Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 1298–1302, February 1996
Genetics

A compositional map of human chromosome band Xq28

(human genome/isochores/yeast artificial chromosomes)

ALBERTINA DE SARIO*†, EVA-MARIA GEIGL*, GIUSEPPE PALMIERI‡, MICHELE D'URSO‡, AND GIORGIO BERNARDI*§

*Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2, Place Jussieu, 75005 Paris, France; and ‡International Institute of Genetics and Biophysics, Naples, Italy

Communicated by Susumu Ohno, Beckman Research Institute of the City of Hope, Duarte, CA, October 10, 1995

ABSTRACT **The molar fractions of guanine plus cytosine (GC) in DNA were determined for 36 yeast artificial chromosomes (YACs) which almost completely cover human chromosome band Xq28, a terminal reverse band, corresponding to about 8 Mb of DNA. This allowed the construction of the most complete compositional map to date of a chromosomal band; three regions were observed: (*i*) a proximal 3.5-Mb region formed by GC-poor L and GC-rich H1 isochores; (*ii*) a middle 2.2-Mb region essentially formed by a GC-rich H2 isochore and a very GC-rich H3 isochore separated by a GC-poor L isochore, YACs from this region being characterized by a striking compositional heterogeneity and instability; and (*iii*) a distal 1.3-Mb region exclusively formed by GC-poor L isochores. Gene and CpG island concentrations increased with the GC levels of the isochores, as expected. Xq28 exemplifies a subset of reverse bands which are different from the two other subsets, namely from telomeric bands, which are characterized by specific cytogenetic properties and by the predominance of H2 and H3 isochores, and from the majority of reverse bands, which do not contain H2 and H3 isochores.**

The human genome is a mosaic of large DNA segments (>300 kb, on the average), the isochores, which are compositionally uniform (above a size of 3 kb), but belong to a small number of families that cover a very wide (30–60%) GC range (where GC is the mole fraction of guanosine plus cytosine in DNA). Hybridization of gene probes on compositional fractions led to the discovery of a strikingly nonuniform gene distribution in the human genome (1–4), which is representative of most mammalian genomes (5, 6). Indeed, gene concentration is low and constant in GC-poor isochore families L1 and L2, which represent 62% of the genome, increases in GC-rich isochores H1 and H2, which correspond to 22% and 9% of the genome, respectively, and reaches the highest value (20 times higher than in GC-poor isochores) in isochore family H3, which represents only 3–4% of the genome, the remaining 4% or so being formed by satellite and ribosomal DNAs.

Because of the correlation between GC levels and gene concentration in isochores (2, 3), compositional mapping of human genome regions is equivalent to mapping gene concentration. So far, compositional mapping (2) has been performed by hybridizing probes for human single-copy DNA sequences (which were already physically mapped) on compositional fractions from human DNA. This approach led to the establishing of partial compositional maps of the long arm of chromosome 21 (7) and of the q26–q28 region of the X chromosome (8), as well as of a complete compositional map of the dystrophin locus, which is located on the short arm of the X chromosome (9). Another approach (10) has provided a compositional map around the cystic fibrosis locus, which is mapped to chromosome 7.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

All these compositional maps were, however, either very fragmentary (7, 8) or limited to loci which were not only relatively short, 1–2 Mb, but also only consisted of very GC-poor isochores (9, 10).

Here, we have taken advantage of the fact that Xq28, an 8-Mb band, is almost completely covered by yeast artificial chromosomes (YACs) assembled in 5 contigs separated only by small gaps (ref. 11; see also *Materials and Methods*), and of a new approach to determine the GC levels of YACs (12) to establish the most complete compositional map to date of Xq28. The terminal reverse band (R band) of the long arm of human X chromosome, Xq28, is one of the most thoroughly studied human chromosomal regions (8, 11, 13–16), largely due to the fact that ≈15% of the ≈170 genetic disease genes localized on the X chromosome map to Xq28 (17), which corresponds to less than 5% of the X chromosome DNA. Interestingly, Xq28 appears to be typical of a subset of R bands distinct from both telomeric bands (T bands), which are essentially formed by H2 and H3 isochores, and from the majority of R bands, which do not contain H2 and H3 isochores.

MATERIALS AND METHODS

YACs. A total of 34 YACs were selected from those mapped to chromosome band Xq28 by Palmieri *et al.* (11). Two additional YACs (79H4 and 519A2) belong to a 600-kb contig (18) that fills the gap between contigs B and C in the map of Palmieri *et al.* (ref. 11; see Fig. 1). Several YACs were analyzed in two to three independent experiments (see Table 1) to verify the reproducibility of the method used to measure buoyant densities. In most cases (see Table 1), the range of buoyant densities found for each YAC was 1 mg/cm^3 (corresponding to a variation in GC content of about 1%).

YACs 79H4, 845, 1692, 660, 1069, 411, and 901 produced a bimodal profile in CsCl with a major and a minor peak. In these cases, the GC levels of both peaks are reported in Table 1. In contrast, only the buoyant density of the major peak was reported in Fig. 1, because it represents the base composition of the majority of the analyzed DNA. Since buoyant densities are additive, the average GC level of the combined major and minor peaks can be calculated by taking into account their relative amounts. The small percentage of the minor peaks (10–20%) make, however, this correction negligible. Indeed, even in the most unfavorable case, that of YAC 845, the minor (20%) 1.700 g/cm^3 component would only lower the buoyant density of the major (80%) component from 1.708 g/cm^3 to

Abbreviations: GC, mole fraction of guanosine plus cytosine in DNA; G band(s), Giemsa band(s); GC_3, average GC level of third codon positions; T band(s), telomeric band(s); R band(s), reverse band(s); R′ band(s), R band(s) not belonging to the subset of T band(s); YAC(s), yeast artificial chromosome(s); SINEs, short interspersed DNA sequence elements; LINEs, long interspersed DNA sequence elements.

†Present address: Laboratoire de Structure, Fonction et Evolution du Génome Eucaryote, Institut de Biologie, 4 Boulevard Henri IV, 34060 Montpellier, France.
§To whom reprint requests should be addressed.

1.7064 g/cm^3, namely by about 1.5% GC. In some cases (YACs 79H4, 845, 1692, and 660), however, the two peaks were useful to understand better some compositional transitions (see *Results*).

Four cocloned YACs were included in the analysis. YACs 526, 540, and 780 contained hamster DNA in addition to human DNA; YACs 1561 was cocloned with noncontiguous sequences mapped to Xq28 (11). These YACs were used because indications existed that in these cases cocloning did not significantly distort the estimates of GC levels. Indeed, these YACs had a base composition very close to those of overlapping, not cocloned YACs. In any case, none of these YACs was of critical importance for the conclusions drawn.

Buoyant Densities and GC Levels of YACs. The buoyant densities and the GC levels of the YACs were estimated according to ref. 12.

GC Levels of Isochores Hosting Sequenced Genes. DNA sequences of 13 coding sequences mapped to Xq28 were obtained from GenBank (release 87; February 1995) and EMBL data bank (release 41; December 1994). The GC levels of the isochores hosting the analyzed genes were estimated by using the equation $GC_3 = 2.74\ GC - 64.5$ that correlates GC_3, the average GC levels of third codon positions, with the GC level of the isochores hosting the corresponding genes (4, 19).

RESULTS

The buoyant densities, GC levels, and sizes of the analyzed YACs are given in Table 1. The 36 YACs analysed in this work are located in a region which begins at ≈7.3 Mb from the telomere and ends at 0.3 Mb from the terminal repeat of the long arm of the X chromosome, thus encompassing 7 Mb out of the ≈8 Mb which can be estimated from ref. 13 to form band Xq28. The compositional map of Xq28 (Fig. 1) comprises three different regions:

(*i*) The proximal region (3.5 Mb) begins 7.3 Mb from the telomere and ends at a site located between the *GABRA* and *DXS1104* loci. This region is preceded by a GC-poor segment. Indeed, YAC 493, which is 42% GC (determined as a result of three independent analyses), largely overlaps with YAC 571, which is 45% GC (a value consistent with the GC level of the distal overlapping clone 873). This indicates that the segment preceding YAC 571 is lower in GC than YAC 493.

The proximal region proper comprises five 0.6- to 0.8-Mb stretches, characterized by GC levels ranging from 39% to 46%. The first stretch (YACs 571 and 873) is 45% GC and is followed, after a transition coinciding with the gap of <100 kb separating contigs A and B (11), by alternating GC-poor and GC-rich stretches.

(*ii*) The middle region extends from the *GABRA* and *DXS1104* loci to the *G6PD* gene and is extremely heterogeneous in base composition. About half of the YACs (7 of 15) from the region show bimodal profiles in CsCl (see *Materials and Methods* and *Discussion*). Thus, the precise location of compositional discontinuities and the isochore sizes are often difficult to define in this region.

The GC level first rises from 39% (at the end of the proximal region) to 49%. The border with the proximal region might be located in the left half of YAC 79H4, which shows a bimodal profile in the CsCl gradient, with a major peak corresponding to 41% GC (2% higher than the preceding YAC) and a minor peak at 46% GC. The latter peak corresponds to the base composition of the distal overlapping YAC 519A2. The profile of YAC 845 was characterized by a major peak at 49% GC and a minor one at 41% GC that might be due to the distal GC-poor DNA. The presence of a compositional change is corroborated by two YACs, 1692 and 660, which overlap YAC 845 and exhibit bimodal profiles with major peaks having a GC level of 47% and minor peaks showing 43% GC. Distal to the *BGN* gene, the GC level decreases to 42% for 0.2 Mb.

Table 1. Properties of the YACs used in this study

YAC*	Size†	Buoyant densities, g/cm^3‡	GC, %§
		Proximal region of Xq28	
493	250	**1.701; 1.700; 1.701**	41.8; 40.8; 41.8
571	360	**1.704**	44.9
873	370	**1.704**	44.9
1561	250	1.701	41.8
388	160	**1.700; 1.702; 1.702**	40.8; 42.9; 42.9
700	190	**1.701; 1.702**	41.8; 42.9
316	280	1.703	43.9
545	360	1.707; 1.705; 1.704	48.0; 45.9; 44.9
119	300	1.705	45.9
415	430	1.703	44.9
1354	230	**1.700**	40.8
813	280	**1.701**	41.8
400	470	**1.703; 1.702**	43.9; 42.9
2413	600	**1.699**	39.8
422	340	**1.698; 1.697**	38.8; 37.7
824	260	**1.698**	38.8
		Middle region of Xq28	
79H4	300	**1.700 + 1.705**¶	40.8 + 45.9
519A2	320	1.705	45.9
1694	100	1.703	43.9
526	250	1.708	49.0
845	490	1.708 + 1.700¶	49.0 + 40.8
1692	300	1.706 + 1.702¶	46.9 + 42.9
660	130	1.706 + 1.702¶	46.9 + 42.9
540	250	**1.701; 1.701**	41.8; 41.8
807	130	**1.702**	42.9
1082	65	**1.711; 1.712**	52.0; 53.1
1069	110	1.713 + 1.705¶; 1.715‖	54.1 + 45.9; 56.1
411	100	1.706; 1.711 + 1.707¶; 1.705‖	46.9; 52.0 + 48.0; 45.9
901	90	1.710 + 1.706¶, 1.709‖	51.0 + 46.9; 50.0
780	120	1.713	54.1
917	90	1.711	52.0
		Distal region of Xq28	
527	360	**1.699**	39.8
705	370	**1.699**	39.8
348	230	**1.699**	39.8
37	245	**1.698**	38.8
364	270	**1.698; 1.698**	38.8; 38.8

*YACs are listed in a centromeric to telomeric order. Co-cloned YACs are underlined.

†YAC sizes are as reported in ref. 11.

‡Buoyant densities of YACs as determined in the present work. Values in boldface type correspond to YACs which were found to have the same or very close buoyant density in different experiments (unless they were co-cloned) or which were overlapping or contiguous with other YACs having the same or a very close buoyant density (vertical lines).

§GC% determined according to procedures given in ref. 12.

¶Bimodal YACs; in these cases, the two values are separated by a + sign and the buoyant density of the major band is underlined.

‖The values of buoyant density measured for the very GC-rich YAC 411 were not reproducible. A GC level higher than 50% (1.709 g/cm^3) was found during a series of preliminary analyses. After some passages in culture, the estimated GC level was 1.705 to 1.707 g/cm^3, and sometimes two DNA peaks were observed (1.707 + 1.711 g/cm^3, the latter being the major peak). YAC 411 shows deletions for internal tandem repeat units (11, 20). Thus, it is possible that initially the buoyant density of the full genomic sequence was measured, while after some passages in culture, a sequence whose GC level was a function of the number of deleted tandem units was analyzed. Similar phenomena were shown by the overlapping clones 1069 and 901.

Proximal to the *L1CAM* gene, the GC level sharply increases again to a very high level (53–56%), which extends for about 0.7 Mb between the *L1CAM* and *G6PD* genes and ends with a very sharp transition (down to 40% GC) localized between the *G6PD* and *F8C3* genes.

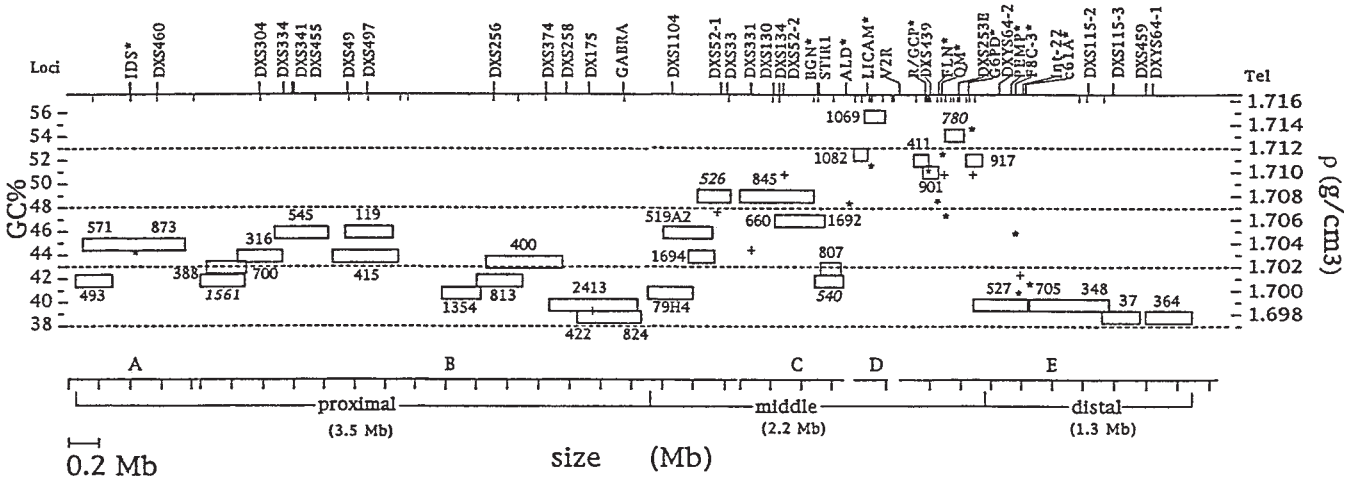

FIG. 1. Compositional map of human chromosome band Xq28. YACs, represented as boxes with a height corresponding to 1% GC, are positioned along the physical map on the horizontal scale, and according to their buoyant density and GC content (right and left ordinates, respectively) on the vertical scale. Asterisks represent GC levels of isochores as calculated from GC_3 values of mapped genes (also indicated by asterisks on gene names on the top part of the figure with the exception, for lack of space, of the gene for *EDMD*). Crosses represent GC levels of isochores as evaluated by hybridization of single-copy probes on compositional DNA fractions (8). The top part of the figure shows a number of mapped genes and loci, as well as the location of CpG islands (thin vertical lines); the region between *R/GCP* and *DXS439* comprises seven CpG islands (unresolved in the figure). The bottom part of the figure shows YAC contigs A–E (11), as well as the proximal, middle, and distal regions described in this paper.

(*iii*) The distal region extends from the *G6PD* gene to the end of the physical map 0.3 Mb from the telomere. In this region, five YACs spanning 1.3 Mb were analyzed. The GC levels of the proximal three YACs were 40%, while those of the distal two YACs were 39%.

The compositional map based on YAC analysis (see Fig. 1) is supported by all other available findings: (*i*) the GC levels of isochores containing 13 genes mapped to Xq28, as calculated from the GC levels of the third codon positions of the genes (GC_3) and from the GC levels calculated for eight regions surrounding probed sequences (8); in the cases of two genes, *QM* and *PEMP*, and the sequence *DXS331*, the calculated GC levels of isochores were, however, different from expectations; these exceptions corresponding to sequences deviating from the general relationship between GC_3 and isochore GC (see ref. 19); (*ii*) the GC levels determined by HPLC analysis of four cosmids mapped to YAC 917 (49.3%, 50.5%, 52.6%, and 52.7%); (*iii*) the GC level (52.7%) of a sequenced, 52-kb region, including, and centromeric to, the *G6PD* gene (21); and (*iv*) the GC level (45%) of a sequenced, 130-kb region around the *IDS* locus (22).

DISCUSSION

The Compositional Map of Xq28. This work has provided the most complete compositional map to date of a whole chromosome band at a high level of resolution. Indeed, apart from some negligible gaps (50 kb or less in size, such as those between YACs 873/1561, 545/415, 526/845, 540/1082, 901/780, 527/705, and 37/354), the compositional map presents only two gaps, between YACs 415 and 1354 and between YACs 1069 and 411, respectively. However, the first one is in a rather uncritical region for the conclusions drawn here, and the second one corresponds to a region in which no YACs are yet available.

Basically, Xq28, which is 43.4% GC (as calculated from the present compositional map), comprises a proximal region (43% GC) formed mostly by H1 isochores, a region (47.2% GC) formed by H2 and H3 isochores, and a distal region (39.5% GC) formed by L isochores. An interesting feature of the map is that GC-rich and very GC-rich regions are all separated, preceded, and followed by GC-poor isochores, which extend farther than the compositional map of Fig. 1 on both ends. Indeed, compositional discontinuities may represent the major difference between Giemsa bands (G bands) and R bands (see also refs. 7 and 8). Globally, L isochores formed 49% of the band, H1 isochores formed 41%, and H2 and H3 isochores formed 4.6% and 4.8%, respectively. These values correspond to the average values in R bands not belonging to the subset of T bands (R′ bands) (23) except for the higher (double) values of H2 and H3.

Isochores being defined as regions of similar GC level, the isochore map of Xq28 can be deduced from Fig. 1 by taking into account the similarity of composition of contiguous YACs. The compositional map allowed us to investigate isochore sizes and borders, the compositional heterogeneity, and the concentration of genes, CpG islands, short interspersed DNA sequence elements (SINEs), and long interspersed DNA sequence elements (LINEs) and to draw some general conclusions on chromosome bands.

Isochore Sizes and Borders in Xq28. Isochore sizes range, in Xq28, from 0.2 Mb in the short GC-poor isochore located between the *BGN* and *ALD* genes to 1.3 Mb in the most distal and very GC-poor isochore but are mostly 0.6–0.8 Mb. These values are in agreement (*i*) with the initial estimate (24) of isochores being ≥0.3 Mb long on the average, (*ii*) with the compositional homogeneity found over almost 1.2 Mb around the cystic fibrosis controlling gene (10), and (*iii*) with the isochore sizes (0.36 to >0.77 Mb) found in the very GC-poor dystrophin gene (9).

The compositional discontinuities (GC transitions) localized on the physical map by the present analysis ranged from 2% to 14% GC. Both transitions delimiting the H3 isochore were very large (10% and 14% GC, respectively). The second of them, which is located between the *G6PD* and *F8C3* genes, had already been described (25) in a study of the base composition of long sequences available in GenBank.

As far as isochore borders are concerned, it should be mentioned that, very recently, the first isochore frontier between an H2 and an L isochore, in the human major histocompatibility locus was sequenced (26). This boundary was found to correspond to a GC-rich cluster of SINEs (*Alu* I sequences) and to a GC-poor cluster of LINEs (L1 sequences) with an associated 650-bp-long element, which shows homology with the PAB1 motif described at the boundary of the pseudoautosomal region (27). Likewise, the Y-specific and

pseudoautosomal portions of the Yp pseudoautosomal boundary differ dramatically in GC and repeat content (28). Interestingly, the PAB1 motif, which is known to be rather conserved in evolution, is present in over 150 copies per haploid human karyotype (T. Ikemura, personal communication).

The High Resolution Cytogenetic Map of Xq28. At this point, it may be of interest to consider the possible correlations between the high resolution compositional and cytogenetic maps of Xq28. Investigations carried out by Yunis and coworkers have depicted Xq28 as a uniform band at a 1250-band resolution (28) or as a complex of three subbands (Xq28.1, Xq28.2, and Xq28.3) with a central small pale G band (29) at a 1000-bands resolution, or a single band at a 2000-band resolution (30). These observations reflect uncertain indications of compositional heterogeneity.

Compositional Heterogeneity of Xq28. The three regions of chromosomal band Xq28 differ not only in base composition but also in their internal compositional heterogeneity. While the proximal region exhibits alternating H1 and L isochores and the distal region exhibits a very homogeneous continuum of L isochores, the middle region shows a number of striking features. (*i*) All seven YACs showing a bimodal CsCl profile are present in this region. Such compositional bimodality can only derive from YAC breakage occurring during the lysis of yeast clones containing compositionally heterogeneous YACs. While YAC instability in this region was already known (11, 31), the correlation between instability and compositional heterogeneity has not previously been reported. (*ii*) The region is also characterized by the highest concentration of gaps in the physical map (three of four), an observation suggesting a possible relationship between DNA unclonability and sequence features associated with compositional heterogeneity. Incidentally, the fourth gap, between contigs A and B, also corresponds to a compositional discontinuity. (*iii*) All the GC-rich ($\geq$50% GC) YACs from the region showed small sizes ($\leq$0.12 Mb). This is true not only for the 6 YACs that were analyzed (1082, 1069, 411, 917, 901, and 780) but also for 15 additional YACs that, although not analyzed here, are mapped to the same genomic region formed by a H3 isochore (11). YACs having a GC level $\leq$50% are occasionally small but have an average size of 0.2–0.3 Mb. This observation might be accounted for by the fact that GC-rich YACs derive from compositionally heterogeneous, *Alu*-rich human genome regions which may preferentially undergo deletions in yeast due to mitotic recombination of commonly occurring *Alu* elements similarly to recombination between homologous chromosomes (32). In this context it should be recalled that compositional heterogeneity increases in isochore families of increasing GC levels (33) and is associated with differences in chromatin structure and accessibility to recombination enzymes (32, 34–36).

Genes, CpG islands, SINEs, and LINEs in Xq28. The three regions of Xq28 also differ in gene and CpG island concentrations. Indeed, few genes have been mapped thus far in the proximal region formed by H1 and L isochores where CpG island density is very low (1/280 kb) (11, 22). CpG islands are much more abundant in the middle region, especially in the H3 isochore located between the *L1CAM* and *G6PD* genes, where a density of about 1/25 kb was observed (11, 14, 15). Finally, the distal region is poor in CpG islands and genes. These results are expected from the correlation between GC levels of isochores and gene concentration (1–3).

The distribution of SINEs and LINEs is also very far from uniform in the three regions. *Alu* sequences are especially scarce in the region between *GABRA* and *DXS1104* (31), whereas LINEs are absent in the region proximal to *DXS497*, as expected (37, 38) from the relatively high GC level of this region. Five GC-rich regions, accounting for 35% of cloned Xq28 (2.6 Mb), have been found to be free or depleted of L1 elements (28), again as expected (37). In particular, L1 sequences are absent in the region between *DXS334* and *DXS49*, in a 0.42-Mb-long region around *DXS305*, in a 0.18-Mb interval proximal to *F8C3*, and in two regions of 0.55 Mb spanning the genes *ALD* and *L1CAM* and the genes *R/GCP* and *G6PD*, respectively (31).

The Compositional Map of Xq28: General Conclusions. The compositional map of Xq28 leads to some general conclusions concerning the correlations between R bands and base composition. *In situ* hybridization of compositional human DNA fractions (22, 39) showed that two sets of R bands should be clearly distinguished: (*i*) T bands, originally described (40) as a subset of R bands which was particularly resistant to heat-denaturation; and (*ii*) R′ bands, namely R bands exclusive of T bands. Indeed, while T bands mainly comprise isochores from the H2 and the H3 family, R′ bands consist, on the average, of GC-poor and GC-rich isochores (mainly of the H1 family) in a ratio close to 50:50. Moreover, GC_3 values of genes contained in R′ bands are closer to GC_3 values of genes from G bands than to the much higher GC_3 values of genes from T bands (41, 42). While the compositional heterogeneity of R′ bands was originally detected in chromosome bands 21q22.1 (7), Xq26, and Xq28 (8), the present work shows that Xq28, an R′ band, contains isochores belonging to all compositional families from L1 to H3.

This conclusion, which contradicts the simple identification of R bands with GC-rich DNA, has two interesting implications. First, the limited presence of H3 isochores is enough to make a band like Xq28 positive (although more weakly positive when compared with classical T bands; ref. 39) by *in situ* hybridization with an H3 isochore fraction but not by T banding (40) by staining with GC-specific fluorochromes (43) or by hybridization with CpG island probes (44). In all likelihood, the small percentage of H3 isochores accounts for this behavior. Second, the weak and medium signals obtained by *in situ* hybridization of H3 isochores on a small number of R′ bands (ref. 39 and unpublished data) also correspond to the limited presence of GC-rich, gene-rich, H3 isochores in those bands. In other words, Xq28 may be characteristic of a subset of R′ bands which differ from the majority of R′ bands in that they contain small amounts of H3 isochores that are absent in the latter. This suggestion may be checked by extending the approach presented here to other R′ bands and to whole chromosomes.

It should be stressed that recent investigatinos (S. Saccone, S. Cacciò, L. Andreozzi, O. Clay, and G.B., unpublished data; see also ref. 3) have shown that a number of R bands have the same features described here for theXq28 band: namely, the presence of very GC-rich isochores. On this basis, these bands can be distinguished from T bands, which are composed predominantly of H2 and H3 isochores, and also from the remaining R bands that do not contain very GC-rich isochores.

We thank Dr. Desgrés (Dijon) for HPLC analysis of cosmids, Dr. G. E. Herman for providing YACs 519A2 and 70H4, and Dr. D. Schlessinger (St. Louis) for discussions. E.-M.G. was supported by a Human Capital and Mobility fellowship of the European Community (ERB4001GT930863). G.P. and M.D. thank Progetto Finalizzato Ingegneria Genetica (Consiglio Nazionale delle Ricerche) and Telethon-Italy (Grant 417).

1. Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M. & Rodier, F. (1985) *Science* **228,** 953–958.
2. Bernardi, G. (1989) *Annu. Rev. Genet.* **23,** 637–661.
3. Bernardi, G. (1995) *Annu. Rev. Genet.* **29,** 445–476.
4. Mouchiroud, D., D'Onofrio, G., Aïssani, B., Macaya, G., Gautier, C. & Bernardi, G. (1991) *Gene* **100,** 181–187.
5. Sabeur, G., Macaya, G., Kadi, F. & Bernardi, G. (1993) *J. Mol. Evol.* **37,** 93–108.
6. Mouchiroud, D. & Bernardi, G. (1993) *J. Mol. Evol.* **37,** 93–108.

7. Gardiner, K., Aïssani, B. & Bernardi, G. (1990) *EMBO J.* **9,** 1853–1858.
8. Pilia, G., Little, R. D., Aïssani, B., Bernardi, G. & Schlessinger, D. (1993) *Genomics* **17,** 456–462.
9. Bettecken, T., Aïssani, B., Müller, C. R. & Bernardi, G. (1992) *Gene* **122,** 329–335.
10. Krane, D. E., Hartl, D. L. & Ochman, H. (1991) *Nucleic Acids Res.* **19,** 5181–5185.
11. Palmieri, G., Romano, G., Ciccodicola, A., Casamassimi, A., Campanile, C., Esposito, T., Cappa, V., Lania, A., Johnson, S., Rinbold, R., Poustka, A., Schlessinger, D. & D'Urso, M. (1994) *Genomics* **24,** 129–158.
12. De Sario, A., Geigl, E. M. & Bernardi, G. (1995) *Nucleic Acids Res.* **23,** 4013–4014.
13. Poustka, A., Dietrich, A., Langenstein, G., Toniolo, D., Warren, S. T. & Lehrach, H. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 8302–8306.
14. Maestrini, E., Tamanini, F., Kioschis, P., Gimbo, E., Marinelli, P., Tribioli, C., D'Urso, M., Palmieri, G., Poustka, A. & Toniolo, D. (1992) *Hum. Mol. Genet.* **1,** 275–280.
15. Bione, S., Tamanini, F., Maestrini, E., Tribioli, C., Poustka, A., Torri, G., Rivella, S. & Toniolo, D. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10977–10981.
16. Sedlacek, Z., Korn, B., Konecki, D. S., Siebenhaar, R., Coy, J. F., Kioschis, P. & Poustka, A. (1993) *Hum. Mol. Genet.* **2,** 1865–1869.
17. McKusick, V. A. (1992) *Mendelian Inheritance in Man* (Johns Hopkins Univ. Press, Baltimore).
18. Rogner, U. C., Kioschis, P., Wilke, K., Gong, W., Pick, E., Dietrich, A., Zechner, U., Hameister, H., Pragliola, A., Herman, G. E., Yates, J. R., Lehrach, H. & Poustka, A. (1995) *Hum. Mol. Genet.* **3,** 2137–2146.
19. Aissani, B., D'Onofrio, G., Mouchiroud, D., Gardiner, K., Gautier, C. & Bernardi, G. (1991) *J. Mol. Evol.* **32,** 497–503.
20. Kohno, K., Wada, M., Schlessinger, D., D'Urso, M., Tanabe, S., Oshiro, T. & Imamoto, F. (1994) *DNA Res.* **1,** 191–199.
21. Zollo, M., Mazzarella, R., Bione, S., Toniolo, D., Schlessinger, D., D'Urso, M. & Chen, E. (1995) *DNA Seq.*, in press.
22. Timms, K. M., Lu, F., Shen, Y., Pierson, C. A., Muzny, D. M., Gu, Y., Nelson, D. L. & Gibbs, R. A. (1995) *Genome Res.* **5,** 71–78.
23. Saccone, S., De Sario, A., Wiegant, J. R., Raap, A., Della Valle, G. & Bernardi, G. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11929–11933.
24. Macaya, G., Thiery, J. P. & Bernardi, G. (1976) *J. Mol. Biol.* **108,** 237–254.
25. Ikemura, T., Wada, K. & Aota, S. (1990) *Genomics* **8,** 207–216.
26. Fukagawa, T., Sugaya, K., Matsumoto, K., Okumura, K., Ando, A., Inoko, H. & Ikemura, T. (1995) *Genomics* **25,** 184–191.
27. Ellis, N. A. & Goodfellow, P. N. (1989) *Trends Genet.* **5,** 406–410.
28. Yunis, J. J. (1976) *Science* **191,** 1268–1270.
29. Yunis, J. J., Sawyer, J. R. & Dunham, K. (1980) *Science* **208,** 1145–1148.
30. Yunis, J. J. (1981) *Hum. Genet.* **56,** 293–298.
31. Whitfield, L. S., Hawkins, T. L., Goodfellow, P. N. & Sulston, T. (1995) *Genomics* **27,** 306–311.
32. Larionov, V., Graves, J., Kouprina, N. & Resnick, M. A. (1994a) *Nucleic Acids Res.* **22,** 4234–4241.
33. Cuny, G., Soriano, P., Macaya, G. & Bernardi, G. (1981) *Eur. J. Biochem.* **115,** 227–233.
34. Tazi, J. & Bird, A. (1990) *Cell* **60,** 909–920.
35. Mendez, M. J., Abderrahim, H., Noguchi, M., David, N. E., Hardy, M. C., Green, L. L., Tsuda, H., Yoast, S., Maynard-Currie, C. E., Garza, D., Gemmill, R., Jakobovits, A. & Klapholz, S. (1995) *Genomics* **26,** 294–307.
36. Larionov, V., Kouprina, N., Nikolaishvili, N. & Resnick, M. A. (1994) *Nucleic Acids Res.* **22,** 4154–4162.
37. Soriano, P., Meunier-Rotival, M. & Bernardi, G. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1816–1820.
38. Zerial, M., Salinas, J., Filipski, J. & Bernardi, G. (1986) *Eur. J. Biochem.* **160,** 479–485.
39. Saccone, S., De Sario, A., Della Valle, G. & Bernardi, G. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4913–4917.
40. Dutrillaux, B. (1973) *Chromosoma* **41,** 395–402.
41. Ikemura, T. & Wada, K. (1991) *Nucleic Acids Res.* **19,** 4333–4339.
42. De Sario, A., Aïssani, B. & Bernardi, G. (1991) *FEBS Lett.* **295,** 22–26.
43. Ambros, P. F. & Sumner, A. T. (1987) *Cytogenet. Cell. Genet.* **44,** 223–228.
44. Craig, T. M. & Bickmore, W. A. (1994) *Nat. Genet.* **7,** 376–382.