

A Genetic Linkage Map of the Human Genome

Helen Donis-Keller,* Philip Green,* Cynthia Helms,* Samuel Cartinhour,* Barbara Weiffenbach,* Karen Stephens,* Tim P. Keith,* Donald W. Bowden,* Douglas R. Smith,* Eric S. Lander,† David Botstein, Gita Akots,* Kenneth S. Rediker,* Thomas Gravius,* Valerie A. Brown,* Marcia B. Rising,* Carol Parker,* Jody A. Powers,* Diane E. Watt,* Erick R. Kauffman,* Angela Bricker,* Pamela Phipps,* Hans Muller-Kahle,* Thomas R. Fulton,* Siu Ng,* James W. Schumm,* Jeffrey C. Braman,* Robert G. Knowlton,* David F. Barker,* Steven M. Crooks,* Steven E. Lincoln,† Mark J. Daly,† and Jeff Abrahamson†

* Department of Human Genetics
Collaborative Research, Inc.
Bedford, Massachusetts 01730

† Whitehead Institute for Biomedical Research
9 Cambridge Center
Cambridge Massachusetts 02142

Summary

We report the construction of a linkage map of the human genome, based on the pattern of inheritance of 403 polymorphic loci, including 393 RFLPs, in a panel of DNAs from 21 three-generation families. By a combination of mathematical linkage analysis and physical localization of selected clones, it was possible to arrange these loci into linkage groups representing 23 human chromosomes. We estimate that the linkage map is detectably linked to at least 95% of the DNA in the human genome.

Introduction

The extensive genetic variation present in humans represents an invaluable resource for molecular biology and for medical investigation (McKusick, 1986). Cloning and characterization of genes underlying human inherited traits has helped elucidate the molecular basis for a wide range of physiological processes in higher organisms (e.g., Goldstein and Brown, 1983). The difficulty with extending this approach to all heritable traits and diseases is that the vast majority of such genes are known only by their phenotype: their protein product is unknown. In experimental higher organisms, the time-honored approach to the study and, more recently, the cloning of such genes is to identify their location in the genome through the use of a genetic linkage map.

The notion of a genetic map dates back to 1911, when Sturtevant, while an undergraduate in T. H. Morgan's lab, realized that linkage information could be used to determine the relative position of genes along a chromosome, and at once produced the first genetic map, comprising five sex-linked loci in *Drosophila* (Sturtevant, 1913, 1965). At the time, the internal consistency of the linear map provided important support for the chromosomal theory of

inheritance. Over the next 75 years, complete genetic linkage maps proved to be essential tools for studying the properties of mutations. Genetic markers gained new importance with the advent of recombinant DNA, since cloned markers provide starting points for cloning closely linked genes by chromosomal walking (Bender et al., 1983). Unfortunately, the construction of complete genetic linkage maps has traditionally required the isolation of hundreds of single-gene mutations with easily scored phenotypes, followed by extensive interbreeding of mutant stocks to ascertain the map position of the mutations. Such an effort has only been practical in a few intensively studied genetic systems, such as *Escherichia coli*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Zea mays*, and *Mus musculus*. In humans, despite great interest and occasional successes in detecting linkage (e.g., Mohr, 1954), construction of a genetic map seemed impractical.

Several years ago, Botstein et al. (1980) argued that it was feasible to construct a complete linkage map of the human genome using common variations in DNA sequence, most conveniently visualized as restriction fragment length polymorphisms (RFLPs), as genetic markers. With such polymorphisms as markers, geneticists could study inheritance in existing pedigrees, since all individuals would be heterozygous at many loci.

Systematic screening has revealed that RFLPs are not uncommon (Willard et al., 1985) in the human genome, although the majority show only a low degree of polymorphism. Through the study of the inheritance of randomly selected RFLPs in human families, linkage has been detected to a number of human diseases, including Duchenne muscular dystrophy (Davies et al., 1983), Huntington's disease (Gusella et al., 1983), cystic fibrosis (Tsui et al., 1985; Knowlton et al., 1985; Wainwright et al., 1985; White et al., 1985b), adult polycystic kidney disease (Reeders et al., 1985), retinoblastoma (Cavenee et al., 1983), familial Alzheimer's disease (St George-Hyslop et al., 1987), bipolar affective disorder (Egeland et al., 1987), von Recklinghausen neurofibromatosis (Barker et al., 1987a; Seizinger et al., 1987), multiple endocrine neoplasia type 2a (Mathew et al., 1987; Simpson et al., 1987) and familial polyposis (Bodmer et al., 1987). The existence of nearby DNA probes has facilitated the molecular cloning of the genes for chronic granulomatous disease (Royer-Pokora et al., 1986), Duchenne muscular dystrophy (Monaco et al., 1986), and retinoblastoma (Friend et al., 1987; Lee et al., 1986).

The availability of a complete linkage map of the human genome would greatly amplify the power of this approach to human molecular genetics. With such a map, (1) the chromosomal location of newly discovered linkages would be known at once; (2) several nearby starting points would be available for efforts to clone disease genes; (3) prenatal or presymptomatic diagnosis of individuals at risk would become more accurate, through the use of markers flanking the disease gene, and more widely available,

since most families would likely be informative for at least some of the markers near the disease; (4) the search for disease genes would become more efficient, because a map would assure that the entire genome had been scanned and because multilocus analysis would decrease the number of meioses required to detect linkage (Lathrop et al., 1984; Lander and Botstein, 1986a,b); (5) it would become possible to map heterogeneous genetic disorders (Lander and Botstein, 1986a) and rare recessive diseases (Lander and Botstein, 1987), as well as to test whether there is a genetic basis for disorders whose inheritance is currently unclear; and (6) one could begin to study the nature of recombination in humans. In addition, the availability of a genetic map would assist efforts to construct a complete physical map of the human genome; islands of overlapping clones could be genetically ordered by means of polymorphisms they detect, even before the intervening DNA had been identified.

In view of the value of a complete human linkage map, an international collaboration called the Centre d'Etude du Polymorphisme Humain (CEPH) was organized in 1984. The CEPH maintains cell lines from 40 three-generation human families, consisting in most cases of four grandparents, two parents, and an average of eight children. Such families are ideal for genetic mapping, because the *cis-trans* relationship (i.e., linkage phase) of alleles in the parents can frequently be inferred from grandparental DNAs (Figure 1) and, once this is determined, crossovers can be counted in the meioses giving rise to the children. The CEPH distributes DNAs from these families to collaborating investigators around the world. Since collaborators use the same reference panel of families, data generated by one group can be used by others to detect new linkages. In particular, crossovers in the families can be mapped, allowing the order of probes to be determined. The collective efforts of the CEPH collaboration will eventually produce a far more detailed map than could any group alone. To date, hundreds of RFLPs have been identified (Willard et al., 1985; Schumm et al., 1987; Nakamura et al., 1987), including an increasing number with a higher degree of polymorphism, but relatively few have been mapped. At present, detailed RFLP linkage maps have been published only for the X chromosome (21 markers spanning 185 cM) (Drayna et al., 1984), a chromosomal region in 6p (four linked loci spanning 20 cM) (Leach et al., 1986), a chromosomal region in 13q (nine loci spanning 70 cM) (Leppert et al., 1986), and preliminary reports for chromosome 7 (Donis-Keller et al., 1986) and chromosome 12 (White et al., 1986).

We report the construction of a linkage map of the human genome, based on the pattern of inheritance of 403 polymorphic loci, including 393 RFLPs, in a panel of DNAs from 21 three-generation families. By a combination of mathematical linkage analysis and physical localization of selected clones, it was possible to arrange these loci into linkage groups representing the 23 human chromosomes. We estimate that the linkage map is detectably linked to at least 95% of the DNA in the human genome.



Figure 1. Inheritance of a RFLP Locus in a CEPH Family
RFLP probe CRI-L1265, which detects a single locus on chromosome 5, displays three alleles on Southern hybridization to DNA from CEPH reference family 1341 digested with the restriction endonuclease *Taq*I. The alleles correspond to single fragments of size 10.0 kb, 7.7 kb, and 6.5 kb, respectively. For each of the parents, one can infer which allele was inherited from the grandmother and which from the grandfather (i.e., linkage phase is known). For each child, one can then infer the grandparental origin of their two alleles.

Results and Discussion

Identification, Characterization and Pedigree Studies of RFLPs

We tested 1680 clones from a Charon 4A phage library of human genomic DNA (Lawn et al., 1978) to see whether they detected restriction fragment length polymorphism by hybridization to Southern blots of DNA from five unrelated individuals, each digested with six to nine restriction enzymes. Over 500 probes were identified that detected variable banding patterns indicative of polymorphism (Schumm et al., 1987). From this collection, a subset of the 180 probes detecting the highest degree of polymorphism were selected for inheritance studies in the 21 CEPH families for which cell lines were available or for which a sufficient quantity of DNA for these studies could be provided by the CEPH. An additional 46 probes detecting polymorphism were similarly identified from human genomic cosmid libraries (D. Bowden, unpublished results), from phage libraries of human genomic DNA grown on recombination deficient hosts (Wyman et al., 1986), and from two chromosome-specific phage libraries, described below. The probes were hybridized to Southern blots of DNAs from the parents of the families to confirm the observed polymorphic pattern and to determine which families were informative for linkage analysis (by virtue of one or both parents being heterozygous for the RFLP). For each informative family, genotypic data were then gathered on all family members. (Probes with fewer than 45 informative meioses in the families were omitted from further analysis.) Table 1 summarizes the probe-enzyme combinations for these markers, together with their heterozygosities, polymorphism information content (PIC; Botstein et al., 1980), and number of informative meioses in the CEPH families studied.

As part of a parallel effort to produce high-resolution linkage maps of chromosomes 7 and 16, inheritance data were similarly gathered for 42 probes on chromosome 7 (Barker et al., 1987) and for 37 probes on chromosome 16 (T. Keith, unpublished results) isolated from libraries enriched for these chromosomes; probes on these chromosomes also arose in the whole genome screen described above, yielding a total of 59 probes on chromosome 7 and 40 on chromosome 16. We also studied the inheritance of 54 further RFLPs detected by probes that had been isolated by other investigators and localized to chromosomes or subchromosomal regions (Table 2). In all, we determined the inheritance of 360 RFLPs. In addition, we included in our analysis previously published data on 46 loci that had been contributed to the CEPH (Table 3).

Construction of the Linkage Map

Multilocus genetic linkage analysis is required for the accurate construction of human linkage maps, including the determination of locus order. Such multilocus analysis has been limited until recently by the inability of available algorithms and computer programs to analyze more than three or four loci simultaneously (Leppert et al., 1986; Drayna et al., 1984; Morton et al., 1986; Smith, 1986). We recently developed faster algorithms, which permit simultaneous maximum likelihood map distance estimation in CEPH families for many loci (Lander and Green, 1987), and have embodied these algorithms in computer programs (Lander et al., 1987; Barker et al., 1987a). This has made possible the use of new analytical strategies for constructing the map.

Construction of Linkage Groups and Physical Assignment to Chromosomes

The traditional measure of support for linkage of a pair of loci, the LOD score, is defined as the \log_{10} of the ratio of the probability that the data would have arisen if the loci are linked to the probability that the data would have arisen if the loci are unlinked (Morton, 1955). The conventional threshold for declaring linkage is a LOD score of 3.0, which would indicate that the observed data are 1000-fold more likely to have occurred for a pair of linked loci than for a pair of unlinked loci. However, since an arbitrary pair of loci is a priori 50-fold more likely to be unlinked than linked, a LOD score of 3.0 provides only about 20:1 odds in favor of linkage (Morton, 1955). In other words, one expects about one in 20 independent linkages with a LOD score of 3.0 to be spurious. Because of the large number of comparisons performed in the present study, we adopted a stricter threshold of LOD 4.0 in order to lower the odds of spurious linkage to about 200:1 against. (In fact, three spurious LOD scores of 4.0 between pairs of loci on separate chromosomes occurred in the data set. The anomalies became obvious upon multipoint analysis and were resolved by physical assignment of various linked loci.)

Throughout the course of the project, linkage data were analyzed on a weekly basis in order to monitor the overall progress toward a linkage map and to select a subset of

the probes for physical assignment to chromosomes. We describe here only the final rounds of analysis, performed on the completed data set. The analysis began with the computation of the recombination fraction and LOD score between each of the 81,003 pairs of loci. Using the threshold of LOD 4.0 for linkage, the loci could be grouped into linkage groups and unlinked loci. Linkage groups were immediately assigned to specific chromosomes if they contained a probe whose chromosomal location had previously been determined. Otherwise, linkage groups were assigned to chromosomes by means of hybridizing one or more probes to panels of rodent-human hybrid cells containing varying human chromosomal complements. Unlinked probes were also assigned in this manner. Each autosome contained at least one linkage group. We then turned to multilocus analysis to detect linkage between syntenic linkage groups, to detect linkage between linkage groups and the unlinked probes, and to construct genetic maps for each chromosome.

Multilocus Analysis of Chromosome Maps

In constructing maps in experimental organisms, one can usually determine the correct order of the loci first (based on data from three-point crosses) and then estimate the map distances between consecutive loci. Individual three-point crosses are less efficient for analyzing human crosses, because different families are informative for different triplets of loci; integrating all of the information requires simultaneous multilocus analysis of the loci. To determine the genetic order of a set of loci, map distances and the corresponding likelihoods should ideally be determined for all possible orders of the loci (Ott, 1985); a genetic order is accepted if it is considerably more likely to have given rise to the data than alternative orders.

For linkage groups containing only six loci, it is straightforward to compute the maximum-likelihood multipoint map for each of the 360 possible orders for the loci and to then compare the likelihoods (Lander and Green, 1987; Lander et al., 1987). Of course, this becomes impossible as the number of loci increases, since the number of possible orders grows exponentially. Instead, maps can be constructed by sequentially adding loci to a map (as indeed is done in experimental genetic organisms). We have employed two such strategies for making maps of the more densely covered chromosomes, both of which led to the same results.

In the first approach, loci were first sorted by their informativeness, and a pair of highly informative linked loci were chosen as the nucleus of the map. The next most informative locus was then selected and placed in each possible position with respect to the two loci. The maximum likelihood distances for each of the orders were computed, and the locus was added to the map if one order was preferred over the others by a 100:1 odds ratio (i.e., if the observed data were 100-fold more likely to have arisen with that order than with any other order). If the locus could not be uniquely placed in this manner, another locus was tested in the same way. Thus, only conservative additions were made to the map. When no remaining lo-

Table 1. (continued)

Probe	Chr.	Enz.	Alleles	Het	PIC	Meioses	Assignment	Probe	Chr.	Enz.	Alleles	Het	PIC	Meioses	Assignment
CRI-R12	7	M	5	.31	.29	102	linkage	CRI-L409	12	E,Hc	4	.60	.45	185	linkage
CRI-R40	7	M	2	.45	.36	143	linkage	CRI-L416	12	E	10	.52	.50	167	hybrid panel
CRI-R53	7	Bg	3	.40	.33	126	linkage	CRI-L809	12	M	3	.45	.40	155	hybrid panel
CRI-R944	7	M	4	.43	.36	141	linkage	CRI-P102	12	M	2	.24	.24	68	linkage
CRI-R967	7	M	4	.40	.33	113	linkage	CRI-P153	12	M	2	.45	.36	129	linkage
								CRI-R102	12	T	4	.60	.52	203	linkage
CRI-C96	8	Bg	4	.29	.24	88	linkage	CRI-V834	12	M	2	.50	.33	117	linkage
CRI-L40	8	Bg	2	.40	.31	125	linkage								
CRI-L186	8	T	4	.57	.55	176	linkage	CRI-R214	13	T	2	.40	.31	131	linkage
CRI-L388	8	T	2	.40	.33	137	linkage	CRI-V1134	13	H	3	.60	.52	188	linkage
CRI-L413	8	P	5	.64	.60	218	hybrid panel								
CRI-L580	8	T	2	.55	.45	174	linkage	CRI-C70	14	M	6	.71	.67	235	linkage
CRI-L1212	8	M	2	.19	.17	58	linkage	CRI-L329	14	T	3	.48	.36	114	linkage
CRI-L1251	8	M	4	.62	.48	207	hybrid panel	CRI-L436	14	T	2	.52	.36	203	hybrid panel
CRI-L1427	8	T	5	.40	.38	134	linkage	CRI-L1013	14	Bg	2	.43	.33	152	linkage
CRI-R150	8	H	2	.50	.31	137	hybrid panel	CRI-L1113	14	P	2	.48	.33	159	linkage
CRI-R191	8	M,T	7	.74	.64	253	linkage								
CRI-R370	8	M	2	.50	.40	169	linkage	CRI-L146	15	H,M	6	.62	.55	183	linkage
CRI-U4	8	M	3	.36	.33	99	linkage	CRI-L389	15	Bg	6	.21	.21	84	hybrid panel
CRI-V822	8	E	2	.57	.40	181	linkage*	CRI-L442	15	M	3	.60	.50	193	linkage
CRI-V1225	8	E	4	.69	.69	200	linkage	CRI-L1204	15	M	2	.57	.43	167	hybrid panel
								CRI-P78	15	B	2	.36	.33	121	linkage
CRI-L659	9	T	2	.45	.38	141	linkage	CRI-P452	15	M	2	.24	.21	82	linkage
CRI-L1022	9	P	2	.26	.24	90	hybrid panel	CRI-R382	15	M	4	.62	.60	191	linkage
CRI-L1263	9	T	4	.64	.52	225	hybrid panel								
CRI-L1424	9	H	2	.43	.31	143	linkage*	CRI-L223	16	T	2	.48	.40	159	linkage
CRI-P110	9	E	2	.45	.33	151	linkage	CRI-L922	16	H,Hc	4	.50	.43	163	linkage
CRI-P111	9	T	2	.55	.36	169	linkage	CRI-R99	16	H,M	4	.64	.62	226	linkage
CRI-R3	9	E	2	.50	.33	154	linkage								
								CRI-L581	17	T	3	.55	.50	196	hybrid panel
CRI-L368	10	H	9	.74	.74	232	hybrid panel	CRI-L946	17	M	3	.43	.40	132	linkage
CRI-L647	10	T	5	.50	.48	152	hybrid panel	CRI-P3	17	M	2	.52	.31	155	linkage
CRI-L893	10	E	2	.40	.31	136	linkage								
CRI-L941	10	M	2	.36	.31	118	linkage	CRI-L84	18	M	2	.36	.31	119	linkage
CRI-L1005	10	H	2	.31	.21	96	linkage	CRI-L159	18	P	23	.74	.74	264	hybrid panel
CRI-L1083	10	M	3	.55	.45	182	hybrid panel	CRI-L261	18	M	2	.52	.40	181	linkage
								CRI-L821	18	M	2	.26	.26	81	linkage*
CRI-L424	11	E,H	4	.57	.55	171	linkage	CRI-L1156	18	M	3	.57	.55	179	linkage
CRI-L451	11	E,M	4	.36	.31	102	linkage	CRI-P30	18	Bg,T	4	.52	.43	161	linkage
CRI-L605	11	M	7	.71	.50	227	hybrid panel	CRI-R397	18	P	5	.57	.55	190	hybrid panel
CRI-L762	11	T	5	.62	.50	206	linkage								
CRI-L834	11	M	2	.45	.33	142	linkage	CRI-L127	20	M	2	.43	.31	130	linkage
CRI-L937	11	Hc	2	.33	.24	97	linkage	CRI-L355	20	Bg	22	.74	.74	257	hybrid panel
CRI-L944	11	T	5	.45	.36	149	linkage*	CRI-L1214	20	Bg	17	.98	.98	334	hybrid panel
CRI-L962	11	M	6	.71	.62	245	hybrid panel	CRI-L1239	20	M	2	.40	.33	115	linkage
CRI-L1382	11	M,T	4	.57	.48	173	linkage								
CRI-R83	11	H,M,T	6	.71	.69	265	linkage	CRI-L427	21	R	25	.95	.93	327	hybrid panel
CRI-R365	11	Bg	5	.48	.48	157	linkage								
CRI-R397	11	R	4	.43	.38	157	linkage	CRI-L518	22	T	2	.38	.31	130	linkage*
CRI-R548	11	P	2	.21	.19	76	linkage	CRI-L1272	22	Bg	3	.40	.31	131	linkage
CRI-R975	11	T	2	.50	.38	159	linkage	CRI-R657	22	T	2	.40	.36	134	hybrid panel
CRI-V928	11	M	3	.48	.40	144	linkage								
								CRI-C88	X	M	2	.14	.14	20	sex-linkage
CRI-C2	12	P,E	8	.71	.69	229	linkage	CRI-L1391	X	M	3	.48	.48	86	sex-linkage
CRI-C86	12	P	4	.17	.14	56	linkage	CRI-R393	X	T	2	.34	.34	49	sex-linkage
CRI-L303	12	M	2	.24	.21	82	linkage	CRI-S232	X	E	>7	.90	.90	138	sex-linkage
CRI-L375	12	T	3	.57	.43	183	hybrid panel								

Enzymes are: B, BamHI; Bg, BglII; E, EcoRI; H, HindIII; Hc, HincII; M, MspI; P, PstI; R, RsaI; T, TaqI. Alleles is the number of different alleles present in the parents of the CEPH families screened, Het is the proportion of these parents that are heterozygotes, and PIC is polymorphism information content (Botstein et al., 1980). Meioses is the number of informative meioses in the parents in the CEPH families studied. Assignment lists whether the locus detected by the probe was initially assigned to the indicated chromosome via hybrid panels, or via linkage to previously assigned loci. An asterisk indicates that the initial two-point LOD score exceeded 3.0, but fell below the more stringent threshold of 4.0. The probes on chromosomes 7 and 16 isolated from the screening of libraries enriched for these chromosomes are described in detail elsewhere (Barker et al., 1987; T. Keith et al., unpublished data). Probes designated with the letter R are random probes from the Charon 4A library; with the letter L, probes from this library screened for single-copy human inserts; with the letter S, probes from the chromosome 7-specific library, some of which did not lie on chromosome 7; with the letters P and O, probes from the phage and cosmid libraries enriched for chromosome 16; with the letter C, probes from a total genomic cosmid library (D. Bowden, unpublished data); with the letters T, U, and V, probes from libraries propagated in recombination-deficient hosts.

Table 2. Physically Localized Anchor RFLP Loci

Locus	Probe	Location	Enzyme(s)	Alleles	Het	PIC	Meioses	Reference
AT3	pAT3	1q23-q25	P	2	.43	.36	143	Prochownic et al., 1983
REN	pHRnES1.9	1p21-qter	H	2	.50	.40	164	Hobart et al., 1984
CRYGP1	p5G1	2q33-q35	M	2	.29	.24	81	Meakin et al., 1985
D2S1	L2.30	2p25	Bg, M	4	.69	.48	223	Lothe et al., 1986
D2S5	IMR32-6	2p16-p15	M	2	.38	.26	116	Shiloh et al., 1985
D2S6	pXG-18	2q32-q36	T	2	.62	.43	219	Davatellis, 1985
D3S1	HS3	3q12	H	2	.40	.31	122	Zabel et al., 1983
D3S2	p12-32	3p21-p14; 3q21-qter	M	2	.38	.29	118	Barker et al., 1984
D3S4	B67	3pter-q21	T, M	5	.45	.38	154	Morie et al., 1984
ADH3	pADH73	4q21-q25	M	2	.26	.24	84	Smith et al., 1984
HEXB	pHexB43	5q13	P	2	.10	.10	32	O'Dowd, 1985
DHFR	CHB203	5q11.1-q13.2; 5q23	M	2	.19	.17	60	Anagnou et al., 1984
HPRTP2	plambda500	5p14	M	2	.31	.24	89	Patel et al., 1984
COL1A2	NJ-3	7q21.3-q22.1	E, M, T	12	.21	.19	87	Tsipouras et al., 1983
MET	pmetD, H	7q21-q31	T	7	.62	.55	223	Cooper et al., 1984
D7S8	pJ3.11	7q22	M	2	.52	.33	183	Cooper and Schmidtke, 1985
TCRB	pT10	7q32; 7q35-q36	Bg	2	.50	.33	153	Berliner et al., 1985
PLAT	pCGE217	8p12	Hc, P	4	.43	.36	136	CRI
ABL	pabIK2	9q34	P	2	.24	.17	74	Srinivasan et al., 1981
D9S1	pHF12-8	9pter-q11	T	2	.26	.21	82	Naylor et al., 1984
IRBP	H.4 IRBP	10p11.2-q11.2	Bg, M	4	.19	.19	67	Liou et al., 1987
PLAU	pCGE194	10q24-qter	Bg	2	.31	.24	101	CRI
VTR.4	VTR4.1	10q26	E	9	.38	.38	120	Colb et al., 1986
D11S16	p32-1	11p13	M	3	.52	.50	173	Feder et al., 1985
SEA	clone 3	11q13	H	2	.33	.26	97	D. Smith, personal comm.
CAT	pINT-800	11p13	T	2	.17	.14	50	Quan et al., 1986
PRB1	pPRPII2.2RP	12p13.2	E	5	.81	.81	256	Azen et al., 1984
D13S1	p7F12	13q12-q14	M, T	4	.64	.55	203	Cavenee et al., 1984
D13S3	p9A7	13q22-qter	H	2	.26	.24	84	Cavenee et al., 1984
D13S4	p1E8	13q31	M	2	.60	.43	179	Cavenee et al., 1984
D14S1	pAW101	14q32.2	E	8	.74	.71	242	Wyman et al., 1982
D15S1	pMS1-14	15q15-q21	M	2	.62	.45	201	de Martinville et al., 1984
D15S2	pDP151	15q15-q22	E	2	.40	.38	131	Brissenden et al., 1986
D15S3	pJU201	15	Bg	3	.50	.38	166	Cooper et al., 1985
HBA1	3'HVR	16p13	M	>10	.93	.93	347	Higgs et al., 1981
D17S2	L1.31	17	Bg	4	.38	.38	115	Schwartz et al., 1985
D17Z1	p17H8	17cen	E	2	.19	.17	70	Willard et al., 1986
MYH2	p10-3	17p13	M	3	.38	.33	109	Leinwand et al., 1983
NGFR	PE51	17q21-q22	Hc	2	.40	.38	133	Chao et al., 1986
D18S1	pHF12-62	18	T	2	.57	.38	163	Naylor et al., 1984
D18S3	B74	18p11.3	M	2	.50	.40	163	Morie et al., 1984
APOC2	pCII-711	19cen-q13.2	T	2	.38	.29	137	Jackson et al., 1984
D19S7	p4.1	19cen-q13.2	M	2	.52	.43	172	Shaw et al., 1986
D19S8	p17.1	19cen-q13.2	T	2	.33	.31	105	Shaw et al., 1986
D19S9	p1J2	19cen-q13.2	E	2	.19	.19	66	Shaw et al., 1986
D20S4	pMB1-27	20q13.2	M	2	.43	.38	148	de Martinville et al., 1984
D20S5	pRI2.21	20p12	M	2	.45	.38	141	Goodfellow et al., 1987
D20S6	pD3H12	20p12	T	2	.48	.38	155	Goodfellow et al., 1987
BCEI	pS2	21q22.3	B	2	.17	.14	59	Jakowlew et al., 1984
D21S15	GMG21S1	21q21.2-qter	M	4	.50	.43	160	Stewart et al., 1984
D21S17	pGSH8	21q21.2-qter	Bg	2	.40	.33	128	Stewart et al., 1984
SOD1	pS61-10	21q22.1	M	2	.12	.12	36	Lieman-Hurwitz et al., 1982
D22S10	D22C1-18	22	T	3	.29	.24	87	Hofker et al., 1985
SIS	pSM-1	22q12.3-q13.1	H	2	.45	.33	144	Clarke et al., 1984

Column headings are as defined in Table 1. Where two conflicting physical localizations have been reported for a locus, both are listed, separated by a semicolon. The reference CRI indicates gene probes that were isolated at Collaborative Research. The locus SIS is also known as PDGFB.

cus could be uniquely placed in the map, then the locus with the smallest number of permissible orders was added to the map, and the entire procedure repeated, with succeeding loci tested in each position with respect to each permissible order. Frequently, the number of permissible orders increased in the initial stages of map con-

struction but decreased when succeeding loci were added, as combined data from many loci resolved ambiguities cooperatively.

In the second strategy, a set of informative loci separated at intervals of roughly 10–20 centiMorgans were chosen, based on two-point distance calculations. All pos-

Table 3. Polymorphic Loci from CEPH Database

Locus	Location	Meioses	Reference
PGM1	1p22.1	199	White et al., 1985a
FY	1p21-q23	154	"
RH	1p36.2-p34	163	"
ACP1	2p25	112	"
GC	4q12-q13	194	"
MNS	4q28-q31	180	"
D6S5	6p24-q12	64	Leach et al., 1986
D6S7	6pter-p24	86	"
D6S8	6p21.3	123	"
D6S10	6p	169	"
GLO1	6p21.31-p21.1	150	"
HLAB	6p21.3	55	"
HLADQA	621.3	136	"
HLADRA	6p21.3	83	"
ABO	9q34.1-q34.2	99	White et al., 1985a
AK1	9q34.1-q34.2	38	"
ORM	9q34.3	190	"
D11S12	11p15.5	110	"
HBBC	11p15.1	107	"
HRAS1	11p15.5	239	"
INS	11p15	207	"
D6S9	11	115	Leach et al., 1986
D13S1	13q12-q14	125	Leppert et al., 1986
D13S2	13q22	153	"
D13S3	13q22-qter	112	"
D13S4	13q31	163	"
D13S5	13q12-q22	86	"
D13S6	13q12-q22	58	"
D13S7	13q12-q22	31	"
D13S10	13q14	19	"
ESD	13q14.11	37	"
HP	16q22.1	165	White et al., 1985a
SE	19	52	"
DXS3	Xq21.1	35	Drayna et al., 1985
DXS9	Xp22	19	"
DXS15	Xq28	72	"
DXS41	Xp22.2-p22.1	45	"
DXS42	Xq24-q26	55	"
DXS43	Xp22.2-p22.1	57	"
DXS51	Xq27-qter	77	"
DXS52	Xq28	117	"
DXS143	Xpter-p22	51	"
HPRT	Xq26-q27.3	31	"
F9	Xq26	56	"
F8C	Xq28	44	"
DXYS1	Xq21.1	64	Page et al., 1982

Meioses is number of informative meioses in the parents in the CEPH families studied. D6S9 is the locus detected by the probe p3C7, which was isolated from a chromosome 6-specific library and thus assigned a designation on chromosome 6. In fact, it lies on chromosome 11 by our linkage studies and by the hybrid panel results of Leach et al. (1986).

sible orders of these loci were tested by multipoint analysis. By trying several possible sets, it was usually possible to find a set of loci for which one order was strongly favored over all alternatives (by an odds ratio of 10,000:1). Remaining markers were then tested in each interval determined by this framework of loci and assigned to an interval when favored by 100:1 odds; markers assigned to the same interval were then ordered by computing likeli-

Family 1331

SIB 2

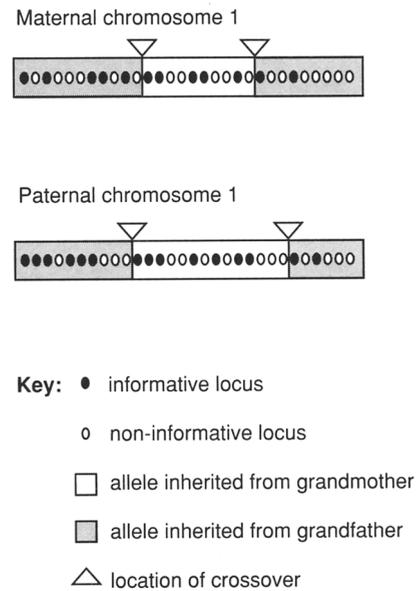


Figure 2. Grandparental Origins of Alleles at 30 Loci on Chromosome 1 in Sibling 2 of CEPH Reference Family 1331

Black ovals correspond to informative loci, white ovals to noninformative loci. A grey stippled region indicates inheritance from the grandfather, while a white region indicates inheritance from the grandmother. Each of the chromosomes 1 inherited by this child apparently underwent two crossover events, the approximate position of which is indicated by the triangle.

hoods for all possible orders. The process was repeated, since the addition of loci to the framework made it possible to place further loci uniquely. Loci that could not be placed uniquely were assigned to all intervals that could not be excluded by the 100:1 odds test, and simultaneous comparisons of all loci within a given region of the map were performed to further resolve orders. As a test of the final order, the maximum likelihood multipoint maps were computed for the various possible permutations of consecutive marker triplets embedded in the order.

Once maps were constructed, we used a computer program to picture the grandparental origins of the alleles in each chromosome under study (Figure 2). These pictures provided strong intuitive support for the genetic orders determined, because they usually indicated fewer than three crossovers per chromosome. When anomalies were observed, such as double-crossovers involving just a single probe, the data were rechecked and, where questions remained, additional Southern blots were prepared and the genotypes determined again; in this way, several data errors were corrected.

For most chromosomes, a majority of the loci fell into a

unique genetic order, while most of the remaining loci were assigned to two or three adjacent intervals in the unique order. Throughout the linkage analysis, sex-specific recombination fractions were allowed for each interval; final maps were also generated in which the recombination fraction in each interval was not permitted to vary between the sexes. Recombination fractions were converted to centiMorgans via the Kosambi mapping function (see Ott, 1985). The distances we report must be regarded as only rough approximations, because the number of meioses is relatively small by the standards of experimental organisms, because the correct mapping function for human is unknown, and because recombination distances probably vary with age and genetic background, as they do in *Drosophila* (Bridges, 1927), maize (Chang and Kikudome, 1974), and other organisms.

As a final step, the linkage maps were oriented relative to the cytogenetic map of the chromosome whenever there were two or more probes whose genetic order could be resolved and whose cytogenetic positions were known and nonoverlapping. The genetic maps did not conflict with the subchromosomal assignments, except in two cases. The location of the locus D2S6 has been reported to lie at chromosome position 2q32–q36 (Davatellis, 1985); our linkage data place it genetically near D2S5, which has been localized to 2p16–p15, and far away from CRYG, which has been localized to 2q33–q35. The conflict between the genetic data and the published physical localization remains to be resolved. Also, ORM has been physically localized to the distal tip of chromosome 9q in 9q34.3, while the ABO blood group has been indirectly assigned to 9q34.1–q34.2 on the basis of tight linkage to a physically assigned locus. Our genetic data place ABO distal to ORM.

Recombinational Map of the Human Genome

The genetic maps of the 23 chromosomes are shown in Figure 3, along with chromosome ideograms to indicate the placement of those loci that have been physically localized to subchromosomal regions. The maps indicate the estimated genetic distances obtained both when separate male and female recombination fractions are allowed for each interval and when only a single recombination fraction is allowed. Uncertainties in the genetic order are indicated.

Sex Differences in Recombination

A striking general feature of the linkage maps is the difference in recombination rates between the sexes: the genetic map of the autosomes is roughly 90% longer in females than in males, although this estimate must be regarded as only approximate. Statistical tests (see Experimental Procedures) confirm that the overall sex-specificity of recombination is highly significant. Of course, male and female meioses occur in entirely different tissues and there is a priori no reason why recombination rates would be expected to be identical. Even in organisms such as maize, in which a single individual produces both male and female gametes, sex differences are observed in recombination rates (Rhoades, 1941; Carlson, 1977; Robertson, 1984). Haldane asserted as a general rule

that, in organisms with a chromosomal mechanism of sex determination, the genetic length is shorter in the heterogametic sex (Haldane, 1922), and this rule applies well to many organisms, including *Drosophila melanogaster*, in which males (XY) show no recombination (Morgan, 1914), the moth *Bombyx mori*, in which females (ZW) show no recombination (Maeda, 1939; Tazima, 1964), and *Mus musculus*, in which certain intervals show up to twice as much recombination in females (XX) as in males (XY) (Davisson and Roderick, 1981). Increased recombination in human females was first observed in linkage studies between the ABO blood group and nail-patella syndrome (Renwick and Schulze, 1965), and has been observed since for a number of intervals in RFLP linkage maps constructed for autosomal regions. Our findings confirm that the phenomenon is quite general in humans: our linkage map indicates a substantial female excess for all autosomes except chromosome 14 (no female excess) and chromosome 10 (an 18% female excess). Substantial sex differences may exist for these chromosomes as well, but may have avoided detection because of limited sample size or incompleteness of the maps.

Despite the strong overall tendency to greater genetic length in females, the increase clearly appears to be nonuniform: many intervals appear much larger in female meiosis, others are consistent with roughly equal lengths, and a handful show increased recombination in males. As an extreme example, the loci R365 and L962 on chromosome 11q, which are jointly informative in a large number of meioses, show 0% recombination in females and 17% recombination in males (chi-squared = 13.1; $p < .0008$). (These recombination fractions are based on two-point analysis of the loci. As shown on the map, multipoint analysis yields similar, but slightly different values: 3% in females and 23% in males, with the additional female distance inferred indirectly from linkage to other loci.) In view of the large number of intervals being studied and the varying amounts of data for different intervals, however, one must be cautious about the significance of a result for any particular interval; a detailed statistical analysis of these data will appear elsewhere. However, the most effective way to test the suggestion of increased male recombination in some intervals will now be to study these same regions in a second, independent set of meioses. Longer male genetic distance has also previously been reported for a distal region of 11p in humans (White et al., 1985a) and for certain intervals in mouse (Davisson and Roderick, 1981). In maize, for which genetic lengths are generally longer in male meiosis, some intervals are reported to be longer in female meiosis (Robertson, 1984). Inhomogeneity in the ratio of female to male genetic length would suggest that, rather than sex differences being the result solely of generalized increase in the *trans*-acting recombinational machinery in human females, sex-specific chromosomal sites for recombination may exist throughout the genome (White et al., 1986). Although a special case, the human pseudoautosomal region provides one example of a clear inhomogeneity: this small region, in which the X and Y chromosomes recombine in male meiosis, measures about 50 cM in male mei-

osis but only about 5 cM in female meiosis (Rouyer et al., 1986; Page et al., 1987).

How Much of the Genome Is Covered?

The probes isolated in the random genomic screen provide an empirical estimate of the fraction of the DNA in the genome linked to the map. Of 208 probes isolated from screening of whole genomic libraries, only five (i.e., 2.4%) fail to show linkage to another locus in the map with a LOD score exceeding 3.0. Assuming that the randomly isolated RFLPs are representative, we therefore estimate that the probability exceeds 95% that a newly isolated RFLP of a similar degree of polymorphism will be detectably linked to the map. Assuming that probes detecting RFLPs are representative of DNA in the human genome (i.e., that there are not vast stretches poor in polymorphism or unclonable), we estimate that 95% of the DNA in the human genome is detectably linked to the map.

The linkage groups themselves also indicate the degree of coverage of the genome. Single, coherent linkage groups emerge for most of the chromosomes, with typically one unlinked probe that presumably lies in distal regions. The largest chromosomes are fairly densely covered, while a few of the smaller chromosomes have more rudimentary maps. We must now direct attention to filling in gaps such as those that occur on chromosome 14, which is represented by two small, unlinked linkage groups, and chromosome 19, on which none of our random probes detected loci and whose map spans only a short genetic distance.

While a random kilobase of DNA has a 95% chance of being linked, it would also be of interest to know the chance that a random centiMorgan of the genome is detectably linked. Since genetic and physical distances need not be proportional, one would expect that a screen of randomly isolated DNA probes should systematically underrepresent regions of genetic map expansion; i.e., regions in which a relatively high degree of recombination occurs in a relatively short physical stretch of DNA. Using cytogenetic distance as a proxy for physical distance (direct determination of which awaits the availability of physical maps of genomes), studies in a number of organisms report or suggest that genetic map expansion occurs in the distal portions of autosomes: in *Drosophila melanogaster*, genetic map expansion is evident in the distal half of each autosomal arm (Morton et al., 1976; Lindsley and Sandler, 1977); in the nematode worm *Caenorhabditis elegans*, genes identified by mutation are tightly clustered in the center of the genetic map and quite sparse toward the ends, strongly suggesting map expansion (Herman, 1987); and in maize, the distal 30% of the cytogenetic length of the short arm of chromosome 4 contains roughly two-thirds of the genetic length, and the distal 50% of the cytogenetic length of the short arm of chromosome 1 has yet even to be placed on the genetic map, with distal map expansion the presumed reason (E. Coe, personal communication). In humans, detailed studies of the distribution of chiasmata (the presumed cytological manifestations of crossing-over) show a significant excess near the telomeres (Laurie and Hulten, 1985a, 1985b). The most completely studied case is chromosome 1, in which the

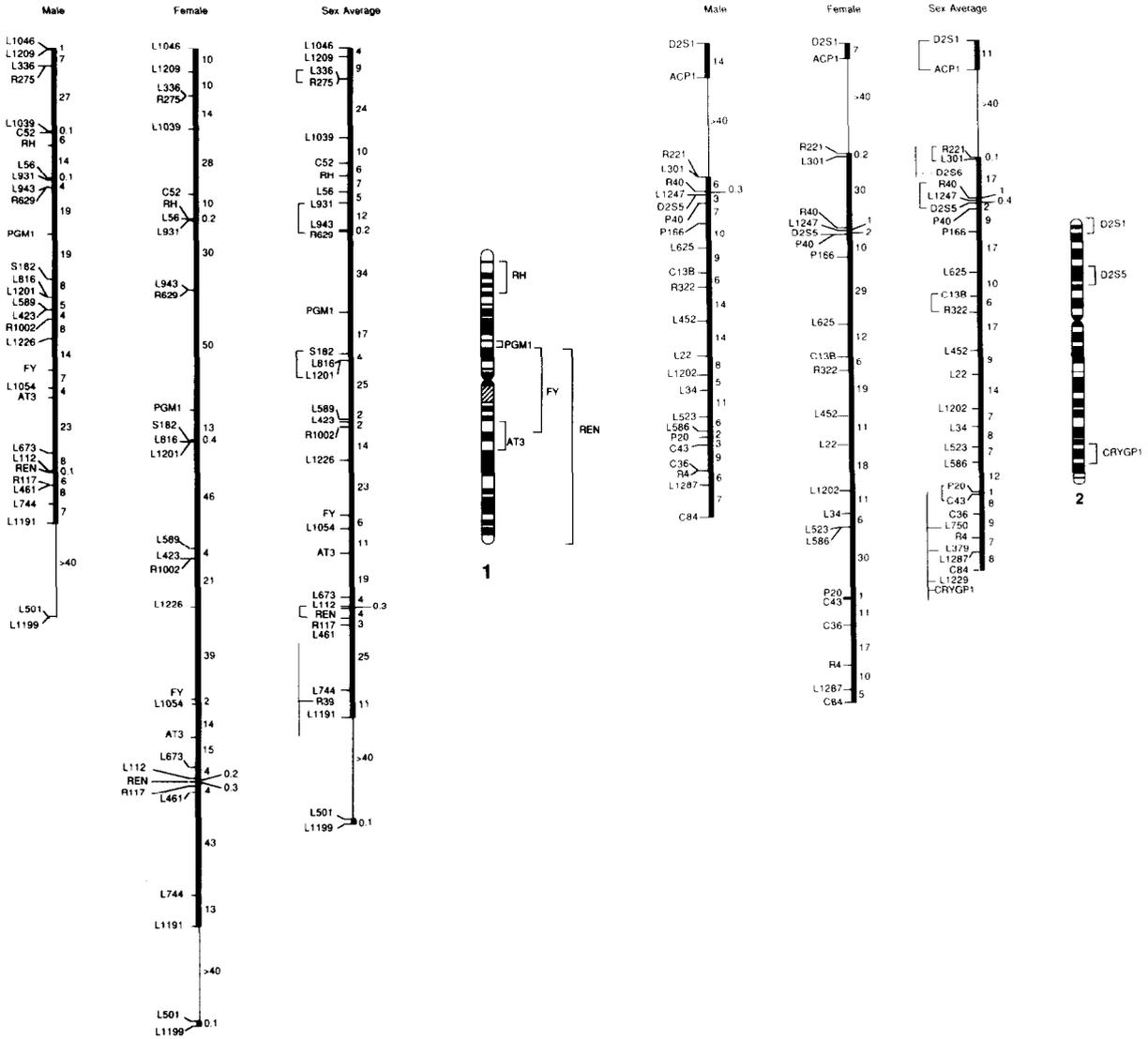
distal 15% of the chromosome contains 40% of all chiasmata (Hulten et al., 1982). Similarly, in the construction of a detailed RFLP linkage map of chromosome 21q, Tanzi and colleagues have recently found that 40% of the observed genetic length corresponds to the distal 10% of the cytogenetic arm (R. Tanzi, personal communication). Our own data for chromosome 21 show a similar effect: the loci BCE1 and SOD1 both lie in 21q2 yet are separated by an estimated genetic distance of 41 cM. Similarly, our map of chromosome 10 suggests map expansion at the distal end of 10q, in the region between PLAU and VTR.

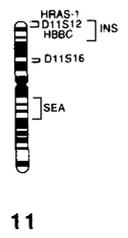
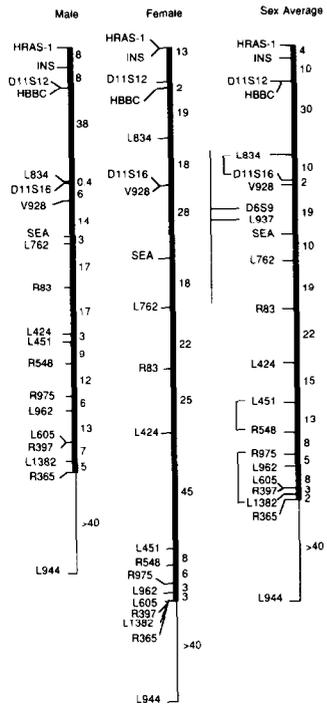
Table 4 shows the estimated minimum genetic distances observed in our map for each autosome, as well as estimates that have been obtained from counts of chiasmata in male meiosis (Morton et al., 1982; Laurie and Hulten, 1985a, 1985b). It is difficult to draw a strict comparison because, on the one hand, our estimates are simply lower bounds, which must be adjusted upward in some way to account for "end effects" and for the presence of some syntenic but as-yet-unlinked loci, and because, on the other hand, substantial methodological uncertainties still exist about counting chiasmata. Nevertheless, a rough comparison can be made. The minimum male genetic length contained within the linkage groups is about 80% of the estimate based on chiasmata counts. The results are broadly consistent with what one would expect if the documented cases of distal map expansion are taken as a rule: namely, single, large linkage groups that somewhat underestimate the total genetic length because they cannot detect linkage to the most distal regions. If we take into account the as-yet-unlinked probes known to lie on certain chromosomes, our estimate of the genetic length spanned by the mapped loci increases. Allowing a somewhat arbitrary distance of 30 cM between the linkage groups and unlinked loci would expand the estimated male length contained within the limits of the map to 2377 cM, or roughly 92% of that expected. Such close agreement is accidental, however, since the genetic lengths of the most densely covered chromosomes exceeds the estimates, while the genetic lengths of the more poorly represented chromosomes fall short.

Distribution of Polymorphism

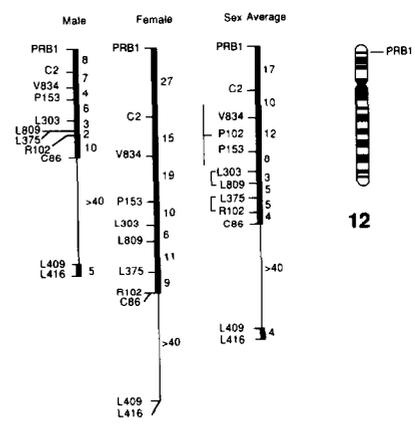
The RFLPs identified in the random genomic screen seem to lie somewhat disproportionately on the larger chromosomes: while chromosomes 1–12 contain about 70% of the cytogenetic length of the genome, some 85% of the randomly isolated RFLPs fell on these chromosomes (Table 5). The same trend is seen for each of the genomic libraries screened, the figures being 84%, 92%, and 92% for the Charon 4A library, the cosmid library, and the library propagated on recombination deficient hosts. The smaller chromosomes may in fact contain fewer polymorphisms per kilobase; alternatively, it is possible that cytogenetic length systematically overestimates the actual physical length in DNA base pairs of the smaller chromosomes.

Only three RFLPs from whole genomic libraries were identified on the X chromosome. This was expected, since the design of the random screen was systematically biased against polymorphisms on X: the libraries under-

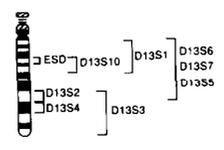
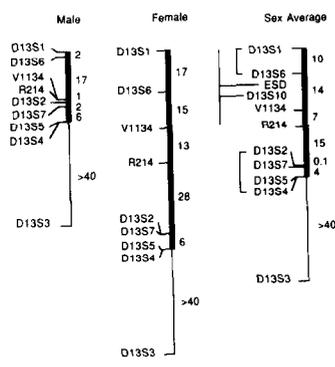




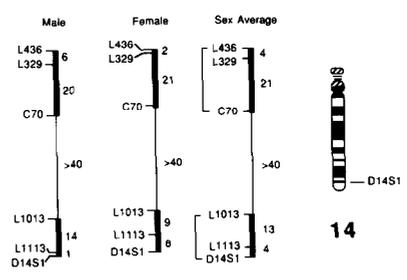
11



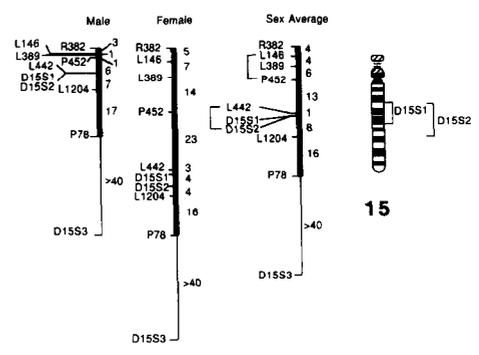
12



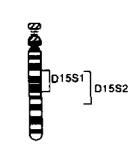
13



14



15



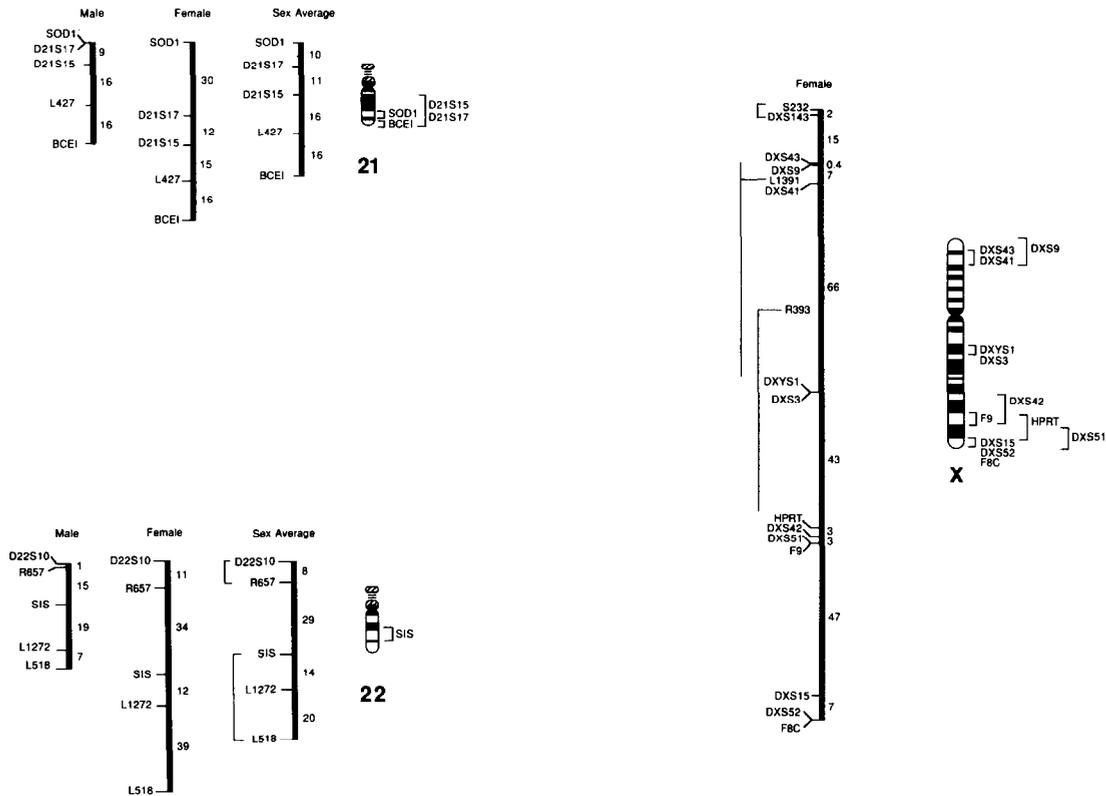


Figure 3. Genetic Linkage Maps for 23 Human Chromosomes

Loci are arranged in their most likely order, with map distances indicated in Kosambi centimorgans. Male and female distances are estimated by allowing the recombination fraction in each interval to vary between the sexes in the linkage analysis. "Sex-average" distances are obtained by repeating the linkage analysis with the constraint that the recombination fraction in each interval be equal for males and females; the "sex-average" distances are thus not simply the mean of the corresponding male and female distances. (Even though two loci are linked, the estimate for the recombination fraction in one sex, typically the female, can occasionally be quite large. Because the mapping function is extremely imprecise for large recombination fractions, we have truncated all distances at a maximum of 50 cM, corresponding to roughly 38% recombination.) The locus order indicated is favored over alternative orders by odds of at least 100:1, with exceptions indicated (on the sex-average map only) as follows: brackets enclose a set of loci whose mutual order cannot be resolved if 100:1 odds are required; when the genetic position of a locus is poorly resolved relative to the neighboring region, a thin line to the left of the chromosome indicates its range of possible positions, using the 100:1 odds test. When the thick bar representing the genetic map narrows to a thin line between two loci, this indicates that no linkage is genetically detectable between the loci; such lines merely record the fact that the loci on either side lie on the same chromosome. The minimum distance between such "unlinked" loci cannot be determined with any accuracy; given the number of meioses and the informativeness of the loci, we estimate that most such loci must be separated by at least 40 cM, and have arbitrarily indicated the genetic distance as > 40. Thirteen of the loci on the X chromosome were previously analyzed and reported (Drayna et al., 1984). Accepting this order, we have determined the location of the three additional CRI probes mapped. Subchromosomal assignment of physically localized loci is indicated to the right of the chromosome ideograms. When there are two or more loci whose genetic and physical positions have been determined and do not overlap, the genetic linkage map has been oriented relative to the physical chromosome. If such information is available for only a single locus but it lies near the telomere, we show the larger part of the linkage map pointing proximally.

Table 4. Predicted and Observed Genetic Lengths

	Predicted		Contained in Map	
	Male	Male	Female	Sex Avg.
1*	195	199	371	282
2*	173	140	236	180
3	150	151	235	180
4	138	108	292	201
5	137	139	252	192
6*	131	122	185	151
7	136	173	275	219
8*	131	111	171	147
9*	115	51	110	83
10*	127	94	111	109
11*	110	166	210	180
12*	137	45	97	68
13*	88	28	79	50
14*	88	42	40	42
15*	96	35	76	52
16	108	164	237	195
17*	109	26	116	56
18*	98	81	210	113
19	97	7	31	13
20	99	52	161	71
21	54	41	73	53
22	60	42	96	71
X	-	-	193	-
TOTAL	2577	2017	3857	2708

Estimated genetic lengths (Morton et al., 1982) are based on counts of chiasmata in male meiosis. The observed genetic lengths are the total distances contained between the loci in linkage groups. An asterisk indicates that the chromosome contains a locus or loci not yet demonstrably linked to the main linkage group; the complete collection of loci on the chromosome must in fact span a larger genetic interval than that shown. All distances are in centiMorgans.

represented X, being derived from males, and the blots used for polymorphism screening underrepresented X, since they included males. In addition, the X chromosome has been shown to contain a several-fold lower density of polymorphism (Hofker et al., 1986).

Interestingly, the most informative of the RFLPs show a tendency to cluster near the ends of the linkage map. Considering those identified in the random screen that have heterozygosity greater than 0.70, we find that 11 of 28 (39%) lie within the terminal 5% at either end of the genetic linkage map, whereas one would expect only two or three. We speculate that increased polymorphism in the most distal regions of the chromosomes may be related to increased recombination in these regions, since the recombination system itself has been implicated in the generation of mutations (Brandenburger et al., 1981; Thomas and Capocchi, 1986). Further support for this notion is provided by the pseudoautosomal region, in which recombination is tremendously enhanced relative to physical distance and in which polymorphism appears to be especially abundant (Cooke et al., 1985; Goodfellow et al., 1986; Page et al., 1987; Simmler et al., 1985).

Many of the issues raised will require further study to resolve. More precise understanding of the differences in

Table 5. Chromosomal Distribution of Loci

Chromosome	CRI RFLPs	"Anchor" Loci		Total
		CRI	Linkage Data Collected At CEPH	
1	27	2	3	32
2	24	4	1	29
3	14	3	0	17
4	16	1	2	19
5	25	3	0	28
6	11	0	8	19
7	59	4	0	63
8	15	1	0	16
9	7	2	3	12
10	6	3	0	9
11	15	3	5	23
12	11	1	0	12
13	2	3	6	11
14	5	1	0	6
15	7	3	0	10
16	40	1	1	42
17	3	4	0	7
18	7	2	0	9
19	0	4	1	5
20	4	3	0	7
21	1	4	0	5
22	3	2	0	5
X	4	0	13	17
Total	306	54	43	403

recombination rate between the sexes, more precise estimates of recombinational distances, and the resolution of local ambiguities of order will require that the genetic markers be studied in many further meioses. Correlation of the genetic and cytogenetic maps, which will shed light on the issue of distal map expansion, will require physical localization of many of the markers on the genetic map via in situ hybridization of subclones containing unique sequence, linkage mapping of many more probes whose physical location is already known, or both.

Conclusion

When we began this project some five years ago, we reasoned that the most efficient way to construct a comprehensive genetic map of the human genome was to isolate and study random polymorphic loci, rather than to proceed chromosome by chromosome. Genetic maps fall together cooperatively: when a sufficiently high density of loci is reached, linkage maps emerge for all the chromosomes at once. Recently, we reached this point. We should emphasize, however, that this genetic linkage map represents only an initial step. Some regions are relatively poorly covered, and must now be filled in. Moreover, considerably more detailed genetic maps should now be constructed, since the power of a genetic map increases considerably with its density.

The current map should be of immediate value for several purposes, including:

(a) Systematic Mapping of Simple, Single-Gene Disorders
By selecting a panel of about 100–200 probes throughout the genome, one can efficiently search nearly all regions

of the genome to locate the cause of any single gene disorder. Given the density and informativeness of the markers in the map and the additional information provided by multipoint linkage analysis, calculations show that there is a high likelihood of being able to map any single-gene disorder for which about 20–25 phase-known meioses are available, provided of course that the locus lies within the 95% of the genome currently linked.

(b) Initial Efforts to Map Diseases with Complex Modes of Inheritance

Without a map, it is difficult to map heterogeneous diseases caused by mutations at any one of several loci, since evidence for linkage to a particular locus in one family is offset by evidence against linkage in another family. Such genetic complexities become more tractable for linkage analysis if inheritance data for the whole genome are analyzed simultaneously (Lander and Botstein, 1986b). The current map is sufficient to begin such approaches, although a denser map would increase their efficiency.

(c) Systematic Deletion Mapping of Recessive Oncogenes, through Study of the Loss of Heterozygosity in Tumor Cells

A number of types of tumors, such as retinoblastoma (Cavenee et al., 1985), Wilm's tumor (Fearon et al., 1984), bilateral acoustic neuroma (Seizinger et al., 1986), and colon cancer (Solomon et al., 1987) are associated with deletions of particular chromosomal regions, which presumably expose recessive oncogenes. The availability of a full panel of probes will facilitate screening of new tumor types, by comparing DNAs from tumor tissue and peripheral blood for the loss of an allele at a polymorphic locus. Systematic screening of the genome will guard against the possibility of being misled by the nonspecific aneuploidy found in many tumors. In addition, the limits of deletions can be defined to within the resolution of the map.

(d) Rapid Assignment of Newly Discovered Polymorphisms to Chromosomes

The data underlying the linkage map characterize the pattern of inheritance (i.e., grandmaternal or grandpaternal allele) at many loci throughout the genomes of the children in the CEPH families studied. The pattern of inheritance of any new probe can now be used to identify at once its approximate chromosomal location (with at least 95% chance of success), just as is done in mapping with recombinant inbred strains in mouse. The approach can serve as a complement or an alternative to methods of physical assignment, such as *in situ* hybridization.

With a sufficiently dense linkage map, rare recessive disease genes can be mapped by using the DNA of inbred affected children to perform homozygosity mapping (Lander and Botstein, 1987), even when few families are available with multiple affected children. The current map is adequate to support preliminary efforts of this sort, but it is sparser and less evenly spaced than desirable (except in the case of chromosomes 7 and 16.) Increasing the density of the markers by a factor of 2–3 would be desirable, in order to take full advantage of these methods.

A genetic map consisting of markers spaced no more than 5 cM apart would be desirable, in order to take full advantage of the power of linkage mapping. In the current

map, the average spacing is roughly 10 cM, and many intervals are considerably larger. Since a similar mapping project of comparable scope using the CEPH families is currently underway in the laboratories of our colleagues in Salt Lake City (White et al., 1985a), we expect that a linkage map of twice the average density will result when the data from the two groups are combined. Together with the data from other groups constructing detailed linkage maps of more limited chromosomal regions, a truly complete genetic linkage map of humans will emerge.

(Note: The probes and data generated in this study will be made available to interested investigators for research purposes.)

Experimental Procedures

Pedigree DNAs and Hybrid Panels

Human DNA samples from three-generation families (typically having four grandparents, both parents, and six to 15 children) were provided by CEPH, the Centre d'Etude du Polymorphisme Humain in Paris, for CEPH families 12, 66, 1332, 1334, 1344, 1346, 1349, 1413, 1416, 1418, and 1421. DNA from families 884, 13291, 13292, 13293, 13294, 1331, 1333, 1340, 1341, 1345, 1350, and 1362 was prepared as described (Bell et al., 1981) from transformed lymphoblast lines, which were purchased from the Camden Cell Repository or were a gift from R. White. These families comprise 310 individuals; the children in them represent a total of 362 meioses. Hybrid panels used for assigning probes to chromosomes were provided by J. Frezal and colleagues and by T. Shows and colleagues.

DNA Probes

DNA probes were isolated from a total human genomic bacteriophage library (Lawn et al., 1978), a total human genomic cosmid library prepared in the vector C2RB (Bates and Swift, 1983), two phage libraries of human DNA flow-sorted for chromosomes 7 and 16, respectively (Van Dilla et al., 1986), a chromosome 16 cosmid library prepared from a rodent–human hybrid cell line bearing chromosome 16 as its only human chromosome, and three bacteriophage libraries, numbered 185, 285, and 1284, grown on hosts designed not to select against repetitive sequences (Wyman et al., 1986). Previously reported gene probes, which were used to assign linkage groups to chromosomes, are described in Table 2.

RFLP Identification and Pedigree Analysis

Human clones were tested for restriction fragment length polymorphism by hybridization to Southern blots of DNA from five unrelated individuals. Each DNA sample was digested with six or more of the restriction endonucleases BamHI, BglII, EcoRI, HincII, HindIII, MspI, PstI, RsaI, and TaqI using the conditions specified by the suppliers (New England Biolabs and BRL) except that 2–5 U/μg of DNA was used in overnight digests and MspI digests were carried out at 20°C to minimize partial digestion. Four micrograms of the resultant DNA fragments were resolved according to size by horizontal gel electrophoresis in 0.8% agarose at 3 V/cm overnight in Tris-acetate buffer, and, after staining with ethidium bromide for direct visual inspection, the DNA fragments were denatured *in situ* with 0.2 N NaOH. The DNA fragments were transferred to a nylon filter support (Zeta-Bind, AMF/Cuno, Meriden, CT) by Southern blotting, and fixed to the filter by baking at 80°C. DNA probes were labeled with [³²P]-dATP or dCTP by nick translation. Bacteriophage lambda clones used as hybridization probes were grown and DNA prepared as described (Helms et al., 1985). Cosmid DNAs were prepared by alkaline lysis of minipreps (Ish-Horowitz and Burke, 1981) or the closed circular DNA purified by CsCl/EtBr density centrifugation. Probes were hybridized to the filters in 10% dextran sulfate, 1 M NaCl, 100 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% sodium pyrophosphate, 0.5% sarkosyl, 0.5 mg/ml heparin, and 30% formamide, except that probes not known to be free of human repeated sequences were prehybridized with unlabeled human DNA (Litt and White, 1985) to block the hybridization of repeated sequences

to the membrane-bound DNA. The filters were washed twice at room temperature in $2\times$ SSC for a total of 30 min, followed by two 65°C washes of 30–60 min each in $0.1\times$ SSC, 0.2% SDS, and were exposed to X-ray film (Kodak XAR-5) with an intensifying screen (Dupont Cronex Lightening Plus) for 1–5 days at -80°C . DNA probes were stripped from the membranes in 0.1 N NaOH. The nylon membranes could be reused up to 25 times.

Candidate probes for polymorphic loci were hybridized to Southern blots of DNAs from parents of CEPH families to confirm the observed polymorphic pattern and to indicate which families were informative for linkage analysis. Analysis of the inheritance of RFLP loci by Southern hybridization was conducted on the informative families. Experimental conditions were identical to those used for RFLP identification and characterization as described above.

Linkage Analysis

Multipoint linkage analysis was performed using two independently written computer programs, CRI-MAP and MAPMAKER (Barker et al., 1987a; Lander et al., 1987). Both programs implement new techniques for maximum likelihood multipoint linkage analysis (Lander and Green, 1987), but use different strategies to find the marker order. Both programs are available from the authors. Multipoint analysis was performed without taking interference into account. The significance of sex differences in two-point crosses was evaluated by means of a likelihood ratio test. The significance of the overall sex difference seen in multipoint maps was evaluated by means of a permutation test: as a robust test of whether the likelihood increase resulting from allowing sex-specific parameters was statistically significant or, instead, simply the result of doubling the number of parameters to be fit, we computed the corresponding likelihood increase for 100 examples in which the sex of the parents had been randomly permuted, but the data were otherwise identical.

Acknowledgments

We express our thanks to Jean Dausset and the members of the CEPH consortium, through whose efforts the international collaboration on the human genetic map has become possible and practical. We thank Ron Davis for support and encouragement. We also thank Jean Frezal, Tom Shows, and their colleagues for providing hybrid panels from which numerous chromosome assignments were made, and Stephen Reeders for making a chromosome 16 hybrid cell line available as a source for generating chromosome 16-specific libraries. We thank Ed Coe, William Gelbart, Newton Morton, and David Page for helpful discussions that aided in the preparation of this manuscript. Probes for RFLP loci were kindly provided by S. Reeders (3'HVR), J. Schmidtke (pJ3.11, pJU201), G. I. Liou and C. D. B. Bridges (H.4IRBP), T. G. Krontinis (VTR4.1), J. G. Seidman (pT10), H. Willard (p17H4), S. Naylor (pHF12.8), W. Cavenee (pIE8, p7F12, p9A7), and M. Chao (pE51). We also thank the Howard Hughes Gene Mapping Library database for providing on-line database searching services. We thank P. Baud for the monumental task of entering all the raw data analyzed in this study. Finally, we thank L. LeJeune for producing the figures shown. This work was supported in part by a grant from the National Institutes of Health (AM37656-01) to H. D.-K. The National Science Foundation (DCB-8611317) and the System Development Foundation (G612) provided support to E. S. L. The participation of several authors (S. E. L., M. J. D., and J. A.) took place under the auspices of the Undergraduate Research Opportunities Program at M. I. T.

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

Received September 2, 1987.

References

Anagnou, N. P., O'Brien, S. J., Shimada, T., Nash, W. G., Chen, M. J., and Nienhuis, A. W. (1984). Chromosomal organization of the human dihydrofolate reductase genes: dispersion, selective amplification, and a novel form of polymorphism. *Proc. Natl. Acad. Sci. USA* **81**, 5170–5174.

Azen, E., Lyons, K. M., McGonigal, T., Barrett, N. L., Clements, L. S.,

Maeda, N., Vanin, E. S., Carlson, D. M., and Smithies, O. (1984). Clones from the human gene complex coding for salivary proline-rich proteins. *Proc. Natl. Acad. Sci. USA* **81**, 5561–5565.

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

Barker, D., Holm, T., and White, R. (1984b). A locus on chromosome 11p with multiple restriction site polymorphisms. *Am. J. Hum. Genet.* **36**, 1159–1171.

Barker, D., Helms, C., Akots, G., and Donis-Keller, H. (1985). RFLP analysis of probes containing repetitive DNA using probe prehybridization. *Cytogenet. Cell Genet.* **40**, 576.

Barker, D., Wright, E., Nguyen, K., Cannon, L., Fain, P., Goldgar, D., Bishop, D. T., Carey, J., Baty, B., Kivlin, J., Willard, H., Waye, J. S., Greig, G., Leinwand, L., Nakamura, Y., O'Connell, P., Leppert, M., Lalouel, J.-M., White, R., and Skolnick, M. (1987a). Gene for von Recklinghausen neurofibromatosis is in the pericentromeric region of chromosome 17. *Science* **236**, 1100–1102.

Barker, D., Green, P., Knowlton, R., Schumm, J., Lander, E., Oliphant, A., Willard, H., Akots, G., Brown, V., Gravius, T., Helms, C., Nelson, C., Parker, C., Rediker, K., Rising, M., Watt, D., Weiffenbach, B., and Donis-Keller, H. (1987b). A genetic linkage map of 63 DNA markers on human chromosome 7. *Proc. Natl. Acad. Sci. USA*, in press.

Bates, P. F., and Swift, R. A. (1983). Double cos site vectors: simplified cosmid cloning. *Gene* **26**, 137–146.

Bell, G. I., Karam, J. H., and Rutter, W. J. (1981). Polymorphic DNA region adjacent to the 5' end of the human insulin gene. *Proc. Natl. Acad. Sci. USA* **78**, 5759–5763.

Bender, W., Spierer, P., and Hogness, D. S. (1983). Chromosomal walking and jumping to isolate DNA from the *ace* and *rosy* loci and bithorax complex in *Drosophila melanogaster*. *J. Mol. Biol.* **168**, 17–33.

Bertiner, N., Duby, A. D., Morton, C. C., Leder, P., and Seidman, J. G. (1985). Detection of a frequent restriction fragment length polymorphism in the human T-cell antigen receptor beta chain locus. *J. Clin. Invest.* **76**, 1283–1285.

Bodmer, W. F., Bailey, C. J., Bodmer, J., Bussey, H. J. R., Ellis, A., Gorman, P., Lucibello, F. C., Murday, V. A., Rider, S. H., Scambler, P., Sheer, D., Solomon, E., and Spurr, N. K. (1987). Localization of the gene for familial adenomatous polyposis in chromosome 5. *Nature* **328**, 614–616.

Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* **32**, 314–331.

Brandenburger, A., Godson, G. N., Radman, M., Glickman, B. W., Van Sluis, C. A., and Doubleday, O. P. (1981). Radiation-induced basepair substitution mutagenesis in single-stranded DNA phage M13. *Nature* **294**, 180–183.

Bridges, C. (1927). The relationship of the age of the female to crossing over in the third chromosome of *Drosophila melanogaster*. *J. Gen. Physiol.* **8**, 689–700.

Brissenden, J. E., Page, D. C., de Martinville, B., Trowsdale, J., Botstein, D., and Francke, U. (1986). Regional assignments of three polymorphic DNA segments on human chromosome 15. *Genet. Epidemiol.* **3**, 231–239.

Carlson, W. R. (1977). The cytogenetics of corn. In *Corn and Corn Improvement*, G. F. Sprague, ed. (Madison, WI: American Society of Agronomy), pp. 225–304.

Cavenee, W. K., Leach, R., Mohandas, T., Pearson, P., and White, R. (1984). Isolation and regional localization of DNA segments revealing polymorphic loci from human chromosome 13. *Am. J. Hum. Genet.* **36**, 10–24.

Cavenee, W. K., Hansen, M. F., Nordernskjold, M., Kock, E., Mau-menu, I., Squire, J. A., Phillips, R. A., and Gallie, B. L. (1985). Genetic origin of mutations predisposing to retinoblastoma. *Science* **228**, 501–503.

Chang, C. C., and Kikudome, G. Y. (1974). The interaction of knobs and B chromosomes of maize in determining the level of recombination. *Genetics* **77**, 45–54.

Chao, M. V., Bothwell, M. A., Ross, A. H., Koprowski, H., Lanahan, A.

- A., Buck, C. R., and Sehgal, A. (1986). Gene transfer and molecular cloning of the human NGF receptor. *Science* 232, 518–521.
- Clarke, M. F., Westin, E., Schmidt, D., Josephs, S. F., Ratner, L., Wong-Staal, F., Gallo, R. C., and Reitz, M. S., Jr. (1984). Transformation of NIH 3T3 cells by a human c-sis cDNA clone. *Nature* 308, 464–467.
- Colb, M., Yang-Feng, T., Francke, U., Mörmer, B., Parkinson, D. R., and Krontiris, T. G. (1986). A variable tandem repeat locus mapped to chromosome band 10q26 is amplified and rearranged in leukocyte DNAs of two cancer patients. *Nucl. Acids Res.* 14, 7929.
- Cooke, H. J., Brown, W. R. A., and Rappold, G. A. (1985). Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. *Nature* 317, 687–692.
- Cooper, C. S., Park, M., Blair, D. G., Tainsky, M. A., Huebner, K., Croce, D. M., and Vande Woude, G. F. (1984). Molecular cloning of a new transforming gene from a chemically transformed human cell line. *Nature* 311, 29–33.
- Cooper, D., and Schmidtke, J. (1985). DNA RFLPs and heterozygosity in the human genome. *Hum. Genet.* 66, 1–16.
- Cooper, D. N., Smith, B. A., Cooke, H. J., Niemann, S., and Schmidtke, J. (1985). An estimate of unique DNA sequence heterozygosity in the human genome. *Hum. Genet.* 69, 201–205.
- Davatelis, G. N. (1985). Isolation and subregional mapping of an arbitrary cloned probe detecting a common RFLP on human chromosome 2. *Am. J. Hum. Genet.* 37, 1015–1021.
- Davies, K. E., Pearson, P. L., Harper, P. S., Murray, J. M., O'Brien, T., Sarfarazi, M., and Williamson, R. (1983). Linkage analysis of two cloned sequences flanking the Duchenne muscular dystrophy locus on the short arm of the human X chromosome. *Nucl. Acids Res.* 11, 2303–2312.
- Davison, M. T., and Roderick, T. H. (1981). Recombination percentages. In *Genetic Variants and Strains of the Laboratory Mouse*, M. C. Green, ed. (Stuttgart: Gustav Fisher Verlag).
- de Martinville, B., Schafer, M., White, R., and Francke, U. (1984). Assignment of RFLP loci D3S3, D20S4, and D15S1 to chromosomes 3p, 20, and 15, respectively. (HGM7). *Cytogenet. Cell Genet.* 37, 531–532.
- Donis-Keller, H., Barker, D. F., Knowlton, R. G., Schumm, J. W., Braman, J. C., and Green, P. (1986). Highly polymorphic RFLP probes as diagnostic tools. *Cold Spring Harbor Symp. Quant. Biol.* 51, 317–324.
- Drayna, D., Davies, K., Hartley, D., Mandel, J. L., Camerino, G., Williamson, R., and White, R. (1984). Genetic mapping of the human X chromosome by using restriction fragment length polymorphisms. *Proc. Natl. Acad. Sci. USA* 81, 2836–2839.
- Egeland, J. A., Gerhard, D. S., Pauls, D. L., Sussex, J. N., Kidd, K. K., Allen, C. R., Hostetter, A. M., and Housman, D. E. (1987). Bipolar affective disorders linked to DNA markers on chromosome 11. *Nature* 325, 783–787.
- Fearon, E. R., Vogelstein, B., and Feinberg, P. (1984). Somatic deletion and duplication of genes on chromosome 11 in Wilms' tumors. *Nature* 309, 176–178.
- Feder, J., Yen, L., Wijsman, E., Wang, L., Wilkens, L., Schroder, J., Spurr, N., Cann, H., Blumenberg, M., and Cavalli-Sforza, L. L. (1985). A systematic approach for detecting high-frequency restriction fragment polymorphisms using large genomic probes. *Am. J. Hum. Genet.* 37, 635–649.
- Friend, S. H., Bernards, R., Rogelj, S., Weinberg, R. A., Rapaport, J. M., Albert, D. M., and Dryja, T. P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 323, 643–646.
- Gilliam, T. C., Tanzi, R. E., Haines, J. L., Bonner, T. I., Faryniarz, A. G., Hobbs, W. J., MacDonald, M. E., Cheng, S. V., Folstein, S. E., Conneally, P. M., Wexler, N. S., and Gusella, J. F. (1987). Localization of the Huntington's disease gene to a small segment of chromosome 4 flanked by D4S10 and the telomere. *Cell* 50, 565–571.
- Goldstein, J. L., and Brown, M. S. (1983). Familial hypercholesterolemia. In *The Metabolic Basis of Inherited Disease*, 5th edition, J. B. Stanbury, J. B. Wyngaarden, D. S. Frederickson, J. L. Goldstein, and M. S. Brown, eds. (New York: McGraw-Hill), pp. 672–712.
- Goodfellow, P. J., White, B. N., Holden, J. J. A., Duncan, A. M. V., Sears, E. V. P., Wang, H.-S., Berlin, L., Kidd, K. K., and Simpson, N. E. (1985). Linkage analysis of a DNA marker localized to 20p12 and multiple endocrine neoplasia type 2A. *Am. J. Hum. Genet.* 37, 890–897.
- Goodfellow, P. J., Darling, S. M., Thomas, N. S., and Goodfellow, P. N. (1986). A pseudoautosomal gene in man. *Science* 234, 740–743.
- Goodfellow, P. J., Duncan, A. M. V., Farrer, L. A., Holden, J. J. A., White, B. N., Kidd, J. R., Kidd, K. K., and Simpson, N. E. (1987). Localization and linkage of three polymorphic DNA sequences on human chromosome 20. *Cytogenet. Cell Genet.* 44, 112–117.
- Gusella, J. F., Keys, C., Varsanyi-Breiner, A., Kao, F.-T., Jones, C., Puck, T. T., and Housman, D. (1980). Isolation and localization of DNA segments from specific human chromosomes. *Proc. Natl. Acad. Sci. USA* 77, 2829–2833.
- Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., Young, A. B., Shoulson, I., Bonilla, E., and Martin, J. B. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306, 234–238.
- Haldane, J. B. S. (1922). Sex ratio and unisexual sterility in hybrid animals. *J. Genet.* 12, 101–109.
- Helms, C., Graham, M. Y., Dutchik, J. E., and Olson, M. V. (1985). A new method for purifying lambda from phage lysates. *DNA* 4, 39–49.
- Herman, R. K. (1987). Genetics. In *The Nematode Caenorhabditis elegans*, W. B. Wood, ed. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), in press.
- Higgs, D. R., Goodbourn, S. E. Y., Wainscoat, J. S., Clegg, J. B., and Weatherall, D. J. (1981). Highly variable regions of DNA flank the human globin genes. *Nucl. Acids Res.* 9, 4213–4224.
- Hobart, P. M., Fogliano, M., O'Connor, B. A., Schaefer, I. M., and Chirgwin, J. M. (1984). Human renin gene: structure and sequence analysis. *Proc. Natl. Acad. Sci. USA* 81, 5026–5030.
- Hofker, M. H., Breuning, M. H., Bakken, E., Ommen, G. J. B., and Pearson, P. L. (1985). An anonymous single copy chromosome 22 clone, D22S10 (22c1-18) identifies an RFLP with PstI. *Nucl. Acids Res.* 13, 7167.
- Hofker, M. H., Skraastad, M. I., Bergen, A. A. B., Wapenaar, M. C., Bakker, E., Millington-Ward, A., Van Ommen, G. J. B., and Pearson, P. L. (1986). The X chromosome shows less genetic variation at restriction sites than the autosomes. *Am. J. Hum. Genet.* 39, 438–451.
- Hulten, M. A., Palmer, R. W., and Laurie, D. A. (1982). Chiasma derived genetic maps and recombination fractions: chromosome 1. *Ann. Hum. Genet.* 46, 167–175.
- Ish-Horowitz, D., and Burke, J. F. (1981). Rapid and efficient cosmid cloning. *Nucl. Acid Res.* 9, 2989–2998.
- Jackson, C. L., Bruns, G. A. P., and Breslow, J. L. (1984). Isolation and sequence of a human apolipoprotein CII cDNA clone and its use to isolate and map to human chromosome 19 the gene for apolipoprotein CII. *Proc. Natl. Acad. Sci. USA* 81, 2945–2949.
- Jakowlew, S. B., Breathnach, R., Jeltsch, J.-M., Masiakowski, P., and Chambon, P. (1984). Sequence of the pS2 mRNA induced by estrogen in the human breast cancer cell line MCF-7. *Nucl. Acids Res.* 12, 2861–2878.
- Knowlton, R. G., Cohen-Haguenaer, O., Van Cong, N., Frezal, J., Brown, V. A., Barker, D., Braman, J. C., Schumm, J. W., Tsui, L.-C., Buchwald, M., and Donis-Keller, H. (1985). A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. *Nature* 318, 381–382.
- Lander, E. S., and Botstein, D. (1986a). Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. *Proc. Natl. Acad. Sci. USA* 83, 7353–7357.
- Lander, E. S., and Botstein, D. (1986b). Mapping complex genetic traits in humans: new strategies using a complete RFLP linkage map. *Cold Spring Harbor Symp. Quant. Biol.* 51, 49–62.
- Lander, E. S., and Botstein, D. (1987). Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 236, 1567–1570.
- Lander, E. S., and Green, P. (1987) Construction of multilocus genetic

- linkage maps in humans. *Proc. Natl. Acad. Sci. USA* 84, 2363-2367.
- Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M., Lincoln, S., and Newburg, L. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, in press.
- Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1984). Strategies for multilocus linkage analysis in humans. *Proc. Natl. Acad. Sci. USA* 81, 3443-3446.
- Laurie, D. A., and Hulten, M. A. (1985a). Further studies on bivalent chiasma frequency in human males with normal karyotypes. *Ann. Hum. Genet.* 49, 189-201.
- Laurie, D. A., and Hulten, M. A. (1985b). Further studies on chiasma distribution and interference in the human male. *Ann. Hum. Genet.* 49, 203-214.
- Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G., and Maniatis, T. (1978). The isolation and characterization of linked δ - and β -globin genes from a cloned library of human DNA. *Cell* 15, 1151-1174.
- Leach, R., DeMars, D., Hasstedt, S., and White, R. (1986). Construction of a map of the short arm of chromosome 6. *Proc. Natl. Acad. Sci. USA* 83, 3909-3913.
- Lee, W.-H., Bookstein, R., Hong, F., Young, L.-J., Shew, J.-Y., and Lee, E.-H. P. (1987). Human retinoblastoma gene: cloning, identification and sequence. *Science* 235, 1394-1399.
- Leinwand, L. A., Saez, L., McNally, E., and Nadal-Ginard, B. (1983). Isolation and characterization of human myosin heavy chain genes. *Proc. Natl. Acad. Sci. USA* 80, 3716-3720.
- Leppert, M., Cavenee, W., Callahan, P., Holm, T., O'Connell, P., Thompson, K., Lathrop, G. M., Lalouel, J.-M., and White, R. (1986). A primary genetic linkage map of chromosome 13q. *Am. J. Hum. Genet.* 39, 425-427.
- Lewis, J. A., Davide, J. P., and Melena, P. W. (1982). Selective amplification of polymorphic dihydrofolate reductase gene loci in Chinese hamster lung cells. *Proc. Natl. Acad. Sci. USA* 79, 6961-6965.
- Lieman-Hurwitz, J., Dafni, N., Lavie, V., and Groner, Y. (1982). Human cytoplasmic superoxide dismutase cDNA clone: a probe for studying the molecular biology of Down Syndrome. *Proc. Natl. Acad. Sci. USA* 79, 2808-2811.
- Lindsley, D. L., and Sandler, L. (1977). The genetic analysis of meiosis in female *Drosophila melanogaster*. *Phil. Trans. Roy. Soc. Ser. B* 277, 295-312.
- Liou, G. I., Li, Y., Wang, C., Fong, S. L., Bhattacharya, S., and Bridges, C. D. B. (1987). BgIII RFLP recognized by a human IRBP cDNA localized to chromosome 10. *Nucl. Acids Res.* 7, 3196.
- Litt, M., and White, R. (1985). A highly polymorphic locus in human DNA revealed by cosmid-derived probes. *Proc. Natl. Acad. Sci. USA* 82, 6206-6210.
- Lothe, R. A., Gedde-Dahl, T., Olaisen, B., Bakker, E., and Pearson, P. (1986). Very close linkage between D2S1 and ACP1 on chromosome 2p. *Ann. Hum. Genet.* 50, 361-367.
- Lyons, K. M., Kim, H. S., Maeda, N., and Smithies, O. (1985). Insertion/deletion polymorphisms in human salivary proline-rich protein genes. *Genetics* 110, 531.
- Maeda, T. (1939). Chiasma studies in the silkworm, *Bombyx mori* L. *Japanese J. Genet.* 15, 118-127.
- Mathew, C. G. P., Chin, K. S., Easton, D. F., Thorpe, K., Carter, C., Liou, G. I., Fong, S.-L., Bridges, C. D. B., Haak, H., Kruseman, A. C. N., Schifter, S., Hansen, H. H., Telenius, H., Telenius-Berg, M., and Ponder, B. A. J. (1987). A linked genetic marker for multiple endocrine neoplasia type 2a on chromosome 10. *Nature* 328, 527-528.
- McKusick, V. A. (1986). Mendelian inheritance in Man: Catalogs of Autosomal Dominant, Autosomal Recessive, and X-Linked Phenotypes, 7th edition (Baltimore: Johns Hopkins University Press).
- Meakin, S. O., Breitman, M. L., and Tsui, L. C. (1985). Structural and evolutionary relationships among five members of the human gamma-crystallin gene family. *Mol. Cell. Biol.* 5, 1408-1414.
- Mohr, J. (1954). A Study of Linkage in Man (Copenhagen: Munksgaard).
- Monaco, A., Neve, R., Colletti-Feener, C., Bertelson, C., Kurnit, D., and Kunkel, L. (1986). Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* 323, 646-650.
- Morgan, T. H. (1914). No crossing over in the male of *Drosophila* of genes in the second and third pair of chromosomes. *Biol. Bull.* 26, 195-204.
- Morle, F., Kloepfer, C., Moisan, J. P., Mandel, J. L., Weil, D., and Grzeschik, K. H. (1984). Assignment of two unique DNA sequences which define restriction fragment length polymorphisms to chromosomes 3 and 18. *Cytogenet. Cell Genet.* 37, 544.
- Morton, N. E. (1955). Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* 7, 277-318.
- Morton, N. E., Rao, D. C., and Yee, S. (1976). An inferred chiasma map of *Drosophila melanogaster*. *Heredity* 37, 405-411.
- Morton, N. E., Lindsten, J., Iselius, L., and Yee, S. (1982). Data and theory for a revised chiasma map of man. *Hum. Genet.* 62, 266-270.
- Morton, N. E., MacLean, C. J., Lew, R., and Yee, S. (1986). Multipoint linkage analysis. *Am. J. Hum. Genet.* 38, 868-883.
- Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* 235, 1616-1622.
- Naylor, S. L., Sakaguchi, A. Y., Barker, D., White, R., and Shows, T. B. (1984). DNA polymorphic loci mapped to human chromosomes 3, 5, 9, 11, 17, 18, and 22. *Proc. Natl. Acad. Sci. USA* 81, 2447-2451.
- O'Dowd, B. F., Quan, F., Willard, H. F., Lamhonwah, A.-M., Korneluk, R. G., Lowden, A. J., Gravel, R. A., and Mahuran, D. J. (1985). Isolation of cDNA clones coding for the B subunit of human B-hexosaminidase. *Proc. Natl. Acad. Sci. USA* 82, 1184-1188.
- Ott, J. (1985). Analysis of Human Genetic Linkage (Baltimore: Johns Hopkins University Press).
- Page, D., Harper, M. E., Love, J., and Botstein, D. (1984). Occurrence of a transposition from the X chromosome long arm to the Y chromosome short arm during human evolution. *Nature* 311, 119-123.
- Page, D., Bieker, K., Brown, L. G., Hinton, S., Leppert, M., Lalouel, J. M., Nystrom-Lahti, M., and White, R. (1987). Linkage, physical mapping and DNA sequence analysis of pseudoautosomal loci on the human X and Y chromosomes. *Genomics*, in press.
- Patel, P., Nussbaum, R. L., Framson, P., Ledbetter, D., Caskey, C. T., and Chinault, A. C. (1984). Organization of the HPRT gene and related sequences in the human genome. *Somat. Cell. Mol. Genet.* 10, 483-493.
- Phillips, J. A., Penny, S. R., Kazazian, H. H., Boehm, C. D., Scott, A. F., and Smith, K. R. (1980). Prenatal diagnosis of sickle cell anemia by restriction endonuclease analysis: Hind III polymorphisms in gamma-globin genes extend test applicability. *Proc. Natl. Acad. Sci. USA* 77, 2853-2857.
- Prochowicz, E. V., Mankham, A. F., and Orkin, S. H. (1983). Isolation of a cDNA clone for human antithrombin III. *J. Biol. Chem.* 258, 8389-8394.
- Quan, F., Korneluk, R. G., Tropak, M. B., and Gravel, R. A. (1986). Isolation and characterization of the human catalase gene. *Nucl. Acids Res.* 14, 5321-5335.
- Reeders, S. T., Breuning, M. H., Davies, K. E., Nicholls, R. D., Jarman, A. P., Higgs, D. R., Pearson, P. L., and Weatherall, D. J. (1985). A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. *Nature* 317, 542-544.
- Renwick, J. H., and Schulze, J. (1965). Male and female recombination fraction of the nail-patella:ABO linkage in man. *Ann. Hum. Genet.* 28, 379-392.
- Rhoades, M. M. (1941). Different rates of crossing over in male and female gametes of maize. *J. Am. Soc. Agron.* 33, 603-615.
- Robertson, D. S. (1984). Different frequency in the recovery of crossover products from male and female gametes of plants hypoploid for B-A translocations in maize. *Genetics* 107, 117-130.
- Rouyer, F., Simmler, M. C., Johnsson, C., Vergnaud, G., Cooke, H. J., and Weissenbach, J. (1986). A gradient of sex-linkage in the pseudoautosomal region of the human sex chromosomes. *Nature* 317, 692-697.

- Royer-Pokora, B., Kunkel, L. M., Monaco, A. P., Goff, S. C., Newburger, P. E., Baehner, R. L., Cole, F. S., Curnutte, J. T., and Orkin, S. H. (1986). Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location. *Nature* 322, 32–38.
- St George-Hyslop, P. H., Tanzi, R. E., Polinsky, R. J., Haines, J. L., Nee, L., Watkins, P. C., Myers, R. H., Feldman, R. G., Pollen, D., Drachman, D., Growdon, J., Bruni, A., Foncin, J.-F., Salmon, D., Frommelt, P., Amaducci, L., Sorbi, S., Piacentini, S., Stewart, G. D., Hobbs, W. J., Conneally, P. M., and Gusella, J. F. (1987). The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science* 235, 885–890.
- Schumm, J., Knowlton, R., Braman, J., Barker, D., Botstein, D., Akots, G., Brown, V., Gravius, T., Helms, C., Hsiao, K., Rediker, K., Thurston, J., and Donis-Keller, H. (1987). Identification of more than 500 restriction fragment length polymorphisms by screening random genomic clones. *Am. J. Hum. Genet.*, in press.
- Schwartz, C. E., McNally, E., Leinwand, L., and Skolnick, M. H. (1985). New highly informative RFLPs at the D17S2 and MYH2 loci and linkage between MYH2 and D17S1. (HGM8). *Cytogenet. Cell Genet.* 40, 740.
- Seizinger, B. R., Martuza, R. L., and Gusella, J. F. (1986). Loss of genes on chromosome 22 in tumorigenesis of human acoustic neuroma. *Nature* 322, 644–647.
- Seizinger, B. R., Rouleau, G. A., Ozelius, L. J., Lane, A. H., Faryniarz, A. G., Chao, M. V., Huson, S., Korf, B. R., Parry, D. M., Pericak-Vance, M. A., Collins, F. S., Hobbs, W. J., Falcone, B. G., Iannazzi, J. A., Roy, J. C., St George-Hyslop, P. H., Tanzi, R. E., Bothwell, M. A., Upadhyaya, M., Harper, P., Goldstein, A. E., Hoover, D. L., Bader, J. L., Spence, M. A., Mulvihill, J. J., Aylsworth, A. S., Vance, J. M., Rossenwasser, G. O. D., Gaskell, P. C., Roses, A. D., Martuza, R. L., Breakfield, X. O., and Gusella, J. F. (1987). Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor receptor gene. *Cell* 49, 589–594.
- Shaw, D. J., Meredith, A. L., Sarfarazi, M., Harley, H. G., Huson, S. M., Brook, J. D., Bufton, L., Litt, M., Mohandas, T., and Harper, P. S. (1986). Regional localisations and linkage relationships of seven RFLPs and myotonic dystrophy on chromosome 19. *Hum. Genet.* 74, 262–266.
- Shiloh, Y., Kanda, N., Kunkel, L. M., Bruns, G., Sakai, K., and Latt, S. A. (1985). Two RFLPs identified by a human chromosome #2 clone at 2p15–2p16 (HGM8 assignment no. D2S5). *Nucl. Acids Res.* 13, 5403.
- Simmler, M. C., Rouyer, F., Vergnaud, G., Nystrom-Lahti, M., Ngo, K. Y., de la Chapelle, A., and Weissenbach, J. (1985). Pseudoautosomal DNA sequences in the pairing regions of the human sex chromosomes. *Nature* 317, 692–697.
- Simpson, N. E., Kidd, K. K., Goodfellow, P. J., McDermid, H., Myers, S., Kidd, J. R., Jackson, C. E., Duncan, A. M. V., Farrer, L. A., Brasch, K., Castiglione, C., Genel, M., Gertner, J., Greenberg, C. R., Gusella, J. F., Holden, J. J. A., and White, B. N. (1987) Assignment of multiple endocrine neoplasia type 2a to chromosome 10 by linkage. *Nature* 328, 528–530.
- Smith, C. A. B. (1986). The development of human linkage analysis. *Ann. Hum. Genet.* 50, 293–311.
- Smith, M., Povey, S., Arredondo-Vega, F. X., Duester, G., Kieley, C., Jeremiah, S., and Hopkinson, D. A. (1984). Mapping of human class 1 alcohol dehydrogenase (ADH). *Cytogenet. Cell Genet.* 37, 586.
- Smith, M., Duester, G., Carlock, L., and Wasmuth, J. (1985). Assignment of ADH1, ADH2 and ADH3 genes (class I ADH) to human chromosome 4q21–4q25. *Cytogenet. Cell Genet.* 40, 748.
- Solomon, E., and Bodmer, W. F. (1979). Evolution of sickle cell variant gene. *Lancet* *i*, 923.
- Solomon, E., Voss, R., Hall, V., Bodmer, W. F., Jass, J. R., Jeffreys, A. J., Lucibello, F. C., Pagtel, I., and Rider, S. H. (1987). Chromosome 5 allele loss in human colorectal carcinomas. *Nature* 328, 616–619.
- Srinivasan, A., Reddy, E. P., and Aaronson, S. A. (1981). Abelson murine leukemia virus: molecular cloning of infectious integrated proviral DNA. *Proc. Natl. Acad. Sci. USA* 78, 2077–2081.
- Stewart, G. D., Gallagher, J., Harris, P., Young, B. D., and Ferguson-Smith, M. A. (1984). Isolation and characterization of cloned DNA from a DNA bank enriched for human chromosome 21 (HGM7). *Cytogenet. Cell Genet.* 37, 593.
- Sturtevant, A. H. (1913). The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.* 14, 43–59.
- Sturtevant, A. H. (1965). *A History of Genetics* (New York: Harper and Row).
- Tazima, Y. (1964). *The Genetics of the Silkworm* (Englewood Cliffs, NJ: Prentice-Hall).
- Thomas, K. R., and Capocchi, M. (1986). Introduction of homologous DNA segments into mammalian cells induces mutations in the cognate gene. *Nature* 324, 34–38.
- Tsipouras, P., Myers, J. C., Ramirez, F., and Prockop, D. (1983). RFLP associated with the proalpha2(I) gene of human type I procollagen. *J. Clin. Invest.* 72, 1262–1267.
- Tsui, L.-C., Buchwald, M., Barker, D., Braman, J. C., Knowlton, R., Schumm, J. W., Eiberg, H., Mohr, J., Kennedy, D., Plavsic, N., Zsiga, M., Markiewicz, D., Akots, G., Brown, V., Helms, C., Gravius, T., Parker, C., Rediker, K., and Donis-Keller, H. (1985). Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science* 230, 1054–1057.
- Van Dilla, M. A., Deaven, L. L., Albright, K. L., Allen, N. A., Aubuchon, M. R., Bartholdi, M. F., Brown, N. C., Campbell, E. W., Carrano, A. V., Clark, L. M., Cram, L. S., Crawford, B. D., Fuscoe, J. C., Gray, J. W., Hildebrand, C. E., Jackson, P. J., Jett, J. H., Longmire, J. L., Lozes, C. R., Luedemann, M. L., Martin, J. C., McNinch, J. S., Meincke, L. J., Mendelsohn, M. L., Meyne, J., Moyzis, R. K., Munk, A. C., Perlman, J., Peters, D.C., Silva, A. J., and Trask, B. J. (1986). Human chromosome-specific DNA libraries: construction and availability. *Biotechnology* 4, 537–552.
- Wainwright, B. J., Scambler, P. J., Schmidtke, J., Watson, E. A., Law, H. Y., Farrall, M., Cooke, H. J., Eiberg, H., and Williamson, R. (1985). Localization of cystic fibrosis locus to human chromosome 7cen-q22. *Nature* 318, 384–386.
- White, R., Leppert, M., Bishop, D. T., Barker, D., Berkowitz, J., Brown, C., Callahan, P., Holm, T., and Jerominski, L. (1985a). Construction of linkage maps with DNA markers for human chromosomes. *Nature* 313, 101–105.
- White, R., Woodward, S., Leppert, M., O'Connell, P., Hoff, M., Herbst, J., Lalouel, J.-M., Dean, M., and Vande Woude, G. (1985b). A closely linked genetic marker for cystic fibrosis. *Nature* 318, 382–384.
- White, R., Leppert, M., O'Connell, P., Nakamura, Y., Julier, C., Woodward, S., Silva, A., Wolff, R., Lathrop, M., and Lalouel, J. M. (1986). Construction of human genetic linkage maps: I. Progress and perspectives. *Cold Spring Harbor Symp. Quant. Biol.* 51, 29–38.
- Willard, H., Skolnick, M., Pearson, P. L., and Mandel, J. L. (1985). Report of the committee on human gene mapping by recombinant DNA techniques. In *Human Gene Mapping 8*. *Cytogenet. Cell. Genet.* 40, 360–489.
- Willard, H. F., Wayne, J. S., Skolnick, M. H., Schwartz, C. E., Powers, V. E., and England, S. B. (1986). Detection of restriction fragment length polymorphism at the centromere of human chromosomes by using chromosome specific alpha satellite DNA probes: implications for development of centromere-based genetic linkage maps. *Proc. Natl. Acad. Sci. USA* 83, 5611–5615.
- Wyman, A., White, R., and Francke, U. (1982). Assignment of the first random restriction fragment length polymorphism (RFLP) locus (D14S1) to a region of human chromosome 14. *Am. J. Hum. Genet.* 34, 216–226.
- Wyman, A., Wertman, K., Barker, D., Helms, C., and Petri, W. H. (1986). Factors which equalize the representation of genome segments in recombinant libraries. *Gene* 49, 263–271.
- Zabel, B. U., Naylor, S. L., Sakaguchi, A. Y., Bell, G. I., and Shows, T. B. (1983). High-resolution chromosomal localization of human genes for amylase, pro-opiomelanocortin, somatostatin, and a DNA fragment (D3S1) by in situ hybridization. *Proc. Natl. Acad. Sci. USA* 80, 6932–6936.

Notes Added in Proof

D. B. is at the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. E. S. L. is also affiliated with Harvard University, Cambridge, Massachusetts 02138.