Hum Genet (1985) 69:201–205

Human Genetics

An estimate of unique DNA sequence heterozygosity in the human genome

David N. Cooper[1,2], Barbara A. Smith[3], Howard J. Cooke[3], Susanne Niemann[1], and Jörg Schmidtke[1]

[1] Institut für Humangenetik der Universität, Gosslerstrasse 12d, D-3400 Göttingen, Federal Republic of Germany
[2] *Present address:* Université de Lausanne, Institut de Biologie Animale, CH-1015 Lausanne, Switzerland
[3] MRC Mammalian Genome Unit, West Mains Road, Edinburgh, Scotland, UK

Summary. Fifteen different restriction fragment length polymorphisms (RFLPs) were detected in the human genome using 19 cloned DNA segments, derived from flow-sorted metaphase chromosomes or total genomic DNA, as hybridization probes. Since these clones were selected at random with respect to their coding potential, their analysis permitted an unbiassed estimate of single-copy DNA sequence heterozygosity in the human genome. Since our estimate ($h = 0.0037$) is an order of magnitude higher than previous estimates derived from protein data, most of the polymorphic variation present in the genome must occur in non-coding sequences. In addition, it was confirmed that enzymes containing the dinucleotide CpG in their recognition sequence detect more polymorphic variation than those that do not contain CpG.

Introduction

Estimates of heterozygosity in the human genome are useful in several respects. Genetic variability is determined both by effective population size and by the mutation rate, in addition to natural selection and genetic drift. Hence measurements of heterozygosity are central to population genetics theory and practice. Furthermore, genetic variants are a useful tool in genome mapping in that they may be used to establish a linkage map with already existing genomic markers. Finally, such variants can be employed directly both in the antenatal and preclinical diagnosis of human genetic disease and the detection of heterozygous carriers, and in paternity testing. Until recently, this kind of approach was constrained by the necessity to utilize the phenotypic effects of genetic variation and was thus limited for two main reasons. First, the number of such variants was low and second many of them were not amenable to genetic analysis since they were not expressed in easily accessible tissues such as amniotic fluid cells or blood. The advent of recombinant DNA technology has promised to circumvent these problems since it permits direct analysis of the genetic material so that one is no longer dependent upon the possession of a specific tissue sample.

Preliminary results have indicated a much higher frequency of polymorphism at the DNA level as compared with that at the level of the protein. Jeffreys (1979) has estimated that "one in every one hundred" base pairs in the β-globin gene region varies polymorphically. So far, this has been the only truly systematic study reported of variation at the level of genomic DNA and it has been unclear as to whether these results could be extrapolated to the rest of the genome. Not only was the investigation confined to a single locus in the genome but also to a region enriched in coding sequences. Most other studies to date employing recombinant DNA methods (Cooper and Schmidtke 1984) have failed to quantify genetic variation at the DNA level in a systematic way. In a recent investigation, Helentjaris and Gesteland (1983) have used cDNA clones in a search for polymorphisms in human DNA. A recent study by Aldridge et al. (1984) was limited to the X chromosome. In our present study, a series of genomic clones selected at random with respect to their coding potential, was used to provide a first unbiased estimate of heterozygosity at the DNA level in the human genome.

The selection of restriction enzymes for use in a search for restriction fragment length polymorphisms (RFLPs) within defined DNA regions is an important consideration for the detection of such variation. Barker et al. (1984) have claimed that RFLPs are detected at a higher frequency in restriction enzyme recognition sites that contain the dinucleotide CpG. The human genome is heavily methylated at this dinucleotide (Van der Ploeg and Flavell 1980; Cooper 1983) and 5-methylcytosine is subject to frequent replacement by thymidine due to the deamination of the methylated base (Coulondre et al. 1978). This explains both the high frequency of the C→T transition (Vogel 1972; Vogel and Kopun 1977) and the resulting CpG deficiency observed in vertebrate genomes (Salser 1978; Bird 1980).

In practical terms, both MspI (CCGG) and Taq I (TCGA) would be expected to detect polymorphisms with a higher efficiency than those enzymes which lack CpG in their recognition sequences. The data presented here permit a quantitative assessment of the relative efficiency of detection of RFLPs with different restriction enzymes.

Materials and methods

Sources of clones. All clones used in this study except pAM 37 were derived from DNA libraries made from flow-sorted chromosomes (Krumlauf et al. 1982; Southern unpublished data). Chromosomal allocation of all clones was confirmed by

Offprint requests to: J. Schmidtke, Institut für Humangenetik der Universität, Gosslerstrasse 12d, D-3400 Göttingen, Federal Republic of Germany

Southern blotting (Southern 1975) to DNA derived from somatic cell hybrids (Krumlauf et al. 1982; Lund et al. 1983; Southern unpublished data). Chromosome 22 clones: λ22.1, λ22.3, λ22.4, λ22.6 are recombinants made by inserting Eco RI fragments into the vector λgt WES (Krumlauf et al. 1982). Chromosome 7 (pJ 3.11, pJ 5.11, pB 79a, pJU 48, pJU 28, pB 78, and pB 74), chromosome 15 (pJU 201), and X (pH 72, pJU 78, pB 20, pB 22) and pB 97 clones: DNA from flow-sorted chromosomes was digested to completion with Bam HI and recombined with λ 1059 (Karn et al. 1980) and λ L47.1 (Maniatis et al. 1982) vectors. Eco RI/Hind III digests of the complete libraries were subcloned into pAT 153 (pJ 3.11 and pJ 5.11) and pBR 328 (pJU 78, pB 97, pB 79a, pJU 48, pJU 201, pJU 28, pB 78, pB 20, pB 22, and pB 74). pAM 37 is an Eco RI/Hind III fragment cloned in pAT 153 (A. Mitchell unpublished data). pX 2.7 is a λ WES clone isolated from a Y-chromosome library (unpublished data). Clones were pre-selected on the basis of their failure to hybridize with total human DNA in colony hybridizations. Individual clones were then hybridized to Southern blots of total human DNA. Those giving clear, unambiguous hybridization results were chromosomally allocated using a panel of human rodent somatic cell hybrids (Lund et al. 1983).

Sources of human DNA samples. DNA was prepared from white blood cells donated by healthy volunteers and from cord-vessel blood from the placentae of new-borns, both from Northern Germany.

DNA preparation. Twenty ml of blood plus 0.2 ml 0.4 *M* EDTA was spun down for 5 min at 5,000 rpm at 4°C. Ten ml 0.8% Triton X-100 was added to the pellet. The cell lysate was then centrifuged at 15,000 rpm for 15 min at 4°C, and 8 ml 7 *M* urea, 1 ml 10× lysis buffer (3.5 *M* NaCl, 0.1 *M* Tris, 0.01 *M* EDTA, pH 8.0), and 1 ml 10% sodium dodecyl sulphate (SDS) were added to the nuclear pellet. After gentle stirring with a glass rod, 30 ml 1× lysis buffer was added. The mixture was then extracted once with phenol (15 ml), once with chloroform (15 ml), and twice with 15 ml phenol/chloroform (1:1). DNA was precipitated with ethanol, washed in 70% ethanol, and dissolved in 300 µl TE buffer (10 m*M* Tris, 1 m*M* EDTA, pH 7.5). DNA was further treated with RNase A (50 µg/ml) and pronase (50 µg/ml for 1 h each at 37°C). After three further phenol/chloroform extractions, the DNA was redissolved in 500 µl TE buffer.

Restriction enzyme digestion. DNA samples were digested with Bam HI, Eco RI, Hind III, Kpn I, Msp I, and Taq I (Boehringer or Bethesda Research Laboratories) using at least a 10-fold excess of enzyme under the manufacturer's recommended conditions.

Gel electrophoresis and Southern blotting. Agarose (0.8%) (BRL) gels were run in phosphate buffer (0.12 *M* Na_2HPO_4, 0.018 *M* NaH_2PO_4, 5 m*M* EDTA, pH 7.4) at a voltage gradient of 0.5–1 V/cm. Gels were stained with ethidium bromide and photographed under UV light, treated with 0.25 *M* HCl for 10 min, denatured in 0.5 *M* NaOH, 1.5 *M* NaCl for 30 min, and neutralized in 1 *M* Tris-HCl, 3 *M* NaCl, pH 4.5 for a minimum of 1 h. Gels were then blotted onto nitrocellulose (Schleicher and Schüll or Millipore) overnight. After transfer, filters were washed in 2× SSC, and baked under vacuum at 80°C for 2 h.

Hybridization conditions. Filters were prehybridized for 2–16 h in 5× SSC, 5× Denhardt's solution (Dehnhardt 1966) at 65–68°C and hybridized in 5× SSC, 5× Dehnhardt's solution, 10% dextran sulphate (Pharmacia or Serva) with 2–20 ng/ml ^{32}P-labelled probe for 4 h at 65–68°C, in the presence of 0.1 mg/ml sonicated salmon sperm competitor DNA (Sigma). Probes were labelled with [α-^{32}P] dATP (Amersham) to a specific activity of $0.5–2.0 \times 10^8$ dpm/µg (Rigby et al. 1977). Filters were washed in 0.1× SSC, 0.1% SDS and subsequently in 0.1× SSC at 50–55°C for 10–20 min. After drying, filters were exposed to preflashed X-ray film (Dupont Cronex-4) at −70°C for between 16 h and 5 days.

Results

Fifteen different RFLPs were detected using 19 recombinant clones as hybridization probes to population samples of between 10 and 15 individuals per probe. Altogether 36 individuals were found to exhibit polymorphic variants in an analysis

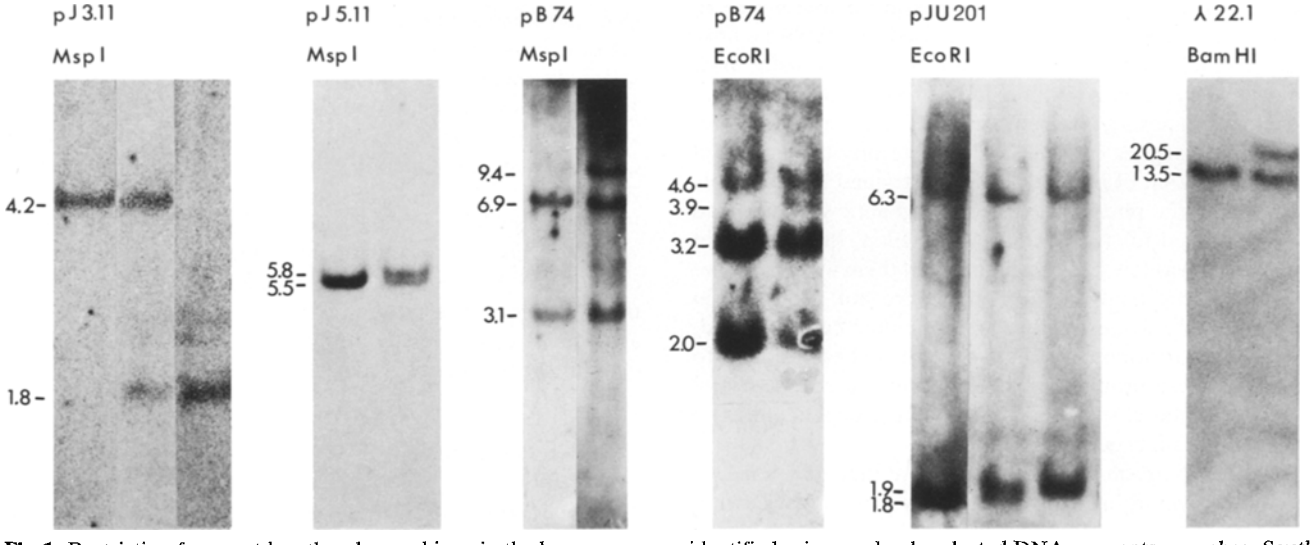

Fig. 1. Restriction fragment length polymorphisms in the human genome identified using randomly selected DNA segments as probes. Southern blots of genomic DNA exhibiting high frequency polymorphisms ($P \geq 0.10$) are illustrated. Heterozygotes are shown together with at least one homozygote for all seven polymorphisms

Table 1. Restriction fragment length polymorphisms detected by cloned single-copy genomic DNA segments

DNA segment	Chromosome	Regional localization	Minimum number of base pairs screened with all enzymes	Restriction fragment length polymorphisms	Variant alleles	Allele frequencies of bands	Allele sizes (kilobase pairs)	Heterozygosity (h)
pJ 3.11	7	–	1636	Msp I	9	0.73/0.27	1.8/ 4.2	0.0122
				Taq I	1	0.97/0.03	6.0/ 3.2	
pJ 5.11	7	–	2116	Msp I	7	0.75/0.25	5.5/ 5.8	0.0066
pB 78	7	–	1160	–	0	–	–	0.0000
pB 74	7+?	–	2496	Msp I	6	0.61/0.39	– / 9.4	0.0088
				Hind III	1	0.98/0.02	– / 7.1	
				Hind III	1	0.98/0.02	– / 3.7	
				Eco RI	2	0.90/0.10	– / 3.9	
				Bam HI	1	0.96/0.04	11.3/13.0	
pB 79a	7	–	1884	–	0	–	–	0.0000
pJU 48	7	–	1472	Taq I	1	0.97/0.03	8.2/15.6	0.0014
pJU 28	7	–	1488	–	0	–	–	0.0000
pJU 201	15	–	1628	Eco RI	11	0.58/0.42	1.9/ 1.8	0.0134
pAM 37	21	–	1456	Msp I	1	0.94/0.06	16.1/ 8.6	0.0014
λ 22.1	22	pter–q11	1680	Bam HI	3	0.90/0.10	13.5/20.5	0.0035
λ 22.3	22	q11–qter	1092	Bam HI	2	0-92/0.08	22.9/20.0	0.0073
				Eco RI	2	0.92/0.08	7.3/ 7.7	
λ 22.4	22	q112–qter	480	–	0	–	–	0.0000
λ 22.6	22	q11–qter	1536	–	0	–	–	0.0000
pH 72	X	q28–qter	1764	–	0	–	–	0.0000
pJU 78	X	q24–q26	1098	–	0	–	–	0.0000
pB 20	X	q24–q26	1260	Msp I	2	0.95/0.05	5.2/–	0.0032
pB 22	X	p11–q12	1168	–	0	–	–	0.0000
pB 97	Autosome	–	1236	–	0	-	-	0.0000
pX 2.7	X	Xq24–qter	360	–	0	–	–	0.0000
			27,010		50			

$$H = \frac{\sum h}{n \text{ loci}} = 0.0030$$

$H_{AUT} = 0.0039$

$H_X = 0.0006$

h over all base pairs examined = 0.0037

which sreened a minimum of 27,010 base pairs of DNA sequence. Southern blots of the high frequency polymorphisms ($q \geqslant 0.10$) are shown in Fig. 1. The complete data are summarized in Table 1 together with the chromosomal allocation of the probes used and the frequencies and fragment sizes of the polymorphic alleles. An allele present in the population at a frequency of 0.15 would in this analysis be detected with a probability greater than 95% (Skolnick and White 1982). Heterozygosities were calculated for the chromosomal regions detected by the cloned probes, using the formula of Nei (1975)

$$h = 1 - \left[\left(\frac{a}{b}\right)^2 + \left(\frac{b-a}{b}\right)^2\right]$$

where a equals the number of polymorphic base pairs found with all enzymes and b equals the total number of base pairs examined. This figure b is a minimum estimate derived from the number of autoradiographic bands, the number of base pairs in the recognition sequence of the enzyme used, and the number of chromosomes scored:

$b = (n \text{ bands} + 1) \times (bp \text{ in restriction site}) \times (n \text{ chromosomes})$

This approach would have overestimated heterozygosity when a probe detected several non-contiguous sequences in the genome, because a greater number of recognition sites would have been screened. Conversely, heterozygosity would have been underestimated either when the fragment length variation lay below the limits of resolution or when small restriction fragments were not detected due to their inefficient binding to the nitrocellulose membrane. It is, however, thought unlikely that the results reported were affected greatly by these possible sources of error. Average heterozygosity (H) is calculated as the arithmetic mean of individual heterozygosities over all the loci examined. For the sequence homologous to the 19 cloned probes used in this study, the average heterozygosity was calculated to be 0.0030. Calculation of the heterozygosity (h) over all base pairs examined gives a slightly higher figure of 0.0037. These figures, however, conceal possible interchromosomal differences in heterozygosity. Average heterozygosities were therefore also calculated separately for the X chromosome and for the autosomes. Although this revealed a lower value for heterozygosity of X-chromosomal DNA (H_X = 0.0006) as compared with autosomal DNA (H_{AUT} =

Table 2. Frequency of detection of restriction fragment length polymorphisms with different restriction enzymes using randomly-selected DNA segments

Enzyme	Number (and proportion) of polymorphisms detected
Msp I	5 (0.33)
Taq I	2 (0.13)
Eco R I	3 (0.20)
Bam H I	2 (0.20)
Hind III	2 (0.13)
Kpn I	0 (0.00)
Total	15 (0.99)

0.0039), this difference was not found to be statistically significant using the Kolmogoroff-Smirnoff test of homogeneity.

Since it is usually assumed that some restriction enzymes detect sequence variation more readily than others, we have also compared the efficiency of polymorphism detection between the different enzymes used. The results of this analysis are presented in Table 2.

Discussion

Previous estimates of heterozygosities in the human genome have been arrived at through the detection and analysis of electrophoretic and immunologic variation of proteins (Nei 1975; Vogel and Motulsky 1982). Nei (1975) has calculated the proportion of heterozygous nucleotides in human protein coding regions to be 4×10^{-4}. However, estimates of variability at the DNA level inferred indirectly from variability at the level of the protein, cannot provide an accurate estimate for heterozygosity in the genome as a whole. It is assumed that only a small proportion of the genome serves a coding function (Lewin 1980). This portion together with putative regulatory sequences is presumed to be less tolerant of mutational events than is the remainder of the genome, which is under a much less stringent selective pressure. Variability detected within coding regions (Jeffreys 1979) is therefore not likely to be representative of the bulk of the genome.

The utilization of a series of DNA probes, selected at random from the genome with respect to coding potential, avoids the problem outlined above. Seventeen of the clones used in this study probably detect 17 distinct loci in the genome. This assertion is based upon the clones themselves being chromosome-specific as well as their being single or low-copy number in the genome. In addition, although multiple (more than two) autoradiographic bands were seen with specific probes with some enzymes, a solitary band was always seen with at least one enzyme. The remaining two clones pB 74 and pX 2.7 are exceptions to this rule. pB 74 is homologous to several sequences in the genome and it is at present unclear as to whether all of these sequences are present on chromosome 7. Similarly, pX 2.7 hybridizes to several homologous sequences in the genome which are located on the X chromosome. As previously reported (Cooper and Schmidtke 1984) clone pB 97, which was tentatively assigned to the X chromosome, exhibited Msp I pattern variants. Subsequent analysis has since shown that this clone is autosomal in origin. Since it is possible that the variation is due to differences in DNA methylation, this Msp I variant has been excluded from the analysis of heterozygosity. Clone pB 20, whose assignment to the X chromosome has been confirmed by others (Wieacker et al. 1984) also detects Msp I variants. The normal pattern, exhibited by both males and females, consists of three fragments of 5.2, 4.9, and 4.6kb respectively. However, two females lack the 5.2kb Msp I fragment. This can be interpreted in terms of a complicated conventional RFLP involving multiple sites of mutation or low repetitiveness of the sequence, but differential DNA methylation cannot yet be ruled out. Family studies which are now in progress to confirm Mendelian inheritance of the polymorphisms detected, should solve the problems of interpretation presented by some of the Msp I variants.

The approach described yielded an estimate for genomic heterozygosity of 0.0037, a value nine times higher than Nei's (1975) estimate (0.0004) derived from structural gene data. The heterozygosity measurements differ slightly from those previously presented (Cooper and Schmidtke 1984) since the sample size is now larger. The data obtained from the analysis of the autosomes ($H = 0.0039$) are in good agreement with the heterozygosity at the β-globin locus (0.0040) calculated from Jeffreys' (1979) results. In this analysis the three detected variants at the β-globin locus were found to be present outside coding regions in intervening sequences. Murray et al. (1983) have recently reported a study of polymorphic variation at the human serum albumin locus. They calculated that some 1 out of 95 nucleotides were affected by an RFLP. This analysis does not, however, provide a direct estimate of heterozygosity since the method employed to calculate the polymorphism frequency ignores the number of individuals screened with each enzyme; this differs some 30-fold between enzymes. Nevertheless, recalculation of their results yielded an estimate of heterozygosity at the albumin locus of $H = 0.0025$, a value not much lower that our estimate and that derived from the data of Jeffreys (1979). That coding DNA sequences reveal less polymorphism than average can also be inferred from the study of Helentjaris and Gesteland (1983). Since measurements of heterozygosity for the β-globin region, the albumin locus, and for the random DNA segments are a factor of between six and ten higher than previous estimates derived from protein data, it can be concluded that most polymorphic variation in the human genome lies outside of coding regions, occurring at a frequency of one in every two or three hundred base pairs. These findings were predicted by the neutral theory of molecular evolution. It had been estimated that the human genome should contain 1.4×10^7 heterozygous nucleotide sites due to neutral mutations (Kimura 1983, p 238); our experimental value is in excellent agreement with this prediction.

As noted above, a lower level of heterozygosity was found on the X chromosome than on the autosomes. While the low level of heterozygosity found on the X chromosome is consistent with the findings of Aldridge et al. (1984), it was not found to be statistically significant and may result from sampling error. If further studies show, however, this difference not to be merely spurious, one could speculate that it might reflect a higher density of structural or regulatory sequences on the X chromosome. This could be interpreted as an extension of "Ohno's law" (Ohno 1967) of the conservation of the X chromosome from identical localization of gene sequences in different species to the conservation of DNA sequences present on this chromosome among individuals of the same species.

Polymorphisms were detected differentially by different restriction enzymes in this study. Our results are consistent with those of previous workers (Barker et al. 1984) who have shown that restriction enzymes such as Msp I and Taq I which contain the methylatable dinucleotide C_PG in their recognition sequences, detect more polymorphic variation than enzymes that do not contain C_PG. As can be inferred from Table 2 the average RFLP detection probability for C_PG recognising enzymes is 0.23, while it is 0.13 for non C_PG recognising enzymes. It should be noted, however, that this effect is markedly less pronounced than indicated by the studies of Barker et al. (1984). These findings should help to place the selection of restriction enzymes for use in studies on DNA polymorphism on a sounder footing.

Limitations of our data include small sample size, both of the number of DNA regions examined and the number of base pairs within each region. Sampling error may for example include linkage disequilibrium with neighbouring genes, selected for or against. A further possible source of error is introduced by the necessary selection of single-copy sequences for use as hybridization probes. Using the method described, it is not possible to measure the heterozygosity associated with sequences flanking highly repetitive elements in the genome. In addition, the presence of these elements to differing extents on different chromosomes could be at least in part responsible for observed interchromosomal differences in heterozygosity. Our estimate for heterozygosity in the human genome therefore needs to be substantiated by further work. Nevertheless our study has placed certain limits upon the level of variation to be expected. In addition, it provides an average figure with which to compare heterozygosities of other DNA sequences in the genome as these data begin to emerge.

Acknowledgements. We are grateful to Bryan Young, Arthur Mitchell, and Juliet Honeycombe for providing cloned probes; to Chris Bostock for cell hybrid DNA; to R. Rauskolb for blood samples; and to Chris Bostock and Ed Southern for advice and discussion. We also wish to thank Bettina Pape for excellent technical assistance. Part of the results were obtained during the medical thesis work of S.N. This work was supported by the Deutsche Forschungsgemeinschaft through grants to J.S.

References

Aldridge J, Kunkel L, Bruns G, Tantravahi U, LaLande M, Brewster T, Moreau E, Wilson M, Bromley W, Roderick T, Latt S (1984) A strategy to reveal high-frequency RFLPs along the human X chromosome. Am J Hum Genet 36:546–564

Barker D, Schäfer M, White R (1984) Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell 36:131–138

Bird AP (1980) DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res 8:1499–1504

Cooper DN (1983) Eukaryotic DNA methylation. Hum Genet 64:315–333

Cooper DN, Schmidtke J (1984) DNA restriction fragment length polymorphisms and heterozygosity in the human genome. Hum Genet 66:1–16

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular basis of base substitution hotspots in *Escherichia coli.* Nature 274:775–778

Denhardt D (1966) A membrane filter technique for the detection of complementary DNA. Biochim Biophys Acta 23:641–646

Helentjaris T, Gesteland R (1983) Evolution of random cDNA clones as probes for human restriction fragment polymorphisms. J Mol Appl Genet 2:237–247

Jeffreys AJ (1979) DNA sequence variants in the $^{G}\gamma$, $^{A}\gamma$, δ- and β-globin genes of man. Cell 18:1–10

Karn J, Brenner S, Barnett L, Cesareni G (1980) Novel bacteriophage cloning vector. Proc Natl Acad Sci USA 77:5172–5176

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge

Krumlauf R, Jeanpierre M, Young BD (1982) Construction and characterisation of genomic libraries from specific human chromosomes. Proc Natl Acad Sci USA 79:2971–2975

Lewin B (1980) Gene expression: vol 2, Eukaryotic chromosomes. John Wiley, New York

Lund E, Bostock C, Robertson M, Christie S, Mitchen JL, Dahlberg JE (1983) U1 Small nuclear RNA genes are located on human chromosome 1 and are expressed in mouse–human hybrid cells. Mol Cell Biol 3:2211–2220

Maniatis T, Fritsch EF, Sambrock J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, New York

Murray JC, Demopulos CM, Lawn RM, Motulsky AG (1983) Molecular genetics of human serum albumin: restriction enzyme fragment length polymorphisms and analbuminemia. Proc Natl Acad Sci USA 80:5951–5955

Nei M (1975) Molecular population genetics and evolution. North Holland Publishing Company, Amsterdam

Ohno S (1967) Sex chromosomes and sex-linked genes. Springer, Berlin

Rigby PWJ, Dieckmann M, Rhodes C, Berg P (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J Mol Biol 113:237–251

Salser W (1978) Globin mRNA sequences: analysis of base pairing and evolutionary implications. Cold Spring Harbor Symp Quant Biol 42:985–1002

Skolnick MH, White R (1982) Strategy for detecting and characterizing restriction fragment length polymorphisms (RFLPs). Cytogenet Cell Genet 32:58–67

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Van der Ploeg LHT, Flavell RA (1980) DNA methylation in the human γδβ-globin locus in erythroid and nonerythroid tissues. Cell 19:947–958

Vogel F (1972) Non-randomness of base replacement in point mutation. J Mol Evol 1:334–367

Vogel F, Kopun M (1977) Higher frequencies of transitions among point mutations. J Mol Evol 9:159–180

Vogel F, Motulsky AG (1982) Human Genetics. Springer, Berlin

Wieacker P, Davies KE, Cooke HJ, Pearson PL, Williamson R, Bhattacharya S, Zimmer J, Ropers HH (1984) Toward a complete linkage map of the human X chromosome: regional assignment of somatic cell hybrids. Am J Hum Genet 36:256–276

Received October 15, 1984 / Revised December 10, 1984