The EMBO Journal vol.9 no.1 pp.25 – 34, 1990

Analysis of human chromosome 21: correlation of physical and cytogenetic maps; gene and CpG island distributions

Katheleen Gardiner, Michel Horisberger[1], Jan Kraus[2], Umadevi Tantravahi[3], Julie Korenberg[4], Veena Rao[5], Shyam Reddy[6] and David Patterson

Eleanor Roosevelt Institute for Cancer Research, Denver, CO, USA, [1]Biotechnology, Ciba-Geigy Ltd, CH-4002, Basle, Switzerland, [2]Department of Pediatrics and Cellular/Structural Biology, University of Colorado Health Sciences Center, Denver, CO, [3]Genetics Division and Mental Retardation Center, Children's Hospital, Boston, MA, [4]Medical Genetics, Cedar-Sinai Medical Hospital, Los Angeles, CA, [5]Fels Institute for Cancer Research, Philadelphia, PA and [6]National Cancer Institute, Frederick Cancer Research Facility, Frederick, MD, USA

Communicated by G.Benardi

Human chromosome 21 has been analyzed by pulsed-field gel electrophoresis using somatic cell hybrids containing limited regions of the chromosome and >60 unique sequence probes. Thirty-three independent *Not*I fragments have been identified, totalling 43 million bp. This must account for essentially the entire long arm, and therefore gaps remaining in the map must be small. The extent of the pulsed-field map has allowed the direct correlation of the physical map with the cytogenetic map: translocation breakpoints can be unambiguously positioned along the long arm and the distances between them measured in base pairs. Three breakpoints have been identified, providing physical confirmation of cytogenetic landmarks. Information on sequence organization has been obtained: (i) 60% of the unique sequence probes are located within 11 physical linkage groups which can be contained in only 20% of the long arm; (ii) 9/21 genes are clustered within 4%; (iii) translocation breakpoints appear to occur within CpG island regions, making their identification difficult by pulsed-field techniques. This analysis contributes to the human genome mapping effort, and provides information to guide the rapid investigation of the biology of chromosome 21.

Key words: chromosome 21/CpG islands/gene maps/pulsed-field mapping/translocation breakpoints

Introduction

One of the aims in developing physical maps of complex genomes is to describe chromosomal sequence organization and to relate features of this organization to biological functions. As such, physical maps must be more than simple ordering of restriction sites at measured distances through a chromosomal region. While the most detailed physical map may be a complete sequence, we still need to understand how sequence relates to function at all levels of genome organization. Interpreting the information contained within complete sequences of large regions will probably require analyses like those presented here to reveal important biological features. For example, the basis for cytogenetically observable banding patterns is not well understood—while R bands are known to be relatively GC and gene rich, how much variation there may be within or between bands, and its significance, is not known. G bands, on the other hand, are postulated to be gene poor, but it is not known if this DNA holds other functions, or if large parts of it are truly 'junk' DNA (Cummings, 1978; Goldman *et al.*, 1984; Ikemura and Aota, 1988). No details are known regarding the organization of genes within the genome, indeed even the number of genes can only be crudely estimated (Ohno, 1986). While the distribution of methylation patterns has been studied for some genes, and is known to vary, again the whole genome picture is missing (Razin and Riggs, 1980; Bird, 1986; Silva and White, 1988). Little is known about the causes of, or predisposition to, the chromosomal rearrangements that are seen in malignant and non-malignant genetic disease. Although a few breakpoints, such as those involving immunoglobulin genes, have been sequenced locally, it cannot be assumed that the long-range environment has no effect on these or other rearrangements (Mitelman and Heim, 1988; Haluska *et al.*, 1987). These and other problems may be well studied at the level of lower resolution physical maps.

Different kinds of chromosomal maps, and the strategies used to produce them, will yield different sets of information. This information ideally will aid in developing new strategies for analysis, pinpoint regions of high biological importance and hasten success in solving genetic problems, such as malignancy and inherited genetic diseases.

For many mapping efforts, chromosome 21 has served as the prototype for the human genome. Recently, a partial physical map of this chromosome was reported (Gardiner *et al.*, 1988). In that report, speculations were raised regarding certain aspects of sequence organization, including the distributions of unique sequences, rare restriction enzyme sites and chromosomal breakpoints. Findings correlated well with those obtained by other approaches (Bernardi *et al.*, 1985; Bird *et al.*, 1985; Bird, 1986; Zerial *et al.*, 1986; Korenberg and Rykowski, 1988). Here, we present a map of considerably more detail, one that orders 15 breakpoints, or clusters of breakpoints, at physically measured distances, and links, by pulsed-field analysis, 60% of the probes used. In total, this map accounts for 43 million bp in *Not*I fragments, and because this is likely to approach spanning the entire long arm (Korenberg and Engels, 1978), the gaps remaining in the map must be small. This has, therefore, allowed the placement, relative to cytogenetic band patterns, of breakpoints, physical linkage groups, unique sequences and CpG islands. In addition to this correlation of physical and cytogenetic maps, partial CpG island and gene maps are emerging. In some regions, notably 21q22.3, CpG islands containing sites for many 'rare' cutting restriction enzymes appear to define clusters of genes, not single genes, and gene density is locally very high. Pulsed-field gel

analysis alone frequently gives a good indication of the sequence content of a region, whether it is within a Giemsa light band or a dark band. However, there is significant heterogeneity among both G and R bands regarding rare CpG islands and gene density.

Results

Hybrid cell panel

The complete panel of rearranged chromosomes 21 used in this study is shown in Figure 1, with the approximate regions that are retained in each hybrid indicated. These regions were originally estimated by cytogenetics, and refinements made after molecular analysis. For example, the order of the 3;21 and the 1;21 breakpoints, and the order and extent of the deletions could not be determined by the resolution obtainable with cytogenetics alone. The figure shows the composite results of both kinds of analysis.

The 14 chromosomes in the panel have been divided into three groups, based upon the region retained and the approximate band location of breaks. The A group retains 21q22, and has breakpoints within 21cen>21q21. The BC group generally retains 21pter>21q22.1, and has breakpoints in the distal part of 21q22. The lower ACEM breakpoint defines the approximate boundary between q21 and q22.

The new members of the chromosome 21 hybrid clone panel are JC6, MRC2 and 6918, and form the D group. In each case, the derivative chromosome 21 is the result of a complex rearrangement, involving other than a simple translocation. Cytogenetic analysis suggested that JC6 and 6918 were deleted within the 21q21 region; MRC2 is a ring chromosome 21. Each chromosome contains one breakpoint near the centromere; JC6 has a second near the boundary of 21q21−q22. MRC2 and 6918 each have two breakpoints within 21q22.

In summary, breakpoints appear to occur as follows: five within 21q11/proximal q21, four within 21q21, four within 21q22.1 and six within 21q22.2−22.3. These positions are, of course, approximate, as discussed below. The 19 breakpoints potentially divide the long arm into 20 regions.

Regional localization of unique sequences and ordering of breakpoints

*Eco*RI and/or *Hin*dIII digests of DNA from each of the 14 members of the hybrid clone panel, as well as from normal controls, were examined for the presence or absence of 67 unique sequence probes. Tables I and II, and Figure 2, column I show a compilation of: (i) the previous mapping data (obtained with 30 probes, and the group A and BC hybrids, Gardiner *et al.*, 1988); (ii) that obtained from the new, group D, hybrids (with all 67 probes); and (iii) that obtained from the previous hybrids with 37 additional probes. This cumulative analysis has differentiated, in total, 11 individual and four clusters of breakpoints, thus defining 16 long arm regions. Breakpoints for the 3;21 and the 1;21, previously not separable, are now distinguished by B14. The region 21q22.1, previously largely uncharacterized, has been divided into four segments by the lower JC6 breakpoint, and breakpoints from MRC2 and 6918.

This analysis also led to the unexpected conclusion that both MRC2 and 6918 are double deletions; both are missing large amounts of 21q22 (distal to D21S65, and approximately ERG respectively), in addition to the interstitial deletion

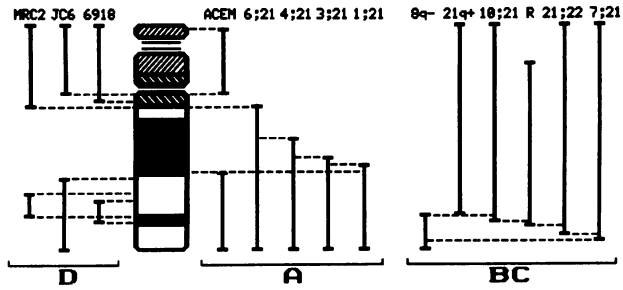

Fig. 1. Chromosome 21 hybrid clone panel. The vertical lines to the left and right of the schematic of the chromosome indicate the regions of the chromosome retained in each of the hybrids listed at the top of the figure. Abbreviated names of the hybrids are used (Gardiner *et al.*, 1988). Brackets at the bottom indicate hybrids that are grouped together based upon region retained and location of the breakpoints.

Table I. Regional localization of unique sequence probes

Probe	Region	Probe	Region	Probe	Region
B37	A2	D51a	A4	26C[a]	A4
C81	A4	C9	A5	D32	A5
C8	A7	D6	D1	B17	D2
B7	D2	C40	D2	B83	D2
B15	D3	D31	B1	D3	B2

[a]HGM locus symbol D21S26.

within 21q21. This is an interesting observation in particular for 6918. This chromosome and that of JC6 had been obtained from patients characterized cytogenetically as possessing the same abnormality, a 21q21 deletion, yet presenting radically different phenotypes: 6918 resembles monosomy 21 (Reynolds *et al.*, 1985), and JC6 is apparently normal (Korenberg *et al.*, 1986). This analysis suggests a possible molecular basis for the observed differences. If the chromosomes present in the hybrids are faithful copies of those in the patients from which they were derived (and the deleted material was not translocated to other chromosomes in the patients), then 6918 is indeed monosomic for a large part of the gene-rich region of the chromosome, 21q22.3, and also a substantial fraction of 21q22.1. JC6, on the other hand, is deleted only for sequences within 21q21, a region that appears relatively low in transcribed sequences (see below). The differences in phenotype could, therefore, be directly ascribed to differences in gene dosage.

In summary, there appears to be a cluster of breaks in the centromere proximal region of q11−q21, while, with the inclusion of the three deletion chromosomes, the distribution now seems more uniform throughout 21q22. While several breakpoints have not been differentiated in this analysis, there is no reason to believe at this point that any are identical. As previously noted, 21q21 remains relatively free of breakpoints.

Sequence distribution

Examination of Tables I and II, and Figure 2, reaffirms observations made previously regarding the non-uniform distribution of the unique and transcribed sequences used as probes (Korenberg *et al.*, 1987; Gardiner *et al.*, 1988; Tantravahi *et al.*, 1988). Overall, twice as many unique sequences map to 21q22 as to 21q21. Fully one-third unambiguously map within 21q22.3, which is believed to

Table II. Pulsed-field restriction fragment sizes (kb)

HMG9	Probe	*Not*I	*Bss*HII	*Mlu*I	*Nru*I	*Sac*II	*Sfi*I	Region
								Centromere
Locus								A1
							Acem,JC6	
D21S16	E9	1500	900	1600		750	600	A2
							6;21,6918	
D21S13	**M21**	1500,700	900,700	1600		700,750	800,400	
D21S4	233F	1200	1150,1000	2700	1950	1200	500	A3
							MRC2	
D21S52	511-1H	1800	1700			1700	850	
D21S59	552-2B	1850	1700	2000		1700	700	A4
D21S110	4U	1850	1700			1700		
							4;21	
D21S11	236B	2350	1800		2350		100	A5
D21S1	228C	2350	1800		2350	1500,1000		
D21S95	**JG90***	3000	2500	2800	3000		350	
							3;21	
	B14[1]	2650,1800	2150				350	A6
							1;21	
D21S111	21pcq	1500*	1125	750	950	300	200	
D21S12	267C	1500*	1125	750	950	300	200	
D21S99	JG108	1500	1125			300	250	A7
	APP	1500*	1125	2500	2500	500	100	
D21S8	245D	6500	6000		7000	500	400	
	B79[1]	1950	2000	2250	2250	2000	825	
							Acem	
D21S54	513-5H	2100	1650	1800	2100	1650		D1
							JC6	
	B88[1]	1600	600	1600	1800	300	250	
D21S93	**JG77**	2000	1800	1500	2000		700,875	D2
	C34[1]	1900	1300	1900	1400	500	400	
	C43[1]	1900	1300	1900			400	
							MRC2	
	B18[1]	1800	1500	2500	2500		400	
D21S63	517-1.7R	400	400	500*		600	50	D3
	SOD1	200	200	350*		100	150	
							6918	
D21S58	**524-5P**	775	100	400	800	100	50	D4
D21S65	525-5H	825	400			200	50	
							8;21,MRC2	
D21S15	**H8**	1125	1050	2000	1800	1125	400	
D21S55	**518**-8B	1500	500*	1500	550	400	50	
D21S60	CP21G1	1500	750*	1500				B1
	ERG	1500	750*				130	
							10;21,6918	
D21S57	**523**-10B	700	500	750	550	350	75	
	ets2	700	500	750	550	350	55	B2
D21S3	**231C**	700	500	750	550	350	60	Ring
D21S71	519-9R	700	500	750	550	350	80	
D21S64	551-8P	775	600	1225,1125	700	400	380	
	MX1	775	600	1225,1125	700	400	380	
	MX2	775	600	1225,1125	700	400	380	
	CBS	775	600	1225,1125	700	400	380	
D21S53	512-5R	775	600	1225,1125	700	400	250	
D21S15	**E8**	1500	1225	1225,1125	1275		200	C1
D21S19	B3	775	300	300	250	450	50	
	BCEI	775	300	300	230			
D21S56	520-10R	775	40	250	270	450		
	PFKL	50	50	700	900		30	
	CRYA	325	380	975	950	350	100	
	E73[1]	550	400				100	
	D22[1]	400	200	500	1400	200		
D21S91	JG63	1125,400	400	400	680	290	100	
D21S25	10.2	200	100	400	350	60	30	
							21;22	
	LFA	100	100	150	150,300		50	C2
	COL6A1	750	120	500	825,500	130	40	
	COL6A2	750	150	500	825,500	150	20	
							7;21	
							Telomere	C3

Information obtained for 52 probes with the enzymes *Not*I, *Bss*HII, *Mlu*I, *Nru*I, *Sac*II and *Sfi*I is given. Horizontal dashed lines indicate chromosomal breakpoints defining the regions as described in the text and Figure 2. Probes for which there is only regional localization data, i.e. no ordering from either genetic linkage (Tanzi *et al.*, 1988) or from pulsed-field analysis, are listed in an arbitrary order within the region. Brackets indicate physical linkage established in this study (* indicates use of methylation differences or polymorphisms); probes in bold type are expressed sequences (anonymous sequences, Neve *et al.*, 1986 R.Neve, unpublished results; Stephani *et al.*, 1988). HGM9 locus: official human gene mapping probe designation (Pearson *et al.*, 1987).

[1]Donated by U.Tantrahavi.

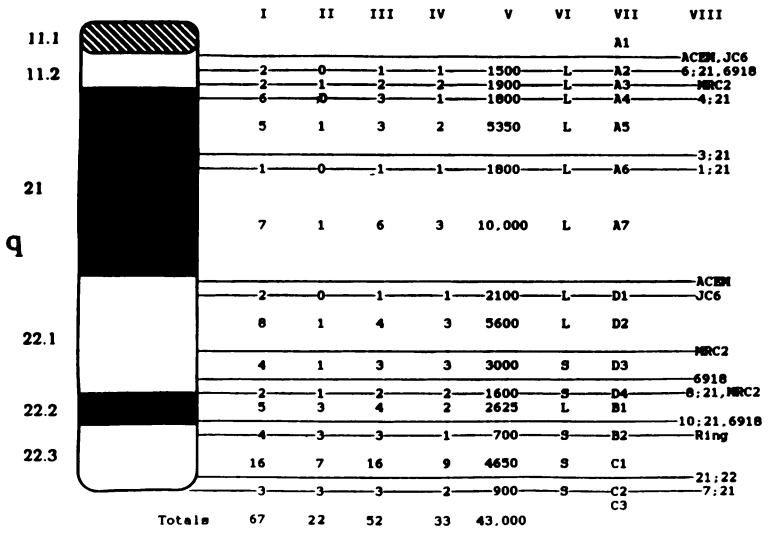

Fig. 2. Summary of mapping information by region. Horizontal lines indicate the positions of the breakpoints of the chromosomes from the hybrids listed in column VIII. Positions were determined by correlating the pulsed-field map with the cytogenetic map. **Column I**, the number of unique sequences (from Tables I and II) localized to each region defined by the breakpoints; **II**, the number of these sequences known to be transcribed; **III**, number of sequences from I analyzed by PFG (Table II); **IV**, number of independent *Not*I fragments obtained by PFG; **V**, total amount of DNA in *Not*I fragments; **VI**, relative sizes of PFG fragments, L (large): 65–100% of the fragments were >200 kb with *Sfi*I, >1000 kb with all other enzymes tested, S (small), the converse; **VII**, region name; **VIII**, abbreviated hybrid name.

contain only 20% of the DNA. These localization data also suggest that regions of human chromosome 21 syntenic with mouse chromosomes 16, 17 and 10 can remain contiguous. Probes E9, 511, APP, 524, SOD1, ETS2, MX1 and MX2 have been mapped to mouse chromosome 16; CBS, 520 and CRYA to mouse 17; and LFA, COLA1 and COLA2 to mouse 10 (Staeheli *et al.*, 1986; Reeves *et al.*, 1987; MacDonald *et al.*, 1988; Nadeau *et al.*, 1989).

Of the known transcribed sequences (Table II and Figure 2, column II), the majority of which are identified genes (see Table II), 85% are within 21q22, with 75% distal to SOD1. 21q22.3 is clearly gene rich, containing 13/21 genes used here, in addition to others such as S100 (Allore *et al.*, 1988). These numbers are somewhat artificial in that not all of these sequences have been examined for expression, and, in any case, lack of observed expression may merely be a result of the wrong tissue, the wrong stage in development or insufficient sensitivity in detection being used in the study. It is, however, noteworthy that while the number of sequences localized here has doubled over the previous report, the pattern of distribution has remained essentially the same. This argues for the relevance of the observation.

Pulsed-field analysis

A subset of probes from Figure 2 have been analyzed by pulsed-field gel electrophoresis. Seven enzymes were routinely used: *Not*I, *Bss*HII, *Mlu*I, *Nru*I and *Sac*II, all of which contain two CpG dinucleotides in their recognition sequences, and give, on average, the largest restriction fragments; *Sfi*I, with an 8 bp G + C recognition sequence but no CpGs, which results in fragments generally much smaller than these; and *Sal*I, with one CpG, which (in our hands) naturally gives partial digests, possibly due to methylation patterns.

Four different electrophoretic analyses were done, as required, and examples are shown in Figure 3.

(i) Digests of normal human DNA (with all enzymes listed above except *Sal*I) were separated under conditions that optimally resolved fragments between 50 and 1500 kb, and the size of hybridization bands determined for probes of interest. (For an example of this resolution, see Figure 3a.) Pulse times were chosen to optimize resolution of the size class studied (Gardiner and Patterson, 1989).

(ii) If this range of resolution did not give interpretable hybridization signals due to the large size of the fragments (which frequently happened in certain regions of the chromosome, see below), digests were then examined under conditions that resolve fragments between 1500 and >6000 kb, and hybridization repeated (Figure 3b). Once probes in the same or adjacent regions had been examined in this way, many physical linkage groups were obvious, simply by comparing the pulsed-field 'fingerprints' (Table II). To confirm linkage, probes within a group were always sequentially hybridized to the same filter, to eliminate any ambiguities due to gel-to-gel mobility variations. Because of the lack of accurate markers between 1600 and 3000 kb, fragment sizes in this range are estimates; below 1600 kb sizes can be determined, conservatively, within 10%.

(iii) Several different DNAs, from both fresh lymphocytes and cultured cells, were digested with the same enzyme, and hybridization with a single probe compared. Frequently, different sized fragments were observed among the DNAs, with the larger fragments most often found in the hybrid cell DNAs (see Figure 3a). Such differences could be due to a sequence change removing an enzyme site, or they could also result from methylation differences due to the presence of CpGs in the enzyme recognition sequences (Drahovsky *et al.*, 1980; Reis and Goldstein, 1982; Nelson and McClelland, 1987). This latter is likely in the experiment shown in Figure 3(a). Probe D21S55 recognizes an intense band of 500 kb in each of the five human DNAs and Hy2 and 3, in addition to (except in Hy3) a faint, larger band at ~1200 kb. In Hy4, the hybridization is solely to the 1200 kb fragment. Variations in the methylation of a *Bss*HII site within the 1200 kb fragment could give rise to these results. Hy4 would be completely methylated at this internal

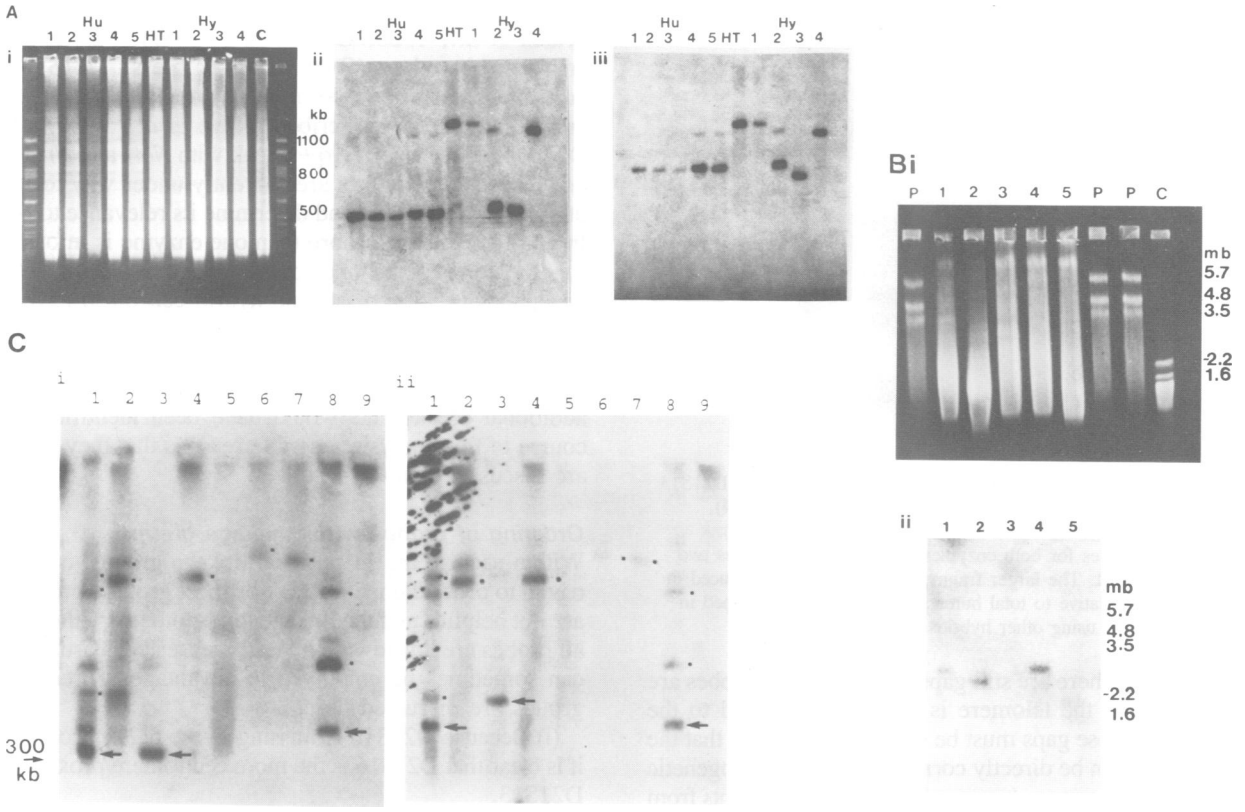

Fig. 3. Examples of pulsed-field gel analysis. (**A**) Resolution of fragments up to 1500 kb and variation in restriction fragment sizes between different sources of DNA. (**i**) Human and hybrid cell line DNA was digested with *Bss*HII and electrophoresed for 18 h with a 60 s pulse time. Outside lanes are *S.cerevisiae* chromosome markers. **Lanes Hu 1–3**, DNA isolated from different samples of human lymphocytes; **lanes 4** and **5**, from fibroblasts; HT, HT1080; **lanes Hy 1–4**, from hybrids 8q–, 10;21, 4;21 and R2-10W; C, CHO ade-C. (**ii**) The gel in (i), after transfer, hybridized with D21S55; (**iii**) the same filter stripped and rehybridized with D21S60. (**B**) Resolution of fragments >1500 kb. (**i**) The chromosomes from *S.pombe* (P) and *S.cerevisiae* (C), and human lymphocyte DNA digested with, **lanes 1–5**, *Not*I, *Bss*HII, *Mlu*I, *Nru*I and *Sac*II. (**ii**) Hybridization with the probe D21S95. These bands are too large to have been observable under the gel conditions in (a). (**C**) *Sal*I partial digestions. DNA from cultured fibroblasts, HT1080, hybrids 4;21, 6;21, 10;21, 7;21, 21;22, R2-10W and 153E7b was digested with *Sal*I and resolved as in (a). (**i**) and (**ii**) show hybridization of this material with the probes D21S11 and D21S95 respectively. Dots indicate bands common to the two probes and serve to establish physical linkage. Arrows in (i) indicate bands that are common to D21S11 and a third probe, D21S1, previously linked to D21S11 (see Table II), but only at distances >1000 kb. Arrows in (ii) are bands unique to D21S95.

site, while HT1080 and Hy1 are completely methylated at one more site, or are genuinely polymorphic. Rehybridization of this filter with probe D21S60 shows the same large fragments in the same lanes as with D21S55, but a fragment of 750 kb in the DNAs previously showing a 500 kb fragment. It is concluded that D21S55 and D21S60 are on adjacent *Bss*HII fragments of 500 and 750 kb respectively, with the interstitial site variably methylated in some DNA sources. However, whether due to methylation or to sequence changes, such differences can sometimes be used to substantiate linkage and occasionally, they can be attributed to the identification of a breakpoint (see Figure 4).

(iv) Several different DNAs were digested with *Sal*I, giving hybridization patterns that are quite complex, with multiple bands of varying intensity (Figure 3c). Because *Sal*I sites occur more frequently than those of the other enzymes used, it can be useful in determining a minimum distance between two probes (e.g. D21S11 and D21S1, Figure 3c). It can also sometimes be used to cross CpG islands, demonstrating linkage that cannot be observed with the enzymes whose recognition sites are of higher GC and CpG content (e.g. D21S95 to D21S11/S1).

The information obtained for 52 probes is discussed here and shown in Table II and Figures 2–4. (i) As noted previously, probes localized to 21q22 identified significantly smaller pulsed-field fragments than those from q21 (Gardiner *et al.*, 1988). This is clearly demonstrated by the necessity of using the large-fragment (1500–6000 kb) resolution conditions for probes in the latter region (see Table II, and Figure 2, column VI). (ii) Six new linkage groups have been identified, and new probes added to three of the previous five. A total of 33 probes (out of the 52 studied) are now linked, and the 11 groups are spread throughout the length of the chromosome. For the majority of these, straightforward pulsed-field 'fingerprinting' was sufficient to demonstrate linkage. *Sal*I digestions were used to determine that D21S1 and D21S11 were only 300 kb apart (the smallest common fragment), and also to link these to D21S95 (Figure 3c). The *Bss*HII polymorphism helped to confirm the linkage of D21S60 and D21S55 (Figure 3a), and an *Mlu*I polymorphism links SOD and D21S63 (data not shown). (iii) A total of 33 independent *Not*I fragments were identified, and account for 43 million bp of DNA. This is a significant number, because the upper estimates for chromosome 21 suggest it contains 60 million bp of DNA with two-thirds to three-quarters of this within the long arm (Korenberg and Engels, 1978). This suggests that these probes, and *Not*I fragments, must approach spanning the entire long arm.

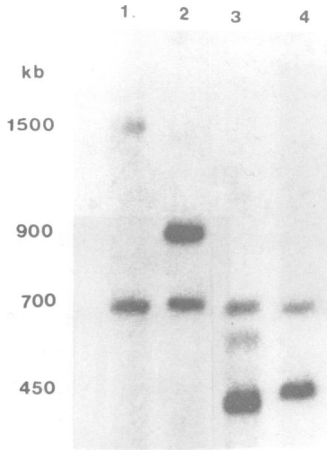

Fig. 4. Identification of the 6;21 breakpoint. Hybridization of D21S13 to human DNA (**lanes 1** and **2**) and 6;21 DNA (**lanes 3** and **4**), digested with *Not*I (**lanes 1** and **3**) and *Bss*HII (**lanes 2** and **4**). D21S13 contains sites for both enzymes and therefore identifies two bands in each digest. The larger fragment in both cases is reduced in the hybrid DNA relative to total human. This result is confirmed in similar experiments using other hybrids.

Also, although there are still gaps in the map (all probes are not linked, and the telomere is not directly linked to the centromere), these gaps must be small. This implies that the physical map can be directly correlated with the cytogenetic map. Figure 2 shows such a correlation; the breakpoints from the clone panel have been positioned along the long arm so that the space between them is proportional to the amount of DNA measured in *Not*I fragments (column V) (for this representation, it has been assumed that the DNA is uniformly distributed throughout the length of the metaphase chromosome). Most breakpoints fall in accord with the cytogenetic determinations made with the hybrid cell lines. Landmarks include the lower ACEM break being near the 21q21−21q22 boundary; the R2-10W, being near the 21q22.2−21q22.3 boundary; the 6;21 being near the centromere. The 8;21 breakpoint moves up from the q22.2−q22.3 boundary to the q22.1−q22.2 boundary, not untenable given the small size of band q22.2 and its translocation to the vicinity of a small dark band on chromosome 8.

Breakpoint identification

Three breakpoints from members of the clone panel have so far been directly identified by pulsed-field analysis. The R2-10W breakpoint has already been described (probe D21S3 spans the break, Wong *et al.*, 1988). ETS2 is physically linked to D21S3 (Table II), and therefore it not only also identifies the R2-10W breakpoint on pulsed-field gels (Gardiner *et al.*, 1988), but the map indicates that it must also be >2000 kb from the 8;21 breakpoint, a translocation frequently seen in acute myelogenous leukemia (AML) subtype M2 (Fourth International Workshop on Chromosomes in Leukemia, 1984). It is unlikely therefore that rearrangements of the ETS2 gene are involved directly in the etiology of this disease. Probe D21S13 detects the 6;21 breakpoint with *Not*I, *Bss*HII and *Sac*II. D21S13 identifies two fragments with each of these enzymes (Table II), because it contains sites for each. As shown in Figure 4, only one fragment is rearranged in the 6;21 chromosome. This chromosome serves to order D21S16 and D21S13, relevant because of their linkage to familial Alzheimer's disease (Tanzi *et al.*, 1988; Goate *et al.*, 1989). This has been confirmed by others (Owen *et al.*, in press). D21S13 is of particular interest because it clearly contains a CpG island and because it is transcribed (Neve *et al.*, 1986). Probe D21S65 identifies the 21q+ break with *Not*I and *Nru*I (data not shown), and efforts are currently under way to isolate this junction fragment and determine its relevance to AML. In all of these cases, more than one enzyme is informative for the breakpoint and many different DNA sources (lymphocytes, fibroblasts and hybrid cell lines) have been compared, lessening the likelihood that it is simply a polymorphism that is being detected. Given the current coverage of the chromosome, it was anticipated that additional breakpoints would have been identified in the course of this analysis; possible reasons that they were not are discussed below.

Ordering of probes within linkage groups

Within each of the 11 physical linkage groups, it is rarely trivial to order the members. For most groups, breakpoints are not helpful, and the linkage tends to be very thorough—all probes are linked with all enzymes tested. Partial digests can sometimes be employed to advantage. Three linkage groups are discussed here.

(i) Because D21S16 is missing in the 6;21 chromosome, it is clear that D21S16 is the more centromere proximal than D21S13.

(ii) For the ETS2 cluster at the ring break, the order must be (ETS2,D21S57)-D21S3-D21S71, as described previously. In partial *Sfi*I digests, D21S71 was found to share bands with ETS2 but not D21S57; ETS2 shares bands with both (data not shown). This suggests the order must be cen-D21S57-ETS2-D21S3-D21S71.

(iii) The four probes in the APP cluster can be ordered less directly. D21S111 was isolated as the sequence adjacent to and distal to the upper break in the ring. [The ring, as described by Wong *et al.* (1988), is believed to have been formed from a dicentric 21 composed essentially of two long arms; this chromosome then broke, in one arm in 21q22.3 (in the middle of D21S3) and, on the other arm, 21q21, immediately proximal to D21S111. The ring was formed by the joining of the two breaks.] If the other members of this cluster—APP, D21S12 and D21S99—were proximal to D21S111, they would each detect two fragments in the ring, normal fragments from the long arm with the break in q22.3 and altered fragments from the long arm portion broken in q21. Only the normal fragments are found (data not shown), indicating that all three must be distal to D21S111. Furthermore, D21S12, D21S111 and D21S99 are all on the same *Sac*II fragment, and D21S12 and D21S111 also on the same *Sfi*I fragment, while APP is on different ones. This implies that the general order must be cen-D21S111-D21S12-D21S99-APP. In addition D21S8 has been very tightly linked in genetic maps to APP (Tanzi *et al.*, 1988). Because D21S8 also does not detect rearrangements in the ring, it is more likely that it also is distal to APP.

Discussion

The physical map of chromosome 21 described here distinguishes 15 breakpoints or breakpoint clusters, and localizes 67 unique sequence probes among 16 regions.

Pulsed-field analysis has identified 11 physical linkage groups and >40 million bp of DNA contained in 33 *Not*I fragments. While all the fragments have not been linked, the gaps this leaves in the map are likely to be small, because of the estimated size of the long arm. The extent of the map allows placement of breakpoints at measured distances along the long arm, thus correlating cytogenetically observable landmarks with molecular ones. Several features of genome organization, pertaining to the clustering of genes, the nature of CpG islands and translocation breakpoints, as they are represented by chromosome 21, are suggested. Some of these conclusions are a direct result of the strategy used and the enzymes selected.

Unique sequence distributions

That approximately two-thirds of the unique sequences examined localize to one-half of the long arm (q22) has been noted previously (Korenberg *et al.*, 1986; Gardiner *et al.*, 1988; Tantravahi *et al.*, 1988). It is worth reiterating that no strong conclusion regarding unique sequence organization overall in the genome can be based on these data. This is due to possible artifacts arising from the methods used in obtaining clones, artifacts due to restriction enzyme site preferences, methylation patterns, poisonous sequences or screening protocols. It also must be kept in mind, however, that even if artifactual, its non-random nature suggests that it reflects some aspect of genome organization, at least as it exists on chromosome 21.

Gene organization

It has become increasingly clear that genes are neither randomly nor uniformly distributed throughout the genome [perusal of the table of cloned human genes from HGM9 suggests that two-thirds to three-quarters of genes are located in Giemsa light bands (Pearson *et al.*, 1987); see also Goldman *et al.* (1984), Kuhn *et al.* (1987), Abe *et al.* (1988), Ikemura and Aota (1988)]. The extent of this non-uniformity, as demonstrated here, is, however, perhaps startling. Consider the four physical linkage groups within 21q22.3. Together, they involve a total of 14 unique sequences, of which seven are cDNAs of identified products (ETS2, MX1, MX2, CBS, COL6A1, COL6A2 and BCEI) and two more are also known to be expressed (R.Neve, unpublished). If one considers the smallest fragment defining each group, these sequences are contained within a maximum of 1700 kb, ~4% of the chromosome. A cluster of seven genes within 170 kb has been observed in the mouse MHC H-2K region (Abe *et al.*, 1988); however, the analysis here spans a larger chromosomal region and includes a less obviously related collection of genes. That nine expressed sequences, out of a total of only 22 so far examined, are found within so limited a region of the chromosome is strongly suggestive of another, at least local, level of genome organization.

Distribution of restriction sites

The recognition sites of the enzymes useful for pulsed-field analysis contain the dinucleotide CpG and/or are high in G + C. The observed distribution of these sites can therefore be used to infer characteristics of local GC content, the presence of HTF islands and the occurrence of methylation. It was shown, for example, in the previous report that these enzymes cut more frequently in the Giemsa light band, 21q22, than in the dark band, 21q21. This is consistent with the use of base-specific dyes in the generation of cytogenetic banding patterns (Cummings, 1978) and with isochore data (reviewed in Bernardi, 1990). The increased detail presented here supports these observations in general, but demonstrates that there exists considerable inter- and intra-band heterogeneity, in particular in Giemsa light bands. For example, 21q22.3 is enriched for these enzyme sites, but the region surrounding the probe E8 is an exception; a more striking dichotomy occurs in q22.1 where the distal part resembles q22.3, but the proximal part resembles 21q21, a dark band. This heterogeneity may be reflecting a higher resolution band pattern (Yunis, 1981) and is explored in detail in the compositional map for chromosome 21 (K.Gardiner *et al.*, in preparation).

CpG islands

Clustering of rare restriction sites is accepted as defining the 5′ regions of genes, but this may be limiting the observations (Gardiner-Garden and Frommer, 1987; Lindsay and Bird, 1987). HTF islands were defined using *Msp*I, *Hpa*II and *Hha*I, enzymes with 4 bp recognition sequences (Cooper *et al.*, 1983). Clearly, the 6 and 8 bp recognition sequences used in pulsed-field work can recognize only subsets of these sites. Indeed this analysis (consider the physical linkage groups in q22.3 discussed above) suggests that some rare sites frequently do not define individual genes, but rather clusters of genes. In particular, the enzyme sites used here (except for *Sfi*I and *Sal*I) contain two CpGs, and thus are the most rare. This suggests that there may exist two classes of CpG islands, the more frequent *Msp*I/*Hpa*II islands defining individual genes, and rare PFG islands defining gene clusters.

Methylation

All CpG dinucleotides can be viewed as potential sites for methylation, but only a few per cent are used by the cell (reviewed by Razin and Riggs, 1980). If the partial digestion exhibited by *Sal*I is due to methylation (Figure 3c), then these rather more common sites seem commonly and variably methylated in lymphocytes as well as in cultured cells. This appears true in all regions of the chromosome, although it may be less extensive in Giemsa light bands, as suggested by the less complex hybridization patterns generated there (data not shown). With other enzymes, lymphocyte and fibroblast DNA exhibited little or no detectable methylation (or polymorphisms). However, DNA from cultured cell lines frequently gave two or more bands, generally equal to or larger in size than the lymphocyte bands, and varying among cell lines in intensity. It is interesting that the ETS2 gene never gave evidence of methylation in eight different cell lines with six enzymes, while SOD1 frequently did. More detailed analysis is necessary to determine the relevance of any putative methylation patterns.

Gaps in the map and breakpoint identification

Because of the amount of DNA accounted for in this analysis, the gaps remaining in this physical linkage map must be small. It follows also (because complete digests were used) that these gaps must occur where the rare restriction sites cluster; these regions would be locally high in CpG and G + C content and might range in size from less than a kb to a few tens of kb. While these gaps no doubt could be spanned using partial digests, or additional enzymes with

possibly somewhat different restriction site distributions (e.g. sites containing a single CpG), their very existence gives additional information. Consider that only three of 19 translocation breakpoints have been detected in this analysis. Where are the remainder? Logic suggests that they must be within the gaps in the map, i.e. within in rare CpG islands. Some support for this suggestion is obtained by inspection of Table II. Two breakpoints occur in or near transcribed sequences, known to contain CpG islands from restriction site analysis. But more interesting is that, of the 11 physical linkage groups, the map shows seven bounded on one or both sides directly by a breakpoint. Postulating that breakpoints tend to occur in regions of locally high G + C content also explains why they occur more frequently in Giemsa light bands than dark bands (Holmberg and Jonasson, 1973; Mitelman and Heim, 1988). If such susceptible regions are extensive (i.e. tens of kb), they may not be evident from the sequencing of 1−2 kb about a breakpoint, as has generally been done.

Correlations with genetic linkage maps

Partial genetic linkage maps of the long arm of chromosome 21 have been published (Tanzi *et al.*, 1988; Goate *et al.*, 1989; Warren *et al.*, 1989), using many of the probes analyzed here. Because of the significant differences between male and female maps, it is not possible to make meaningful correlations between physical and genetic distances, other than to observe that in some cases, e.g. between EST2 and D21S3, the physical distance is quite small in which measurable recombination has taken place (<300 kb, with >2% recombination, Warren *et al.*, 1989). On the other hand, there are some conflicts between the genetic map and the physical map presented here. Specifically, Warren *et al.* have placed BCEI distal to COLA1, whereas the 21;22 breakpoint would dictate that the reverse order must be correct. In addition, Warren *et al.* give the order D21S8-APP-D21S111. The data presented above suggest the order D21S111-APP-D21S8. The probabilities for reverse genetic orders between BCEI and COLA1, and D21S8 and APP were <1000, which possibly accounts for those discrepancies.

Phenotypic correlations

With a map spanning the entire long arm, it becomes possible to attempt direct correlations between observed phenotypes and deletions or rearrangements of specific gene regions. This is of relevance in cases of apparent monosomy 21, such as the 6918 patient studied here. Because the regions missing from the abnormal 21 can be well defined, the question of their presence, and possible other location, in the patient's genome can be addressed. This could be done by *in situ* fluorescence, using a mixture of probes from 21q21 and from 21q22.3 (Lichter *et al.*, 1988). Hybridization to chromosomes other than the normal 21 would be easily detected. Such investigations would help to establish if monosomy 21 is a *bona fide* syndrome. In another deletion chromosome, JC6, if a similar analysis yields negative results, it would be strong support for the suggestion that 21q21 is indeed gene poor, and that its deletion (or partial deletion) is compatible with a normal phenotype. Similarly, in cases of Down's syndrome due to translocations or small duplications, where only a portion of the chromosome is present in triplicate (Sinet *et al.*, 1976; Perez-Castello *et al.*, 1983), it now is possible to define more precisely the region responsible (McCormick *et al.*, 1989; Rahmani *et al.*, 1989). Again, in 'modified' Down's syndrome (Habedank and Rodewald, 1982; Jenkins *et al.*, 1983), it may be possible to assign specific aspects of the phenotype to specific regions. Such investigations are difficult due to the complexity and the variability of the Down's syndrome phenotype, and for any conclusions obtained from them to be reliably informative, it is absolutely essential to start from a map as detailed as possible.

Conclusion

Pulsed-field analysis of chromosome 21 has identified a collection of *Not*I fragments and sets of linked probes. It has allowed the direct correlation of cytogenetic landmarks with molecular ones, and made possible an unambiguous comparison of the features of Giemsa light and dark bands. It has shown that, while Giemsa light bands are generally enriched for rare restriction sites and genes, there is significant heterogeneity within and between bands. 21q22.3 stands out as possessing several unique features. It contains ~6000 bp, or 15%, of the DNA, half the expressed sequences and one-third of the unique sequences examined. It is also known that 40% of the recombination occurs here (Tanzi *et al.*, 1988). Pulsed-field restriction fragments tend to be smaller in the distal part of the band, methylation appears to decrease and the G + C content becomes the highest observed for the chromosome (K.Gardiner *et al.*, in preparation). Are these features reflections of the higher gene content, or are they fundamental to telomeric regions? Band q22.1, in contrast, resembles q22.3 only in the distal part, displaying features of a dark band in the proximal part. With the framework that has been laid here, determinations of gene density, patterns of methylation in rare CpG islands and their tissue specificity, isolation and characterization of translocation breakpoints, in addition to physical linkage, can all now be carried out more rapidly and precisely. As mapping efforts progress, it will be interesting to see which of these emerging features of human chromosome 21 are peculiar to it and which are general to mammalian genome organization.

Materials and methods

Cell lines

Cell lines HT1080, HeLa, normal long arm controls, 153E7b and 2Fur-1, and all members of the hybrid clone panel have been described previously (VanKeuren *et al.*, 1986, 1989; Gardiner *et al.*, 1988) with the exception of JC6, MRC2 and 6918. Hybrids JC6 and 6918 contain the deletion chromosome 21s described by Korenberg *et al.* (1986) and Reynolds *et al.* (1985) respectively; MRC2 contains a ring chromosome 21, isolated from the mother of two children with Down's syndrome (A.Robinson and B.Ward, unpublished results). Chinese hamster ovary (CHO)/human hybrids are constructed by fusion of the original cell line or fresh lymphocytes with CHO ade-C, using UV-inactivated Sendai virus (Moore *et al.*, 1977). Hybrids are selected for growth and maintained in F12D + 8% fetal calf serum that has been dialyzed to remove purines. This medium, which does not allow growth of the CHO parental cells, selects for the maintenance of any human chromosome containing 21q22.1 (Chadefaux *et al.*, 1984). The chromosome complement of all cell lines is monitored by cytogenetic analysis, with Giemsa−trypsin banding and Giemsa 11 staining (Seabright, 1971; Friend *et al.*, 1976; Morse *et al.*, 1982).

Chromosome 21-specific sequences

Abbreviated probe names, together with the Human Genome Mapping IX designation are given in Tables I and II. The genomic clones pGS 4U, B3, H8, E8 and E9 are from G.Stewart (Stewart *et al.*, 1985a,b); pJG 63, 66,

90 and 108 from J.Galt (Pearson *et al.*, 1987); pG-PFKL3.3 (PFKL), liver-type phosphofructokinase from Y.Groner (Levanon *et al.*, 1987); pCW21pcq from C.Wong (Wong *et al.*, 1989); the 16 probes of the pPW 200 and 500 series are subclones in pBR328 and pBR322 respectively of previously described sequences (Watkins *et al.*, 1985a,b); clones in Table I, and in Table II marked 1 are from U.Tantravahi (Tantravahi *et al.*, 1988). The following are cDNAs: Hu-ets2, the human EST2 oncogene, from N.Sacchi (Sacchi *et al.*, 1986); FB130 (APP), the β-amyloid precursor protein, from J.Gusella and R.Tanzi; pSC61-10 (SOD1), cytoplasmic superoxide dismutase, from Y.Groner (Sherman *et al.*, 1983); pS2 (BCEI), a breast cancer estrogen-induced mRNA, from ATCC (Moisan *et al.*, 1985); LFA9.1.1 (LFA), a lymphocyte function-associated antigen, from T.Springer and T.Kishimoto (Kishimoto *et al.*, 1988); pMalphaACr2-Pst (CRYA), murine α-crystallin, from J.Hawkins (Hawkins *et al.*, 1987); MX1 and MX2, two interferon-induced cDNAs, p78 and p78-related, from M.Horisberger (Horisberger *et al.*, 1988; and M.Horisberger, unpublished; for nomenclature, see Carritt and Litt, 1988; Aebi *et al.*, 1989); COL6A1 and COL6A2, collagen 6 genes (Weil *et al.*, 1988); ERG, the ETS2-related oncogene, from V.Rao (Rao *et al.*, 1987).

Standard gel electrophoresis, Southern transfer and hybridizations were all by established methods (Maniatis *et al.*, 1982). Probes were labelled by the method of random oligo priming (Feinberg and Vogelstein, 1984).

Pulsed-field gel procedures

Preparation of DNA in agarose plugs, restriction enzyme digestions, size markers and resolution of fragments up to 1500 kb were all as described previously (Gardiner *et al.*, 1986, 1988). Size markers: <50–300 kb, λ DNA (New England Biolabs) ligated to form concatamers; 200–2200 kb, chromosomes from *Saccharomyces cerevisiae* strain AB1380 (M.V.Olson); 3000–6000 kb, chromosomes from *Schizosaccharomyces pombe* 972-h. Conditions for the separation of *S.pombe* chromosomes: 24 h, 40 V, 60 min pulse time, followed by 60 h, 60 V, 40 min pulse time. Pulsed-field gel electrophoresis was carried out in the TAFE system (Beckman Instruments). Restriction enzymes were from Beckman Instruments and New England Biolabs.

Acknowledgements

The authors wish to thank Iris Hart, Barbara Flook, Diane Vannais and Reginald Williams for excellent technical assistance. This is publication no. 1034 from the Eleanor Roosevelt Institute for Cancer Research, Inc. and Florence R.Sabin Laboratories for Genetic and Development Medicine. It was supported by grant no. HD17449 from the National Institutes of Health.

References

Abe,K., Wei,J.-F., Wei,F.-S., Hau,Y.-C., Uehara,H., Arbgt,K. and Bennett,D. (1988) *EMBO J.*, **7**, 3441–3449.

Aebi,M., Fah,J., Hurt,N., Samual,C., Thomis,D., Bayyigher,L., Pavlovic,J., Haller,O. and Staeheli,P. (1989) *Mol. Cell Biol.*, **8**,

Allore,R., O'Hanlon,D., Price,R., Neilson,K., Willard,H., Cox,D., Marks,A. and Dunn,R. (1988) *Science*, **239**, 1311–1313.

Bernardi,G. (1990) *Annu. Rev. Genet.*, in press.

Bernardi,G., Olofsson,B., Filipski,J., Zerial,M., Salinas,J., Cuny,G., Meunier Rotival,M. and Rodier,F. (1985) *Science*, **228**, 953–957.

Bird,A. (1986) *Nature*, **321**, 209–213.

Bird,A., Taggart,M., Frommer,M., Miller,O.J. and MacCleod,D. (1985) *Cell*, **40**, 91–95.

Carritt,B. and Litt,M. (1988) *Cytogenet. Cell Genet.*, **49**, 103–104.

Chadefaux,B., Allard,O. and Rethore,M. *et al.* (1984) *Hum. Genet.*, **66**, 190–192.

Cooper,D., Taggart,M. and Bird,A. (1983) *Nucleic Acids Res.*, **11**, 647–658.

Cummings,D.E. (1978) *Annu. Rev. Genet.*, **12**, 25–46.

Drahovsky,D., Kaul,S., Boehm,T. and Wacker,A. (1980) *Int. J. Biochem.*, **13**, 565–570.

Feinberg,A. and Vogelstein,B. (1984) *Anal. Biochem.*, **137**, 266–267.

Fourth International Workshop on Chromosomes in Leukemia, Chicago, IL (1982) *Cancer Genet. Cytogenet.*, **11**, 284–287.

Friend,K., Chen,S. and Ruddle,F. (1976) *Somat. Cell Mol. Genet.*, **2**, 183–188.

Gardiner,K. and Patterson,D. (1989) *Electrophoresis J.*, **10**, 296–302.

Gardiner,K., Laas,W. and Patterson,D. (1986) *Somat. Cell Mol. Genet.*, **12**, 185–195.

Gardiner,K., Watkins,P., Munke,M., Drabkin,H., Jones,C. and Patterson,D. (1988) *Somat. Cell Mol. Genet.*, **14**, 623–638.

Gardiner-Garden,M. and Frommer,M. (1987) *J. Mol. Biol.*, **196**, 261–282.

Goate,A., Haynes,A.R., Owen,M., Farrall,M., James,L., Lai,L., Mullan,M., Roques,P., Rossor,M., Williamson,R. and Hardy,R. (1989) *Lancet*, **i**, 352–355.

Goldman,M.A., Holmquist,G.P., Cray,M.C., Caston,L.A. and Nog,A. (1984) *Science*, **224**, 686–692.

Habedank,M. and Rodewald,A. (1982) *Hum. Genet.*, **60**, 74–77.

Haluska,F., Tsiyimoto,Y. and Croce,C. (1987) *Trends Genet.*, **3**, 11–15.

Hawkins,J., VanKeuren,M., Piatigorsky,J., Law,M.L., Patterson,D. and Kao,F.T. (1987) *Hum. Genet.*, **76**, 375–380.

Holmberg,M. and Jonasson,J. (1973) *Hereditas*, **74**, 57–68.

Horisberger,M.A., Wathelet,M., Syperer,J., Islam,Q., Levan,G. and Huey,G. (1988) *Content J. Somat. Cell Mol. Genet.*, **14**, 123–131.

Ikemura,T. and Aota,S. (1988) *J. Mol. Biol.*, **203**, 1–13.

Jenkins,E., Duncan,C., Wright,C., Geodeano,F., Wilbur,L., Wisniewski, K., Sklower,S., French,J., Jonea,C. and Brow,W. (1983) *Clin. Genet.*, **24**, 97–102.

Kishimoto,T.K., O'Connor,K., Lee,A., Roberts,T.M. and Springer,T.A. (1988) *Cell*, **48**, 681–690.

Korenberg,J. and Engels,W.R. (1978) *Proc. Natl. Acad. Sci. USA*, **74**, 3382–3386.

Korenberg,J. and Rykowski,R. (1988) *Cell*, **53**, 391–400.

Korenberg,J., Kalousek,D., Anneren,K., Hał,J., Epstein,C. and Cox,D. (1986) *Am. J. Hum. Genet.*, **39**, A119.

Korenberg,J., Croyle,M. and Cox,D. (1987) *Am. J. Hum. Genet.*, **41**, 963–978.

Kuhn,F., Herman,E. and Denniston,C. (1985) *Hum. Genet.*, **70**, 1–5.

Levanon,D., Danciger,E., Dafni,N. and Groner,Y. (1987) *Biochem. Biophys. Res. Commun.*, **147**, 1182–1187.

Lindsay,S. and Bird,P. (1987) *Nature*, **327**, 336–338.

Lichter,P., Cremer,T., Tang,C.-C., Watkins,P.C., Manuelidis,L. and Ward,D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 9664–9668.

MacDonald,G., Price,E., Chu,M.-L., Timple,R., Allore,R., Marlas,A., Dunn,R. and Cox,D. (1988) *Am. J. Hum. Genet.*, **??**, A.151.

Maniatis,T., Fritsch,E. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

McCormick,M.K., Schinzel,A., Petersen,M.B., Stetten,G., Driscoll,D.J., Cantu,E.S., Tranebjaerg,L., Mikkelsen,M., Watkins,P.C. and Antonarakis,S.E. (1989) *Genomics* **5**, 325–331.

Mitelman,F. and Heim,S. (1988) *Cancer Res.*, **48**, 7115–7119.

Moisan,J.P., Mattei,M.G., Baeteman-Voekel,M.A., Mattei,J.F., Brown,A., Garnier,J., Jeltsch,M., Masiakowsky,P., Roberts,M. and Mandel,J. (1985) *Cytogenet. Cell Genet.*, **40**, 701–702.

Moore,E., Jones,C., Kao,F. and Oates,D. (1977) *Am. J. Hum. Genet.*, **29**, 389–396.

Morse,H.G., Hays,T., Patterson,D. and Robinson,A. (1982) *Hum. Genet.*, **61**, 141–144.

Nadeau,J., Eppig,J. and Reiner,A. (1989) *Linkage and Synteny Homologies between Mouse and Man*. Jackson Laboratory, Bar Harbor, ME.

Nelson,M. and McClelland,M. (1987) *Nucleic Acids Res.*, **15**, r215–r230.

Neve,R., Stewart,G., Newcomb,P., VanKeuren,M., Patterson,D., Drabkin,H. and Kurnit,D. (1986) *Gene*, **49**, 361–369.

Ohno,S. (1986) *Trends Genet.*, **2**, 8.

Owen,M., James,L., Hardy,J., Williamson,R. and Goate,A. *Am. J. Hum. Genet.*, in press.

Pearson,P., Kidd,K. and Willard,H. (1987) *Cytogenet. Cell Genet.*, **46**, 523–524.

Perez-Castello,A., Del Mayo,J. and Abrisqueta,J. (1983) *Annal. Genet.*, **26**, 123–128.

Rahmani,Z., Blouin,J.-L., Creau-Goldberg,N., Watkins,P.C., Mattei,J.-F., Poissonnier,M., Prieur,M., Chettouh,Z., Nicole,A., Aurias,A., Sinet,P.-M. and Delabar,J.-M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5958–5962.

Rao,V.N., Papas,T.S. and Reddy,S.P. (1987) *Science*, **237**, 635–639.

Razin,A. and Riggs,A.D. (1980) *Science*, **210**, 604–610.

Reeves,R., Robalais,N., Oster-Granite,M.L., Wisniewski,H., Coyle,J. and Gearhart,J. (1987) *Mol. Brain Res.*, **2**, 215–221.

Reis,R. and Goldstein,S. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3949–3953.

Reynolds,J., Wyandt,H. and Kelly,T. (1985) *Am. J. Hum. Genet.*, **20**, 173–180.

Sacchi,N., Watson,D.K., Guerts vanKessel,A.H., Hagemeijer,S., Kelsey,J., Drabkin,H.D., Patterson,D. and Papas,T.S. (1986) *Science*, **231**, 279–382.

Seabright,M. (1971) *Lancet*, **ii**, 971–972.

Sherman,L., Dafnie,N., Lieman-Hurwitz,J. and Groner,Y. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5465–5469.

Silva,A. and White,R. (1988) *Science*, **210**, 604–610.
Sinet,P.-M., Couturier,J., Dutrillaux,B., Poissonnier,M., Raoul,O., Rethere,M.-O., Allard,D., Lyeune,J. and Jerome,H. (1976) *Exp. Cell Res.*, **97**, 47–55.
Stefani,L., Galt,J., Palmer,A., Mullington-Ward,A. and Nevin,M. (1988) *Nucleic Acids Res.*, **16**, 2885–2890.
Stewart,G., Harris,P., Galt,J. and Ferguson-Smith,M.A. (1985a) *Nucleic Acids Res.*, **13**, 4125–4132.
Stewart,G., Tanzi,R. and Gusella,J.F. (1985b) *Nucleic Acids Res.*, **13**, 7168.
Tantravahi,U., Stewart,G., VanKeuren,M., McNeil,G., Roy,S., Patterson,D., Drabkin,H., Lalande,M., Kurnit,D. and Latt,S. (1988) *Hum. Genet.*, **79**, 196–202.
Tanzi,R., Haines,J., Watkins,P., Stewart,G., Wallace,M., Halewell,R., Wong,C., Wexler,N., Conneally,P. and Gusella,J. (1988) *Genomics*, **3**, 129–136.
VanKeuren,M., Watkins,P., Drabkin,H., Jabs,E., Gusella,J. and Patterson,D. (1986) *Am. J. Hum. Genet.*, **38**, 793–804.
VanKeuran,M., Stewart,G., Bradley,C., Kurnit,D., Neve,R., Watkins,P., Tanzi,R., Gusella,J. and Patterson,D. (1989) *Am. J. Med. Genet.*, **33**, 369–375.
Warren,A., Slaugenhaupt,S., Lewis,J., Chakravarti,A. and Antonarakis,S. (1989) *Genomics*, **4**, 579–591.
Watkins,P., Tanzi,R. and Gibbons,K.T. *et al.* (1985a) *Nucleic Acids Res.*, **13**, 6075–6088.
Watkins,P., Watkins,P.A., Hoffman,N. and Stanislovitis,P. (1985b) *Cytogenet. Cell Genet.*, **40**, 701–702.
Weil,D., Mattei,M.-G., Passage,E., Cong,N., Conway,D., Mann,K., Deutymann,R., Timpl,R. and Chu,M.-L. (1988) *Am. J. Hum. Genet.*, **42**, 435–445.
Wong,C., Kazazian,H., Setton,G., Earnshaw,W., VanKeuren,M. and Antonarakis,S. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1914–1918.
Yunis,J.J. (1981) *Hum. Genet.*, **56**, 293–298.
Zerial,M., Salinas,J., Filipski,J. and Bernardi,G. (1986) *Eur. J. Biochem.*, **160**, 474–485.

Received on August 16, 1989; revised on August 16, 1989