Downloaded from genome.cshlp.org on June 15, 2026 . Published by Cold Spring Harbor Laboratory Press

Naturally occurring antisense: Transcriptional leakage or real overlap?

Dvir Dahary, Orna Elroy-Stein and Rotem Sorek

Genome Res. 2005 15: 364-368
Access the most recent version at doi:10.1101/gr.3308405

References This article cites 23 articles, 10 of which can be accessed free at:
http://genome.cshlp.org/content/15/3/364.full.html#ref-list-1

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on June 15, 2026 . Published by Cold Spring Harbor Laboratory Press

Letter

Naturally occurring antisense: Transcriptional leakage or real overlap?

Dvir Dahary,[1,2,3] Orna Elroy-Stein,[2] and Rotem Sorek[1]

[1]*Compugen Ltd., Tel Aviv 69512 Israel;* [2]*Department of Cell Research & Immunology, George S. Wise Faculty of Life Science, Tel Aviv University, Tel Aviv 69978, Israel*

Naturally occurring antisense transcription is associated with the regulation of gene expression through a variety of biological mechanisms. Several recent genome-wide studies reported the identification of potential antisense transcripts for thousands of mammalian genes, many of them resulting from alternatively polyadenylated transcripts or heterogeneous transcription start sites. However, it is not clear whether this transcriptional plasticity is intentional, leading to regulated overlap between the transcripts, or, alternatively, represents a "leakage" of the RNA transcription machinery. To address this question through an evolutionary approach, we compared the genomic organization of genes, with or without antisense, between human, mouse, and the pufferfish *Fugu rubripes*. Our hypothesis was that if two neighboring genes overlap and have a sense–antisense relationship, we would expect negative selection acting on the evolutionary separation between them. We found that antisense gene pairs are twice as likely to preserve their genomic organization throughout vertebrates' evolution compared to nonantisense pairs, implying an overlap existence in the ancestral genome. In addition, we show that increasing the genomic distance between pairs of genes having a sense–antisense relationship is selected against. These findings indicate that, at least in part, the abundance of antisense transcripts observed in expressed data represents real overlap rather than transcriptional leakage. Moreover, our results imply that natural antisense transcription has considerably affected vertebrate genome evolution.

[Supplemental material is available online at www.genome.org.]

Naturally occurring antisense (*cis*-encoded) describes a genomic locus in which two partially overlapping genes are transcribed from opposite strands of the DNA. In such a case, RNA transcribed from the sense gene may interact with the antisense RNA, possibly leading to various cellular responses. Until recently, naturally occurring antisense was described mainly in viruses and prokaryotes (Wagner and Simons 1994). However, studies of individual antisense transcripts have shown them to regulate key gene expression mechanisms in eukaryotes (Knee and Murphy 1997; Kumar and Carmichael 1998) including genomic imprinting (Rougeulle and Heard 2002), RNA interference and translational regulation (Brantl 2002), transcriptional interference (Prescott and Proudfoot 2002), alternative splicing (Hastings et al. 2000), X-inactivation (Ogawa and Lee 2002), and RNA editing (Peters et al. 2003). In addition, increasing evidence suggests that natural antisense transcription (also called NAT) may play a key role in a range of human diseases (for review, see Lavorgna et al. 2004).

Several independent studies recently reported that antisense transcription is widespread in mammals. In humans, between 5% and 10% of all genes were found to have an antisense counterpart (Lehner et al. 2002; Shendure and Church 2002; Yelin et al. 2003). Similar results were reported for the mouse genome, where ~2400 sense–antisense gene pairs have been identified (Kiyosawa et al. 2003).

Interestingly, in most cases of antisense overlap between two protein-coding genes, the overlap is restricted to their untranslated regions (UTRs). In addition, in many of the cases, the antisense overlap involves alternative polyadenylation, creating several variants of the transcript that differ in their 3′ termini length. For example, in the *TP53BP1-76P* sense–antisense locus (Yelin et al. 2003), the abundant transcript of *TP53BP1* is of 6.3 kb, with no potential overlap with *76P* transcripts (Fig. 1). The transcripts can overlap only when a less abundant, 10.5-kb alternatively polyadenylated *TP53BP1* transcript, or a longer, 6.8-kb alternatively polyadenylated *76P* transcript, is generated (Yelin et al. 2003). This phenomenon is apparent also in the *CCNE2-FLJ20530* locus described by Yelin et al. (2003) and in the Hs.125819 locus described by Shendure and Church (2002).

The large heterogeneity of 3′ and 5′ ends in human transcripts has been reported before. Firstly, many overlapping genes exhibit complex 5′ UTR and promoter structures (for review, see Boi et al. 2004). Secondly, it was suggested that at least half of all human genes encode multiple transcripts with alternative 3′ termini (Iseli et al. 2002). However, it was not established whether this alternative 3′ end processing is intentional, leading to regulated overlap between the transcripts, or, alternatively, represents a "leakage" of the RNA transcription machinery. Indeed, failure of the transcription machinery to recognize the correct polyA site (for example through mutations in the polyA site) can lead to transcription read-through into downstream genes (Connelly and Manley 1988). In addition, when several closely spaced polyA sites reside in the same transcript, they compete for polyadenylation (the most upstream one chosen preferentially, but downstream sites are also active) (Batt et al. 1994). Such polyA sites can easily be added in evolution: The L1 retrotransposon, which accounts for ~17% of the human genome, contains a strong polyA site in its antisense orientation (Han et al. 2004). It was hypothesized that such L1, when inserted downstream to a certain gene, can compete with the original polyA site and

[3]Corresponding author.
E-mail dvir@compugen.co.il; fax 972-3-765-8555.
Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.3308405. Article published online ahead of print in February 2005.

15:364–368 ©2005 by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/05; www.genome.org

Downloaded from genome.cshlp.org on June 15, 2026 . Published by Cold Spring Harbor Laboratory Press

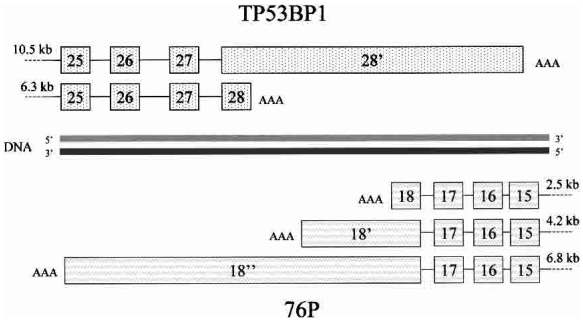

Figure 1. The *TP53BP1-76P* sense–antisense locus (Yelin et al. 2003). Two alternatively polyadenylated transcripts of *TP53BP1* (*above* DNA) and three alternatively polyadenylated transcripts of *76P* (*below* DNA). The abundant transcripts of both genes are the short variants; overlap is only possible when the longer form of one of the genes is produced.

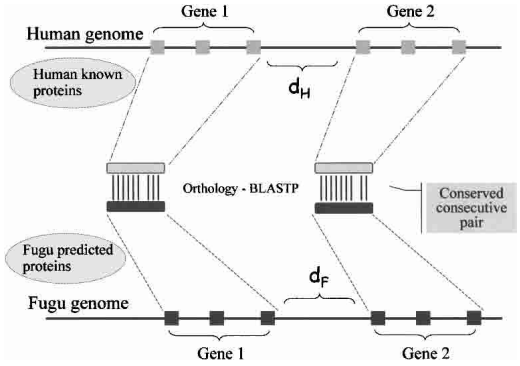

Figure 2. Identification of conserved consecutive gene pairs between human and *Fugu* genomes. An orthology between human and *Fugu* proteins (light and dark boxes, respectively) was defined using BLASTP as described in Methods; mappings of proteins to the human and *Fugu* genomes (light and dark boxes, respectively) were used to define a consecutive pair and to calculate the distance between the coding sequence coordinates in each pair (d_H and d_F for human and *Fugu*, respectively).

cause the elongation of some of the transcripts through alternative polyadenylation, leading to an overlap with a proximate downstream gene (Han et al. 2004). Presumably, such a read-through into an oppositely oriented gene will be represented as antisense overlap between the two genes. Whether this overlap has a biological relevance is questionable.

In this study we employed an evolutionary approach to address this question by comparing the genome organization between human and the pufferfish *Fugu rubripes*. Although their gene repertoire is similar, the 450 million years of evolution caused a considerable scrambling in gene order between these two genomes (Aparicio et al. 2002). In addition, there has been a significant genome expansion in the mammalian lineage (mostly due to transposable element activity) together with a possible compaction in the teleost fish lineage, so that the human genome is eightfold larger than that of *Fugu* (Aparicio et al. 2002). From an evolutionary point of view, if two neighboring genes overlap and have a sense–antisense relationship, we would expect the separation between them, either by rearrangement or by genome expansion, to be selected against. It was therefore appealing to test whether such a selection could be observed.

We show here that antisense gene pairs tend to preserve their genome organization significantly more than nonantisense pairs, suggesting that the overlap observed in the human genome may be conserved throughout vertebrate evolution. This conservation implies that the overlap is "real" rather than transcriptional leakage, for a substantial number of human sense–antisense gene pairs.

Results

Gene pairs with conserved linkage between human and *Fugu*

To identify pairs of genes that remained consecutive in both human and *Fugu*, we first assigned one-to-one orthologies between genes. For this, we used a method similar to that described by Aparicio et al. (2002). Using that method, 33,609 predicted *Fugu* peptides were compared to 26,309 known human peptides to identify 9156 human–*Fugu* orthologous genes (see Methods).

We mapped these 9156 genes to the human and *Fugu* genomes, and further analyzed only pairs of consecutive genes (see Methods). We found 2737 such pairs on the human genome. Of these, 453 pairs (16.5%) were found to be consecutive on the *Fugu* genome as well (Fig. 2). This set represents gene pairs with conserved linkage between human and *Fugu*. Our results are comparable to those of Aparicio et al. (2002), who reported on similar rates of linkage in human–*Fugu* genomic comparison.

Antisense transcription and gene order evolution

To examine the influence of antisense transcription on the human–*Fugu* gene order evolution, we first used the "Antisensor" algorithm (Yelin et al. 2003) to detect natural antisense transcription in human genes. Out of the 2737 consecutive gene pairs in human, 236 (8.6%) had a sense–antisense relationship with each other, similar to the percentage predicted for all human genes (Lehner et al. 2002; Shendure and Church 2002; Yelin et al. 2003). Note that we could not detect antisense overlaps in *Fugu*, as there are very few full-length cDNAs and/or fish ESTs in the public databases.

Of these 236 human antisense pairs, 55 (23.3%) remained consecutive and preserved their orientation in *Fugu* (Table 1). As a control set we used pairs of genes that are consecutive and transcribed from the same strand in human, and are thus unable to possess a sense–antisense relationship (denoted "same-strand"). Remarkably, only 13.5% (170/1257) of the "same-strand" gene pairs remained consecutive in the *Fugu* genome (Table 1). This indicates that gene pairs with an antisense re-

Table 1. Conservation of gene order and orientation between human and *Fugu*

Human pairs classification[a]	Consecutive pairs in human[b]	Pairs with human–*Fugu* linkage[c]	Percent conserved pairs[d]
Antisense	236	55	23.3%
Same-strand	1257	170	13.5%
Total	2737	453	16.3%

[a]Classification of groups of gene-pairs. Antisense, pairs for which a sense–antisense relationship was observed in human cDNAs and ESTs. Same-strand, pairs of consecutive genes that are on the same strand and are therefore unable to possess sense–antisense relationship.
[b]Pairs of human genes that are consecutive on the human genome and have orthologs in the *Fugu* genome.
[c]Pairs of consecutive orthologous genes that preserved their gene order and orientation in human and *Fugu*.
[d]Percentage of linked pairs.

Downloaded from genome.cshlp.org on June 15, 2026 . Published by Cold Spring Harbor Laboratory Press

Table 2. Genomic distances of gene pairs with human–*Fugu* linkage

	Antisense		Same-strand		Total	
Distance[a]	Human	*Fugu*	Human	*Fugu*	Human	*Fugu*
Average (kb)	5.4	2.2	78	5.8	56.8	5
Median (kb)	1.8	1.4	26.6	3.2	16.9	2.5
# of pairs	55		170		453	

[a]Distances were calculated between the coding sequence (CDS) coordinates of each pair of genes on each genome.

lationship tend to maintain their gene order significantly more than genes that originate from the same strand (*P*-value < $1*10^{-4}$, by Fisher's exact test), suggesting that some of the pairs had overlap in the *Fugu* genome, so that antisense transcription restricted gene order shuffling throughout vertebrate genome evolution.

It is known that the probability of rearrangement can depend on the distance between a pair of genes in the ancestral genome (Kent et al. 2003). Indeed, while the average distance of antisense pairs on *Fugu* was 2.2 kb, the distance of "same-strand" pairs was 5.8 kb (Table 2). To rule out the possibility that the observed rearrangement differences between antisense and "same-strand" gene pairs resulted from differences in their original distances, we analyzed the group of the closest 450 "same-strand" pairs that have an average distance comparable to that of the antisense pairs. Still, only 13.5% (61/450) "same-strand" pairs had their gene order and orientation conserved.

Note that we eliminated from this analysis pairs of genes that are oppositely oriented but do not show overlap in EST databases. Intriguingly, 18.3% (228/1244) of these pairs showed conserved linkage in *Fugu*. This is less than the fraction of antisense pairs (23.3%) but significantly more than that of the "same-strand" pairs (13.5%; $P < 6*10^{-4}$, by Fisher's exact test). These results imply that some of these gene pairs may possess a sense–antisense relationship (that could not be identified using the currently available expressed sequences), and suggest that the fraction of antisense gene pairs in vertebrates is higher than previously proposed.

Antisense transcription and genome expansion in vertebrates

The *Fugu* genome is the smallest known vertebrate genome, comprising ~365 Mb (eight times smaller than human; Aparicio et al. 2002). The compactness of the *Fugu* genome mostly stems from the fact that its introns are short, and its intergenic regions are highly compressed compared to the human genome. Much of the relative expansion of human introns and intergenic regions is attributed to the abundant repetitive elements in human (Lander et al. 2001). Such repetitive elements are rare in *Fugu* (Aparicio et al. 2002).

To investigate whether antisense transcription influences the nature of gene-distance expansion, we calculated the distances between the 453 pairs of genes that preserved the same order and orientation in human and *Fugu*. As the rarity of expressed sequences in the *Fugu* EST/cDNA databases prevents the correct annotation of the full lengths of the *Fugu* genes, the distances were measured between the protein-coding sequences (CDS) of each gene pair (see Methods).

Table 2 summarizes the distances for each group of gene pairs. For the entire group of 453 pairs, the average distance between genes was 11-fold larger in human than in *Fugu* (average distances 56.8 kb and 5 kb, respectively; Fig. 3A). This difference of distances was more pronounced in the "same-strand" group of pairs (averaging 78 kb and 5.8 kb in human and *Fugu*, respectively). Remarkably, only a 2.5-fold difference was observed for antisense gene pairs, with average distance of 2.2 kb in *Fugu* and 5.4 kb in human (medians 1.4 kb and 1.8 kb, respectively; Fig. 3A). These results suggest a negative selection on the separation of sense–antisense genes, implying real ancestral overlap.

To show that this effect does not stem from the smaller distance of antisense pairs on the *Fugu* genome, we further studied the distance relationships of the "antisense" and the "same-strand" pairs that are up to 5 kb apart in *Fugu* (50 and 120 "antisense" and "same-strand" pairs, respectively). Figure 3B illustrates the distribution of distance differences for these two groups. While most of the "antisense" gene pairs maintain a short distance in both genomes, many of the "same-strand" pairs are separated by tens of kilobases in human ($P < 6*10^{-6}$, by Student's *t*-test). This further implies that an antisense relationship between a pair of genes imposes a stringent restriction on the expansion of the distance between them, most probably due to the deleterious effects of transposable element insertion between the genes.

Antisense transcription and genome expansion in mammals

To check whether the evolutionary effect of antisense seen between human and fish also influences mammalian genome evo-

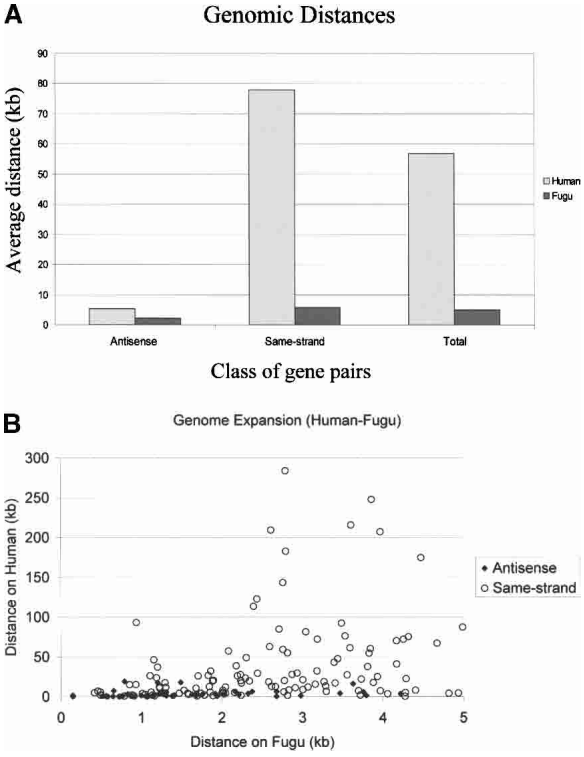

Figure 3. Antisense and gene distance expansion. (*A*) Average genomic distances (in kb) between pairs of genes in human and *Fugu* genomes. (*B*) Relationship between the human and *Fugu* genomic distances. For this analysis, only the 50 "antisense" and 120 "same-strand" pairs with distance on *Fugu* <5 kb were taken. While the "same-strand" group shows large distance expansion in human, the distance of pairs in the "antisense" group is almost unvaried between the two genomes. Distances appear in kb. Note the scale difference between the two axes.

Downloaded from genome.cshlp.org on June 15, 2026 . Published by Cold Spring Harbor Laboratory Press

lution, we compared the human and mouse genomes. Due to the relatively short evolutionary distance, gene order remains largely similar between human and mouse (Waterston et al. 2002; Kent et al. 2003). There is, however, a difference in the total length of the euchromatic portion of these two genomes (2.9 Gb and 2.5 Gb in human and mouse, respectively), probably due to the different composition of lineage-specific repetitive elements, as well as different rates of deletion (Waterston et al. 2002).

Using the HomoloGene database and the mapping of the proteins on each genome (see Methods), we compiled a set of 4036 pairs of consecutive genes that maintain the same gene order and orientation in both human and mouse. This number of pairs is relatively small due to the strict orthology definition (see Methods). We observed that 338 and 287 of these gene pairs have an antisense relationship in human and mouse, respectively. For the human–mouse analysis, we classified as "antisense" only the 135 pairs of genes that were identified as antisense both in the human and the mouse genomes.

As summarized in Table 3, the distance between consecutive pairs on the human genome is about 10% longer in human ($P < 1e\text{-}15$, Student's *t*-test), more or less as expected from the differences in the genome sizes (~16%). For pairs of genes with an antisense relationship, however, there is no significant ($P > 0.5$) distance difference between human and mouse (4.1 and 4.5 kb, respectively). This further demonstrates that NAT affects mammalian genome evolution in a manner similar to that observed for human–fish.

Discussion

Taken together, our results suggest a negative selection on the separation of sense–antisense pairs across vertebrate evolution. This indicates the ancestral existence of overlap, at least for some of the detected antisense loci. For these pairs, therefore, the overlap observed from EST and cDNA analyses is "real" and not the result of random transcriptional leakage. Naturally, separation between such a pair would have affected the two genes (as opposed to the neutral effect of separation between nonoverlapping genes), and hence the negative selection observed.

We found that antisense gene pairs are almost twice as likely to show linkage in *Fugu* compared to pairs transcribed from the same strand. The antisense pairs with linkage comprised 23% of all antisense pairs; however, the actual number of conserved pairs is probably much higher, for several reasons. First, the assembly of the *Fugu* genome is fragmented (comprised of 9011 scaffolds), preventing many pairs from being observed as consecutive in *Fugu*. Second, orthology assignment was not possible for all genes, with fast-evolving genes probably lacking sufficient similarity between organisms. Third, the annotation of the *Fugu* genome is preliminary, and artifacts, e.g., a dubious open reading frame between a pair of real genes, can lead to the misidentification of some actually conserved pairs.

On the other hand, some of the human predicted NATs might be false predictions. Artifacts, such as intron contamination in EST databases, can lead to misidentification of antisense overlaps. The comparative method we presented can assist in evaluating whether a predicted antisense is "real," by checking its human–*Fugu* linkage as well as its human–*Fugu* distance difference. For instance, the *TP53BP1-76p* antisense locus (Yelin et al. 2003) shows linkage as well as conserved genomic distances (0.6 kb and 2 kb in *Fugu* and human, respectively; Fig. 1).

Our results imply that NAT gene pairs had an effect on the evolution of additional vertebrate genomes. For example, analysis of the human–chicken relationship shows that while nonantisense genes have a sixfold increased CDS distance in human, the distance between NAT pairs was increased only twofold. We predict that such an effect would be observed for many other vertebrate genomes.

We note that, due to the remoteness of genomes compared, our study is confined to NAT pairs in which both genes code for proteins. This is because amino acid sequences are generally more conserved, and are thus more reliable for orthology assignment, than the nucleic acid sequences coding for them. The coding-coding class of antisense pairs is estimated to comprise ~30% of the antisense cases in mammalian genomes (Kiyosawa et al. 2003; Yelin et al. 2003). Alternative methods should be used to assess the conservation and biological functions of noncoding antisense transcripts in the process of vertebrate evolution.

In summary, our results imply that generally, antisense observed in expressed sequences represents real overlap rather than transcriptional leakage. Our results also indicate that natural antisense transcription might significantly influence the evolution of vertebrate genomes, an effect so far largely overlooked. Indeed, nonrandom gene order in eukaryotes was observed before (Hurst et al. 2004). As antisense gene pairs might comprise ~10% of all human genes, they can impose a major restriction on gene-order evolution and gene-distance expansion.

Table 3. Genomic distances of gene pairs with human–mouse linkage

	Antisense[a]		Same-strand		Total[b]	
Distance[c]	Human	Mouse	Human	Mouse	Human	Mouse
Average (kb)	4.1	4.5	115	104	110	100
Median (kb)	1.4	1.5	28	22	26	21
Average difference (kb)	−0.4		11		10	
# of pairs	135		1908		3681	

[a]Pairs of genes that were identified as antisense both in the human and the mouse genomes.
[b]4036 pairs of consecutive genes that maintain the same gene order and orientation in both human and mouse were analyzed. Pairs that show antisense overlap in only one organism were discarded, resulting in a final set of 3681 pairs.
[c]Distances were calculated between the coding sequence (CDS) coordinates of each pair of genes on each genome (see Methods).

Methods

A data set of 33,609 predicted *Fugu* peptides and their mapping on the *Fugu* genome scaffolds was downloaded from the *Fugu* Genome Project (v3 assembly; ftp://ftp.jgi-psf.org/pub/JGI_data/Fugu). A data set of 26,309 known human peptides was downloaded from Ensembl (build 34; ftp://ftp.ensembl.org/pub/current_human/data/fasta/pep).

Orthology relationships between human and *Fugu* genes were determined as described by Aparicio et al. (2002). Briefly, the *Fugu* proteins were searched against the human proteins and vice versa using BLASTP (BLOSUM62; e-score ≤ 1e-07; identity ≥ 30%). The reciprocal best hits (9156) were taken as the likely orthologs.

The coding sequence (CDS) coordinates of human genes on the human genome were downloaded from the UCSC Genome Browser database (http://genome.ucsc.edu/goldenPath/hg16/database/knownGene.txt). The coding sequence (CDS) coordi-

Downloaded from genome.cshlp.org on June 15, 2026 . Published by Cold Spring Harbor Laboratory Press

nates of mouse genes on the mouse genome were downloaded from the UCSC Genome Browser database (http://genome.ucsc.edu/goldenPath/mm4/database/knownGene.txt).

Orientations (same-strand, opposite-strand) and distances (between CDS coordinates) between consecutive pairs of genes were calculated using the above mappings on the *Fugu*, mouse, and human genomes. Human consecutive pairs with orthology in *Fugu* (2737) were extracted for subsequent analysis. Orthologous pairs with the same orientation (i.e., consecutive both in human and *Fugu*) were regarded as conserved pairs (453 gene pairs; see Fig. 2 for illustration). The two data sets (of 2737 and 453 pairs) are given as Supplemental Tables 1 and 2, respectively.

Similar analyses were conducted for the chicken genome, using 28,416 predicted chicken peptides that were downloaded from ftp://ftp.ensembl.org/pub/current_chicken/data/fasta/pep/.

For the Antisensor analysis, human ESTs and cDNAs were obtained from NCBI GenBank version 136 (www.ncbi.nlm.nih.gov/dbEST) and aligned to the human genome build 32 (April 2003) (www.ncbi.nlm.nih.gov/genome/guide/human), as described in Sorek et al. (2002). Sense–antisense pairs were identified using the same methods described in Yelin et al. (2003).

The same process was performed with the mouse data: ESTs and cDNAs from NCBI GenBank version 136 (www.ncbi.nlm.nih.gov/dbEST) and build 30 (February 2003) of the mouse genome (www.ncbi.nlm.nih.gov/genome/guide/mouse).

To link between the human and mouse data sets, we used the HomoloGene database of orthologous loci (www.ncbi.nlm.nih.gov/HomoloGene/). Cases in which a locus in the human genome was assigned two or more orthologous loci in the mouse genome, or vice versa, were discarded from the final set of orthologous loci.

Acknowledgments

We thank Amit Novik, Yaron Kinar, and Ami Haviv for algorithmic support, and Eli Eisenberg, Erez Levanon, Amir Toporik, Tali Lailah, and Galit Rotman for useful discussion.

References

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J.M., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A., et al. 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297:** 1301–1310.

Batt, D.B., Luo, Y., and Carmichael, G.G. 1994. Polyadenylation and transcription termination in gene constructs containing multiple tandem polyadenylation signals. *Nucleic Acids Res.* **22:** 2811–2816.

Boi, S., Solda G., and Tenchini, M.L. 2004. Shedding light on the dark side of the genome: Overlapping genes in higher eukaryotes. *Curr. Genomics* **5:** 509–524.

Brantl, S. 2002. Antisense-RNA regulation and RNA interference. *Biochim. Biophys. Acta* **1575:** 15–25.

Connelly, S. and Manley, J.L. 1988. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes & Dev.* **2:** 440–452.

Han, J.S., Szak, S.T., and Boeke, J.D. 2004. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* **429:** 268–274.

Hastings, M.L., Ingle, H.A., Lazar, M.A., and Munroe, S.H. 2000. Post-transcriptional regulation of thyroid hormone receptor expression by *cis*-acting sequences and a naturally occurring antisense RNA. *J. Biol. Chem.* **275:** 11507–11513.

Hurst, L.D., Pal, C., and Lercher, M.J. 2004. The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* **5:** 299–310.

Iseli, C., Stevenson, B.J., de Souza, S.J., Samaia, H.B., Camargo, A.A., Buetow, K.H., Strausberg, R.L., Simpson, A.J., Bucher, P., and Jongeneel, C.V. 2002. Long-range heterogeneity at the 3′ ends of human mRNAs. *Genome Res.* **12:** 1068–1074.

Kent, W.J., Baertsch, R., Hinrichs, A., Miller, W., and Haussler, D. 2003. Evolution's cauldron: Duplication, deletion, and rearrangement in the mouse and human genomes. *Proc. Natl. Acad. Sci.* **100:** 11484–11489.

Kiyosawa, H., Yamanaka, I., Osato, N., Kondo, S., and Hayashizaki, Y. 2003. Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Res.* **13:** 1324–1334.

Knee, R. and Murphy, P.R. 1997. Regulation of gene expression by natural antisense RNA transcripts. *Neurochem. Int.* **31:** 379–392.

Kumar, M. and Carmichael, G.G. 1998. Antisense RNA: Function and fate of duplex RNA in cells of higher eukaryotes. *Microbiol. Mol. Biol. Rev.* **62:** 1415–1434.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409:** 860–921.

Lavorgna, G., Dahary, D., Lehner, B., Sorek, R., Sanderson, C.M., and Casari, G. 2004. In search of antisense. *Trends Biochem. Sci.* **29:** 88–94.

Lehner, B., Williams, G., Campbell, R.D., and Sanderson, C.M. 2002. Antisense transcripts in the human genome. *Trends Genet.* **18:** 63–65.

Ogawa, Y. and Lee, J.T. 2002. Antisense regulation in X inactivation and autosomal imprinting. *Cytogenet. Genome Res.* **99:** 59–65.

Peters, N.T., Rohrbach, J.A., Zalewski, B.A., Byrkett, C.M., and Vaughn, J.C. 2003. RNA editing and regulation of *Drosophila* 4f-rnp expression by sas-10 antisense readthrough mRNA transcripts. *RNA* **9:** 698–710.

Prescott, E.M. and Proudfoot, N.J. 2002. Transcriptional collision between convergent genes in budding yeast. *Proc. Natl. Acad. Sci.* **99:** 8796–8801.

Rougeulle, C. and Heard, E. 2002. Antisense RNA in imprinting: Spreading silence through Air. *Trends Genet.* **18:** 434–437.

Shendure, J. and Church, G.M. 2002. Computational discovery of sense–antisense transcription in the human and mouse genomes. *Genome Biol.* **3:** RESEARCH0044.

Sorek, R., Ast, G., and Graur, D. 2002. *Alu*-containing exons are alternatively spliced. *Genome Res.* **12:** 1060–1067.

Wagner, E.G. and Simons, R.W. 1994. Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* **48:** 713–742.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420:** 520–562.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., et al. 2003. Widespread occurrence of antisense transcription in the human genome. *Nat. Biotechnol.* **21:** 379–386.

Web site references

http://www.ncbi.nlm.nih.gov/dbEST/; Expressed Sequence Tags database Web site.

http://www.ncbi.nlm.nih.gov/genome/guide/human/; Human Genome Resources Web site.

http://www.ncbi.nlm.nih.gov/genome/guide/mouse/; Mouse Genome Resources Web site.

http://www.ncbi.nlm.nih.gov/RefSeq/; NCBI Reference Sequences home page.

http://www.ncbi.nlm.nih.gov/HomoloGene/; NCBI Orthologs database Web site.

http://www.labonweb.com/antisense/; LEADS-Antisensor Web site.

http://genome.ucsc.edu/; UCSC genome browser home page.

ftp://ftp.jgi-psf.org/pub/JGI_data/Fugu; *Fugu* Genome Project home page.

Received October 3, 2004; accepted in revised form December 13, 2004.