Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 14771–14775, December 1996
Medical Sciences

A metric map of humans: 23,500 loci in 850 bands

(map integration/cytogenetic band/interference)

A. COLLINS*, J. FREZAL†, J. TEAGUE*, AND N. E. MORTON*

*Human Genetics, Level G, Princess Anne Hospital, Southampton SO16 5YA, United Kingdom; and †Genatlas, Service de Génétique Médicale, Hôpital des Enfants Malades, 149 Rue de Sèvres, 74743 Paris Cedex 15, France

Contributed by N. E. Morton, September 12, 1996

ABSTRACT **High-resolution maps integrated with the enhanced location data base software (LDB+) give improved estimates of genetic parameters and reveal characteristics of cytogenetic bands. Chiasma interference is intermediate between Kosambi and Carter–Falconer levels, as in *Drosophila* and the mouse. The autosomal genetic map is 2832 and 4348 centimorgans in males and females, respectively. Telomeric T-bands are strikingly associated with male recombination and gene density. Position and centromeric heterochromatin have large effects, but nontelomeric R-bands are not significantly different from G-bands. Several possible reasons are discussed. These regularities validate the maps, despite their high resolution and inevitable local errors. No other approach has been demonstrated to integrate such a large number of loci, which are increasing at about 45% per year. The maps and the data and software from which they are constructed are available through the Internet (http://cedar.genetics.soton.ac.uk/public_html). Successive versions of this location data base may also be accessed on CD-ROM.**

Resolution of the human genome has increased by about 45% per annum during the past 10 years (Fig. 1). If this exponential growth continues, the map will include a million STS-based loci by the year 2005, when sequencing the human genome may be complete (4). Despite these advances, most current maps are metric only for linkage and radiation hybrids, the physical map providing order but not distance (3). This leads to graphical alignment of a small subset of maps instead of metric integration into a composite map that gives distance as well as order. To remedy this deficit, we have developed algorithms that make the map fully metric by integrating linkage, radiation hybrid, and physical evidence (5). Applying these algorithms to data accessible through Internet doubles map resolution and reveals characteristics of chromosomes and cytogenetic bands that validate the map, improve estimates of genetic parameters, and elucidate aspects of chromosome structure and function.

METHODS

The enhanced *l*ocation *d*ata *b*ase (LDB+) comprises a suite of software for map construction together with high-resolution metric maps of the human genome (6). Integrated summary maps in LDB+ contain fields for the composite location (comp) and physical location (phymb) as megabases (Mbs) from the p telomere, radiation hybrid location (rhmb) scaled to the physical length, genetic map (mcM, male centimorgans; fcM, female centimorgans), and cytogenetic assignment. Sex-specific linkage maps were generated from pairwise lods computed on Centre d'Etude du Polymorphisme Humain data base version 7.1 (7), the LODSOURCE data base (8), and from an incomplete literature search, using the MAP+ program (9). MAP+ uses an algorithm for locus-oriented improvement of a trial map and estimates the Rao mapping parameter p (10),

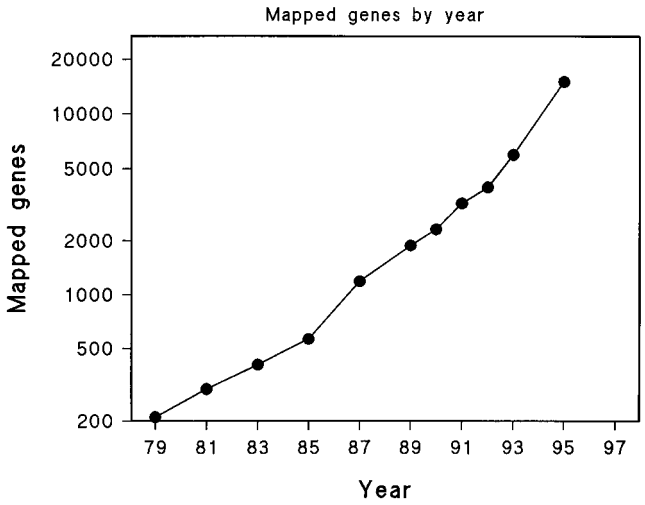

FIG. 1. Mapped genes by year. During this time mapping standards and gene definitions have changed: 1979–1983, autosomal confirmed and provisional (1); 1985–1993, polymorphic genes and D segments (2); 1995, yeast artificial chromosome (YAC) sequence tagged sites (STS) content map (3).

where 1-p is a measure of interference (Table 1), together with the typing error frequency $\in$ (13). These high-resolution maps update earlier estimates of genetic parameters (11). There are few discrepancies in order between these maps and the recently published Généthon map of 5264 microsatellites (12). Additional locations in published linkage maps for which lods were not available were entered by interpolation using the LDB+ program (5). Radiation hybrid maps were incorporated from publicly available data (3, 14) and by using the radiation hybrid mapping option in MAP+, which uses an algorithm that considers error, polysomy, and preferential retention of markers close to the centromere (15). Cytogenetic assignments and mouse homology were captured from the tabular output of Genome data base and Genome Interactive data base (16). Physical data came from two Internet sites (3, 7), and an incomplete literature search. Hudson *et al.* (3) presented an STS content map of 10,850 loci as a number of YAC contigs per chromosome. Although there were likely to be local errors due mostly to gaps and overlaps, we treated the contigs as contiguous for the purposes of entering the data into LDB+.

To place physically ordered loci on a physical (megabase) scale, we considered using radiation hybrid maps. However, the possibility of radiation breakage hotspots (15) and large "holes" in the centromere region (3) made this unattractive at a time when radiation hybrid maps are in flux and show

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

Abbreviations: STS, sequence tagged site; YAC, yeast artificial chromosome; Mb, megabase; cM, centimorgan; MD and FD, male and female cM/Mb, respectively; LK, log (female cM/male cM); LM and LF, loci in linkage map/male and female cM, respectively; PD, loci in linkage map/Mb; PC, principle component of MD, FD, LK, LM, LF, and PD.

substantial variation (15). We chose therefore to adopt a method whereby cytogenetic band assignments provide physical locations on the assumption that band width is proportional to DNA content (17). The digitized idiograms of Francke (18) at 850 band resolution provided an update to the LBD+ band files. The band file gives estimates of left and right cytogenetic band borders in megabases from the p telomere, scaling each chromosome arm to physical estimates on the assumption of 3200 Mb per genome (11). From the band file and cytogenetic assignments for loci in the integrated linkage map, we obtained a crude estimate of physical location for each locus as the cytogenetic band midpoint. The sex-averaged genetic map was then transformed onto a physical scale by fitting a double logistic curve to each chromosome arm using the band midpoint as the dependent variable and weighting by the inverse of the band width (1/S). As many cytogenetic assignments in Genome data base are inferences from preliminary linkage data, we rejected $1/S^2$ as a weight; although appropriate to a uniform distribution, it gives excessive influence to narrow band assignments that may be in error. A number of erroneous assignments for disease genes in Genome data base were corrected in Genome Interactive data base (16). From this projection, a set of reliably ordered markers was obtained by selecting a subset where the linkage map order agreed with YAC-based contigs (3). The resulting map was entered in the LDB+ phymb field to provide a metric scale for order-only maps and connectivity for smaller segments. From the available data, a composite map on the physical scale was created by interpolation with priority physical > radiation hybrid > genetic > cytogenetic (5).

A location data base must provide the best current estimates, but keep them rigorously separated from data. To evaluate cytogenetic band characteristics, we constructed a file (gmap) for each chromosome that defines regions, cytogenetic bands, and genetic locations for all loci in the summary map. We divided each chromosome into 26 equal regions designated A to Z. Each region was divided into 10 equal subregions designated 0 to 9. This two-character code generated $24 \times 26 \times 10 = 6240$ genome subregions of average size about 500 kb, an order of magnitude greater than high-resolution banding, and appending a third character would give 162,240 microregions of average size about 20 kb. Assignment of loci to a subregion depends on the estimated physical location and therefore must improve. To assign loci to bands by projection of genetic location, we constructed an anchor map for which locus order was confirmed by linkage and YAC-based maps. In each region, one or more STSs was selected as anchors, giving preference to D numbers over local symbols and to polymorphic loci in the Généthon map, without regard to expression. Rank (rk) is 3 for regional anchors, 2 for subregional anchors, 1 for any other loci in the projected genetic map, 0 for other loci in the genetic, physical, or radiation hybrid maps, and c for loci with no location except assignment cytogenetically to a band of width 10 Mb or less. Regions, subregions, and anchors provide connectivity between mapping and sequencing, and anchors are used for interpolation of missing data in the gmap file. Despite the precision that fluorescence *in situ* hybridization makes possible, most of the current cytogenetic assignments are so broad as to be of limited use for mapping. We therefore estimated from the composite location the likely assignment of an anchor and its associated interval to a single cytogenetic band (bd). All these projections are regenerated from the summary map after each significant accession of data; the old projection is discarded. The quality of these estimates has improved dramatically during the past year, and the greater resolution that they provide is invaluable for selecting markers

Table 1. Parameters of the human genome

					Centromere location							
			Genetic map, cM			Genetic map, cM			Genethon map, cM (12)			
Chromosome	No. of loci	Physical map, Mb*	Male	Female	Physical, Mb	Male	Female	Chiasma male, cM (11)	Male	Female	Interference, 1-p†	Typing error frequency
1	1,693	263	221	376	128	118	199	201	220	358	0.68	0.008
2	1,515	255	193	297	99	76	115	185	211	325	0.69	0.002
3	1,602	214	186	289	99	90	153	163	183	269	0.66	0.009
4	1,705	203	157	274	56	51	82	141	157	271	0.71	0.006
5	1,264	194	149	267	52	53	94	144	147	242	0.60	0.012
6	1,157	183	142	222	65	68	106	146	135	265	0.46	0.003
7	1,034	171	147	241	65	61	94	151	178	187	0.72	0.005
8	972	155	135	226	50	61	74	138	113	221	0.71	0.004
9	813	145	130	176	51	53	63	129	138	194	0.75	0.004
10	1,049	144	144	192	44	57	60	129	146	210	0.53	0.005
11	1,386	144	125	189	58	59	70	116	122	180	0.52	0.007
12	949	143	136	232	39	50	71	142	126	212	0.54	0.002
13	695	114	112	152	16	0	0	100	97	132	0.55	0.007
14	659	109	106	151	16	0	0	100	104	154	0.75	0.007
15	570	106	84	150	17	0	0	100	92	131	0.66	0.003
16	942	98	110	152	39	57	54	111	98	169	0.75	0.005
17	779	92	108	152	28	54	46	114	104	145	0.73	0.007
18	702	85	111	149	20	50	53	113	93	151	0.78	0.004
19	711	67	113	121	30	57	53	100	98	115	0.78	0.012
20	494	72	104	120	31	56	49	100	73	120	0.73	0.007
21	656	50	62	81	11	0	0	58	47	71	0.71	0.009
22	395	56	78	89	13	0	0	70	47	75	0.73	0.005
X	1,767	164	—	193	62	—	90	50‡	—	198	0.59	0.002
Y	—	59	—	—		—	—	50‡	—	—		
Autosomes	21,742	3063	2832	4348				2751	2729	4197		
Total	23,509											

*Assuming 3200 Mb for the sex-averaged genome (11).
†*P* is the mapping parameter in the Rao function, where p = 0 represents complete interference and p = 1 null interference (10).
‡Pseudoautosomal region in males.

Table 2. Definition of variables

Category	Variable name	Definition
Biological property	MD	Male cM/Mb
	FD	Female cM/Mb
	LK	Log (female cM/male cM)
	LM	Loci in linkage map/male cM
	LF	Loci in linkage map/female cM
	PD	Loci in linkage map/Mb
	PC	Principal component of above = 0.563 MD + 0.441 FD − 0.422 LK − 0.373 LM + 0.013 LF + 0.415 PD
Band	RBD	R-bands (1, 0) (ref. 18)
	GBD	G-band shades from light grey to black (2, 3, 4, and 5, elsewhere = 0) (ref. 18)
	POS	Midband position in range 0, 1, where 0 is the centromere and 1 the telomere
	TEL	Telomeric band (1, 0)
	HET	Variable centromeric heterochromatin (1, 0); acrocentric heterochromatin rejected
	CBD	Nonvariable C-band (1, 0)
	AVM	Centromeric band = 1 in acrocentrics, −1 in metacentrics, 0 elsewhere
Chromosome	CLG	Length of chromosome in Mb
	MCR	Acrocentric (0) or metacentric (1) chromosome

in regions of interest. Discrepancies with recently published maps are few and small.

The projected summary maps defined in this way provide an opportunity to examine properties of the genome at 850 band resolution. We constructed a datafile characterizing each band in terms of recombination rate, sex ratio for recombination, loci per centimorgan, loci per megabase, G and R banding, various positional correlates, and chromosome size and shape characteristics (Table 2). The location of band borders on the physical scale was defined by the band file and on the genetic scale by linear interpolation between the closest flanking markers for each band border. For the loci per cM variables (LM, males; LF, females), where a band showed less than 1 cM recombination in either sex an interval ± 0.5 cM around band midpoint was selected and the variables recalculated. For the sex ratio of recombination K, a logarithmic transformation was used to damp extreme values where recombination in the band is very low. Giemsa bands were classified according to Francke (18), and the presence or absence of R-bands was denoted by 1 or 0, respectively.

Biological Properties of Chromosomes. Table 1 gives current estimates of mapping parameters, which supersede earlier estimates. Physical map lengths based on 3200 Mb for the sex-averaged genome remain unchanged, as do estimates of male genetic length based on chiasmata (11). There is good agreement among estimates of genetic lengths from these high-resolution maps, the male chiasma map, and the Généthon maps of 5264 microsatellites, although the latter tend to be slightly smaller (12). This is consistent with a paucity of microsatellites near centromeres and telomeres (19). Genome parameters are much more accurate than they were 5 years ago, when female lengths were based on fragmentary data.

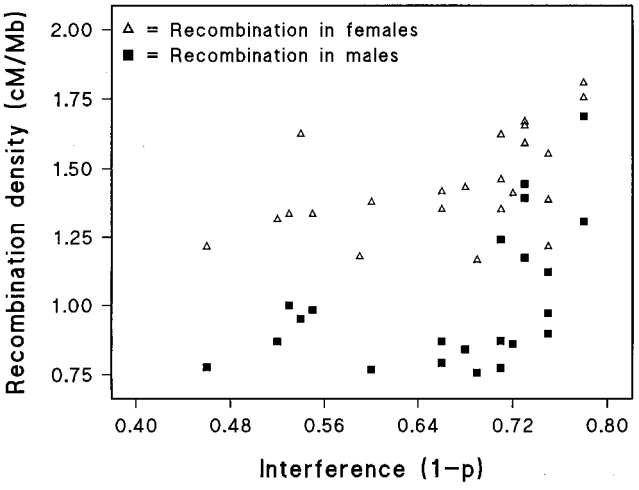

FIG. 2. Interference by recombination density.

The mean error frequency is estimated to be 0.0056 per marker. Interference as the complement of the Rao mapping parameter is 0.67, in remarkably good agreement with the value of 0.65 estimated from male chiasmata (10) and linkage data in mouse and *Drosophila* (20). This represents greater interference than for the Kosambi function (0.5) but less than for the Carter–Falconer function (0.75) (10). We examined correlations of these parameters with the measured variables (Table 2), weighting by band width to give a mean for each chromosome. The error frequency is positively correlated with locus density (PD, r = 0.45, $P = 0.030$), perhaps because dense maps are less carefully checked or the frequency of local errors in order increases with density and simulates error. Error and interference appear to be independent (r = 0.20, $P = 0.36$). Interference is positively correlated with recombination density in both sexes (FD, r = 0.51, $P = 0.01$; MD, r = 0.49, $P = 0.02$). This suggests that there is a minimal physical distance within which meiotic double recombinants cannot occur, the importance of this constraint increasing with recombination density (Fig. 2). The ratio of female/male recombination (LK) is strongly correlated with chromosome length (CLG, r = 0.70, $P = 0.0003$), presumably because the telomeric band with its low value of LK is an increasing fraction of chromosome arm length as the latter decreases. This is consistent with a significant negative correlation between LK and the density variables PD, MD, and FD, all of which are positively associated with telomeres.

Table 3. Values of R^2 for dependent variables

	MD	FD	LK	LM	LF	PD	PC
R^2	0.641	0.459	0.220	0.179	0.158	0.136	0.324

See Table 2 for definitions.

Table 4. Regression models for bands (variables MD, FD, LM, and LF)

		Male			Female		
Dependent variable	Independent variable	Coefficient	SE	Single variable correlation	Coefficient	SE	Single variable correlation
Centimorgan density (cM/Mb) (MD, FD)	INTERCEPT	−0.016	0.069	—	0.580	0.109	—
	TEL	3.237	0.132	0.704***	0.751	0.139	0.345***
	PD	0.197	0.014	0.501***	0.305	0.015	0.616***
	POS	0.502	0.117	0.362***	0.389	0.116	0.238***
	AVM	0.889	0.186	0.101**	—	—	0.037
	CBD	0.761	0.196	−0.035	—	—	0.029
	MCR	—	—	−0.022	−0.189	0.091	−0.015
	HET	—	—	−0.103**	−0.516	0.171	−0.158***
Loci per centimorgan (LM, LF)	INTERCEPT	8.401	0.805	—	1.817	0.119	—
	POS	−4.149	0.801	−0.310***	−0.726	0.201	−0.090*
	CLG	−0.022	0.004	−0.145***	—	—	0.014
	MCR	2.746	0.684	0.070*	—	—	0.032
	TEL	−3.972	0.912	−0.181***	−0.645	0.238	−0.019
	AVM	−3.693	1.063	−0.207***	—	—	−0.094**
	HET	6.392	1.567	0.116**	—	—	−0.024
	PD	0.305	0.095	0.033	0.305	0.025	0.358***

*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.

The probability that a significant linkage between two random loci is true depends on ϕ, the probability that they lie on the same chromosome. Neglecting ϕ leads to a Bonferroni calculus in which all significant results are type 1 errors (21). At a time when there were no critical observations on human chromosomes (and the diploid number was still thought to be 48), the available data on experimental organisms suggested $\phi = 0.05$ for human autosomes (22). The basic formula is $\phi = \Sigma L_i^2/(\Sigma L_i)^2$, where L_i is proportional to the number of loci on the ith chromosome. If gene number is proportional to physical length, the estimate for autosomes is 0.054 (23, 24), becoming 0.051 if the X chromosome is included as is appropriate for complex inheritance. We have seen that gene density decreases as physical length increases. If gene number is proportional to genetic length (25), the corresponding estimate from Table 1 is 0.051. For mapped loci, $\phi = 0.050$. A reasonable consensus is $\phi = 0.051$. From this value and assumptions about power, it seems that a lod of 3 if correctly calculated is necessary but not sufficient to make a linkage test reliable, and this applies to complex inheritance as well as major loci (26).

Biological Properties of Bands. Analysis of bands encounters three problems. First, assignment of a locus is based on position in the summary map and must entail significant error, especially for small bands. Second, high-resolution bands are poorly characterized with respect to GC-content and other parameters (27). Third, loci in the location data base genetic map are predominantly microsatellites, but other loci are also included. Properties of this unresolved mixture are not necessarily shared by all constituents. Given these reservations, forward stepwise regression was carried out with band width as a weight and significance level 0.05 for retention in the model. The squared multiple correlation R^2 ranged from 0.1 to 0.6, indicating fair to good prediction (Table 3). The "inde-

Table 5. Regression models for bands (variables LK, PD, and PC)

Dependent variable	Independent variable	Coefficient	SE	Single variable correlation
Log K the ratio female/male recombination (LK)	INTERCEPT	1.132	0.105	—
	TEL	−1.021	0.122	−0.344***
	POS	−0.580	0.109	−0.345***
	HET	0.679	0.214	0.105**
	CLG	−0.002	0.0005	−0.070
	MCR	0.358	0.094	0.071*
	AVM	−0.287	0.146	−0.154***
Loci per megabase (PD)	INTERCEPT	2.571	0.250	—
	TEL	2.312	0.314	0.266***
	HET	−2.194	0.384	−0.156***
	POS	0.570	0.269	0.148***
	CLG	−0.003	0.001	−0.038
Principal component (PC) of above	INTERCEPT	−1.962	0.410	—
	TEL	5.174	0.460	0.437***
	POS	2.801	0.423	0.397***
	AVM	3.437	0.840	0.192***
	HET	−3.609	0.806	−0.127***
	CLG	0.007	0.002	0.042
	MCR	−1.270	0.355	−0.084*
	CBD	2.343	0.864	−0.115**

*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.

pendent" variables are intercorrelated, and the centromeric band variable (CBD) takes on different signs in single and multiple regression for the principal component and male centimorgan density.

Recombination density (MD and FD) is highest in telomeric bands, especially in males, and lowest in centromeric heterochromatin (Tables 4 and 5). Euchromatic bands adjacent to centromeres of acrocentric (but not metacentric) chromosomes have high recombination density, which increases toward the telomere independently of other variables. Bands with high locus density have high recombination density. This analysis recovers one of Haldane's rules: that chiasmata are more localized in the heterogametic sex. Preference for telomeric recombination in males must facilitate crossing over in the pseudoautosomal region, which is necessary for normal X/Y disjunction.

Locus density on the genetic map (LM and LF) has the larger number of significant correlates in males, the positive associations with acrocentric chromosomes, centromeric heterochromatin, and metacentric centromeres being nonsignificant in females. Apparently, acrocentric centromeres in males do not act as obligatory chiasmata, permitting a genetic length that more closely approaches the female length as in the mouse.

At the band level, the female/male recombination ratio (LK) is significantly associated with other variables besides chromosome length, decreasing with relative distance from the centromere but increasing in heterochromatin and for metacentric chromosomes, supporting the suggestion from recombination density that the nearly equal recombination in the male and female mouse is associated with centromere position. Perhaps Robertsonian translocations as in *Mus poschiavinus* approach the higher female/male ratio of man.

Loci per megabase (PD) is significantly associated with the four variables that are important for other biological properties; of these, only physical length (CLG) is nonsignificant for any dependent variable. In addition to the four variables significant for physical density, the principal component of the six dependent variables (PC) shows an effect of chromosome structure, with an excess for acrocentrics and especially acrocentric centromeres. None of the metacentrics but all the acrocentrics except chromosome 21 have a region of high recombination and marker density close to the centromere.

In all these analyses, only the telomeric and heterochromatic bands are significant; the Giemsa and R-band differences as measured do not enter into the multiple regressions. There are several possible reasons for this nonsignificance. R-bands at high resolution are poorly characterized with respect to GC content and isochore composition (27); recombination may be increased at R/G junctions and therefore not assignable to either class (28); assignment of genes to single bands is currently inaccurate, especially for small bands; band composition is confounded with physical density and recombination and position, which are more accurately measured; and effects of band differences may be relatively small except at the telomeric and heterochromatic extremes. Rigorous discrimination among these possibilities must await more accurate assignment of loci and properties to high-resolution bands.

DISCUSSION

The location data base LDB+ was developed after the World Wide Web had demonstrated its superiority over alternative access to data bases. We have therefore avoided graphics, which present only a small fraction of the information in a form that precludes integration. Graphs of high-density maps on Internet have slow access and are almost as illegible as recent publications (3, 14). The LDB+ provides an integrated summary map for each chromosome with a hypertext link for each locus to deliver a report in catalog format. This report provides hypertext links to deeper structures, including partial maps, the largest lods, clones positive for that locus, and entries in other data bases. Successive versions of summary maps are also available on CD-ROM (16).

The formerly largest map had 10,850 loci ordered but not metrically integrated, with an additional 4236 loci in aligned linkage and radiation hybrid maps, neither ordered with respect to the STS-content map not metrically integrated (3). The larger number of loci in this report is merely a continuation of exponential growth that has outstripped the resources and imagination of the mapping community. The interest of these high-resolution maps is therefore not in the number of loci, but in development of methods for map construction and integration and their use to elucidate biological problems, including chromosome structure and function. There are many errors, especially of data omission, in such a pilot project. Nevertheless, it has demonstrated its validity and capacity to handle a huge number of loci. It is more informative than the costly data bases that merely align a small number of maps with no attempt at metrical integration, and compares favorably with consensus maps produced by many hands equipped only with artistic license.

1. McKusick, V. A. (1990) in *Genetic Maps*, ed. O'Brien, S. J. (Cold Spring Harbor Lab. Press, Plainview, NY), 5th Ed., pp. 5047–5133.
2. Bowcock, A. M., Bakker, B., Ceverha, P., Chipperfield, M. A., Minter-Morrison, A., Porter, C. J. & Pearson, P. L. (1994) *Report of the DNA Committee: Human Gene Mapping 1993* (Johns Hopkins Univ. Press, Baltimore), pp. 893–972.
3. Hudson, T. J., Stein, L. D., Gerety, S. S., Ma, J., Castle, A. B., *et al.* (1995) *Science* **270,** 1945–1954.
4. Anonymous (1995) *Human Genome News* **7,** 1–9.
5. Morton, N. E., Collins, A., Lawrence, S. & Shields, D. C. (1992) *Ann. Hum. Genet.* **56,** 223–232.
6. Morton, N. E., Collins, A. & Teague, J. (1995) *Genome Digest* **2,** 14.
7. Dausset, J., Cann, H., Cohen, D., Lathrop, M., Lalouel, J. M. & White, R. (1990) *Genomics* **8,** 575–577.
8. Keats, B. J. B., Sherman, S. L., Morton, N. E., Robson, E. B., Buetow, K. H., Cartwright, P. E., Chakravarti, A., Francke, U., Green, P. P. & Ott, J. (1991) *Ann. Hum. Genet.* **55,** 1–6.
9. Collins, A., Teague, J., Keats, B. J. & Morton, N. E. (1996) *Genomics* **36,** 157–162.
10. Rao, D. C., Morton, N. E., Lindsten, J., Hulten, M. C. & Yee, S. (1977) *Hum. Hered.* **27,** 99–104.
11. Morton, N. E. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 7474–7476.
12. Dib, C., Fauré, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J. & Weissenbach, J. (1996) *Nature (London)* **380,** 152–154.
13. Shields, D. C., Collins, A., Buetow, K. H. & Morton, N. E. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 6501–6505.
14. Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-Czarny, N., Spillet, D., Muselet, D., Prud-Homme, J.-F., Dib, C., Auffray, C., Morissette, J., Weissenbach, J. & Goodfellow, P. N. (1996) *Hum. Mol. Genet.* **5,** 339–346.
15. Teague, J., Collins, A. & Morton, N. E. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11814–11818.
16. Frezal, J. (1995) *Genome Digest* **2,** 10–11.
17. Collins, A., Keats, B. J., Dracopoli, N., Shields, D. C. & Morton, N. E. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4598–4602.
18. Francke, U. (1994) *Cytogenet. Cell Genet.* **65,** 206–219.
19. Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millaseau, P., Voysseix, G. & Lathrop, M. (1992) *Nature (London)* **359,** 794–801.
20. Morton, N. E., MacLean, C. J. & Lew, R. (1985) *Genet. Res.* **45,** 279–286.
21. Lander, E. S. & Krugylak, L. (1995) *Nat. Genet.* **11,** 241–247.
22. Morton, N. E. (1955) *Am. J. Hum. Genet.* **7,** 277–318.
23. Renwick, J. (1969) *Br. Med. Bull.* **25,** 65–73.
24. Elston, R. C. & Lange, K. (1975) *Ann. Hum. Genet* **38,** 341–350.
25. Haldane, J. B. S. & Smith, C. A. B. (1947) *Ann. Eugen.* **14,** 20–31.
26. Morton, N. E. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 3471–3476.
27. De Sario, A., Geigl, E.-M., Palmieri, G., D'Urso, M. & Bernardi, G. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 1298–1302.
28. Bernardi, G. (1995) *Annu. Rev. Genet.* **29,** 445–476.