–1444.

Lukacsovich, T., and A. S. Waldman, 1999 Suppression of intrachromosomal gene conversion in mammalian cells by small degrees of sequence divergence. Genetics **151:** 1559–1568.

Lynch, M., and J. S. Conery, 2000 The evolutionary fate and consequences of duplicate genes. Science **290:** 1151–1155.

Martinsohn, J. T., A. B. Sousa, L. A. Guethlein and J. C. Howard, 1999 The gene conversion hypothesis of MHC evolution: a review. Immunogenetics **50:** 168–200.

Mondragon-Palomino, M., and B. S. Gaut, 2005 Gene conversion and the evolution of three leucine-rich repeat gene families in *Arabidopsis thaliana*. Mol. Biol. Evol. **22:** 2444–2456.

Nei, M., and A. P. Rooney, 2005 Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. **39:** 121–152.

Ohno, S., 1970 *Evolution by Gene Duplication*. Springer-Verlag, Berlin.

Ohta, T., 1997 Role of gene conversion in generating polymorphisms at major histocompatibility complex loci. Hereditas **127:** 97–103.

Posada, D., 2002 Evaluation of methods for detecting recombination from DNA sequences: empirical data. Mol. Biol. Evol. **19:** 708–717.

Posada, D., and K. A. Crandall, 2001 Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. USA **98:** 13757–13762.

Sawyer, S. A., 1989 Statistical tests for detecting gene conversion. Mol. Biol. Evol. **6:** 526–538.

Semple, C., and K. H. Wolfe, 1999 Gene duplication and gene conversion in the *Caenorhabditis elegans* genome. J. Mol. Evol. **48:** 555–564.

She, X., G. Liu, M. Ventura, S. Zhao, D. Misceo *et al.*, 2006 A preliminary comparative analysis of primate segmental duplications shows elevated substitution rates and a great-ape expansion of intrachromosomal duplications. Genome Res. **16:** 576–583.

Wang, H. Y., H. C. Chien, N. Osada, K. Hashimoto, S. Sugano *et al.*, 2007a Rate of evolution in brain-expressed genes in humans and other primates. PLoS Biol. **5:** 335–342.

Wang, X., H. Tang, J. E. Bowers, F. A. Feltus and A. H. Paterson, 2007b Extensive converted evolution of rice paralogs and the road to regaining independence. Genetics **177:** 1753–1763.

Weiss, E. H., A. Mellor, L. Golden, K. Fahrner, E. Simpson *et al.*, 1983 The structure of a mutant H-2 gene suggests that the generation of polymorphism in H-2 genes may occur by gene conversion-like events. Nature **301:** 671–674.

Winderickx, J., L. Battisti, Y. Hibiya, A. G. Motulsky and S. S. Deeb, 1993 Haplotype diversity in the human red and green opsin genes: evidence for frequent sequence exchange in exon 3. Hum. Mol. Genet. **2:** 1413–1421.

Yang, Z. H., 2007 PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. **24:** 1586–1591.

Zhang, J., 2003 Evolution by gene duplication: an update. Trends Ecol. Evol. **18:** 292–298.

Zhou, Y.-H., and W.-H. Li, 1996 Gene conversion and natural selection in the evolution of X-linked color vision genes in higher primates. Mol. Biol. Evol. **13:** 780–783.

Communicating editor: Z. H. Humayun

GENETICS

Supporting Information
http://www.genetics.org/cgi/content/full/genetics.109.101428/DC1

Minimal Effect of Ectopic Gene Conversion Among Recent Duplicates in Four Mammalian Genomes

Casey L. McGrath, Claudio Casola and Matthew W. Hahn

DOI: 10.1534/genetics.109.101428

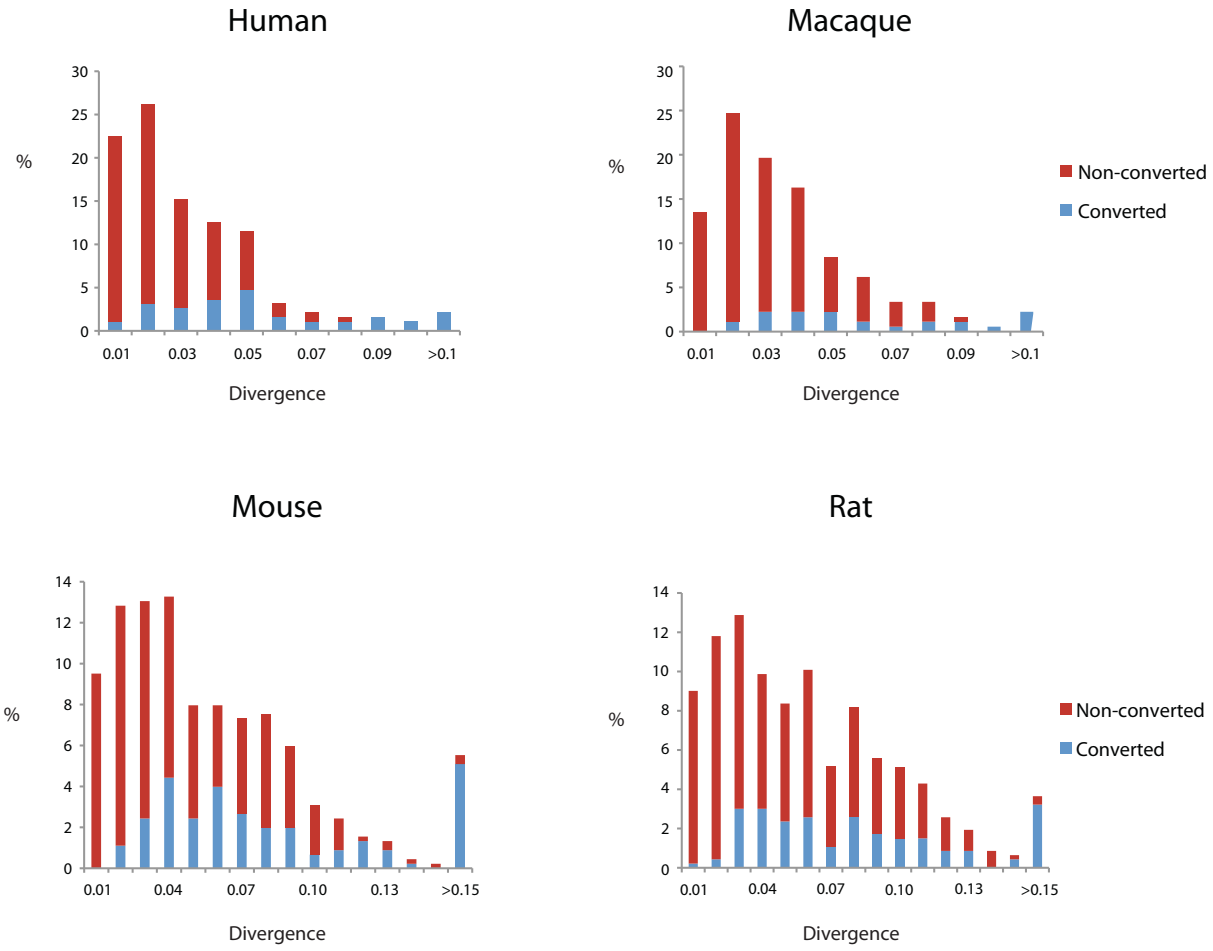

FIGURE S1.—Histogram of the divergence of paralogs from alignments with two genes, showing those that have undergone conversion and those experiencing no conversion. Conversion tracts were excluded when calculating the divergence of converted pairs. The low power of GENECONV to detect conversion when gene pairs are highly similar is demonstrated by the underrepresentation of converted pairs with low divergence.

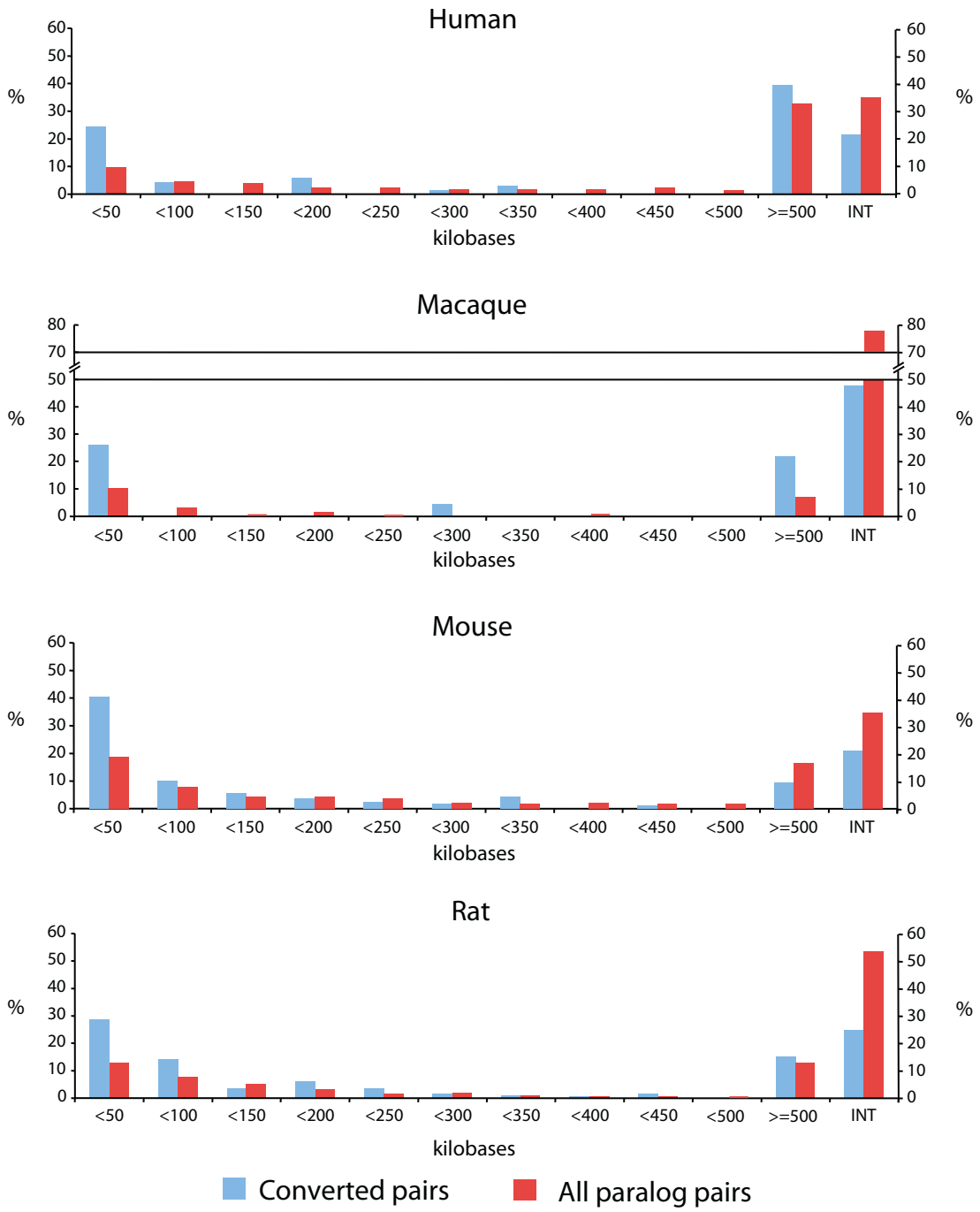

FIGURE S2.—Histogram of the intrachromosomal distance (kb) between genes in duplicate pairs that have undergone conversion compared to all gene pairs. Duplicates that are close together (<50 kb apart) demonstrate a higher rate of conversion. There is a significant (P<0.05) negative correlation between conversion and intrachromosomal distance for human, mouse, and rat. Interchromosomal percentages are also shown ("INT").

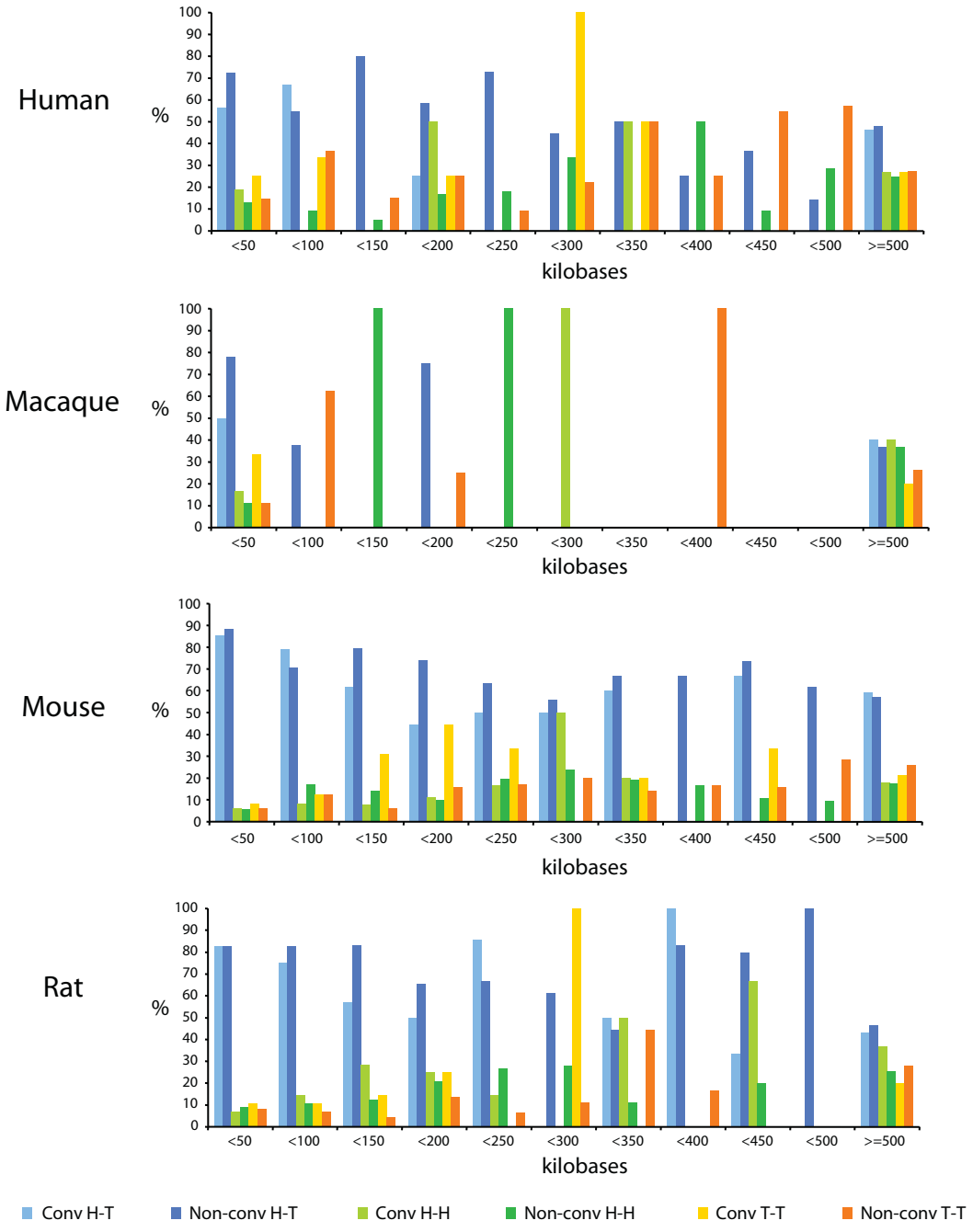

FIGURE S3.—Histogram of the respective orientation of duplicate pairs that have undergone conversion compared to all gene pairs, ordered by distance between genes in each pair. Proportion of pairs of converted genes and pairs of non-converted paralogs oriented in the three possible arrangements: "head-to-tail", "head-to-head" and "tail-to-tail". Pairs with members separated by 500 kilobases or more are grouped together in the interval ">=500". Percentages are shown on the y axis.

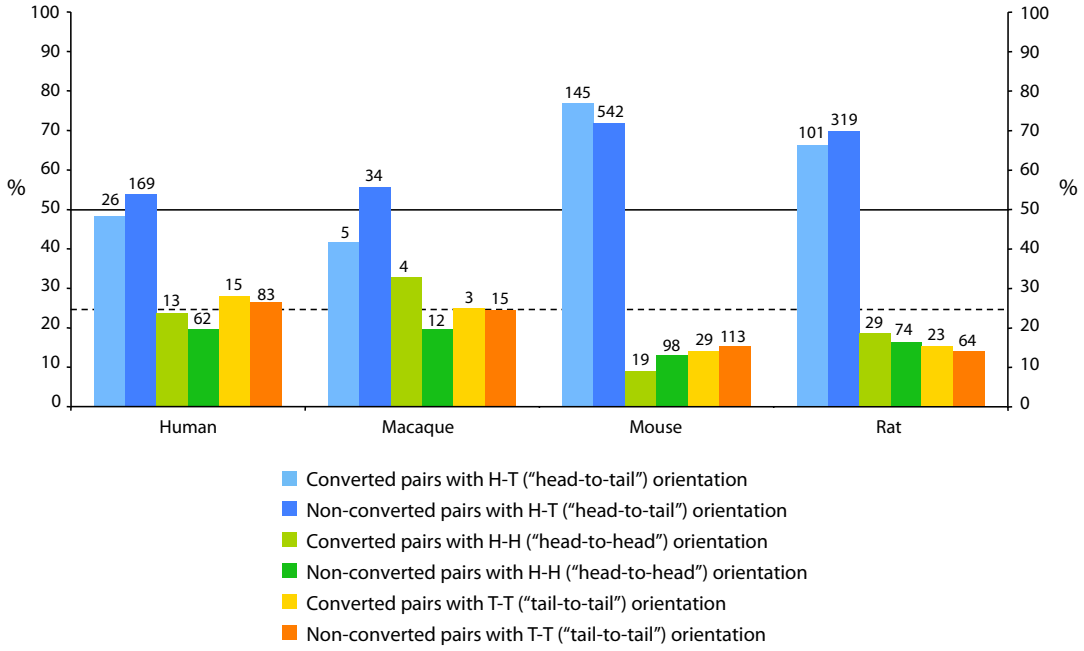

FIGURE S4.—Histogram of the respective orientation of duplicate pairs that have undergone conversion compared to all gene pairs in the four mammalian species. Proportion of pairs of converted genes and pairs of non-converted paralogs oriented in the three possible arrangements: "head-to-tail", "head-to-head" and "tail-to-tail". Percentages are shown on the y axis. The two lines show the 50 and 25 percentage thresholds. Numbers of duplicate pairs are shown above each column.

TABLE S1

Power of GENECONV to detect simulated gene conversion events of various sizes between two sequences (1500 bp each) at varying levels of divergence in 1000 simulations

		Divergence					
		0.01	0.02	0.05	0.075	0.1	0.18
Length of conversion tract (bp)	501	37.1%	85.5%	99.9%	99.9%	99.9%	100.1%
	402	27.0%	66.9%	100.1%	100.0%	100.0%	100.1%
	252	15.2%	27.9%	98.5%	100.5%	100.6%	101.0%
	150	9.6%	12.4%	40.8%	90.3%	101.3%	102.3%
	90	6.7%	7.2%	13.3%	22.7%	49.2%	103.7%
	45	5.3%	5.2%	7.5%	7.1%	7.5%	21.6%

TABLE S2

Predicted conversion events between gene pairs in (A) human, (B) macaque, (C) mouse, and (D) rat lineages, with pairwise *P*-values from GENECONV

A

Gene 1	Gene 2	*P*-value
ENSP00000308764	ENSP00000369796	0.0401
ENSP00000317447	ENSP00000363911	0.0001
ENSP00000355342	ENSP00000358387	0.0439
ENSP00000358224	ENSP00000358387	0.0439
ENSP00000358370	ENSP00000358387	0.0439
ENSP00000244519	ENSP00000366937	0.0048
ENSP00000341961	ENSP00000371940	0.0283
ENSP00000366693	ENSP00000371940	0.0314
ENSP00000358156	ENSP00000358157	<0.0001
ENSP00000294342	ENSP00000334246	0.0013
ENSP00000310860	ENSP00000329355	0.0443
ENSP00000306535	ENSP00000308080	0.0328
ENSP00000240189	ENSP00000332134	0.0001
ENSP00000330156	ENSP00000365363	0.0376
ENSP00000365328	ENSP00000365363	0.0376
ENSP00000228226	ENSP00000348170	0.0006
ENSP00000228226	ENSP00000348170	0.0446
ENSP00000348864	ENSP00000359307	<0.0001
ENSP00000349891	ENSP00000366697	0.0052
ENSP00000295450	ENSP00000295453	0.0167
ENSP00000238651	ENSP00000311224	0.0363
ENSP00000348646	ENSP00000349942	<0.0001
ENSP00000288911	ENSP00000289105	<0.0001
ENSP00000309233	ENSP00000337310	<0.0001
ENSP00000344876	ENSP00000367226	0.0210
ENSP00000261609	ENSP00000315224	0.0129
ENSP00000261609	ENSP00000320293	0.0154
ENSP00000339793	ENSP00000340787	0.0008
ENSP00000256733	ENSP00000348918	0.0040
ENSP00000327703	ENSP00000372100	0.0238
ENSP00000372020	ENSP00000372100	0.0001
ENSP00000372100	ENSP00000372101	0.0002
ENSP00000260309	ENSP00000364858	0.0028
ENSP00000272546	ENSP00000366540	<0.0001
ENSP00000284154	ENSP00000345796	0.0001
ENSP00000348915	ENSP00000366573	0.0200
ENSP00000334952	ENSP00000372866	0.0132
ENSP00000302745	ENSP00000371877	0.0123

ENSP00000371743	ENSP00000371877	0.0123
ENSP00000371802	ENSP00000371877	0.0022
ENSP00000368280	ENSP00000368282	0.0107
ENSP00000368282	ENSP00000368284	0.0372
ENSP00000355119	ENSP00000372521	0.0127
ENSP00000194530	ENSP00000354433	0.0481
ENSP00000355218	ENSP00000365117	0.0246
ENSP00000364309	ENSP00000364438	0.0286
ENSP00000344220	ENSP00000371763	0.0218
ENSP00000217933	ENSP00000328001	0.0355
ENSP00000290422	ENSP00000311682	0.0394
ENSP00000332724	ENSP00000369752	0.0236
ENSP00000370088	ENSP00000372505	0.0301
ENSP00000342609	ENSP00000363544	0.0095
ENSP00000328178	ENSP00000350575	0.0468
ENSP00000351530	ENSP00000364026	0.0004
ENSP00000184183	ENSP00000251776	0.0015
ENSP00000226798	ENSP00000278882	0.0314
ENSP00000266775	ENSP00000370757	0.0390
ENSP00000326538	ENSP00000341051	0.0161
ENSP00000321876	ENSP00000370076	0.0262
ENSP00000251152	ENSP00000333522	0.0214
ENSP00000319520	ENSP00000328223	0.0241
ENSP00000184266	ENSP00000244249	0.0464
ENSP00000215794	ENSP00000292729	<0.0001
ENSP00000281871	ENSP00000311500	0.0226
ENSP00000266604	ENSP00000351888	0.0255
ENSP00000371102	ENSP00000371227	<0.0001
ENSP00000329663	ENSP00000337144	0.0248
ENSP00000348463	ENSP00000366715	<0.0001
ENSP00000283507	ENSP00000355792	0.0223
ENSP00000349714	ENSP00000352732	0.0339

B

Gene 1	Gene 2	*P*-value
ENSMMUP00000025509	ENSMMUP00000037565	0.0006
ENSMMUP00000037561	ENSMMUP00000041384	0.0018
ENSMMUP00000021879	ENSMMUP00000038632	0.0488
ENSMMUP00000003419	ENSMMUP00000006164	0.0022
ENSMMUP00000022059	ENSMMUP00000026144	<0.0001
ENSMMUP00000037167	ENSMMUP00000037174	0.0003
ENSMMUP00000023369	ENSMMUP00000041127	0.0236
ENSMMUP00000013015	ENSMMUP00000023533	0.0332
ENSMMUP00000040908	ENSMMUP00000040912	<0.0001
ENSMMUP00000034259	ENSMMUP00000004409	<0.0001
ENSMMUP00000034259	ENSMMUP00000004409	0.0272

ENSMMUP00000034858	ENSMMUP00000008029	0.0200
ENSMMUP00000024380	ENSMMUP00000030346	0.0425
ENSMMUP00000018393	ENSMMUP00000018394	0.0010
ENSMMUP00000013687	ENSMMUP00000036701	0.0001
ENSMMUP00000022513	ENSMMUP00000025146	0.0126
ENSMMUP00000013923	ENSMMUP00000007467	0.0335
ENSMMUP00000021715	ENSMMUP00000023181	0.0115
ENSMMUP00000039002	ENSMMUP00000039003	0.0485
ENSMMUP00000032253	ENSMMUP00000008346	<0.0001
ENSMMUP00000018617	ENSMMUP00000004417	0.0415
ENSMMUP00000010473	ENSMMUP00000031421	0.0016
ENSMMUP00000025499	ENSMMUP00000025735	0.0222
ENSMMUP00000012867	ENSMMUP00000025730	0.0454
ENSMMUP00000036500	ENSMMUP00000039740	0.0008
ENSMMUP00000015284	ENSMMUP00000009090	0.0141
ENSMMUP00000021021	ENSMMUP00000032233	0.0367
ENSMMUP00000019692	ENSMMUP00000040064	0.0164
ENSMMUP00000016914	ENSMMUP00000016915	0.0008
ENSMMUP00000015569	ENSMMUP00000037121	0.0211

C

Gene 1	Gene 2	*P*-value
ENSMUSP00000025322	ENSMUSP00000047766	0.0198
ENSMUSP00000049819	ENSMUSP00000074958	0.0064
ENSMUSP00000049819	ENSMUSP00000080597	0.0021
ENSMUSP00000074958	ENSMUSP00000080597	0.0298
ENSMUSP00000040319	ENSMUSP00000095797	0.0358
ENSMUSP00000040319	ENSMUSP00000095797	0.0438
ENSMUSP00000095811	ENSMUSP00000095813	0.0021
ENSMUSP00000051280	ENSMUSP00000052396	0.0065
ENSMUSP00000060602	ENSMUSP00000088194	0.0083
ENSMUSP00000097403	ENSMUSP00000097405	0.0012
ENSMUSP00000097386	ENSMUSP00000097388	<0.0001
ENSMUSP00000099021	ENSMUSP00000099023	0.0021
ENSMUSP00000068282	ENSMUSP00000072598	0.0001
ENSMUSP00000068282	ENSMUSP00000072598	0.0001
ENSMUSP00000068282	ENSMUSP00000075255	0.0101
ENSMUSP00000072016	ENSMUSP00000073233	0.0186
ENSMUSP00000076282	ENSMUSP00000089616	0.0001
ENSMUSP00000071488	ENSMUSP00000080106	0.0170
ENSMUSP00000078739	ENSMUSP00000096707	0.0093
ENSMUSP00000072743	ENSMUSP00000079096	0.0343
ENSMUSP00000001088	ENSMUSP00000045527	0.0380
ENSMUSP00000041636	ENSMUSP00000074358	0.0250
ENSMUSP00000075398	ENSMUSP00000092426	0.0191
ENSMUSP00000077246	ENSMUSP00000081869	0.0014

ENSMUSP00000055181	ENSMUSP00000066270	0.0417
ENSMUSP00000058027	ENSMUSP00000096676	0.0306
ENSMUSP00000060524	ENSMUSP00000075190	0.0304
ENSMUSP00000071064	ENSMUSP00000086528	0.0159
ENSMUSP00000072555	ENSMUSP00000086528	0.0004
ENSMUSP00000073251	ENSMUSP00000079451	<0.0001
ENSMUSP00000074242	ENSMUSP00000077220	0.0441
ENSMUSP00000071263	ENSMUSP00000071824	0.0206
ENSMUSP00000071824	ENSMUSP00000087273	0.0346
ENSMUSP00000071604	ENSMUSP00000072947	<0.0001
ENSMUSP00000076912	ENSMUSP00000080646	0.0033
ENSMUSP00000071372	ENSMUSP00000077635	0.0042
ENSMUSP00000073558	ENSMUSP00000073602	0.0385
ENSMUSP00000077615	ENSMUSP00000079205	<0.0001
ENSMUSP00000078700	ENSMUSP00000096472	0.0484
ENSMUSP00000075102	ENSMUSP00000080385	0.0099
ENSMUSP00000090259	ENSMUSP00000090260	0.0285
ENSMUSP00000077521	ENSMUSP00000092002	0.0222
ENSMUSP00000080740	ENSMUSP00000095745	0.0065
ENSMUSP00000078814	ENSMUSP00000087516	0.0352
ENSMUSP00000073588	ENSMUSP00000079881	0.0160
ENSMUSP00000056586	ENSMUSP00000092387	0.0004
ENSMUSP00000048118	ENSMUSP00000096266	<0.0001
ENSMUSP00000048118	ENSMUSP00000096266	0.0002
ENSMUSP00000048118	ENSMUSP00000096266	0.0172
ENSMUSP00000020535	ENSMUSP00000098936	0.0007
ENSMUSP00000076203	ENSMUSP00000079121	0.0405
ENSMUSP00000015588	ENSMUSP00000080742	0.0026
ENSMUSP00000071067	ENSMUSP00000092442	<0.0001
ENSMUSP00000071067	ENSMUSP00000092448	0.0406
ENSMUSP00000071067	ENSMUSP00000096428	0.0406
ENSMUSP00000072978	ENSMUSP00000092448	0.0063
ENSMUSP00000072978	ENSMUSP00000096428	0.0063
ENSMUSP00000073963	ENSMUSP00000077962	0.0028
ENSMUSP00000076390	ENSMUSP00000076407	0.0049
ENSMUSP00000023469	ENSMUSP00000064161	0.0001
ENSMUSP00000077546	ENSMUSP00000092515	0.0147
ENSMUSP00000005077	ENSMUSP00000093512	0.0010
ENSMUSP00000032206	ENSMUSP00000080469	<0.0001
ENSMUSP00000032206	ENSMUSP00000080469	0.0004
ENSMUSP00000003416	ENSMUSP00000076827	0.0232
ENSMUSP00000076671	ENSMUSP00000089622	<0.0001
ENSMUSP00000081855	ENSMUSP00000082009	<0.0001
ENSMUSP00000067114	ENSMUSP00000093990	0.0352
ENSMUSP00000067114	ENSMUSP00000093991	0.0475
ENSMUSP00000093988	ENSMUSP00000093990	0.0016

ENSMUSP00000093988	ENSMUSP00000093990	0.0423
ENSMUSP00000074799	ENSMUSP00000083024	0.0013
ENSMUSP00000062542	ENSMUSP00000078800	0.0058
ENSMUSP00000097191	ENSMUSP00000097192	0.0134
ENSMUSP00000051938	ENSMUSP00000058587	0.0187
ENSMUSP00000073416	ENSMUSP00000074546	0.0035
ENSMUSP00000072207	ENSMUSP00000087282	0.0058
ENSMUSP00000073594	ENSMUSP00000097213	0.0329
ENSMUSP00000087798	ENSMUSP00000097213	0.0170
ENSMUSP00000036258	ENSMUSP00000058567	0.0001
ENSMUSP00000036258	ENSMUSP00000058567	0.0460
ENSMUSP00000058567	ENSMUSP00000063005	0.0265
ENSMUSP00000092586	ENSMUSP00000092591	0.0164
ENSMUSP00000092586	ENSMUSP00000092596	0.0160
ENSMUSP00000092591	ENSMUSP00000092596	0.0486
ENSMUSP00000055035	ENSMUSP00000091467	0.0135
ENSMUSP00000092423	ENSMUSP00000096413	0.0008
ENSMUSP00000072222	ENSMUSP00000075537	0.0004
ENSMUSP00000072222	ENSMUSP00000089314	0.0001
ENSMUSP00000075392	ENSMUSP00000093657	0.0038
ENSMUSP00000015595	ENSMUSP00000069418	<0.0001
ENSMUSP00000092174	ENSMUSP00000098671	0.0449
ENSMUSP00000095083	ENSMUSP00000095088	0.0230
ENSMUSP00000095083	ENSMUSP00000098585	0.0403
ENSMUSP00000095083	ENSMUSP00000098587	0.0439
ENSMUSP00000095083	ENSMUSP00000098671	0.0203
ENSMUSP00000095088	ENSMUSP00000095889	0.0230
ENSMUSP00000095889	ENSMUSP00000098585	0.0403
ENSMUSP00000095889	ENSMUSP00000098587	0.0439
ENSMUSP00000095889	ENSMUSP00000098671	0.0203
ENSMUSP00000051550	ENSMUSP00000075380	<0.0001
ENSMUSP00000051550	ENSMUSP00000081421	0.0218
ENSMUSP00000053912	ENSMUSP00000062113	0.0063
ENSMUSP00000053912	ENSMUSP00000092098	0.0063
ENSMUSP00000062113	ENSMUSP00000092103	0.0050
ENSMUSP00000092098	ENSMUSP00000092103	0.0050
ENSMUSP00000081204	ENSMUSP00000081210	0.0019
ENSMUSP00000047435	ENSMUSP00000056820	0.0372
ENSMUSP00000071989	ENSMUSP00000096106	0.0342
ENSMUSP00000037665	ENSMUSP00000069932	0.0495
ENSMUSP00000071526	ENSMUSP00000087292	<0.0001
ENSMUSP00000067008	ENSMUSP00000073525	0.0093
ENSMUSP00000073286	ENSMUSP00000079579	0.0379
ENSMUSP00000071622	ENSMUSP00000088787	0.0286
ENSMUSP00000069042	ENSMUSP00000072232	0.0319
ENSMUSP00000029463	ENSMUSP00000064673	0.0058

ENSMUSP00000064673	ENSMUSP00000088246	0.0006
ENSMUSP00000021649	ENSMUSP00000082306	0.0016
ENSMUSP00000021649	ENSMUSP00000082306	0.0092
ENSMUSP00000080256	ENSMUSP00000092702	0.0056
ENSMUSP00000080256	ENSMUSP00000092702	0.0119
ENSMUSP00000064828	ENSMUSP00000091504	0.0107
ENSMUSP00000027612	ENSMUSP00000092060	0.0289
ENSMUSP00000066677	ENSMUSP00000092060	0.0144
ENSMUSP00000092060	ENSMUSP00000092065	0.0004
ENSMUSP00000092060	ENSMUSP00000092065	0.0392
ENSMUSP00000053398	ENSMUSP00000093060	0.0003
ENSMUSP00000085130	ENSMUSP00000085132	0.0011
ENSMUSP00000085130	ENSMUSP00000085132	0.0120
ENSMUSP00000053016	ENSMUSP00000095750	<0.0001
ENSMUSP00000029717	ENSMUSP00000039583	0.0013
ENSMUSP00000014476	ENSMUSP00000045595	0.0436
ENSMUSP00000059936	ENSMUSP00000085336	0.0042
ENSMUSP00000094379	ENSMUSP00000097824	0.0487
ENSMUSP00000071797	ENSMUSP00000097624	0.0214
ENSMUSP00000095757	ENSMUSP00000095758	0.0109
ENSMUSP00000095760	ENSMUSP00000095761	0.0009
ENSMUSP00000026093	ENSMUSP00000096864	0.0144
ENSMUSP00000078796	ENSMUSP00000087343	0.0188
ENSMUSP00000073405	ENSMUSP00000089328	0.0027
ENSMUSP00000073121	ENSMUSP00000074422	<0.0001
ENSMUSP00000086839	ENSMUSP00000086840	0.0190
ENSMUSP00000093561	ENSMUSP00000093562	0.0327
ENSMUSP00000074442	ENSMUSP00000078830	0.0179
ENSMUSP00000047032	ENSMUSP00000091780	0.0011
ENSMUSP00000029481	ENSMUSP00000088014	0.0285
ENSMUSP00000029929	ENSMUSP00000054526	0.0449
ENSMUSP00000037127	ENSMUSP00000096010	0.0221
ENSMUSP00000026537	ENSMUSP00000092660	0.0002
ENSMUSP00000071454	ENSMUSP00000089465	0.0151
ENSMUSP00000080854	ENSMUSP00000089465	0.0151
ENSMUSP00000061130	ENSMUSP00000083426	0.0259
ENSMUSP00000056676	ENSMUSP00000087132	<0.0001
ENSMUSP00000051716	ENSMUSP00000068997	0.0308
ENSMUSP00000073107	ENSMUSP00000075822	0.0415
ENSMUSP00000073107	ENSMUSP00000078323	<0.0001
ENSMUSP00000075822	ENSMUSP00000078323	0.0384
ENSMUSP00000086700	ENSMUSP00000094192	0.0001
ENSMUSP00000072405	ENSMUSP00000079048	0.0049
ENSMUSP00000061900	ENSMUSP00000087888	0.0108
ENSMUSP00000000264	ENSMUSP00000009340	0.0001
ENSMUSP00000000264	ENSMUSP00000009340	0.0007

ENSMUSP00000000264	ENSMUSP00000009340	0.0242
ENSMUSP00000062409	ENSMUSP00000009340	<0.0001
ENSMUSP00000062409	ENSMUSP00000009340	0.0075
ENSMUSP00000054351	ENSMUSP00000078365	<0.0001
ENSMUSP00000054351	ENSMUSP00000078365	0.0090
ENSMUSP00000038678	ENSMUSP00000088827	0.0001
ENSMUSP00000074386	ENSMUSP00000097830	0.0277
ENSMUSP00000074386	ENSMUSP00000098255	0.0277
ENSMUSP00000010745	ENSMUSP00000086894	0.0218
ENSMUSP00000080468	ENSMUSP00000086875	0.0016
ENSMUSP00000077588	ENSMUSP00000095659	0.0221
ENSMUSP00000095643	ENSMUSP00000095651	0.0253
ENSMUSP00000095643	ENSMUSP00000095652	0.0253
ENSMUSP00000095651	ENSMUSP00000095658	0.0253
ENSMUSP00000095652	ENSMUSP00000095658	0.0253
ENSMUSP00000044587	ENSMUSP00000093430	0.0045
ENSMUSP00000044587	ENSMUSP00000093430	0.0082
ENSMUSP00000028691	ENSMUSP00000098644	0.0390
ENSMUSP00000043926	ENSMUSP00000093615	0.0436
ENSMUSP00000036682	ENSMUSP00000093059	0.0474
ENSMUSP00000052236	ENSMUSP00000088461	0.0009
ENSMUSP00000075514	ENSMUSP00000085969	0.0048
ENSMUSP00000000724	ENSMUSP00000092184	0.0217
ENSMUSP00000067611	ENSMUSP00000082366	0.0010
ENSMUSP00000004055	ENSMUSP00000097893	<0.0001
ENSMUSP00000071783	ENSMUSP00000076665	0.0312
ENSMUSP00000073950	ENSMUSP00000096494	0.0260
ENSMUSP00000093075	ENSMUSP00000096488	0.0443
ENSMUSP00000082500	ENSMUSP00000094240	0.0059
ENSMUSP00000071834	ENSMUSP00000089610	0.0037
ENSMUSP00000089610	ENSMUSP00000098019	0.0175
ENSMUSP00000053555	ENSMUSP00000066311	<0.0001
ENSMUSP00000082544	ENSMUSP00000089457	0.0284
ENSMUSP00000029754	ENSMUSP00000093818	0.0155
ENSMUSP00000042431	ENSMUSP00000075018	0.0007
ENSMUSP00000032474	ENSMUSP00000032512	0.0084
ENSMUSP00000000327	ENSMUSP00000048568	0.0007
ENSMUSP00000079233	ENSMUSP00000093999	0.0351
ENSMUSP00000040224	ENSMUSP00000095291	0.0007
ENSMUSP00000043112	ENSMUSP00000090401	0.0265
ENSMUSP00000066276	ENSMUSP00000090644	0.0001
ENSMUSP00000055006	ENSMUSP00000094933	0.0007
ENSMUSP00000027650	ENSMUSP00000027654	0.0008
ENSMUSP00000027650	ENSMUSP00000027654	0.0036
ENSMUSP00000067043	ENSMUSP00000097656	0.0046
ENSMUSP00000077223	ENSMUSP00000093161	0.0305

ENSMUSP00000093166	ENSMUSP00000098727	0.0399
ENSMUSP00000098728	ENSMUSP00000098729	0.0382
ENSMUSP00000034934	ENSMUSP00000056476	<0.0001
ENSMUSP00000023589	ENSMUSP00000097623	0.0248
ENSMUSP00000023589	ENSMUSP00000097623	0.0288
ENSMUSP00000074910	ENSMUSP00000076604	0.0226
ENSMUSP00000074910	ENSMUSP00000076604	0.0311
ENSMUSP00000073768	ENSMUSP00000074593	0.0008
ENSMUSP00000073768	ENSMUSP00000080123	0.0002
ENSMUSP00000073768	ENSMUSP00000096443	0.0193
ENSMUSP00000096316	ENSMUSP00000096318	0.0402
ENSMUSP00000096316	ENSMUSP00000096343	0.0402
ENSMUSP00000096318	ENSMUSP00000096340	0.0402
ENSMUSP00000096340	ENSMUSP00000096343	0.0402
ENSMUSP00000077552	ENSMUSP00000096298	0.0110
ENSMUSP00000096297	ENSMUSP00000096298	<0.0001
ENSMUSP00000096297	ENSMUSP00000096300	0.0007
ENSMUSP00000058650	ENSMUSP00000095329	<0.0001
ENSMUSP00000075590	ENSMUSP00000098717	0.0065
ENSMUSP00000031788	ENSMUSP00000055390	0.0418
ENSMUSP00000091323	ENSMUSP00000098858	<0.0001
ENSMUSP00000077960	ENSMUSP00000081772	0.0244
ENSMUSP00000079460	ENSMUSP00000082051	0.0209
ENSMUSP00000024727	ENSMUSP00000096816	0.0280
ENSMUSP00000024727	ENSMUSP00000096817	0.0155
ENSMUSP00000024727	ENSMUSP00000096820	0.0280
ENSMUSP00000096818	ENSMUSP00000096821	0.0022
ENSMUSP00000096818	ENSMUSP00000096822	0.0022
ENSMUSP00000048111	ENSMUSP00000087012	0.0003
ENSMUSP00000022142	ENSMUSP00000070827	0.0003
ENSMUSP00000022142	ENSMUSP00000070827	0.0037
ENSMUSP00000022142	ENSMUSP00000070827	0.0372
ENSMUSP00000023525	ENSMUSP00000096977	0.0043
ENSMUSP00000023525	ENSMUSP00000096977	0.0241
ENSMUSP00000030665	ENSMUSP00000093007	<0.0001
ENSMUSP00000073689	ENSMUSP00000075435	0.0156
ENSMUSP00000047347	ENSMUSP00000079660	0.0026
ENSMUSP00000081154	ENSMUSP00000092069	0.0008
ENSMUSP00000081154	ENSMUSP00000095399	0.0453
ENSMUSP00000020262	ENSMUSP00000020266	0.0001
ENSMUSP00000020262	ENSMUSP00000020266	0.0218
ENSMUSP00000006687	ENSMUSP00000074810	0.0096
ENSMUSP00000045141	ENSMUSP00000059379	0.0019
ENSMUSP00000094082	ENSMUSP00000094083	0.0017
ENSMUSP00000094082	ENSMUSP00000094083	0.0037
ENSMUSP00000023502	ENSMUSP00000086148	0.0446

ENSMUSP00000046909	ENSMUSP00000072512	0.0005
ENSMUSP00000062023	ENSMUSP00000093045	0.0016
ENSMUSP00000058437	ENSMUSP00000062098	<0.0001
ENSMUSP00000058437	ENSMUSP00000062098	0.0015
ENSMUSP00000096964	ENSMUSP00000097279	0.0268
ENSMUSP00000040240	ENSMUSP00000053845	0.0113
ENSMUSP00000077473	ENSMUSP00000080384	0.0473
ENSMUSP00000015051	ENSMUSP00000093427	0.0492
ENSMUSP00000048041	ENSMUSP00000087912	0.0278
ENSMUSP00000078641	ENSMUSP00000086746	0.0282
ENSMUSP00000063842	ENSMUSP00000065612	0.0331
ENSMUSP00000065612	ENSMUSP00000096086	0.0173
ENSMUSP00000096086	ENSMUSP00000096087	0.0035
ENSMUSP00000026912	ENSMUSP00000092780	0.0008
ENSMUSP00000021728	ENSMUSP00000082966	0.0166
ENSMUSP00000027202	ENSMUSP00000039936	0.0133
ENSMUSP00000075226	ENSMUSP00000097216	0.0170
ENSMUSP00000027769	ENSMUSP00000082548	0.0003
ENSMUSP00000075957	ENSMUSP00000089332	0.0349
ENSMUSP00000023474	ENSMUSP00000094267	0.0208
ENSMUSP00000075425	ENSMUSP00000096555	0.0073
ENSMUSP00000094130	ENSMUSP00000094131	0.0004
ENSMUSP00000055935	ENSMUSP00000070138	0.0004
ENSMUSP00000055935	ENSMUSP00000070138	0.0449
ENSMUSP00000066137	ENSMUSP00000091392	0.0416

D

Gene 1	Gene 2	*P*-value
ENSPNOP00000041288	ENSPNOP00000046154	0.0270
ENSPNOP00000016061	ENSPNOP00000041368	0.0443
ENSPNOP00000040143	ENSPNOP00000051198	0.0282
ENSPNOP00000048041	ENSPNOP00000051198	0.0344
ENSPNOP00000040410	ENSPNOP00000051204	0.0012
ENSPNOP00000034877	ENSPNOP00000039547	0.0029
ENSPNOP00000034877	ENSPNOP00000039547	0.0289
ENSPNOP00000039547	ENSPNOP00000042674	0.0003
ENSPNOP00000045457	ENSPNOP00000057983	0.0397
ENSPNOP00000029986	ENSPNOP00000055060	0.0071
ENSPNOP00000044813	ENSPNOP00000055153	0.0045
ENSPNOP00000033356	ENSPNOP00000039965	0.0148
ENSPNOP00000041343	ENSPNOP00000045073	0.0300
ENSPNOP00000039066	ENSPNOP00000048160	0.0341
ENSPNOP00000044363	ENSPNOP00000048160	0.0050
ENSPNOP00000042088	ENSPNOP00000048940	0.0180
ENSPNOP00000042322	ENSPNOP00000046245	0.0058
ENSPNOP00000048074	ENSPNOP00000049109	0.0232

ENSPNOP00000034723	ENSPNOP00000043239	0.0074
ENSPNOP00000004503	ENSPNOP00000057716	0.0006
ENSPNOP00000047137	ENSPNOP00000048662	0.0011
ENSPNOP00000047137	ENSPNOP00000048662	0.0307
ENSPNOP00000042053	ENSPNOP00000046782	0.0038
ENSPNOP00000047762	ENSPNOP00000049956	0.0060
ENSPNOP00000041683	ENSPNOP00000058205	0.0086
ENSPNOP00000048263	ENSPNOP00000049798	0.0154
ENSPNOP00000058205	ENSPNOP00000058219	0.0170
ENSPNOP00000049221	ENSPNOP00000058171	0.0351
ENSPNOP00000050205	ENSPNOP00000058253	0.0428
ENSPNOP00000039954	ENSPNOP00000044774	0.0118
ENSPNOP00000041991	ENSPNOP00000051478	0.0077
ENSPNOP00000042222	ENSPNOP00000042970	0.0235
ENSPNOP00000042222	ENSPNOP00000045509	0.0041
ENSPNOP00000042222	ENSPNOP00000053198	<0.0001
ENSPNOP00000042970	ENSPNOP00000045509	0.0150
ENSPNOP00000042970	ENSPNOP00000051478	0.0156
ENSPNOP00000042970	ENSPNOP00000053198	0.0005
ENSPNOP00000042816	ENSPNOP00000049182	0.0073
ENSPNOP00000049182	ENSPNOP00000049212	0.0073
ENSPNOP00000020822	ENSPNOP00000046951	0.0053
ENSPNOP00000045184	ENSPNOP00000046700	0.0013
ENSPNOP00000046329	ENSPNOP00000047888	0.0349
ENSPNOP00000043696	ENSPNOP00000049692	0.0013
ENSPNOP00000043696	ENSPNOP00000049692	0.0227
ENSPNOP00000031107	ENSPNOP00000049307	0.0412
ENSPNOP00000020952	ENSPNOP00000041128	0.0074
ENSPNOP00000040266	ENSPNOP00000041174	0.0369
ENSPNOP00000047467	ENSPNOP00000049399	0.0254
ENSPNOP00000050575	ENSPNOP00000055214	0.0083
ENSPNOP00000056854	ENSPNOP00000007684	0.0436
ENSPNOP00000056843	ENSPNOP00000056844	<0.0001
ENSPNOP00000056844	ENSPNOP00000056848	0.0005
ENSPNOP00000041702	ENSPNOP00000045349	0.0218
ENSPNOP00000039551	ENSPNOP00000049209	0.0257
ENSPNOP00000043172	ENSPNOP00000044756	<0.0001
ENSPNOP00000044208	ENSPNOP00000044950	0.0454
ENSPNOP00000041311	ENSPNOP00000048202	0.0003
ENSPNOP00000002046	ENSPNOP00000045076	0.0459
ENSPNOP00000040213	ENSPNOP00000043037	0.0104
ENSPNOP00000041076	ENSPNOP00000046998	0.0487
ENSPNOP00000043505	ENSPNOP00000044377	0.0095
ENSPNOP00000027213	ENSPNOP00000046654	0.0176
ENSPNOP00000054798	ENSPNOP00000054809	0.0050
ENSPNOP00000042475	ENSPNOP00000050449	0.0148

ENSPNOP00000042475	ENSPNOP00000050449	0.0312
ENSPNOP00000039457	ENSPNOP00000043726	0.0147
ENSPNOP00000040006	ENSPNOP00000045094	0.0112
ENSPNOP00000040845	ENSPNOP00000042173	0.0249
ENSPNOP00000047521	ENSPNOP00000048950	0.0297
ENSPNOP00000029828	ENSPNOP00000056392	0.0010
ENSPNOP00000037783	ENSPNOP00000056467	0.0119
ENSPNOP00000027918	ENSPNOP00000056483	0.0012
ENSPNOP00000022268	ENSPNOP00000046188	0.0042
ENSPNOP00000013896	ENSPNOP00000013961	0.0002
ENSPNOP00000013896	ENSPNOP00000013961	0.0131
ENSPNOP00000052820	ENSPNOP00000052821	0.0025
ENSPNOP00000017042	ENSPNOP00000043746	0.0184
ENSPNOP00000041843	ENSPNOP00000053460	0.0050
ENSPNOP00000045781	ENSPNOP00000053460	0.0063
ENSPNOP00000013328	ENSPNOP00000043504	0.0046
ENSPNOP00000041727	ENSPNOP00000058527	0.0231
ENSPNOP00000039430	ENSPNOP00000047683	0.0010
ENSPNOP00000039430	ENSPNOP00000047683	0.0010
ENSPNOP00000041735	ENSPNOP00000050588	<0.0001
ENSPNOP00000041735	ENSPNOP00000050588	0.0015
ENSPNOP00000049679	ENSPNOP00000056921	0.0107
ENSPNOP00000042221	ENSPNOP00000049850	0.0001
ENSPNOP00000016305	ENSPNOP00000058522	0.0002
ENSPNOP00000041890	ENSPNOP00000044743	0.0151
ENSPNOP00000042226	ENSPNOP00000043440	0.0450
ENSPNOP00000041551	ENSPNOP00000049573	0.0346
ENSPNOP00000039211	ENSPNOP00000048145	<0.0001
ENSPNOP00000041610	ENSPNOP00000048145	0.0278
ENSPNOP00000015370	ENSPNOP00000056423	0.0027
ENSPNOP00000039667	ENSPNOP00000047949	0.0121
ENSPNOP00000040706	ENSPNOP00000047401	0.0121
ENSPNOP00000045107	ENSPNOP00000049417	0.0222
ENSPNOP00000029029	ENSPNOP00000047504	<0.0001
ENSPNOP00000040450	ENSPNOP00000049773	0.0001
ENSPNOP00000032852	ENSPNOP00000040764	<0.0001
ENSPNOP00000032852	ENSPNOP00000057351	0.0001
ENSPNOP00000052553	ENSPNOP00000052555	0.0279
ENSPNOP00000001291	ENSPNOP00000041138	0.0005
ENSPNOP00000046439	ENSPNOP00000047124	0.0232
ENSPNOP00000058398	ENSPNOP00000008319	0.0037
ENSPNOP00000018934	ENSPNOP00000026003	<0.0001
ENSPNOP00000018934	ENSPNOP00000026003	0.0420
ENSPNOP00000001415	ENSPNOP00000043856	0.0105
ENSPNOP00000052099	ENSPNOP00000052103	0.0047
ENSPNOP00000050478	ENSPNOP00000051367	0.0008

ENSPNOP00000050478	ENSPNOP00000051367	0.0017
ENSPNOP00000016364	ENSPNOP00000016391	0.0060
ENSPNOP00000056638	ENSPNOP00000056642	<0.0001
ENSPNOP00000034875	ENSPNOP00000040305	<0.0001
ENSPNOP00000049464	ENSPNOP00000050871	0.0435
ENSPNOP00000056776	ENSPNOP00000056797	0.0301
ENSPNOP00000044387	ENSPNOP00000045937	<0.0001
ENSPNOP00000045937	ENSPNOP00000048609	<0.0001
ENSPNOP00000013303	ENSPNOP00000043859	<0.0001
ENSPNOP00000030037	ENSPNOP00000046752	0.0007
ENSPNOP00000043859	ENSPNOP00000046752	0.0002
ENSPNOP00000043598	ENSPNOP00000052660	0.0043
ENSPNOP00000020647	ENSPNOP00000039702	0.0247
ENSPNOP00000018462	ENSPNOP00000021237	0.0324
ENSPNOP00000017619	ENSPNOP00000053039	0.0014
ENSPNOP00000015398	ENSPNOP00000049742	0.0074
ENSPNOP00000015398	ENSPNOP00000054999	0.0114
ENSPNOP00000049742	ENSPNOP00000054999	<0.0001
ENSPNOP00000054012	ENSPNOP00000054040	0.0065
ENSPNOP00000043326	ENSPNOP00000047657	0.0149
ENSPNOP00000045860	ENSPNOP00000050262	0.0028
ENSPNOP00000018325	ENSPNOP00000057443	0.0186
ENSPNOP00000018325	ENSPNOP00000057443	0.0424
ENSPNOP00000010891	ENSPNOP00000056632	0.0234
ENSPNOP00000028590	ENSPNOP00000044949	0.0458
ENSPNOP00000000134	ENSPNOP00000038260	0.0260
ENSPNOP00000023404	ENSPNOP00000051319	0.0427
ENSPNOP00000011102	ENSPNOP00000035155	0.0351
ENSPNOP00000040740	ENSPNOP00000049710	0.0320
ENSPNOP00000028681	ENSPNOP00000009815	0.0193
ENSPNOP00000024311	ENSPNOP00000047004	<0.0001
ENSPNOP00000042734	ENSPNOP00000048101	0.0204
ENSPNOP00000042738	ENSPNOP00000050006	0.0090
ENSPNOP00000040923	ENSPNOP00000049412	0.0323
ENSPNOP00000043009	ENSPNOP00000049863	0.0248
ENSPNOP00000044806	ENSPNOP00000058358	0.0167
ENSPNOP00000021645	ENSPNOP00000039262	0.0020
ENSPNOP00000043156	ENSPNOP00000053105	0.0130
ENSPNOP00000042061	ENSPNOP00000046328	0.0310
ENSPNOP00000039557	ENSPNOP00000049161	0.0102
ENSPNOP00000047140	ENSPNOP00000051286	0.0004
ENSPNOP00000011092	ENSPNOP00000011312	0.0188
ENSPNOP00000001227	ENSPNOP00000043587	0.0234
ENSPNOP00000001227	ENSPNOP00000049252	0.0154
ENSPNOP00000028496	ENSPNOP00000041346	0.0077
ENSPNOP00000028496	ENSPNOP00000041346	0.0401

ENSPNOP00000025888	ENSPNOP00000037426	0.0235
ENSPNOP00000031121	ENSPNOP00000037426	0.0235
ENSPNOP00000040648	ENSPNOP00000050854	0.0135
ENSPNOP00000034293	ENSPNOP00000034917	0.0020
ENSPNOP00000037267	ENSPNOP00000054846	0.0185
ENSPNOP00000048400	ENSPNOP00000049567	0.0146
ENSPNOP00000009594	ENSPNOP00000009904	0.0001
ENSPNOP00000009594	ENSPNOP00000009904	0.0146
ENSPNOP00000011314	ENSPNOP00000050051	0.0377
ENSPNOP00000058201	ENSPNOP00000058213	0.0258
ENSPNOP00000044116	ENSPNOP00000056331	0.0099
ENSPNOP00000045542	ENSPNOP00000054888	0.0140
ENSPNOP00000047960	ENSPNOP00000054448	<0.0001
ENSPNOP00000016036	ENSPNOP00000056901	<0.0001
ENSPNOP00000015017	ENSPNOP00000055188	0.0402
ENSPNOP00000043597	ENSPNOP00000044739	0.0088
ENSPNOP00000043597	ENSPNOP00000044739	0.0088
ENSPNOP00000043597	ENSPNOP00000044739	0.0144
ENSPNOP00000043597	ENSPNOP00000044739	0.0291
ENSPNOP00000042228	ENSPNOP00000043220	0.0401
ENSPNOP00000024176	ENSPNOP00000054515	0.0006
ENSPNOP00000039452	ENSPNOP00000044766	0.0417
ENSPNOP00000049533	ENSPNOP00000056134	0.0047
ENSPNOP00000042883	ENSPNOP00000043314	0.0359
ENSPNOP00000001589	ENSPNOP00000043021	0.0434
ENSPNOP00000026968	ENSPNOP00000007288	0.0063
ENSPNOP00000026968	ENSPNOP00000007288	0.0081
ENSPNOP00000026968	ENSPNOP00000007288	0.0272
ENSPNOP00000044976	ENSPNOP00000004563	<0.0001
ENSPNOP00000004563	ENSPNOP00000058455	<0.0001
ENSPNOP00000040559	ENSPNOP00000041418	0.0353
ENSPNOP00000051763	ENSPNOP00000005658	0.0043
ENSPNOP00000020749	ENSPNOP00000030860	<0.0001
ENSPNOP00000020749	ENSPNOP00000030860	0.0117
ENSPNOP00000056282	ENSPNOP00000058785	0.0365
ENSPNOP00000031104	ENSPNOP00000036409	0.0012
ENSPNOP00000032200	ENSPNOP00000056137	<0.0001
ENSPNOP00000032200	ENSPNOP00000056137	0.0011
ENSPNOP00000032200	ENSPNOP00000056137	0.0072
ENSPNOP00000042976	ENSPNOP00000004958	0.0166
ENSPNOP00000055320	ENSPNOP00000056021	0.0391
ENSPNOP00000012263	ENSPNOP00000051333	0.0293
ENSPNOP00000017601	ENSPNOP00000050070	0.0031
ENSPNOP00000042710	ENSPNOP00000009039	0.0484
ENSPNOP00000043556	ENSPNOP00000058245	0.0052
ENSPNOP00000019667	ENSPNOP00000031889	0.0008

ENSPNOP00000013287	ENSPNOP00000015467	<0.0001
ENSPNOP00000046710	ENSPNOP00000048648	<0.0001
ENSPNOP00000016089	ENSPNOP00000006734	0.0372
ENSPNOP00000045183	ENSPNOP00000047964	0.0113
ENSPNOP00000012102	ENSPNOP00000048440	0.0401
ENSPNOP00000023070	ENSPNOP00000049711	0.0342
ENSPNOP00000015146	ENSPNOP00000000911	0.0117
ENSPNOP00000028476	ENSPNOP00000028483	0.0162
ENSPNOP00000028807	ENSPNOP00000056065	0.0020
ENSPNOP00000031413	ENSPNOP00000058055	0.0180
ENSPNOP00000035857	ENSPNOP00000056230	0.0016
ENSPNOP00000037955	ENSPNOP00000044764	0.0151
ENSPNOP00000042398	ENSPNOP00000053992	0.0117
ENSPNOP00000031710	ENSPNOP00000000580	0.0213
ENSPNOP00000047671	ENSPNOP00000049951	0.0037
ENSPNOP00000039674	ENSPNOP00000043243	0.0035
ENSPNOP00000043243	ENSPNOP00000046225	0.0035
ENSPNOP00000043419	ENSPNOP00000053395	0.0001
ENSPNOP00000029603	ENSPNOP00000030175	0.0119
ENSPNOP00000044356	ENSPNOP00000051547	0.0424