Chromosome Research 7: 379–386, 1999.

Identification of the gene-richest bands in human prometaphase chromosomes

Salvatore Saccone[1], Concetta Federico[1], Irina Solovei[2], Marie-Françoise Croquette[3], Giuliano Della Valle[4] & Giorgio Bernardi[5]*

[1]*Dipartimento Biologia Animale, University of Catania, via Androne 81, 95124 Catania, Italy;* [2]*Institut fuer Anthropologie und Humangenetik, Ludwig Maximilian Universitaet, Muenchen, Germany;* [3]*Hôpital Saint-Antoine, 329 Bld. Victor-Hugo B.P. 255, 59019 Lille, Cedex;* [4]*Dipartimento Biologia Evoluzionistica Sperimentale, University of Bologna, via Selmi 4, Bologna, Italy;* [5]*Laboratorio di Evoluzione Molecolare, Stazione Zoologica, Villa Comunale, 80121 Napoli, Italy; Tel: (+39) 081 5833300; Fax: (+39) 081 2457294; E-mail: Bernardi@alpha.szn.it*

* *Correspondence*

Received 10 March 1999; received in revised form and accepted for publication by B. Dutrillaux 4 May 1999

Key words: chromosomal bands, genome, *in-situ* hybridization, isochores

Abstract

The human genome is a mosaic of long, compositionally homogeneous DNA segments, the isochores, that can be partitioned into five families, two GC-poor families (L1 and L2), representing 63% of the genome, and three GC-rich families (H1, H2 and H3), representing 24%, 7.5% and 4–5% of the genome, respectively. Gene concentration increases with increasing GC levels, reaching a level 20-fold higher in H3 compared with L isochores. *In-situ* hybridization of DNA from different isochore families provides, therefore, information on the chromosomal distribution of genes. Using this approach, three subsets of reverse or Giemsa-negative bands, $H3^+$, $H3^*$ and $H3^-$, containing large, moderate, and no detectable amounts, respectively, of the gene-richest H3 isochores were identified at a resolution of 400 bands. $H3^+$ bands largely coincide with the most heat-denaturation-resistant bands, the chromomycin-A3-positive, DAPI-negative bands, the bands with the highest CpG island concentrations, and the earliest replicating bands. Here, we have defined the $H3^+$ bands at a 850-band resolution, and have thus identified the human genome regions, having an average size of 4 Mb, that are endowed with the highest gene density.

Introduction

The human genome is a mosaic of isochores, long DNA segments which are compositionally homogeneous and can be partitioned into five families, namely two GC-poor families (L1 and L2, collectively called L; GC is the molar fraction of guanine + cytosine in DNA), representing about 63% of the genome, and three GC-rich families (H1, H2 and H3) representing about 24%, 7.5% and 4–5% of the genome, respectively, the remaining DNA being formed by satellite and ribosomal DNAs (Bernardi *et al.* 1985, Bernardi 1989, 1995). Gene concentration increases with increasing GC levels (Bernardi *et al.* 1985, Mouchiroud *et al.* 1991, Zoubak *et al.* 1996). *In-situ* hybridization of compositional DNA fractions corresponding to different isochore families on metaphase chromosomes (chromosomal compositional mapping) provides, therefore, information not only on the correlation between isochores and chromosomal bands, but also on the gene distribution in chromosomes.

In the first investigation of this kind (Saccone *et al.* 1992), it was shown that the hybridization of a DNA fraction derived from H3 isochores produced the highest concentration of signals on two largely coincident subsets of R (reverse) bands: (1) the T (telomeric) bands (Dutrillaux 1973), which are the most heat-denaturation-resistant R bands; and (2) the chromomycin A3-positive, DAPI-negative bands (Ambros & Sumner 1987), which are the GC-richest bands of human chromosomes (DAPI is 4,6-diamino-2-phenylindole). *In-situ* hybridization of H3 isochore DNA established that T bands comprise GC-rich gene-rich single-copy DNA because the contribution of repetitive DNAs was suppressed by competition with excess unlabeled total human DNA and Alu sequences (or SINEs). This rules out the possibility that T bands correspond to GC-rich satellite DNA(s).

In subsequent investigations, the chromosomal locations of the other isochore families L1+L2, H1 and H2 were studied (Saccone *et al.* 1993). This established: (1) that T bands contain not only H3 isochores, but also H2 and some H1 isochores; (2) that R′ bands (namely R bands exclusive of T bands) are formed, on the average and to almost equal extents, by H1 and L isochores, H2 and H3 isochores being only rarely present; and (3) that G bands essentially consist of L isochores, H1 isochores being present at low levels.

In more recent work (Saccone *et al.* 1996), it was shown: (1) that the number of bands characterized by strong hybridization signals, which were called T or $H3^+$, is 28; (2) that 31 additional R bands, called T′ or $H3^*$ bands, also contain H3 isochores, although at a lower concentration than $H3^+$ bands; (3) that H2 and H3 isochores, as a rule, appear to have the same band distribution; and (4) that the remaining R bands (about 140 out of 200, at a resolution of 400 bands), called $H3^-$ bands, do not contain any detectable H3 isochores.

The existence of three distinct sets of R bands was further supported: (1) by the different compositional features of genes located in them; (2) by the very low gene density of chromosomes 13 and 18, in which all R bands are $H3^-$ bands; (3) by the compositional map of a $H3^*$ band, Xq28, and of a $H3^-$ band, 21q11.2 (see below); (4) by the overwhelming presence of long (> 50 kb) GC-rich sequences in $H3^+/H3^*$ bands and GC-poor sequences in $H3^-/G$ bands; and (5) by the large degree of coincidence of $H3^+$ and $H3^*$ bands with CpG island-positive bands (Saccone *et al.* 1996, Craig & Bickmore 1994). These observations have already been extended to the genomes of other mammals (Cacciò *et al.* 1994, Saccone *et al.* 1997).

A complementary approach, molecular compositional mapping (Bernardi 1989, Gardiner *et al.* 1990, Bettecken *et al.* 1992, Pilia *et al.* 1993, De Sario *et al.* 1996, 1997), provided more detailed information on some human genome regions; some of them, such as Xq28 and the 21cen-21q21 region, were investigated using a novel methodology (De Sario *et al.* 1995). More recently, we found (Federico *et al.* 1998) that the chromosomal bands that contain H3 isochores replicate mainly (in the case of $H3^+$ bands), or largely (in the case of $H3^*$ bands), at the onset of S phase, whereas chromosomal bands containing no H3 isochores ($H3^-$ bands) replicate later, and G bands replicate even later. Expectedly, CpG islands, which are most abundant in H3 isochores (Aïssani & Bernardi 1991a, 1991b, Jabbari & Bernardi 1998) were found to replicate coordinately early in the S phase (Delgado *et al.* 1998).

Now we report *in-situ* hybridization results obtained at higher resolution (about 850 bands) and show that hybridization signals are generally localized on R bands derived from the $H3^+$ and $H3^*$ bands, but with different features. Moreover, we show the existence of other $H3^+$ bands, not previously identified, characterized by the presence of very small amounts of H3 isochores.

Materials and methods

DNA preparation

High-molecular-weight DNA was prepared from a human placenta and fractionated in a Cs_2SO_4/BAMD density gradient, as previously reported (Cuny *et al.* 1981); BAMD is 3,6-bis(acetato-mercurimethyl) 1,4-dioxane. Compositional DNA fractions were analyzed by ultracentrifugation in CsCl to identify the fraction(s) containing H3 isochores. The CsCl profile of fraction 8, which consisted of DNA derived from H3 isochores, is presented (together with those of the other fractions) in Saccone *et al.* (1996). DNA from fraction 8, as such and as cloned in a lambda vector (Saccone *et al.* 1996), was used in the present work.

Chromosome preparation and in-situ hybridization

Metaphase and prometaphase chromosomes were obtained by phytohemagglutinin/stimulated peripheral blood lymphocytes and prepared using standard cytogenetic procedures. Cells were synchronized by using thymidine/BrdU standard techniques. *In-situ* hybridization and detection were performed according to a previous protocol (Saccone *et al.* 1996).

Hybridization signals were assigned to specific prometaphase chromosomal bands by measuring the relative distances between them and the reference points of each chromosome arm, as described by Francke (1994; Table 1).

Results and discussion

DNA from the H3 isochore family was hybridized to human prometaphase chromosomes using the biotin/avidin system under conditions in which repeated sequences were competed out (Figure 1). After hybridization, each chromosome was identified by its characteristic H3 banding. The distribution of the hybridization signals on chromosomal bands was determined, taking as a reference the band ideogram and the band sizes reported by Francke (1994; Table 1). The distribution of the hybridization signals on chromosomes derives from the analysis of at least 14 hybridized chromosomes for each chromosome pair. These results indicate that the H3 isochores are only located on a small number of R_{850} bands (Figure 2) and on none of the G_{850} bands (R_{400}, G_{400}, R_{850}, G_{850} indicate the R and G bands at resolutions of 400 and 850, respectively). In fact, H3 hybridization signals covered almost all, many, and only a few of the R_{850} bands derived from R_{400} $H3^+$ bands, $H3^*$ bands, and $H3^-$ bands, respectively.

Indeed, 23 out of the 28 R_{400} $H3^+$ bands only yielded R_{850} bands containing H3 isochores (Figure 3), whereas only some of the R_{850} bands originating from the other five R_{400} $H3^+$ bands showed H3 hybridization signals. For example (Figure 4), the R_{400} $H3^+$ band 11q13 is resolved into three R_{850} $H3^+$ bands (q13.1, q13.3, and q13.5), and two G_{850} bands (q13.2, and q13.4), whereas the $H3^+$ band 11p15 was one of the five exceptions, with only one (11p15.5) of the three derived R_{850} bands showing hybridization signals. Globally, the coverage of the R_{850} bands derived from the 28 R_{400} $H3^+$ bands was 91% (this calculation was based on band sizes, as assessed in Francke, 1994). In no case did G_{850} bands derived from the R_{400} $H3^+$ bands show hybridization signals.

In contrast, only some of the R_{850} bands derived from 23 out of 31 R_{400} $H3^*$ bands showed hybridization signals (see Figures 2 and 3). For example, the $H3^*$ band 11q23 (Figure 4) yielded only one R_{850} $H3^+$ band (11q23.3; in fact, only the distal part of it was $H3^+$), whereas the other R_{850} band (11q23.1) was $H3^-$. The remaining eight $H3^*$ bands showed the features observed in the vast majority of $H3^+$ bands, in that all the derived R_{850} bands were $H3^+$. Globally, 66% (in terms of size; see above) of the R_{850} bands derived from the R_{400} $H3^*$ bands contained H3 isochores.

As far as the R_{400} $H3^-$ bands are concerned, the higher resolution allowed the identification of 20 bands containing H3 isochores that had not been detected at the lower resolution (Saccone *et al.* 1996). The majority of these bands were located close to other $H3^+$ or $H3^*$ bands (see bands 5q33.1, 6p21.1, and 12q24.13) and were very thin (see bands 1p13.3, and 7p13). Only some of the R_{850} bands derived from these 20 R_{400} $H3^-$ bands (see Figures 2 and 3) exhibited hybridization signals (see band 11p11.2 in Figure 4). Moreover, in a number of cases, the signal was thinner than the corresponding R_{850} bands (see Figure 2), indicating that only part of the R_{850} band contained H3 isochores. Globally, only 9% (in terms of size; see above) of the R_{850} bands derived from the $H3^-$ bands contained H3 isochores. If only the R_{850} bands derived from the 20 $H3^-$ bands (listed in Figure 3) were taken into consideration, the coverage was 47%. Incidentally, previous work (Saccone *et al.* 1996) had shown that H2 and H3 isochores colocalize on metaphase chromosomes, with only four exceptions (the telomeric bands, 3q29, 6q27, 13q34, and 20p13) which were $H2^+$ and $H3^-$. Now these bands were shown to be $H3^+$, indicating a colocalization of H3 and H2 isochores in these cases as well.

Finally, G bands did not reveal the presence of H3 isochores, the only exceptions being two G_{400} bands, 1p36.2 and 19q13.4, which yielded two R_{850} $H3^+$ bands, 1p36.22, and 19q13.42, respectively (Figure 2).

On the basis of the present results and of the estimated band sizes (Francke 1994), it can be calculated that about 17% of all bands at a 850-band resolution contain H3 isochores and that about 9%,

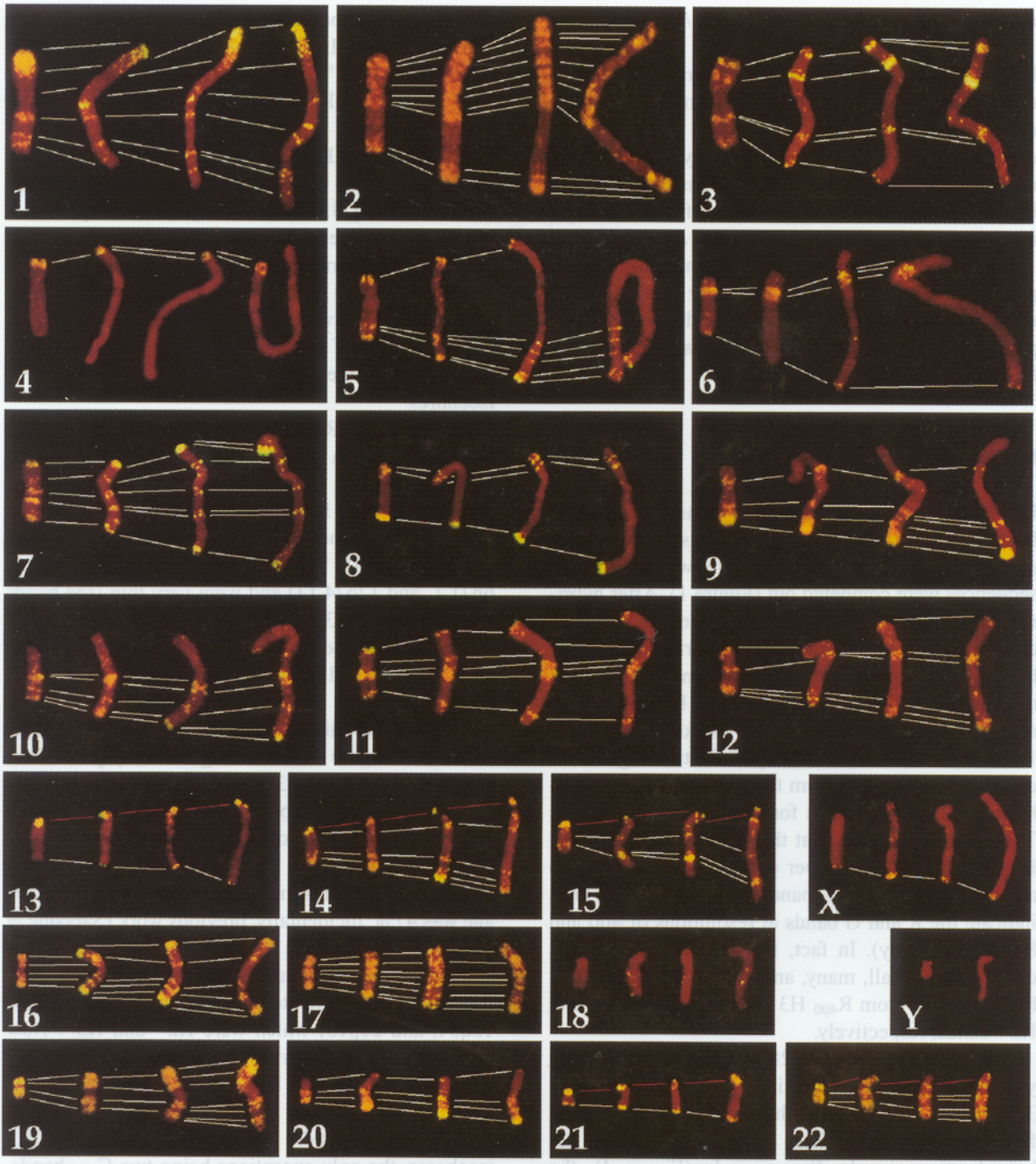
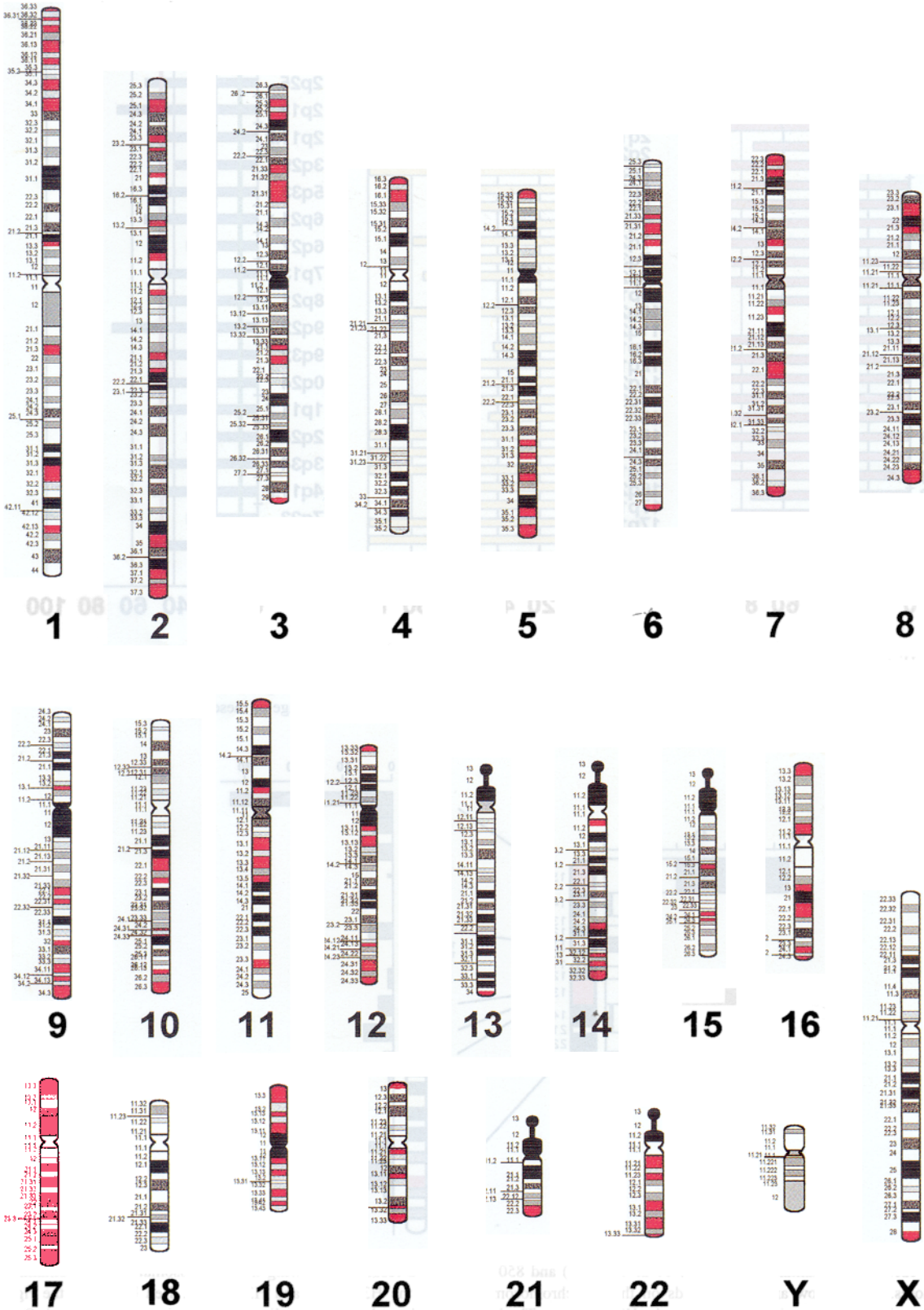


Figure 1. Human chromosomes hybridized with the biotin-labeled DNA from the H3 isochore family, at different levels of resolution. The hybridized regions were visualized by fluorescein (yellow signals) and chromosomes were red-stained with propidium iodide. Each panel presents chromosomes with a band resolution ranging from about 300 to about 850.

Figure 2 (opposite). Ideogram of human chromosomes at a 850 band resolution (Francke 1994) showing the $H3^+$ bands as red bands.


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Y
X

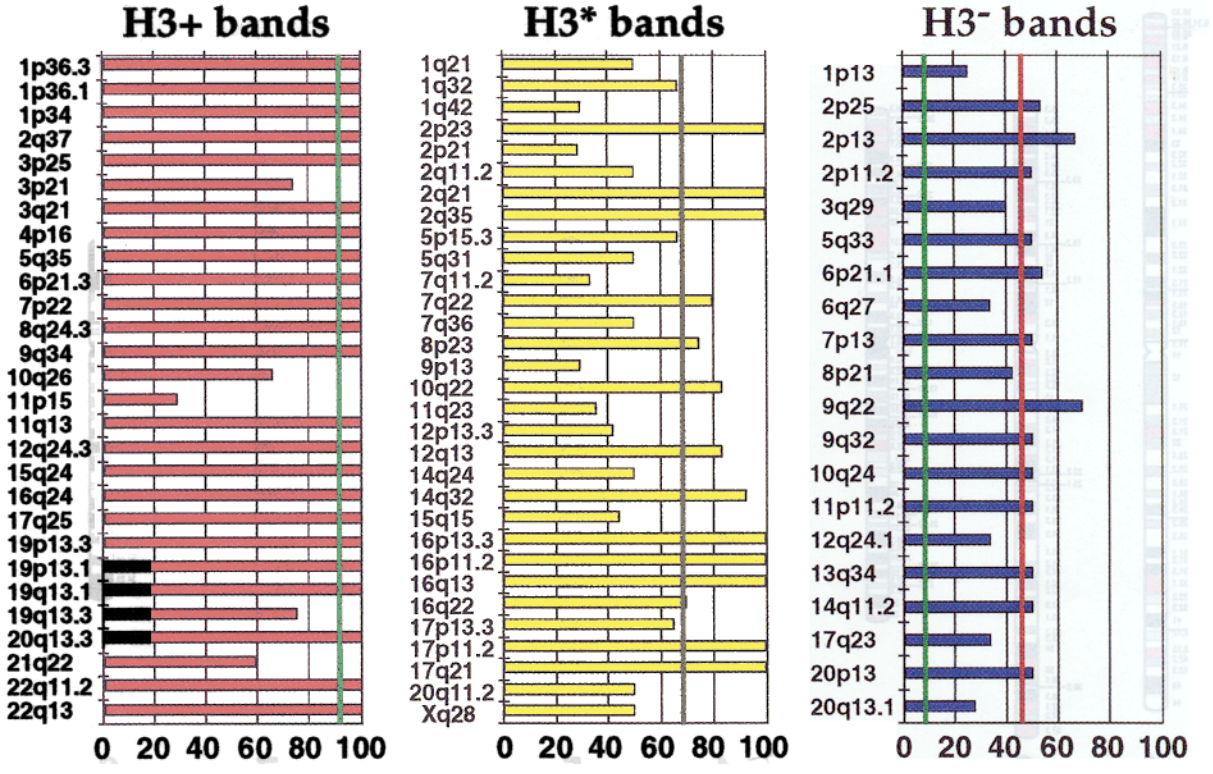


Figure 3. Histograms showing the relative hybridization coverage of R_{850} bands derived from $H3^+$, $H3^*$, and $H3^-$ bands. The global percentage (average of all bands belonging to each class) is indicated by a green linè in each histogram. In the case of $H3^-$ bands, only those containing H3 isochores were reported, and the red line is the average of hybridization coverage of these 20 bands.

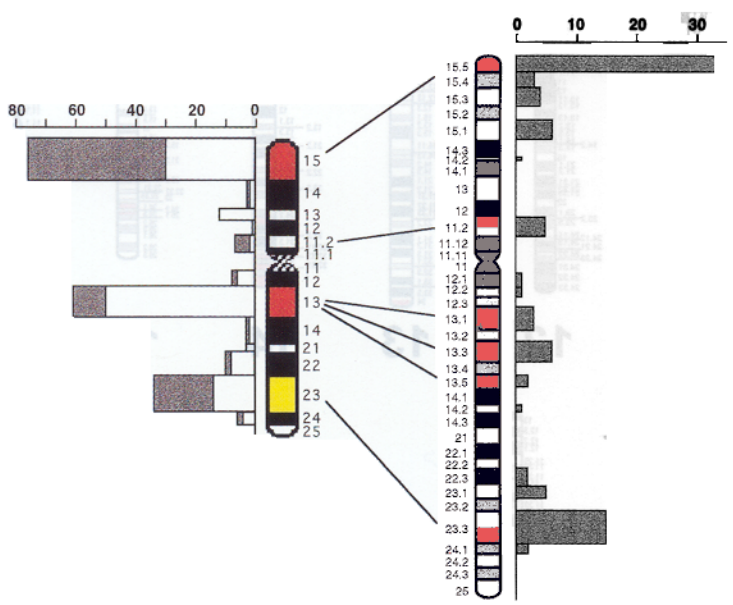


Figure 4. Ideogram of human chromosome 11 at 400 (left) and 850 (right) band resolution showing the chromosomal regions containing H3 isochores. Red, yellow and white bands on the left chromosome indicate the $H3^+$, $H3^*$ and $H3^-$ bands. Red bands on the right chromosome indicate the regions hybridizing the H3 isochores. The histograms show the genes localized on each chromosomal band. Only genes localized on single bands were listed, those assigned to more than one band not being taken into consideration. In the left histogram, the white and gray bars indicate the genes localized at low and high resolution, respectively. In the right histogram, only the genes localized at high resolution are indicated. The numbers of genes are indicated by the scales of the histograms.

6%, and 2% of them are derived from R_{400} $H3^+$, H3* and $H3^-$ bands, respectively.

The number of genes localized on chromosomal bands considerably increased in recent years, especially at high resolution. An analysis of 1302 localized genes showed a correlation between the human GC-rich regions and the chromosomal bands characterized by the highest numbers of genes (Saccone *et al.* 1996). This was done at a resolution of 400 bands per haploid genome. In the present work, we analyzed the genes localized at high resolution (Collins *et al.* 1996, and data listed in the web site: http://cedar.genetics.soton.ac.uk/public_html), and observed, in the majority of cases, a very high degree of coincidence between the regions characterized by the highest gene levels and the bands containing the H3 isochores described in Figure 2. As an example, Figure 4 shows the situation found in chromosome 11: (1) at low resolution (the left part of Figure 4), the $H3^+$ and H3* bands correspond to the highest gene-level regions, as previously observed using a lower number of genes (Saccone *et al.* 1996); whereas (2) at higher resolution (right part of Figure 4), the highest gene levels were in bands 11p15.5 and 11q23.3 (derived from the 11p15 and 11q23 bands, respectively), where H3 isochores were located. The $H3^+$ band 11q13 contains a small number of genes localized at high resolution but, in this case, all the three R_{850} bands contain a comparable number of genes (as expected considering the distribution of the H3 isochores) and no genes were localized in the G_{850} bands.

The present results lead to several conclusions:

1. Since the colocalization of H2 and H3 isochores (which represent 12% of the human genome) in R_{850} $H3^+$ bands appears now to be the rule, the fraction of these isochores in those bands (which represent 17% of the total genome) correspond to the majority, 70%, of the DNA contained in them.
2. In some cases, however, the coverage of R_{850} $H3^+$ bands by hybridization signals is overestimated; for example, the present experiments suggest that almost 50% of band Xq28 is $H3^+$, whereas compositional mapping has shown that only 5% is formed by H3 isochores (De Sario *et al.* 1996).
3. In a number of R_{850} $H3^+$ bands (Figure 2), H3 hybridization coverage was limited to a fraction of the band. This indicates that the present results provide information concerning a resolution higher than 850 bands; thus, they may correspond, in many cases, to the practical highest resolution that can be attained, namely 1250 bands (Drouin *et al.* 1991, 1994).
4. 83% of the bands shown in Figure 2, namely the R_{850} $H3^-$ and the G_{850} bands, exhibit low or very low gene concentrations. Since genome size is remarkably constant in mammals and since such regions are conserved in syntenic regions of chromosomes from mammalian orders that diverged about 100 millions years ago (Scherthan *et al.* 1994, Raudsepp *et al.* 1996, Saccone *et al.* 1997, Morescalchi *et al.* 1997, Chowdhary *et al.* 1998), this suggests some functional role for the gene-poor majority of the genome.
5. Finally, the present results are relevant for the choice of the regions of the human genome that deserve sequencing priority; interestingly, these regions correspond to gaps in the physical map of the human genome (Chumakov *et al.* 1995, see Saccone *et al.* 1996). The difficulty experienced in cloning these regions into YACs and/or in avoiding high levels of chimerism and deletion is most probably related to their high recombination level, a property which apparently is conserved when these regions are cloned in yeast.

References

Aïssani B, Bernardi G (1991a) CpG islands: features and distribution in the genomes of vertebrates. *Gene* **106:** 173–183.

Aïssani B, Bernardi G (1991b) CpG islands, gene and isochores in the genomes of vertebrates. *Gene* **106:** 185–195.

Ambros PF, Sumner AT (1987) Metaphase bands of human chromosomes, and distinctive properties of telomeric regions. *Cytogenet Cell Genet* **44:** 223–228.

Bernardi G (1989) The isochore organization of the human genome. *Annu Rev Genet* **23:** 637–661.

Bernardi G (1995) The human genome: organization and evolutionary history. *Ann Rev Genet* **29:** 445–476.

Bernardi G, Olofsson B, Filipski J *et al.* (1985) The mosaic genome of warm-blooded vertebrates. *Science* **228:** 953–958.

Bettecken T, Aïssani B, Mueller CR, Bernardi G (1992) Compositional mapping of the human dystrophin gene. *Gene* **122:** 329–335.

Cacciò S, Perani P, Saccone S, Kadi F, Bernardi G (1994) Single-

copy sequence homology among the GC-richest isochores of the genomes from warm-blooded vertebrates. *J Mol Evol* **39:** 331–339.

Chowdhary BP, Raudsepp T, Frönicke L, Scherthan H (1998) Emerging patterns of comparative genome organization in some mammalian species as revealed by Zoo-FISH. *Genome Res* **8:** 577–589.

Chumakov IM and 60 other authors (1995) A YAC contig map of the human genome. *Nature* **377** (suppl): 175–297.

Collins A, Frezal J, Teague J, Morton NE (1996) A metric map of humans: 23,500 loci in 850 bands. *Proc Natl Acad Sci USA* **93:** 14771–14775.

Craig M, Bickmore WA (1994) The distribution of CpG islands in mammalian chromosomes. *Nature Genet* **7:** 376–382.

Cuny G, Soriano P, Macaya G, Bernardi G (1981) The major components of the mouse and human genomes: preparation, basic properties and compositional heterogeneity. *Eur J Biochem* **111:** 227–233.

Delgado S, Gómez M, Bird A, Antequera F (1998) Initiation of DNA replication at CpG islands in mammalian chromosomes. *EMBO J* **17:** 2426–2435.

De Sario A, Geigl E-M, Bernardi G (1995) A rapid procedure for the compositional analysis of yeast artificial chromosomes. *Nucl Acids Res* **23:** 4013–4014.

De Sario A, Geigl E-M, Palmieri G, D'Urso M, Bernardi G (1996) A compositional map of human chromosome band Xq28. *Proc Natl Acad Sci USA* **93:** 1298–1302.

De Sario A, Roizës G, Allegre N, Bernardi G (1997) A compositional map of the cen-q21 region of human chromosome 21. *Gene* **194:** 107–113.

Drouin R, Lemieux N, Richer CL (1991) Chromosome condensation from prophase to late metaphase: relationship to chromosome bands and their replication time. *Cytogenet Cell Genet* **57:** 91–99.

Drouin R, Holmquist G, Richer CL (1994) High resolution replication bands compared with morphologic G- and R-bands. *Adv Hum Genet* **22:** 47–115.

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

Federico C, Saccone S, Bernardi G (1998) The gene-richest bands of human chromosomes replicate at the onset of the S-phase. *Cytogenet Cell Genet* **80:** 83–88.

Francke W (1994) Digitized and differentially shaded human chromosome ideograms for genomic applications. *Cytogenet Cell Genet* **65:** 206–219.

Gardiner K, Aïssani B, Bernardi G (1990) A compositional map of human chromosome 21. *EMBO J* **9:** 1853–1858.

Jabbari K, Bernardi G (1998) CpG doublets, CpG islands and Alu repeats in long human DNA sequences from different isochore families. *Gene* **224:** 123–127.

Morescalchi MA, Schempp W, Consigliere S, Bigoni F, Wienberg J, Stanyon R (1997) Mapping chromosomal homology between humans and the black-handed spider monkey by fluorescence *in situ* hybridization. *Chromosome Res* **5:** 527–536.

Mouchiroud D, D'Onofrio G, Aïssani B, Macaya G, Gautier C, Bernardi G (1991) The distribution of genes in the human genome. *Gene* **100:** 181–187.

Pilia G, Little RD, Aïssani B, Bernardi G, Schlessinger D (1993) Isochores and CpG islands in YAC contigs in human Xq26.1-qter. *Genomics* **17:** 456–462.

Raudsepp T, Frönicke L, Scherthan H, Gustavsson I, Chowdhary BP (1996) Zoo-FISH delineates conserved chromosomal segments in horse and man. *Chromosome Res* **4:** 218–225.

Saccone S, De Sario A, Della Valle G, Bernardi G (1992) The highest gene concentrations in the human genome are in T-bands of metaphase chromosomes. *Proc Natl Acad Sci USA* **89:** 4913–4917.

Saccone S, De Sario A, Wiegant J, Raap AK, Della Valle G, Bernardi G (1993) Correlations between isochores and chromosomal bands in the human genome. *Proc Natl Acad Sci USA* **90:** 11929–11933.

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

Saccone S, Cacciò S, Perani P *et al.* (1997) Compositional mapping of mouse chromosomes and identification of the gene-rich regions. *Chromosome Res* **5:** 293–300.

Scherthan H, Cremer T, Arnason U, Weier HU, Lima-de-Faria A, Fronicke L (1994) Comparative chromosome painting discloses homologous segments in distantly related mammals. *Nature Genet* **6:** 342–347.

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