INTRACHROMOSOMAL MAPPING OF THE HUMAN ALPHA-GLOBIN GENE CLUSTER

João Manuel Lopes Borges Lavinha

Thesis submitted for the degree of Master of Science
(Medical Science)

University of Glasgow
Department of Medical Genetics
Duncan Guthrie Institute of Medical Genetics

August 1983

Table of contents

Abbreviations. 9
Acknowledgements. 12
Summary. 13

1. Introduction. 17
1.1. Human gene mapping. 17
1.1.1. Historical perspective of linkage analysis in man. 17
1.1.2. Approaches to gene mapping. 18
1.1.3. Prospects. 24
1.1.4. Current status of the gene map of human chromosome 16. 32
1.2. The object - the human globin gene family. 37
1.2.1. Anatomy of the human globin gene family. 37
1.2.2. Evolution of the globin gene family. 46
1.3. The method- DNA restriction fragment analysis combined with interspecfic somatic cell hybridization. 54
1.3.1. The aim and the strategy. 54
1.3.2. Discussion of the strategy and of some methodologies utilized in this study. 55
1.4. Analysis of the human genome and genetic disease - the haemoglobinopathies. 65

2. Materials and methods. 72
2.1. Production of man x mouse somatic cell hybrids. 72
2.1.1. Parental cells. 72
2.1.2. Cell fusion. 75
2.1.3. Selection of hybrid clones carrying HC16q. 75
2.1.4. Subcloning of primary clones. 75
2.1.5. Mass culture of cloned hybrid cell populations. 75
2.1.6. Counterselection against HC16q. 76
2.1.7. Mycoplasma testing. 77
2.2. Cytogenetic analysis of the hybrids. 78
2.2.1. Chromosome preparations. 78
2.2.2. Staining procedures. 78
2.2.3. Microscopy and micrography. 79
2.3. Biochemical analysis of the hybrids. 80
2.3.1. Cell pelleting and storage. 80
2.3.2. Cell lysis and protein estimation in the lysate. 80
2.3.3. Electrophoretic detection of biochemical markers. 80
2.3.4. The markers. 81
2.4. DNA analysis of the hybrids. 85
2.4.1. Preparation and characterization of high molecular weight genomic DNA. 85
2.4.2. Preparation and characterization of plasmid DNA to be used as hybridization probe. 86
2.4.3. Molecular hybridization on solid phase. 88

3. Results. 93
3.1. Cytogenetic and biochemical characterization of man x mouse somatic cell hybrids. 93
3.2. Detection of human alpha-globin gene cluster in DNA from somatic cell hybrids. 100
3.3. Comments on the methods. 108
3.4. Results versus aims. Conclusions. 111

4. Discussion. 114
4.1. The methods put into perspective. 114
4.2. A side line - intrachromosomal localization of APRT, DIA4 and PGP. 118
4.3. Prospects for further work. 119
4.3.1. Setting up a bank of somatic cell hybrids. 119
4.3.2. Molecular characterization and prenatal diagnosis of haemoglobinopathies. 120

References. 122

Figures

1.1. Evolution of the number of loci and their chromosomal localization. 19
1.2. Genetic mapping resolution by different methods. 23
1.3. Regional assignments of confirmed loci on human chromosome 16. 33
1.4. Globin chain synthesis at different stages of fetal maturation. 38
1.5. Linkage arrangement of human alpha- and non-alpha-globin gene clusters. 41
1.6. Structure of human globin genes. 43
1.7. The aminoacid sequence of the alpha- and beta-globin chains of human Hb A. 44
1.8. The distribution of sequence homologies within a region of the human alpha-globin gene cluster. 45
1.9. Calibration curve for divergence for globin coding sequences. 50
1.10. An evolutionary tree for human globins. 51
1.11. Intrachromosomal localization of the human alpha-globin gene cluster. 57
1.12. Mapping of cloned genes in human x rodent somatic cell hybrids by SOUTHERN blotting. 59
1.13. Purine biosynthesis, reutilization, interconversion and catabolic pathways. 61
1.14. Deletions in the various alpha-thalassaemia haplotypes resulting in phenotypes alpha^{+} and alpha^{o} thalassaemia. 67

Figures (cont)

2.1. Rearranged human chromosomes used in this study. 73
2.2. Map of recombinant plasmid αP7(α1) used as hybridization probe in this study. 87

3.1. Comparison of the heterogeneity of two hybrid cell populations, at similar stages of growth. 97
3.2. Instability of OAA9 5RII hybrid cell population. 98
3.3. Characterization of recombinant plasmid αP7(α1) on an agarose gel. 102
3.4. Typical elution profile of column chromatography of nick-translation reaction mixture. 103
3.5. Pvu II digestion patterns of genomic DNA. 105

Tables

1.1. Methods of human gene mapping. 21
1.2. Prenatal diagnosis by the linkage principle. 28
1.3. Chromosomes and genes involved in the expression and regulation of interferon-mediated antiviral state. 34
1.4. Gene map of human chromosome 16. 35
1.5. Globin chains and haemoglobins produced at different stages of human development. 39
1.6. Analysis of the organization of human globin genes. 42
1.7. Calibration of the evolutionary clock for globin genes. 49
1.8. Chromosomal localization of the human alpha-globin gene cluster. 56
1.9. Enzymatic markers used in this study. 63
1.10. Inherited disorders of haemoglobin. 66

2.1. Composition of the prehybridization and hybridization mixtures. 91
2.2. Standard posthybridization procedure. 92

3.1. Characterization of informative clones from somatic cell hybrids derived from DE human cells. 94
3.2. Idem from LO human cells. 95
3.3. Idem from OA human cells. 95
3.4. Idem from SY human cells. 96
3.5. Genomic DNA preparations. 101
3.6. Screening of DNA from somatic cell

Tables (cont)

hybrids for the human alpha-globin gene cluster. 107

3.7. Data on the screening of DNA from somatic cell hybrids for the alpha-globin gene cluster. 112

Abbreviations

Chromosome nomenclature and notation according to the Paris Conference (1971), Supplement (1975).The symbols for the gene loci conform to those recommended by the Committee on Standardized Human Gene Nomenclature, 1981.

AA	Adenine/Alanosine (selection system)
bp	Base pair (10^3 bp = 1 kb = 0.65 Mdalton); physical map unit
cDNA	Complementary deoxyribonucleic acid
Ci	Curie (1 Ci = 3.7×10^{10} becquerel); unit of radioactivity
cM	Centimorgan (= 1% recombination; on average corresponds to 10^6 bp); genetic map unit.
cpm, cps	Counts per minute, counts per second (1 cps = 1 becquerel); unit of radio-activity
cRNA	Complementary ribonucleic acid
DAP	2,6-Diaminopurine (selection system)
dATP	2'-Deoxyadenosine-5'-triphosphate
dCTP	2'-Deoxycytidine-5'-triphosphate
dGTP	2'-Deoxyguanosine-5'-triphosphate
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dTTP	2'-Deoxythymidine-5'-triphosphate
EC	Enzyme Comission (systematic classification of enzymes issued by the International Union of Biochemistry; IUB Nomenclature Committee, 1978)
FACS	Fluorescence-activated cell (or chromosome) sorter
FAD	Flavine adenine dinucleotide

Abbreviations (cont)

HAT	Hypoxanthine/Aminopterine/Thymidine (selection system)
Hb	Haemoglobin (e.g., Hb A = adult human haemoglobin)
HC	Human chromosome
Hind III	Restriction endonuclease produced by *Haemophilus influenzae* R_d (restriction site: 5'-A↓AGCTT-3') 3'-TTCGAA↑-5'
HPFH	Hereditary persistence of fetal haemoglobin
IMP	Inosine-5'-monophosphate
IS	Intervening sequence (= intron = non-coding sequence)
kb	Kilobase pair
mRNA	Messenger ribonucleic acid
MTT	Methyl thiazolyl tetrazolium
NAD/NADH	Nicotinamide adenine dinucleotide (oxidized/reduced)
NADP/NADPH	Nicotinamide adenine dinucleotide phosphate (oxidized/reduced)
p, pp	Page, pages
PEG	Polyethyleneglycol
PMS	Phenolmethane sulphate
PND	Prenatal diagnosis
poly(A)	Polyadenylic acid (5')
PRPP	Phosphoribosyl pyrophosphate
Pvu II	Restriction endonuclease produced by *Proteus vulgaris* (restriction site: 5'-CAG↓CTG-3') 3'-GTC↑GAC-5'

Abbreviations (cont)

RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
SDS	Sodium dodecylsulphate
SOR	Shortest region of overlap (in gene assignments)
SSC	Standard saline citrate (150 mM sodium chloride, 15 mM trisodium citrate, pH=7)
Tris	Tris(hydroxymethyl)aminomethane

Acknowledgements

The author wishes to express his sincere appreciation to:

Professor Malcolm A. Ferguson-Smith for suggesting the subject of this thesis, for the use of the facilities in the Duncan Guthrie Institute of Medical Genetics and for his supervision throughout the course of this study.

Norma Morrison and Linda Glasgow for their friendly and constant collaboration particularly in the production of somatic cell hybrids and the performance of isozyme analyses.

Allan Caine for the numerous exciting discussions regarding the cytogenetic analysis of the hybrids.

George Lanyon for the alpha-globin probe and for my initiation in the 'magic' of practical molecular genetics.

My colleagues in the postgraduate research laboratory and all the staff of the Duncan Guthrie Institute of Medical Genetics for their kind cooperation.

The British Council for financial support from 1981 to 1983.

The Instituto Nacional de Saúde in Lisbon for their support during my stay abroad.

Summary

The aim of this study was twofold. Firstly, to confirm and, if possible, to narrow the regional localization of the human alpha-globin gene cluster on the short arm of chromosome 16. Secondly, to test the value of an experimental approach combining somatic cell genetics and recombinant DNA procedures for human gene mapping.

Fourteen informative clones from somatic cell hybrids (derived from human fibroblasts fused with mouse L A9 cells) were produced and phenotypically characterized in terms of their human chromosome content by cytogenetic and isozyme analysis.

The four human parental cell types used in this study carried a balanced translocation involving chromosome 16 and chromosomes 2, 5, 9 and 17. These had been cytogenetically interpreted as follows: t(2:16)(2q37:16q11), t(5:16)(5p11:16p11), t(9:16)(9q22:16q22), and t(16:17)(16q12:17p11).

The expression of the selectable marker APRT enabled the clones from somatic cell hybrids to be selected for and against the long arm of HC 16 (AA selection and DAP selection).

Combining the cytogenetic and biochemical data on this panel of hybrid clones made it possible to assign APRT and DIA4 to 16q12→q22 and PGP to 16p11→pter.

The DNAs from the hybrid clones, as well as those from both parental origins, were screened for the presence of

the human alpha-globin gene cluster. Molecular hybridization on solid phase (Southern blotting) with a ^{32}P-labelled human genomic DNA probe (the recombinant plasmid αP7(α1)) was used in the screening.

A positive association (at a level of significance of 5%) between the presence/absence of the short arm of HC 16 and the presence/absence of the human alpha-globin gene cluster was found.

It is therefore concluded that the experimental approach combining somatic cell hybridization and DNA hybridization on solid phase can be successfully applied to the intrachromosomal mapping of cloned single copy genes.

The value of different methods such as molecular hybridization in solution, in situ molecular hybridization and trisomy mapping, in assigning single copy genes to chromosome regions is discussed taking the human alpha-globin gene cluster as the model system.

The setting up of a bank of somatic cell hybrids and the molecular characterization and prenatal diagnosis of haemoglobinopathies are suggested as prospects for further work.

To

Geno
Hilary
Ivone
João

and

Nicas

my
friends.

'I am neere to th'place where they should meet, if Pisanio haue mapp'd it truely.'

W. Shakespeare

'DNA, you know, is Midas' gold. Everybody who touches it goes mad.'

M. Wilkins

1. Introduction.

1.1. Human gene mapping.

1.1.1. Historical perspective of linkage analysis in man

The attempt to understand how genetic information is packaged and distributed throughout the genome is as old as Genetics itself. The following sequence of landmark events indicates (recalling the cartographic metaphor used by MCKUSICK and RUDDLE, 1977) that in the exploration of the 'genetic planet' that is the cell nucleus of man, the broad outlines of the 'continents' (the chromosomes) and some of the gross details of their topography have been known for sometime. However, it is only since the use of somatic cell hybrids in human linkage analysis - namely during the last ten years - that regional cartographic details, the location of specific genes to specific chromosomes or chromosome regions, have been determined.

SUTTON, 1903 - Establishment of a relationship between genetic (the Mendelian factors) and cytologic entities (the chromosomes).
CORRENS, 1905 - First report of (complete) linkage in plants (genus Matthiola).
MORGAN, 1911 - Genes are linked as a result of being carried on the same chromosome and those which are close together will be coupled (i. e., in cis-configuration) more frequently than those more distant from one another.
WILSON, 1911 - First gene assignment to a specific human chromosome (CB to HC X).
STURTEVANT, 1913 - The frequency of recombination of phenotypic traits is an index of relative position of the correspondent genes (construction of gene maps in

Drosophila showing that the genes are linearly arranged along the chromosomes).
BERNSTEIN, 1931 - Earliest quantitative approach to linkage analysis in man (y-statistic; recombination fraction).
PAINTER, 1933 - Assignment of genes to specific areas on chromosomes (regional mapping) in salivary gland nuclei of Drosophila.
BELL and HALDANE, 1937 - Establishment of the first X-linked linkage group in man (CB/HEMA) by family studies.
MOHR, 1951 - Establishment of the first autosomal linkage group in man (Lu/Le blood groups) by family studies.
NABHOLZ et al, 1969 - First human genes assigned using somatic cell hybrids (G6PD and HPRT to HC X).
HENDERSON et al, 1972 - First human gene family assigned using in situ molecular hybridization (rRNA genes to HC groups D and F).
FERGUSON-SMITH et al, 1973 - First human gene assigned by deletion mapping (ACP1 to HC 2p).
In quantitative terms, it is informative to compare the evolution of the identification of gene loci in man, mainly by mendelizing phenotypes (MCKUSICK, 1983), and the rate of chromosomal assignments (Fig 1.1). Although the number of gene assignments increased dramatically during the last ten years (by 200%), the fraction of the identified loci that have been mapped increased much more slowly (50% increase from 12% in 1973 to 18% in 1981). Incidentally, one must keep in mind that the total number of stuctural genes in man has been estimated to be 50,000 - 100,000 (MCKUSICK and RUDDLE, 1977).

1.1.2. Approaches to gene mapping.

In essence, the approaches to human gene mapping can be

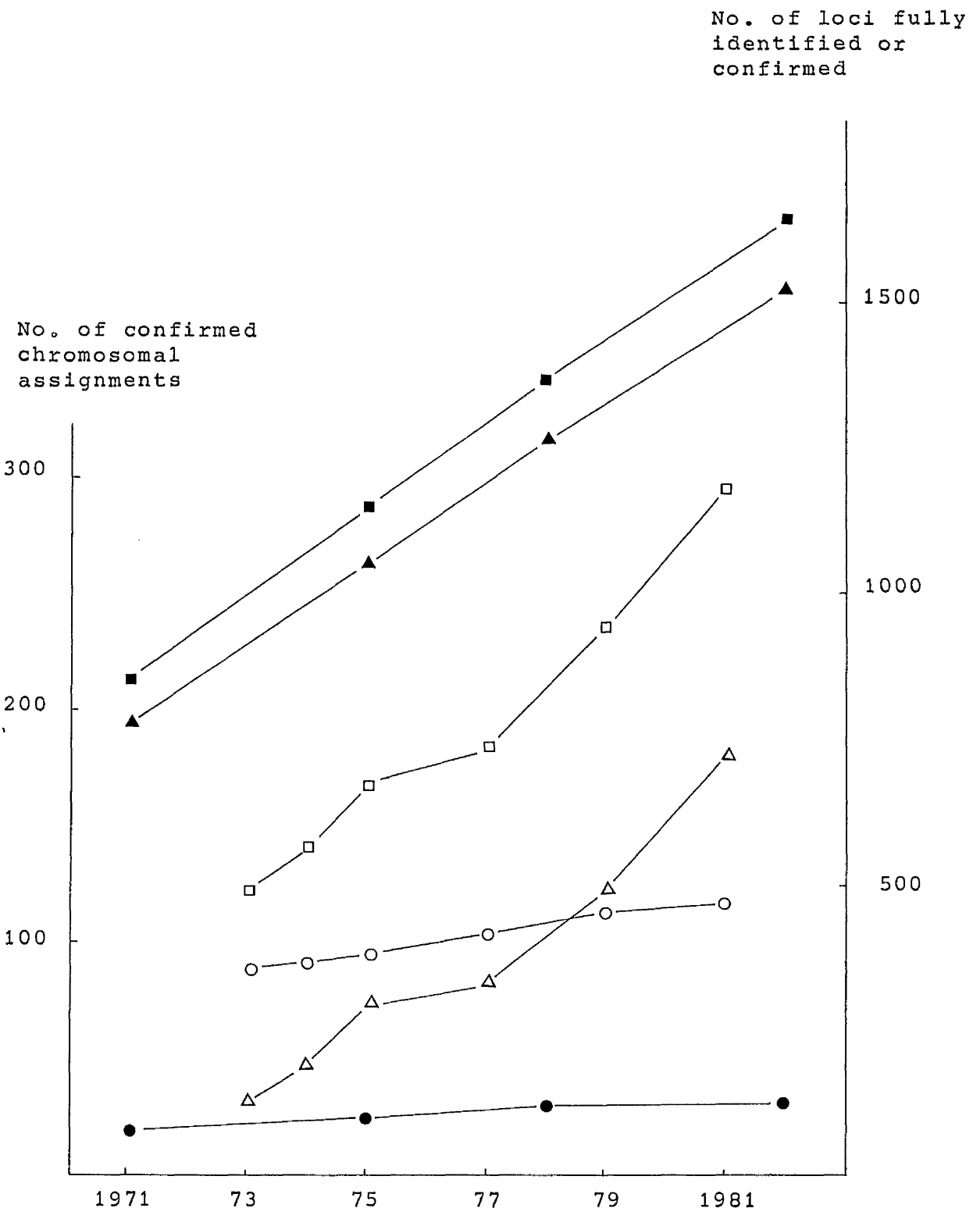


Fig 1.1. Evolution of the number of loci and their chromosomal localization. (▲)Autosomal loci, (●) X-linked loci and (■) = ▲ + ●. (△) Autosomal assignments, (○) X-linked assignments and (□) = △+○ . (MCKUSICK,1982).

categorized into three main groups: (i) those studying the offspring of sexual reproduction (e.g., family studies, deletion mapping); (ii) those depending on parasexual gene transfer and segregation, between donor and recipient somatic cells (somatic cell genetics); and (iii) those dealing with the human genome at the molecular level (either by in situ or in solution or on solid phase DNA/RNA, DNA/DNA or DNA/cDNA hybridization; nucleotide sequencing of nucleic acids). Thus, the observations involved can be directed either to the organismal or cellular phenotypes (as expression of gene activity) or directly to the genotype. Table 1.1 summarizes and exemplifies the use of different approaches to human gene mapping. In a number of cases successful mapping depends on the conjugation of different approaches, e.g., the human alpha-globin gene cluster has been mapped to HC 16 by a combination of somatic cell hybridization and DNA/cDNA molecular hybridization in solution (DEISSEROTH et al, 1977). The contribution of the different types of approach to the current status of the map is as follows for (confirmed, provisional and inconsistent) autosomal assignments (MCKUSICK, 1982):

By family studies	75	19%
By somatic cell hybridization (including M, R, HS, REa)	233	59%
Independently by both of above methods	25	6%
By other methods (including A, REb, D, AAS, LD, V, Ch, OT, EM, H)	59	15%
Total	392	

Genes are assigned to the HC X by studying the

Table 1.1. Methods of human gene mapping.

Method (symbol)	Observation	Example* (HC)	Ref
Study of linkage between traits in families (F)	Recombination fraction in families	ABO/NPS 1	(1)
--- one trait being a chromosome heteromorphism or rearrangemant (Fc)	Recombination fraction in families	Fy/1qh	(2)
Cosegregation of cellular traits and chromosomes (or their fragments) in clones from somatic cell hybrids (S)	Chromosomes and gene products (enzymes, antigenes, anonymous peptides)	TK 1 (17)	(3)
Microcell-mediated gene transfer (M)	Chromosomes and gene products	COLM (17)	(4)
Chromosome-mediated gene transfer (can be used in conjunction with FACS) (C)	Chromosomes and gene products	GALK/TK 1 (17)	(5)
Cotransference (radiation induced gene segregation) (R)	Gene products	Order of X-linked genes	(6)
In situ hybridization (DNA/RNA annealing) (A)	Chromosomes	RNR (D and F groups) IGK (2)	(7) (20)
Molecular hybridization in solution (Cot analysis) (HS)	DNA	HBA (16)	(8)
Restriction endonuclease techniques (RE)	DNA (including detection of RFLPs)	NAG fine structure	(9, 10)
--- combined with somatic cell hybridization (REa)	Chromosomes and DNA (including detection of anonymous nucleotide sequences)	NAG (11)	(11)
--- combined with FACS (REb)	Chromosomes and DNA	INS (11)	(12)
Deletion mapping, trisomy mapping, gene dosage effect (D)	Chromosomes and gene products	ACP 1 (2)	(13)
Aminoacid sequence of proteins (AAS)	Aminoacid sequences	HBD/HBB	(14)
Linkage disequilibrium (LD)	Polymorphic traits (population dynamics of)	HBD/HBB	(15)
Induction of chromosome change by adenovirus (activation of kinases) (V)	Chromosomes	A12M4 (17)	(16)
Chromosome change associated with phenotype (Ch)	Chromosomes (in cultured cells) and phenotypic traits	FSXq27	(17)
Centromere mapping (ovarian teratomas) (OT)	Gene products	PGM 3 (6)	(18)
Exclusion mapping (Lod score analysis) (EM)	Recombination fraction	MNSs (4)	(19)
Interspecies homology (H)	Linkage groups in other species	LDHC (12)	

(*) Loci symbols as recommended by the Committee on Standardized Human Gene Nomenclature, 1981.

References: (1) RENWICK and LAWLER, 1955. (2) DONAHUE et al, 1968. (3) MILLER et al, 1971. (4) SUNDAR RAJ et al, 1977. (5) KLOBUTCHER and RUDDLE, 1979. (6) GOSS and HARRIS, 1977. (7) HENDERSON et al, 1972. (8) DEISSEROTH et al, 1977. (9) FRITSCH et al, 1980. (10) BOTSTEIN et al, 1980. (11) JEFFREYS et al, 1979. (12) LEBO et al, 1981. (13) FERGUSON-SMITH et al, 1973. (14) BAGLIONI, 1962. (15) BOYER et al, 1963. (16) MCDOUGALL et al, 1973. (17) TURNER et al, 1980. (18) OTT et al, 1976. (19) BOOTSMA and MCALPINE, 1979. (20) MALCOLM et al, 1981.

characteristic pattern of inheritance of X-linked traits (mainly by family studies). Only a small number of genes have been regionally localized on HC X using somatic cell hybrids carrying part of the chromosome translocated to an autosome (HAMERTON et al, 1975). The method of radiation induced gene segregation was used in relation to HC X, to determine the relative order of a set of X-linked genes (GOSS and HARRIS, 1977).

The different mapping techniques spread over a wide range of resolving power (Fig 1.2). Conventional low-resolution somatic cell genetics procedures can be expected to map genes to chromosomes, and then within chromosomes, to a resolution of 5-10cM (5-10 Mbp)*. At the other end of the scale, high-resolution recombinant DNA procedures, together with restriction site and nucleotide sequence analysis, can be expected to provide mapping data from the level of the single base pair to the 0.1 Mbp (0.1 cM). Thus, in mammalian linkage analysis a resolution gap in the gene mapping techniques exists over the range 0.1-5 cM. This is a significant gap, because many functionally related clusters of genes will fit into a space of this dimension. Gene transfer systems provide possible approaches to intermediate level resolution analysis. These methods depend on fragmenting linkage groups and require a selectable marker in the linkage group of interest. The frequency of recovery of genes in linkage with the selectable marker provides information on the distance and order of genes with respect to the selectable marker. Donor chromosome breakage by irradiation and subsequent

---

(*) Total 'genetic' length of human genome: 3,000 cM; total 'physical' length of human genome: 3,000 Mbp; therefore, on average, 1 cM corresponds to 1 Mbp.

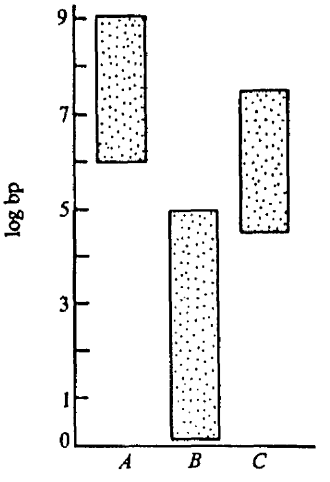


Fig 1.2. Genetic mapping resolution by different methods. A, Gene mapping by somatic cell genetics. B, Molecular genetic techniques. C, gene transfer methods. Ordinate = logarythm of base pair number (from RUDDLE, 1981).

transfer to recipient cells by hybrid formation (GOSS and HARRIS, 1977) and transfer of isolated chromosomes (KLOBUTCHER and RUDDLE, 1979) are two gene transfer systems suitable for intermediate level resolution analysis. These techniques can be subjected to vigorous evaluation only when numerous unique nucleic acid probes have been isolated which map into appropriately short segments of mammalian genomes amenable to analysis (RUDDLE, 1981).

1.1.3. Prospects.

1.1.3.1. Applications of human gene mapping - the value of mapping information.

a) Understanding of biological evolution by interspecies comparative mapping: In Primates, there is a good agreement between chromosome homology, as inferred from chromosome morphology and banding, and data on comparative gene mapping. This approach has even succeeded in the demonstration of interspecific chromosome homologies in situations in which the implications of chromosome morphology and banding were not clearly evident. This was the case of HC 9 that had no recognizable homologue in the gorilla and the orangutan. Two gene loci (ACO1 and AK3) have been reported which segregate with HC 9 and with its corresponding homologue in the chimpanzee (PTR 11) as well as an acrocentric chromosome in the gorilla (GGO 13), a chromosome that was not recognized as being homologous to any chromosome in man or other ape species (SEUÁNEZ, 1979, pp 111-127). Combining the information on chromosome morphology and banding, syntenic groups and protein and nucleic acid homologies, made it possible to understand the phylogeny of Primates and to refine the evolutionary tree exclusively based on the

fossil record.
No evolutionary change has occurred in the mammalian X chromosome in nearly 85 Myear (since the mammalian radiation). This 'frozen accident' also involved the development of a means of dosage compensation in females ('lyonization'). It is now generally accepted that the 'inactive' X chromosome of the human female is not wholly inactivated, so that some loci escape inactivation in a non-random way. It has been suggested, as a possible cause, the fact that these loci are within the pairing segment between Xp and Yp, hence present in double dose in both male and female. This peculiarity of gene expression might be a remnant of the (common) evolutionary history of the sex chromosomes (POLANI, 1982).

b) Understanding of chromosomal organization in relation to genetic control mechanisms and cell differentiation: Point mutations in regulatory genes and/or changes in gene order due to chromosomal rearrangements can account for important organismal differences between closely related species such as man and chimpanzee. Ontogenic significance of gene clusters: genes that are sequentially activated during development are linked while genes coding for enzymes in successive steps of metabolic pathways are not syntenic probably because many enzymes function in more than one pathway making coordinated function (like the 'operon' model of prokaryotes) undesirable (MCKUSICK, 1980).

c) Understanding of pathogenesis of human genetic disease: A genetic understanding of disease is gained by knowing where the affected gene is located and its relationship to other and adjacent genes. If a structural gene's function, expression, or product is dependent on additional genes, such as processing genes,

developmentally regulated genes, or architectural genes, then a mutation in any of these steps in expression could possibly result in a phenotype similar to one involving a structural gene defect. Mutations of one gene that affect the expression of another gene identify loci that function in the final realization of a structural gene product. Thus, if mutant genes involved in a sequence of physiological steps can be chromosomally located, it can be investigated whether diseases affecting the same tissue, pathway or organelle or that result in similar pathology are the result of mutations to genes that are closely linked, clustered on the same chromosome, or under coordinate control (SHOWS et al, 1982). Close linkage between immunoglobulin genes and the breakpoints of abnormal chromosomes associated with myeloproliferative diseases has been established (BERNHEIM et al, 1980; LENOIR et al, 1982). In particular, in chronic myelocytic leukaemia, a cellular oncogene (c-abl), normally localized on HC 9, is translocated to the Philadelphia chromosome, the breakpoint being indistinguishable from the loci of immunoglobulin lambda light chain gene cluster (DE KLEIN et al, 1982). The gene content of chromosomes may suggest mechanisms for the onset of phenotypic changes in trisomy and deletion syndromes. In relation to the pathogenesis of diabetes a deleted insulin gene has been found in a case of stucturally abnormal human insulin (KWOK et al, 1981).

d) Prenatal and premorbid diagnosis, and carrier detection of Mendelian disorders not biochemically identifiable: If loci closely linked to the disease locus are identified, and if allelic variation at those loci can be scored in amniotic cells or fluid, there is a chance of prenatal diagnosis of the disease by the linkage principle (EDWARDS, 1956). So far, this

principle has been applied successfully in a number of different conditions (Table 1.2). The applicability of the linkage principle to prenatal diagnosis depends on (i) the tightness of linkage (less than 1 cM), (ii) finding appropriate genotypes in the parents, and (iii) the knowledge of linkage phase (cis- or trans-configuration) of the linkage group in the parents.

The availability of many DNA restriction fragment length polymorphisms, RFLP's, (BOTSTEIN et al, 1980 suggest the need for linked markers about 20 cM apart which makes necessary 150 markers/haploid genome; however, taking into account that the markers will fall randomly along the genome, LANGE and BOEHNKE, 1982 showed that a much larger number of such loci will have to be isolated and tested before the goal of a saturated gene map is reached) will increase the efficiency of prenatal diagnosis, provided the probability of recombination between the restriction site and the disease locus tends to zero. In the limit, the RFLP may fall within the disease locus or even coincide with the mutated codon, as it is the case of Dde I or Mst II polymorphisms in sickle-cell anaemia (CHANG and KAN, 1982). A genetic distance greater than 1 cM is unlikely to be acceptable for prenatal diagnosis, so that in many instances cloning of the disease gene may be a necessary step in the development of a prenatal diagnosis test based on the detection of RFLP's. Alternatively, the predictive value of the test can be significantly improved by mapping RFLP's to positions flanking the disease gene hence reducing the frequency of false negative results due to double crossing-over (10^{-4} for flanking markers located 1 cM far from the disease gene) (RUDDLE, 1981). Thus, many disorders can be detected even though the precise gene-determined biochemical defect is not yet identified or cannot be recognized in the cellular phenotype (MCKUSICK, 1980).

Table 1.2. Prenatal diagnosis by the linkage principle.

Linkage group (marker gene/disease gene)	HC	Ref
G6PD/HEMA	X	(1)
Se/DM	19	(2)
Hpa I RFLP/Hb S	11	(3)
HLA/CAH (21-hydroxylase deficiency)	6	(4)
Bam HI, Hind III RFLP's/Beta^{o} thalassaemia	11	(5)

RFLP = DNA Restriction Fragment Length Polymorphism.

References: (1) MCCURDY, 1971; (2) INSLEY et al, 1976; (3) KAN and DOZY, 1978 a, b; (4) LEVINE et al, 1978; (5) LITTLE et al, 1980.

1.1.3.2. New developments in methodology (DAVIES, 1981; RUDDLE, 1981; JUNIEN, 1982; PUCK and KAO, 1982).

a) Eliciting the expression of differentiated functions in cultured non-differentiated cells: Interspecific somatic cell hybrids are an useful model to study the regulation of the expression of genes coding for differentiated gene products. So far, at least two types of rodent recipient cells have been used in performing cell fusion experiments: mouse erytholeukaemia cells (DEISSEROTH and HENDRICK, 1978) and mouse hepatoma cells (DARLINGTON et al, 1982). In the first of these systems, the expression of human alpha-globin genes was assessed by molecular hybridization assaying of RNA, for human alpha-globin sequences, and by measuring the level of synthesis of human alpha-globin in the hybrids.
The mouse hepatoma cell hybridization system has been used to produce hybrids with several human cell types (lymphocytes, fibroblasts, amniotic fluid cells and hepatocytes). DARLINGTON et al suggest that the frequency of hybrid clones expressing human hepatic gene products, and the array of proteins produced, are influenced by the histogenic state of the human parental cell type. A particularly interesting aspect of this work concerns the expression of fetal genes in somatic cell hybrids: thus, the failure to observe alpha-fetoprotein in the hybrids which expressed human albumin might suggest either that this fetal gene is not syntenic to the albumin gene or that it is not accessible to activation in somatic cell hybrids.

b) Gene transfer procedures: Microcell-mediated gene transfer, chromosome-mediated gene transfer, DNA-mediated gene transfer, fusion of human haploid

spermatids with non-human somatic cells. Systems suitable for intermediate level resolution analysis of the human genome (10^4-10^8 bp).

c) Fine structure mapping by DNA analysis and genetic manipulation: Construction, selection, amplification and purification of nucleic acid probes for use in mapping experiments. Restriction endonuclease analysis by blot finger printing of DNA (SOUTHERN, 1975). Nucleic acid sequencing. Construction of cDNA, genomic DNA, and chromosome-specific DNA banks and isolation of cloned unique DNA sequences (KAO et al, 1982).

d) Co-operative interaction between Mendelian Genetics, somatic cell Genetics and DNA studies: At the present time, family studies can be informative in mapping traits for which no cellular pnenotype is known and in the measurement of the genetic distance separating loci shown to be syntenic by parasexual methods. However, the chromosome assignment of a mutant or polymorphic trait segregating in families can be facilitated by a linkage association with a second gene which had been assigned previously by somatic cell Genetics. Family studies are likely to become more useful and easier to perform as more and more RFLP's are localized over the chromosome map: simplification of the analysis of family data, less dependency on statistical methods, increased efficiency (small bodies of family data, even single families, will suffice for mapping genes associated with complex organismal traits).

e) Correlation between genetic and physical gene maps: Progress in the identification of chromosome banding patterns through high resolution techniques (YUNIS, 1981) together with somatic cell hybridization have resulted in the assignment of many genes to their

physical locations on the chromosomes (see 1.1.2). At the same time, family studies giving estimates of the recombination fraction between pairs of loci and data on the frequency distribution of male chiasmata (LAURIE et al, 1981) have been accumulating. A number of attempts has been made to synthesize these different types of evidence. KEATS et al (1981) suggest a method for obtaining a linear genetic map using both recombination data and physical assignments of the loci, assuming that genetic and physical distances are proportional over small intervals, but generally not over the whole chomosome. Once such a 'synthetic map' is produced, the precision of the statements defining chromosome rearrangements will be greatly increased.

1.1.4. Current status of the gene map of human chromosome 16 (Table 1.4).

Human chromosome 16 corresponds to 3.11% of human haploid genome length (FERGUSON-SMITH, 1974) and to 1.55% of human autosomal DNA content (MENDELSOHN, 1973). Eight loci have been consistently assigned to it (Fig 1.3). Fourteen provisional assignments are waiting for confirmation. One assignment is still tentative.

The gene content of HC 16 displays a great heterogeneity: apart from the alpha-globin gene cluster, two transaminases (GOT2, GPT1), and two genes coding for enzymes of purine (APRT) and pyrimidine (TK2) salvage pathways, no other functionally related linkage group has been detected so far. This is in accordance with the above discussion of the biological significance of linkage and synteny: gene clustering appears to be the structural expression of sequencial and mutually exclusive gene activation during ontogeny rather then the coordinate activation of genes coding for enzymes catalysing different steps of a metabolic pathway (see 1.1.3.1.b)).

A certain degree of homology of gene content (e.g., TK2 maps on HC 16 while TK1 maps on HC 17) and general morphology between HC's 16 and 17 led certain authors to suggest they had a common origin and resulted from tetraploidization (COMINGS, 1972). HC 16 also carries two (?) of a multichromosome set of genes necessary for expression and regulation of interferon-mediated antiviral state (Table 1.3).

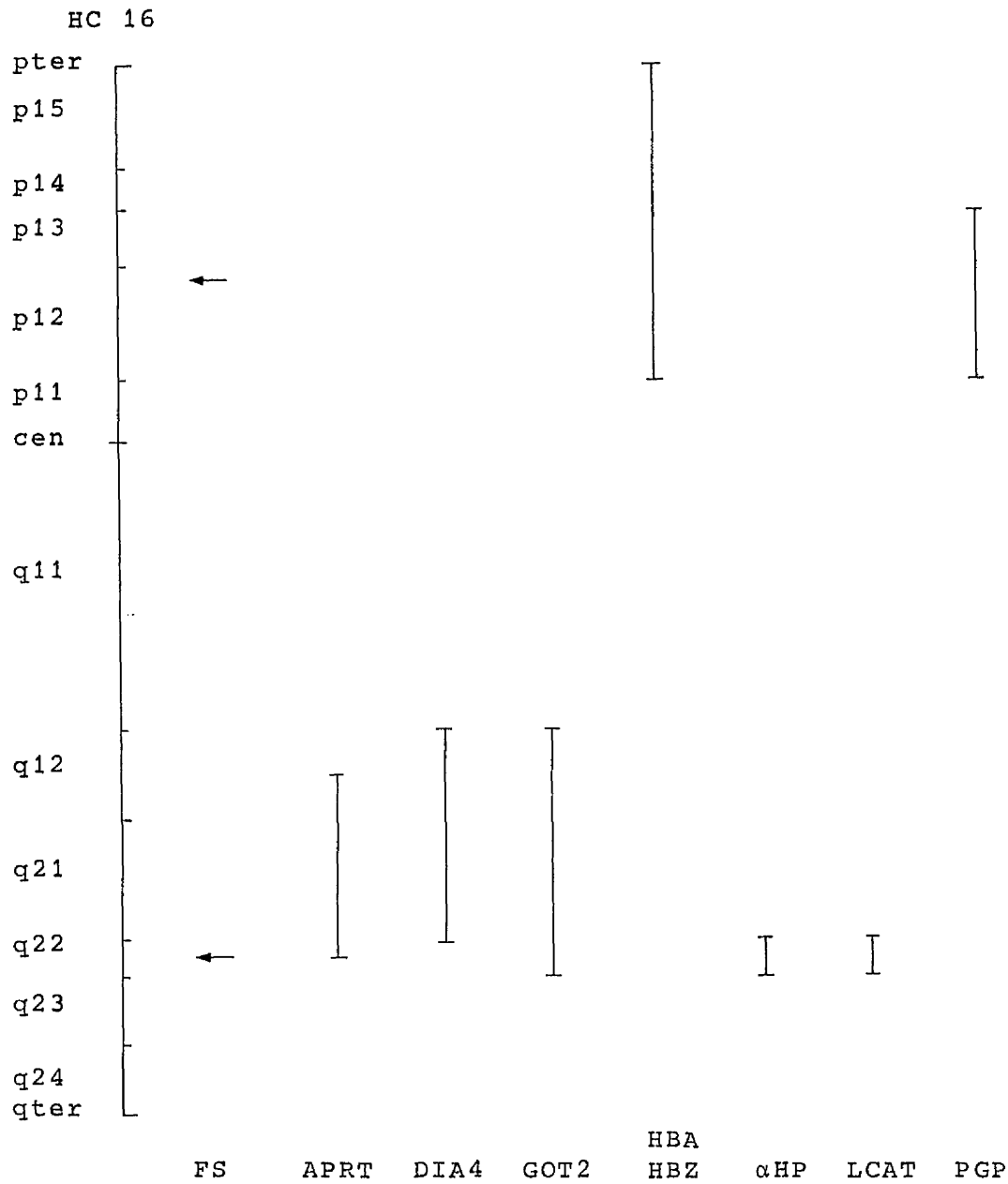


Fig 1.3. Regional assignments (shortest region of overlap) of confirmed loci on HC 16.

Table 1.3. Chromosomes and genes involved in the expression and regulation of interferon-mediated antiviral state.

HC	Gene (symbol)
2	Interferon 1 (IF1)
5	Interferon 2 (IF2) Antiviral state repressor regulator (AVRR)
9	Interferon, fibroblast (IFF) ---- , leukocyte (IFL)
12	Interferon, immune (IFI)
16	Antiviral state regulator (AVR) Defective interfering particle induction, control of (DIPI)
21	Interferon receptor (antiviral protein) (IFRC)

Table 1.4. Gene map of human chromosome 16 (MCKUSICK, 1982).

Marker (certainty*)	Biological significance	Method	SOR	References
APRT (C) adenine phosphodibosyl transferase, EC 2.4.2.7	Purine salvage pathway (2,8-dihydroxy adenine urolithiasis)	S,D	q12-q22	(1,2,3,4)
AVR (P) antiviral state regulator	Regulation of antiviral response	S,D		(5)
BCS (L) breast cancer susceptibility	Genetic predisposition for cancer	F		(6)
CTH (P) cystathionase, EC 4.4.1.1	Cystheine synthesis (cystathioninuria)	S		(7)
CTRB (P) chymotrypsinigen B, EC 3.4.21.1	Proteolysis	S,REa		(8)
D16S1, D16S2 (P) anonymous DNA fragments	Unknown	RE		(9)
DIA4 (C) diaphorase 4, EC 1.6.2.2	Electron transport	S	q12-q21	(10,11)
DIPI (P) defective interfering particle induction, control of	Regulation of antiviral response	S		(27)
ESB1 (P) epidermolysis bullosa, Ogna type	Skin malignancy	F		(12)
ESB3 (P) esterase B, EC 3.1.1.1	Esterolysis	S		(13)
FS16p124, FS16q22 (C) fragile sites	Unknown	Ch		(14,15)
GCF2 (P) growth rate contolling factor	Regulation of growth rate	S		(16)
GOT2 (C) glutamate oxaloacetate transaminase, mitochondrial, EC 2.6.1.1	Aminoacid interconversion	S	q12-q22	(17,18,19)
GPT1 (P) glutamate pyruvate transaminase, soluble, red cell, EC 2.6.1.2	Aminoacid interconversion	S		(20)
HAGH (P) hydroxyglutathione hydrolase, EC 3.1.2.6	Detoxyfication of methylglyoxal? Regulation of cell division? Microtubule assembly?	S		(21)
HBA1, HBA2, HBZ1, HBZ2 (C) α-globin gene cluster	Oxygen transport (α-thalassaemias, HbH related mental retardation)	S,REa,A,D	p12-pter	(22,23,24,25)
HP (C) α-haptoglobin	Haemoglobin elimination	F	q22	(26)
LCAT (C) lecithin-cholesterol acyl-transferase, EC 2.3.1.43	Phospholipid metabolism (NORUM's disease: LCAT deficiency)	F, LD	q22	(28)
LIPB (P) lysosomal acid lipase B, EC 3.1.1.3	Lysosomal system	S		(29)
NHCP2 (P) non-histone chromosome protein	Regulation of gene expression?	S		(30)
PGP (C) phosphoglycolate phosphatase, EC 3.1.3.18	Regulation of oxygen transport?	S	p12-p13	(24,31,32,33, 34)
TK2 (P) thymidine kinase, mitochondrial, EC 2.7.1.21	Pyrimidine salvage pathway	S		(35)
VMD (P) macular distrophy, atypical vitelliform		F	16p? (5 cM from GPT1)	(36)

(*) Degrees of certainty of assignment: C = Confirmed (at least two laboratories or two families). P = Provisional (one laboratory or family). L = in 'Limbo' (less than provisional).

References: (1) TISCHFIELD et al, 1973. (2) KAHAN et al, 1975. (3)BARG et al, 1981. (4)RETHORE et al, 1982. (5) CHANY et al, 1975. (6) KING et al, 1980. (7) DONALD et al, 1981a. (8) SAKAGUCHI et al, 1981. (9) PEARSON et al, 1981. (10) GRZESCHIK, 1979. (11) POVEY et al, 1980. (12) OLAISEN and GEDDE-DAHL, 1973. (13) ASTRIN et al, 1981. (14) MAGENIS and CHAMBERLIN, 1979. (15) SCMID et al ,1980. (16) DONALD et al, 1981b. (17) TOLLEY et al, 1980. (18) POVEY et al, 1981. (19) JEREMIAH et al, 1982. (20) WIJNEN and MEERE-KHAN, 1981. (21)HONEY and SHOWS, 1981. (22) DEISSEROTH et al, 1977. (23) BARG et al, 1981. (24) KOEFFLER et al, 1981. (25) BARTON et al, 1982. (26) ROBSON et al, 1969. (27) CREAGAN et al, 1975. (28) TEISSBERG and GJONE, 1974. (29) NGUYEN VAN CONG et al, 1980. (30) BODE et al, 1980. (31) BLANKENSTEIN-WIJNE et al, 1979. (32) DONALD et al, 1979. (33) POVEY et al, 1979. (34) SPARKES et al, 1979. (35) WILLECKE et al, 1977. (36) FERRELL et al, 1983.

1.2. The object - the human globin gene family.

1.2.1. Anatomy of the human globin gene family.

1.2.1.1. The ontogeny of globin gene expression.

Adult human haemoglobin (Hb A), the major oxygen carrier in human blood, is an iron (II) metalloprotein with a molecular weight of 64,450. It is a tetramer in which the four polypeptide subunits (two alpha- and two beta-globin chains with a total of 574 aminoacid residues) are combined with four haem molecules.
In addition to Hb A, other globin chains are synthesized, and other haemoglobin tetramers assembled, at different stages of normal development (Fig 1.4 and Table 1.5). The primary difference between the gene clusters coding for alpha- and non-alpha-globin chains is that two 'switches' in gene expression are observed for non-alpha-globin genes (epsilon ⟶ gamma ⟶ beta) while a single 'switch' results in activation of adult alpha-globin production early in fetal life (MANIATIS et al, 1980). The observed sequencial gene activation during development implies globin genes are organized in a developmentally regulated way (see 1.2.1.2). Regulation of globin chain synthesis seems to be operated at three levels (WEATHERALL and CLEGG, 1981): (i) Principal control: mRNA synthesis (transcription and RNA processing). (ii) Minor control: chain initiation (translation). (iii) Final overall balance: proteolysis of the chain in excess (post-translation). The latter is particularly important when the synthesis of one of the globins is reduced as in thalassaemias (TESTA et al, 1981;BALLAS and BURKA, 1982). The mechanism of 'switching' from fetal to adult haemoglobin synthesis is unknown. It seems to be regulated more by gestational age than by extrauterine environment (BUNN et al, 1977;

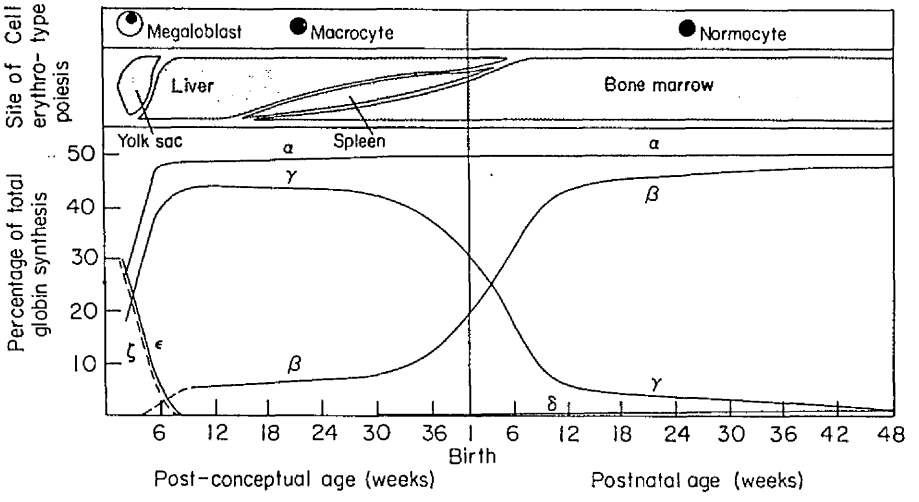


Fig 1.4. Globin chain synthesis at different stages of fetal maturation (from WEATHERALL and CLEGG, 1981, p 64).

Table 1.5. Globin chains and haemoglobins produced at different stages of human development.

Stage	α-like chain	Non-α-like chain	Haemoglobins (subunits)
Embryo	ζ, α	ε, γ	Gower 1 ($\zeta_2\varepsilon_2$) Portland ($\zeta_2\gamma_2$) Gower 2 ($\alpha_2\varepsilon_2$)
Fetus	α	γ	F ($\alpha_2\gamma_2$)
Adult	α	β, δ	A_1 ($\alpha_2\beta_2$) A_2 ($\alpha_2\delta_2$)

NIENHUIS and STAMATOYANNOPOLOUS, 1978).

1.2.1.2. Chromosomal arrangement of human globin genes.

The globin gene family is distributed in two linkage groups, which have been assigned to HC 16 (alpha-globin gene cluster) and to HC 11 (non-alpha-globin gene cluster). The current picture of these two gene clusters is shown in Fig 1.5 and the sequence of the major contributions leading to this picture is summarized in Table 1.6. It should be emphasized at this stage that the chromosome arrangement of the two globin gene clusters parallels the order in which the individual genes are expressed during ontogenesis (see 1.2.1.1). Alpha- and non-alpha-globin genes display a marked similarity of internal organization (Fig 1.6). Although the function of introns is not known, it has been proposed that they play a role in evolution by joining in different DNA sequences encoding protein structural (and functional) domains. In fact, the distribution of 'functional' aminoacid residues (e.g., two histidine residues that interact with the haem, Fig 1.7) in alpha- and beta-globin polypeptides can be correlated with the arrangement of coding and non-coding sequences in these genes (CRAIK et al, 1980; MANIATIS et al, 1980).

Heteroduplex analysis of the alpha 1 - alpha 2 region of the alpha-globin gene cluster (LAUER et al, 1980) demonstrated an extensive homology between alpha 1 and alpha 2 regions (Fig 1.8). The observed maintenance of homology between a pair of structurally and functionally related genes will be discussed later in more detail (see 1.2.2).

Related DNA sequences have been found in both globin gene clusters (Figs 1.5 and 1.8) which seem to belong to

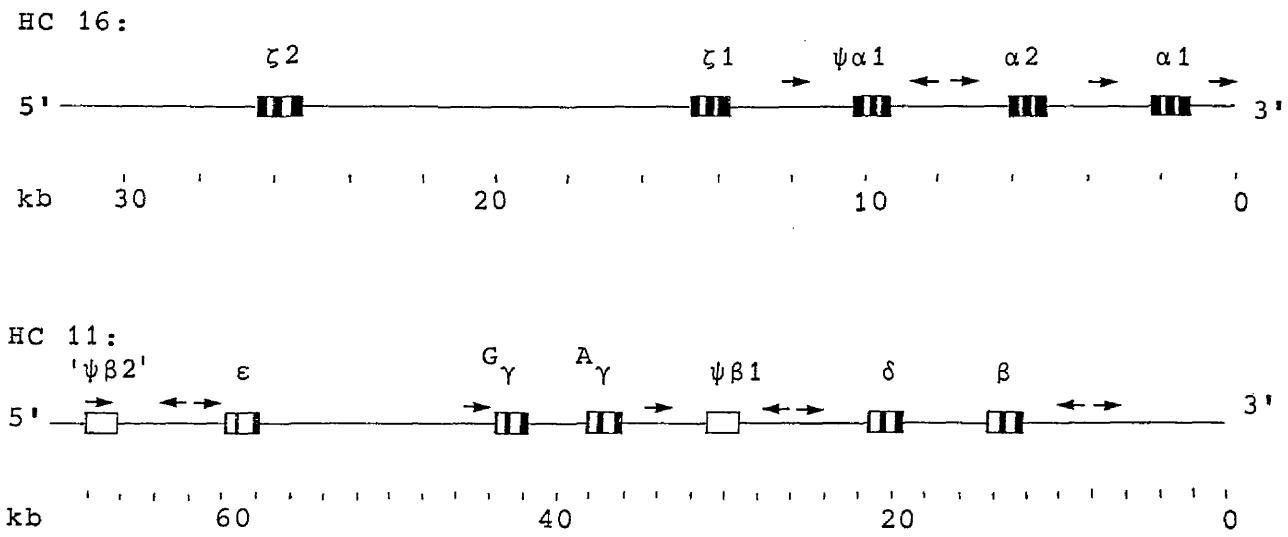


Fig 1.5. Linkage arrangement of human α- and non-α-globin gene clusters. For each gene the black and white boxes represent coding (exon) and noncoding (intron) sequences (from MANIATIS et al, 1980). The arrows (→) indicate the position of related DNA sequences (about 300 bp) detected in both clusters by LAUER et al (1980) and FRITSCH et al (1980) respectively. ψβ2 is not a globin-related sequence (SHEN and SMITHIES, 1982).

Table 1.6. Analysis of the organization of human globin genes.

Fact	Method	Ref
Two globin genes (α,β)	Family studies	(1)
δ/β linkage	Family studies (linkage disequil) Globin analysis (Hb Lepore)	(2,3)
γ/β linkage and $^{G}\gamma$-$^{A}\gamma$-δ-β arrangement	Globin analysis (Hb Kenya)	(4)
Two linked α-globin genes (α_1,α_2)	Family studies DNA analysis	(5,6)
Asinteny of globin genes and chromosome assignment (α on HC 16, non-α on HC 11)	Somatic cell hybr. Molecular hybr. in solution (Cot analysis)	(7)
Non-α-globin gene cluster (5'-$\psi\beta_2$-ε-$^{G}\gamma$-$^{A}\gamma$-$\psi\beta_1$-δ-β-3')	DNA analysis and cloning	(8,9, 10)
α-globin gene cluster (5'-ζ_2-ζ_1-$\psi\alpha_1$-α_2-α_1-3')	DNA analysis and cloning	(11,12)
$\psi\beta_2$ is not globin related	DNA sequencing	(13)

References: (1) BRADLEY et al, 1961. (2) BOYER et al, 1963. (3) BAGLIONI, 1962. (4) HUISMAN et al, 1972. (5) HOLLAN et al, 1972. (6) ORKIN, 1978. (7) DEISSEROTH et al, 1976, 1977, 1978. (8) FLAVELL et al, 1978. (9) LAWN et al, 1978. (10) FRITSCH et al, 1980. (11) PROUDFOOT and MANIATIS, 1980. (12) LAUER et al, 1980. (13) SHEN and SMITHIES, 1982.

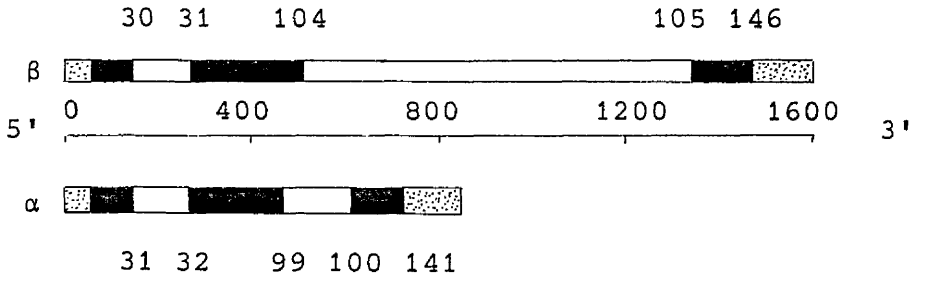


Fig 1.6. Structure of human globin genes. The canonical structures for the human α-like and β-like globin genes are drawn to approximate scale. Solid and open boxes represent coding and noncoding sequences, respectively. The α-like globin genes contain introns of approximately 95 and 125 bp, located between codons 31 and 32 and 99 and 100, respectively. The β-like globin genes contain introns of approximately 125 - 150 and 800 - 900 bp, located between codons 30 and 31 and 104 and 105, respectively (from MANIATIS et al, 1980).

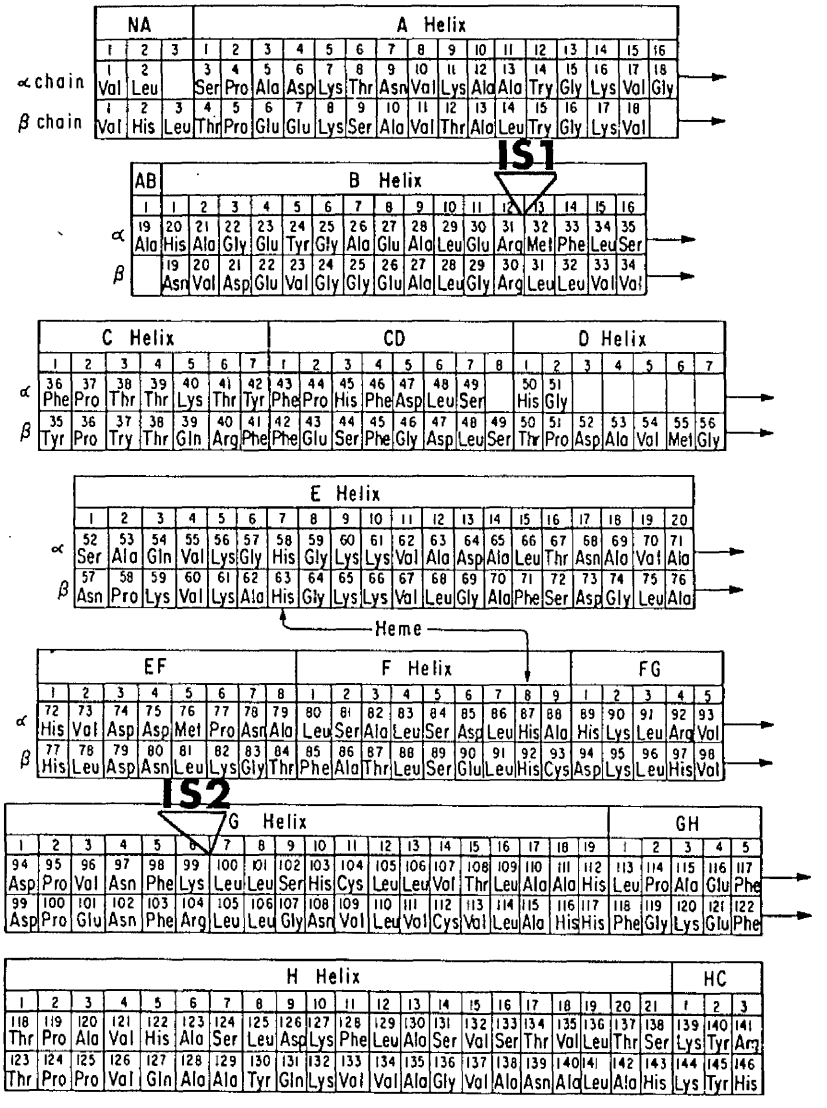


Fig 1.7. The aminoacid sequence of the α- and β-globin chains of human Hb A. The helices are given the letters A-H and the areas between the helices labelled AB, CD, etc. NA and HC represent the three amino and carboxyl terminal residues, respectively. IS1 and IS2 indicate the location of introns (adapted from WEATHERALL and CLEGG, 1981, p 22).

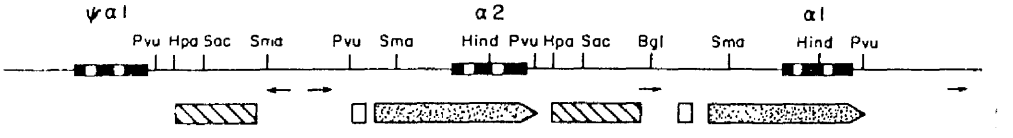


Fig 1.8. The distribution of sequence homologies within a region of the human α-globin gene cluster. Regions of sequence homology are indicated by cross-hatched boxes, white boxes, or stippled arrows. These homologies were detected by heteroduplex analysis of the cloned α-globin gene cluster (LAUER et al, 1980). (→) indicates the location of related DNA sequences.

a large family of DNA sequences (about 300 bp long, defined by Alu I restriction sites) reiterated approximately 500,000 times in the human genome (HOUCK et al, 1979; JELINEK et al, 1980; SCHMID and JELINEK, 1982).

1.2.2. Evolution of the globin gene family (JEFFREYS, 1982).

In cyclostomes only one globin chain has been detected. Human globin gene family includes eleven loci distributed in two linkage groups (alpha- and non-alpha-globin gene clusters). The aim of this section is to answer the question of how have these genes emerged during evolution. It has been possible to construct an evolutionary tree for the human globin gene family (Fig 1.10) and to design a scenario of the genetic events responsible for the evolutionary process. This has been accomplished basically by measuring the degree of homology of aminoacid and nucleotide sequences, respectively among globins and globin genes from different species or within species. The sequence of events could be summarized as follows:

a) Duplication of the primordial globin gene: Since the cyclostomes only have one globin chain and the bony fishes have both alpha and beta chains, it is assumed that the alpha/beta duplication occurred some time during the evolution of bony fish, approximately 500 Myear ago (DAYHOFF et al, 1972; DICKERSON and GEIS, 1980). The alpha-beta-globin gene linkage in the amphibian Xenopus suggests that alpha- and beta-globin genes evolved by a tandem duplication of the ancestral globin gene due to unequal crossing-over following mispairing of repetitive sequences flanking the duplicated gene (JEFFREYS et al, 1980).

b) Disruption of alpha-beta-globin gene linkage: This must have occurred during reptile evolution before the appearance of the mammalian line (300 Myear ago) but after the divergence of amphibians and reptiles (350 Myear ago). The unlinking of alpha- and beta-globin genes could have occurred either by translocation between the two genes or by a mechanism involving chromosome duplication (tetraploidization? OHNO, 1970), in which case each duplicate but unlinked alpha-beta-globin gene cluster could then evolve towards a mammalian like alpha- or non-alpha-gene cluster by tandem alpha- or beta-globin gene duplication plus silencing of the linked beta- or alpha-globin genes within each cluster (JEFFREYS et al, 1980; HOSBACH et al, 1983).

c) Evolution of two different globin gene clusters in birds and mammals: Tandem duplication, point mutation and concerted evolution (i. e., parallel and coincidental evolution of two genes which correct against each other via homologous but unequal crossing-over, SLIGHTOM et al, 1980; ZIMMER et al, 1980; LIEBHABER et al, 1981), are the mechanisms assumed to have brought about the evolution leading to the present alpha- and non-alpha-globin gene clusters. Regions with higher homology within the clusters (e. g., alpha 1-alpha 2 and gamma G-gamma A) have undergone more recent correcting genetic exchange.

In order to elucidate the detailed timing of the evolution of globin genes, some sort of time scale must be established. The rate of nucleotide substitutions resulting in aminoacid replacements (which appear to accumulate linearly with divergence time as documented by the fossil record) provide a reliable 'evolutionary

clock' (WILSON et al, 1977; PERLER et al, 1980). To calibrate the evolutionary clock, data included in table 1.7 have been considered. Based on this type of evidence a calibration curve for divergence for globin coding sequences can be drawn (Fig 1.9). This curve serves as a molecular clock. Its slope measures the rate of fixation of replacement substitutions in globin genes. Entering in the curve with the values of percentage divergence between a given pair of globin genes, an estimate of the time since the two genes begun to diverge can be obtained. Using this estimates EFSTRATIADIS et al (1980) constructed an evolutionary tree for the human globin gene family. Contributions from a number of different sources (ZIMMER et al, 1980; CZELUSNIAK et al,1982) resulted in the picture shown in Fig 1.10.
Branch points represent times at which gene duplication events took place. It is possible that the initial products of a given duplication were corrected against each other for an unknown period of time: alpha 1-alpha 2 and gamma G-gamma A gene pairs appear to be still at this stage of concerted evolution, while delta-beta gene pair seems to have evolved to a sequence divergence stage (MARTIN et al, 1983).

Gene clusters are inherently unstable entities that can expand, contract, and disperse during the course of evolution. Of all the possible shifts in rearrangement, natural selection will eliminate the unfavourable, leaving neutral and advantageous to be fixed in evolution (JEFFREYS and HARRIS, 1982). The present arrangement of human globin genes displays some of the remnants of such a process - the pseudogenes (DNA sequences that have originally derived from a functional globin gene, that are incapable of producing functional globin protein and that are fixed in the population,

Table 1.7. Calibration of the evolutionary clock for globin genes.

Gene pair	Average divergence (%)	Evolutionary event	Date from fossil record (Myear ago)	Ref
α-globin between mammals	10	mammalian radiation	85	(1)
β-globin between mammals	10			
α-globin birds/mammals	21	bird/mammal divergence	250-300	(2)
β-globin birds/mammals	24			
α/β within species	50	evolution of bony fish	500	(3)

References: (1) ROMERO-HERRERA et al, 1973. (2) WILSON et al, 1977. (3) DICKERSON and GEIZ, 1980.

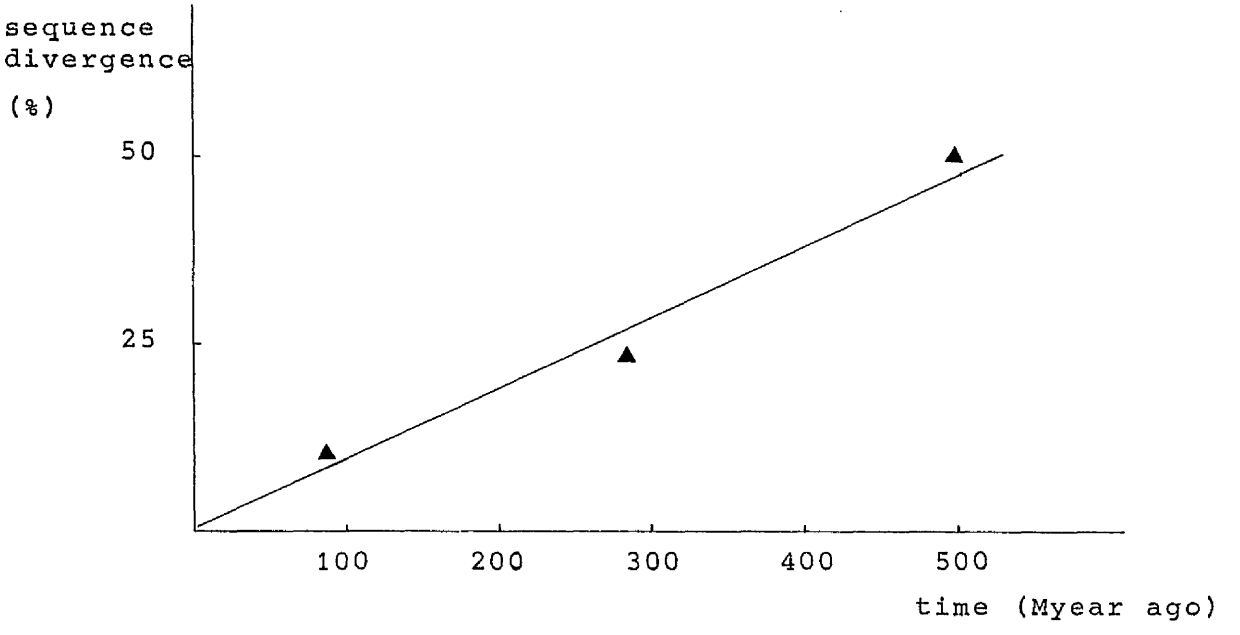


Fig 1.9. Calibration curve for divergence for globin coding sequences. The UEP (= Unit Evolutionary Period, time for the fixation of 1% sequence divergence) is 10.3 Myear for replacement substitutions (resulting in aminoacid change) and 1.3 Myear for silent substitutions (not resulting in aminoacid change because either of having occurred in non-coding sequences or of the degeneracy of the genetic code).

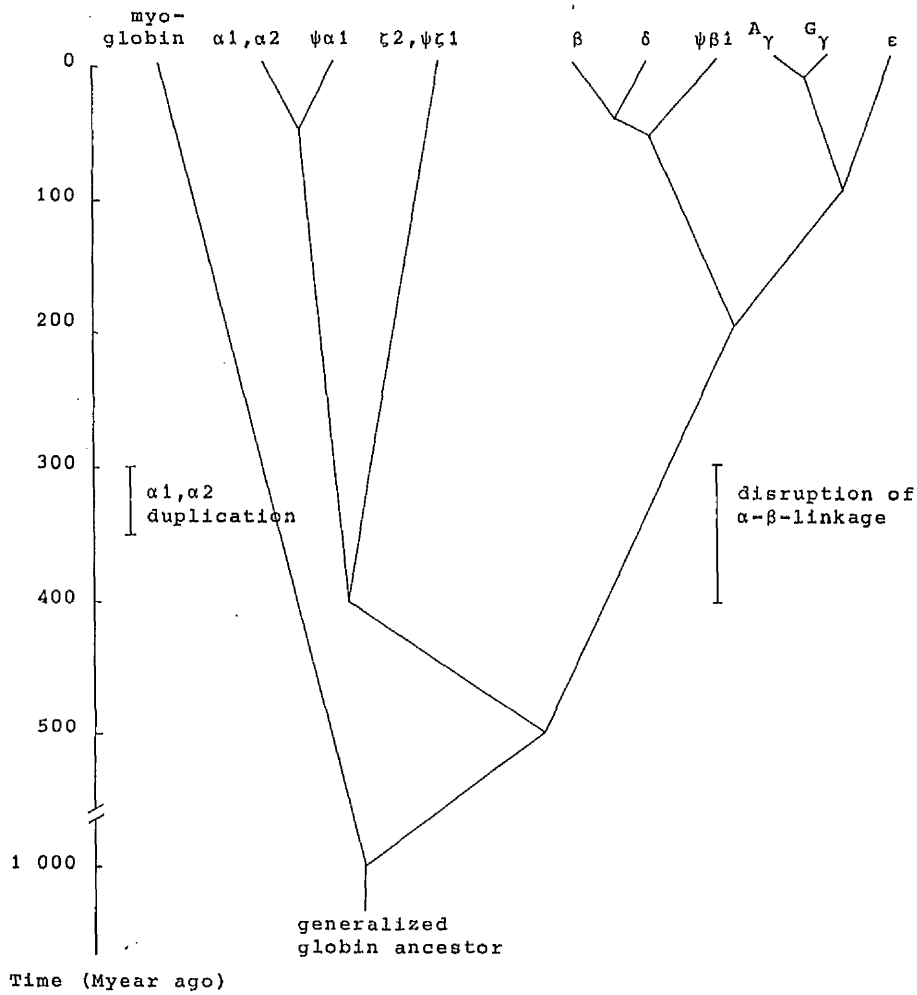


Fig 1.10. An evolutionary tree for human globins.

LITTLE, 1982*). Globin pseudogenes may have been generated in a two-stage process:

a) Gene duplication followed by a period (14 Myear for pseudo-beta 2, 15 Myear for pseudo-alpha 1, >40 Myear for delta; in Old World monkeys delta-globin gene is already a recently silenced pseudo-delta gene, MARTIN, 1980) of concerted evolution during which exchange of information occurs between the duplicated genes resulting in a mutual correction of mutations.

b) Sequence divergence following the disruption of the correcting mechanism. The time at which this occurs depends on the relative rates of mutation and correction. Once a gene has become incapable of producing a functional globin chain, it will then accumulate mutations more rapidly, since it is no longer under the constraints of selection in favour of functional protein sequences (LITTLE, 1982).

The feature of pseudogenes being evolutionary relics of duplicated genes does not exclude the possibility of expression of nonglobin proteins encoded by them. However, so far no new protein product has been suggested or identified for these pseudogenes (PROUDFOOT, 1980; JEFFREYS, 1981).
Sometimes a single nucleotide substitution is responsible for pseudogene disfunction. A 'stop' codon at aminoacid 6 indicates that the human gene previously

---

(*) An alternative definition is given by PROUDFOOT (1980): regions of DNA that display significant homology (>70%) to functional globin genes but have mutations which prevent their expression.

known as zeta 1 is, in fact, a pseudogene pseudo-zeta 1 (PROUDFOOT et al, 1982; LITTLE, 1982). This changes the current ideas about functionality of embryonic alpha-like globin genes. In their review of the molecular genetics of human globins, MANIATIS et al (1980) state that evidence suggests both zeta-globin genes are functional. On one side, detection of zeta-globin in an infant carrying a homozygous deletion that removes zeta 1, but spares zeta 2, indicates the latter is a functional gene (PRESSLEY et al, 1980b). On the other hand, the agreement between nucleotide sequence of zeta 1 gene and the aminoacid sequence of zeta-globin, suggests that the zeta 1 gene is probably also functional.

The above description of globin gene clusters portrays a dynamic gene family undergoing rapid evolutionary changes. This family includes genes which have specialized to function at different developmental stages (embryonic, fetal, and adult globin genes), genes for which no globin polypeptide has been identified and which may therefore no longer be functional (pseudo-genes), and genes which are presently undergoing mutual correction and/or deletion and duplication in the human population (gamma G-gamma A and alpha 1-alpha 2) (LAUER et al, 1980; SANCAR et al, 1982).

1.3. The method - DNA restriction fragment analysis combined with interspecific somatic cell hybridization.

1.3.1. The aim and the strategy.

The aim of this study was (i) to confirm and, if possible, to strengthen the regional localization of the human alpha-globin gene cluster on the short arm of HC 16, and (ii) to test the value in human gene mapping of an experimental approach combining somatic cell Genetics and recombinant DNA procedures.

The followed strategy consisted of:

- Production of man x mouse somatic cell hybrids, the human parental cells being carriers of balanced translocations involving HC 16.

- Selection and expansion of hybrid cell populations carrying HC 16q, taking advantage of the expression of a selectable marker (APRT) under restrictive conditions (AA selection).

- Counterselection and expansion of the hybrid cell populations which carry an intact HC 16 under restrictive conditions (DAP selection) that select against the same marker.

- Characterization of the cellular phenotype of the somatic cell hybrids in terms of human chromosome content and expression of relevant human enzymatic markers.

- Screening of the DNA from the different types of hybrid cell populations (either carrying HC 16, or a 16 fragment, or not carrying 16 at all) for the presence of the human alpha-globin gene cluster, by molecular hybridization on solid phase with a human genomic DNA probe after restriction endonuclease treatment.

- Establishing a correlation between the presence or absence of HC 16p and the presence or absence of alpha-globin gene in each hybrid cell population.

Previous attempts to localize the human alpha-globin gene cluster, using a number of different approaches, are summarized in Table 1.8 and Fig 1.11.

1.3.2. Discussion of the strategy and of some methodologies utilized in this study.

a) Until 1977 the method of choice in gene assignment was the correlation of the presence or absence in a hybrid cell population of a particular gene product with the presence or absence of a chromosome or chromosome fragment. This method has severe limitations because it depends on the expression of genes and on the ability to distinguish human and rodent gene products and chromosomes. Thus, genes which are expressed only in differentiated cells, as is the case of globin genes, cannot be mapped by this method. This dificulty can be overcome by direct detection of the genes using nucleic acid hybridization techniques. The assignment of alpha- and beta-globin genes by molecular hybridization in solution of alpha- and beta-globin cDNA probes with

Table 1.8. Chromosomal localization of human α-globin gene cluster.

Method	Hybridization probe (origin; label)	Localization	Ref
Molecular hybridization in solution (Cot analysis) of DNA from man x mouse hybrids	cDNA (abnormal human reticulocyte mRNA; ^{32}P)	16	(1)
Molecular hybridization on solid phase (Southern blotting) of DNA from man x mouse hybrids (human cells: t(11:16)(11q25::16q22) and t(3:16)(3q21:16q12))	cDNA (^{32}P)	16p12-pter	(2)
Molecular hybridization in situ	cRNA (human fetal liver DNA; ^{3}H)		
Molecular hybridization in situ	cDNA (human adult or premature infant reticulocyte RNA;^{125}I)	16	(3)
Trisomy mapping (16p12-pter) by assessment of red cell α/β chain synthesis ratio		16p12-pter	(4)
Molecular hybridization on solid phase of DNA from man x mouse hybrids (human cells: t(11:16)(11q13:16p11))	cDNA (human adult or premature infant reticulocyte RNA;^{32}P)	16p	(5)
Molecular hybridization in situ	cRNA (human fetal liver DNA; ^{3}H)	16p12-pter	(6)

References: (1) DEISSEROTH et al, 1977. (2) BARG et al, 1981. (3) GERHARD et al, 1981. (4)WAINSCOAT et al, 1981. (5) KOEFFLER et al, 1981. (6) BARTON et al, 1982.

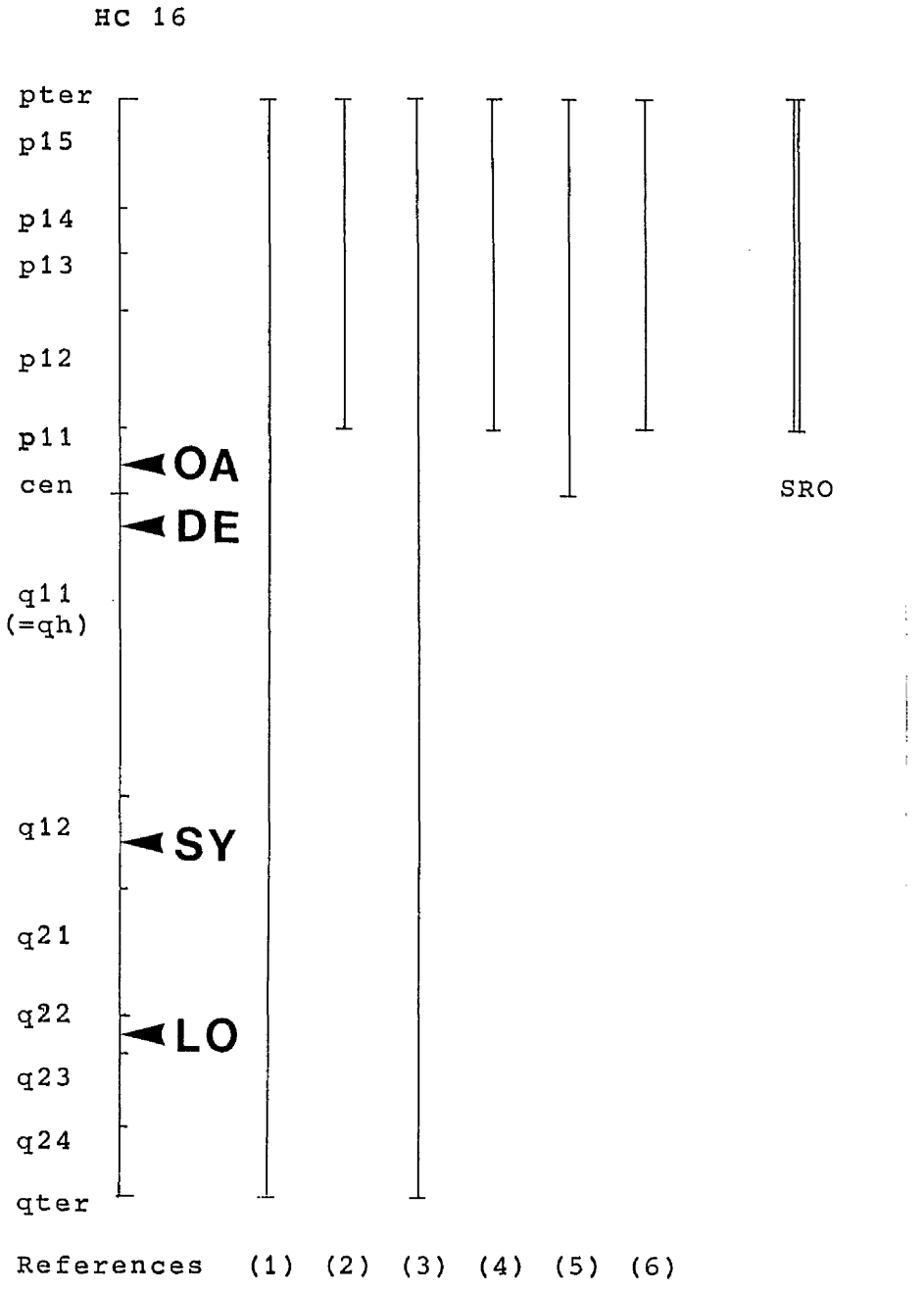


Fig 1.11. Intrachromosomal localization of the human α-globin gene cluster. SRO = shortest region of overlap. (◀) indicates the breakpoints of the rearranged HC 16s used in this study. DE, LO, OA and SY identify the subjects carrying these rearranged chromosomes. For references see Table 1.8.

genomic DNA preparations from hybrid cell populations, demonstrated the feasability of this new approach (DEISSEROTH et al, 1977 and 1978). The success of this method depends on the specificity of hybridization of the human gene probe with the homologous human genomic sequence, and so, appreciable interspecific cross-reaction with rodent DNA sequences will limit the use of Cot analysis as a gene mapping technique. This problem was solved by exploiting human and rodent differences in restriction endonuclease sites within or around the gene of interest and using a procedure of molecular hybridization on solid phase (SOUTHERN, 1975). In this method the presence of homologous DNA fragments of rodent origin is not a disadvantage, since the cross-hybridization signal can be used as an internal control for the method (Fig 1.12).

Restriction fragment mapping makes possible the mapping of any unique, or low copy number, nucleotide sequence for which there is a suitable hybridization probe. Among others, the human insulin gene (OWERBACH et al, 1980), and the human alpha- and non-alpha-globin gene clusters (KOEFFLER et al, 1981, and JEFFREYS et al, 1979, respectively) have been mapped using this approach. Another advantage of restriction fragment mapping is that it needs a much smaller amount (10 - 100 ug) of DNA than Cot analysis (1 - 2 mg). Regional mapping can be accomplished by using donor cells which carry reciprocal translocations involving a particular chromosome. In this study, four different translocations involving HC 16 have been used.

b) In the production of interspecific somatic cell hybrids the critical steps include the use of a 'hybridogen', to aid the cell fusion, and the choice of selective systems under whose pressure hybrid cells (carrying or lacking a particular donor chromosome) show

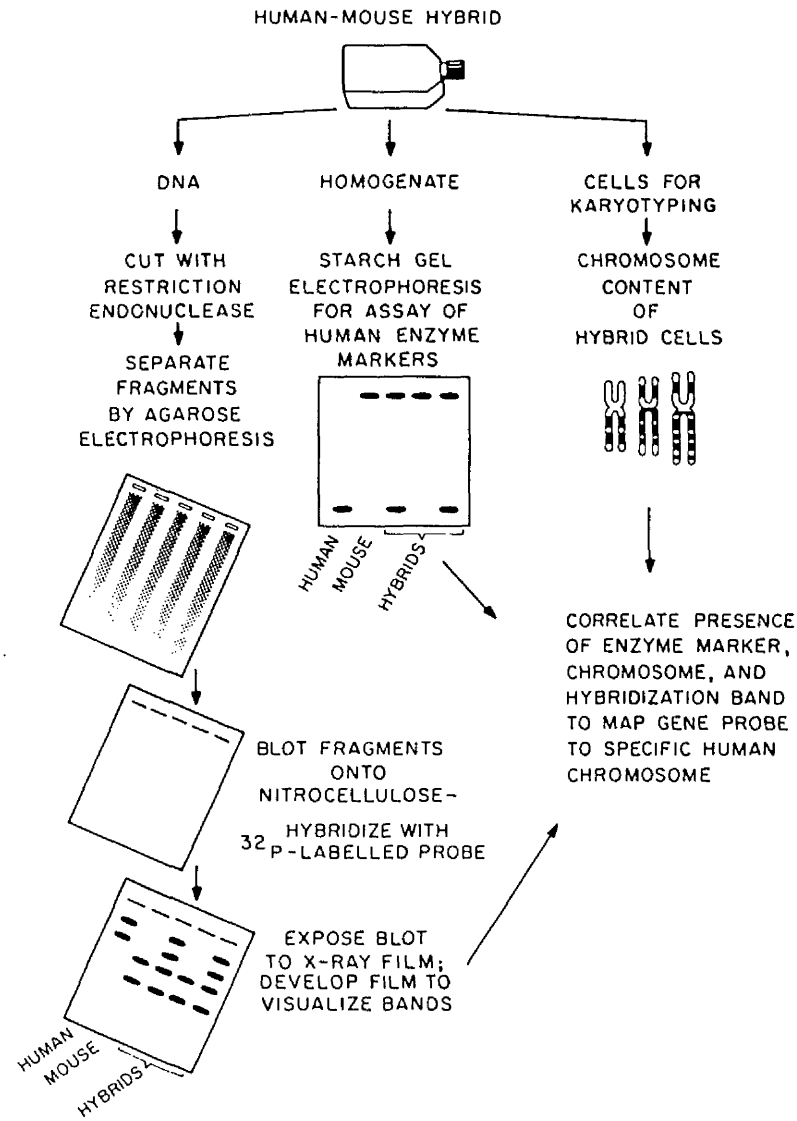


Fig 1.12. Mapping of cloned genes in human × rodent somatic cell hybrids by SOUTHERN blotting (from SHOWS et al, 1982).

a remarkable adaptive advantage. In this study, polyethyleneglycol (PEG, a nonionic, polymeric, surfactant which acts on the interacting cell membranes leading to the formation of heterokaryons) was used to increase the frequency of cell fusion events (PONTECORVO, 1976). To select hybrid cells carrying HC 16q, a mixture of ouabain (to eliminate human parental cells), adenine and alanosine (to eliminate APRT-deficient mouse parental cells) was used. The use of ouabain takes advantage of the differential toxicity of the drug (which inhibits plasma-membrane mediated, ATPase-dependent, K^+ intake and Na^+ output): human cells are sensitive to ouabain concentrations ranging from $3x10^{-8}$ to 10^{-7} M while mouse L cells show sensitivity at 10^{-3} M (CHU and POWELL, 1976). To select hybrid cells carrying no HC 16 at all, diaminopurine (DAP) was used. To understand the mechanism of action of AA and DAP selection systems it is necessary to keep in mind the biochemistry of purine biosynthesis, interconversion, reutilization, and catabolism (Fig 1.13). Mouse L A9 cells are APRT-deficient which means they cannot reuse ('salvage') adenine from nucleotide catabolism. For the same reason they are resistant to adenine analogues (e.g., DAP) which inhibit APRT. Their requirements for adenine nucleotides are supplied by the de novo synthesis of adenine via inosine-monophosphate. One of the enzymes involved in this pathway (adenylosuccinate synthetase) is inhibited by an aspartate analogue (alanosine). Now, if the growth medium contains both adenine (which APRT-deficient cells cannot utilize) and alanosine (which blocks the de novo synthesis of adenine), mouse L A9 cells will not survive. On the contrary, hybrid cells which carry the human APRT gene (located on HC 16q) will show adaptive advantage and therefore will preferentially survive by a mechanism of enzyme deficiency complementation (KUSANO et al, 1971).

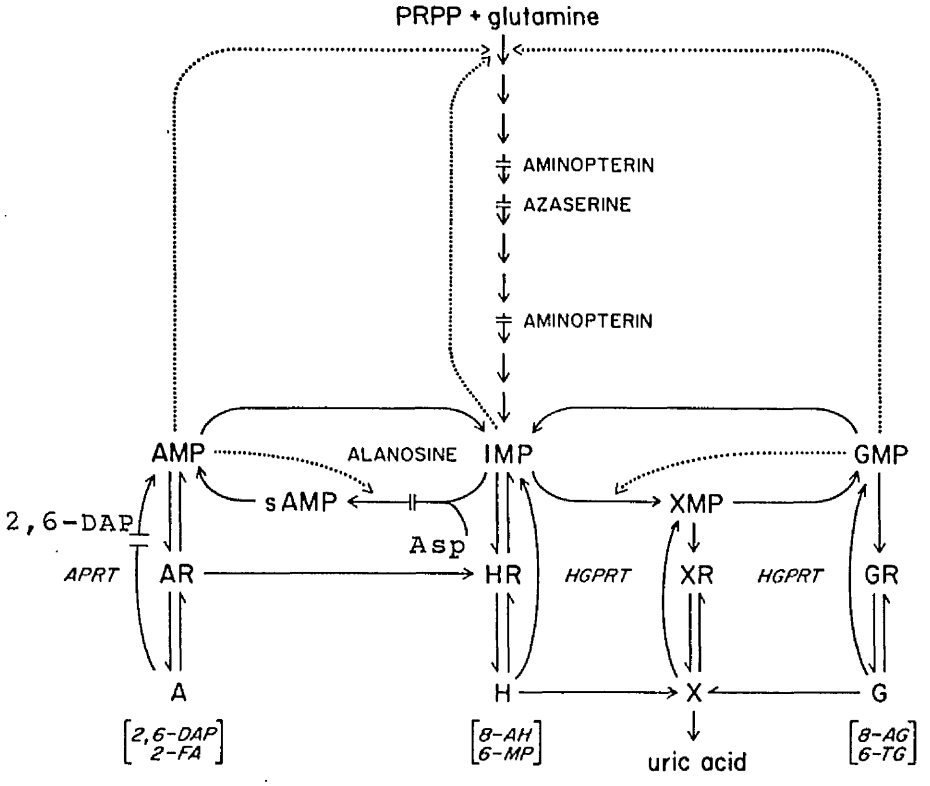
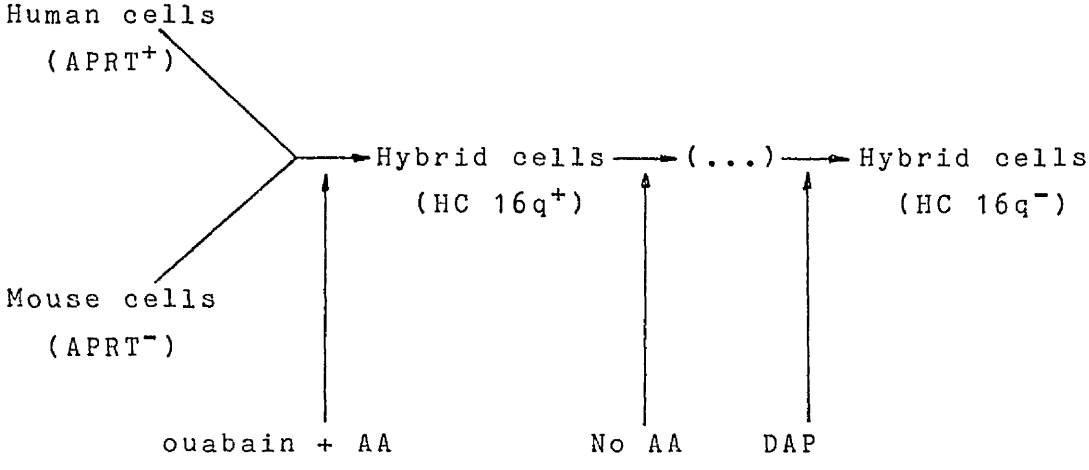


Fig 1.13. Purine biosynthesis, reutilization, interconversion, and catabolic pathways. APRT: adenine phosphoribosyltransferase (EC 2.4.2.7). HGPRT: hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8). A: adenine, AR: adenosine, AMP: adenosine 5'-monophosphate. Asp: aspartate. G: guanine, GR: guanosine, GMP: guanosine 5'-monophosphate. H: hypoxanthine, HR: inosine, IMP: inosine 5'-monophosphate. PRPP: 5-phosphoribosyl-1-pyrophosphate. sAMP: adenylosuccinic acid. X: xanthine, XR: xanthosine, XMP: xanthosine 5'-monophosphate. 2,6-DAP: 2,6-diaminopurine. 2-FA: 2-fluoroadenine. 8-AH: 8-azahypoxanthine. 6-MP: 6-mercaptopurine. 8-AG: 8-azaguanine. 6-TG: 6-thioguanine. Dotted arrows indicate feed-back inhibition (adapted from CHU and POWELL, 1976).

If, instead of adenine and alanosine, the growth medium contains DAP, the hybrids cells carrying HC 16q will be in selective disadvantage. The overall process of selection (for HC 16q) and counterselection (against HC 16q) can be summarized as follows:



c) The characterization of the cellular phenotype of the hybrid cells relies on the cosegregation of human chromosomes and human enzymatic markers (Table 1.9) on basis of species-specific banding patterns and other distinctive staining features (MILLER, 1972b), and electrophoretic behaviour respectively.

d) The detection of restriction DNA fragments depends on their specific hybridization with a labelled DNA probe (SOUTHERN, 1975). Increased sensitivity of hybridization techniques can be achieved using hybridization probes prepared from genomic recombinants that hybridize to flanking and intervening sequences as well as to the coding sequences (MALCOLM et al, 1981b). The recombinant plasmid αP7(α1) used as hybridization probe in this study carries a genomic DNA fragment including human alpha 1-globin gene and its intervening and surrounding sequences to a total length of 3.8 kb. It is a subclone

Table 1.9. Enzymatic markers used in this study.

Chromosome localization	Symbol	EC No*	Subunit structure
2p23	MDH1	1.1.1.37	dimeric
2q32→qter	IDH1	1.1.1.42	dimeric
5cen→q13 (β subunit)	HEXB	3.2.1.30	olygomeric
9q34	AK1	2.7.4.3	monomeric
16q12→q21	DIA4	1.6.2.2	dimeric
16q12→q22	APRT	2.4.2.7	dimeric
16p12→p13	PGP	3.1.3.18	dimeric
17q210→q220	GALK	2.7.1.6	dimeric

(*) IUB Nomenclature Committee, 1978.

of an element of a bacteriophage lambda bank of human fetal liver DNA (LAWN et al, 1978) screened with an alpha-globin cDNA probe (WILSON et al, 1978). The subcloned plasmid was constructed by inserting a PvuII restriction DNA fragment, containing the alpha 1-globin gene into pBR322 as cloning vector (Fig 2.2).

As mentioned above (see 1.2.1.2), a family of related DNA sequences of about 300 bp is scattered throughout the human genome (500,000 times/haploid genome according to SCHMID and JELINEK, 1982). Elements of this family have been found interspersed within the non-alpha-globin gene cluster (FRITSCH et al, 1980) and it has been suggested that they are also present within the alpha-globin gene cluster (LAUER et al, 1980). For this reason, specially stringent posthybridization washing must be used to remove nonspecific molecular hybrids, taking advantage of the fact that this family of related DNA sequences displays about 20% divergence from the consensus sequence.

1.4. Analysis of the human genome and genetic disease - the haemoglobinopathies.

The assignment of genes to specific regions on chromosomes, in conjunction with other levels of gene mapping, provides an insight into the organization and regulation of the expression of genetic information (see 1.1.3). In particular, a more detailed knowledge of the structure and mode of action of the alpha-globin genes (which have now been localized, isolated, cloned, and partly sequenced) has as a consequence developed the understanding, and the means of diagnosis, of inherited disorders of haemoglobin - the haemoglobinopathies (for classification see Table 1.10).

Thalassaemias are genetic disorders characterized by a reduced rate of production of one or more of the globin chains of haemoglobin (WEATHERALL and CLEGG, 1981). Although the majority of alpha-thalassaemias appears to be due to gene deletion, some nondeletion types have been reported in which no gross gene deletion can be detected. At least six deletions resulting in alpha-thalassaemia have been described (Fig 1.14).

The severity of the disease depends on the extent of the deletion rather than on the degree of impairment of a single alpha-globin gene. It is remarkable that the locations of the breakpoints of deletions naturally associated with alpha^{+}-thalassaemia are indistinguishable from those which occur in vitro during propagation of recombinant bacteriophages, containing alpha-globin genes, in *E. coli* (LAUER et al, 1980).

The nondeletion types of alpha-thalassaemia appear to be due to a number of different gene defects: (i) Point mutation in the terminator codon resulting in the trans-

Table 1.10. Inherited disorders of haemoglobin (WEATHERALL and CLEGG, 1981, p 85).

In the structure of globin chains (structural variants)
- no clinical abnormality
- haemolytic anaemia
- abnormal oxygen transport
In the rate of synthesis of globin chains (reduced rate)
- thalassaemia
In the structure and rate of synthesis of globin chains
- thalassaemia (interaction with structural variants)
Hereditary persistence of fetal haemoglobin (HPFH)

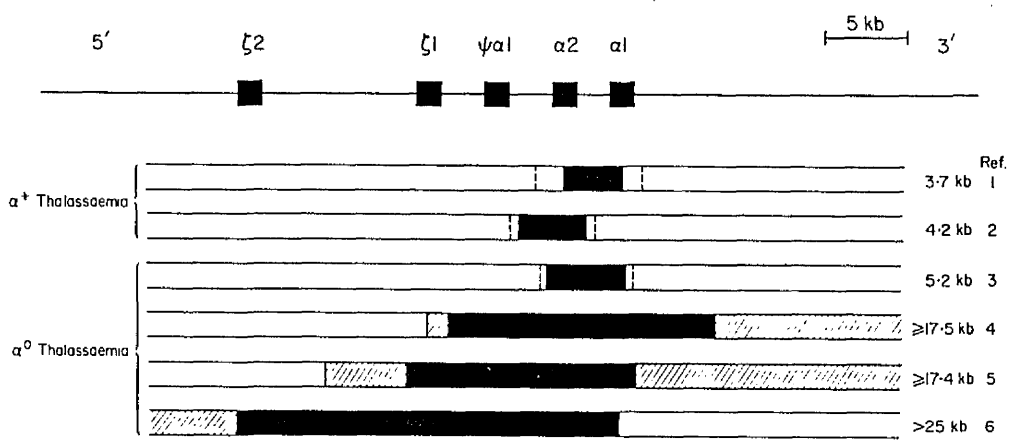


Fig 1.14. Deletions in the various α-thalassaemia haplotypes resulting in phenotypes α^+ (minimal red-cell abnormalities and little detectable globin chain imbalance) and α° thalassaemia (severe red-cell changes and an appreciably reduced α-chain output). Black areas represent the amount of DNA that has been deleted from within the area defined by the vertical dotted lines. In the last three cases areas known to be deleted are shown in black. Where the limits of the deletion are not defined their maximum extremes are indicated by hatched areas. Refs: (1) ORKIN, 1978. (2) EMBURY et al, 1979. (3) PRESSLEY et al, 1980a. (4,5) PRESSLEY et al, 1980b. (6) ORKIN and MICHELSON, 1980. From WEATHERALL and CLEGG, 1981, p 776.

location of 3', normally, untranslated region into an elongated (nonfunctional) alpha-globin chain (WEATHERALL and CLEGG, 1975; MICHELSON and ORKIN, 1980). (ii) Short deletion in an intervening sequence resulting in abnormal RNA splicing (ORKIN et al, 1981; FELBER et al, 1982); (iii) Single nucleotide substitution in the coding region associated with alpha/beta interaction during haemoglobin assembly impeding $alpha_1$-$beta_1$ formation (GOOSSENS et al, 1982).

The diagnosis of disfunctional globin genes is accomplished by a combined approach: (i) clinical and haematological analysis of patients and their families, including studies of any interaction between the genetic determinant for thalassaemia with those for different structural haemoglobin variants; and (ii) analysis at the molecular level including the measurement of rates of globin chain synthesis, the determination of the ratio of alpha- and beta-globin mRNA in red-cell precursors, the detection of deletions, and the quantification of globin genes.

Regarding the prenatal diagnosis of haemoglobinopathies, three alternative routes of fetal sampling can be considered: chorion biopsy, amniocentesis and puncture of umbilical chord vessels. These methods may be assessed in terms of their timing, fetal loss rate (over the base line for the gestational age of sample collection), ability to provide information and extent of contamination by maternal tissue:

a) Trophoblast: Chorion biopsy is performed during the first trimester of pregnancy under endoscopy (and optional ultrasound) guidance. It involves a fetal loss rate estimated in 1% and it is relatively less invasive than the other two methods. It provides enough cells to

prepare DNA for direct analysis of globin gene structure (WILLIAMSON et al, 1981; OLD et al, 1982).

b) Amniotic fluid cells: Amniocentesis is performed during the second trimester of pregnancy under ultrasound guidance. It involves a fetal loss rate of less than 1%, although it is more invasive than chorion biopsy. DNA can be extracted from fresh or cultured amniotic fluid cells and the globin gene structure analysed (WONG et al, 1978; CHAN and KAN, 1982).

c) Fetal blood: Puncture of umbilical chord vessels is performed during the second trimester of pregnancy under fetoscopy guidance. It involves a much higher fetal loss rate (4 - 10%), being the most invasive of the three methods. However, it is the only one to give access to the study of the actual expression of globin genes at the level of RNA processing, globin synthesis and haemoglobin assembly (ALTER, 1979).

At the moment, the extent of maternal contamination of fetal samples is still higher in chorion biopsy than in the other two methods. Nevertheless, the fact that chorion biopsy can be performed safely at an early stage of pregnancy, suggests that it may become the method of choice in the prenatal diagnosis of haemoglobinopathies (GILMORE, D, unpublished review).

Prenatal diagnosis of homozygous alpha-thalassaemia was achieved first by direct 'titration' of the alpha-globin genes in a DNA/cDNA hybridization assay of the DNA extracted from cultured amniotic fluid fibroblasts (WONG et al, 1978).
As it was pointed out before (see 1.1.3.1.b)), DNA RFLPs segregating in the population can be used as linkage markers in the indirect detection of disfunctional genes

by the linkage method. Alternatively, suitable restriction endonucleases can detect mutated restriction sites within the abnormal gene itself as it is the case of the Dde I polymorphism in sickle-cell mutation (LITTLE, 1981; BOEHM et al, 1983).

Recently, WEATHERALL et al (1981) published their findings of new mutations in the alpha-globin gene cluster in three patients who coincidentally (?) suffer from a severe form of alpha-thalassaemia (Hb H disease) and are mentally retarded. The Authors suggest the new mutations may be related to the developmental changes in the patients either (i) because novel mutations affecting the synthesis of alpha-, and probably zeta-, chains induce intrauterine hypoxia at a very early stage of embryonic or fetal development or (ii) because of the effect of an abnormality of HC 16 (encompassing other than the globin genes) on fetal maturation. The Authors are now carrying out a survey of the mentally retarded in order to bring to light other cases of thalassaemia. If analysis of rate of production of a gene product, such as alpha-globin chain, can identify subtle chromosomal abnormalities that are not definable by current cytogenetic methods, the application of this strategy could provide an useful approach to the study of unexplained cases of mental retardation. Also, the demonstration of the feasability of such an approach would strengthen the conviction that analysis at the molecular level (either of the genes or of their detectable products) can be informative of chromosomal rearrangements otherwise nondetectable.

Finally, just a few words about the prospects of gene therapy of haemoglobinopathies (as the paradigm of well understood single gene defects leading to disease). Essentially, two different strategies can be devised for

human gene therapy: (i) Treatment of affected tissues from patients with a genetic disease (e.g., blood and bone marrow in thalassaemic patients); this kind of treatment would only affect the patient itself. (ii) Insertion of a normal gene into early embryos; in this case the patient's progeny would be altered as well. It is possible now to isolate normal genes, to introduce them into mammalian cells (e.g., of rabbit or mouse) at any stage of development, including the fertilized embryo, and to obtain gene function in a random way and at a low level (STEWART et al, 1982). However, it is not yet possible to ensure that the newly inserted genes function under normal control in the animal, in time, space and level of expression. Technical difficulties and ethical problems of gene replacement in the therapy of human genetic disease must be balanced against antenatal diagnosis/termination of affected pregnancies and conventional postnatal therapies (WILLIAMSON, 1982). So, until more is known about the potential for the expression of transferred human genes in human cells in an in vivo model, it is far too premature to attempt gene replacement experiments in patients (WEATHERALL and CLEGG, 1981).

As an alternative to 'genetic surgery', it seems possible to achieve clinical management of deficient gene expression by molecular manipulation of DNA in vivo. This has been tryed by LEY at al (1982) using a methylation inhibitor (5-azacytidine) to reverse transiently haemoglobin synthesis to the fetal mode in beta-thalassaemic patients.

2. Materials and methods.

2.1. Production of man x mouse somatic cell hybrids.

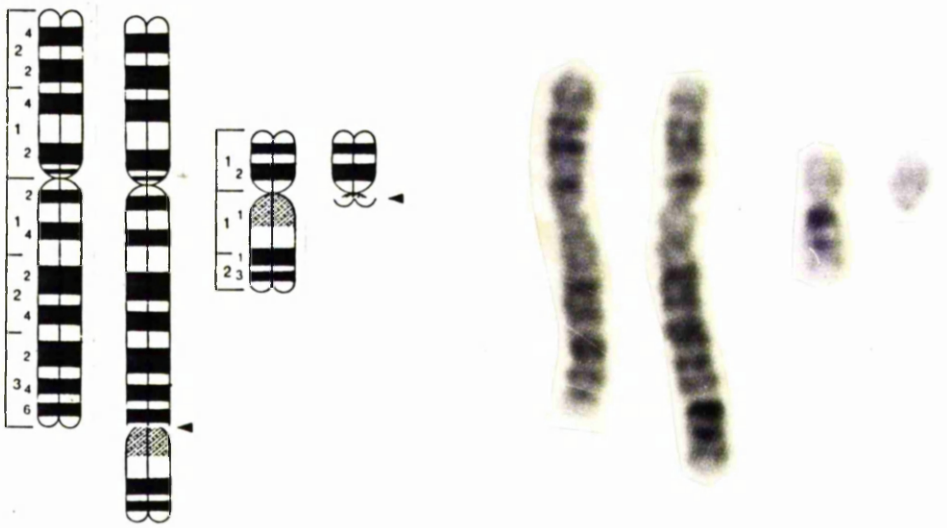
2.1.1. Parental cells.

Somatic cell hybrids were obtained by fusion of mouse L A9 cells with human fibroblasts. The L A9 mouse cell line had been established from connective tissue. It bears a heteroploid chromosome number (mean = 70; range: 40 - 115) and resistance to 8-azaguanine (HPRT deficiency) and to 2,6-diaminopurine (APRT deficiency). Human primary cell cultures originated from skin biopsies and were obtained from individuals with both normal and abnormal karyotypes (Fig 2.1). The chromosome abnormalities (balanced translocations involving HC 16) had been detected and characterized in the diagnostic cytogenetic laboratories: SY in the Duncan Guthrie Institute of Medical Genetics, Glasgow, and DE, LO and OA in the Human Genetics Laboratory, Instituto Nacional de Saúde, Lisbon.

Mouse cells were grown in Eagle's medium (Glasgow modification; MACPHERSON and STOKER, 1961; STOKER and MACPHERSON, 1962) supplemented with L-glutamine (2 mM) and fetal bovine serum (10%(v/v)). The pH of the medium was adjusted to 7 with 5%(w/v) sodium bicarbonate. Human cells were grown in the same medium further supplemented with non-essential aminoacids (Ala, Asn, Asp, Gly, Glu, Pro, Ser; Gibco, Paisley), pyruvate (1 mM) and fetal bovine serum (20%(v/v)).

Both cell types grow in monolayer culture.

DE 46,XX,t(2:16)(2q37:16q11)



LO 46,XX,t(9:16)(9q22:16q22)

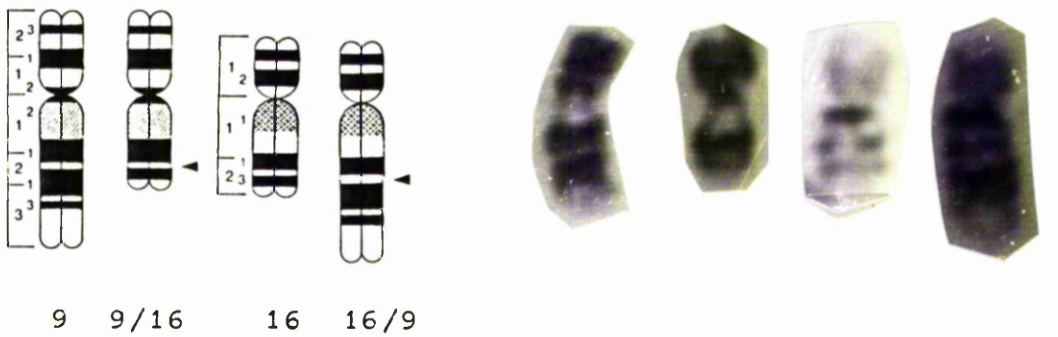
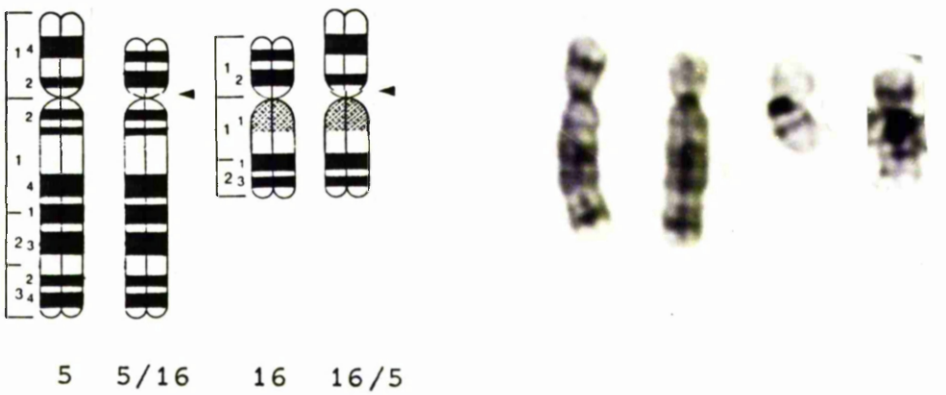


Fig 2.1. Rearranged human chromosomes used in this study.

OA 46,XX,t(5:16)(5p11:16p11)



SY 46,XY,t(16:17)(16q12:17p11)

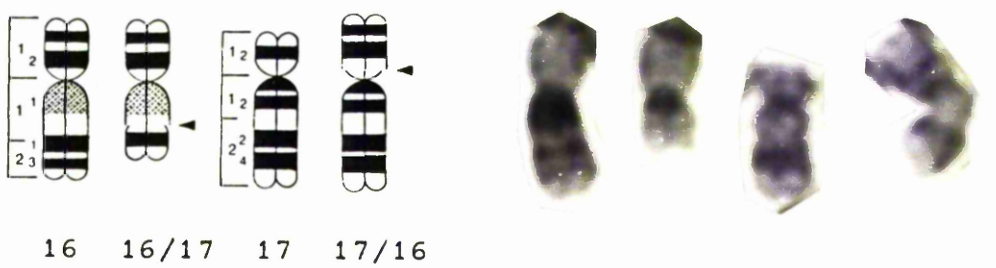


Fig 2.1. Cont.

2.1.2. Cell fusion (modified from PONTECORVO, 1975).

Two million cells from each species were mixed together pelleted and treated with 1 ml of 50%(w/w) PEG 6 000 (average molecular weight: 6 000 - 7 500) in Eagle's medium, for 1 minute. A 40-fold dilution of PEG with Eagle's medium followed immediately. After pelleting ($1o^3$ rpm, 5 minutes), cells were plated in 6 cm-diameter Petri dishes, at a density of 10^5 cells/dish, and fed with Ham's F12 medium supplemented with L-glutamine (2 mM), non-essential aminoacids, pyruvate (1 mM) and fetal bovine serum (10%(v/v)), and incubated at 37°C, in an atmosphere containing 5%(v/v) carbon dioxide.

2.1.3. Selection of hybrid clones carrying HC 16q (KUSANO et al, 1971).

24 - 48 hours after fusion, cells were fed (twice during the first week and then weekly) with the same medium plus the selective agents: 50 μM adenine, 50 μM alanosine and 2.5 μM ouabain. Well defined, active colonies were picked up using a Pasteur pipette and transferred into cell culture bottles containing the same medium. After the first passage, ouabain was excluded from the selective system.

2.1.4. Subcloning of primary clones.

Cytogenetically heterogeneous cell populations were subcloned by plating a diluted cell suspension (approximately 5 cells/ 20 mm-diameter dish). Colonies were picked up as described before.

2.1.5. Mass culture of cloned hybrid cell popuplations.

Cells from well defined, active colonies were grown in

small plastic bottles (25 cm^2 growth area) until confluency. Then cells were trypsinized and tranferred to medium sized bottles (80 cm^2). At this stage a medium sized bottle for cytogenetic analysis was set up. If the hybrid cell population proved to be informative for the purpose of this study, it was further expanded in larger bottles (170 cm^2) to the level of 10^7 - 10^8 cells. These were used for enzyme and DNA analysis as well as for confirmation of the human chromosome complement. Thus, all the results shown refer to the same transfer number of the cell cultures. At each passage, cells were frozen in 10%(v/v) DMSO and stored in liquid nitrogen.

2.1.6. Counterselection against HC 16q.

Those hybrid clones which carried an intact HC 16 were grown in non-selective Eagle's medium supplemented with L-glutamine (2mM), non-essential aminoacids, pyruvate (1mM) and fetal bovine serum (10%(v/v)) for three weeks. Then they were grown in the same medium plus the counterselective agent (0.1 mM 2,6-diaminopurine) and human chromosome content monitored cytogenetically until complete segregation of HC 16 was observed. The counter-selected cell population was subsequently expanded according to the procedures described above.

2.1.7. Mycoplasma testing (CHEN, 1977).

Mycoplasma testing was performed regularly during the production of somatic cell hybrids and immediately before pelleting for biochemical and DNA analyses.

The cells were grown in Petri dishes for at least four days and tested before they were confluent. Cells were fixed in situ with methanol/acetic acid (3+1). The staining solution consisted of 0.05 µg/ml bisbenzimid -

Höechst 33258 in Dulbecco's phosphate buffered saline (DULBECCO and VOGT, 1954). Fixed cells were stained for 10 minutes, washed with distilled water, mounted in McIlvainie's (phosphate/citrate) buffer pH = 5.5, and observed immediately under UV light using a fluorescence microscope.

2.2. Cytogenetic analysis of the hybrids.

2.2.1. Chromosome preparations.

Harvest was carried out 24 hours after trypsinization of hybrid cells. Cell cycle was arrested in metaphase with colchicine (0.2 μg/ml, 37°C, 90 minutes). Cells were trypsinized, washed with versene buffer pH = 7.4, treated with hypotonic salt solution (0.075 M potassium chloride, 37°C, 15 minutes) and fixed in methanol/acetic acid (3+1) at 4°C for, at least, 30 minutes. After two more fixing cycles, fixed cells were dripped onto wet glass slides and left to air dry (MOORHEAD et al, 1960).

2.2.2. Staining procedures.

a) Trypsin/Leishman (modified from SEABRIGHT, 1972): Slides were soaked in 2xSSC at 60°C for two hours. Dehydration was carried out in water/ethanol mixtures of increasing ethanol concentration (50%(v/v), 70%, 90%, 95%, 100%; 2 - 3 minutes each). After drying in the air, the dehydrated slides were treated with a trypsin solution whose concentration, temperature and time of action depended on the cell type. A concentration range from 0.025% to 0.5% of Bacto-Trypsin Difco (1:250) and a time range from 10 s to 90 s, at room temperature, were used in this study. Enzyme action was stopped by extensively washing the slides with saline. The staining solution consisted of a 1:4 dilution of Leishman's stain (BDH, Poole) in phosphate buffer, pH = 6.8 (G T Gurr). Slides were flooded with the staining solution for 3 - 5 minutes, extensively washed with phosphate buffer, pH = 6.8, and left to air dry.

b) Quinacrine/Bisbenzimid - Höechst 33258 (KUCHERLAPATI

et al, 1975): When appropriate, chromosome preparations were destained with methanol, rehydrated by reversing the series of water/ethanol mixtures (from 100% to 50%; 2 - 3 minutes each), and soaked in McIlvainie's buffer, pH = 7.0. The staining solution consisted of 50 µg/ml quinacrine hydrochloride and 0.1 µg/ml Höechst 33258 in McIlvainie's buffer, pH = 7.0. Slides were stained for 20 - 30 minutes, washed extensively with the same buffer, mounted in McIlvainie's buffer, pH = 4.5, and analysed immediately under UV light in a fluorescence microscope.

2.2.3. Microscopy and micrography.

10 to 30 mitoses of each cell line were analysed (at least 20 mitoses if the cell line proved to be informative for this study) and the best and/or critical chromosome spreads photographed (3 to 10 photographs/cell line). All the cases which could not be conclusively characterized by G-banding were also analysed using the differential staining procedure to determine the parental origin of the chromosomes (quinacrine/Höechst 33258).

2.3. Biochemical analysis of the hybrids.

2.3.1. Cell pelleting and storage.

Cells were pelleted when the cultures reached confluency in large culture bottles (approx 10^7 cells). The medium was removed and cells trypsinized, centrifuged (1 000 - 2 000 rpm), washed three times with Dulbecco's phosphate buffered saline - part A, at 4°C and the final dry pellet stored in liquid nitrogen.

2.3.2. Cell lysis and protein estimation in the lysate.

Pelleted cells were resuspended in a lysis buffer (5 mM tris/HCl, pH = 7.5, 1 mM Na_2EDTA, 1 mM dithiothreitol) and sonicated (4°C) until homogenization. The lysate was then centrifuged (10 000 g, 4°C, 30 minutes) and the clear supernatant recovered and kept in liquid nitrogen. A protein estimation (LOWRY, 1951) was carried out on the supernatant.

2.3.3. Electrophoretic detection of biochemical markers.

Essentially two modes of electophoresis of proteins were used in this study - cellulose acetate gel (Cellogel) and starch gel electophoresis (HARRIS and HOPKINSON, 1976).

a) Cellulose acetate gel: Gel sheets were stored in 30%(v/v) methanol in a sealed container and kept at 4°C. Before use, the gels were blotted in filter paper (but not left dry), soaked for 30 minutes in the running buffer and blotted again. Then they were mounted in the electophoresis tank and pre-run at 200 V for 10 minutes. 10 - 15 µl of cell lysate were loaded on the gels and allowed to be absorbed before running the electo-

phoresis. After the electophoretic run the gels were blotted, and treated with the appropriate reaction mixture at 37°C, for the time specified in each case, in a closed box, protected from light.

b) Starch gel: A 12%(w/v) starch gel was prepared by suspending hydrolysed starch in the gel buffer at 100°C. The molten gel was poured into a perspex slab and allowed to solidify. 10 to 15 µl-samples were soaked on small pieces of filter paper and inserted into slits left open in the gel. The samples were then sealed with molten vaseline and the electophoresis run vertically (4°C). After the electophoretic run the gels were allowed to cool for 30 minutes, sliced, and treated with the appropriate reaction mixture at 37°C, for the time specified in each case. When the detection of biochemical markers involved the use of radioactively labelled substrates, stained gels were fixed in 0.1 M lanthanum chloride/0.1 M tris pH = 7.0 (20 hours, 4°C). After fixation the gels were washed with running water (3 - 4 hours), dried in a vacuum drier (4.5 hours with heat), between filter paper and a dialysis membrane, and exposed to X-ray film for the appropriate time.

2.3.4. The markers.

a) EC 1.1.1.37. Malate dehydrogenase 1 (HARRIS and HOPKINSON, 1976):

- Starch gel electophoresis (3 V/cm, 17 hours, 4°C).
- Bridge buffer: 0.245 M monosodium phosphate/0.15 M citric acid, pH = 5.9.
- Gel buffer: bridge buffer diluted 1:40.
- Reaction mixture: 50 mM L-malic acid, 0.2 mg/ml NAD, 0.15 mg/ml MTT, 0.1 mg/ml PMS, 1%(w/v) agar in 50mM Tris/HCl buffer, pH = 8. Incubation time: 30 minutes.
- Detection: colour.

b) EC 1.1.1.42. Isocitrate dehydrogenase 1 (HARRIS and HOPKINSON, 1976):
- Starch gel electrophoresis (3 V/cm, 17 hours, 4°C).
- Bridge buffer: 0.245 M monosodium phosphate/0.15 M citric acid, pH =5.9.
- Gel buffer: bridge buffer diluted 1:40.
- Reaction mixture: 1.3 mM isocitric acid (trisodium salt, dihydrate), 0.1 mg/ml Na_2NADP, 0.1 mg/ml MTT, 0.1 mg/ml PMS, 1%(w/v) agar in 0.1 M Tris/HCl buffer, pH = 8.0, 57 mM magnesium chloride. Incubation time: 30 minutes.
- Detection: colour.

c) EC 1.6.2.2. Diaphorase 4 (modified from GRZESCHIK, 1980 and EDWARDS et al, 1980):
- Cellulose acetate gel electrophoresis (200 V, 1.5 hour, room temperature).
- Buffer: 0.1 M tris/50 mM borate, ph = 8.6, 70 M NADH.
- Reaction mixture: 2.7 mg/ml NADH, 0.54 mg/ml FAD, 27 µg/ml 2,6-dichlorophenolindophenol, 54 µg/ml MTT in 1 M Tris/HCl buffer, pH = 8.6, 0.004 M Na_2EDTA. Incubation time: 10 - 30 minutes.
- Detection: colour.

d) EC 2.4.2.7. Adenine phosphoribosyltransferase (modified from HARRIS and HOPKINSON, 1976):
- Starch gel electophoresis (5 V/cm, 17 hours, 4°C).
- Bridge buffer: 27 mM citric acid/167 mM dipotassium phosphate, pH = 6.8.
- Gel buffer: 1.21 mM citric acid/ 6.07 mM dipotassium phosphate, pH = 6.8.
- Reaction mixture: 8.9 µM ^{14}C-8-adenine (specific activity: 40 - 50 mCi/mmole), 0.25 mM PRPP in 0.1 M Tris/HCl buffer, pH = 7.4, 10 mM magnesium chloride. Incubation time: 1.5 hour.

- Detection: autoradiography.

e) EC 2.7.1.6. Galactokinase (HARRIS and HOPKINSON, 1976):
- Starch gel electophoresis (5 V/cm, 17 hours, 4°C).
- Bridge buffer: 27 mM citric acid/167 mM dipotassium phosphate, pH = 6.8.
- Gel buffer: 1.21 mM citric acid/6.07 mM dipotassium phosphate/0.5mM DL-dithiothreitol, pH = 6.8.
- Reaction mixture: 13 μM ^{14}C-D-galactose (specific activity: 45 - 50 mCi/mmole), 3.61 mM Na_2ATP in 0.2 M Tris/HCl buffer, pH=7.2 7.4 mM magnesium chloride. Incubation time: 1 - 2 hours.
- Detection: autoradiography.

f) EC 2.7.4.3. Adenylate kinase 1 (modified from HARRIS and HOPKINSON, 1976):
- Starch gel electrophoresis (10 V/cm, 4.5 hours, room temperature).
- Bridge buffer: 0.1 M Tris/0.025 M citric acid, pH=8.0.
- Gel buffer: bridge buffer diluted 1:5.
- Reaction mixture ('forward positive' staining): 1 mM ADP (disodium salt), 10 mM magnesium chloride, 2 mg/ml glucose, 0.25 mg/ml Na_2NADP, 125 mU/ml hexokinase, 0.125 mg/ml MTT, o.125 mg/ml PMS, 1%(w/v) agar in 0.1 M Tris/HCl buffer, pH = 8.0. Incubation time: 30 minutes.
- Detection: colour.

g) EC 3.1.3.18. Phosphoglycollate phosphatase (modified from HARRIS and HOPKINSON, 1976):
- Starch gel electrophoresis (3 V/cm, 17 hours, 4°C).
- Bridge buffer: 0.1 M tris/0.1 M maleate/0.01 M Na_2EDTA/0.01 M megnesium chloride, pH = 7.2.
- Gel buffer: bridge buffer diluted 1:10, 0.1%(v/v) 2-mercaptoethanol.
- Reaction mixture:

Stage I: 25 mM 2-phosphoglycollic acid, 0.8 mM magnesium sulphate, 1%(w/v) agar in 0.05 M tris/HCl buffer, pH = 7.5. Incubation time: 2 hours.
Stage II (after removal of the previous agar overlay): 1.25%(w/v) ammonium molybdate in 2N sulphuric acid, 25 mg/ml ascorbic acid, 1%(w/v) agar. Incubation time: 10 - 30 minutes.
- Detection: colour.

h) EC 3.2.1.30. Hexosaminidase B (VAN SOMEREN and VAN HENEGOUWEN, 1973):
- Cellulose acetate gel electrophoresis (20 - 25 mA, 4 hours, room temperature).
- Buffer: 0.025 M citrate/citric acid, pH = 5.6.
- Reaction mixture: 0.5 mg/ml 4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside in 0.1 M citrate/citric acid buffer, pH = 4.5. Incubation time: 20 minutes, 35°C.
- Detection: fluorescence under long-wavelength (360 nm) UV light, after treatment of the gel with 1M carbonate/bicarbonate buffer, pH = 10.

2.4. DNA analysis of the hybrids.

2.4.1. Preparation and characterization of high molecular weight genomic DNA (modified from GROSS-BELLARD et al, 1973).

Two different procedures were used in this study: (i) cell pelleting prior to lysis (as in 2.3.1) and (ii) cell lysis in situ. Both procedures aim to dissociate DNA from chromosome and membrane proteins and to inhibit nuclease activity. This was achieved using a lysis buffer consisting of 10 mM tris/HCl, pH=8, 10 mM Na_2EDTA, 10 mM sodium chloride, 0.5%(w/v) sodium dodecylsulfate and 100 - 200 µg/ml proteinase K (fungal, 20 mAnson units/mg, BRL). Cells were incubated with lysis buffer overnight at 37°C with gentle shaking. Remaining protein was precipitated by adding one volume of freshly distilled phenol saturated with 10 mM Tris/HCl, pH = 8 10 mM Na_2EDTA, 10 mM sodium chloride. Phenol extraction (10 minutes of gentle mixing at room temperature) was repeated until interphase was clear. One tenth volume of 3 M sodium acetate, pH = 7.5, and two volumes of ice-cold absolute ethanol were added to the aqueous phase and DNA immediately spooled out with a closed Pasteur pipette. DNA was dissolved in 10 mM Tris/1 mM Na_2EDTA, buffer, pH = 8.0, and dialysed against water, for, at least, 24 hours, at 4°C. Alternatively, the aqueous phase resulting from phenol extraction was further extracted with one volume of chloroform/octan-2-ol (24+1), DNA precipitated twice with salt and ethanol and finally dissolved in buffer, at 4°C. Whenever necessary DNA preparations were freeze-dryed and stored at -20°C. DNA solutions of suitable concentration were stored at -20°C.

2.4. DNA analysis of the hybrids.

2.4.1. Preparation and characterization of high molecular weight genomic DNA (modified from GROSS-BELLARD et al, 1973).

Two different procedures were used in this study: (i) cell pelleting prior to lysis (as in 2.3.1) and (ii) cell lysis in situ. Both procedures aim to dissociate DNA from chromosome and membrane proteins and to inhibit nuclease activity. This was achieved using a lysis buffer consisting of 10 mM tris/HCl, pH=8, 10 mM Na_2EDTA, 10 mM sodium chloride, 0.5%(w/v) sodium dodecylsulfate and 100 - 200 µg/ml proteinase K (fungal, 20 mAnson units/mg, BRL). Cells were incubated with lysis buffer overnight at 37°C with gentle shaking. Remaining protein was precipitated by adding one volume of freshly distilled phenol saturated with 10 mM Tris/HCl, pH = 8 10 mM Na_2EDTA, 10 mM sodium chloride. Phenol extraction (10 minutes of gentle mixing at room temperature) was repeated until interphase was clear. One tenth volume of 3 M sodium acetate, pH = 7.5, and two volumes of ice-cold absolute ethanol were added to the aqueous phase and DNA immediately spooled out with a closed Pasteur pipette. DNA was dissolved in 10 mM Tris/1 mM Na_2EDTA, buffer, pH = 8.0, and dialysed against water, for, at least, 24 hours, at 4°C. Alternatively, the aqueous phase resulting from phenol extraction was further extracted with one volume of chloroform/octan-2-ol (24+1), DNA precipitated twice with salt and ethanol and finally dissolved in buffer, at 4°C. Whenever necessary DNA preparations were freeze-dryed and stored at -20°C. DNA solutions of suitable concentration were stored at -20°C.

The DNA preparations were characterized in terms of A^{260}, A^{260}/A^{280}, UV scan (320 - 230 nm) and electrophoretic behaviour on an agarose gel.

2.4.2. Preparation and characterization of plasmid DNA to be used as hybridization probe (modified from BIRNBOIM and DOLY, 1979).

The hybridization probe was a 3.8 kb PvuII fragment, isolated from the clone $\lambda H\alpha G2$ of a bacteriophage lambda bank of human fetal DNA (LAWN et al, 1978), including alpha 1-globin gene and its flanking sequences. This fragment was subcloned into pBR322/E.coli HB101 (SPANDIDOS, D and LANYON, G, unpublished results).
Bacteria, carrying the recombinant plasmid $\alpha P7(\alpha 1)$ (Fig 2.2), were grown in L-Broth (1%(w/v) tryptone, 0.5%(w/v) yeast extract, 1%(w/v) sodium chloride, 100 µg/ml L-ampicillin, pH 7.2 - 7.4) at 37°C, with vigorous shaking, until A^{650} fell within the range 0.7 - 1.0. The amplification of the plasmid copy number/cell was carried out by the addition to the exponentially growing culture of 170 µg/ml L-chloramphenicol and incubation overnight at 37°C with shaking. Bacteria were then killed with 0.2%(v/v) chloroform (10 minutes, 37°C, shaking). Cells were pelleted (7 000 g, 4°C, 21 minutes), resuspended in 0.05 M tris/HCl, ph = 8.0, and repelleted under the same conditions. The pellet was resuspended in lysis solution (50 mM glucose, 25 mM Tris/HCl, pH = 8.0, 10 mM Na_2EDTA, 5 mg/ml lysozyme) and incubated at 4°C for 30 minutes. 1%(w/v) sodium dodecylsulphate, 0.2 M sodium hydroxide solution was then mixed with the lysate (5 minutes, 4°C). Bacterial chromosome DNA was then precipitated by the addition of 3M potassium acetate, pH = 4.8 (15 minutes followed by centrifugation: 7 000 g, 6.5 minutes, 10 - 20°C). The

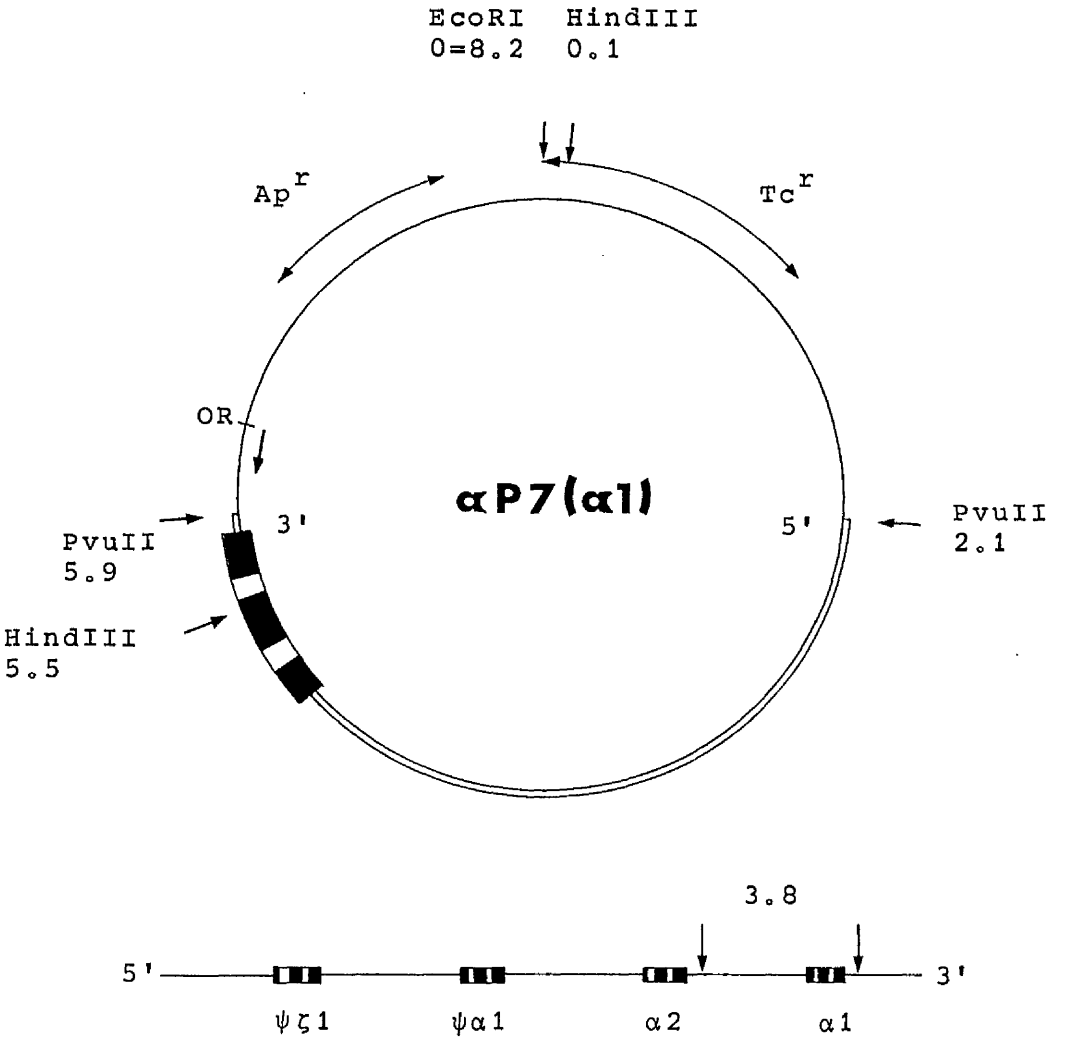


Fig 2.2. Map (in kb) of recombinant plasmid αP7(α1) used as hybridization probe in this study. The vector is pBR322. OR = origin and direction of replication. Ap^r = ampicillin resistance. Tc^r = tetracyclin resistance. The localization of the 3.8 kb Pvu II insert is shown in the map of α-globin gene cluster.

clear supernatant was recovered and added to 0.6 volumes of cold isopropanol, mixed and centrifuged immediately (8 000 g, 5 minutes, 10 - 20°C). The resulting pellet was recovered in 10 mM Tris/10 mM Na_2EDTA buffer, pH = 8.0, and, once dissolved, run in a caesium chloride gradient containing 1 mg/ml ethidium bromide (60 000 g, 18 hours, 20°C). The band corresponding to a higher density was detected with short-wavelength UV light, harvested through the tube wall, extracted with one volume of isopropanol (until aqueous phase was colourless) and dialysed against 10 mM tris/1 mM Na_2EDTA buffer, pH = 8.0 (4°C, overnight). The final DNA solution was stored at -20°C.
The plasmid DNA preparation was characterized in terms of A^{260}, electrophoretic behaviour on an agarose gel, and restriction pattern after treatment with PvuII and HindIII.

2.4.3. Molecular hybridization on solid phase (modified from SOUTHERN, 1975).

a) DNA restriction: 10 - 30 µg of genomic DNA of both parental and hybrid origin were restricted in the presence of the appropriate restriction endonuclease. A 5 to 10-fold excess of enzyme was used. The enzyme was allowed to act for 4 - 5 hours at 37°C. The buffers were those recommended by the manufacturer (Bethesda Research Laboratories, Cambridge). The enzymatic reaction was stopped by cooling the reaction mixture to 4°C. Phage lambda DNA, restricted with HindIII, was used as fragment size marker.

b) Separation of DNA restriction fragments according to size: Restricted DNA samples were loaded on a 0.8%(w/v) agarose gel. The electrophoresis buffer was 40 mM Tris/acetate, pH = 7.5, 20 mM sodium acetate, 1 mM

Na_2EDTA, and 1 µg/ml ethidium bromide. The electrophoreses ran overnight under an electric field of 1 V/cm.

c) Transfer and immobilisation of the separated DNA restriction fragments onto nitrocellulose filters (WAHL et al, 1979): After electrophoresis the DNA fragments were partly depurinated in situ (0.25 M hydrochloric acid, 30 minutes), denatured and hydrolysed at the apurinic sites (0.5 M sodium hydroxide/1.5 M sodium chloride, 30 minutes) and, finally, the gel was neutralized (0.5M Tris/HCl pH = 7.4, 1.5 M sodium chloride, 90 minutes). This treatment allowed the efficiency of transfer of the DNA fragments to be independent of their size. The transfer was carried out by allowing a steady flow of 20xSSC to pass through the gel and a nitrocellulose filter (Schleicher and Schull, Dassel, BA85, 0.45 µm pore size) on which the DNA fragments eluted from the gel became entrapped. The immobilisation of the DNA fragments was completed by baking the filter at 80°C for 4 hours.

d) Labelling of the hybridization probe by nick-translation (RIGBY et al, 1977): Approximately 0.5 µg of plasmid DNA was labelled to a specific activity of 1 - $2x10^8$cpm/ µg DNA. The reaction mixture consisted of the nick-tranlation buffer (50 mM Tris/HCl pH = 7.5, 5 mM magnesium chloride, 0.4 µg/ml 2-mercaptoethanol, 50 µg/ml bovine serum albumin), a pool of 'cold' nucleotides (dTTP, dGTP, dATP if the 'hot' nucleotide was ^{32}P-dCTP) at a concentration of 10 mM, 0.2 µg/ml pancreatic DNase I (3 136 units/mg; Miles Laboratories, Goodwood), 225 U/ml E.coli DNA polymerase I (RICHARDSON et al, 1964; BRL, Cambridge), and the relevant ^{32}P-labelled nucleotide (usually 50 µCi). The nick-translation reaction was allowed to proceed for 2 - 3 hours at 15°C and then stopped with 0.1 M Na_2EDTA.

The labelled DNA was separated from the unincorporated labelled nucleotides by column chromatography (Sephadex G50 equilibrated with 3xSSC which was also the elution buffer) using 0.2 mg/ml E.coli DNA as a carrier in the separation. The fractions of eluate were monitored with a Geiger counter for radioactivity: the leading peak of radioactivity consists of labelled nucleotides incorporated into DNA, while the trailing peak consists of unincorporated nucleotides (see Fig 3.4). The 'hottest' fractions were pooled and the actual radioactivity measured by liquid scintillation counting (using 40%(v/v) Lumax in xylene as the scintillator).

e) Molecular hybridization on solid phase (WAHL et al, 1979): Nitrocellulose filters with immobilised DNA fragments were treated with a prehybridization mixture ($42^{o}C$, for at least 4 hours with shaking) prior to the actual hybridization ($42^{o}C$, overnight with shaking). Prehybridization and hybridization were performed inside heat-sealed polythene bags in the presence of 5 - 10 ml of the relevant solution. 10 - $30x10^{6}$ cpm (approx 10^{5} cpm/cm^{2} filter) were used for each filter. The composition of the prehybridization and hybridization mixtures is described in Table 2.1. The final concentration of the probe was 20 ng/ml. The post-hybridization procedure consisted of a series of washes of increasing stringency (Table 2.2). After having been washed, the nitrocellulose filters were left to air dry enclosed in polythene bags and exposed on both sides to X-ray film, between intensifying screens, at $-70^{o}C$. One of the films was procesed after overnight exposure and the other one after a period of time considered to be long enough to produce a good signal/noise ratio.

Table 2.1. Composition of the prehybridization and hybridization mixtures.

Component	Concentration	
	Prehybridization	Hybridization
Formamide (freshly deionised)	50%(v/v)	50%(v/v)
Denhardt's solution *	5x	1x
SSC	5x	5x
Phosphate buffer, pH = 6.7	50 mM	20 mM
Salmon sperm DNA (sonicated and heat denatured)	380 µg/ml	100 µg/ml
Poly(A)	10 µg/ml	20 µg/ml
SDS	0.2%(w/v)	-
Dextran sulphate (sodium salt)	-	10%(w/v)
Labelled probe (heat denatured)	-	$1-2x10^7$ cpm

(*) Denhardt's solution: 0.2 mg/ml ficoll, 0.2 mg/ml bovine serum albumin, 0.2 mg/ml polyvinylpyrrolidone.

Table 2.2. Standard posthybridization procedure.

Wash	Composition	Time	Temperature
I	3xSSC 0.1%(w/v) SDS 50mM phosph buff pH=6.7 1xDenhardt's soln	2-4 changes for 5 minutes	room temp
II	3xSSC 0.1% SDS	1 change for 30 minutes	65°C
III	1xSSC 0.1% SDS	1 change for 30 minutes	65°C
IV	0.1xSSC 0.1% SDS	1 or 2 changes for 30 minutes or longer	65 - 70°C

3. Results.

3.1. Cytogenetic and biochemical characterization of man x mouse somatic cell hybrids.

Sixty four different clones from somatic cell hybrids (derived from four distinct human parental cell lines) were analysed for their human chromosome content. Fourteen of these appeared to be informative for the purpose of this study. This was confirmed by electrophoretic detection of enzymatic markers. The results obtained in both cytogenetic and biochemical analysis are summarized in Tables 3.1 to 3.4. They contain information relevant to the intrachromosomal mapping of the human alpha-globin gene cluster as well as the APRT, DIA4 and PGP genes.

As a measure of the heterogeneity, in terms of their human chromosome content, of the hybrid cell populations, the following parameters have been considered: (i) range of human chromosome number/cell counted, and (ii) its mean and standard deviation. The heterogeneity of two clones is graphically compared (MILLER, 1972a) in Fig 3.1. The instability of one particular clone (OAA9 5RII) is depicted in Fig 3.2.

Two clones of the OAA9 family proved to carry no cytogenetically detectable human genetic material. However, the fact that they had survived in the AA selective system indicated that some APRT activity might be expressed in these clones. It was, eventually, detected in both OAA9 2A (human form) and OAA9 4C (mouse form) clones. These findings may be interpreted, in the first case, as the result of a minor human-mouse chromosome rearrangement involving the human APRT gene and, in the second case, as a reversion of the

Table 3.1. Characterization of informative clones from somatic cell hybrids derived from DE human cells.

Hybrid ref[(a)]	Human chromosomes																					Heterogeneity[(b)]		Human biochemical markers					Not
	No of cells counted	2	(2/16)	(16/2)	16	1	3	4	5	6	7	8	9	10	11	12	X	13	14	15	17	18	19	20	21	22	mean±SD	range	APRT	DIA4	PGP	MDH1	IDH1	
DEA9 3t5	20	1	0	0	20			3		2	15	1			21	3	4				8	2	2	1	1		4.2±2.3	2 - 9						(c)
-- t6	20	0	0	0	18																								+	+	+	+	+	(d)
DEA9 3RIXt13(3)	20	5	0	0	0					5	17				17												2.8±1.2	1 - 5						
-- t13(7)	11	2	0	0	0																								-	-	-	+	+	
DEA9 5t6	21	0	17	0	0						20										2	4		8	1		3.3±1.7	1 - 7						
-- t7	8	0	8	0	0																								+	+	-	+	+	

Notes

(a) Hybrid notation The first group of four characters identifies human and mouse cells fused for the production of hybrids. The digit which follows indicates the order of isolation of primary clones. Roman numerals indicate the order of isolation of subclones. 'R' denotes that a counterselective process has been carried out. The stage of growth of the hybrid cell population is expressed by 't', the transfer (i. e., trypsinization) number. Transfer numbers between brackets show the number of passages since subcloning.

(b) The heterogeneity of hybrid cell populations was measured by the number of human chromosomes per cell counted in terms of (i) range, (ii) mean and (iii) standard deviation.

(c) Biochemical tests were only performed in the latest stage of hybrid population growth. At this stage chromosome preparations were made and cells tested for biochemical markers and α-globin gene.

(d) Details of human chromosome content not relevant for the purpose of this study are not reported at this stage of hybrid population growth.

Table 3.2. Characterization of informative clones from somatic cell hybrids derived from LO human cells.

Hybrid ref	Human chromosomes																								Human biochemical markers				Notes
	No of cells counted	9	(9/16)	(16/9)	16	1	2	3	4	5	6	7	8	10	11	12	X	13	14	15	17	18	19	20	21	22	Heterogeneity mean±SD	Heterogeneity range	APRT	DIA4	PGP	AK1	
LOA9 8It5(3)	20	1	0	18	0				10						19	18	8		5	2	12	10	10	14	10		8.0±2.6	3 - 13					
-- t5(6)	10	0	0	10	0																								+	+	+	+	

Table 3.3. Characterization of informative clones from somatic cell hybrids derived from OA human cells.

Hybrid ref	Human chromosomes																								Human biochemical markers				Notes
	No of cells counted	5	(5/16)	(16/5)	16	1	2	3	4	6	7	8	9	10	11	12	X	13	14	15	17	18	19	20	21	22	Heterogeneity mean±SD	Heterogeneity range	APRT	DIA4	PGP	HEXB	
OAA9 3Vt5(3)	20	11	0	20	0					18	1			1	15	1	15	2				7	3	7	4		5.4±2.0	2 - 10					
-- t5(6)	11	5	0	10	0																								+	+	-	-	
OAA9 5t4	21	0	0	0	21				1	8	20		4				8				17	5	6	7			5.3±1.9	2 - 9					
-- t5	10	0	0	0	10																								+	+	+	-	
OAA9 5RIIt7(3)	31	0	0	0	0		11		12	12											18	5		11			2.4±2.2	0 - 9					
-- t7(7)	22	0	0	0	0																								-	-	-	+	
OAA9 5RIXt7(3)	20	0	0	0	0					5																							
-- t7(7)	30	0	0	0	0																								-	-	-	-	
-- t7(11)	10	0	0	0	0																								-	-	-	+	

Table 3.4. Characterization of informative clones from somatic cell hybrids derived from SY human cells.

Hybrid ref	Human chromosomes																					Heterogeneity		Human biochemical markers				Notes
	No of cells counted	16	(16/17)	(17/16)	17	1	2	3	4	5	6	7	8	9	10	11	12	X	13	14	15	18	19	20	21	22	mean±SD	range	APRT	DIA4	PGP	GALK	
SYA9 1IVt5(3)	30	26	0	1	23				4		19					22	7	4				4	1		1		4.0±1.4	2 - 7					
-- t5(6)	20	19	0	0	13																								+	+	+	+	
SYA9 1IVRt5(11)	20	0	0	0	16				1		12			1		13	16	6		5		4	16	8			5.6±2.2	3 - 10					
-- t5(14)	11	0	0	0	6																								-	-	-	+	
SYA9 1Xt5(3)	21	19	0	0	1				2	6	2	9			1	19	16	3		1	1		1	2	1		4.4±1.5	2 - 7					
-- t5(6)	12	11	0	0	1																								+	+	+	-	
SYA9 1XRt5(8)	20	0	0	0	0				1		5	16				15	15				2	13		5	1		4.4±2.0	1 - 8					
-- t5(11)	10	0	0	0	0																								-	-	-	-	
SYA9 2At4	20	0	0	17	0											2	7					1		3			1.7±1.3	0 - 6					
-- t8	10	0	0	10	0																								+	+	-	-	
SYA9 8t4	21	0	0	19	14						1	10				12	7	2				10	2		3		3.9±1.7	1 - 7					
-÷ t5	10	0	0	9	9																								+	+	-	+	

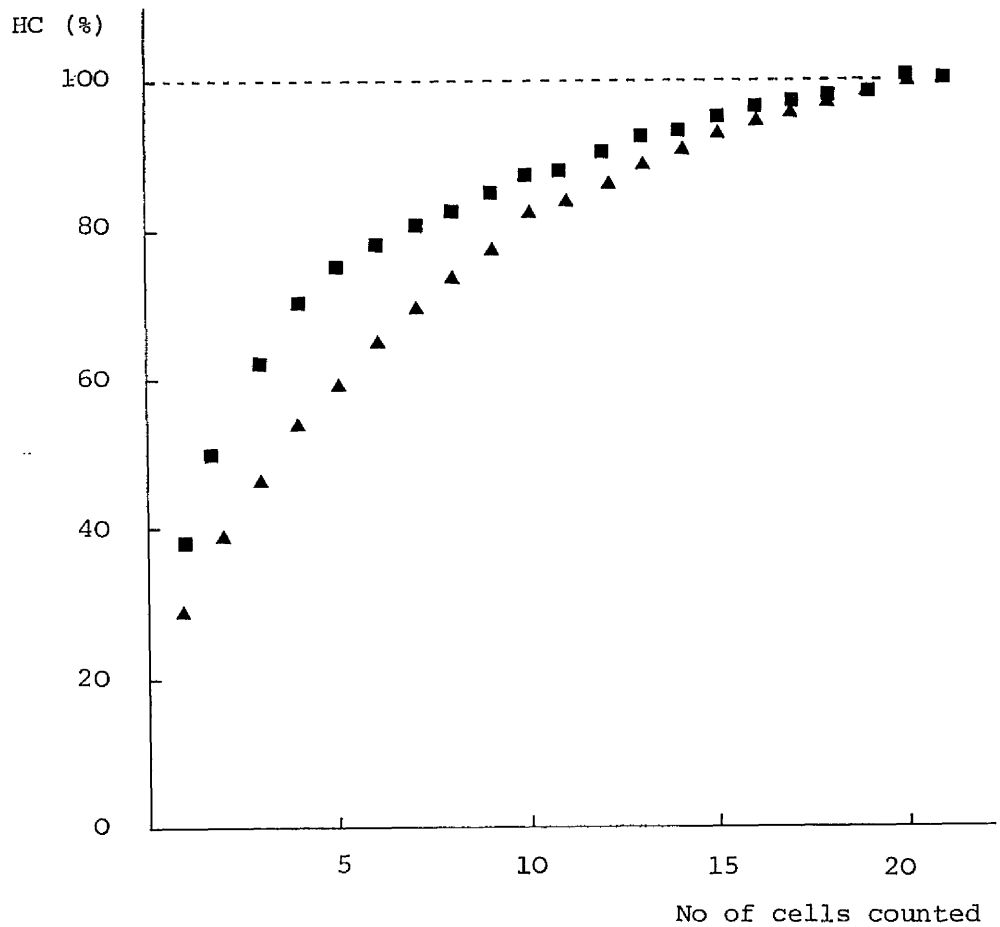


Fig 3.1. Comparison of the heterogeneity of two hybrid cell populations, at similar stages of growth, following a modification of the method proposed by MILLER (1973). DEA9 3t5 (▲) appears to be more heterogeneous than SYA9 8t4 (■) and these findings are in agreement with other measures of heterogeneity (see Tables 3.1 and 3.4). The heterogeneity curves were drawn considering the cells in 20 (DEA9 3t5) and 21 (SYA9 8t4) different orders. HC (%) = (mean cumulative number of different human chromosomes / total number of different human chromosomes) x 100. The horizontal line at HC = 100% represents complete homogeneity.

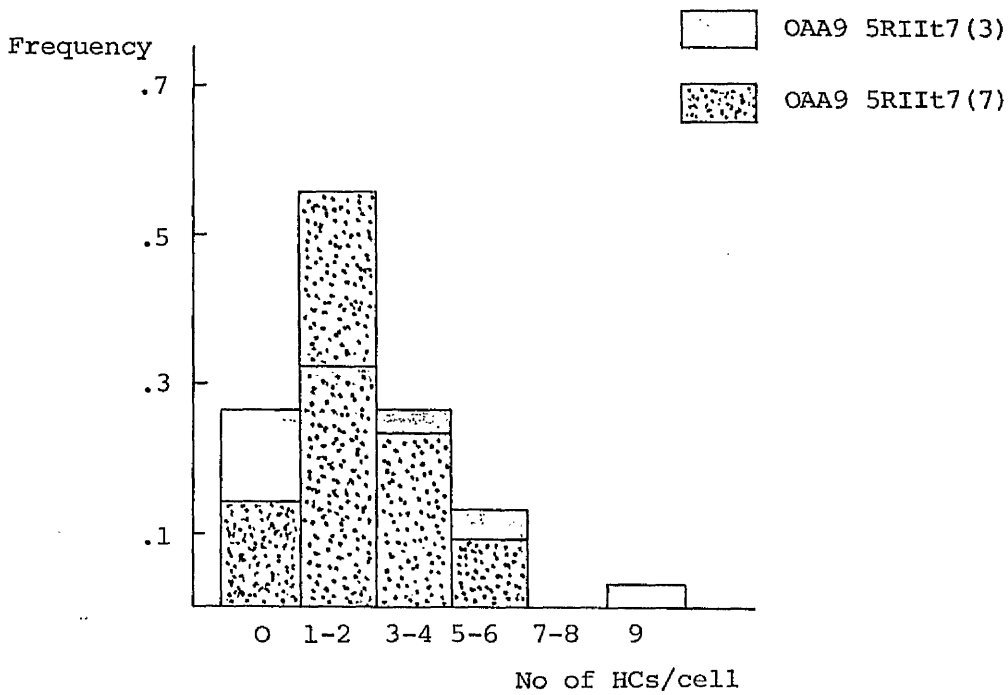


Fig 3.2. Instability of OAA9 5RII hybrid cell population. The range narrows from 0 - 9 in t7(3) to 0 - 6 in t7(7). The mean$\pm$SD also varies: 3.1$\pm$3.9 in t7(3) as opposed to 2.1$\pm$1.6, respectively. Cells were grown in the presence of 0.1 mM DAP.

mutated mouse APRT$^-$ gene (TISCHFIELD et al, 1973; TISCHFIELD and RUDDLE, 1974; CHU and POWELL, 1976).

All the cell lines used in this study were *Mycoplasma* free when tested by the method of CHEN (1977).

3.2. Detection of human alpha-globin gene cluster in DNA from somatic cell hybrids.

Total genomic DNAs from five human fibroblast lines (four parental lines and one normal control), one mouse fibroblast line (parental) and fifteen hybrid cell lines were extracted and characterized (Table 3.5).

Plasmid DNA to be used as an hybridization probe was cut with two restriction endonucleases (Pvu II and Hind III) to identify the human DNA fragment within recombinant plasmid αP7(α1). The result, shown in Fig 3.3 (lanes 3 and 5), demonstrates the presence of the alpha 1-globin region, and its flanking regions, in the vector pBR322. The electrophoretic behaviour of the non-digested plasmid (lane 2) shows that the preparation consists largely of the supercoiled form (apparent size: 4 kb approx) in the presence of trace amounts of the open circular and linear forms. The radioisotope labelling of this plasmid DNA by nick-translation led to hybridization probes with a specific activity ranging from $5x10^7$ to $2.5x10^8$ cpm/ μg DNA, with a reaction yield of 30 - 50%. A typical elution profile of the nick-tranlation reaction mixture on a column of Sephadex G50 (3xSSC as the eluent) is shown in fig 3.4.

The molecular hybridization on solid phase (Southern blot) presented several problems resulting from (i) genomic DNA degradation and (ii) non-specific hybridization due to the presence in the probe of one Alu sequence. The first of these problems was solved by extracting DNA under milder conditions and by using a restriction endonuclease (PvuII) which gives rise to smaller fragments (than, say, Eco RI) in the human alpha-globin gene region: 3.8 kb, 2.1 kb and 2.3 kb corresponding to alpha 1, alpha 2 and pseudo-alpha 1 /

Table 3.5. Genomic DNA preparations.

Cell line	No of cells ($x10^{-7}$)	Dilution (a)	A^{260}	A^{260}/A^{280}	Amount of DNA ($x10^{-2}\mu g$)
DE (b)	3	1:100	.077	1.8	7.6
OA (b)		1:100	.132	1.9	13
SY (b)		1:100	.103	1.6	10
SY (c)	.3	1:1	.840	1.7	2.5
821589 (b,d)	6	1:100	.066	2.0	6.6
A9 (b)		1:100	.238	1.7	20
DEA9 3 (b)	5	1:100	.316	1.8	32
DEA9 3RIX (b)	7	1:100	.257	1.8	26
DEA9 5 (b)	5	1:100	.138	1.8	14
LOA9 8I (c)	1	1:10	.244	1.5	15
OAA9 3V (b)	5	1:100	.231	1.8	23
OAA9 5 (b)	5	1:100	.110	1.6	11
OAA9 5RII (c)	6	1:200	.124	1.6	25
OAA9 5RIX (b)	10	1:100	.315	1.8	32
OAA9 5RIX (c)	6	1:200	.096	1.7	18
SYA9 1IV (b)	5	1:100	.089	2.0	8.8
SYA9 1IVR (b)	6	1:100	.573	1.8	58
SYA9 1X (b)	7	1:100	.230	1.9	24
SYA9 1XR (b)	10	1:100	.338	1.8	34
SYA9 2A (b)	5	1:100	.201	2.0	20
SYA9 8 (b)	5	1:100	.205	1.7	20

Notes: (a) Dilution of the preparations when A^{260} and A^{280} were determined. (b) Cells pelleted prior to lysis. (c) Cells lysed *in situ*. (d) Normal human control.

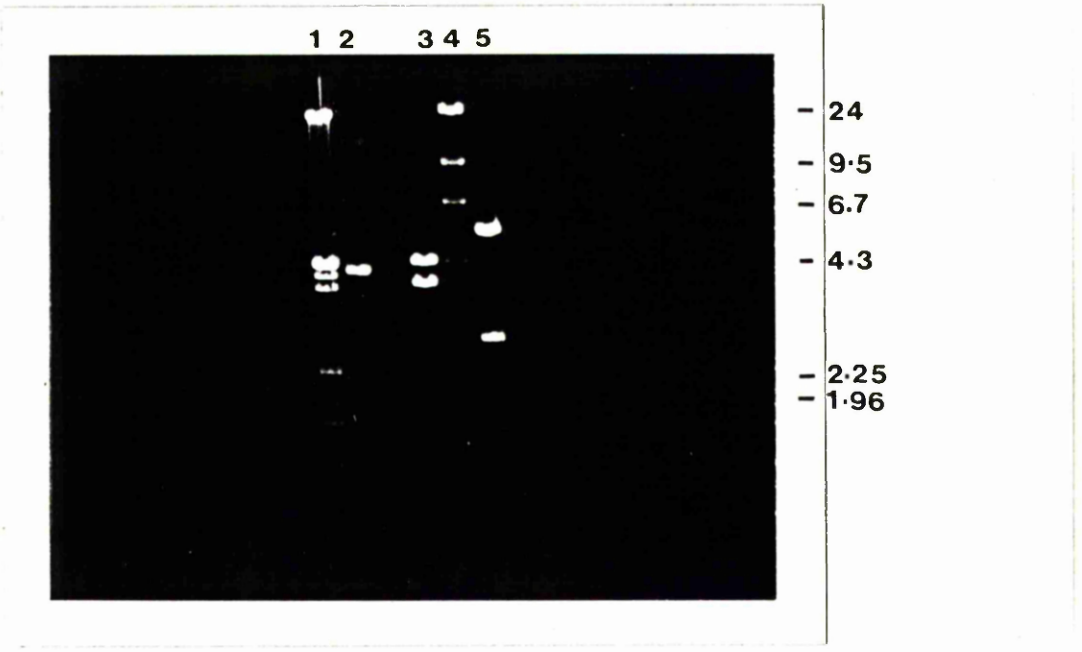


Fig 3.3. Characterization of recombinant plasmid αP7(α1) on a 0.8% (w/v) agarose gel. Lanes 1 and 4 contain the products of digestion of 1.5 µg of phage λ DNA with Pvu II and Hind III, respectively. Lane 2 contains 0.5 µg of non-digested recombinant plasmid DNA. Lanes 3 and 5 contain the products of digestion of 0.5 µg of recombinant plasmid DNA with Pvu II and Hind III, respectively. The samples were incubated at 37 °C for 2.5 hours and ran on a horizontal gel, overnight, under an electric field of 1 V/cm. For interpretation see Fig 1.8.

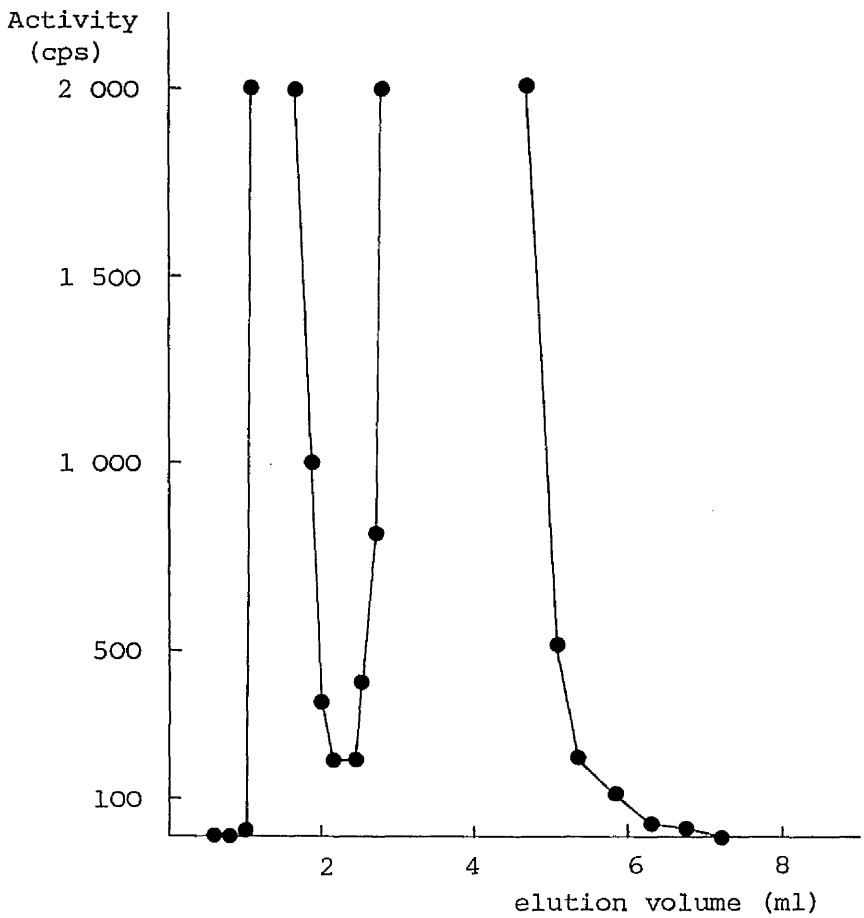


Fig 3.4. Typical elution profile of Sephadex-G50 column chromatography of nick-tranlation reaction mixture. 0.5 µg of recombinant plasmid αP7(α1) DNA was labelled in the presence of (^{32}P)dCTP (50 µCi for 2.5 h, at 15 ^{o}C).

alpha 2 intergenic regions, respectively (see Fig 1.8). The second problem was partially overcome when the post-hybridization washing of the nitrocellulose filter was performed with 0.1xSSC, at 70 °C, for 2 hours. However, most of the blots still present a fairly intense background (Fig 3.5.A and B). The human alpha-globin gene cluster was detected in all tested DNAs (DE, OA, SY, 821589).Mouse alpha-globin gene could not be detected with this human genomic probe even at low stringency conditions. The correlation between human chromosome content and human alpha-globin gene in a panel of somatic cell hybrids is summarized in Table 3.6.

The presence of the human alpha-globin gene cluster in the DNA from the hybrid clones carrying an intact HC 16 (DEA9 3, OAA9 5, SYA9 1IV and SYA9 1X) and its absence in the revertant clones which lost HC 16 (DEA9 3RIX, OAA9 5RII and SYA9 1IVR) confirm the assignment of this gene cluster to HC 16. The positive result in LOA9 8I (carrying t(16:9)) narrows the assignment to 16q22→pter. The negative result in DEA9 5 (carrying t(2:16)) excludes the loci from 16q11→qter. Similarly, the negative results observed in OAA9 3V (carrying t(16:5)), and SYA9 2A and SYA9 8 (both carrying t(17:16)) exclude the loci from 16p11→qter and 16q21→qter, respectively. In conclusion, the data suggest that the human alpha-globin gene cluster maps to HC 16p.

In section 3.4, the statistical analysis of the data will be reported and the discordant results observed in two 'revertant' clones (OAA9 5RIX and SYA9 1XR) discussed.

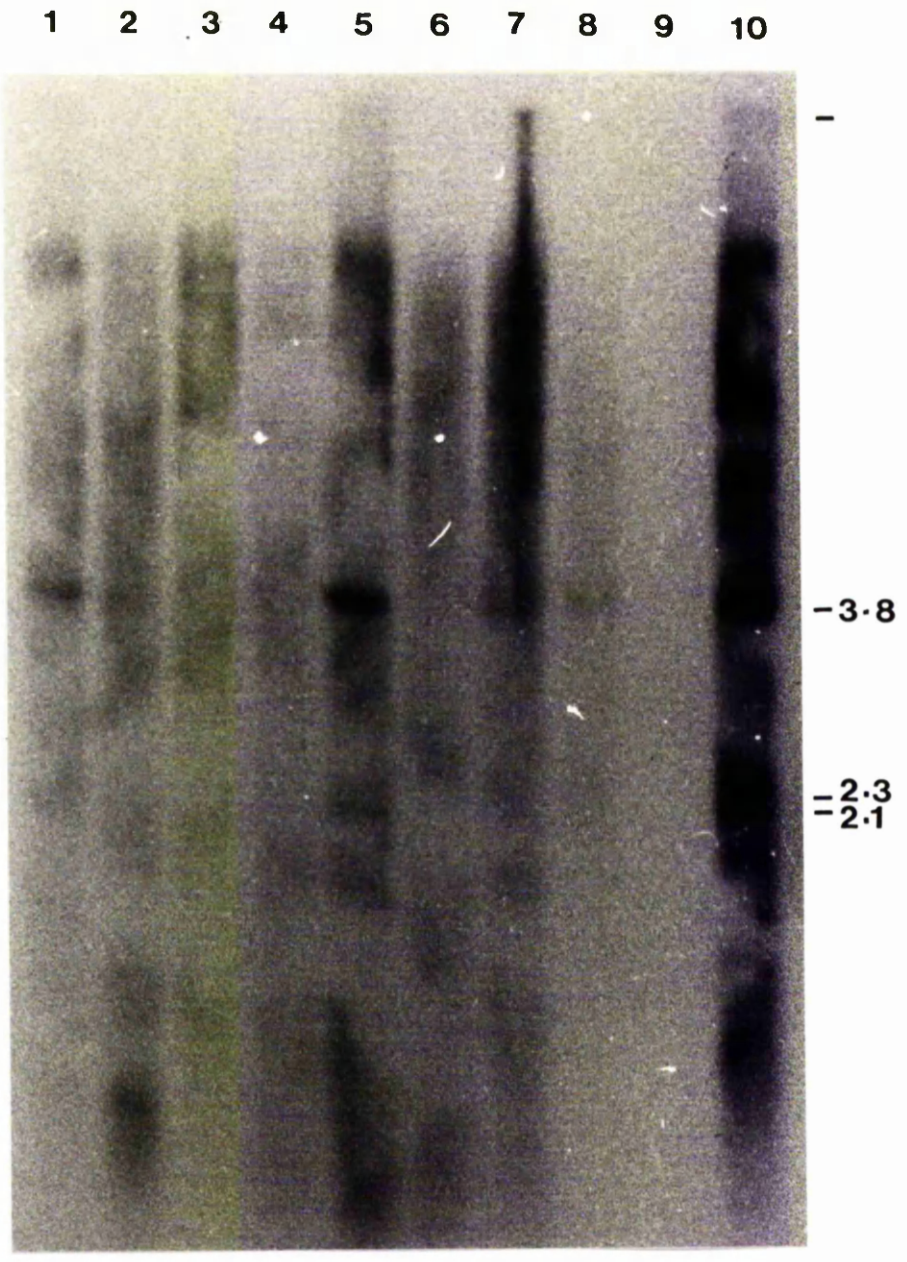


Fig 3.5.A. Pvu II digestion patterns of genomic DNA from human fibroblasts (DE, lane 1; OA, lane 5; 821589, lane 10), mouse L A9 cells (lane 9), hybrid clones with HC 16p (DEA9 3, lane 2; OAA9 5, lane 7) and hybrid clones without HC 16p (DEA9 3RIX, lane 3; DEA9 5, lane 4; OAA9 3V, lane 6; OAA9 5RIX, lane 8). Size scale is shown in kilobase pair. Exposure time: 6 days.

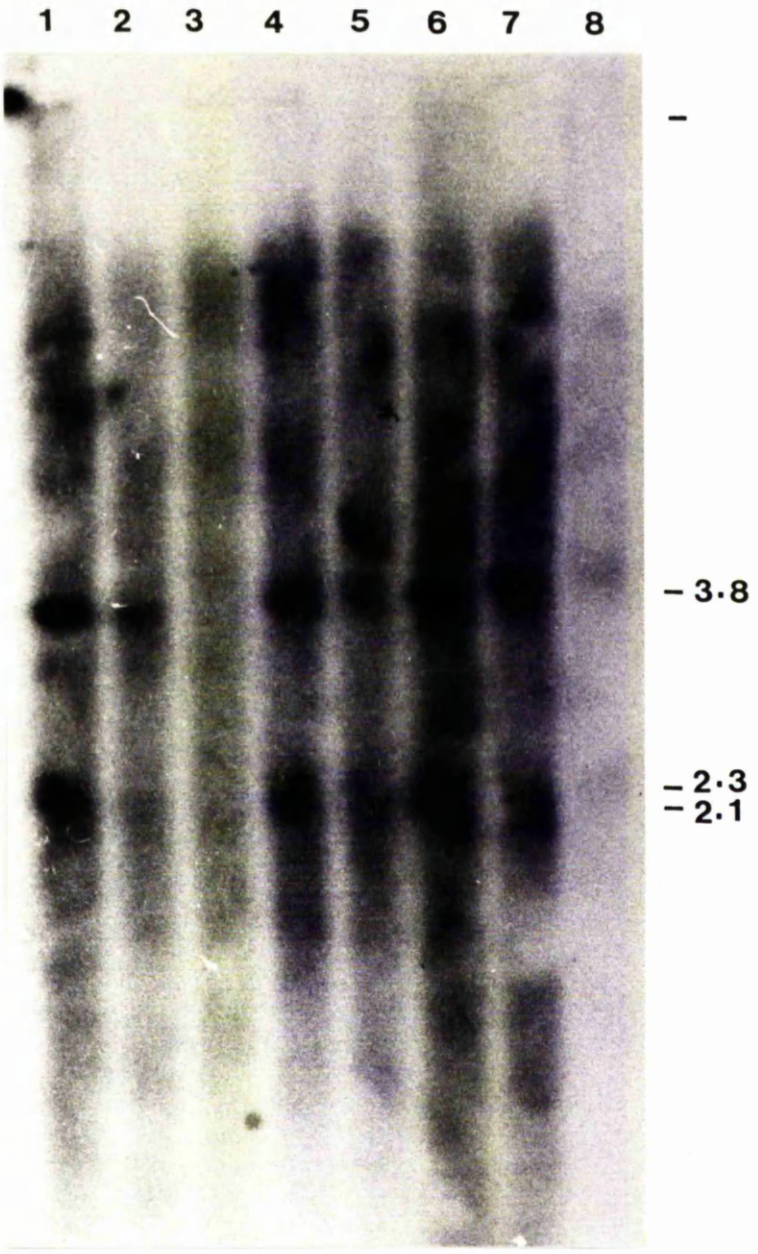


Fig 3.5.B. Pvu II digestion patterns of genomic DNA from human fibroblasts (SY, lane 1; OA, lane 6), hybrid clones with HC 16p (SYA9 1IV, lane 2; SYA9 1X, lane 4; OAA9 5, lane 7) and hybrid clones without HC 16p (SYA9 1IVR, lane 3; SYA9 1XR, lane 5; OAA9 5RIX, lane 8). Size scale is shown in kilobase pair. Exposure time: 7 days.

Table 3.6. Screening of DNA from somatic cell hybrids for human alpha-globin gene cluster.*

Hybrid ref	Hum Chrom 16q	16p	α-globin gene cluster	Comments
DEA9 3	+	+	+	
DEA9 3RIX	-	-	-	
DEA9 5	+	-	-	16q11 → qter
LOA9 8I	+	+	+	16qcen → q22
OAA9 3V	+	-	-	16p11 → qter
OAA9 5	+	+	+	
OAA9 5RII	-	-	-	
OAA9 5RIX	-	-	+	cells pelleted prior to lysis
OAA9 5RIX	-	-	+	cells lysed <u>in situ</u>
SYA9 1IV	+	+	+	
SYA9 1IVR	-	-	-	
SYA9 1X	+	+	+	
SYA9 1XR	-	-	+	
SYA9 2A	+	-	-	16q21 → qter
SYA9 8	+	-	-	16q21 → qter

(*) DNAs from all hybrid clones were restricted with PvuII. In addition DNAs from OAA9 5RIX and SYA9 1XR were restricted with Hind III as well.

3.3. Comments on the methods.

3.3.1. Production of man x mouse somatic cell hybrids.

Seven cell fusion experiments were performed yielding thirty one primary hybrid clones: 4.4 clones/fusion (range: 1 - 10). The colonies were picked up, on average, 5.4 weeks after the fusion (range: 2.8 - 9.7). Thus, the AA selection appears to be significantly less efficient than HAT in selecting hybrid clones (higher cell toxicity? lower rate of reversion?). An yield of 21 clones/fusion and a mean selection time of 4.3 weeks are typical figures for a HAT selection experiment.

After the production of the hybrid cell populations used in this study was completed, KLEBE and MANCUSO (1982) reported that many compounds (e.g., acetamide, potassium bromide, urea, glycerol, ...) which induce erythroleukaemia cell differentiation also promote PEG-mediated cell membrane fusion. It was therefore not possible to incorporate this refinement of the method in the cell fusion protocol.
Counterselection of hybrid cell populations, combining DAP selection and subcloning, resulted in the cytogenetic and biochemical loss of HC16 (see Tables 3.1 to 3.4) in five hybrid clones.

3.3.2. Characterization of the cellular phenotype of the somatic cell hybrids.

Trypsin/Leishman staining of chromosome preparations provided good evidence, in the overwhelming majority of the cases, of the human chromosome content of the hybrid cell populations, since the size, shape and banding pattern of both human and mouse chromosomes are

distinctly specific. Nevertheless, the fluorescent differential staining method made possible the unequivocal identification of the parental origin of the chromosomes whenever the former method had been inconclusive. In a small number of cases, the differential staining pointed out the presence of previously non-detected human chromosomes.

Due to the heterogeneity of the cell populations used in this study (see 3.1), there is a finite probability of missing any human chromosome, present in a fraction of the whole population, in the cytogenetic analysis. If the frequency of a given chromosome in the population is q, the probability P of missing it after having analysed n cells is $P = (1 - q)^n$. This is the probability of false negatives for this particular chromosome. On the other hand, a negative result in the biochemical analysis of the hybrids does not necessarily mean that the gene which codes for the tested enzyme is absent. A false negative enzymatic result can occur when there is (i) a low level of gene expression; (ii) a null allele; (iii) a high turnover in vivo; and/or (iv) a high degradation rate in vitro. These are putative causes for the disagreement between cytogenetic and biochemical results which can be observed in some hybrid clones, e.g., OAA9 3V and SYA9 2A (see Tables 3.1 to 3.4). False positive results can be explained mainly in terms of lack of specificity of the test. Hence the importance of employing well established, standardized methods.

3.3.3. Detection of human alpha-globin gene cluster in the DNA from somatic cell hybrids.

The yield of genomic DNA extraction proved to be extremely variable, ranging from 11 to 150 $\mu g/10^6$ cells. The two methods used in this study (cells pelleted and

frozen before lysis; cells lysed *in situ*) gave markedly different yields: on average, 37.7 $\mu g/10^6$ cells and 76.2 $\mu g/10^6$ cells, respectively. The quality of the preparations also displayed noticeable differences. When the cells were pelleted and frozen before lysis, the protein contamination was lower (on average, A^{260}/A^{280} = 1.83) than in the case of lysing the cells *in situ* (on average, A^{260}/A^{280} = 1.62). On the contrary, the latter method produced less RNA-contaminated and higher molecular weight preparations.

Interspecific DNA cross-hybridization could not be detected in any blotting experiment. Mouse restricted DNA did not hybridize stably with the human DNA probe at the level of stringency used in the posthybridization washing. This may be due to the fact that the intervening and flanking sequences of alpha 1-globin gene, present in the human genomic probe, accumulate mutations at a higher rate than the coding sequences (JEFFREYS, 1982). Therefore the man/mouse sequence divergence, in the region of the alpha-globin gene cluster, may account for the lack of interspecific cross-hybridization. However, when a cDNA probe is used, the mouse alpha-globin gene cluster can be detected in Southern blots (KOEFFLER et al, 1981).

3.4. Results versus aims. Conclusions.

Data shown in table 3.6 can be summarized in a 2 x 2 contingency table (Table 3.7.A). As some of the numbers in this table are very small (< 5), the chi-square test cannot be used, even after introducing Yates'correction (BAILEY, 1959, p 61). In such a situation an exact test of significance must be employed. The rationale is to take the Null Hypothesis (no association whatever between HC16p and human alpha-globin gene cluster) and then consider the experimental results that might be observed. The actual result is judged significant if it belongs to a sufficiently extreme (in probability terms) class of events.

The whole series of possible results, for marginal totals the same as observed, are represented in Table 3.7.B and the probability of each possible result is indicated below each contingency table. A decision as to significance rests on comparing the probability of occurrence of the observed table with the chosen level of significance (5% in this case). Thus, the actually observed result, whose probability of occurrence is 1.04%, is significant at the 5% level. This means that the Null Hypothesis is to be rejected and that the level of confidence in a positive association between HC16p and human alpha-globin gene cluster is 95%.

The deviant results (negative for the short arm of chromosome 16 but positive for the alpha-globin gene cluster) occurred in the analysis of 'revertant' hybrid clones: OAA9 5RIX and SYA9 1XR. A possible explanation for these findings is that both cytogenetic and biochemical characterization of these clones failed do detect the presence of HC16p (possibly present in a small fraction of the hybrid cell population) while the

Table 3.7. Data on the screening of DNA from somatic cell hybrids for the human alpha-globin gene cluster.

A.

		Human alpha-globin gene cluster		Total
		+	-	
HC16p	+	5	0	5
	-	2	7	9
Total		7	7	14

B.

5 | 0 2 | 7	4 | 1 3 | 6	3 | 2 4 | 5	2 | 3 5 | 4	1 | 4 6 | 3	0 | 5 7 | 2
1.04%*	12.2%	36.7%	36.7%	12.2%	1.04%

(*) The probability of occurrence of a contingency table $\begin{array}{c|c} a & b \\ \hline c & d \end{array}$ is given by

$$P = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

where $n = a+b+c+d$.
Or, using logarithms,

$$\log P = \log(a+b)!+\log(c+d)!+\log(a+c)!+\log(b+d)!-$$
$$-(\log n!+\log a!+\log b!+\log c!+\log d!)$$

DNA test succeeded to detect human alpha-globin gene cluster. Thirty and twenty cells from these clones were counted, respectively. The corresponding probability of missing a chromosome, present in 5% of the cells, is 21% and 36% respectively which is in sharp contrast with 4.2% and 12% if the chromosome is present in, say, 10% of the cells (see 3.3.2). If this interpretation is correct, one may suggest that the elimination of HC16p, by means of DAP selection and subcloning, was not complete and that, under these conditions, a residual fraction of the cell population retained HC16p. Provided the counterselective conditions are kept for longer and a larger number of cells (approx 60) is counted, it seems possible to achieve and monitor complete segregation of HC16p. It should be pointed out that the sensitivity of the different methods involved in the characterization of the hybrid populations must be carefully assessed and balanced.

In conclusion, it may be said (i) that, at a level of significance of 5%, the human alpha-globin gene cluster maps on the short arm of human chromosome 16 and (ii) the experimental approach combining somatic cell hybridization and DNA hybridization on solid phase to a genomic probe is feasible and useful for the mapping of single copy genes.

4. Discussion.

4.1. The methods put into perspective.

The evidence reported in this thesis (see section 3) demonstrates that the experimental approach combining somatic cell hybridization and DNA hybridization on solid phase, with a human genomic DNA probe, can be successfully applied to the intrachromosomal mapping of human genes.

The human alpha-globin gene cluster was chosen as the model system to assess the feasability and the usefulness of the above experimental approach in the localization of unique gene sequences. The choice of the model was made on several grounds: (i) The alpha-globin gene cluster had been consistently mapped to the short arm of HC 16 (Table 1.8). (ii) A series of human cell lines, carrying different balanced translocations involving HC 16 were available (iii) The availability of selective systems which select either for (AA selection) or against (DAP selection) HC 16q. (iv) A recombinant plasmid had been cloned which contains a human genomic DNA insert, including the alpha 1-globin gene and its intervening and flanking sequences (SPANDIDOS, D and LANYON, G, unpublished results).

A number of different experimental approaches have been employed in previous attempts to map the human alpha-globin gene cluster (Table 1.8). Historically, the first method to be used successfully, involved a DNA/cDNA annealing assay (Cot analysis) performed on total genomic DNA extracted from a panel of clones from somatic cell hybrids (DEISSEROTH et al, 1977). The Authors were able, using this method, to assign the human alpha-globin structural gene to HC 16. This was

an important advance, since it made possible the mapping of genes coding for specialized gene products, therefore only expressed in differentiated cells. As, at that time, molecular hybridization *in situ* had not yet been shown to map single copy genes, this species-specific molecular hybridization assay also represented a marked increase in specificity and sensitivity in the detection of unique DNA sequences. It success depends on a number of conditions not always easily met: It uses a relatively large amount of DNA (1 - 2 mg/assay). The purity of the cDNA probe (prepared by reverse transcription of alpha-globin-enriched mRNA) is very important to avoid cross-hybridization with beta-globin sequences. The molecular hybridization must be performed at a temperature high enough to prevent the formation of man/mouse interspecific molecular hybrids. The sensitivity of the DNA/cDNA assay must be similar to the sensitivity of the methods used for characterization of the human chromosomes in the hybrid clones.

Four years later, a new generation of methods have been applied to confirm and refine the previous localization of the human alpha-globin gene cluster: molecular hybridization *in situ* (BARG et al, 1981; GERHARD et al, 1981), trisomy mapping (WAINSCOAT et al, 1981) and molecular hybridization on solid phase (BARG et al, 1981; KOEFFLER et al, 1981). All these new contributions not only unanimously confirmed the assignment of human alpha-globin gene to HC 16, they also led to the definition of a shortest region of overlap covering the region within 16p12 and 16pter.

The success of molecular hybridization *in situ* in the detection of single copy genes has been the result of improving the sensitivity of the method. This was accomplished either by using recombinant cDNA probes

that form networks at the site of hybridization in the presence of dextran sulphate (GERHARD et al, 1981) or by using hybridization probes made from genomic recombinants that hybridize to adjacent and intervening sequences as well as to the structural gene (MALCOLM et al, 1981b).

WAINSCOAT et al (1981) reported the detection of imbalance of globin chain synthesis (alpha/non-alpha =1.6) in the red cells of an infant with partial trisomy of HC 16p (p12→pter). These workers interpreted their findings as a gene dosage effect. In fact, trisomy for the chromosome segment containing the alpha-globin genes would, assuming full gene expression, lead to an excess of alpha- to non-alpha-globin chain production in the ratio of 3:2, close to the observed ratio of 1.6. The possibility of this observation being related to developmental abnormalities (gene regulatory rather then gene dosage effect) could only be ruled out by direct 'titration' of the alpha-globin genes using a molecular hybridization assay.

Finally, the third new experimental approach combines the production of somatic cell hybrids with molecular techniques: restriction endonuclease digestion, DNA/cDNA hybridization on solid phase and cloning of recombinant DNA containing specific human gene sequences (KOEFFLER et al, 1981). This method closely parallels the type of approach used in this study. The main difference is that KOEFFLER et al have used as hybridization probe a cDNA sequence obtained by reverse transcription of a fraction of human reticulocyte RNA, while the probe used in this study is a subclone of a human genomic DNA bank. The different nature of the probes may account for the absence of man/mouse interspecific cross-hybridization found in this study, in contrast with the clear, yet

somewhat weaker, mouse band reported by KOEFFLER et al. The alpha-globin gene coding sequences of man and mouse share approximately 90% homology which is more than enough to produce interspecific stable cross-hybridization if a cDNA probe is used. On the other hand, the mam/mouse sequence divergence in the intervening and flanking regions of the alpha-globin gene cluster may be sufficiently large to prevent significant cross-hybridization to a genomic DNA probe.

The agreement of the results obtained using these four different experimental approaches suggests that, in spite of their peculiar advantages and/or limitations, they all can give a valuable contribution to the construction of the human gene map. It is, therefore, difficult to conclude in terms of what is the best way of assigning cloned genes to chromosome regions. Considerations of cost, time consumption, ease in performance, resolving power, applicabilty, etc, would clearly lead to ambivalent answers. As SHOWS et al (1982) point out 'the enormous complexity of the human genome will be tamed and unraveled by a combination of several experimental approaches'.

4.2. A side line - intrachromosomal localization of APRT, DIA4 and PGP.

This study provides further evidence for the regional mapping of APRT, DIA4 and PGP - three biochemical markers on chromosome 16 (Tables 3.1 to 3.4). Combining the results from cytogenetic and isozyme analysis performed on a panel of fourteen informative clones from somatic cell hybrids, it was possible to assign both APRT and DIA4 to region 16q12→q22 and PGP to 16p. These assignments confirm previously reported results (BARG et al, 1981 on APRT and DIA4; KOEFFLER et al, 1981 on PGP).

However, agreement is still to be found concerning the regional localization of APRT (16q22.2→qter according to RÉTHORÉ et al, 1982, based on the detection of a gene dosage effect in one case or partial trisomy of HC 16q) and the relative position of both APRT and DIA4 loci (DIA4 distal to APRT according to POVEY et al, 1980; DIA4 proximal to APRT according to POVEY et al, 1981 and JEREMIAH et al, 1982). The data presented here provides further evidence for the first of these open questions. Nevertheless, a definite assignment depends on a better definition of the breakpoint at band 16q22. As DIA4 and APRT cosegregated in all tested clones, nothing can be said regarding the order of the two loci on HC16q.

4.3. Prospects for further work.

4.3.1. Setting up a bank of somatic cell hybrids (SHOWS et al, 1982).

Somatic cell hybrids have proved to be, during the last fifteen years or so, the most powerful tool to assign genes to chromosomes and/or to chromosome regions (see 1.1.2). It is now becoming obvious that the applications of somatic cell hybridization are extending to a wide range of new fields, namely in the screening of chromosome-specific DNA banks (originated from sorted chromosomes), in the intrachromosomal mapping of cloned anonymous DNA sequences (potentially useful to probe RFLPs linked to 'disease' loci), in the genetic dissection of disease, in the elucidation of the regulation of gene expression, in the production of monoclonal antibodies etc. It is therefore outstandingly important to set up a bank of clones from somatic cell hybrids well characterized in terms of their human chromosome content.

Such a bank should include chromosome assignment as well as regional assignment panels of somatic cell hybrids. The chromosome assigment panels consist of a number of hybrid clones, each of which has a unique human chromosome content, in order to allow unambiguous correlation between a single chromosome and the marker being mapped. Regional assignment panels are constructed from human cells carrying chromosome rearrangements (translocations or deletions detected and characterized in the diagnostic cytogenetic practice). The ultimate goal is to set up a series of twenty four panels, each of which consists of a series of hybrid clones carrying different regions of a particular human chromosome. Such panels would contribute to the linear ordering of genes

on chromosomes, as was the case in this study.

Ultimately, it would be extremely useful to generate a panel of single chromosome hybrids. This can be achieved in a number of different ways: (i) subcloning cell hybrids containing few human chromosomes; (ii) subcloning and counterselecting hybrids with selectable markers; and (iii) producing microcells with few human chromosomes (EGE and RINGERTZ, 1974) for fusion with rodent cells. Single chromosome hybrids represent a much easier way of achieving fractionation of the human genome than, for instance, fluorescence-activated chromosome sorting (CARRANO et al, 1979).

4.3.2. Molecular characterization and prenatal diagnosis of haemoglobinopathies.

As it has been suggested in the Introduction (see 1.1.3 and 1.4), the natural development of the work reported here would be the application of Molecular Genetics information to clinical practice. In fact, some of the methodologies utilized in this study (DNA extraction, use of restriction endonucleases, Southern blotting) as well as a specific strategy for direct gene analysis (RFLP mapping) can be transferred to the diagnostic laboratory concerned with genetic disease.

In close connection with the subject of this study, and particularly badly characterized in my country (Portugal), the haemoglobinopathies (largely beta-thalassaemia) appear as the obvious group of diseases to start with. Research on this field should develop along two different, yet complementary, lines:

(i) Characterization, at the molecular level, in the population, of the portuguese forms of thalassaemia and

establishment of a correlation between globin gene RFLP haplotypes and thalassaemia syndromes (ORKIN et al, 1982).

(ii) Implementation of the prenatal diagnosis of beta-thalassaemia in fetuses at risk of being homozygous for the gene mutation. In some forms of beta-thalassaemia the mutated codon coincides with a restriction site. This unusual fact makes the diagnosis *in utero* completely reliable. For the vast majority of beta-thalassaemias, however, prenatal diagnosis depends on the demonstration of linkages between the thalassaemia genes and particular RFLPs within families (WEATHERALL, 1982, p 82). The safety and acceptability of the prenatal diagnosis will be greatly improved when earlier, less invasive, techniques of fetal sampling (e.g., chorion biopsy during the first trimester of pregnancy) become generally available.

References.

ALTER, BP (1979) Prenatal diagnosis of haemoglobinopathies and other haematologic diseases. J Pediatr 95: 501.

ASTRIN, KH, DESNICK, RJ, SMITH, M (1981) Provisional assignment of esterase B3 (ESB3) to human chromosome 16. Cytogenet Cell Genet 32: 250.

BAGLIONI, C (1962) The fusion of two peptide chains in haemoglobin Lepore and its interpretation as a genetic deletion. Proc Natl Acad Sci USA 48: 1880.

BAILEY, NTJ (1959) Statistical methods in biology. London. Hodder and Stroughton.

BALLAS, SK, BURKA, ER (1982) Proteolysis of excess alpha chains in beta-thalassaemia. J Lab Clin Med 99: 754.

BARG, R, BARTON, P, CAINE, A, CLEMENTS, RL, FERGUSON-SMITH, MA, MALCOLM, S, MORRISON, N, MURPHY, CS (1981) Regional localization of human alpha-globin gene to short arm of chromosome 16 (16p12-pter) using both somatic cell hybrids and in situ hybridization. Cytogenet Cell Genet 32: 252.

BARTON, P, MALCOLM, S, MURPHY, CS, FERGUSON-SMITH, MA (1982) Localization of the human alpha-globin gene cluster to the short arm of chromosome 16 (16p12-pter) by hybridization in situ. J Molec Biol 156: 269-278.

BELL, J, HALDANE, JBS (1937) The linkage between the genes for colour blindness and haemophilia in man. Proc Roy Soc 123B: 119-150.

BERNHEIM, A, BERGER, R, LENOIR, G (1980) Translocation t(2:8) et t(8:22) dans les lignees cellulaires continues de lymphome de Burkitt africain. C R Acad Sci (Paris) 291: 237-239.

BERNSTEIN, F (1931) Zur Grundlegung der Chromosomentheorie der Vererbung beim Menschen. Zeitschr Abst Vererb 57: 113-138.

BIRNBOIM, HC, DOLY, J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl Acid Res 7: 1513-1523.

BLANKENSTEIN-WIJNEN, LMM, SCHMIEDER, G, GRZESCHIK, K-H, PEARSON, PL, WESTERVELD, A, MEERA KHAN, P (1979) Chromo-

somal localization of PGP. Cytogenet Cell Genet 25: 138.

BODE, U, DEISSEROTH, A, HENDRICK, D (1980) Expression of non-histone chromosomal proteins (NHCP) on human chromosome hybrid mouse erythroleukaemia (MEL) cells which contain human alpha-globin gene (HAG) from non-erythroid cells. Fed Proc 39: 2198.

BODMER, WF (1981) Gene clusters, genome organization and complex phenotypes. When the sequence is known, what will it mean? Am J Hum Genet 33: 664-682.

BOEHM, CD, ANTONARAKIS, SE, PHILIPS III, JA, STETTEN, G, KAZAZIAN JR, HH (1983) Prenatal diagnosis using DNA polymosphisms: Report on 95 pregnancies at risk for sickle-cell disease and beta-thalassaemia. N Engl J Med 308: 1054.

BOOTSMA, D, MCALPINE, PJ (1979) Report of the Committee on the genetic constitution of chromosomes 2, 3, 4 and 5. Cytogenet Cell Genet 25: 21-31.

BOTSTEIN, D, WHITE, RL, SKOLNICK, M, DAVIS, RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314-331.

BOYER, SH, RUCKNAGEL, DL, WEATHERALL, DJ, WATSON-WILLIAMS, EJ (1963) Further evidence for linkage between beta and delta loci governing human haemoglobin and the population dynamics of linked genes. Am J Hum Genet 15: 438.

BRADLEY, TB, BOYER, SH, ALLEN, FH (1961) Hopkins-2-haemoglobin: a revised pedigree with data on blood and serum groups. Bull Johns Hopkins Hosp 108: 75.

BUNN, HF, FORGET, BG, RANNEY, HM (1977) Human haemoglobins. Philadelphia. W B Saunder Co.

CARRANO, AV, GRAY, JW, LANGLOIS, RG, BURKHART-SCHULTZ, KJ, VAN DILLA, MA (1979) Measurement and purification of human chromosomes by flow cytometry and sorting. Proc Natl Acad Sci USA 76: 1382-1384.

CHANG, JC, KAN, YW (1982) A sensitive prenatal test for sickle-cell anaemia. N Engl J Med 307: 30-32.

CHANY, C, VIGNAL, M, COUILLIN, P, NGUYEN VAN CONG, BOUE, J, BOUE, A (1975) Chromosomal localization of human genes governing the interferon-induced antiviral-state. Proc Natl Acad Sci USA 72: 3129-3133.

CHEN, TR (1977) In situ detection of mycoplasma contamination in cell cultures by fluorescent Hoechst

33258 stain. Exp Cell Res 104: 255.

CHU, EHY, POWELL, SS, (1976) Selective systems in somatic cell genetics. Adv Hum Genet 7: 189-258.

COMINGS, DE (1972) Evidence for ancient tetraploidy and conservation of linkage groups in mammalian chromosomes. Nature 238: 455-467.

Committee on Standardized Human Gene Nomenclature (1981) The 1981 catalogue of assigned human genetic markers and report of the nomenclature committee. Cytogenet Cell Genet 32: 221-245.

CONNEALLY, PM, RIVAS, ML (1980) Linkage analysis in man. Adv Hum Genet 10: 209-266.

CORRENS, C (1905) Über Vererbungsgesetze. Berlin. G Borntträger.

CRAIK, CS, BUCHMAN, SR, BEYCHOK, S (1980) Characterization of globin domains: haem binding to the central exon product. Proc Natl Acad Sci USA 77: 1384-1388.

CREAGAN, RP, TAN, YH, CHEN, S, RUDDLE, FH (1975) Somatic cell genetic analysis of the interferon system. Fed Proc 34: 2222-2226.

CZELUSNIAK, J, GOODMAN, M, HEWETT-EMMETT, D, WEISS, ML, VENTA, PJ, TASHIAN, RE (1982) Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. Nature 298: 297-300.

DARLINGTON, CJ, RANKIN, JK, SCHLANGER, G (1982) Expression of human hepatic genes in somatic cell hybrids. Somat Cell Genet 8: 403-412.

DAVIES, KE (1981) The application of DNA recombinant technology to the analysis of the human genome and genetic disease. Hum Genet 58: 351-357.

DEISSEROTH, A, VELEZ, R, NIENHUIS, AW (1976) Haemoglobin synthesis in somatic cell hybrids: independent segregation of the human alpha- and beta-globin genes. Science 191: 1262-1264.

DEISSEROTH, A, NIENHUIS, AW, TURNER, P, VELEZ, R, ANDERSON, WF, RUDDLE, FH, LAWRENCE, J, CREAGAN, R, KUCHERLAPATI, R (1977) Localization of the human alpha-globin structural gene to chromosome 16 in somatic cell hybrids by molecular hybridization assay. Cell 12: 205-218.

DEISSEROTH, A, HENDRICK, D (1978) Human alpha-globin gene expression following chromosomal dependent gene

transfer into mouse erythroleukaemia cells. Cell 15: 55-63.

DEISSEROTH, A, NIENHUIS, AW, LAWRENCE, J, GILES, R, TURNER, P, RUDDLE, FH (1978) Chromosomal localization of human beta-globin gene on human chromosome 11 in somatic cell hybrids. Proc Natl Acad Sci USA 75: 1456-1460.

DE KLEIN, A, VAN KESSEL, AG, GROSVELD, G, BARTRAM, CR, HAGEMEJER, A, BOOTSMA, D, SPURR, NK, HEISTERKAMP, N, GROFFEN, J, STEPHENSON, JR (1982) A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature 300: 765-767.

DICKERSON, RE, GEIS, I (1980) Proteins: structure, function and evolution. Menlo Park, California. Benjamin Cummings Publ.

DONAHUE, RP, BIAS, WB, RENWICK, JH, MCKUSICK, VA (1968) Probable assignment of the Duffy blood group locus to chromosome 1 in man. Proc Natl Acad Sci USA 61: 949-955.

DONALD, LJ, WANG, HS, HAMERTON, JL (1979) Assignment of phosphoglycollate phosphatase locus to chromosome 16. Cytogenet Cell Genet 25: 147.

DONALD, LJ, WANG, HS, HAMERTON, JL (1981a) Assignment of the gene for cystathionase to human chromosome 16. Cytogenet Cell Genet 32: 268.

DONALD, LJ, WANG, HS, HOLLIDAY, NJ, HAMERTON, JL (1981b) Are there growth-rate-controlling factors on chromosomes 7 and 16? Cytogenet Cell Genet 32: 268.

DULBECCO, R, VOGT, M (1954) Plaque formation and isolation of pure lines with polyomyelitis viruses. J Exp Med 99: 167-182.

EDWARDS, JH (1956) Antenatal detection of hereditary disorders. Lancet I: 579.

EDWARDS, Y, POTTER, J, HOPKINSON, DA (1980) Human FAD-dependent NAD(P)H diaphorase. Biochem J 187: 429-436.

EFSTRATIADIS, A, POSAKONY, JW, MANIATIS, T, LAWN, RM, O'CONNELL, C, SPRITZ, RA, DERIEL, JK, FORGET, BG, WEISSMAN, SM, SLIGHTOM, JL, BLECHL, AE, SMITHIES,O, BARALLE, FE, SHOULDERS, CC, PROUDFOOT, NJ (1980) The structure and evolution of the human beta-globin gene family. Cell 21: 653-668.

EGE, T, RINGERTZ, NR (1974) Preparation of microcells by enucleation of micronucleate cells. Exp Cell Res 87: 378-382.

EMBURY, SH, LEBO, RV, DOZY, AM, KAN, YW (1979) Organization of the alpha-globin genes in the Chinese alpha-thalassaemia syndromes. J Clin Invest 63: 1307.

FELBER, BK, ORKIN, SH, HAMER, DH (1982) Abnormal RNA splicing causes one form of alpha-thalassaemia. Cell 29: 895-902.

FERGUSON-SMITH, MA, NEWMAN, BF, ELLIS, PM, THOMSON, DMG, RILEY, ID (1973) Assignment by deletion of human red cell acid phosphatase gene locus to the short arm of chromosome 2. Nature 243: 271-273.

FERGUSON-SMITH, MA (1974) Techniques of human chromosome analysis. La Ricerca Clin Lab 4: 297-335.

FERRELL, RE, HITTNER, HM, ANTOSZYK, JH (1983) Linkage of atypical vitelliform macular dystrophy (VMD1) to the soluble glutamate pyruvate transaminase (GPT1) locus. Am J Hum Genet 35: 78-84.

FLAVELL, RA, KOOTER, JM, DEBOER, E, LITTLE, PFR, WILLIAMSON, R (1978) Analysis of the beta-delta-globin gene loci in normal and Hb Lepore DNA: direct determination of gene linkage and intergene distance. Cell 15: 25-41.

FRITSCH, EF, LAWN, RM, MANIATIS, T (1980) Molecular cloning and characterization of the human beta-like globin gene cluster. Cell 19: 959-972.

GERHARD, DS, KAWASAKI, ES, CARTER BANCROFT, F, SZABO, P (1981) Localization of a unique alpha-globin gene by direct hybridization *in situ*. Proc Natl Acad Sci USA 78: 3755-3759.

GOOSSENS, M, LEE, KY, LIEBHABER, SA, KAN, YW (1982) Globin structural mutant alpha^{125}Leu→Pro in a novel case of alpha-thalassaemia. Nature 296: 864-865.

GOSS, SJ, HARRIS, H (1977) Gene transfer by means of cell fusion. I. Statistical mapping of the human X chromosome by analysis of radiation induced gene segregation. J Cell Sci 25: 17-38.

GROSS-BELLARD, M, OUDET, P, CHAMBON, P (1973) Isolation of high-molecular-weight DNA from mammalian cells. Eur J Biochem 36: 32-38.

GRZESCHIK, K-H (1979) Assignment of a structural gene for a fourth diaphorase to chromosome 16 in human x mouse somatic cell hybrids. Cytogenet Cell Genet 25: 162.

GRZESCHIK, K-H (1980) Assignment of a structural gene

for a fourth human diaphorase (DIA4) to chromosome 16 in man x mouse somatic cell hybrids. Hum Genet 53: 189-193.

HAMERTON, JL, MOHANDAS, T, MCALPINE, PJ, DOUGLAS, GR (1974) Localization of gene loci on chromosome 12 and the X in man. Birth Defects: Original Article Series XI(3): 179-182.

HARRIS, H, HOPKINSON, DA (1976) Handbook of enzyme electophoresis in human genetics. Amsterdam. North-Holland Publ Co.

HENDERSON, AS, WARBURTON, D, ATWOOD, KC (1972) Localization of ribosomal DNA in the human chromosome complement. Proc Natl Acad Sci USA 69: 3394.

HOLLAN, SR, SZELENYI, JG, BRIMHALL, B et al (1972) Multiple alpha-chain loci for human haemoglobin. HbJ-Buda and HbG-Pest. Nature 235: 47.

HONEY, NK, SHOWS, TB (1981) Assignment of glyoxalase II gene (HAGH) to human chromosomse 16. Hum Genet 58: 358-361.

HOSBACH, HA, WYLER, T, WEBER, R (1983) The Xenopus laevis globin gene family: chromosomal arrangement and gene structure. Cell 32: 45-53.

HOUCK, CM, RINEHART, FP, SCHMID, CW (1979) A ubiquitous family of repeated DNA sequences in the human genome. J Molec Biol 132: 289-306.

HUISMAN, THJ, WRIGHTSTONE, RN, WILSON, JB, SCHROEDER, WA, KENDALL, AG (1972) Haemoglobin Kenya, the product of fusion of gamma and beta polypeptide chains. Arch Biochem Biophys 153: 850.

INSLEY, J, BIRD, GWG, HARPER, PS, PEARCE, GW (1976) Prenatal diagnosis of myotonic dystrophy. Lancet I: 806.

IUB Nomenclature Committee (1978) Enzyme nomenclature. New York. Academic Press.

JEFFREYS, AJ, CRAIG, IW, FRANCKE, U (1979) Localization of gamma-G-, gamma-A-, delta- and beta-globin genes on the short arm of human chromosome 11. Nature 281: 606-608.

JEFFREYS, AJ, WILSON, V, WOOD, D, SIMONS, JP, KAY, RM, WILLIAMS, JG (1980) Linkage of adult alpha- and beta-globin genes in X. laevis and gene duplication by tetraploidization. Cell 21: 555-564.

JEFFREYS, AJ (1981) Recent studies of gene evolution using recombinant DNA. In WILLIAMSON, R (ed) Genetic engineering 2. London. Academic Press. Pp 1-48.

JEFFREYS, AJ (1982) Evolution of globin genes. In DOVER, GA, FLAVELL, RB (eds) Genome evolution. London. Academic Press. Pp 157-176.

JEFFREYS, AJ, HARRIS, S (1982) Processes of gene duplication. Nature 296: 9-10.

JELINEK, WR, TOOMEY, TP, LEINWAND, L, DUNCAN, CH, BIRO, PA, CHOUDARY, PV, WEISSMAN, SM, RUBIN, CM, HOUCK, CM, DEININGER, PL, SCHMID, CW (1980) Ubiquitous, interspersed repeated sequences in mammalian genomes. Proc Natl Acad Sci USA 77: 1398-1402.

JEREMIAH, SJ, POVEY, S, BURLEY, MW, KIELTY, C, LEE, M, SPOWART, G, CORNEY, G, COOK, PJL (1982) Mapping studies on human mitochondrial glutamate oxaloacetate transaminase. Ann Hum Genet 46: 145-152.

JUNIEN, C (1982) Les techniques de recombinaison génétique et la carte chromosomique humaine. Ann Génét 25: 71-86.

KAHAN, B, HELD, KR, DEMARS, R (1975) The locus of adenine phosphoribosyltransferase on chromosome no. 16. Genetics 78: 1143-1156.

KAN, YW, DOZY, AM (1978a) Antenatal diagnosis of sickle-cell anaemia by DNA analysis of amniotic fluid cells. Lancet II: 910-912.

KAN, YW, DOZY, AM, (1978b) Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. Proc Natl Acad Sci USA 75: 5631-5635.

KAO, F-T, HARTZ, JA, LAW, ML, DAVIDSON, JN (1982) Isolation and chromosomal localization of unique DNA sequences from a human genomic library. Proc Natl Acad Sci USA 79: 865-869.

KEATS, BJB, MORTON, NE, RAO, DC (1981) Reduction of physical assignments to a standard lod table: chromosome 1. Hum Genet 56: 353-359.

KING, M-C, GO, RCP, ELSTON, RC, LYNCH, HT, PETRAKIS, NL (1980) Allele increasing susceptibility to human breast cancer may be linked to the glutamate-pyruvate transaminase locus. Science 208: 406-408.

KLEBE, RJ, MANCUSO, MG (1982) Enhancement of polyethylene glycol-mediated cell hybridization by inducers of erythroleukaemia cell differentiation. Somat Cell Genet 8: 723.

KLOBUTCHER, LA, RUDDLE, FH (1979) Phenotype stabilization and integration of transferred material in chromosome-mediated gene transfer. Nature 280: 657-660.

KOEFFLER, HP, SPARKES, RS, STANG, H, MOHANDAS, T (1981) Regional assignment of genes for human alpha-globin and phosphoglycollate phosphatase to the short arm of chromosome 16. Proc Natl Acad Sci USA 78: 7015-7018.

KUCHERLAPATI, RS, HILWIG, I, GROPP, A, RUDDLE, FH (1975) Mammalian chromosome identification in interspecific hybrid cells using Hoechst 33258. Humangenetik 27: 9-14.

KUSANO, T, LONG, C, GREEN, H (1971) A new reduced human x mouse somatic cell hybrid containing the human gene for adenine phosphoribosyltransferase. Proc Natl Acad Sci USA 68: 82-86.

KWOK, SCM, CHAN, SJ, RUBENSTEIN, AH, POUCHER, R, STEINER, DF (1981) Loss of a restriction endonuclease cleavage site in the gene of a structurally abnormal human insulin. Biochem Biophys Res Commun 98: 844-849.

LANGE, K, BOEHNKE, M (1982) How many polymorphic marker genes will it take to span the human genome? Am J Hum Genet 34: 842-845.

LAUER, J, CHE-KUN, JS, MANIATIS, T (1980) The chromosomal arrangement of human alpha-like globin genes: sequence homology and alpha-globin gene deletions. Cell 20: 119-130.

LAURIE, DA, HULTÉN, M, JONES, GH (1981) Chiasma frequency and distribution in a sample of human males: chromosomes 1, 2 and 9. Cytogenet Cell Genet 31: 153-166.

LAWN, RM, FRITSCH, EF, PARKER, RC, BLAKE, G, MANIATIS, T (1978) The isolation and characterization of linked delta- and beta-globin genes from a cloned library of human DNA. Cell 15: 1157-1174.

LEBO, RV, YU, LC, CORDELL, B, CHEUNG, MC, CHANG, JC, CARRANO, AV, GOODMAN, HM, KAN, YW (1981) Subchromosomal assignment of the human insulin gene to the terminal portion of the short arm of chromosome 11 by chromosome sorting and DNA restriction enzyme analysis. Clin Res 29: 114A (abstract).

LENOIR, GM, PREUD'HOMME, JL, BERNHEIM, A, BERGER, R (1982) Correlation between immunoglobulin light chain expression and variant translocation in Burkitt's lymphoma. Nature 298: 474-476.

LEVINE, LS, ZACHMAN, M, NEW, MI, PRADER, A, POLLACK, MS,

O'NEILL, GJ, YANG, SY, OBERFIELD, SE, DUPONT, B (1978) Genetic mapping of the 21-hydroxylase-deficiency gene within the HLA linkage group. N Engl J Med 299: 911-915.

LEY, TJ, DESIMONE, J, ANAGNOU, NP, KELLER, GH, HUMPHRIES, RK, TURNER, PH, YOUNG, NS, HELLER, P, NIENHUIS, AW (1982) 5-Azacytidine selectively increases gamma-globin synthesis in a patient with beta$^+$-thalassaemia. N Engl J Med 307: 1469-1475.

LIEBHABER, SA, GOOSSENS, M, KAN, YW (1981) Homology and concerted evolution at the alpha 1 and alpha 2 loci of human alpha-globin. Nature 290: 26-29.

LITTLE, PFR, ANNISON, G, DARLING, S, WILLIAMSON, R, CAMBA, L, MODELL, B (1980) Model for antenatal diagnosis of beta-thalassaemia and other monogenic disorders by molecular analysis of linked DNA polymorphisms. Nature 285: 144-147.

LITTLE, PFR (1981) DNA analysis and the antenatal diagnosis of haemoglobinopathies. In WILLIAMSON, R (ed) Genetic engineering 1. London. Academic Press. Pp 62-102.

LITTLE, PFR (1982) Globin pseudogenes (minireview). Cell 28: 683-684.

LOWRY, OH, ROSEBROUGH, NJ, FARR, AL, RANDALL, RJ (1951) Protein measurement with the Folin phenol reagent. J Biol Chem 193: 265-275.

MACPHERSON, I, STOKER, MGP (1961) Studies on transformation of hamster cells by polyoma virus in vitro. Virology 14: 359-370.

MAGENIS, E, CHAMBERLIN, J (1979) Location of fragile site: 16q22.00. Cytogenet Cell Genet 25: 180.

MALCOLM, S, BARTON, P, BENTLEY, DL, FERGUSON-SMITH, MA, MURPHY, CS, RABBITTS, TH (1981a) Assignment of an IGKV locus for immunoglobulin light chains to the short arm of chromosome 2 (2p13→cen) by in situ hybridization using a cRNA probe of H_k101lambdaCh4A. Cytogenet Cell Genet 32: 296.

MALCOLM, S, BARTON, P, MURPHY, C, FERGUSON-SMITH, MA (1981b) Chromosomal localization of a single copy gene by in situ hybridization - human beta-globin genes on the short arm of chromosome 11. Ann Hum Genet 45: 135-141.

MANIATIS, T, FRITSCH, EF, LAUER, J, LAWN, RM (1980) The molecular genetics of human haemoglobins. Ann Rev Genet 14: 145-178.

MARTIN, SL, ZIMMER, EA, KAN, YW, WILSON, AC (1980) Silent delta-globin gene in Old World monkeys. Proc Natl Acad Sci USA 77: 3563-3566.

MARTIN, SL, VINCENT, KA, WILSON, AC (1983) Rise and fall of the delta-globin gene. J Molec Biol 164: 513-528.

MCCURDY, PR (1971) Use of genetic linkage for the detection of female carriers of haemophilia. N Engl J Med 285: 218-219.

MCDOUGALL, JK, KUCHERLAPATI, R, RUDDLE, FH (1973) Localization and induction of the thymidine kinase gene by adenovirus 12. Nature 245: 172-173.

MCKUSICK, VA, RUDDLE, FH (1977) The status of the gene map of the human chromosomes. Science 196: 390-405.

MCKUSICK, VA (1980) The anatomy of the human genome. J Hered 71: 370-391.

MCKUSICK, VA (1982) The human gene map (20 October 1982). Clin Genet 22: 359-391.

MCKUSICK, VA (1983) Mendelian inheritance in man. 6th edition. Baltimore. Johns Hopkins University Press.

MENDELSOHN, ML, MAYALL, BH, BOGART, E, MOORE, DH, PERRY, BH (1973) DNA content and DNA-based centromeric index of the 24 human chromosomes. Science 179: 1126-1129.

MICHELSON, AM, ORKIN, SH (1980) The 3' untranslated regions of the duplicated human alpha-globin genes are unexpectedly divergent. Cell 22: 371-377.

MILLER, OJ, ALLERDICE, PW, MILLER, DA, BREG, WR, MIGEON, BR (1971) Human thymidine kinase gene locus: assignment to chromosome 17 in a hybrid of man and mouse cells. Science 173: 244-245.

MILLER, OJ (1972a) Analysis of heterogeneity. In CASPERSSON, T, ZECH, L (eds) Chromosome identification - technique and applications in biology and medicine. 23rd Nobel Symposium. New York. Academic Press. Pp 177-178.

MILLER, OJ (1972b) The karyotype and chromosome map of the mouse. In CASPERSSON, T, ZECH, L (eds) Chromosome identification - technique and applications in biology and medicine. 23rd Nobel Symposium. New York. Academic Press. Pp 132-144.

MOHR, J (1951) Estimation of linkage between the Lutheran and the Lewis blood groups. Acta Path Microbiol Scand 29: 339-344.

MOORHEAD, PS, NOWELL, PC, MELLMAN, WJ, BATTIPS, DM, HUNGERFORD, DA (1960) Chromosome preparations of leukocytes cultured from human peripheral blood. Exp Cell Res 20: 613-616.

MORGAN, TH (1911) Random segregation versus coupling in Mendelian inheritance. Science 34: 384.

NABHOLZ, M, MIGGIANO, V, BODMER, W (1969) Genetic analysis with human x mouse somatic cell hybrids. Nature 223: 358.

NGUYEN VAN CONG, WEIL, D, HORS-CAYLA, MC, GROSS, MS, HEUERTZ, S, FOUB, C, FREZAL, J (1980) Assignment of the genes for human lysosomal acid lipases A and B to chromosomomes 10 and 16. Hum Genet 55: 375-381.

NICHOLS, EA, RUDDLE, FH (1973) A review of enzyme polymorphism, linkage and electrophoretic conditions for mouse and somatic cell hybrids in starch gels. J Histochem Cytochem 21: 1066-1081.

NIENHUIS, AW, STAMATOYANNOPOULOS, G (1978) Haemoglobin switching. Cell 15: 307.

OHNO, S (1970) Evolution by gene duplication. Heidelberg. Springer.

OHNO, S (1973) Ancient linkage groups and frozen accidents. Nature 244: 259-262.

OLAISEN, B, GEDDE-DAHL, Jr, T (1973) GPT-epidermolysis bullosa simplex (EB 5 Ogna) linkage in man. Hum Hered 23: 189-196.

OLD, JM, WARD, RHT, PETROV, M, KARAGOZLU, F, MODELL, B, WEATHERALL, DJ (1982) First trimester fetal diagnosis for haemoglobinopathies. Lancet II: 1413.

ORKIN, SH (1978) The duplicated human alpha-globin genes lie close together in cellular DNA. Cell 21: 555-564.

ORKIN, SH, MICHELSON, A (1980) Partial deletion of the alpha-globin structural gene in human alpha-thalassaemia. Nature 286: 538.

ORKIN, SH, GOFF, SC, HECHTMAN, RL (1981) Mutation in an intervening sequence splice junction in man. Proc Natl Acad Sci USA 78: 5041-5045.

ORKIN, SH, KAZAZIAN, HH, ANTONARAKIS, SE, GOFF, SC, BOEHM, CD, SEXTON, JP, WABER, PG, GIARDINA, PJV (1982) Linkage of beta-thalassaemia mutations and beta-globin gene polymorphisms with DNA polymorphisms in human beta-

globin gene cluster. Nature 296: 627-631.

OTT, J, LINDER, D, MCCAW, BK, LOVRIEN, EW, HECHT, F (1976) Estimating distances from the centromere by means of benign ovarian teratomas in man. Ann Hum Genet 40: 191-196.

OWERBACH, D, BELL, GT, RAUTTER, WJ, SHOWS, TB (1980) The insulin gene is located on chromosome 11. Nature 286: 82-84.

PAINTER, TS (1933) A new method for the study of chromosome rearrangements and the plotting of chromosome mappings. Science 78: 585-586.

Paris Conference (1971), Supplement (1975) Standardization in Human Genetics. Birth Defects: Original article series, XI, 9, 1975. The National Foundation, New York.

PEARSON, PL, BAKKER, E, FLAVELL, RA (1981) Considerations in designing an efficient strategy for localizing unique sequence DNA fragments to human chromosomes. Cytogenet Cell Genet 32: 308.

PERLER, F, EFSTRATIADIS, A, LOMEDICO, P, GILBERT, W, KOLODNER, R, DODGSON, J (1980) The evolution of genes: the chicken preproinsulin gene. Cell 20: 555-565.

POLANI, PE (1982) Paring of X and Y chromosomes, non-inactivation of X-linked genes and the maleness factor. Hum Genet 60: 207-211.

PONTECORVO, G (1975) Production of mammalian somatic cell hybrids by means of polyethyleneglycol treatment. Somat Cell Genet 1: 397-400.

PONTECORVO, G (1976) Polyethyleneglycol (PEG) in the production of mammalian somatic cell hybrids. Birth Defects: Original Article Series XII: 399-400.

POVEY, S, JEREMIAH, SJ, BARKER, RF, HOPKINSON, DA, SOLOMON, E, BOBROW, M, CARRITT, B, ROBSON, EB, COOK, PJL (1979) The current position of PGP. Cytogenet Cell Genet 25: 199.

POVEY, S, WILSON, D, EDWARDS, YH (1980) Assignment of human diaphorase (DIA4) to chromosome 16. Ann Hum Genet 43: 349-353.

POVEY, S, JEREMIAH, SJ, BURLEY, MW, KIELTY, C, SPOWART, G, CORNEY, G, COOK, PJL (1981) Contributions to the mapping of chromosome 16. Cytogenet Cell Genet 32: 311.

PRESSLEY, L, HIGGS, DR, ALDRIGE, B, METAXATOU-MAVROMATI,

A (1980a) Characterization of a new alpha-thalassaemia 1 defect due to a partial deletion of the alpha-globin gene complex. Nucl Acid Res 8: 4889.

PRESSLEY, L, HIGGS, DR, CLEGG, JB, WEATHERALL, DJ (1980b) Gene deletion in alpha-thalassaemia prove that the 5' zeta locus is functional. Proc Natl Acad Sci USA 77: 3586-3589.

PROUDFOOT, N (1980) Pseudogenes. Nature 286: 840-841.

PROUDFOOT, N, MANIATIS, T (1980) The structure of a human alpha-globin pseudogene and its relationship to alpha-globin gene duplication. Cell 21: 537-544.

PROUDFOOT, N, SHANDER, MHM, MANLEY, JL, GETTER, ML, MANIATIS, T (1980) Structure and in vitro transcription of human globin genes. Science 209: 1329-1336.

PROUDFOOT, N, GIL, A, MANIATIS, T (1982) The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. Cell 31: 553-563.

PUCK, TT, KAO, F-T (1982) Somatic cell genetics and its applications to medicine. Ann Rev Genet 16: 225-271.

RENWICK, JH, LAWLER, SD (1955) Genetic linkage between the ABO and nail-patella loci. Ann Hum Genet 19: 312-331.

RÉTHORÉ, M-O, LAFOURCADE, J, COUTURIER, J, HARPEY, JP, HAMET, M, ENGLER, R, ALCINDOR, LG, LEJEUNE, J (1982) Augmentation d'activité de APRT chez un enfant trisomique 16q22.2→qter par translocation t(16:21) (q22.2:q22.2)pat. Ann Génét 25: 36-42.

RICHARDSON, CC, SCHILDKRAUT,CL, APOSHIAN, HV, KORNBERG, A (1964) Enzymatic synthesis of deoxyribonucleic acid. XIV. Further purification and properties of deoxyribonucleic acid polymerase of Escherichia coli. J Biol Chem 239: 222-232.

RIGBY, PWJ, DIECKMANN, M, RHODES, C, BERG, P (1977) Labelling DNA to high specific activity in vitro by nick translation with DNA polymerase I. J Molec Biol 113: 237-251.

ROBSON, EB, POLANI, PE, DART, SJ, JACOBS, PA, RENWICK, JH (1969) Probable assignment of the alpha locus of haptoglobin to chromosome 16 in man. Nature 223: 1163.

ROMERO-HERRERA, AE, LEHMANN, H, JOYSEY, KA, FRIDAY, AE (1973) Molecular evolution of myoglobin and the fossil record: a phylogenetic synthesis. Nature 246: 389-395.

RUDDLE, FH (1981) A new era in mammalian gene mapping: somatic cell genetics and recombinant DNA methodologies. Nature 294: 115.

SAKAGUCHI, AY, NAYLOR, SL, QUINTO, C, RUTTER, WJ, SHOWS, TB (1981) The chymotrypsinogen B gene (CTRB) is on human chromosome 16. Cytogenet Cell Genet 32: 313.

SANCAR, GB, CEDENO, MM, BELLEVUE, R, RIEDER, RF (1982) Interaction of chromosomes bearing 1, 2 or 3 alpha-globin genes in an American black family with alpha-thalassaemia. Haemoglobin 6: 99.

SCHMID, CW, JELINEK, WR (1982) The Alu family of dispersed repetitive sequences. Science 216: 1065-1070.

SCHMID, M, KLETT, C, NIEDERHOFFER, A (1980) Demonstration of a heritable fragile site in human chromosome 16 with distamycin A. Cytogenet Cell Genet 28: 87-94.

SEABRIGHT, M (1972) The use of proteolytic enzymes for the mapping of structural rearrangements in the chromosomes of Man. Chromosoma 36: 204-210.

SEUÁNEZ, HN (1979) The phylogeny of human chromosomes. Berlin. Springer-Verlag.

SHEN, S-H, SLIGHTON, JL, SMITHIES, O (1981) A history of the human fetal globin gene duplication. Cell 26: 191-203.

SHEN, S-H, SMITHIES, O (1982) Human globin pseudobeta 2 is not a globin-related sequence. Nucl Acid Res 10: 7809-7818.

SHOWS, TB, SAKAGUCHI, AY, NAYLOR, SL (1982) Mapping the human genome, cloned genes, DNA polymorphisms and inherited disease. Adv Hum Genet 12: 341-452.

SLIGHTON, JL, BLECHL, AE, SMITHIES, O (1980) Human fetal gamma G- and gamma A-globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21: 627-638.

SOUTHERN, EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Molec Biol 98: 503-519.

SPARKES, RS, MOHANDAS, T, SPARKES, MC, PASSAGE, MB, SHULKIN, JD (1979) Expression of human phosphoglycollate phosphatase in somatic cell hybrids. Cytogenet Cell Genet 25: 207.

STEWART, TA, WAGNER, EF, MINTZ, B (1982) Human beta-

globin gene sequences injected into mouse eggs, retained in adults and transmitted to progeny. Science 217: 1046-1048.

STOKER, MGP, MACPHERSON, I (1962) Polyoma transformation of hamster cell clones - an investigation of genetic factors affecting cell competence. Virology 16: 147-151.

STURTEVANT, AH (1913) The linear arrangement of six sex-linked factors in Drosophila as shown by their mode of association. J Exp Zool 14: 43-59.

SUNDAR RAJ, CV, CHURCH, RL, KLOBUTCHER, LA, RUDDLE, FH (1977) Genetics of the connective tissue proteins: assignment of the gene for human type I procollagen to chromosome 17 by analysis of cell hybrids and microcell hybrids. Proc Natl Acad Sci USA 74: 4444-4448.

SUTTON, WS (1903) The chromosomes in heredity. Biol Bull 4: 231-251.

TEISBERG, P, GJONE, B (1974) Probable linkage of LCAT locus in man to the alpha-haptoglobin locus on chromosome 16. Nature 249: 550-551.

TESTA, U, HINARD, N, BEUZARD, Y, TAPSIS, A, GALACTEROS, F, THOMPOULOS, P, ROSA, J (1981) Excess alpha chains are lost from beta-thalassaemic reticulocytes by proteolysis. J Lab Clin Med 98: 352-363.

TISCHFIELD, JA, CREAGAN, RP, RUDDLE, FH (1973) Assignment of a selectable gene: adenine phosphoribosyltransferase to chromosome 16. Birth Defects: Original Article Series X: 167-172.

TISCHFIELD, JA, RUDDLE, FH (1974) Assignment of the gene for APRT to human chromosome 16 by mouse-somatic cell hybridization. Proc Natl Acad Sci USA 71: 45-49.

TOLLEY, E, VAN HEYNINGEN, V, BROWN, R, BOBROW, M, CRAIG, IW (1980) Assignment to chromosome 16 of a gene necessary for the expression of human mitochondrial glutamate oxaloacetate transaminase (aspartate aminotranferase) (EC 2.6.1.1). Biochem Genet 18: 947-954.

TURNER, G, BROOKWELL, R, DANIEL, A, SELIKOWITZ, M (1980) Heterozygous expression of mental retardation and X chromosome marker fra(X)(q27). N Engl J Med 303: 662-664.

VAN SOMEREN, H, VAN HENEGOUWEN, HB (1973) Independent loss of human hexosaminidase A and B in man x Chinese hamster somatic cell hybrids. Humangenetik 18: 171-174.

WAHL, GM, STERN, M, STARK, GR (1979) Efficient transfer of large DNA fragments from agarose gels to diazobenzyl-oxymethyl-paper and rapid hybridization by using dextran sulfate. Proc Natl Acad Sci USA 76: 3683-3687.

WAINSCOAT, JS, KANAVAKIS, E, WEATHERALL, DJ, WALKER, J, HOLMES-SEIDLE, M, BOBROW, M, DONNISON, AB (1981) Regional localization of the human alpha-globin genes. Lancet II: 301-302.

WEATHERALL, DJ, CLEGG, JB (1975) The alpha-chain-termination mutants and their relationship to alpha-thalassaemia. Philos Trans R Soc London 271: 411-455.

WEATHERALL, DJ, CLEGG, JB, (1981) The thalassaemia syndromes. 3rd ed. Oxford. Blackwell Scientific Publ.

WEATHERALL, DJ, HIGGS, DR, BUNCH, MB, et al (1981) Haemoglobin H disease and mental retardation: a new syndrome or a remarkable coincidence? N Engl J Med 305: 607-612.

WEATHERALL, DJ (1982) The new genetics and clinical practice. London. Nuffield Provincial Hospitals Trust.

WIJNEN, LMM, MEERA-KHAN, P (1981) Assignment of GPT1 to human chromosome 16. Cytogenet Cell Genet 32: 327.

WILLECKE, K, REBER, T, KUCHERLAPATI, RS, RUDDLE, FH (1977) Human mitichondrial thymidine kinase is coded for by a gene on chromosome 16 of the nucleus. Somat Cell Genet 3: 237-245.

WILLIAMSON, R, ESKDALE, J, COLEMAN, DV, NIAZI, M, LOEFFLEY, FE (1981) Direct gene analysis of chorionic villi: a possible technique for first-trimester antenatal diagnosis of haemoglobinopathies. Lancet II: 1125-1127.

WILLIAMSON, R (1982) Gene therapy. Nature 298: 416-418.

WILSON, AC, CARLSON, SS, WHITE, TJ (1977) Biochemical evolution. Ann Rev Biochem 46: 573-639.

WILSON, EB (1911) The sex chromosomes. Arch Mikrobiol Anat 77: 249-271.

WILSON, TJ, WILSON, LB, DE RIEL, JK, VILLAKOMAROFF, L, EFSTRATIADIS, A, FORGET, BG, WEISSMAN, SM (1978) Insertion of synthetic copies of human globin genes into bacterial plasmids. Nucl Acid Res 5: 563-582.

WONG, V, MA, HK, TODD, D, GOLBUS, MS, DOZY, AM, KAN, YW (1978) Diagnosis of homozygous alpha-thalassaemia in cultured amniotic-fluid fibroblasts. N Engl J Med 298:

669-670.

YUNIS, JJ (1981) Mid-prophase human chromosomes. The attainment of 2 000 bands. Hum Genet 56: 293-298.

ZIMMER, EA, MARTIN, SL, BEVERLEY, SM, KAN, YW, WILSON, AC (1980) Rapid duplication and loss of genes coding for the alpha-chains of haemoglobin. Proc Natl Acad Sci USA 77: 2158-2162.