Comparative genome mapping in mammals: the shrew map

Denis M. LARKIN, Oleg L. SEROV, Pavel M. BORODIN, Natalia S. ZHDANOVA and Jeremy B. SEARLE


Comparative mapping has shown that the genome evolution of eutherian mammals mostly involves movement of large blocks of genes. Zoo-FISH analysis has revealed that 23 to 56 blocks of human chromosome material are conserved in different mammalian genomes. Several syntenic groups of genes are found in most mammalian orders: some involve complete or partial human chromosomes and others involve segments that are found in a fused state in most mammals but separated in the human genome. It can be hypothesized that these conserved syntenic groups were present in the ancestral genome of all eutherian mammals. There is a need to obtain data from as wide a taxonomic spread of mammals as possible to further elucidate this ancestral genome. We have initiated studies on the order Insectivora and have located 45 genetic markers in the common shrew Sorex araneus Linnaeus, 1758. Zoo-FISH analysis shows conservation of human chromosome material in shrews in 32 blocks and that the shrew genome shares many of the conserved chromosome segments found in other mammalian orders.

Institute of Cytology and Genetics, Russian Academy of Sciences, Novosibirsk 630090, Russia, e-mail: lark@bionet.nsc.ru (DML, OLS, PMB, NSZ); Department of Biology, University of York, PO Box 373, York YO10 5YW, U.K. (JBS)

Key words: Sorex araneus, genome map, comparative mapping

In all mammals the genome is basically the same size, comprising approximately \(3.3 \times 10^9\) base pairs and \(7 \times 10^4\) genes (Graves 1998a). Other than genes, mammalian genomes are composed of noncoding, mostly repetitive sequences, of variable quantities, thereby generating the small differences in genome size that do occur among species. The genomes of birds are only about one third the size of those of mammals, and the genome of the pufferfish (Fugu) is especially compact at about one eighth the mammalian genome size (Brenner et al. 1993). It has been suggested that several rounds of genome duplication took place early during the diversification of vertebrates (Ohno 1970, Nadeau and Sankoff 1997).

The equivalent genome is divided differently in different mammalian species. Haploid chromosome numbers range from \(n = 3\) large chromosomes in the Indian muntjac to \(n = 67\) small chromosomes in the black rhinoceros. Because of this variation it was long considered that many rearrangements had mixed gene orders
beyond recognition in different mammalian lineages. However, comparative mapping, which is the assignment of homologous loci to physical and genetic maps in different species, has shown that the mammalian genome is much more conserved than was earlier realised. This conclusion is being constantly reinforced by gene mapping data from new species.

Currently genome maps have been developed for more than 50 mammalian species from 10 orders. The human genome map contains about 7000 genes and 14,000 markers (Human Genome Database, see http://www.gdb.org), the mouse map contains about 6000 genes and 13,000 markers (Davisson et al. 1998). The rat genome map contains about 1000 genes and 4000 markers (Levan et al. 1998). Genome maps of cattle and pig contain approximately 1400 genetic markers each (Animal Genome Mapping Database, see http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/browser.sh). Genome maps of other species contain fewer markers.

The reasons for constructing a genome map are varied. The emerging human genome map is clearly useful in localizing genes important in disease, physiology and development. Mice and rats are animal models whose genome maps also aid in the study of genes that are important from a biomedical perspective. Genetic maps of livestock species are of particular interest to locate economically important traits (quantitative trait loci: QTLs), such as fatness in pigs (Andersson et al. 1994) and milk-yield in cattle (Womack 1998). The genome maps of other mammals have been constructed specifically to aid our understanding of human, mammalian and vertebrate genome evolution. Comparative genome mapping can also make use of the gene maps constructed for biomedical and agricultural purposes, as "wittingly or unwittingly, all comparative geneticists - whether they are working on sheep improvement or development of mutants in zebrafish - are contributing in our knowledge about how the vertebrate genome evolved" (Graves 1998a).

To reconstruct the ancestral mammalian genome, as many mammalian species as possible, representing different orders, should be mapped. To facilitate cross-species comparison, their maps should contain the same set of homologous genetic markers. A set of such markers, including 321 genes containing sequences conserved among humans, cattle and cat and uniformly distributed among human chromosomes, was recommended for comparative mapping (O'Brien et al. 1993). These conservative markers were called "reference anchor loci" (type I markers) as opposed to species-specific microsatellites (type II markers). For "reference anchor loci" consensus PCR primers called "CATS" (Comparative Anchor Tagged Sequences) were designed for mapping in different mammals (Lyons et al. 1997).

A recent development that has greatly aided the process of genome mapping in mammals is the application of the zoo-FISH technique (Jauch et al. 1992). In this method, chromosomal painting probes of one species are hybridised with the metaphase chromosomes of another. The resulting map delineates relatively large regions of chromosomal homology between the two species. Chromosomal painting probes have been developed from human flow-sorted and microdissected chromosomes and have been applied to the metaphase chromosomes of various mammalian...
Comparative genome mapping: the shrew map

species. Zoo-FISH maps have been established for species from several different orders, including Artiodactyla: pigs, cattle, sheep, Indian muntjac (Rettenberger et al. 1995b, Solinas-Toldo et al. 1995, Fronicke and Scherthan 1997, Iannuzzi et al. 1999), Carnivora: cat, American mink, harbor seal (Rettenberger et al. 1995a, Fronicke et al. 1997, Hameister et al. 1997), Perissiodactyla: horse (Raudsepp et al. 1996), Cetacea: dolphin (Bielec et al. 1998), Chiroptera: flower-visiting bat (Volleth et al. 1999), Insectivora: common shrew (Dixkens et al. 1998) and at least 14 primate species (Muller et al. 1996). A zoo-FISH comparison between the genomes of humans and rodents has not been possible because of a large number of rearrangements in the rodent genome in comparison to humans. From gene maps of mouse and humans it appears that there are about 150 segments of conserved synteny (Davisson et al. 1998). It seems that rodents are an unusual order in having highly rearranged genomes in comparison to other mammals. A similar level of genome rearrangements has only been found in some primates (Jauch et al. 1992). All other eutherian mammalian orders that have been examined are more conserved. For instance, human and cat genomes share 32 conserved blocks (that correspond to 10 chromosomal rearrangements which would convert a human karyotype into a cat karyotype) and other carnivores show about the same number (mink 34 and harbor seal 31). Cetaceans show a very similar situation, with 39 autosomal segments conserved between humans and dolphins while artiodactyls are somewhat more distinct: human and bovine genomes share 56 autosomal blocks (corresponding to 34 chromosomal rearrangements). Zoo-FISH studies have also shown that some large regions of the mammalian genome are exceptionally conserved. For example, most of human chromosome 9 is represented by cat chromosome D4, dolphin 6, common shrew bc and mink 9. The same conservation has been found for human chromosomes 13 and 17; the latter is present as a single block of conserved synteny or even a single chromosome in most mammalian species (excluding dogs which have two conserved blocks: Werner et al. 1997). The presence of the same gene blocks in genomes of different species from different orders suggests that this state represents the ancestral genome of all eutherian mammals. Sometimes chromosome material is found in separate locations in humans but combined together in many other mammal orders, such as 3-21, 12-22, 14-15 and 16q-19q (Chowdhary et al. 1998). Likewise the combinations of human chromosomes 2-20, 3-19, 4-8-4 and 18-22 are found in the genomes of cat, mink and harbor seal and it is likely that these chromosome blocks represents elements of the ancestral karyotype of carnivores. Human chromosomes 14-15 and 5-19 are found combined in all Artiodactyla and likewise would presumably have joined together during the ancestry of that order. These various data are shown in a tentative evolutionary scheme (Fig. 1).

It should be noted that the X chromosome appears to be almost completely conserved in all eutherian mammals. Only 3 of hundreds of genes mapped to the human X chromosome are autosomal in the mouse (Graves 1998a). The conservation of the X chromosome in eutherians can be ascribed to selection against disruption
Fig. 1. Schematic representation of the karyotypic fission/fusion events that may have occurred in the evolution from the common eutherian ancestor to twelve species of mammal, which together with a number of primates have been studied by zoo-FISH. The common ancestor is presented at the top. Since all zoo-FISH experiments were performed with the use of human chromosomal painting probes, the numbers represented relate to human chromosomes. In the eutherian ancestor two types of chromosomes are identified: undisrupted segments represent those human chromosomes or chromosomal arms the entire homologues of which were found in at least one of eleven nonprimate species; NSCs (neighbouring segment combinations) represent human chromosomes that were found to be contiguously syntenic in at least five species. From the proposed ancestor the probable fission events (rectangles) and fusion events (ellipses) leading to the karyotypes can be followed along the arrows. Human chromosome 22 material is present as two blocks in most eutherian mammals which are marked as 22* and 22** in the putative ancestral karyotype. As it can be seen, two fused chromosomes (19q-19p and 22*-22**) and separated versions of chromosomes 12 and 22 are common to both shrews and primates and may have been present in a common ancestral form [from Chowdhary et al. (1998) with our additions].
Comparative genome mapping: the shrew map

of the system of dosage compensation (Ohno 1967). On the other hand, X-linked human genes occur as two groups in marsupials: human Xq and pericentric genes map onto the X chromosome in marsupials but distal genes on human Xp are autosomal. It can be suggested that the ancestral X chromosome common to both marsupials and eutherians was small, like that currently found in marsupials, but that the eutherian X gained a new part derived from the autosomes. This suggestion is confirmed by the mapping of human Xp markers in monotremes: they are autosomal in the platypus and echidna (Graves 1998b).

In contrast to the conservation of the X chromosome in eutherians, the Y chromosome has evolved very rapidly, losing and gaining different, mostly repetitive sequences in different species.

Of the various mammalian orders that have been used for gene mapping, we have a particular interest in the Insectivora, and have generated the first gene map for that order involving the common shrew Sorex araneus Linnaeus, 1758. Given that the Insectivora are almost certainly a polyphyletic lineage (Nowak 1991), our data are best considered as representative of the Soricidae, rather than insectivores in general (Fig. 1). Other distinctive insectivores, such as hedgehogs and tenrecs, also need to be studied. At present the shrew gene map contains 42 genes and 3 human YAC probes (Fig. 2). Sixteen of them are “reference anchor loci” recommended for comparative mapping in mammals (O’Brien et al. 1993). Although the whole number of assignments in shrews is rather small in comparison with some other species, it is enough to argue that most conservative blocks of the mammalian genome are preserved in shrews. This contention is strongly supported by a zoo-FISH comparative map of shrew and human genomes, which showed that human chromosomes are found as 32 blocks in shrews (Dixkens et al. 1998). No human homologues were revealed for two elements: the proximal short arm of chromosome go and the short arm of chromosome qr.

The current shrew gene map covers the chromosome de (which is a tandem fusion of the standard eutherian X and an autosome) and 7 out of 9 metacentric autosomes of the Novosibirsk chromosome race, and was established with the use of a shrew-rodent somatic cell hybrid panel (Pack et al. 1995). Thirty-five out of 42 genes mapped are enzymes and it seems that the number of biochemical markers suitable for mapping in shrews is exhausted with this set. The restriction in the number of this type of markers explains why two of the shrew chromosomes (bc and tu) remained unmarked. Only two markers – ornithine transcarbamylase (OTC) and growth hormone (GH) (Malchenko et al. 1996) were assigned by Southern blotting and 5 additional markers have recently been mapped by PCR (Larkin et al. 2000). The nature of the shrew-rodent panel allowed most markers to be mapped to the level of the chromosome. But some clones in the panel contain the arms m or p of chromosome mp and q or r of chromosome qr. Using this opportunity, markers PEPB and ACY1 were mapped to the m arm of chromosome mp (Matyakhina et al. 1996) and GOT1 to the q arm of chromosome qr (Matyakhina et al. 1997). To make the shrew genome map more useful for comparative analysis, the distribution of the
rest of the markers need to be mapped to the level of chromosomal arms. Currently the only way to reach this goal is by the comparison of the gene and zoo-FISH maps of common shrew. This comparison shows that most localizations are in complete agreement but 10 diverge (Fig. 2). These divergent cases are presumably explained by small, undetected, translocations of chromosomal material, and indicate that the true number of gene blocks in shrews that are conserved relative to humans are best viewed as 32 large blocks plus an unknown number of small blocks.

As can be seen from the map (Fig. 2), the shrew genome has many of the conserved gene blocks found in most mammalian orders. These are the whole human chromosomes: 9 (part of c), 13 (part of b) and 17 (part of h), and the fused chromosomes: part 3-21 (part of α), 14-15 (ɪ, confirmed by the markers IDH2 and NP), 16q-19q (part of n, confirmed by the markers GOT2 and PEPD), but not 12-22, which had previously been found in all non-primate species studied (Chowdhary et al. 1998). The fact that chromosomes 12 and 22 are not found in a fused state in either shrews or humans, and that chromosomes 19 and 22 are only present as whole chromosomes in shrews and primates, suggests that primates and shrews may derive from a common lineage within the mammals, as indicated in Fig. 1.

Another gene block common to both humans and shrews is the entire human chromosome 2 (Dixkens et al. 1998). It is suggested that the present human chromosome 2 resulted from the recent fusion of the chimpanzee chromosomes PTR12 and PTR13 at a position corresponding to HSA2q13 (Haaf and Bray-Ward 1993). In most mammalian species studied, human chromosome 2 is found as at least 2 blocks of conserved synteny. They can be identified as 2pter-ql3 and 2q13-qter that are the homologous equivalents of PTR12 and PTR13. Therefore, it can be suggested that this chromosome was found in the ancestral mammalian karyotype in two blocks. While in humans the blocks fused to form chromosome 2, in shrews they fused independently to form the d arm of chromosome de. This is a remarkable case of convergent evolution (Dixkens et al. 1998).

The number of loci mapped in shrews is not very high and mostly restricted to biochemical markers. To change this situation our last efforts have been directed to mapping a set of PCR markers, which are numerous and theoretically only

---

Fig. 2. Shrew consensus genome map. The colour painted chromosomes represent the result of zoo-FISH analysis with human chromosomal probes according to Dixkens et al. (1998) (the extent of homology with segments of specified human chromosomes are indicated by coloured rectangles). Those gene mapping data which agree with the results from chromosome painting are listed to the right of each G-banded shrew chromosome, and the corresponding human localization is indicated. Inconsistent gene assignments are listed in rectangles under the chromosomes. Thus, the relevant genes were assigned to human chromosomes as follows: PEPA: 18; AKI and AK3: 9; MPI: 15; PGM1: 1; GPT and GSR: 8; GUS2: 7; ADA: 20. The PEPI gene was assigned to the m arm of chromosome mp by Matyakhina et al. (1996) but zoo-FISH has revealed homology to human chromosome 12 and an expected localisation on the p arm (Dixkens et al. 1998). The human YAC clones 802D8, 820C10 and 962C4 were assigned by Dixkens et al. (1998) and their positions on shrew chromosome de are indicated by black dots.
Fig. 3. The conservation of the human chromosome 17 (HSA17) gene association in 10 mammalian species. The data are taken from Wakefield and Graves (1996) and the Mouse Genome Database (http://www.informatics.jax.org). The new genes mapped to shrew chromosome hn are marked by asterisks.

restricted by the number of conservative DNA markers sequenced in at least one mammalian species. For this analysis, we have selected one of the most conservative mammalian gene associations, as found in human chromosome 17 (Fig. 3). As was mentioned above, this gene association is almost fully conserved to a single chromosome in most mammalian species studied, except dogs and some primates. In rodents this gene association is presented as a single block of conserved synteny interrupted by several genes from other syntenic groups (Davisson et al. 1998). Previously two loci from this gene group were assigned to shrew chromosome hn (TK1, GH). Zoo-FISH analysis has shown the conservation of the material from
Comparative genome mapping: the shrew map

human chromosome 17 to the distal part of the h arm (Dixkens et al. 1998). We used a set of about 30 heterologous primer pairs for genes located in humans to chromosome 17 for a PCR mapping effort in shrews. As a result we assigned 5 additional genes to shrew chromosome hn (Larkin et al. 2000).

We showed that the use of heterologous primers (primers designed for other species) is suitable for shrew gene mapping. This was confirmed by sequencing data for two out of the 5 genes mapped (Larkin et al. 2000). It was found to have about 90% homology with the sequences of homologous genes in other species. It looks like this approach will be the most appropriate for future mapping of the shrew genome as well as other species which have few or no gene sequences extracted and no inbred lines for linkage analysis.

So, comparative mapping of the mammalian genome shows that chromosome evolution mostly involves rearrangement of large gene blocks. The detailed analysis of zoo-FISH data and genome maps can help us to reconstruct the ancestral karyotype and genome of at least eutherian mammals. This reconstruction benefits from data from as many orders of eutherian mammals as possible. Clearly, further efforts in the mapping of monotremes, marsupials and more distantly related vertebrates would also help in establishing not only the ancestral eutherian genome, but the ancestral vertebrate genome.

Acknowledgements: We appreciate the useful comments on the manuscript by J. Hausser. This study was supported by a grant from Russian Foundation for Basic Research (99-04-49979).

References


Comparative genome mapping: the shrew map


Received 19 October 1999, accepted 3 February 2000.