Synteny Conservation and Chromosome Rearrangements During Mammalian Evolution

Jason Ehrlich,*† David Sankoff† and Joseph H. Nadeau*‡

*Jackson Laboratory, Bar Harbor, Maine 04609 and Centre de recherches mathématiques, Université de Montréal, Montréal, Québec, H3C 3J7 Canada

Manuscript received December 13, 1996
Accepted for publication June 4, 1997

ABSTRACT

An important problem in comparative genome analysis has been defining reliable measures of synteny conservation. The published analytical measures of synteny conservation have limitations. Nonindependence of comparisons, conserved and disrupted syntenies that are as yet unidentified, and redundant rearrangements lead to systematic errors that tend to overestimate the degree of conservation. We recently developed methods to estimate the total number of conserved syntenies within the genome, counting both those that have already been described and those that remain to be discovered. With this method, we show that ~65% of the conserved syntenies have already been identified for humans and mice, that rates of synteny disruption vary ~25-fold among mammalian lineages, and that despite strong selection against reciprocal translocations, inter-chromosome rearrangements occurred approximately fourfold more often than inversions and other intra-chromosome rearrangements, at least for lineages leading to humans and mice.

Genomes evolve through changes in DNA sequence as a result of nucleotide substitution and through changes in DNA organization because of chromosome rearrangement. The types, rates, and patterns of nucleotide substitution have been documented for many genes and a sound theoretical basis has been developed for analyzing these changes (Sankoff and Kruskal, 1984; Nei, 1987; Li and Graur, 1991; Waterman, 1995). Relatively less evidence is available for chromosome rearrangements and the analytical basis for interpreting these changes is modest (Sankoff, 1993; Hannenhalli et al., 1995). An important method to describe changes in genome organization involves comparing chromosome banding patterns between different but closely related species (Jauch, 1992; Reid et al., 1993; Scherthan et al., 1994; Rettenberger et al., 1995). However, the number of bands that can be discriminated is inherently limited and assessing similarities in banding patterns becomes increasingly ambiguous for distantly related species. Comparative genetic maps are a powerful alternative. By comparing the chromosomal location of homologous genes in different species, the extent of chromosome conservation and rearrangement can be described even for distantly related species (Ohno, 1970; Lundin, 1979; Nadeau and Taylor, 1984; O'Brien et al., 1993; Andersson et al., 1996). Homologous genes are readily identified through cross-hybridization and DNA sequence analysis (Andersson et al., 1996). Moreover, the number of homologous genes that are being identified is large (tens of thousands) relative to the number of chromosome rearrangements that are likely to have occurred during mammalian evolution (hundreds) (Nadeau and Taylor, 1984; Nadeau et al., 1995).

An important problem in comparative genome analysis has been reliably measuring the extent of chromosome rearrangement during evolution. These measures could be used to calculate rates of chromosome rearrangements, to determine whether some lineages are prone to rearrangement, and to test whether some chromosome segments are preferentially conserved. Nadeau and Taylor (1984) developed a measure of linkage conservation based on the genetic length of each conserved segment and the number of homologous genes that have been mapped to each segment. With this measure, it was estimated that since divergence of lineages leading to humans and mice, the rate of lineage disruption has been approximately one event per million years. However, this method is applicable only to species with genetic maps, as are most measures of genome distance (Sankoff et al., 1992; Hannenhalli and Pevzner, 1995). For most species, genes have been assigned to chromosomes but the location of most genes on the chromosome has not been determined, i.e., synteny but not linkage has been established. As a result, methods are needed to exploit these data for understanding questions of genome organization and evolution.
Several measures of synteny conservation have been proposed (Zakharov and Valeev 1988; Zakharov et al. 1992, 1995; Bengtsson et al. 1993; Zakharov 1993), but they have limitations. Disproportionate weight is given to segments in which many genes have been mapped, thereby overestimating the degree of conservation. More importantly, these measures are based only on the observed number of conserved and disrupted syntenies, and do not fully exploit implicit assumptions of uniform randomness both of genomic rearrangement events and of gene discovery patterns. Sankoff and Nadeau (1996) derived methods to estimate the total number of conserved syntenies within the genome, counting both those that have already been described and those that remain to be discovered. In this article, we apply this method to comparative maps for a variety of mammalian species and use these estimates to calculate lineage-specific rates of chromosome rearrangement.

Conserved syntenies among mammalian species: To ensure adequate sample size for rigorous analysis, we selected nine of ~50 species in MGD (Nadeau et al. 1995) with which to quantify the extent of synteny conservation (Figure 1). These species were baboon, cat, chimpanzee, Chinese hamster, cow, human, mink, mouse, and rat. At least 20 homologous genes have been mapped in each pair-wise species comparison. This group represented two major orders (primates and rodents), while also including selected carnivores and artiodactyls.

Synteny refers to the occurrence of two or more genes on the same chromosome, whereas conserved synteny refers to two or more homologous genes that are syntenic in two or more species, regardless of gene order on each chromosome, i.e., synteny but not necessarily gene order is conserved (Figure 2; see also Nadeau 1989). Conserved linkage pertains to the conservation of both synteny and order of homologous genes between species (Figure 2; see also Nadeau 1989). A disrupted synteny refers to circumstances where a pair of genes are located on the same chromosome in one species but their homologues are located on different chromosomes in another species, i.e., the genes are syntenic in only one of the two species. Syntenic genes can be identified by examining published genetic maps and conserved segments can be identified by comparing

### Table: Conserved Syntenies among Mammalian Species

<table>
<thead>
<tr>
<th>Species</th>
<th>baboon</th>
<th>Capuchin monkey</th>
<th>cat</th>
<th>chimp</th>
<th>Chinese hamster</th>
<th>cow</th>
<th>deermouse</th>
<th>dog</th>
<th>giraffe</th>
<th>gorilla</th>
<th>Human</th>
<th>lemur</th>
<th>marmoset</th>
<th>mink</th>
<th>mouse</th>
<th>orangutan</th>
<th>pig</th>
<th>rabbit</th>
<th>rhesus</th>
<th>sheep</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48</td>
<td>19</td>
<td>26</td>
<td>32</td>
<td>25</td>
<td>34</td>
<td>2</td>
<td>23</td>
<td>24</td>
<td>24</td>
<td>23</td>
<td>47</td>
<td>21</td>
<td>22</td>
<td>29</td>
<td>40</td>
<td>20</td>
<td>14</td>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td>Conserved</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 1.**—Numbers of genes examined in each species (on diagonal), number of homologous genes between species (above diagonal), number of conserved syntenies (below diagonal). Data were derived from Mouse Genome Database (MGD) as posted in March 1995. Guidelines for identifying homologous loci have been established by the Human Genome Database (GDB) and HUGO Comparative Mapping Committees (Anderson et al. 1996). The MGD contains information about homologous genes for nearly 50 mammalian species. From the gene list for each species, mitochondrial genes were excluded, as were autosomes. Including these data would inflate measures of synteny conservation. Including these data would inflate measures of synteny conservation.
maps that show homologous genes in different species (Lundin 1979; Nadeau and Taylor 1984; O’Brien et al. 1993; Lyons et al. 1997).

Measures of synteny conservation and disruption: The Bengtsson (Bengtsson et al. 1993) and Zakharov (Zakharov et al. 1992, 1993) measures of synteny distance focus on the observed syntenies without considering how the statistical regularities in these observations are generated by randomness both in the process of genome rearrangement and the discovery and attribution of genes to chromosomes. Rather than just normalizing the observed number of conserved syntenies by the number of syntenies within the two species being compared, it is preferable to estimate the total number of conserved syntenies given the observed number of conserved syntenies and how many genes they contain. The estimated number of conserved segments should then be a direct measure of the total number of disruptions, without recourse to normalization.

As a probabilistic model according to which we constructed our estimator, we assumed that evolutionary disruptions resulted in \( n + c \) conserved segments in the genome of a species, separated by \( n \) breakpoints scattered at random (i.e., according to the uniform distribution) along the lengths of the \( c \) chromosomes, and that the \( m \) mapped genes were also distributed at random, independently of the location of the breakpoints. Observed lengths of conserved linkages between human and mouse support these assumptions (Nadeau and Taylor 1984; Nadeau 1989).

It is a property of the uniform distribution on the unit interval that if we draw \( k \) independent samples, then the probability density of the length \( x \) of the interval between any two adjacent sample points is

\[
p(x) = k(1-x)^{k-1}.
\]

Setting \( k = n + c - 1 \), the probability density of segment length in our model is then

\[
p(x) = (n + c - 1)(1-x)^{n+c-2},
\]

where \( x \) is the proportion of total chromosomal length in the genome represented by a segment, \( n \) is the number of breakpoints resulting from chromosome rearrangements, and \( c - 1 \) is the number of "artificial" breakpoints introduced by concatenating the \( c \) chromosome for analytical purposes. Analysis of chromosome lengths in many species shows that treating the concatenation points as also drawn independently from the uniform distribution is a reasonable assumption (Sankoff et al. 1996).

Now, for a segment of length \( x \), the probability that a gene randomly placed on the genome will be found in that segment is just \( x \), because we consider the total genome to be of length 1, and the probability that \( r \) randomly and independently located genes out of \( m \) observed genes will fall in the segment is just the binomial probability

\[
B(m,x;r) = \binom{r}{x} x^r (1-x)^{m-r}.
\]

Then the probability \( P(r) \) of finding \( r \) genes in any segment is found by integrating this binomial probability times the density function \( p \) over all possible segment lengths. This gives

\[
P(r) = \frac{(n + c - 1) m! (n + c + m - r - 2)!}{(n + c - 1 + m)! (m - r)!}
\]

for \( r = 0, 1, \ldots, m \). Given this theoretical distribution, our knowledge of \( m, c \), and the observed frequency distribution of \( r \) for nonzero values of \( r \), we could then approximate the maximum likelihood estimator of \( n \) (Sankoff and Nadeau 1996).

To illustrate this method, we evaluated data on 1152 homologous autosomal genes in humans and mice, observed to be distributed among 91 conserved syntenies. As a first approximation, we equate one conserved segment to one conserved synteny, although this is somewhat of an underestimate. Based on the formula for \( P(r) \) and applying maximum likelihood methods (Sankoff and Nadeau 1996), we estimate \( n = 122 \) breakpoints. The observed frequency distribution for the number of genes per conserved synteny is compared in Figure 3 with the theoretical distribution \( P \) based on \( n = 122 \) breakpoints and \( c = 19 \) autosomes. The estimated total number of conserved syntenies is 141. Given that 91 have already been identified, we estimate an additional 51 remain to be found, a result suggesting that ~65% (~91/141) of the conserved syntenies have already been identified for humans and mice.

Phylogenies: Phylogenetic analysis was used to evaluate data quality. If homologies and genetic map localizations are correctly determined, the phylogenetic tree based on synteny distance should be similar to the known evolutionary relations among these nine species. A phylogenetic tree based on estimates of \( n \) was constructed using the neighbor-joining (NJ) method as implemented in PHYLIP version 3.54c (Figure 4). Other tree-building algorithms such as UPGMA and Fitch-Margoliash/Least-Squares, as well as other measures of synteny distance (Zakharov et al. 1992, 1995; Bengtsson et al. 1993; J. Ehrlich, D. Sankoff and J. H. Nadeau, unpublished results), were also tested and the results were highly consistent (not shown).

It should be noted that our method underestimates the number of conserved syntenies yet to be found for species in which few genes have been mapped. As a result, our model must be modified to take into account that at early stages of mapping, experimentalists tend to report more genes that are linked in two or more species than would be predicted from our assumptions of randomness.
**Rates of synteny disruption:** We next calculated rates of synteny disruption for selected lineages and examined these rates for homogeneity. The analysis partitions the total synteny distance separating two species into the distance since divergence for each lineage (SARICH and WILSON 1967). At least three pair-wise distances are needed to estimate the three lineage-specific rates. This analysis does not require knowledge of divergence times, which are generally not known accurately (ROMER 1966).

We define the following relations and terms:

\[ R_{ref} + R_h = N_{ref,h}, \]
\[ R_{ref} + R_m = N_{ref,m}, \]
\[ R_h + R_m = N_{h,m}, \]

where \( N_{ref,h} \), \( N_{ref,m} \) and \( N_{h,m} \) represent the synteny distance between pairs of species, and \( R_{ref}, R_h \) and \( R_m \) designate the lineage-specific number of synteny disruptions for reference species (cat, cow or mink), human and mouse, respectively.

We assumed that the NJ tree (Figure 4) was correct and used it, rather than the raw estimates of number of conserved syntenies between pairs of species, as a source of distance estimates of \( N \); the tree structure imposes a degree of consistency on the distances and reduces statistical fluctuations. To calculate synteny distance for a species pair, the estimated number of synteny disruptions was summed from the branch tip for one species to the branch tip for the other. The cat lineage had the lowest rate of disruption; chimp, cow and mink had intermediate rates, and humans and mice showed had the highest rates (Table 1). Rates for lineages leading to human, mouse and rat did not differ substantially. It is important to emphasize that these comparisons are not independent, cannot be tested for statistical significance, and should be considered as qualitative generalizations only.

To test whether the variable number of genes used in the comparisons contributed to the heterogeneous rates, we replaced mouse with rat as the representative rodent and human with chimp as the representative primate (Table 1). The cat lineage still showed the lowest rate of synteny disruption; chimp, cow and mink lineages showed intermediate rates, and the rat lineage showed a high rate. Thus sample size did not bias rate estimates.

Finally, we examined rates of disruption for the three primate species and separately for the three rodent species (Table 1). Baboon and chimp lineages showed comparable rates of disruption; the rate for the human lineage was approximately twice the rate for these other primates. Among rodents, the rat and mouse lineages showed comparable rates, whereas the hamster showed a low rate of disruption that was similar to the rate for the cat lineage. Thus rates varied considerably even within the primate and rodent lineages.

**DISCUSSION**

A concern in the use of comparative maps for genome analysis is the modest number of genes that have been mapped in most species and the disproportionate influence that errors might consequently have. A possible example of this problem is the apparent excess of synteny segments marked by a single gene: ~16 are expected and 35 are observed (Figure 3). Explanations for this excess include incorrect homology determination, mistaken synteny assignment, or questionable assumptions in the estimation procedures (SANKOFF and NADEAU 1996). To further illustrate the nature of the problem, consider the following. In April 1996, only a single gene marked 28 of the 100 conserved syntenies in the mouse and human comparative map (MGD, http://www.informatics.jax.org). By August 1996, five of these genes had been removed from the mouse or human genetic map, four genes had been reassigned to another chromosome in one or both species, two segments were confirmed by the addition of a new gene, and six new single gene segments had been identified. These data support the experience of the mapping community that apparent conserved syntenies marked by a single gene are suspect; some may be subsequently confirmed with additional data, and some may be wrong. The apparent excess of single gene segments (Figure 3) is therefore not surprising. More compelling is the excellent fit between the observed and predicted number of segments marked by two or more genes.

Despite the inherent ambiguities in some gene mapping data, three arguments suggest that most mapping and homology data are largely correct and that, while improvements might be made, estimation procedures are robust to a modest gene mapping error rate and to small sample sizes. The first is that the phylogenetic trees (Figure 4 and not shown) based on various measures of synteny distance gave similar results, each of which reflected expectations. If synteny distances were based on too small sample sizes, or if many errors existed in the database, the resulting trees should have been unreliable. The second argument is that replacing “map rich” species (humans and mice) with “map poor” species (chimps and rats) did not substantially influence estimates of lineage-specific rates of synteny disruption (Table 1). Finally, experience with conserved linkages suggests that measures of genome rearrangement may be remarkably good even with small sample sizes. The original estimate for the average length of all autosomal linkages that have been conserved in the genomes of mouse and human was 8.1 ± 1.6 CM (NADEAU and TAYLOR 1984). This estimate was based on 85 genes and 13 conserved linkages. Nine
**A. Genetic map in reference species**

- **Gene arrangement:**
  
  Chromosome 1 (Chr 1) (A, B, C, D, E, F)

- **Definition:** Genes that are located on the homologous chromosome; order of genes is not known or is ignored for analytical purposes.

- **Count:**
  - One conserved synteny involving genes A, B, C, D, E, F.

**Possible causes:**
- No inter-chromosomal rearrangement.
- No evidence for intra-chromosomal rearrangement.

**B. No genetic map in comparison species**

- **Disrupted synteny**
  
  Gene arrangement:
  
  Chromosome 1 (Chr 1) (A, B, C, D)  Chromosome 2 (Chr 2) (E, F)

- **Definition:** Genes that are on the same chromosome in one species are on different chromosomes in another species.

- **Count:**
  - One disrupted synteny involving genes A, B, C, D vs genes E, F;
  - One conserved synteny involving genes A, B, C, D;
  - One conserved synteny involving genes E, F.

**Possible causes:**
- An inter-chromosomal rearrangement, e.g. reciprocal translocation.
- No evidence for intra-chromosomal rearrangement.

**C. Genetic map in comparison species**

- **Conserved synteny and linkage**
  
  Gene arrangement:
  
  A - B - C - E - F

- **Definition:** Same gene order and similar genetic distances.

- **Count:**
  - One conserved linkage involving genes A, B, C, D.

**Possible causes:**
- No inter-chromosomal rearrangement.
- No intra-chromosomal rearrangement.

- **Conserved synteny, disrupted synteny, conserved linkage, disrupted linkage**
  
  Gene arrangement:
  
  A - B - C - D - E - F

- **Count:**
  - One conserved linkage involving genes B, C, D;
  - One disrupted linkage involving genes E, F vs A.
  - One conserved synteny involving genes A, B, C, D, E, F.

**Possible causes:**
- An inter-chromosomal rearrangement, such as a paracentric inversion.

---

**FIGURE 2.**—Definition of conserved and disrupted syntenies and linkages. Examples are given for a chromosome marked by genes A, B, C, D, E, and F. Spacing reflects approximate recombination distances. In the comparative map (parts A-C), one of the species is referred to as the reference species, the other as the comparison species. In the comparison species, two chromosomes are illustrated, Chr 1 and Chr 2. Conserved linkages are marked by heavy black lines. Count: the number of conserved or disrupted syntenies and linkages for each example.
years later, and with ~11-fold more data, the average length was 8.8 cM (Copeland et al. 1993). These correspond to 178 ± 39 and 144 chromosome rearrangements, respectively. Thus modest genetic maps and possible mapping errors do not seem to influence results dramatically.

Given rates of linkage and synteny disruption, occurrence of inter- vs. intra-chromosome rearrangements can be characterized. Both intra-chromosomal rearrangements, e.g., inversions, and inter-chromosomal rearrangements, e.g., reciprocal translocations, disrupt linkage, whereas only inter-chromosomal rearrangements disrupt synteny. We estimated that ~140 synteny disruptions and ~180 linkage disruptions (Nadeau and Taylor 1984; Nadeau 1989; Copeland et al. 1993) have occurred since divergence of lineages leading to humans and mice. The difference between these two numbers is ~40 intra-chromosome rearrangements. For these two species, inter-chromosome rearrangements have been fixed during evolution approximately fourfold (~140 vs. ~40) more often than intra-chromosome rearrangements. It will be interesting to determine whether a similar bias is found for other mammalian lineages.

The ability to use synteny maps for estimating rates of chromosome rearrangements for many mammalian lineages provides an opportunity to study lineage-specific rates of chromosome rearrangement. An important limitation of the method proposed by Nadeau and Taylor (1984) is that it depends on the availability of genetic maps. Although rapid progress is being made in developing genetic maps for several mammalian species, most maps still have more synteny assignments than linkage assignments (Andersson et al. 1996). An important consequence is that conserved linkages cannot be used to calculate lineage-specific rearrangement rates. Synteny data partially resolve this problem because distances can be calculated for many pairs of species (Figure 1).

We found that rates of synteny disruption differ substantially among mammalian lineages, with cats and hamster showing the slowest rates (range 3.1–4.2), humans, mice and rats showing the highest rates (range 48.8–79.9), and baboons, chimps, cow, and mink showing intermediate rates (range 14.7–25.1). These rates correspond to ~0.05, 0.90, and 0.30 synteny disruptions per million years, respectively. Zakarov (1993) also found substantial variation in rates of synteny disruption among mammalian lineages. The measures of synteny conservation that were used have statistical limitations, however (Sankoff and Nadeau 1996). Various cytogenetic methods have previously shown variability in chromosome rearrangements among lineages (Wilson et al. 1974, 1977; Bush et al. 1977). The advantage of comparative maps for this purpose is that the data have high resolution, the analysis is based on a mathematical model, and the results are quantitative. With these rate estimates, it will now be possible to construct and evaluate more detailed models of the patterns of chromosome rearrangement during mammalian evolution.

Many important problems remain to be resolved in using comparative maps for genome analysis. Examples include using comparative mapping information for inferring the organization of the "primordial" mammalian genome (Ferretti et al. 1996), defining segment identities (Sankoff et al. 1997) and developing methods for quantitating comparative mapping information resulting from chromosome painting (Jauch 1992; Reid et al. 1993; Scherthan et al. 1994; Rettenberger et al. 1995).

**Figure 3.** Distribution of the number of genes per conserved synteny. Two data points for conserved syntenies with 62 and 64 genes are not on scale. There is no data point for conserved syntenies with zero genes, but the random disruption model predicts 15 or 29, depending on whether the total number of conserved syntenies is 141 (heavy line), the maximum likelihood estimate, or 200 (thin line), an arbitrarily chosen number of the same order of magnitude, illustrating the resolution of the method. Note the apparent empirical excess of conserved syntenies with just one gene. This excess could result from incorrect homology determinations or mistaken linkage assignments. Note also the excellent fit between the observed and predicted numbers of segments marked by two or more genes.

**Figure 4.** Neighbor-joining tree based on n, the estimated number of conserved syntenies.
# Synteny Conservation

## TABLE 1

Rates of synteny disruption among representative mammalian species

<table>
<thead>
<tr>
<th>Synteny distance</th>
<th>Lineage-specific rate</th>
<th>Relative rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>cat-human-mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cat-human:</td>
<td>57.7</td>
<td>cat: 4.2</td>
</tr>
<tr>
<td>cat-mouse:</td>
<td>79.4</td>
<td>human: 53.5</td>
</tr>
<tr>
<td>human-mouse:</td>
<td>128.7</td>
<td>mouse: 75.2</td>
</tr>
<tr>
<td>cow-human-mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cow-human:</td>
<td>73.9</td>
<td>cow: 25.1</td>
</tr>
<tr>
<td>cow-mouse:</td>
<td>105.0</td>
<td>human: 48.8</td>
</tr>
<tr>
<td>human-mouse:</td>
<td>128.7</td>
<td>mouse: 79.9</td>
</tr>
<tr>
<td>human-mink-mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>human-mink:</td>
<td>76.9</td>
<td>human: 59.3</td>
</tr>
<tr>
<td>human-mouse:</td>
<td>128.7</td>
<td>mink: 17.6</td>
</tr>
<tr>
<td>mink-mouse:</td>
<td>87.0</td>
<td>mouse: 69.4</td>
</tr>
<tr>
<td>Test for choice of reference species: cat, cow or mink vs human and mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cat-human:</td>
<td>57.7</td>
<td>mouse:cat 17.9</td>
</tr>
<tr>
<td>cow-human:</td>
<td>73.9</td>
<td>mouse:cow 3.2</td>
</tr>
<tr>
<td>human-mink:</td>
<td>76.9</td>
<td>mouse:mink 3.9</td>
</tr>
<tr>
<td>Test for sample size effect: replacing rat for mouse and chimp for human</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cat-human-rat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cat-human:</td>
<td>57.7</td>
<td>cat: 4.2</td>
</tr>
<tr>
<td>cat-rat:</td>
<td>69.1</td>
<td>human: 53.5</td>
</tr>
<tr>
<td>human-rat:</td>
<td>118.4</td>
<td>rat: 64.9</td>
</tr>
<tr>
<td>cow-human-rat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cow-human:</td>
<td>78.6</td>
<td>human: 51.2</td>
</tr>
<tr>
<td>human-rat:</td>
<td>118.4</td>
<td>cow: 27.5</td>
</tr>
<tr>
<td>rat:</td>
<td>94.7</td>
<td>rat: 67.3</td>
</tr>
<tr>
<td>mink-human-rat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mink-human:</td>
<td>76.9</td>
<td>human: 59.3</td>
</tr>
<tr>
<td>human-rat:</td>
<td>118.4</td>
<td>mink: 17.6</td>
</tr>
<tr>
<td>rat:</td>
<td>76.7</td>
<td>rat: 59.1</td>
</tr>
<tr>
<td>cat-chimp-mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cat-chimp:</td>
<td>27.6</td>
<td>cat: 4.2</td>
</tr>
<tr>
<td>cat-mouse:</td>
<td>76.3</td>
<td>chim: 23.4</td>
</tr>
<tr>
<td>chim-mouse:</td>
<td>95.5</td>
<td>mouse: 72.1</td>
</tr>
<tr>
<td>cow-chimp-mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chimp-cow:</td>
<td>48.5</td>
<td>chim: 21.1</td>
</tr>
<tr>
<td>chimp-mouse:</td>
<td>95.5</td>
<td>cow: 27.5</td>
</tr>
<tr>
<td>cow-mouse:</td>
<td>101.9</td>
<td>mouse: 74.5</td>
</tr>
<tr>
<td>mink-chimp-mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chimp-mink:</td>
<td>46.8</td>
<td>chim: 29.2</td>
</tr>
<tr>
<td>chimp-mouse:</td>
<td>95.5</td>
<td>mink: 17.6</td>
</tr>
<tr>
<td>mouse:</td>
<td>83.9</td>
<td>mouse: 66.3</td>
</tr>
<tr>
<td>Variation within mammalian orders: primates and rodents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>baboon-chimp-human</td>
<td></td>
<td></td>
</tr>
<tr>
<td>baboon-chimp:</td>
<td>32.7</td>
<td>baboon: 14.7</td>
</tr>
<tr>
<td>baboon-human:</td>
<td>51.0</td>
<td>chim: 18.0</td>
</tr>
<tr>
<td>chim-human:</td>
<td>36.3</td>
<td>human: 36.3</td>
</tr>
<tr>
<td>hamster-mouse-rat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hamster-mouse:</td>
<td>45.7</td>
<td>hamster: 3.1</td>
</tr>
<tr>
<td>hamster-rat:</td>
<td>38.5</td>
<td>mouse: 42.6</td>
</tr>
<tr>
<td>mouse-rat:</td>
<td>78.0</td>
<td>rat: 35.4</td>
</tr>
<tr>
<td>baboon-chimp:</td>
<td></td>
<td>human:baboon 2.5</td>
</tr>
<tr>
<td>baboon-human:</td>
<td></td>
<td>human:chimp 2.0</td>
</tr>
<tr>
<td>chim-human:</td>
<td></td>
<td>chimp:baboon 1.2</td>
</tr>
<tr>
<td>hamster-mouse:</td>
<td></td>
<td>mouse:hamster 13.7</td>
</tr>
<tr>
<td>hamster-rat:</td>
<td></td>
<td>rat:hamster 11.4</td>
</tr>
<tr>
<td>mouse-rat:</td>
<td></td>
<td>mouse:rat 1.2</td>
</tr>
</tbody>
</table>
We thank Joe Felsenstein for providing the PHYLIP software package and Alex Smith for help in designing some of the computer programs used here. Part of this work was completed at the Jackson Laboratory. This work was supported by National Institutes of Health grant HG-00390, grant BIR / 00379 from the National Science Foundation Research Experiences for Undergraduates Program, a grant from the Burroughs Wellcome Fund to the Jackson Laboratory, grants from the Natural Sciences and Engineering Research Council of Canada and the Canadian Genome Analysis and Technology program. D.S. is a fellow of the Canadian Institute for Advanced Research.

LITERATURE CITED

ANDERSSON, L., A. ARCHIBALD, M. ASHIBURNER, S. AUDUN, W. BAR- 
ENDZEK ET AL., 1996 Comparative genome organization of verte- 

BENGSTSSON, B. O., K. K. LEVAN and G. LEVAN 1993 Measuring gene 

BUSH, G. L., S. M. CASE, A. C. WILSON and J. L. PATTON 1977 Rapid 
Acad. Sci. USA 74: 3942–3946.

COPELAND, N. G., N. A. JENKINS, D. J. GILBERT, J. T. EPPIG, L. M. 
JALTAIS ET AL., 1993 A genetic linkage map of the mouse: current 

DISTCHET, C. M.; C. I. BRANNAN, A. LARSEN, D. A. ADLER, D. F. SCHORD- 
ERET ET AL., 1992 The human pseudoautosomal GM-CSF receptor 

FERRETTI, V., J. H. NADEAU and D. SANKOFF, 1996 Original synteny, 
pp. 159–167, in Combinatorial Pattern Matching, 7th Annual Symp- 
sium, Lecture Notes in Computer Science, edited by D. HIRSCHBERG 
and G. MEYERS. Springer-Verlag, New York.

HANNENHALLI, S., and P. A. PEVZNER 1995 Transforming men into 
178–189.

HANNENHALLI, S., C. CHAPPY, E. V. KOONIN, and P. A. PEVZNER 1995 
Genome sequence comparison and scenarios for gene rearrange- 

JAUCH, A., 1992 Reconstruction of genomic rearrangements in great 
USA 89; 8611–8615.

Sinauer Associates, Sunderland, MA.

LUINING, L.-G. 1979 Evolutionary conservation of large chromosomal 
segments reflected in mammalian gene maps. Clin. Genet. 16: 
72–81.

LYONS, L. A., T. F. LAUGHLIN, N. G. COPELAND, N. A. JENKINS, J. E. 
WOMACK ET AL., 1997 Comparative anchor tagged sequences 
(CATS) for integrative mapping of mammalian genomes. Nat. 

NADEAU, J. H., 1989 Maps of linkage and synteny homologies be- 
 tween mouse and man. Trends Genet. 5: 82–86.

NADEAU, J. H., and B. A. TAYLOR, 1984 Lengths of chromosomal 
segments conserved since divergence of man and mouse. Proc. 

NADEAU, J. H., P. L. GRANT, S. MANKALA, A. H. REINER, J. E. RICHARD- 
SON ET AL., 1995 A Rosetta Stone for mammalian genetics. Na- 
ture 373: 363–365.

NEI, M., 1987 Molecular Evolutionary Genetics. Columbia University 
Press, New York.

et al., 1995 Anchored reference loci for comparative genome 

OHNO, S. 1970 Evolution by Gene Duplication. Springer-Verlag, New- 
York.

PALMER, S., J. PERRY and A. ASHWORTH 1995 A contravention of 

REID, T., N. ARNOLD, D. C. WAARD and J. WEINBERG 1993 Compari- 
tive high-resolution mapping of human and primate chromo- 
somes by fluorescence in situ hybridization. Genomics 18; 381– 
386.

REITENBERGER, G., C. KLETT, U. ZECHNER, J. KUNZ, W. VOGEL ET 
al., 1995 Visualization of the conservation of synteny between 
humans and pigs by heterologous chromosome painting. Geno- 

ROMER, A. S., 1966 Vertebrate Paleontology. University of Chicago 
Press, Chicago.

RUGARI, F. L., D. A. ADLER, G. BORSANI, K. TSUCHIYA, B. FRANCO ET 
al., 1995 Different chromosomal localization of the Cln6 gene in 
Mus spretus and C57BL/6J. Nat. Genet. 10: 466–471.

SANKOFF, D., 1993 Analytical approaches to genomic evolution. Bio- 
chemistry 75: 409–415.

SANKOFF, D., V. FERRETTI and J. H. NADEAU 1997 Conserved seg- 
ment identification. RECOMB 97, Proceedings of the First An- 
nual International Conference on Computational Molecular Bi- 

SANKOFF, D., and J. B. KRUSKAL (EDITORS), 1984 Time Warps, String 
Edits and Macromolecules. Addison Wesley, Reading, MA.

SANKOFF, D., G. LEDUC, N. ANTOINE, B. PAQUIN, B. F. LANG ET AL., 
6575–6579.

SANKOFF, D., and J. H. NADEAU 1996 Conserved synteny as a mea- 
257.

SARICH, V. M., and A. C. WILSON 1967 Immunological time scale for 

SCHERTHAN, H., T. GREMER, U. ARNASON, H. U. WEIER, A. LIMA-DE- 
FARIA ET AL., 1994 Comparative chromosome painting discloses 
homologous segments in distantly related mammals. Nature 
Genet. 6: 342–347.

WATERMAN, M. 1995 Introduction to Computational Biology. Chapman 
Hall, New York.

WILSON, A. C., V. M. SARICH and L. R. MAXSON 1974 The impor- 
tance of gene rearrangement in evolution: evidence from studies 
on rates of chromosomal, protein and anatomical evolution. 

Molecular evolution and cytogenetic evolution, pp. 375–393 in 
Molecular Human Cytogenetics, ICN-UCLA Symposia 
Molecular and Cellular Biology, Vol. VII, edited by R. S. 

ZAKHAROV, I. A., 1993 Measurements of similarity of synteny groups 
and an analysis of genome rearrangements in the evolution of 
mammals, pp. 107–113 in Bioinformatics, Supercomputing and Com- 
plicated Genome Analysis, edited by H. Lim, J. Ficket and C. Cantor. 
World Scientific, New York.

ZAKHAROV, I. A., and A. K. VALEEV 1988 Quantitative analysis of 
the mammal genome evolution via the comparison of genetic maps. 

ZAKHAROV, I. A., V. I. NIKIFOROV and E. V. STEPANOV 1992 Holom- 
omy and evolution of gene orders: combination measurement of 
synteny group similarity and simulation of the evolutionary 

ZAKHAROV, I. A., V. I. NIKIFOROV and E. V. STEPANOV 1995 Interval 
estimates of combinatorial measures of similarity for orders of 

Communicating editor: A. G. CLARK.