Am. J. Hum. Genet. 54:516-525, 1994

Genetic Mapping of the BRCA1 Region on Chromosome 17q21

Hans Albertsen,* Rosemarie Plaetke,* Linda Ballard,* Esther Fujimoto,* Judith Connolly,* Elizabeth Lawrence,* Pilar Rodriguez,* Margaret Robertson,*,† Paige Bradley,* Bruce Milner,* David Fuhrman,* Andy Marks,* Robert Sargent,* Peter Cartwright,* Nori Matsunami,*,† and Ray White*,†

*Eccles Institute of Human Genetics and †Howard Hughes Medical Institute, University of Utah, Salt Lake City

Summary

Chromosome 17q21 harbors a gene (BRCA1) associated with a hereditary form of breast cancer. As a step toward identification of this gene itself we developed a number of simple-sequence-repeat (SSR) markers for chromosome 17 and constructed a high-resolution genetic map of a 40-cM region around 17q21. As part of this effort we captured genotypes from five of the markers by using an ABI sequencing instrument and stored them in a locally developed database, as a step toward automated genotyping. In addition, YACs that physically link some of the SSR markers were identified. The results provided by this study should facilitate physical mapping of the BRCA1 region and isolation of the BRCA1 gene.

Introduction

Breast cancer is an often fatal neoplastic disease of mammary tissue that causes ~50,000 deaths/year in the United States alone. Most cases are sporadic, with no apparent genetic lineage. However, a hereditary form of the disease, observed in approximately 5% of all cases and characterized by early age at onset, has been genetically linked to marker loci on chromosome 17q21 (Hall et al. 1990). A summary of 13 published reports located the cancer-predisposing locus, BRCA1, within the interval defined by D17S250 (mfd15) and D17S588 (42D6)(Easton et al. 1993). Several authors, however, had suggested a more narrow localization defined proximally by THRA1 and distally by D17S579 (mfd188) (Bowcock et al. 1993; Devilee et al. 1993; Simard et al. 1993; Smith et al. 1993). The genetic distance separating these two markers is ~5 cM (O'Connell et al. 1993). In the study reported here, we found supporting evidence for this estimate and identified 25 simple-sequence-repeat (SSR) markers, many of which fall within this region or flank it closely on either side. To facilitate efforts to identify the BRCA1 gene itself, we have integrated these new markers with previously known markers to form a highly resolved genetic map of this region of the long arm of chromosome 17.

The new SSR markers were developed as part of a comprehensive effort to generate large numbers of highly informative DNA markers with which a high-resolution genetic map of the entire human genome might be constructed. Such loci are easily typed by the PCR, using unique primers flanking each variable-repeat region (Saiki et al. 1988; Weber and May 1989; Orita et al. 1990). That markers of this type are abundant and highly informative has been shown by several studies describing di-, tri-, and tetranucleotide repeats (Litt and Luty 1989; Economou et al. 1990; Weissenbach et al. 1992; Melis et al. 1993). As part of an effort to automate SSR genotyping, five of our new markers in the BRCA1 region were fluorescently labeled, gel separated, and analyzed on an automated ABI373A DNA sequencer (Ziegle et al. 1992). Specialized software was developed to facilitate capture and storage of allelic information.

In addition to providing genetic information, the new SSR markers can serve as anchor points for a physical map of the BRCA1 region. By screening the CEPH library of YACs (Albertsen et al. 1990), we have identified several YACs that provide evidence for physical

Received July 27, 1993; accepted for publication October 25, 1993.

Address for correspondence and reprints: Dr. Hans Albertsen, Eccles Institute of Human Genetics, Building 533, Room 2100, Salt Lake City, UT 84112.

0002-9297/94/5403-0015$02.00

linkage among some of the SSR markers on the genetic map reported here.

Material and Methods

Preparation and Screening of Sau3A-digested Genomic M13 Library

Genomic DNA from lymphocytes of a human male was digested with *Sau*3AI, partially filled in with Klenow large fragment and dGTP/dATP, and size fractionated on a 1.2% agarose gel. DNA fragments of 400–900 bp were recovered from the gel and ligated with M13mp18 vector that had been digested with *Sal*I and partially filled in with dTTP/dCTP. The library was plated on Luria-Bertani-medium plates, transferred to nylon membranes and UV-cross-linked, and hybridized sequentially with end-labeled oligonucleotide probes $(dG\text{-}dT)_{10}$, $(dA_3\text{-}dT)_6$, $(dA_3\text{-}dG)_6$, and $(dA\text{-}dG\text{-}dA\text{-}dT)_6$. Hybridization was carried out for 3–4 h at 65°C in 10% polyethylene glycol, 7% SDS, and 1.5 × SSPE ($NaCl\text{-}NaH_2PO_4$–EDTA buffer). Membranes were washed in 6 × SSC and 0.1% SDS at 65°C.

Sequencing of Positive Clones

Positively hybridizing plaques were directly picked into 100-μl PCR cocktails containing 10 mM Tris-HCl pH 8.8, 40 mM NaCl, 1.5 mM $MgCl_2$, 5 pmol of each vector primer (A—TGT AAA ACG ACG GCC AGT CGC CAG GGT TTT CCC AGT CAC GAC; and B—CAG GAA ACA GCT ATG ACC AGC GGA TAA CAA TTT CAC ACA GGA); 2.5 μmol each of dNTPs, and 2 U of *Taq* DNA polymerase (Boehringer-Mannheim, Indianapolis). The reaction mixtures were heated at 94°C for 6 min, and the PCR was performed on a GeneAmp PCR System 9600 (Perkin-Elmer) for 25 cycles as follows: 95°C for 20 s (denaturation), 60°C for 20 s (annealing), and 72°C for 20 s (extension). The amplified inserts were purified with a Centricon-100 microconcentrater (Amicon, Danvers, MA), and the sequencing reactions were carried out on ABI Catalyst (Applied Biosystems, Foster City, CA) by the dideoxy chain-termination method using fluorescently labeled M13 sequencing primers: −21M13—TGT AAA ACG ACG GCC AGT; and M13RP1—CAG GAA ACA GCT ATG AC. The sequences were collected on an ABI383A sequencer (Applied Biosystems, Foster City, CA).

Primer Synthesis

Primers were synthesized in 40-nmol reactions by using an ABI394 DNA/RNA synthesizer (Applied Biosystems, Foster City, CA). After lyophilization, the primers were resuspended in 300 μl of TE_{-4} (10 mM Tris-Cl pH 7.8, 0.1 mM EDTA pH 8.0). Primer concentrations were determined by densitometry at 260 nm, and working stocks were prepared at 25 pmol/μl.

Primer Selection and Development

PCR primers were developed using a locally developed computer program, OLIGO (J.-M. Lalouel and T. Elsner, personal communication), on the basis of genomic sequence flanking the repeats. This program allows sequence comparisons between oligonucleotide primers and Alu consensus sequences, to detect homologies and to exclude primers with homologies above a user-defined threshold. PCR conditions were optimized with respect to $MgCl_2$ concentration (1.0, 1.25, 1.5, 2.0, 3.0, and 4.0 mM) and annealing temperatures. The size of each PCR product was determined from the sequence and verified, first through gel electrophoresis in 5% NuSieve 3:1 agarose (FMC BioProducts, Rockland, ME) and later through acrylamide gel electrophoresis. The primer sequences and their characteristics are listed in table 1 for each of the 25 new SSR markers on chromosome 17.

Radioactive Genotyping

The primer of each pair that showed the least homology to Alu was labeled with ^{32}P in a kinase reaction. PCR was carried out for a total of 30 cycles. Denaturation was done at 95°C for 6 min in the first cycle and for 10 s in each of the subsequent 29 cycles; annealing was done for 10 s at the temperature specific to each set of primers; and extension was done at 72°C for 20 s. The reactions were carried out in a 96-well Techne MW-2 thermal cycler (Techne, Cambridge). PCR products were mixed with a formamide dye solution (98% formamide, 0.1% xylene cyanol, 0.1% bromophenol blue, and 10 mM EDTA) and heated to 94°C for 4 min. A 3-μl aliquot of each sample was electrophoresed through a 7% denaturing polyacrylamide sequencing gel with 32% of deionized formamide, 5.6 M urea, and 1 × TBE (Tris–boric acid–EDTA buffer). After electrophoresis, the gels were exposed to X-ray film without drying, at −70°C for 12–24 h, with intensifying screens.

Fluorescent Genotyping

As with radioactive genotyping, the primer showing the least homology to Alu was labeled, albeit with a fluorochrome, for detection on an ABI373A sequencing instrument. Of the three different dyes, the blue dye

Table 1

Characteristics of the 25 New SSR Markers

Name (locus)	Primers	Product Size (bp)	No. of Alleles	Heterozygosity (%)	Mg^{++} (mM)/ Annealing Temperature (°C)
UT7 (D17S752)	GGTGACAGAGTGAGATCCTGTC CCCATGTTTGAGAGATCAAACCTT[a]	180	4	77	1.0/60
UT8 (D17S761)	GGGGTAGGGCGGGAATAGAGG[a] TGATAAAAGCAGTGAGACATTGGTA	152	4	50	1.0/60
UT26 (D17S700)	GATGACAGAGTGAGACCTTGTCC[a] GTGGGCAGACAGAGCCAAACCA	80	3	44	1.0/60
UT50 (D17S733)	TGGGCAACAGAGCAAAACTCTGTT[a] GACAGAGGGGAGGGGAGACGG	201	6	89	1.5/60
UT52 (D17S735)	TGGAGGATGCAGTGAGCTGAAATG GACCTCCCCAAAGTACTGAGTTAC[a]	238	7	62	1.5/60
UT62 (D17S746)	TCTATGGGGGCCTTTTCCCAGAG[a] CAAGTGTATATCCCTGAGGTGTG	108	3	44	2.0/58
UT67 (D17S750)	AGCGAAGAGCACTCTGACTAGAATC[a] ATGGGGTAAGGTCTACAGAATGCCT	175	5	66	1.5/65
UT71 (D17S754)	TGGATTCACTGACTCAGCCTGC[a] GCGTGTCTGTCTCCATGTGTGC	145	3	25	1.0/55
UT77 (D17S758)	CCCCGCCCCCATACAACAGGCT[a] CCTGGCCATGTTTTTTTTTAATGTC	76	4	44	1.0/55
UT159 (D17S620)	GGGAAGGTGTCTGAAACCCAAGG CCCACCACTACCTATTGTTCTATAC[a]	143	3	51	1.0/60
UT165 (D17S627)	ACGAACAGAGCAAGACTCTGTCTC TCCAACAAACCAGAAAGTCCACTAT[a]	145	3	44	1.0/60
UT185 (D17S648)	TGGGCGACAAGGCAAGACTTCGT CCTGAGTGAAGGTGGCTTAAAAAAG[a]	351	2	29	2.0/60
UT189 (D17S652)	TTTCTTATCACCTTCCCTCCCACTG[a] CTGCAATCTATCAGTGTCCAAGATG	159	5	63	1.5/65
UT205 (D17S656)	CTCCAGCCTGGGGGACAAC TGTGGGGAACCACTTTTGTTCGTA[a]	138	5	69	1.25/60
UT217 (D17S669)	CAGCCTGGGTGACGGAGTG TGTCATTTAACCTCTTTGAGCTCTA[a]	235	6	85	1.5/60
UT224 (D17S677)	CTCCATCATAGGTGACAAATTG TTTGAGCTTTGCCACTTGCCAGC[a]	275	6	86	1.5/60
UT267 (D17S693)	GACATAGCGAGACTCCCTGTCCA GGCACATAGTAGGAGTGCTACATG[a]	100	3	42	1.5/65
UT287 (D17S708)	TGGGTGACAGAGTGAGACGCTG CCTCCAGGCCCTTCTTCCTGG[a]	116	2	50	1.25/60
UT290 (D17S966)	GGGTGACAGAGTGAGACTCCATC GCTCTGTGTCAGGGATGAGTTCA[a]	319	8	81	2.0/60
UT385 (D17S967)	AACAAGAGCGAAACTCCGTCTCAA TCCTCCTTGTTCAAAGTCTGCATG[a]	93	4	75	1.5/60
UT394 (D17S702)	AGCAACACATATCAGGGGC TGTAGGTTGACCTTAAGGC[a]	350	8	88	1.5/52
UT401 (D17S964)	GTTCTTTCCTCTTGTGGGG[a] AGTCAGCTGAGATTGTGCC	224	2	44	1.25/54
UT573 (D17S902)	GAGGTTGCAGTGAGTTGAGA GGAACATCCTCCTTCACTCTT[a]	155	6	78	1.5/62
UT752 (D17S907)	ACTCCAACGTGGGCGACAG[a] CCTTCGTTTTATGTCCCGAG	338	7	91	1.0/60
UT956 (D17S965)	GAAGTGCCAAAGGCCAGGAACC CTGCACTTCATTCAGCCTGGGTG[a]	185	5	75	1.5/65

[a] This primer, of the pair, should be labeled.

(FAM) can be incorporated directly onto the 5′ end of an oligonucleotide on the DNA synthesizer; labeling with the yellow and green dyes (TAMRA and JOE) requires synthesis of oligonucleotides with aminolink 2 (400808, ABI) attached 5′, and subsequent addition of the appropriate dye-NHS ester. Unreacted dye is removed by means of a PD-10 column (17-0851-01, Pharmacia) (for technical details, see Genescan 672 Software User's Manual, Appendix D, ABI). Unlabeled oligonucleotides are removed on a purification cartridge, OPC (400771, ABI; User Bulletin 51, ABI).

To establish working conditions for each primer pair, PCR was performed on a few samples of DNA from the CEPH reference panel in a GeneAmp PCR System 9600. Reaction volumes of 100 μl contained 400 ng of template, 200 μM each dNTP, 0.5 μM each primer, 2.7 U of *Taq* DNA polymerase, 0.24 mM spermidine, 10 mM Tris-HCl pH 8.7, and appropriate concentrations (1.0–4.0 mM) of $MgCl_2$. Annealing temperatures were estimated on the basis of primer sequence and were adjusted where necessary. The fluorescent PCR products were also tested for optimal signal on the sequencing instrument, by analyzing aliquots taken at 14, 17, 20, and 23 cycles to determine the number of cycles necessary to observe a specific product with minimal formation of spurious product. Following the establishment of working conditions, the PCR reactions were proportionally scaled down to 25 μl and were used for genotyping individuals in the CEPH panel.

Linkage Analysis

All linkage analyses described in this paper were performed using four programs from the LINKAGE package (version 5.1): CFACTOR, CLODSCORE, CILINK, and CMAP (Lathrop et al. 1984).

Linkage Data

Linkage data used in this study were derived partly from O'Connell et al. (1993), partly from the CEPH database (version 6), and partly from genotypic analysis of new SSR markers that were labeled with ^{32}P or with a fluorochrome.

Identification of YAC Clones

To initiate the development of physical representation from the BRCA1 region, the CEPH library of YACs was screened by PCR according to a protocol described by Green and Olson (1990). Some of the YACs isolated in this way identified close physical linkage for several of the SSR markers described here.

Results and Discussion

To supplement the existing archive of SSR markers, we have developed >2,000 genomic-sequence-tagged markers based predominantly on tetranucleotide repeats, as this type of repeat is, in general, highly informative and tends to show less susceptibility to PCR artifacts such as laddering than the dinucleotide $(CA)_n$ repeats (Litt and Luty 1989; Tautz 1989). To augment the marker density on chromosome 17 specifically, a flow-sorted cosmid library (a gift from Dr. L. Deaven, Los Alamos National Laboratory) was subcloned into the M13 sequencing vector and screened for the presence of selected di- and tetranucleotide repeats; appropriate SSR loci were sequenced for development of primers. Approximately 80 new SSR markers for chromosome 17 were obtained in this way (data not shown). We have now mapped 25 of those markers, by genetic linkage analysis, to a 40-cM region surrounding the BRCA1 locus.

Localization of New SSR Markers to the BRCA1 Region

Strategies to reduce the effort involved in linkage analyses, by reducing the number of genotypes required for map construction, are being implemented in our laboratory. One of these strategies begins by genotyping each new marker in only four CEPH pedigrees (884, 1331, 1332, and 1362) and comparing two-point lod scores obtained with selected loci, to determine an approximate chromosomal location; the marker is then genotyped on a panel of CEPH individuals with known meiotic breakpoints in that region. An initial rough localization of each new SSR marker to the BRCA1 region was based on information from the four selected CEPH pedigrees. The power of a two-point lod score analysis in these pedigrees, when testing SSR markers against the Genethon markers (Weissenbach et al. 1992), depends on the degree of informativeness of both markers in the comparison, as well as on their genetic distance. For two completely informative markers <5 cM apart, the lod scores can approach 20. However, in light of the fact that markers in the comparisons rarely reveal complete genetic information, most SSR markers are assigned to particular chromosomal regions with lod scores of 6–15 and recombination fractions of .05–.20.

Selection of Markers for the BRCA1 Map

The initial stage of the genetic analysis of markers from the BRCA1 region was in part based on results published elsewhere (O'Connell et al. 1993), which had

Table 2

Previously Characterized Markers Used in the Map Construction

Locus (probe)	Polymorphism[a]	Reference
D17S36 (CRI-L946)	HAP	Green et al. 1987
D17S37 (CRI-L581)	*Taq*I	Green et al. 1987
D17S41 (LEW102)	*Taq*I	Nakamura et al. 1988
D17S74 (pCMM86)	*Hin*fI	Nakamura et al. 1988
D17S80 (YNZ94)	*Pst*I	O'Connell et al. 1993[b]
D17S83 (2C11.1)	HAP	O'Connell et al. 1993[b]
D17S85 (p2F9.8)	HAP	O'Connell et al. 1993[b]
D17S183 (SCG43)	PCR	Black et al. 1993
D17S250 (mfd15)	PCR	Weber and May 1989
D17S507 (cLS17.13)	HAP	O'Connell et al. 1993[b]
D17S508 (cLB17.20)	HAP	O'Connell et al. 1993[b]
D17S509 (fLB17.1)	*Hae*III	O'Connell et al. 1993[b]
D17S511 (fLB17.18)	*Taq*I	O'Connell et al. 1993[b]
D17S579 (mfd188)	PCR	Weber and May 1989
D17S791 (AFM155xd12)	PCR	Weissenbach et al. 1992
D17S800 (AFM200zf4)	PCR	Weissenbach et al. 1992
D17S810 (AFM248yg1)	PCR	Weissenbach et al. 1992
NF1 (four markers from NF1)	HAP	O'Connell et al. 1993[b]
Hox2B (Hox2B)	*Sac*I	Bentley et al. 1989
NME1 (NME1)	*Bgl*II	Varesco et al. 1992
PPY (PPY)	*Msp*I	Takeuchi et al. 1985
THRA1 (THRA1)	*Msp*I	Sheer et al. 1985

[a] HAP = haplotype.
[b] Reference provides verified and haplotyped genetic data rather than the original descriptions of these loci.

established the relative order and genetic distances between loci, NF1(haplotype)–D17S250(mfd15)–D17S509-(fLB17.1)–D17S508(LS17.13)–D17S74(pCMM86). Using these five markers in two-point comparisons with all chromosome 17 markers from a local database and from the CEPH database (version 6) allowed us to identify 42 additional markers that showed linkage to the BRCA1 region. This group of RFLP and SSR loci generated a set of 47 informative markers (tables 1 and 2).

Generation of New Genetic Data

Since the initial localization of the SSR markers to the BRCA1 region was based on genetic information from only four CEPH pedigrees, it was necessary to widen the data set to increase precision. For markers UT67, UT573, UT752, and D17S183 this was done by first genotyping all 120 CEPH parents to identify completely informative pedigrees; all members of the pedigrees so indicated were then genotyped. For the five markers genotyped on the ABI instrument (UT7, UT217, UT224, UT394, and D17S250; see below) most CEPH pedigrees were genotyped. The remaining markers were typed on arbitrarily selected CEPH pedigrees or on small panels of individuals selected for the presence of meiotic breakpoints. The radioactive signatures of 13 of these markers are shown in figure 1.

Capturing Genotypes by Using an ABI373A Sequencing Instrument

As a test of an alternative method of genotyping, five selected SSR markers were visualized and captured with the ABI373A Laser Scanning Unit developed by Applied Biosystems. To capture genotypes on this instrument, a fluorescent tag is attached to the 5′ end of one of the two primers that define an SSR locus. The tagged primer is then used in a standard PCR. The tag can be blue (FAM), green (JOE), or yellow (TAMRA); the fourth color, red (ROX), is reserved for the internal standard. After the PCR product is mixed with the internal standard and loaded on the instrument, the alleles are captured as fluorescent peaks (fig. 2). The internal standard precisely determines the size, in base pairs, of a given genotype peak, permitting comparison and scoring of alleles across gels. The potential advantages to the fluorescent/electronic system compared with the radioactive system of capturing genotypes are

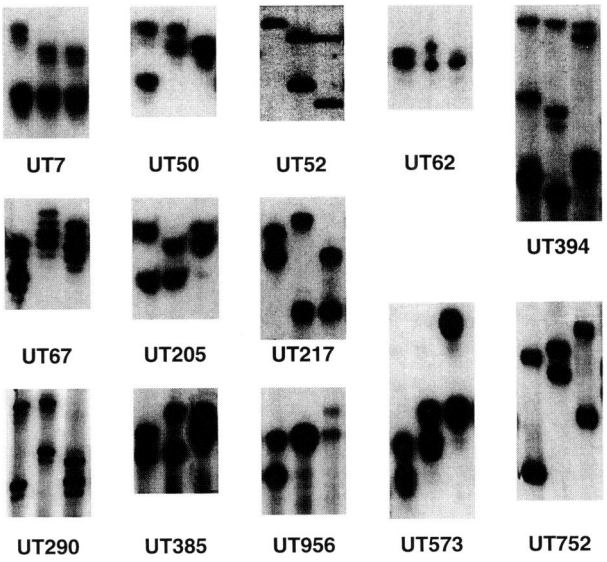

Figure 1 Radioactive signatures of 13 of the new SSR markers, selected in part for their high degree of heterozygosity. Most of these markers amplify tetranucleotide repeats; however, UT67 represents a typical dinucleotide repeat with the laddering effect frequently associated with this type of repeat. UT394 shows a complex PCR pattern that most likely originated from amplification of two separate loci. However, as segregational analysis of UT394 in the CEPH reference panel did not disclose any recombination between the two loci, a very close genetic relationship between the two systems was indicated. Further evidence for physical linkage between the two systems was obtained when two distinct PCR products were visualized when YAC 151E11 was used as template (data not shown).

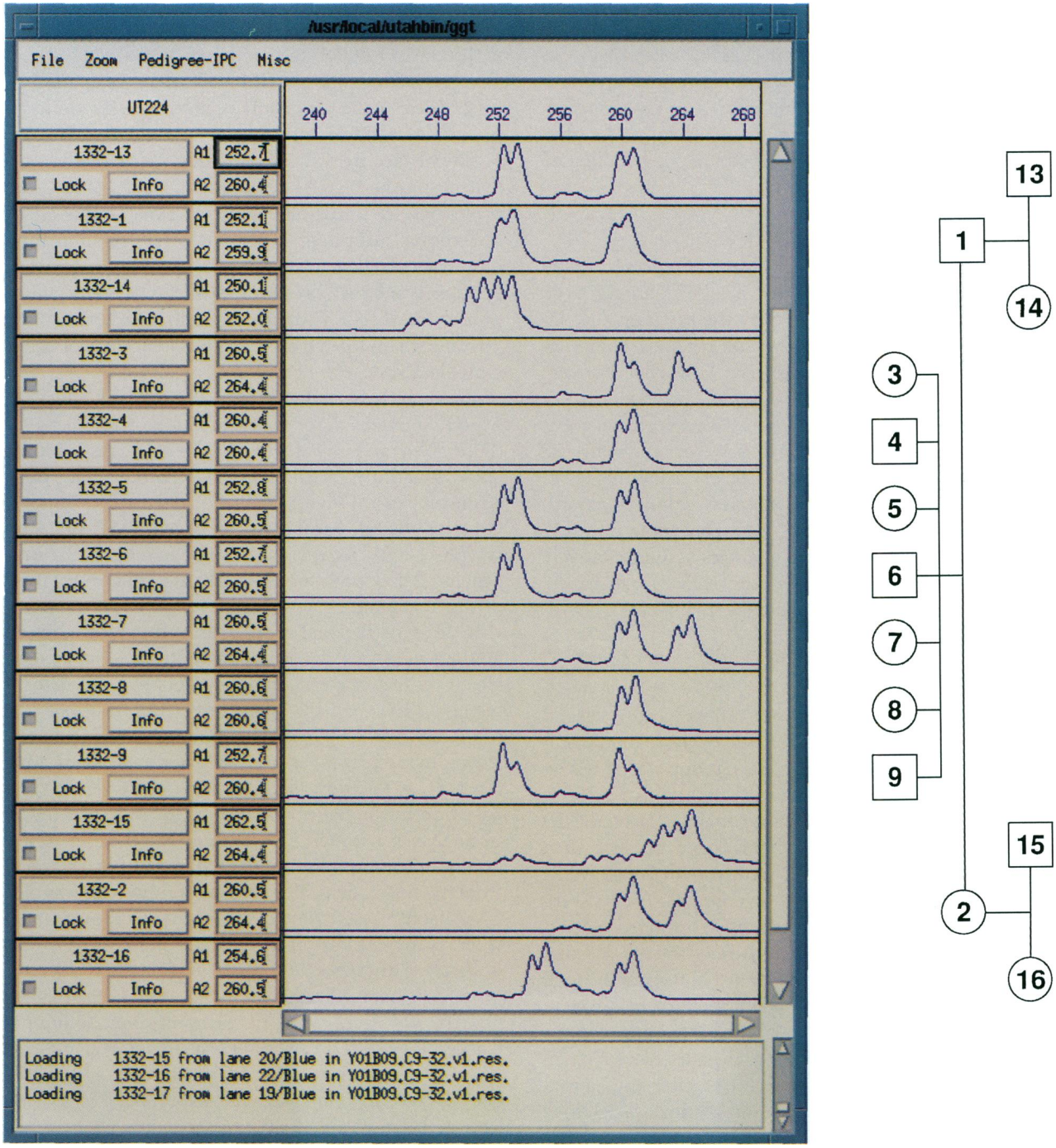

Figure 2 Data traces from fluorescent allelotyping of UT224 labeled with the FAM fluorochrome, as an alternative to traditional visualization of SSR markers with ^{32}P. Each blue trace represents the PCR-amplification product of a single individual who is identified by pedigree number and individual number (e.g., 1332-13) to the left of each trace. Also indicated at left are allele sizes (in base pairs) observed for each individual. The pedigree structure shown to the right is added here to aid interpretation of the traces. The split peaks 1 bp apart, seen in each allele, could be interpreted as an effect of the *Taq* DNA polymerase, which is known to unspecifically add a single A nucleotide to the 3′ terminus of an extension product (Clark 1988).

several: (1) fluorescently labeled primers have a significantly longer shelf life, usually >1 year; (2) a high degree of multiplexing with different dyes and allele sizes can be achieved; (3) electronic recording of genotypes reduces transcription errors and allows for rapid and accurate scoring; and (4) archiving, retrieving, and verifying genotypes does not rely on storage or handling of autoradiograms.

Selection of the Strongest Markers

The map was constructed by means of tools from the LINKAGE package. As a first step, all 47 markers in the BRCA1 region were analyzed by the program CLODSCORE to determine all possible pairwise combinations of two-point lod scores. To construct a strong map (i.e., with odds >1,000:1 against inversion of any pair of loci), 20 of the 47 markers were selected according to the sole criterion of having one or more pairwise lod scores >30. The rationale for choosing this criterion for selection, rather than heterozygosity or number of informative offspring, is that mapping information that is shared by two markers is quantitatively more significant than the extent to which a single marker has been typed and is informative in the data set.

Generation of a Trial Map and Verification against Inversion

The 20 selected markers were divided into several groups of six, each containing one or more markers in common with another group. Each group was tested by CILINK to establish a linear order with odds >1,000:1 against alternative orders within that set; 3 of the 20 markers were eliminated at this stage because they could not be ordered consistently with those odds. Where the six-member maps overlapped, orders were retested in new six-marker groups so arranged that information could be combined. Approximately 15 such runs were sufficient to establish a trial map of 17 loci.

To verify the trial map, likelihoods for order were calculated with CILINK for all possible permutations of the central three markers in each series of seven loci, the first series beginning with NF1. After the central triplet of a set was tested in the context of its four nearest flanking markers, the next group of seven was constituted by dropping the most centromeric marker of the prior set and adding the marker next in line; the central triplet of the new series was tested, and so on down the entire map, until every marker had been evaluated for possible local inversions. (Note, however, that in the case of markers at the ends of the map, a tested triplet could not be "central"; three or four of the other markers would lie on one side and one or none on the other.) In this way, all 17 markers were linearly ordered with odds >1,000:1 against inversions.

Subsequent analyses showed that one marker, UT224, from the group of weaker markers could be placed in a single interval on the map, with odds >1,000:1 (see below). The "strong map" of these 18 markers in the BRCA1 region, with odds against local inversions, is shown in figure 3. Since inclusion of breakpoint information will bias the recombination frequencies and thereby the genetic distance between markers typed on a breakpoint panel, it is important to note here that UT205 is the only marker, on the "strong" map, for which a significant amount of genotypic data had been obtained as breakpoint data.

Two markers, mfd188 (D17S579) and HOX2B, which showed zero recombination with fLB17.1 (D17S509) and cLS17.13 (D17S507), respectively, could substitute for the latter two markers with odds against inversion slightly reduced, albeit still >1,000:1. On this basis, haplotypes between D17S509 and D17S579 and between D17S507 and HOX2B were generated and were subsequently used in construction of the final map.

In light of the linear order and distances between markers in the map of strong markers, it became possible to localize the remaining markers to particular intervals in the map by using the CMAP routine. The result of this analysis usually defined a region in the strong map, encompassing two or three neighboring intervals, as the most likely location of a given marker. Based on this initial observation, a location was tested by analyzing with CILINK all specific positions of the test locus among the seven to nine surrounding markers from the strong map. The reason for doing this was that CMAP only tests discrete positions in a given map, whereas the CILINK analysis is not limited to a subset of discrete distances and therefore is likely to identify genetic distances with higher likelihoods. Using this approach, we were able to position one additional marker, UT224 (D17S677), in a single interval and add it to the strong map. Other markers that could not be placed in a single interval were each tested in at least five of the most likely consecutive intervals by using CILINK to identify the set of intervals to which the marker could be localized with odds >1,000:1. These positions are shown as confidence intervals to the right of the map in figure 3.

Errors in the Genetic Data and Map Inconsistencies

The genetic data developed in this study, as well as data included from a previous study, were extensively

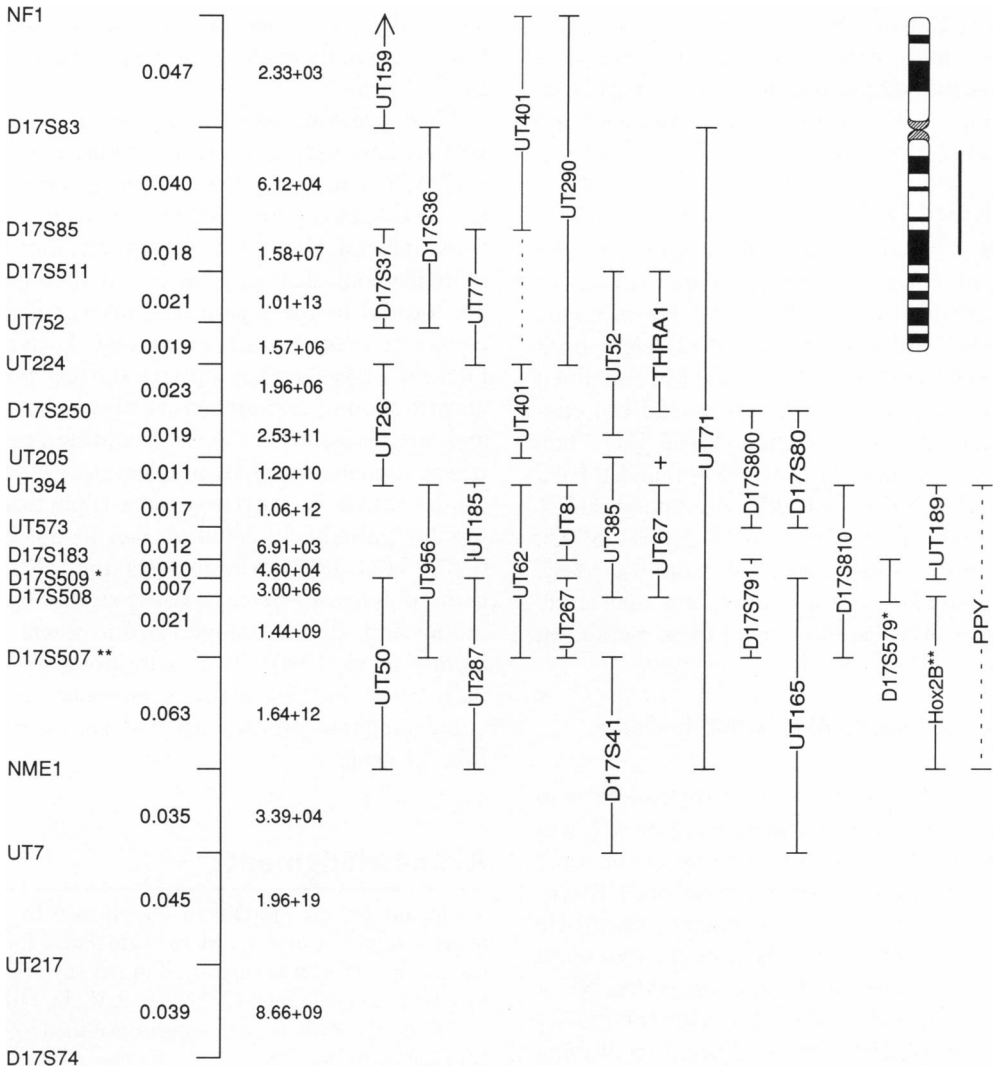

Figure 3 *Upper right,* Karyogram of chromosome 17, with a vertical line to indicate the approximate coverage of the map developed in this study. *Left,* All markers that could be linearly ordered with odds >1,000:1, starting, at the top, with the most centromeric marker, NF1. D17S74 marks the map boundary on the telomeric side. Recombination frequencies between neighboring markers are indicated on the left side of the map, and the odds against inversion in each interval are on the right. The markers that could not be placed in a single interval on the map are placed within confidence intervals of 1,000:1 odds and are illustrated by vertical lines. The two pairs of loci that were haplotyped (see text) are indicated by single and double asterisks (* and **). Marker UT159 is drawn with an arrow pointing up to indicate that the confidence interval extends beyond NF1. Marker UT401 is drawn in two intervals connected by a dashed line to indicate a noncontinuous confidence interval; however, the location of UT401 around D17S250 is favored with >40:1 odds over the other location. The location of THRA1, shown by its confidence interval, is based on the genetic information; however, since this location is in disagreement with observations made with YACs and pulsed-field gel electrophoresis, its physical location is indicated by a plus sign (+). PPY could not be placed into an unambiguous confidence interval with 1,000:1 odds and is therefore shown on the map by a dashed line to indicate its most likely location.

analyzed for errors and, where possible, were tested by secondary typings (see O'Connell et al. 1993). Inconsistency between the genetic and the physical map locations of THRA1 suggested that the genetic data were probably incorrect, in that this gene had been physically linked to UT205 on YAC 44F8 (280 kb) (and on other YACs; data not shown); yet the genetic analysis excluded UT205 from the 1,000:1-odds interval of THRA1. The physical location of THRA1 in the immediate vicinity of the UT205 locus is supported by observations using pulsed-field gel electrophoresis (Lemons et al. 1990). We attempted to place another marker,

PPY, within a confidence interval on the map, but this locus showed widely varying confidence boundaries when tested against different groups of markers from the strong map. Both of these inconsistencies are shown graphically in figure 3.

SSR Markers Linked by YACs

Several YACs identified by PCR screening of the CEPH library were found to contain more than one of the SSR markers described in this study. For example, UT67 (D17S750) and UT189 (D17S652) were both present on 416F6 (290 kb); 151E11 (600 kb) contained UT394 (D17S702) and D17S800; 44F8 (280 kb) contained THRA1, UT205 (D17S656), and UT26 (D17S700); 233F6 (310 kb) contained UT62 (D17S746) and UT385 (D17S967); and 500D9 (360 kb) contained UT71 (D17S754), UT205 (D17S656), and UT26 (D17S700). The physical observations obtained from the YAC screening suggest that certain markers are sufficiently close together to allow haplotyping of these markers in future studies.

Comparison of the Genetic Map with a Radiation Hybrid Map

In parallel with the construction of the genetic map of the BRCA1 region, a complementary approach was taken to construct a map by using many of the same SSR markers, ordering them on the basis of PCR typings in radiation hybrids (see accompanying article [O'Connell et al. 1994]). Comparison of the two maps shows agreement for the following linear order: NF1–UT71–UT394–UT573–D17S183–D17S509/D17S579–D17S507/Hox2B–NME1. In addition to having markers in common, the two maps complement each other and augment map information by contributing additional markers to the composite map.

Conclusion

Twenty-five newly developed SSR markers have been genotyped and integrated with 22 previously mapped markers from chromosome 17q11.2-21.33, to form a highly resolved genetic map. Most of the markers were genotyped by traditional means, but technological developments allowed us to automate sample handling and to adapt five of the markers for fluorescent genotyping. Automated procedures promise not only higher throughput but also minimal risk of human errors. The entire map, ranging from NF1 on the centromeric side to D17S74 on the telomeric side, spans ~46 cM. However, most of the markers are grouped in the central 25 cM of the map, where they provide a high concentration of reagents in the region believed to contain the BRCA1 gene.

Most previous reports on the fine localization of BRCA1 have indicated a region flanked by THRA1 and D17S579. The map we present here places at least 6 of the new markers, and possibly as many as 13, in this 5-cM interval. Under the assumptions that 1 cM represents the equivalent of 1 Mb and that 10 markers truly are located in the region, the marker density would equal one every 500 kb, on average. Such a density of markers makes an appropriate starting point for the construction of comprehensive physical maps in YACs and other large-insert clones. In addition to their qualities as reagents for physical mapping, the SSR markers (or derivatives thereof) may serve as physical probes to test for pulsed-field gel anomalies in BRCA1 patients (Joslyn et al. 1991). Alternatively, the SSR markers can be used in loss-of-heterozygosity studies by comparing blood and tumor samples from single individuals (Cropp et al. 1990). The many possible applications mean that the SSR markers presented in this study could facilitate progress toward the isolation of the BRCA1 gene.

Acknowledgments

The authors are grateful to Ed Meenen for synthesizing oligonucleotide primers and to Ruth Foltz for editing the manuscript. H.A. was supported in part by Danish Medical Research Council grant 12-9631-2, R.W. by NIH grant R01 HG00367-04. R.W. is an Investigator of the Howard Hughes Medical Institute.

References

Albertsen HM, Abderrahim H, Cann HM, Dausset J, le Paslier D, Cohen D (1990) Construction and characterization of a human yeast artificial chromosome library containing seven haploid genome equivalents. Proc Natl Acad Sci USA 87:4256–4260

Bentley KL, Ferguson-Smith AC, Miki T, Kidd KK, Ruddle F (1989) Physical linkage of Hox 2.1 and nerve growth factor receptor. Cytogenet Cell Genet 51:961

Black DM, Nicolai H, Borrow J, Solomon E (1993) A somatic cell hybrid map of the long arm of human chromosome 17, containing the familial breast cancer locus (BRCA1). Am J Hum Genet 52:702–710

Bowcock AM, Anderson LA, Friedman LS, Black DM, Osborne-Lawrence S, Rowell SE, Hall JM, et al (1993) THRA1 and D17S183 flank an interval of <4 cM for the

breast-ovarian cancer gene (BRCA1) on chromosome 17q21. Am J Hum Genet 52:718–722

Clark JM (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res 16:9677–9686

Cropp C, Liderau R, Campbell G, Champene HM-H, Callahan R (1990) Loss of hererozygosity on chromosome 17 and 18 in breast carcinoma: two new regions identified. Proc Natl Acad Sci USA 87:7737–7741

Devilee P, Cornelis RS, Bootsma A, Bardoel A, van Vliet M, van Leeuwen I, Cleton FJ, et al (1993) Linkage to markers for the chromosome region 17q12-q21 in 13 Dutch breast cancer kindreds. Am J Hum Genet 52:730–735

Easton DF, Bishop DT, Ford D, Crockford GP, The Breast Cancer Linkage Consortium (1993) Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. Am J Hum Genet 52:678–701

Economou EP, Bergen AW, Warren AC, Antonarakis SE (1990) The polyadenylate tract of Alu repetitive elements is polymorphic in the human genome. Proc Natl Acad Sci USA 87:2951–2954

Green ED, Olson MV (1990) Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. Proc Natl Acad Sci USA 87:1213–1217

Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith T, Bowden DW, et al (1987) Towards a primary genetic linkage map of the human genome: a progress report on linkage studies with more than 300 new RFLPs. Cytogenet Cell Genet 46:623

Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King M-C (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250:1684–1689

Joslyn G, Carlson M, Thliveris A, Albertsen H, Gelbert L, Samowitz W, Groden J, et al (1991) Identification of deletion mutations and three new genes at the familial polyposis locus. Cell 66:601–613

Lathrop GM, Lalouel J-M, Ott J (1984) Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA 81:3443–3446

Lemons RS, Eilender D, Waldman RA, Rebentisch M, Frej A-K, Ledbetter DH, Willman C, et al (1990) Cloning and characterization of the t(15;17) translocation breakpoint region in acute promyelocytic leukemia. Genes Chromosom Cancer 2:79–87

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397–401

Melis R, Bradley P, Elsner T, Robertson M, Lawrence E, Gerken S, Albertsen H, et al (1993) Polymorphic SSR (simple-sequence-repeat) markers for chromosome 20. Genomics 16:56–62

Nakamura Y, Lathrop M, O'Connell P, Leppert M, Barker D, Wright E, Skolnick M, et al (1988) A mapped set of DNA markers for human chromosome 17. Genomics 2:302–309

O'Connell P, Albertsen H, Matsunami N, Taylor T, Hundley JE, Johnson-Pais TL, Reus B, et al (1994) A radiation hybrid map of the BRCA1 region. Am J Hum Genet 54:526–534 (in this issue)

O'Connell P, Plaetke R, Matsunami N, Odelberg S, Jorde L, Chance P, Leppert M, et al (1993) An extended genetic linkage map and an "index" map for human chromosome 17. Genomics 15:38–47

Orita M, Suzuki Y, Sekiya T, Hayashi K (1990) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5:874–879

Saiki RF, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Erlich HA (1988) Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Sheer D, Sheppard DM, le Beau M, Rowley JD, San Roman C, Solomon E (1985) Localization of the oncogene c-erbA1 immediately proximal to the acute promyelocytic leukemia breakpoint on chromosome 17. Ann Hum Genet 49:167–171

Simard J, Feunteun J, Lenoir G, Tonin P, Norman T, The VL, Vivier A, et al (1993) Genetic mapping of the breast-ovarian cancer syndrome to a small interval on chromosome 17q12-21: exclusion of candidate genes EDH17B2 and RARA. Hum Mol Genet 2:1193–1199

Smith SA, Easton DF, Ford D, Peto J, Anderson K, Averill D, Stratton M, et al (1993) Genetic heterogeneity and localization of a familial breast-ovarian cancer gene on chromosome 17q12-q21. Am J Hum Genet 52:767–776

Takeuchi T, Gumucio D, Eddy R, Meisler M, Minth C, Dixon J, Yamada T, et al (1985) Assignment of the related pancreatic polypeptide (PPY) and neuropeptide Y (NPY) genes to regions on human chromosomes 17 and 7. Cytogenet Cell Genet 40:759

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17:6463–6471

Varesco L, Caligo MA, Simi P, Black DM, Nardini V, Casario L, Rocchi M, et al (1992) The NM23 gene maps to human chromosome band 17q22 and shows a restriction fragment length polymorphism with BglII. Genes Chromosom Cancer 4:84–88

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, et al (1992) A second-generation linkage map of the human genome. Nature 359:794–801

Ziegle JS, Su Y, Corcoran KP, Nie L, Mayrand E, Hoff LB, McBride LJ, et al (1992) Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics 14:1026–1031