

Analysis of distribution in the human, pig, and rat genomes points toward a general subtelomeric origin of minisatellite structures.

V Amarger, D Gauguier, Martine M. Yerle, F Apiou, P Pinton, F Giraudeau, S Monfouilloux, M Lathrop, B Dutrillaux, J Buard, et al.

▶ **To cite this version:**

V Amarger, D Gauguier, Martine M. Yerle, F Apiou, P Pinton, et al.. Analysis of distribution in the human, pig, and rat genomes points toward a general subtelomeric origin of minisatellite structures.. Genomics, 1998, 52 (1), pp.62-71. 10.1006/geno.1998.5365 . hal-01160637

HAL Id: hal-01160637

https://ensta-paris.hal.science/hal-01160637

Submitted on 1 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GENOMICS **52,** 62–71 (1998)
ARTICLE NO. GE985365

Analysis of Distribution in the Human, Pig, and Rat Genomes Points toward a General Subtelomeric Origin of Minisatellite Structures

Valérie Amarger,*,[1] Dominique Gauguier,†,‡ Martine Yerle,§ Françoise Apiou,¶ Philippe Pinton,§ Fabienne Giraudeau,*,[1] Sylvaine Monfouilloux,* Mark Lathrop,† Bernard Dutrillaux,¶ Jérôme Buard,*,[2] and Gilles Vergnaud*,‡,[1,3]

*Laboratoire de Recherche en Génétique des Espèces, Institut de Biologie des Hôpitaux de Nantes, 9, Quai Moncousu, 44035 Nantes Cedex, France; †The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom; ‡Laboratoire de Génétique Moléculaire, Centre d'Etudes du Bouchet, BP 3, 91710 Vert le Petit, France; §Laboratoire de Génétique Cellulaire, Centre de Recherche INRA de Toulouse Auzeville, BP 27, 31326 Castanet Tolosan Cedex, France; and ¶Institut Curie Recherche–CNRS UMR 147, 26 Rue d' Ulm, 75248 Paris Cedex, France

Received February 2, 1998; accepted May 4, 1998

We have developed approaches for the cloning of minisatellites from total genomic libraries and applied these approaches to the human, rat, and pig genomes. The chromosomal distribution of minisatellites in the three genomes is strikingly different, with clustering at chromosome ends in human, a seemingly almost even distribution in rat, and an intermediate situation in pig. A closer analysis, however, reveals that interstitial sites in pig and rat often correspond to terminal cytogenetic bands in human. This observation suggests that minisatellites are created toward chromosome ends and their internalization represents secondary events resulting from rearrangements involving chromosome ends.
© 1998 Academic Press

INTRODUCTION

Tandem repeats are classified, according to the overall size of the repeat sequence, among satellites (largest, megabases long), minisatellites (in the kilobase size range), and microsatellites (10 to 40 nucleotides long). Satellite DNA is preferentially associated with centromeric heterochromatin, and microsatellites are distributed throughout the genome. Similarly, in the mouse (Jeffreys *et al.*, 1987; Julier *et al.*, 1990; Mariat *et al.*, 1993), bovine (Georges *et al.*, 1991), and rat genomes (Pravenec *et al.*, 1996) at least, minisatellites are spread over the genome. In contrast, the minisatellites characterized so far in human show a marked clustering in the terminal cytogenetic bands of each chromosome (Royle *et al.*, 1988; Wells *et al.*, 1989; Vergnaud *et al.*, 1991).

This apparent discrepancy may be an artifact due to the cloning procedures that have been used: clustering in human is suggested by a variety of independent approaches performed in different laboratories, whereas the work in the mouse, rat, and bovine relies in each case essentially on a single approach. Furthermore, a very limited number of the supposed minisatellites detected (Jeffreys *et al.*, 1987; Julier *et al.*, 1990; Mariat *et al.*, 1993) or cloned (Georges *et al.*, 1991) from animal genomes have been characterized at the sequence level to confirm their minisatellite nature. Alternatively, this discrepancy could reflect differences in the evolutionary forces that shape mammalian genomes.

To investigate this issue, we cloned minisatellites in human, pig, and rat using identical techniques and compared their chromosomal localization in the three genomes. Clones containing minisatellites were identified in genomic libraries using both synthetic tandem repeats as screening probes (Vergnaud *et al.*, 1991) and a new procedure that is based only on the structural characteristics of the minisatellites and does not require a hybridization step. The cosmids were assigned to a cytogenetic band by *in situ* hybridization assisted when appropriate by linkage analyses. Striking differences were observed in the distribution patterns of these sequences, with subtelomeric clustering confirmed in human, a majority of internal loci in the rat,

Sequence data for this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AJ001151 to AJ001156.

[1] Present address: IECH, Bat 400, Institut de Génétique et Microbiologie, Université Paris-Sud, 91405 Orsay Cedex, France.

[2] Present address: Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK.

[3] To whom correspondence should be addressed at IECH, Bat 400, Institut de Génétique et Microbiologie, Université Paris-Sud, 91405 Orsay Cedex, France. Telephone: +33 1 69 15 62 08. Fax: +33 1 69 15 66 78. E-mail: gilles@igmors.u-psud.fr.

0888-7543/98 $25.00
Copyright © 1998 by Academic Press
All rights of reproduction in any form reserved.

and an intermediate situation in pig. The comparative mapping of the minisatellite-containing regions in the three genomes is consistent with the hypothesis that minisatellites have a subtelomeric origin. These results underline the interest of minisatellites for comparative genome analysis.

MATERIALS AND METHODS

Minisatellite Cloning

Cosmid clones from commercially available libraries (Clontech Refs. HL 1145y, RL1032m, and PL1008m for human, rat, and pig, respectively, on pWE15 vector) were plated at a low density and picked in 96- or 384-well microtiter plates. Replicas were made on nylon filters with a comb.

Procedure without hybridization screening. The clones were grown individually in 2-ml LB medium cultures with shaking at 37°C. After 20 h growth, 24 cultures were pooled, and the cosmid DNA was extracted by the alkaline SDS lysis method as described in Sambrook *et al.* (1989). The cosmid DNA pools were suspended in 300 μl of 10 mM Tris, pH 8, 1 mM EDTA (TE).

Procedure with hybridization screening using synthetic tandem repeat probes. This procedure was described in detail in Vergnaud *et al.* (1991). The synthetic probes used were 13C1, 14C3, 14C5, 14C21, 14C24, 14C31, 15C3, 16C2, 16C4, 16C7, 16C17, 16C20, 16C22, 16C23, 16C24, 16C26, and 16C27. DNA from positive clones was extracted by the alkaline SDS lysis method from 50 ml MLS (Sambrook *et al.,* 1989) medium culture and suspended in 300 μl of TE. In both procedures, three aliquots of 20 (for pools of clones) or 1 μg (for individual clones) of cosmid DNA were digested respectively with the restriction endonucleases *Alu*I and *Hae*III, *Alu*I and *Hin*fI, and *Hin*fI and *Hae*III in a total volume of 100 μl following the supplier's instructions (New England Biolabs or Appligene). After an overnight digestion, the samples were ethanol precipitated, suspended in 5 μl of TE, and loaded for electrophoresis on a 1% horizontal agarose gel. The gel was then stained with ethidium bromide. Fragments above 1.5 kb (from pools) or above 1kb (from individual clones) were sliced out of the gel and recovered by centrifugation as described by Heery *et al.* (1990). One to five nanograms of insert DNA was then directly [^{32}P] labeled by random priming (Feinberg and Vogelstein, 1983) ([^{32}P]dCTP from New England Nuclear; random priming kit from Boehringer) and hybridized at 65°C as described in Vergnaud *et al.* (1991) to a Southern blot carrying the DNA from two unrelated reference individuals digested by *Hae*III, *Hin*fI, and *Pvu*II. This provided an identification pattern for each minisatellite probe, and the profile was encoded by sizing each allele in the three digests. In the case of a pool of clones, the cosmid from which a fragment originates was identified by hybridizing back to the replica of the 384-well plate.

Chromosomal Mapping by Linkage Analysis

Polymorph loci were mapped by linkage analysis using the CEPH panel of human families, the PiGMaP panel of pig reference families, and a F2 (BN × GK) rat population (Gauguier *et al.,* 1996). Genotypes were managed using GENBASE, developed by Jean-Marc Sebaoun. Linkage file output from the human or pig data was converted to CRIMAP file format using the LINK2CRI utility software written by John Attwood. CRIMAP version 2.4 was used for the analyses.

Fluorescence in Situ Hybridization

Rat chromosomes were prepared from normal rat fibroblast cultures. The probes were labeled by nick translation with digoxigenin-11–dUTP (Boehringer Mannheim) and hybridized at a final concentration varying from 2 to 10 ng/μl in presence of an excess of sonicated rat DNA (between 100 and 300 ng/μl) used as competitor. A standard hybridization was performed, as described in Apiou *et al.* (1996). Immunochemical detection of hybridization was performed using mouse anti-digoxygenin antibodies and sheep FITC-conjugated anti-mouse antibodies (Boehringer Mannheim). Rat metaphases were stained with DAPI, and chromosomes were identified with computer-generated reverse DAPI banding. Metaphases were observed under a fluorescence microscope (DMRB, Leica, Germany). Images were captured using a cooled Sensys CCD camera and Quips-Smat capture software (Vysis).

Pig metaphase chromosome spreads were obtained from peripheral blood lymphocyte cultures by standard procedures. Chromosomes were G-banded before hybridization with the G-T-G technique, and pictures were taken using a video printer (Mitsubishi). They were classified according to their G-band pattern as defined by the Committee for the Standardized Karyotype of the Domestic Pig (Gustavsson, 1988). Whole cosmids containing minisatellite sequences were used as probes and were labeled by incorporation of biotinylated 16–dUTP (Boehringer Mannheim) by the random priming method (Feinberg and Vogelstein, 1983). For suppression of background caused by repetitive sequences, the cosmid probe (final concentration 2 μg/ml), sonicated pig genomic competitor DNA (200 μg/ml), and salmon sperm DNA were ethanol precipitated together, dissolved in 25 ml of hybridization mixture, and prehybridized for 3 h at 37°C. *In situ* hybridization, detection of the signals, and interpretation of the results were performed as described previously (Yerle *et al.,* 1992).

Minisatellite Subcloning and Sequencing

Minisatellite inserts were subcloned into PUC 18 or 19 and partially sequenced from both ends using [^{33}P]-labeled universal and reverse M13 primers and the cycle sequencing kit deltaTaq from USB/Amersham.

Identification of Conserved Regions

Restriction fragments from pig cosmids were purified from agarose gel as described before and used as probes to screen a human chromosome 1 library obtained from the Max-Planck Institute for Molecular Genetics: c112 Reference Library, ICRF (Lehrach, 1990). This library consists of two high-density gridded filters of 20,000 clones on each membrane. Each clone is named by the number of the library, c112, and a specific number (for example, c112 F0525). Hybridization on the library filters was at 60°C as described by Vergnaud (1989). Cross-hybridizing fragments from the human cosmids were identified by hybridization with the pig fragments on restriction digests of human cosmids.

Conserved fragments from both species were subcloned in PUC 18 or 19 and sequenced as described before. Direct (universal) and reverse M13 primers were used for sequencing of both ends, and specific primers were ordered to continue the sequence reaction. The accession numbers of the sequences generated in the course of this project are AJ001151 to AJ001156.

RESULTS

Cloning of Minisatellites

Minisatellites were isolated from human, pig, and rat cosmid libraries by two procedures. The first procedure screens cosmid libraries with synthetic tandem repeat probes (Vergnaud, 1989) and the second ("large fragment procedure") does not need a hybridization step (see Materials and Methods). Briefly, hybridization-selected individual cosmids or pools of random cosmids are digested with two frequently cutting restriction enzymes, and fragments above 1 (from hybridization selected cosmids) or 1.5 kb (from pools of random cosmids) are tested by Southern blot hybridization for the presence of a minisatellite.

The number of minisatellites isolated with each method is given in Table 1. Sixty-six, 34, and 43 mini-

TABLE 1

Number of Minisatellite-Containing Cosmids Isolated with Each of the Two Methods

Species	Number of minisatellite-containing cosmids isolated: Method 1[a]	Method 2[b] (clones analyzed)	Number of cosmids localized by FISH[c]
Human	30	36 (5000)	39
Pig	27	7 (3500)	33
Rat	15	28 (7400)	34

[a] Method 1: screening with synthetic tandem repeat probes.
[b] Method 2: the large fragment procedure with no hybridization screening of clones (indicating total number of cosmid clones analyzed in pools).
[c] Number of clones localized by FISH.

satellites were isolated from human, pig, and rat libraries, respectively. The efficiency of the large fragment procedure varies from one species to another, with yields of 0.72% in human, 0.41% in rat, and only 0.2% in pig.

The efficiency of synthetic probes is different in the three species. Probes allowing the detection of minisatellites were 13C1, 14C21, 16C2, 16C4, 16C17, and 16C27 in human; 14C24, 15C3, 16C2, 16C7, 16C22, 16C23, 16C24, and 16C26 in pig; and 13C1, 14C3, 14C5, 14C31, 16C2, 16C4, 16C17, 16C20, 16C24, and 16C27 in rat. The most successful STR probes in the detection of minisatellites were 16C27 in human and 16C23 and 16C26 in pig, allowing the detection of 14, 6, and 6 minisatellites, respectively.

Identification of Minisatellite Structures and Sequences

The fragments detected were identified as minisatellites according to their hybridization pattern on Southern blots, i.e., a strong hybridization as expected for a tandem repeat probe. To confirm minisatellite structure, we subcloned and sequenced 20, 15, and 15 fragments from human, rat, and pig, respectively. A tandem repeat structure is observed in the three species for each fragment sequenced. In all but four cases, a repeat unit, with length ranging from 9 to 128 nucleotides, can be defined. Four fragments detecting multilocus hybridization patterns on Southern blots even under high stringency conditions are composed of successive di-, tri-, or tetranucleotide repeats interspersed with a large number of variant repeats. Consensus sequences of four minisatellites from each of the three species are presented in Table 2, together with the length of the repeat unit and internal variations among units. The majority of variants is due to transition events (represented by letters Y and R). Transversions (S, K, or W) and sometimes three possible variants at the same position (represented by letters D, H, V, and B) are also observed. The number of variant sites in repeat units differs between the minisatellites from 2 variants/51 nucleotides in human minisatellite CEB84 and 30 variants/50 nucleotides in pig minisatellite ADGE S1. A length heterogeneity of the units resulting from insertion/deletion is sometimes observed [repeat $(AGGG)_n$ in CEB1].

Chromosomal Assignment of the Minisatellite Loci

The cosmids were assigned to a cytogenetic band by a combination of fluorescence *in situ* hybridization (FISH) and segregation analyses. This second approach was most appropriate for human loci because a physical assignment of highly polymorphic loci can usually be deduced from segregation data of CEPH families (Spurr *et al.*, 1994) and the knowledge of the physical assignment of close-by markers (Bray-Ward *et al.*, 1996) from the CEPH database. Thirty-nine, 33, and 34 cosmids were mapped by FISH in human, pig, and rat, respectively. The minisatellite localization is shown on Figs. 1A (hu-

TABLE 2

Nucleotide Sequence of a Set of Minisatellites from Human, Pig, and Rat Genomes

Name[a]	Unit sequence[b]	Unit length[c]
CEB 1	GRGGGAGGGARGG(AGGG)$_n$TGG(C)CTGCRGAGGTCCCYGGGYTGA	37–43 bp
CEB 8	GGGGRAGBDT	10 bp
CEB 62	AYCTACCCCRGGS(C)AGCAGGSAGGAGGCWGC	30–31 bp
CEB 84	GGTCACATGTGGGGGGGCTGGRGTCAGASAGACCAAGGTTGGAACCCGGCA	51 bp
ADGE S1	AYGWGDKYTYRARYCHCARTMTRKHTTYWKRYTGKGWTYRAGTYCYRRS(Y)	49–50 bp
ADGE S2	GGATGRGGWCRGGGACR(GGG(R)CA)	17–23 bp
ADGE S12	HAG(G)TTCAGGTGATGGRVG	19–20 bp
ADGE S40	Y(A)GARR(R)CAGGVCWGTGTYTCYCYTCAGCCYGGGG(Y)TC(YC)	35–40 bp
CEB R10	CCC(AAT)AT	5–8 bp
CEB R15	AYBRKGBH(C)T(T)CC	11–13 bp
CEB R21	AGRYRCCASRGYYRRR	16 bp
CEB R31	DGCYCATCAGTGGGAGYGTCC(G)AKTGAGA	28–29 bp

[a] Minisatellite name (CEB, human minisatellites; CEB R, rat minisatellites; ADGE S, pig minisatellites).
[b] Sequence of the elementary unit of the minisatellite loci. Variant positions are underlined. (R = A or G; Y = C or T; M = A or C; K = G or T; S = G or C; W = A or T; H = A, T, or C; B = G, T, or C; D = G, A, or T; N = A, G, C, or T; V = G, A, or C; IUB coding rules). The sequence presented is a consensus sequence deduced from the sequence of different motifs.
[c] Observed length of the different units.

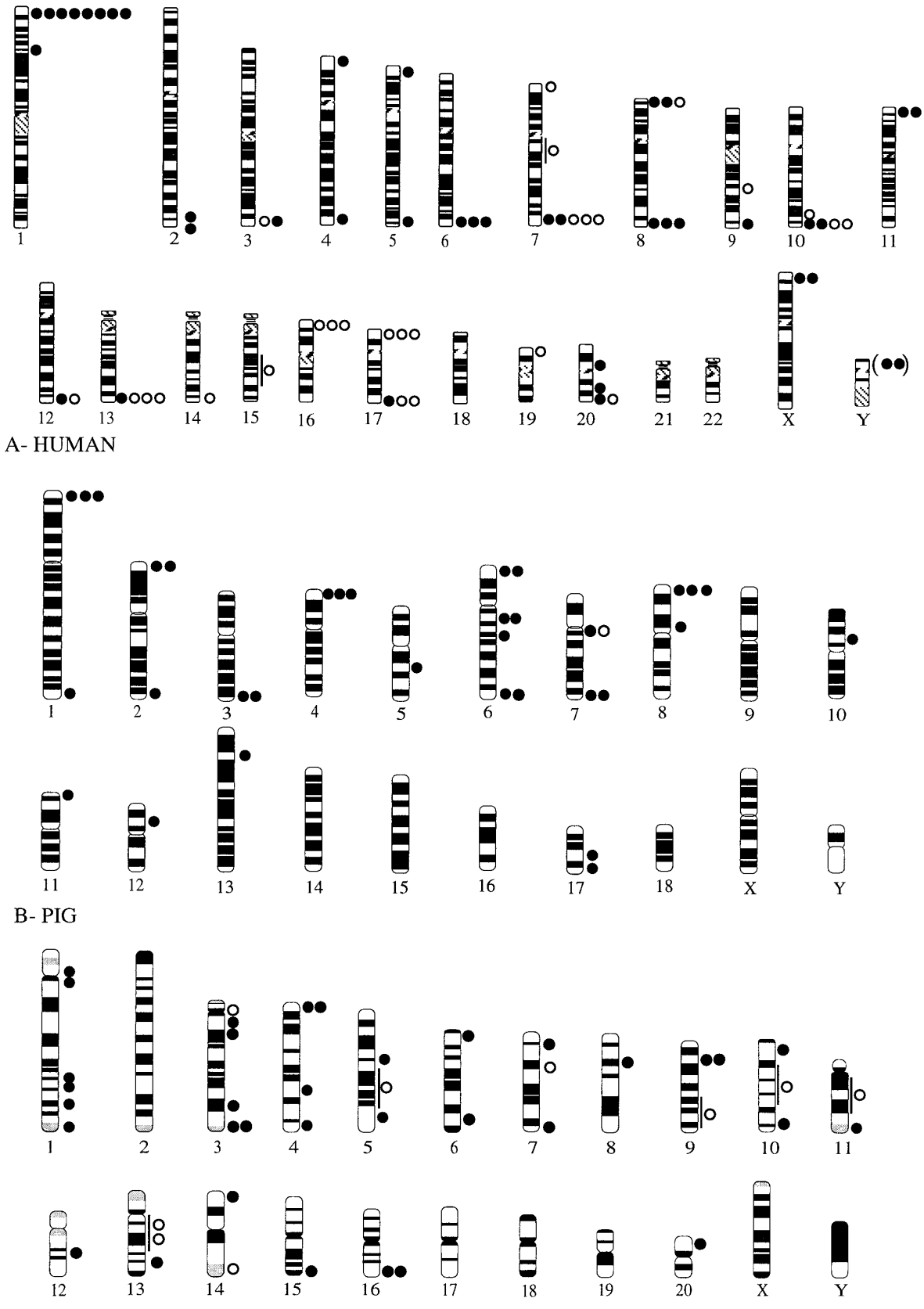


FIG. 1. Chromosomal localization of minisatellites loci on human (**A**), pig (**B**), and rat (**C**) chromosomes (Giemsa banding). Minisatellites localized by both FISH and linkage analyses are represented by black circles. Minisatellites localized only by linkage analysis to physically assigned polymorphic markers are represented by open circles. The two loci represented at the tip of the short arm of the X chromosome are within the X–Y pseudoautosomal region and are also represented at the tip of Yp.

man loci), 1B (pig loci), and 1C (rat loci). A prominent subtelomeric clustering of the minisatellites is a common feature of human and pig genomes. In human, 59 among 66 loci isolated (90%) are in the terminal cytogenetic band of a chromosome. Of 23 terminal human bands with minisatellites, 10 are "major clusters," which we define as the presence of three or more minisatellites. These 10 terminal bands alone contain 60% of minisatellites reported here. The most striking cluster is found on chromosome 1p36.3. The lower proportion of subtelomeric minisatellites in pig (23 of 34; 66%) results from a few interstitial clustered or isolated loci (on chromosomes 5, 6, 7, 8, 10, 12, 13, and 17). In the rat genome, the situation is different, with a much lower rate of terminal band clustering (30%).

Comparative Localization of the Minisatellites in the Three Genomes

Human–Pig Comparison

Chromosome 6 alone accounts for 3 of the 11 interstitial sites detected in pig. Interestingly, the physical assignment of the minisatellite cluster at 6q2.1–q2.2 corresponds to the position of an ancestral chromosome fusion event with remnants of telomere arrays (Gu *et al.*, 1996). The proximal side of the fusion event is homologous to part of human chromosome 19, and the distal side is homologous to human chromosome 1p36–p31 as deduced from heterologous chromosome painting experiments (Goureau *et al.*, 1996) and a number of gene assignments (Yerle *et al.*, 1997). We have identified conserved regions within cosmids containing the pig chr 6q minisatellites to characterize more precisely the homology between pig chr 6q minisatellite loci and human (Fig. 2).

The different fragments obtained after an *Eco*RI digestion of the four pig cosmids containing minisatellites ADGE S26, S37, S16, and S3 were used to screen a human chromosome 1 cosmid library at moderate stringency. Clones giving a hybridization signal with at least two fragments from the same pig cosmid and no signal with fragments from another cosmid were selected. No clone satisfying this criterion was obtained when fragments from the ADGE S26 cosmid were used. In the three other cases, cross-hybridizing fragments were identified and sequenced, and the human homologous cosmids were assigned to a chromosome 1 cytogenetic band by FISH.

Following is a characterization of the conserved sequences.

Pig ADGE S37-containing cosmid is homologous to human cosmid icrf-c112 F0525 located at chromosome 1p36. The conserved region was included in a 0.9-kb *Sma*I fragment in the pig cosmid (Accession No. AJ001154) and in a 1.6-kb *Pvu*II fragment in the human cosmid c112 F0525 (Accession No. AJ001151). The fragments show two regions of high homology (88–90%) separated by a short fragment of low homology (54%). No homology was found within the databases, and no splicing sites were observed. No open reading frame common to the two fragments was found.

Pig ADGE-S16 minisatellite-containing cosmid is homologous to human cosmid icrf-c112 B1018 and corresponds to the P5CDH locus. A part of the pig restriction fragment (*Pvu*II—1.8 kb; Accession No. AJ001155) and a part of the human restriction fragment (*Bam*HI—0.5 kb; Accession No. AJ001152) showed 88–90 and 100% homology, respectively, each with a different contiguous segment of the human expressed sequence tag HSU24267, located on chromosome 1p36 and encoding the human δ 1-pyrroline-5-carboxylate dehydrogenase (Hu *et al.*, 1996). Splice acceptor and donor site consensus sequences were found at the extremities of each exon.

Pig ADGE-S3 minisatellite-containing cosmid is homologous to human cosmid icrf-c112 D2131 located at chromosome 1p31–p32. The cross-hybridizing fragments in the pig cosmid (*Bam*HI—1.6 kb; Accession No. AJ001156) and in the human cosmid (*Sma*I—0.7 kb; Accession No. AJ001153) showed two regions each of about 100 bp with 74% homology separated by about 200 bp of unrelated sequence and followed by a region of intermediate homology (66–70%).

Human–Rat Comparison

We focused our analysis on rat chromosome 1 from which six minisatellite-containing cosmids were isolated and which benefits from relatively intensive gene mapping efforts. Minisatellites R12 and R21 were localized by FISH in the centromeric region of chromosome 1 and R19 in the telomeric region, whereas R10, R15, and R30 are interstitial. The minisatellites' positions with respect to genes and other markers (Fig. 3) have been determined by typing a rat F2 cross with microsatellites isolated from the cosmids (F. Giraudeau, unpublished). Minisatellites R10, R12, R15, R19, and R30 are closely linked with genes Drd4, Prkcg, Igf2, Cyp2c12, and Pygm, respectively (Fig. 3).

DISCUSSION

Minisatellite Cloning

We present here a differential organization of minisatellite loci from the human, pig, and rat genomes. The combination of two independent procedures was highly efficient for the cloning of these loci from genomic cosmid libraries. The method involving synthetic tandem-repeat hybridization probes has already proved its efficiency in detecting polymorphic loci in complex genomes (Mariat and Vergnaud, 1992) and was initially tested on a smaller scale for the cloning of such loci in human (Vergnaud *et al.*, 1991). The species-specific efficiency of a given synthetic probe in minisatellite cloning in the human, rat, and pig genomes parallels our earlier observations on Southern

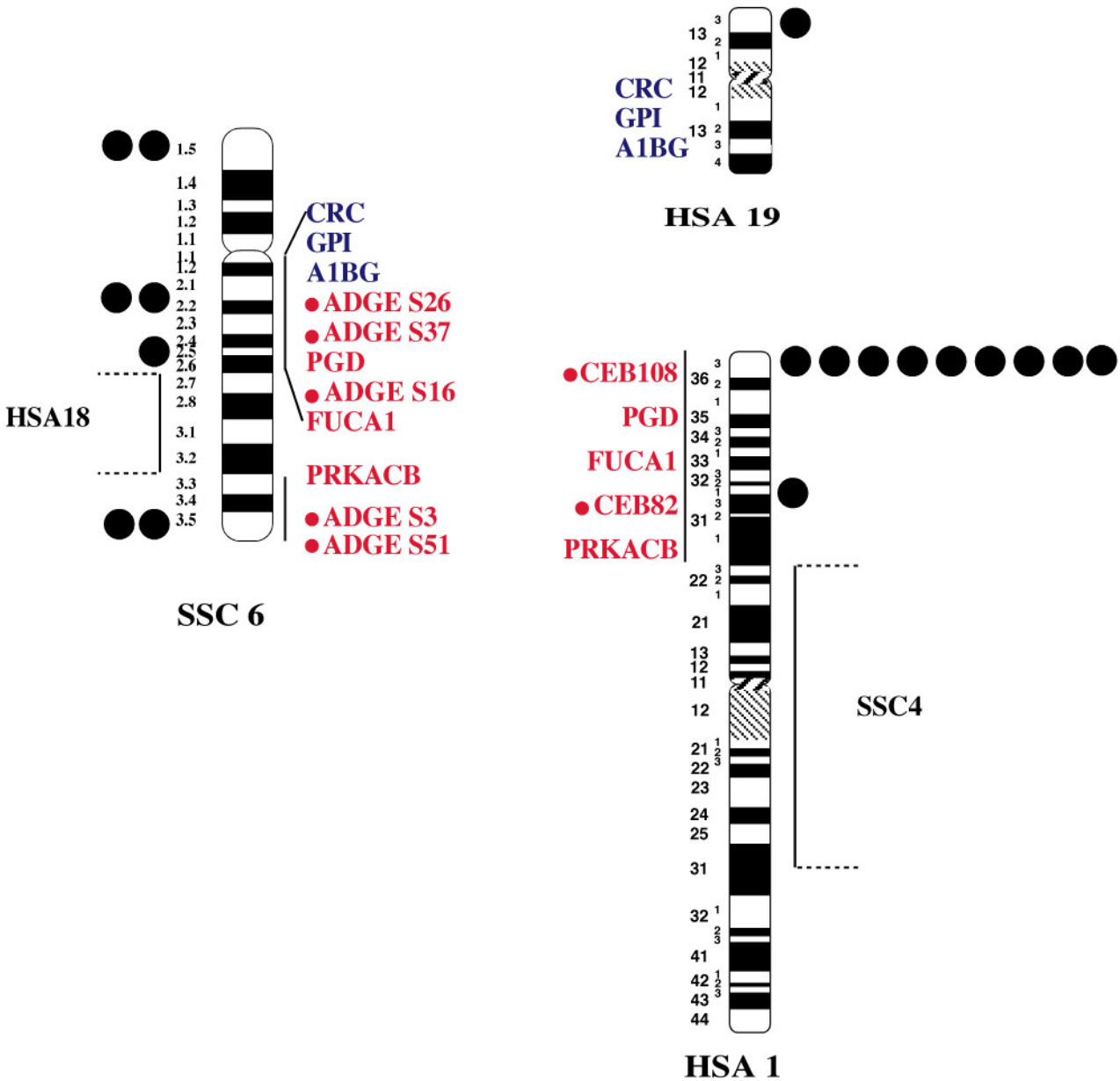


FIG. 2. Synteny relationships between pig chromosome 6 and human chromosomes 1 and 19. Large black circles represent minisatellites as localized by FISH. Small circles represent the same minisatellites (names given) as localized by linkage analysis among several genes (for reasons of convenience, only one (CEB108) representative minisatellite of the human chromosome 1p terminal band cluster is shown). Black lines along the chromosomes represent homologous segments between human and pig chromosomes as suggested by heterologous painting experiments (Rettenberger *et al.*, 1995; Goureau *et al.*, 1996). In addition to homologous segments between HSA 1 and 19 and SSC 6, those from HSA 18 and SSC 4, contributing to SSC 6 and HSA 1, respectively, are shown.

blots of genomic DNA in a larger range of species (Mariat *et al.*, 1992) and might indicate the existence of species-specific minisatellite families. The second method uses a repeat structure criterion and avoids the selection bias introduced by cross-hybridization. This method showed different isolation efficiencies for human, rat, and pig (0.7, 0.4, and 0.2%, respectively, of the cosmid clones analyzed; Table 1), suggesting that the frequency of minisatellites is highest in human or, alternatively, that on average human minisatellite alleles are longer.

Chromosomal Localization

Results reported here are in agreement with earlier ones conducted in human and suggesting minisatellite clustering in subtelomeric regions of human chromosomes (Royle *et al.*, 1988; Vergnaud *et al.*, 1991, 1993). Because the same methodology was here applied in a single laboratory to the three genomes, it is clear that the distribution of minisatellite localizations reported now is not an artifact introduced by cloning bias. Three different situations are observed, with 90% of subtelo-

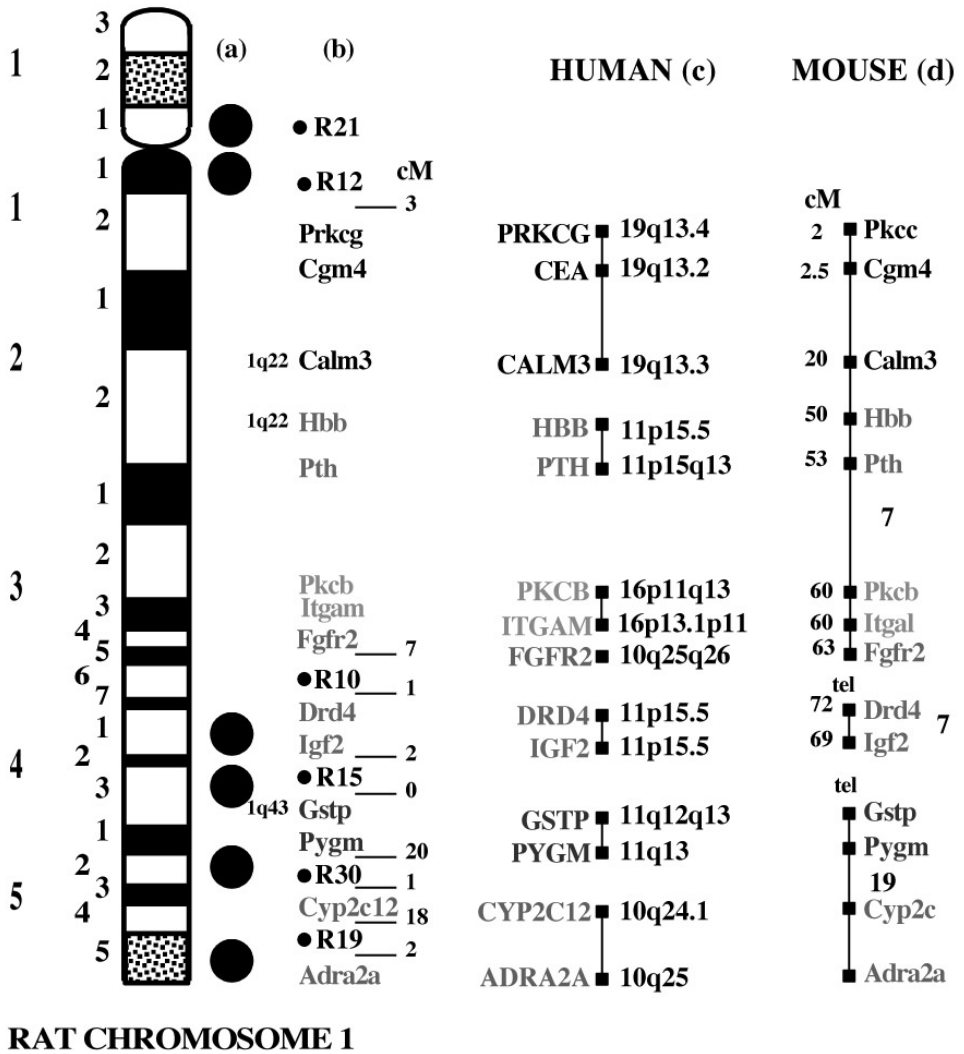


FIG. 3. Comparative mapping between rat chromosome 1 and human and mouse chromosomes. (**a**) Localization of rat minisatellites (represented by black circles) by FISH. (**b**) Localization of rat minisatellites (small black circle and name R-) among genes as determined by linkage analysis. Genetic distances in cM between minisatellites and closest genes are indicated. (**c**) Localization of the homologous genes in human (physical localization). (**d**) Localization of the homologous genes in mouse (genetic localization: on the right of the black line are gene names and chromosome numbers and on the left the position of the gene in centimorgans on the genetic map). The relative positions of genes in human and mouse were obtained from NCBI (http://www.ncbi.nlm.nih.gov/SCIENCE 96/) for human and from The Jackson Laboratory (http://www.jax.org/) for mouse.

meric clustering for human minisatellite loci, 66% for pig loci, and only 30% for rat loci. In addition, we demonstrate that minisatellites are not found equally at all chromosome ends. In this respect, human 1pter is most remarkable compared to 1qter. It may be relevant to note that 1p36 is a very strong T band (Dutrillaux, 1973) characterized by a high concentration of H3 isochores, CpG islands, and genes (Saccone *et al.,* 1996; Zoubak *et al.,* 1996). The much more homogeneous distribution of minisatellites observed in rat would then also be correlated with GC level, which is distributed more homogeneously in rodent chromosomes (Sabeur *et al.,* 1993). However, a number of discrepancies exist between the location of minisatellite clusters (for instance human 6qter is a minisatellite cluster; Fig. 1) and the location of T bands (6qter is not a T band; Dutrillaux, 1973). More generally, what is the reason for the relative paucity of minisatellites at some chromosome ends: does it reflect a more recent telomeric location or rather a specific behavior of some subtelomeric domains promoting the creation of minisatellites (either at a higher density or at a comparable density in a larger domain). Further studies and especially sequencing of a number of subtelomeric regions may help address these issues.

Comparative Analysis of Minisatellite-Containing Regions

Comparative gene mapping data available show a correspondence between pig interstitial minisatellite-containing regions and human chromosome extremi-

ties. The genes RNR2 and VWF (Von Willebrand Factor) localized respectively on pig chr 10p12–p13 and chr 5q21 (Miyake *et al.,* 1988; Sjoberg *et al.,* 1996) in the vicinity of interstitial pig minisatellites (Fig. 1) map on human terminal bands on chromosome 14p12 and chromosome 12p13.3, respectively (the VWF gene is also subtelomeric in cattle (chr 5q35), sheep (chr 3q35), and goat (chr 5q35); Janel *et al.,* 1996).

Pig chromosome 6 results from a telomeric fusion between two chromosomes (Gu *et al.,* 1996). Heterologous painting experiments (Rettenberger *et al.,* 1995; Goureau *et al.,* 1996) and comparative gene mapping reveal that the two extremities involved are homologous with human chromosome 1p and chromosome 19q (Fig. 2). Pig minisatellites ADGE S26 and ADGE S37 are localized in the band containing the telomere fusion between the two synteny groups, and the human homologue of the ADGE S37 locus is indeed located in 1p36, showing that the pig interstitial chromosome 6 minisatellite cluster is conserved syntenically with the human chromosome 1pter minisatellite cluster. Strikingly, the human homologue of the 6qter pig cosmid containing minisatellite ADGE S3 is located at 1p31–p32, from which we independently characterized a minisatellite locus (Fig. 1).

Although the situation in rat is expectedly much more complex, results from the localization of rat minisatellites provide further support for the hypothesis of a subtelomeric origin of minisatellite loci. As a general rule, comparative analysis between human and rat genomes is hampered by the paucity of mapped genes in rat (Bihoreau *et al.,* 1997) and by the large number of chromosomal rearrangements between rodent and human genomes. The number of conserved segments between human and mouse is estimated to be at least 180 (DeBry and Seldin, 1996), whereas it is estimated at less than 50 between human and pig (Rettenberger *et al.,* 1995; Goureau *et al.,* 1996). Comparative mapping of genes in human, mouse, and rat chromosomes X showed the extent of complex rearrangements within a single chromosome (Millwood *et al.,* 1997). The rat chromosome 1 provides a good model for comparative mapping and chromosomal evolution because it is the longest rat chromosome (about 10% of the genome), and the presence of susceptibility loci for multifactorial diseases (Gauguier *et al.,* 1996) has led to a dramatic improvement in the density of the linkage map (Bihoreau *et al.,* 1997).

There is evidence of conservation between rat chromosome 1 and mouse chromosome 7 with an inversion of a segment at the extremity of mouse chromosome 7 (genes Drd4 and Igf2). Starting from the extremity of the long arm, rat chr 1 is homologous to segments from human chromosomes 10q24–q25, 11q12–q13, 11p15.5, 10q25q26, 16p13.1–p11, 11p15, and 19q13. R30 is located between PygM and Cyp2c12, which are on two different chromosomes in human. Moreover, minisatellite loci R10 and R15 are located immediately proximal and immediately distal to the two genes Drd4 and Igf2, which are in the terminal 11p15.5 human band.

This collection of observations, together with the striking distribution of minisatellites in human, supports the hypothesis of a subtelomeric formation of minisatellites followed by internalization through chromosomal rearrangements. This hypothesis predicts that interstitial minisatellites are markers of ancestrally telomeric regions of the genome. Not exclusively, one could postulate movements of minisatellites within the genome by a number of mechanisms, as reported, for instance, in Wong *et al.* (1990), Eichler *et al.* (1996), and Monfouilloux *et al.* (1998). However, the events described in these reports produce duplications, triplications, etc., of relatively small domains (less than a few hundred kilobases in Monfouilloux *et al.* (1998) and 26 kb in Eichler *et al.* (1996)), whereas the interstitial pig minisatellite cluster on chromosome 6 described here is part of a very large conserved segment between human and pig and has not been duplicated. Interestingly, in the two independent examples reported in Wong *et al.* (1990) and Eichler *et al.* (1996), one of which contains a minisatellite, the ancestral site of the transposition event is telomeric (at respectively 6pter and Xq28) and both target sites are within band 16p11.

No general function has been attributed to minisatellites, but some of them do have a biological role. They can be coding sequences (Boguski *et al.,* 1986; Paulsson *et al.,* 1992) or regulatory elements (Krontiris *et al.,* 1993; Bennett *et al.,* 1995). The abnormal expansion of a minisatellite can create a fragile site (Virtaneva *et al.,* 1997; Yu *et al.,* 1997). The analysis of neomutated alleles at other, nonpathological hypermutable minisatellites has shown that they evolve predominantly by gene-conversion-like mechanisms, presumably initiated by double-strand breaks (Buard and Vergnaud, 1994; Jeffreys *et al.,* 1994), whereas microsatellites most likely mutate by replication slippage mechanisms (Heale and Petes, 1995). Therefore, a good knowledge of domains in which minisatellites are likely to be found within the human genome may accelerate the identification of genetic disorders associated with this type of tandem repeat element.

ACKNOWLEDGMENTS

The authors thank Professor Howard Cann for helpful comments on the manuscript. This work was supported by the PiGMaP project (EC Contract BIO2-CT94-3044), by the EUROGEM project (EC Contract GENE-CT93-0101), and by a rat genome project (EC Contract BI04-CT96-0372). Research at Oxford was supported by the Wellcome Trust. M.L. holds a Wellcome Trust Principal Fellowship. More detailed information on individual markers can be accessed at our Web site (http://www.igmors.u-psud.fr/iech).

REFERENCES

Apiou, F., Flagiello, D., Cillo, C., Malfoy, B., Poupon, M. F., and Dutrillaux, B. (1996). Fine mapping of human HOX gene clusters. *Cytogenet. Cell Genet.* **73:** 114–115.

Bennett, S. T., Lucassen, A. M., Gough, S. C. L., Powell, E. E., Undlien, D. E., Pritchard, L. E., Merriman, M. E., Kawaguchi, Y., Dronsfield, M. J., Pociot, F., Nerup, J,. Bouzekri, N., Cambon-Thomsen, A., Ronningen, K. S., Barnett, A. H., Bain, S. C., and Todd, J. A. (1995). Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat. Genet.* **9:** 284–292.

Bihoreau, M. T., Gauguier, D., Kato, N., Hyne, G., Lindpaintner, K., Rapp, J. P., James, M. R., and Lathrop, G. M. (1997). A linkage map of the rat genome derived from three F2 crosses. *Genome Res.* **7:** 434–440.

Boguski, M. S., Birkenmeier, E. H., Elshourbagy, N. A., Taylor, M. T., and Gordon, J. I. (1986). Evolution of the apolipoproteins. *J. Biol. Chem.* **261:** 6398–6407.

Bray-Ward, P., Menninger, J., Lieman, J., Desai, T., Mokady, N., Banks, A., and Ward, D. C. (1996). Integration of the cytogenetic, genetic, and physical maps of the human genome by FISH mapping of CEPH YAC clones. *Genomics* **32:** 1–14.

Buard, J., and Vergnaud, G. (1994). Complex recombination events at the hypermutable minisatellite CEB1 (D2S90). *EMBO J.* **13:** 3203–3210.

DeBry, R. W., and Seldin, M. F. (1996). Human/mouse homology relationships. *Genomics* **33:** 337–351.

Dutrillaux, B. (1973). Nouveau système de marquage chromosomique: Les bandes T. *Chromosoma* **41:** 395–402.

Eichler, E. E., Lu, F., Shen, Y., Antonacci, R., Jurecic, V., Doggett, N. A., Moyzis, R. K., Baldini, A., Gibbs, R. A., and Nelson, D. L. (1996). Duplication of a gene-rich cluster between 16p11.1 and Xq28: A novel pericentromeric-directed mechanism for paralogous genome evolution. *Hum. Mol. Genet.* **5:** 899–912.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132:** 6–13.

Gauguier, D., Froguel, P., Parent, V., Bernard, C., Bihoreau, M. T., Portha, B., James, M. R., Penicaud, L., Lathrop, M., and Ktorza, A. (1996). Chromosomal mapping of genetic loci associated with the GK rat. *Nat. Genet.* **12:** 38–43.

Georges, M., Gunawardana, A., Threadgill, D. W., Lathrop, M., Olsaker, I., Mishra, A., , Sargeant, L. L. Schoeberlein, A., Steele, M. R., Terry, C., Threadgill, D. S., Zhao, X., Holm, T., Fries, R., and Womack, J. E. (1991). Characterization of a set of variable number of tandem repeat markers conserved in Bovidae. *Genomics* **11:** 24–32.

Goureau, A., Yerle, M., Schmitz, A., Riquet, J., Milan, D., Pinton, P., Frelat, G., and Gellin, J. (1996). Human and porcine correspondence of chromosome segments using bidirectional chromosome painting. *Genomics* **36:** 252–262.

Gu, F., Hindkjaer, J., Gustavsson, I., and Bolund, L. (1996). A signal of telomeric sequences on porcine chromosome 6q21–q22 detected by primed in situ labelling. *Chromosome Res.* **4:** 251–252.

Gustavsson, I. (1988). Standard karyotype of the domestic pig. Committee for the Standardized Karyotype of the Domestic Pig. *Hereditas* **109:** 151–157.

Heale, S. M., and Petes, T. D. (1995). The stabilization of repetitive tracts of DNA by variant repeats requires a functional mismatch repair system. *Cell* **83:** 539–545.

Heery, D. M., Gannon, F., and Powell, R. (1990). A simple method for subcloning DNA fragments from gel slices. *Trends Genet.* **6:** 173.

Hu, C. A., Lin, W. W., and Valle, D. (1996). Cloning, characterization, and expression of cDNAs encoding human delta 1-pyrroline-5-carboxylate dehydrogenase. *J. Biol. Chem.* **271:** 9795–9800.

Janel, N., Schibler, L., Oustry, A., Kerbiriou-Nabias, D., Cribiu, E. P., and Vaiman, D. (1996). The localization of the von Willebrand factor gene on cattle, sheep and goat chromosomes illustrates karyotype evolution in mammals. *Mamm. Genome* **7:** 633–634.

Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L., and Armour, J. A. L. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nat. Genet.* **6:** 136–145.

Jeffreys, A. J., Wilson, V., Kelly, R., Taylor, B. A., and Bulfield, G. (1987). Mouse DNA 'fingerprints': Analysis of chromosome localization and germ-line stability of hypervariable loci in recombinant inbred strains. *Nucleic Acids Res.* **15:** 2823–2836.

Julier, C., de Gouyon, B., Georges, M., Guénet, J. L., Nakamura, Y., Avner, P., and Lathrop, G. M. (1990). Minisatellite linkage maps in the mouse by cross-hybridization with human probes containing tandem repeats. *Proc. Natl. Acad. Sci. USA* **87:** 4585–4589.

Krontiris, T. G., Devlin, B., Karp, D. D., Robert, N. J., and Risch, N. (1993). An association between the risk of cancer and mutations in the HRAS1 minisatellite locus. *N. Engl. J. Med.* **329:** 517–523.

Lehrach, H. (1990). "Genome Analysis: Genetic and Physical Mapping" (K. E. Davies and S. M. Tilghman, Eds.), Vol. 1, pp. 39–81, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mariat, D., De Gouyon, B., Julier, C., Lathrop, M., and Vergnaud, G. (1993). Genetic mapping through the use of synthetic tandem repeats in the mouse genome. *Mamm. Genome* **4:** 135–140.

Mariat, D., Guérin, G., Bertaud, M., and Vergnaud, G. (1992). Modulation of polymorphic loci detection with synthetic tandem repeat variants. *Mamm. Genome* **3:** 546–549.

Mariat, D., and Vergnaud, G. (1992). Detection of polymorphic loci in various genomes with synthetic tandem repeats. *Genomics* **12:** 454–458.

Millwood, I. Y., Bihoreau, M. T., Gauguier, D., Hyne, G., Levy, E. R., Kreutz, R., Lathrop, G. M., and Monaco, A. P. (1997). A gene-based genetic linkage and comparative map of the rat X chromosome. *Genomics* **40:** 253–261.

Miyake, Y. I., O'Brien, S., and Kaneda, Y. (1988). Regional localization of rDNA gene on pig chromosome 10 by in situ hybridization. *Jpn. J. Vet. Sci.* **50:** 341–345.

Monfouilloux, S., Avet-Loiseau, H., Amarger, V., Balazs, I., Pourcel, C., and Vergnaud, G. (1998). Recent human-specific spreading of a subtelomeric domain. *Genomics* **51,** in press.

Paulsson, G., Höög, C., Bernholm, K., and Wieslander, L. (1992). Balbiani ring 1 gene *in Chironomus tentans.* Sequence organisation and dynamics of a coding minisatellite. *J. Mol. Biol.* **225:** 349–361.

Pravenec, M., Gauguier, D., Schott, J.-J., Buard, J., Kren, V., Bila, V., Szpirer, C., Szpirer, J., Wang, J.-M., Huang, H., St. Lezin, E., Spence, M. A., Flodman, P., Printz, M., Lathrop, G. M., Vergnaud, G., and Kurtz, T. W. (1996). A genetic linkage map of the rat derived from recombinant inbred strains. *Mamm. Genome* **7:** 117–127.

Rettenberger, G., Klett, C., Zechner, U., Kunz, J., Vogel, W., and Hameister, H. (1995). Visualization of the conservation of synteny between humans and pigs by heterologous chromosomal painting. *Genomics* **26:** 372–378.

Royle, N. J., Clarkson, R. E., Wong, Z., and Jeffreys, A. J. (1988). Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. *Genomics* **3:** 352–360.

Sabeur, G., Macaya, G., Kadi, F., and Bernardi, G. (1993). The isochore patterns of mammalian genomes and their phylogenetic implications. *J. Mol. Evol.* **37:** 93–108.

Saccone, S., Caccio, S., Kusuda, J., Andreozzi, L., and Bernardi, G. (1996). Identification of the gene-richest bands in human chromosomes. *Gene* **174:** 85–94.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sjoberg, A., Seaman, W. T., Bellinger, D. A., Griggs, T. R., Nichols, T. C., and Chowdhary, B. P. (1996). FISH mapping of the porcine vWF gene to chromosome 5q21 extends synteny homology with human chromosome 12. *Hereditas* **124:** 199–202.

Spurr, N. K., Bryant, S. P., Attwood, J., Nyberg, K., Cox, S. A., Mills, A., Bains, R., Warne, D., Cullin, L., Povey, S., Sebaoun, J.-M., Weissenbach, J., Cann, H. M., Lathrop, M., Dausset, J., Marcadet-

Troton, A., and Cohen, D. (1994). European Gene Mapping Project (EUROGEM): Genetic maps based on the CEPH reference families. *Eur. J. Hum. Genet.* **2:** 193–203.

Vergnaud, G. (1989). Polymers of random short oligonucleotides detect polymorphic loci in the human genome. *Nucleic Acids Res.* **17:** 7623–7630.

Vergnaud, G., Gauguier, D., Schott, J.-J., Lepetit, D., Lauthier, V., Mariat, D., and Buard, J. (1993). Detection, cloning, and distribution of minisatellites in some mammalian genomes. *In* "DNA Fingerprinting: State of the Science" (S. D. J. Pena, R. Chakraborty, J. T. Epplen, and A. J. Jeffreys, Eds.), Vol. 67, pp. 47–57,. Birkhäuser, Basel/Switzerland.

Vergnaud, G., Mariat, D., Apiou, F., Aurias, A., Lathrop, M., and Lauthier, V. (1991). The use of synthetic tandem repeats to isolate new VNTR loci: Cloning of a human hypermutable sequence. *Genomics* **11:** 135–144.

Virtaneva, K., D'Amato, E., Miao, J., Koskiniemi, M., Norio, R., Avanzini, G., Franceschetti, S., Michelucci, R., Tassinari, C. A., Omer, S., Pennachio, L. A., Myers, R. M., , Dieguez-Lucena, J. L. Krahe, R., de la Chapelle, A., and Lehesjoki, A. E. (1997). Unstable minisatellite expansion causing recessively inherited myoclonus epilepsy, EPM1. *Nat. Genet.* **15:** 393–396.

Wells, R. A., Green, P., and Reeders, S. T. (1989). Simultaneous genetic mapping of multiple human minisatellite sequences using DNA fingerprinting. *Genomics* **5:** 761–772.

Wong, Z., Royle, N. J., and Jeffreys, A. J. (1990). A novel human DNA polymorphism resulting from transfer of DNA from chromosome 6 to chromosome 16. *Genomics* **7:** 222–234.

Yerle, M., Galman, O., Lahbib-Mansais, Y., and Gellin, J. (1992). Localization of the pig luteinizing hormone/choriogonadotropin receptor gene (LHCGR) by nonradioactive in situ hybridization. *Cytogenet. Cell Genet.* **59:** 48–51.

Yerle, M., Lahbib-Mansais, Y., Pinton, P., Robic, A., Goureau, A., Milan, D., and Gellin, J. (1997). The cytogenetic map of the domestic pig (*Sus scrofa domestica*). *Mamm. Genome* **8:** 592–607.

Yu, S., Mangelsdorf, M., Hewett, D., Hobson, L., Baker, E., Eyre, H. J., Lapsys, N., Le Paslier, D., Doggett, N. A., Sutherland, G. R., and Richards, R. I. (1997). Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. *Cell* **88:** 367–374.

Zoubak, S., Clay, O., and Bernardi, G. (1996). The gene distribution of the human genome. *Gene* **174:** 95–102.