The Human Y Chromosome: A 43-Interval Map Based on Naturally Occurring Deletions

Douglas Vollrath, Simon Foote, Adrienne Hilton, Laura G. Brown, Peggy Beer-Romero, Jonathan S. Bogan, David C. Page

A deletion map of the human Y chromosome was constructed by testing 96 individuals with partial Y chromosomes for the presence or absence of many DNA loci. The individuals studied included XX males, XY females, and persons in whom chromosome banding had revealed translocated, deleted, isodicentric, or ring Y chromosomes. Most of the 132 Y chromosomal loci mapped were sequence-tagged sites, detected by means of the polymerase chain reaction. These studies resolved the euchromatic region (short arm, centromere, and proximal long arm) of the Y chromosome into 43 ordered intervals, all defined by naturally occurring chromosomal breakpoints and averaging less than 800 kilobases in length. This deletion map should be useful in identifying Y chromosomal genes, in exploring the origin of chromosomal disorders, and in tracing the evolution of the Y chromosome.

Among human chromosomes, the Y is unusual in that most of the chromosome does not participate in meiotic recombination. This precludes construction of a genetic linkage map for most of the chromosome. Instead, identification of Y-linked genes has depended on physical mapping based on naturally occurring deletions. Such deletion mapping of the Y chromosome is practical because individuals with deletions of portions of the chromosome are viable and occur at a reasonable frequency in the human population.

Initial attempts to map Y-chromosomal genes by deletion analysis were based on correlations of cytologically detectable Y anomalies with abnormal phenotypes. For example, cytogenetic studies of six sterile men with long arm (Yq) deletions led to the hypothesis that a gene associated with spermatogenesis was located on proximal Yq (*1*). Similar attempts were made to define the portion of the chromosome related to the determination of gonadal sex (*2*). These early cytogenetic efforts suffered from the limited resolution and accuracy of chromosome banding patterns visualized by light microscopy.

Later, Y-chromosomal deletion maps were constructed by hybridizing Y-specific probes to immobilized genomic DNA's (*3*, *4*). Because the Y is a haploid chromosome, the ability to determine precisely and accurately the extent of Y-chromosomal DNA in individuals with informative phenotypes has been limited only by the number of probes used. If we assume that each deleted chromosome has suffered a single break, with loss of all sequences on one side and retention of all sequences on the other, DNA loci can be ordered according to their presence or absence in the genome of a given individual. Analysis of a collection of such individuals, each harboring a different deletion, can yield a self-consistent map comprising a series of ordered intervals. The boundaries of the intervals are defined by Y-chromosomal breakpoints in the individuals used to construct the map. Correlation of the phenotypes of these individuals with their content of Y-chromosomal DNA can localize genes. Such deletion mapping has resulted in (i) identification of the sex determining gene (*5*) *SRY*, as well as a candidate gene for Turner syndrome (*6*), *RPS4Y*, and (ii) localization of other genes, including one responsible for the expression of the minor histocompatibility antigen, H-Y (*7*, *8*).

Creation of a deletion map results in the ordering of DNA loci along the chromosome. The ordered loci constitute a physical map with other virtues, apart from its utility in locating genes. A physical map can be used to compare the structure of the Y chromosome to that of the X chromosome and to study structural diversity of the Y within the human population and among primates, thereby gaining information on both the evolution of chromosome structure and the evolution of the human species through paternal lineages. A physical map can also be used to elucidate the mechanisms by which abnormal Y chromosomes are generated.

We set out to produce a deletion map of the Y chromosome based largely on detection of sequence-tagged sites (STS's). An STS is a short stretch of genomic sequence that can be detected by the polymerase chain reaction (PCR) (*9*) and mapped to a particular point in the genome, where the STS then serves as a landmark (*10*). The speed, sensitivity, and flexibility of PCR make it the method of choice for both construction and application of such a deletion map. In addition, STS's and their corresponding PCR assays can be readily disseminated through electronic databases. This provides a common pool of loci for map construction and allows comparison of maps made by different investigators. Moreover, the ability to order Y-chromosomal STS's suggested a strategy for constructing an overlapping set of yeast artificial chromosome (YAC) clones encompassing the euchromatic Y chromosome, as described by Foote *et al.* (*11*). The same PCR assays used to construct the deletion map also provided a facile means of identifying Y chromosomal YACs within a total genomic library (*12*). A correspondence between ordered STS's and YAC clones was immediately created, simplifying the problem of ordering YAC's based on their STS content.

Y-chromosomal STS's. To generate a collection of STS's providing a near-random sampling of the chromosome, sequence was obtained from several hundred previously uncharacterized Y-chromosomal DNA fragments. The fragments were derived from two recombinant lambda phage libraries constructed with Y chromosomes

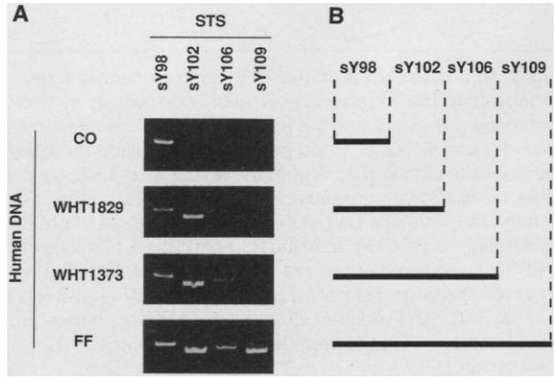


Fig. 1. Ordering of Y-chromosomal breakpoints and STS's by deletion mapping. (**A**) Results of testing genomic DNA's from four individuals with partial Y chromosomes for the presence or absence of four Y-specific STS's. PCR products were separated by polyacrylamide gel electrophoresis and visualized by staining with ethidium bromide. (**B**) Inferred order of breakpoints and STS's, that is, the simplest interpretation of the results in (A).

The authors are at the Howard Hughes Research Laboratories at Whitehead Institute, and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142.

that had been purified by flow sorting. The first library consisted of complete-digest Hind III fragments averaging 4 to 5 kb in length (*13*). The second library, constructed with Y chromosomes flow-sorted from a human-hamster hybrid cell line, consisted of partial-digest Mbo I fragments averaging 20 kb in length. To counter possible biases, two different strategies were used to select clones for sequencing. For the first library, clones deficient in common interspersed repeats were identified by probing with total human DNA; clones showing minimal hybridization were selected. For the second library, clones showing strong hybridization with repetitive human DNA were selected; these should be enriched in interspersed repeats. In all, 296 distinct sequences were obtained, comprising 120 kb, or nearly 2 percent of the Y chromosome (*14*).

Computer matching algorithms were used to identify sequences that contained common, interspersed repeats (*15*). Such sequences are not suitable for creating STS's because they are unlikely to tag a specific site in the genome. As expected, the percentage of sequences judged free of repeats was significantly higher in the first library than in the second (Table 1). Still, 41 percent of clones from the first library contained a repetitive element. This observation is consistent with the hypothesis that repeats should accumulate on the Y chromosome as a result of its restricted meiotic exchange and postulated low gene content (*16*). Similar data from other human chromosomes would provide a useful comparison. In any case, the presence of repeats rendered nearly half the Y sequences unsuitable for STS generation.

A computer algorithm (*17*) was used to select PCR primer pairs from 155 sequences judged free of repeats as well as from the sequences of 40 previously characterized Y-chromosomal DNA probes. Nine additional primer pairs were selected from known Y sequences (Table 2). Primer pairs were initially tested on normal male and female genomic DNA's under a single set of buffer and thermal cycling conditions. More than 95 percent of primer pairs yielded a product of expected size with these standardized conditions, facilitating subsequent STS mapping and YAC library screening (*18*).

For most primer pairs, one of two results was obtained. Either the STS was Y-specific as demonstrated by a product of expected size from male DNA and its absence in female DNA, or the STS was classified as "male-female common" because DNA from both sexes yielded the expected product. A male-female common STS could derive from portions of the Y that share sequence similarity with the X chromosome (*19–21*), from regions of the Y that share similarity with autosomes (*22*), from contamination of the flow-sorted libraries with X or autosomal DNA, or from failure of the matching algorithms to detect an interspersed repeat. As a means of distinguishing among these possibilities, the chromosomal location of male-female common STS's was determined by scoring DNA's from the following hybrid cell lines: (i) a line containing the Y as its sole human chromosome (*23*), (ii) two lines together containing 11 different human autosomes (half the human genome) (*24*), and, as necessary, (iii) a line containing the X as its sole human chromosome (*25*). The results revealed that, of 155 STS's from anonymous phage inserts, 88 were Y-specific, 30 were common to the X and Y, 25 were common to both the Y and at least one autosome, 10 were autosomal, and 2 were X-specific. The 30 X-Y common STS's were regionally localized on the X by means of four hybrid cell lines containing partial X chromosomes (*26*). Because of the highly specific nature of PCR assays, some of the Y-specific STS's may actually derive from regions of X homology. Thus, the 21 percent of Y-derived STS's that are X-Y common may be an underestimate of the fraction of the euchromatic Y that is homologous to the X. The 182 STS's ultimately used in deletion mapping or YAC contig construction are listed in Table 2.

Creation of a deletion map. Deletion mapping of the STS's was achieved by scoring their presence or absence in individuals with partial Y chromosomes. Three hundred individuals suspected of having Y deletions

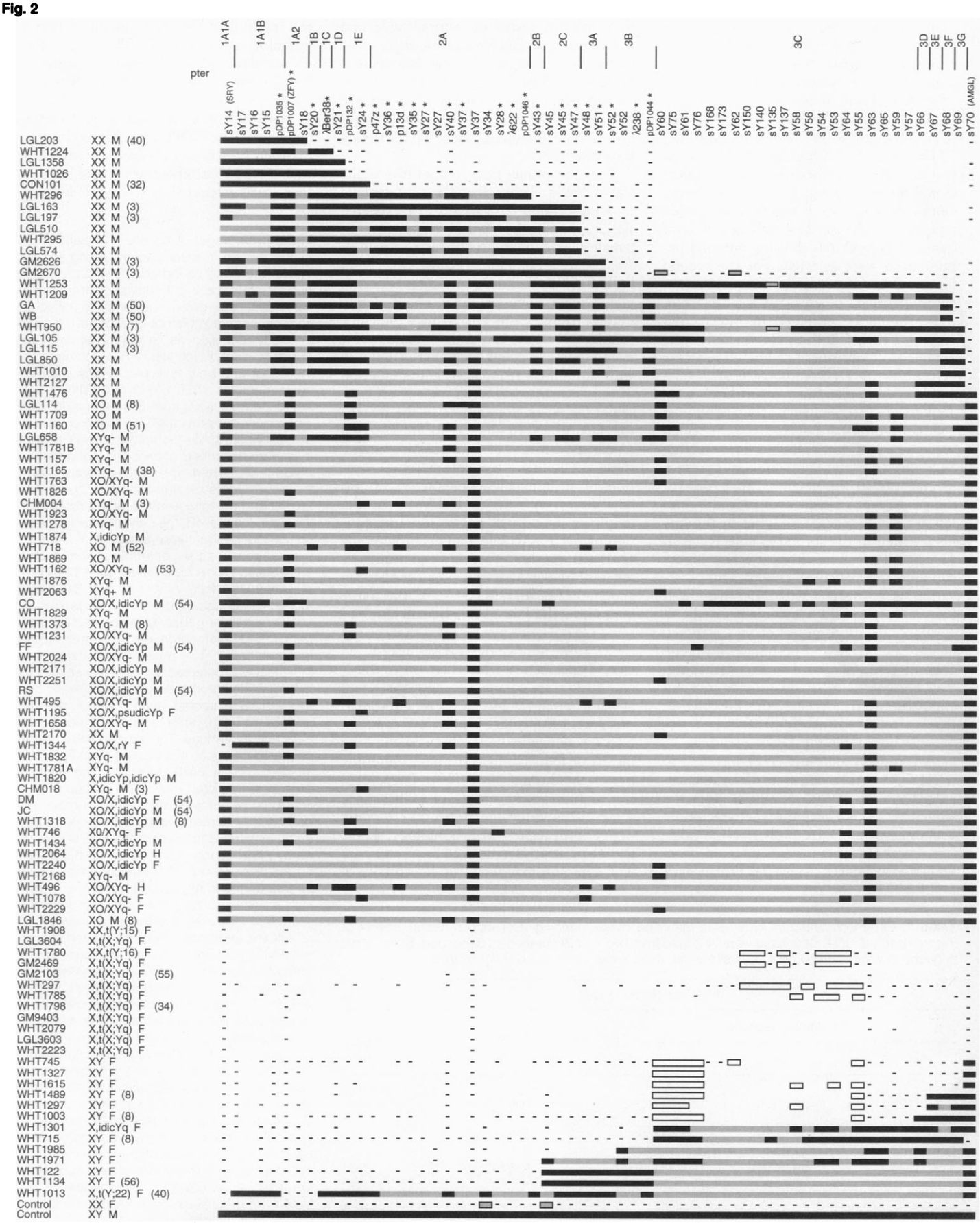
Table 1. Common interspersed repeats identified by sequencing 296 anonymous segments of Y-chromosomal DNA. Segments were obtained from two Y-DNA libraries as described. Some of the sequences were derived from opposite ends of the same clone but did not overlap.

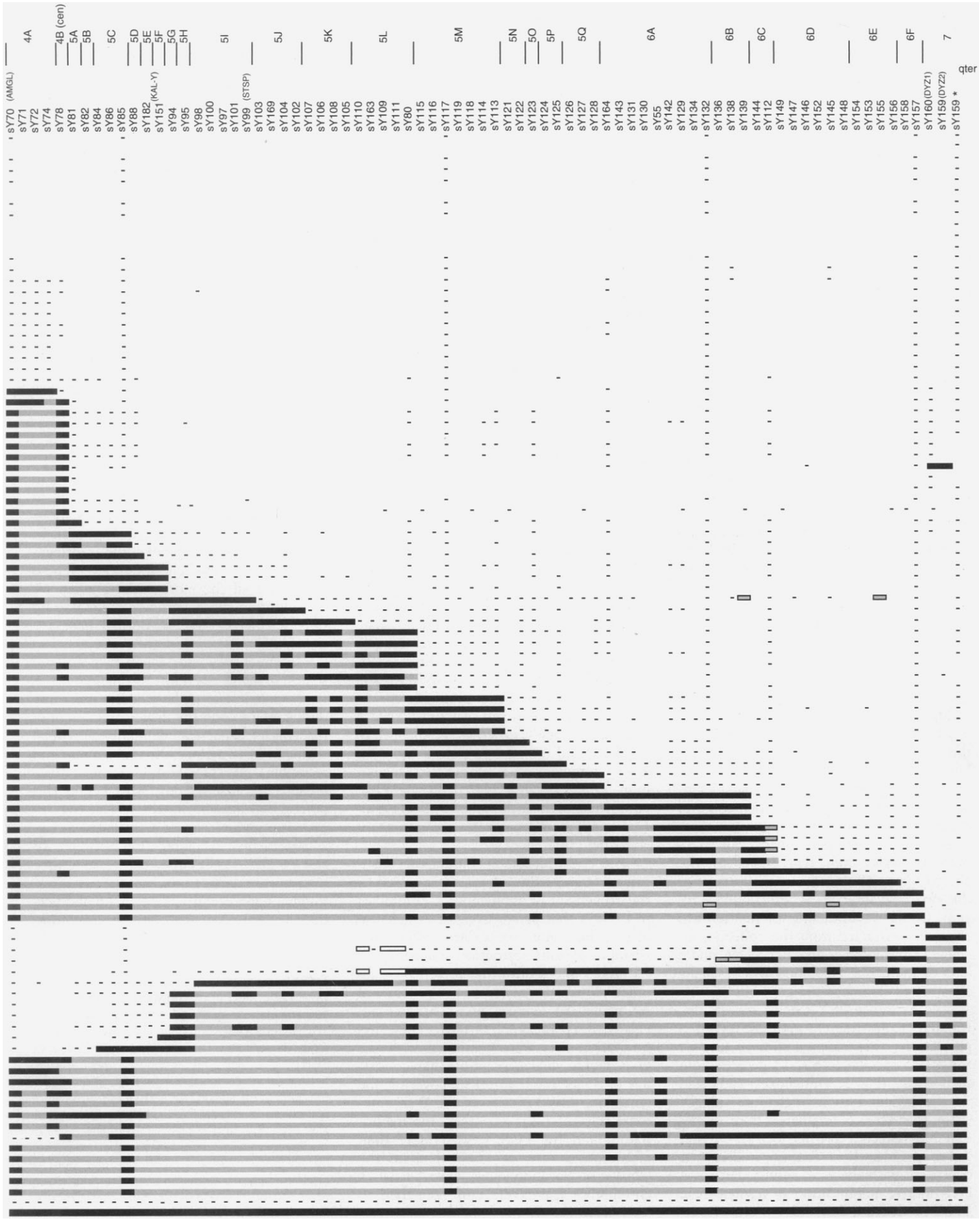
Number of sequences				
Containing repeats			Free of obvious repeats	Total†
Alu	L1	Other*		
First				
21 (12%)	32 (18%)	22 (12%)	108 (59%)	182
Second				
29 (25%)	27 (24%)	10 (9%)	50 (44%)	114
Combined				
50 (17%)	59 (20%)	32 (11%)	158 (53%)	296

*Includes matches to human transposable elements, retroviruses, simple sequence repeats, alpha and beta satelites, mitochrondrial DNA, and unidentified repeats. †One sequence from the first library and two from the second contained portions of both Alu and L1 repeats. These were counted only once in the total.

Fig. 2 (next page). A 43-interval deletion map of the human Y chromosome. Along the left border are listed 96 individuals who carry part but not all of the Y chromosome [abbreviated karyotypes are given; the precise nature of some abnormal Y's is not known, having been originally identified as "markers" of unknown origin; M, male; F, female; H, hemaphrodite (an individual with both testicular and ovarian tissue)]. Samples WHT 1781A and WHT 1781B are cloned cell lines from the same individual. References are provided in the case of individuals for whom some Y-chromosomal DNA findings were described previously. Along the top border are listed deletion intervals 1A1A through 7. The short arm telomere (Ypter) is to the left and the long arm telomere (Yqter) is to the right. Interval 4B, the only segment present on all independently segregating Y chromosomes, contains the centromere (cen). Listed immediately below the intervals are 132 Y-chromosomal DNA loci comprising 122 STS's and ten unsequenced plasmid or phage clones. Locus names for genes, unprocessed pseudogenes, and heterochromatic repeats are given in parentheses. The presence or absence of most loci was detected by PCR; loci scored by hybridization are indicated by an asterisk; and five loci were scored by both methods. The body of the figure represents both experimental data and inferences. The experimentally demonstrated presence of a locus in an individual is indicated by a black segment; the inferred presence (by interpolation) of a locus in an individual is indicated by a gray segment; experimentally demonstrated absence is indicated by a minus, and inferred absence is indicated by the absence of any symbol. White boxes represent positive PCR results, and these must be interpreted in the context of the Y-specific–repeat nature of the sequences being considered. It is very likely that these positive results reflect the existence of closely related, cross-amplifying sequences in other portions of the Y chromosome. Gray boxes represent a few PCR results for repeated or X-Y homologous loci that are positive but of reduced strength relative to results obtained with normal males. Such reduced signals could result from contamination of genomic DNA's, from chromosomal breakage within a repeat array, or from closely related, cross-amplifying sequences elsewhere in the genome. Within an interval, the order of loci is not known. Deletion interval nomenclature was based on the seven-interval map of Vergnaud *et al.* (*3*) and subsequent refinements (*40, 42*). Inclusion of many individuals studied by Vergnaud *et al.* ensured correspondence between the original seven intervals and the 43 intervals shown here.

Fig. 2


1A1A
1A1B
1A2
1B
1C
1D
1E
2A
2B
2C
3A
3B
3C
3D
3E
3F
3G
pter
sY14 (SRY)
sY17
sY16
sY15
pDP1035 *
pDP1007 (ZFY) *
sY18
sY20 *
λBer38*
sY21 *
pDP132 *
sY24 *
p47z *
sY36 *
p13d *
sY35 *
sY27 *
sY27
sY40 *
sY37 *
sY37
sY34
sY28 *
λ622 *
pDP1046 *
sY43 *
sY45
sY45 *
sY47 *
sY48 *
sY51 *
sY52 *
sY52
λ238 *
pDP1044 *
sY60
sY75
sY61
sY76
sY168
sY173
sY62
sY150
sY140
sY135
sY137
sY58
sY56
sY54
sY53
sY64
sY55
sY63
sY65
sY59
sY57
sY66
sY67
sY68
sY69
sY70 (AMGL)
LGL203 XX M (40)
WHT1224 XX M
LGL1358 XX M
WHT1026 XX M
CON101 XX M (32)
WHT296 XX M
LGL163 XX M (3)
LGL197 XX M (3)
LGL510 XX M
WHT295 XX M
LGL574 XX M
GM2626 XX M (3)
GM2670 XX M (3)
WHT1253 XX M
WHT1209 XX M
GA XX M (50)
WB XX M (50)
WHT950 XX M (7)
LGL105 XX M (3)
LGL115 XX M (3)
LGL850 XX M
WHT1010 XX M
WHT2127 XX M
WHT1476 XO M
LGL114 XO M (8)
WHT1709 XO M
WHT1160 XO M (51)
LGL658 XYq- M
WHT1781B XYq- M
WHT1157 XYq- M
WHT1165 XYq- M (38)
WHT1763 XO/XYq- M
WHT1826 XO/XYq- M
CHM004 XYq- M (3)
WHT1923 XO/XYq- M
WHT1278 XYq- M
WHT1874 X,idicYp M
WHT718 XO M (52)
WHT1869 XO M
WHT1162 XO/XYq- M (53)
WHT1876 XYq- M
WHT2063 XYq+ M
CO XO/X,idicYp M (54)
WHT1829 XYq- M
WHT1373 XYq- M (8)
WHT1231 XO/XYq- M
FF XO/X,idicYp M (54)
WHT2024 XO/XYq- M
WHT2171 XO/X,idicYp M
WHT2251 XO/X,idicYp M
RS XO/X,idicYp M (54)
WHT495 XO/XYq- M
WHT1195 XO/X,psudicYp F
WHT1658 XO/XYq- M
WHT2170 XX M
WHT1344 XO/X,rY F
WHT1832 XYq- M
WHT1781A XYq- M
WHT1820 X,idicYp,idicYp M
CHM018 XYq- M (3)
DM XO/X,idicYp F (54)
JC XO/X,idicYp M (54)
WHT1318 XO/X,idicYp M (8)
WHT746 X0/XYq- F
WHT1434 XO/X,idicYp M
WHT2064 XO/X,idicYp H
WHT2240 XO/X,idicYp F
WHT2168 XYq- M
WHT496 XO/XYq- H
WHT1078 XO/XYq- F
WHT2229 XO/XYq- F
LGL1846 XO M (8)
WHT1908 XX,t(Y;15) F
LGL3604 X,t(X;Yq) F
WHT1780 XX,t(Y;16) F
GM2469 X,t(X;Yq) F
GM2103 X,t(X;Yq) F (55)
WHT297 X,t(X;Yq) F
WHT1785 X,t(X;Yq) F
WHT1798 X,t(X;Yq) F (34)
GM9403 X,t(X;Yq) F
WHT2079 X,t(X;Yq) F
LGL3603 X,t(X;Yq) F
WHT2223 X,t(X;Yq) F
WHT745 XY F
WHT1327 XY F
WHT1615 XY F
WHT1489 XY F (8)
WHT1297 XY F
WHT1003 XY F (8)
WHT1301 X,idicYq F
WHT715 XY F (8)
WHT1985 XY F
WHT1971 XY F
WHT122 XY F
WHT1134 XY F (56)
WHT1013 X,t(Y;22) F (40)
Control XX F
Control XY M


4A
4B (cen)
5A
5B
5C
5D
5E
5F
5G
5H
5I
5J
5K
5L
5M
5N
5O
5P
5Q
6A
6B
6C
6D
6E
6F
7
qter
sY70 (AMGL)
sY71
sY72
sY74
sY78
sY81
sY82
sY84
sY86
sY85
sY88
sY182
sY151 (KAL-Y)
sY94
sY95
sY98
sY100
sY97
sY101
sY99 (STSP)
sY103
sY169
sY104
sY102
sY107
sY106
sY108
sY105
sY110
sY163
sY109
sY111
sY80
sY115
sY116
sY117
sY119
sY118
sY114
sY113
sY121
sY122
sY123
sY124
sY125
sY126
sY127
sY128
sY164
sY143
sY131
sY130
sY55
sY142
sY129
sY134
sY132
sY136
sY138
sY139
sY144
sY112
sY149
sY147
sY146
sY152
sY145
sY148
sY154
sY153
sY155
sY156
sY158
sY157
sY160(DYZ1)
sY159(DYZ2)
sY159 *

were studied in our laboratory over 10 years. These individuals were ascertained either by microscopic detection of an aberrant chromosome or by discordance between sex chromosome constitution and sex phenotype. At the time of DNA collection, a coarse map of the sex-linked portion of the chromosome (*3*) was used to localize grossly the breakpoint in each individual by screening with a small number of probes. A Y-chromosomal breakpoint was detected in about 150 of the individuals tested (*27*). More detailed information on the chromosomal breakpoints in some of these individuals and on about 80 of the clones from which STS's were derived was available because of efforts to refine the deletion map by Southern blot hybridization prior to initiation of the PCR-based strategy. This information served as the starting point for construction of the PCR-based deletion map.

We mapped 104 Y-specific STS's by scoring the presence or absence of a band of expected size after resolution of PCR products by gel electrophoresis (Fig. 1). Frequently, an STS was localized to a single interval with as few as ten reactions, by successively scoring small numbers of chromosomes selected to provide maximum information. Most STS's were initially localized to one of two regions, corresponding roughly to Yp (the short arm) and Yq (the long arm), and then further localized by scoring deleted chromosomes known from hybridization results to subdivide Yp or Yq. When STS's originated from probes present on the hybridization map, the process was abbreviated. As mapped STS's accumulated, additional chromosomes were recruited from among the 150 to enhance the resolution of the map. Selected STS's mapping to the same region were scored against chromosomes with possible breaks in that region. New intervals were created when some of these STS's were present on a partial chromosome and others were absent. All STS's from the region were then scored on the chromosome and thereby assigned to one of the two new intervals subdividing the region. In this way, ordering information was efficiently extracted from the collection of deleted chromosomes and 104 STS's, and a map of increasing resolution was gradually constructed (Fig. 2).

The existence of Y-specific repeats distributed about the chromosome posed a problem for this mapping strategy (*3*). Of the 104 Y-specific STS's, 18 derived from such repeats (Fig. 2, white boxes) and could not be localized to a single interval. Of these 18 STS's, 5 originated from probes that had detected distinct loci on Yp and Yq in previous hybridization studies (*3*), consistent with the PCR results. For some PCR assays, amplification of normal male DNA produced a complex pattern of heteroduplexes, indicative of repeated sequences, and in one case (sY55), the pattern was used to map both a Yp and a Yq locus. Of the 18 Y-specific repetitive STS's, 15 were conclusively localized to interval 3C, on Yp, with evidence for at least three other Y-specific repeat blocks, one near the centromere (intervals 4A to 5H), and two on Yq (intervals 5I to 6A and 6B to 7). Three Y-specific repetitive STS's were not present on Yp but mapped to two regions of Yq (intervals 5L and 6B to 7). An additional two assays (not among the 18) were obviously derived from Y-specific repeats because each generated two products of different sizes that mapped to separate intervals, defining a total of four STS's (sY63, sY164, sY80, and sY112).

Certain other regions of the Y chromosome were not amenable to PCR-based mapping because they share a high degree of sequence similarity with the X chromosome. Many of the PCR assays from these regions produced indistinguishable products from the X and Y chromosomes and, because the X chromosome is present in all individuals, could not be mapped with the panel of Y deletions. Instead, regional localization of the 30 X-Y common STS's on the X chromosome (see above) made it possible to deduce their likely location on the Y on the basis of prior knowledge of X-Y homologies. Of these STS's, 23 mapped to proximal Xq, suggesting that they originate from a portion of Yp with more than 98 percent sequence identity to Xq21 (*19*). Seven other STS's mapped to distal Xp and probably derive from the pseudoautosomal region or two sex-linked regions of Xp22.3 and Y homology (*21*). One STS mapped to distal Xq and may derive from a region of Xq-Yq homology (*22*). Of the 30 X-Y common STS's, 29 were placed on the YAC contig map of the Y described by Foote *et al.* (*11*), and their locations were consistent with these inferences.

The Yp-Xq21 homologous region constitutes a large portion of Yp, as suggested by the 16 percent of anonymous Y-chromosomal STS's that derive from it. A deletion map of this region was constructed by identifying Y-specific restriction fragments for 25 probes and scoring the presence or absence of these fragments in the genomes of XX males, XY females, and other individuals with Yp breakpoints (*28*). These hybridization data were combined with the PCR data to yield a composite map consisting of 43 deletion intervals covering the entire sex-linked portion of the Y chromosome (Fig. 2) (*29*). Together, the hybridization and the PCR yielded more than 2900 data points for 132 Y loci on the DNA's of 96 individuals with partial Y chromosomes, an average of more than 30 loci scored per chromosome.

The 97 well-characterized Y chromosomes (Fig. 2) derive from human individuals, many with abnormal phenotypes. They represent a valuable resource for localizing phenotypes to portions of the Y chromosome by phenotype-karyotype correlations. The Y-chromosomal phenotypes that have been identified but for which candidate genes have yet to be discovered include a gene necessary for the expression of H-Y antigen; a spermatogenesis factor; a gene contributing to gonadoblastoma, a rare neoplasia; and stature determinants (*1, 7, 8, 30*). Because 37 of the 43 intervals can be scored by PCR alone, the extent of Y-chromosomal DNA present in any individual can be assessed in a matter of hours with minimal consumption of DNA. The ability to automate PCR should allow the scoring of a large number of chromosomes for many loci, facilitating detection of small rearrangements, the type most useful in localizing genes.

The nature of Y-chromosomal rearrangements. The mechanisms that created the aberrant chromosomes shown in Fig. 2 are not clear. Of 95 breakpoints falling in the euchromatic region (intervals 1A1A to 6F), 54 are the result of translocations between the Y and other chromosomes. In all but 10 of these 54 cases, the translocation partner is the X chromosome. These results are all the more striking when we consider that the X chromosome makes up only 2.5 percent of the normal male genome (*31*).

The frequency with which two chromosomes undergo aberrant recombination is likely a function of both their proximity during meiosis and the degree and extent of nucleotide sequence similarity between them. In male meiosis, Xp and Yp pair and recombine at their distal extremes (the pseudoautosomal region). Mistakes in this process may account for the 34 X;Y translocations (in XX males and XY females) in Fig. 2 that involve recombination between Xp and Yp (*32*). As discussed earlier, extensive sequence similarity (outside of the pseudoautosomal region) exists between Xp and portions of Yp and Yq. Of 42 Xp;Y translocation

Table 2. Y-chromosomal STS's. In all, 182 STS's were generated from known sequences or by sequencing anonymous fragments and ends of YAC inserts (*45*). For each STS are listed the locus designation, PCR primer sequences, and PCR product size. The "left primer" was arbitrarily defined as corresponding to the sequenced (or published) strand. All sequences are listed 5′ (left) to 3′ (right). PCR products for several X-Y common and Y-specific repetitive STS's displayed heteroduplex bands when resolved by polyacrylamide gel electrophoresis due to cross-amplifying sequences on the X or elsewhere on the Y. These bands were used to score the presence or absence of Y loci.

Table 2.

STS	Locus	Left Primer	Right Primer	Size (bp)	STS	Locus	Left Primer	Right Primer	Size (bp)
sY 1*	DXYS14	ACCCGCTAATCTCGGCTATC	ATGACGGAAAGTGACGGAGA	109	sY 92	DXYS126	TAGCACTCAGAGTGATGTCACC	AGTAAAACCAACAGCAAGAACA	211
sY 2*	DXYS20	GTCCTCAATGAGACCTAGGCC	TTAGAGTCTGCATTGGGCCT	219	sY 93	DXYS127	CCAGATATGTCCACATGTTCC	TTCTAAAACCTGTACATTTGAAGG	320
sY 3*	DXYS28	GGAACTGGCTTTTCATTTCC	AGGTCACCTGGATGGTCAGT	200	sY 94	DYS279	TCATGACAGCCAGGGTATTT	TTTGGACATAGTTTTTTGGTCC	150
sY 4*	DXYS87	CTGCTCTCATCTGGAAGCG	GGTGCTGAGGAATGAGACTTG	139	sY 95	DYS280	TCCTACAGATGTCCAAAGTGC	GATGAGTGACCCCAGAATTG	303
sY 5*	DXYS86	AGAAGTTTCTTCTCCACCTGC	TCTGCCATCCTAATGATGGT	144	sY 96*	DXYS3	TTCCCATACAATGGTGTGTG	AGGTGTAAGGTGGGCAAACT	93
sY 6*	DXYS15	TATTTATGGAAATTGCCCCC	TAATACAAGCCAGACGAGCC	281	sY 97	DYS281	AACTTCATCAGTGTTACATCAAGG	TGTGGCATTTTGTTATGTGG	104
sY 7*	DXYS85	ACCACAGGGCCTATCGTG	TTTGCTGAGCACCTAGAAGG	201	sY 98	DYS282	TGTCAGCAGGCTTAGTTCCT	CCTCTTCCCCACTACTTCAA	266
sY 8	DXYS125	CCCAGGGTCTAATCACTTGA	CAGGGATAAAGGAATTTGTCC	138	sY 99*†	STSP	GACTCAGGGATCCAGGTTG	GCACTGCAACTTTTATGCCT	357
sY 9*	CSF2RA	AAATTCAATGGCTTCACTCCA	TTCACCAATGGAGTTTTAAACC	100	sY 100	DYS196	TAAAGGAACTTCTGTGTGTAAACA	TAAGCCAGATAGGGGCTTCT	111
sY 10	DXYS91	ATTCCTTCTCGTGGACATCA	GTCTGAGCCATCTCACACCT	107	sY 101	DYS197	TTTTGCACCTTCACAGATGA	TTCATTGAAAGCTGACACGA	131
sY 11	DXYS92	CATGTGAACAGTACACATCTCTG	ATAATAATTTTCTACACGCAGTTCC	103	sY 102	DYS198	CACTACCACATTTCTGGTTGG	CGCTGAGTCCATTCTTTGAG	218
sY 12	DXYS93	CCGAAGATTCAGGTTCTGAA	GCTTTTGTTTCTTCACTGGG	300	sY 103	DYS199	TAATCAGTCTCCTCCCAGCA	AAAATTGTGAATCTGAAATTTAAGG	241
sY 13*	DXYS77	GTGACACACAGACTATGCTTC	TCAAGGTTGTTGTTTAAGCT	187	sY 104	DYS200	CAGCAAAAAGGACTAAGGCA	TGGATCTCTGGAATTTGGAA	130
sY 14*	SRY	GAATATTCCCGCTCTCCGGA	GCTGGTGCTCCATTCTTGAG	472	sY 105	DYS201	AAGGGCTTCTTCTCTTGCTT	AGGGAGCTTAAACTCACCGT	301
sY 15	DYS234	TCCTAGGTTATGATTACAGAGCG	TGCAGAACATTTGTACTGTTCC	273	sY 106	DYS202	AACACGCATTTCCTTTTCAA	TTTAGCCATAATTATCTGTGGAAGA	231
sY 16	DYS242	CCTGGTGCTTCTGTGAAAAA	TGAAAGGAGCATAGTCCTGC	210	sY 107	DYS203	TTCTCTGGCTTCCTATTCCA	ATCCTCCAACAACTGTGGAA	173
sY 17	DYS250	CAGACGGAACTATCTCACAGG	GCTGAGAACAGTGCTAAGGG	329	sY 108	DYS204	TTGTGGATTGTTGTTTTTGTTG	AAGACAATGTTGTACCGGCA	360
sY 18	DYS251	ATTTGGCTTGGAGTCAGTCA	TATGAACCAATCTCACCCCC	144	sY 109	DYF43S1	AGGAGATGTCAGGACTATCAGC	TCCATCCAGCTGGTCATATT	233
sY 19	DYS252	AAACCCATTCTCACCCC	TTCCCAAGGTTGCTAGTACA	220	sY 110	DYF44S1	GCTCACAGTCCAAAAGGAAA	GTCTTGTGCAAACTCTGGCT	179
sY 20*	DXYS42	TCATAGCCCAAGACAAGGAG	ATGGTAAATCCAGCACAGGA	295	sY 111	DYF45S1	AAAAGCTGAGCTCACCAATG	CATTCCACAAGCAATGAATG	200
sY 21*	DXYS69	CTCTGGGAAACTGACCAAGA	CATTAGAGTTATAACCGAGTGCG	322	sY 112¶#	DYF46S1	CATTTGCTCAGTGGGAGTCA	ACCTTTCAGTGAGGGTGGAA	1200§
sY 22	DXYS106	TGAGTTGGCTTGAGCATGTA	TACACTTGGAAGGGATCCAA	301	sY 113#	DYS205	GTTCTTTCCACAGCCCATAG	TGGAACACAATCCAAAATTG	290
sY 23	DXYS107	AAGTTATCAAAATTTGAGGAAAGC	TTGAATTTAACGTGCCCTGT	139	sY 114#	DYS206	TGCACTCATGGAGACAACAG	AACCAGGGTTTTCACTGAAA	1450
sY 24*	DXYS5	TATGATGTTAGCCATGGGCT	ATGAGGCCAGCATTACTCTG	373	sY 115	DYS207	TTTCACACTGTTAGAAATCAGGC	TGGGACACTATGGAGGCTAA	115
sY 25	DXYS108	GCATCAACACAGACCAACAA	GGTTGGCTACAGAAAGAAGG	326	sY 116	DYS208	TGTGTCATTGCACTTTAGCC	CATTCCCCATGAAGTCAAAC	147
sY 26	DXYS109	AAAAATACAGATGTTCATAGGTTCG	TTTACCTATGATCTGGGCCA	170	sY 117	DYS209	GTTGGTTCCATGCTCCATAC	CAGGGAGAGAGCCTTTTACC	262
sY 27*†	DXYS10	AGAAGAAAAACCACAAGCCC	TCAACTATGGGGACAGTGGT	240	sY 118	DYS210	GTACCTCTGCAGGCACTGAT	ACACAATCCAACCTGGCTAA	218
sY 28	DXYS110	ACTATTTTGCCCATGTGGAC	ATGAGCTGCTGTCTCAATCC	311	sY 119	DYS211	AGGCTCCATCTGTAGCACAC	TAACCTTATAGACCAACCCCG	191
sY 29	DXYS111	CCCTATCAGCTCTGTAGCCA	AGCAGCACAGAACCCTTATG	283	sY 120	DYF47S1	ACATTGTCAATCAGTGTCAAGG	AAAAGGCAAGTGAATGAAGC	176
sY 30	DXYS112	CTTCAGATCAGATTAAGGTGCTCT	GGGAAGCATTGACTGCATTA	231	sY 121	DYS212	AGTTCACAGAATGGAGCCTG	CCTGTGACTCCAGTTTGGTC	190
sY 31	DXYS113	TGGCACAATCTCAGCATATG	CCTGTCTTTGAGTGGGAGAA	216	sY 122	DYS213	AGAATGAGTGCACCTGTAACA	CACCCCTAATAATACAGCTTCC	201
sY 32	DXYS114	TTTTTCATGAAGTTTGACAAGC	ATTGTAAAATATTAAATAGCTGGCC	140	sY 123‡	DYS214	CCGTGTGTTGCTGGGCTGTC	GGGGTTTATACTGACCTGCC	649
sY 33	DXYS115	TGAAAACTGGCTCTCCTCAC	AATCAACATTGAGAACATCTGAGA	198	sY 124	DYS215	CAGGCAGGACAGCTTAAAAG	ACTGTGGCAAAGTTGCTTTC	109
sY 34‡	DYS253	ACTGTGAGCGAGCTGAAAAT	GCAGCCTTGTGAACCAATTA	280	sY 125	DYS216	GGGATAGGGAAAGGGTACAA	CCGGGAGAAAAAAAACTGAA	200
sY 35	DXYS116	CATTCTCGTGAGCTTCCTGT	CGCTGTGACTCGGATAACTT	328	sY 126	DYS217	AAAAATGAGTGGCACTATGTACA	CTTTCAAGTGCATACATGGTG	323
sY 36*	DXYS2	GAGTAACGTTGTAACATAGAATTGC	ATTGAGATGTGTACTTTTCCACA	199	sY 127	DYS218	GGCTCACAAACGAAAAGAAA	CTGCAGGCAGTAATAAGGGA	274
sY 37*†	DXYS6	TTTCTGTGTATGTGTCACAGAGC	TGTGTTAGCCCTTCTTTTGC	238	sY 128	DYS219	GGATGAGACATTTTTGTGGG	AGCCCAATGTAAACTGGACA	228
sY 38	DXYS117	ATGCTGCTGTGTTTCTGTGT	AAAATTGAGGCTTCCTTGC	184	sY 129	DYS220	AGCTTCAGGAGGTTCAAAAC	AAGTGGGACCTAAGCTACGA	194
sY 39	DXYS118	TGAGAAAAATATGAACAAAGCTG	TGCTGATGGAAAGGTTATATAGG	330	sY 130	DYS221	AGAGAGTTTTCTAACAGGGCG	TGGGAATCACTTTTGCAACT	173
sY 40*	DXYS8	CTTGGCAGGACTGCTAATTG	TGCTTAAAAAAGCCATCTGG	241	sY 131	DYS222	ACATATCCCTTGCCACTTCA	TCAGGTACCTTCTGCCTGAG	143
sY 41	DXYS119	CCCAATCTCTCCTGGCTATT	TGTAACCATGCAACCATGAG	171	sY 132*	DYS7	GAGAGTCATAATGCCGACGT	TGGTCTCAGGAAGTTTTGC	159
sY 42	DXYS120	TCTTGATTTAACTCTCAACCTGG	CCATTGTACTGGAAGCCCTA	260	sY 133	DYS223	ATTTCTCTGCCCTTCACCAG	TGATGATTGCCTAAAGGGAA	177
sY 43	DYS254	CCTTTCATGACGTATAGCTGG	AAAGTTTTGAATACCAAGAAGAAGG	300	sY 134	DYS224	GTCTGCCTCACCATAAAACG	ACCACTGCCAAAACTTTCAA	301
sY 44	DXYS121	GAAATTGACAGATCAGCTACTGG	GAGGGTAAGGTGGCTGAATT	259	sY 135	DYS225	CATTTCCACATTATGATAATTATGC	ACCCAGAGAGTAGAAACAGTGC	253
sY 45‡	DYS255	CTCATGCAAATATGACTTTAATAGC	TGATGATTTTCTCAATGTGAGG	170	sY 136	DYS226	CACATGAAGCACTGGAACTG	GTTGTCTGGAAATCCCTGTG	235
sY 46	DXYS122	TATTTGTGCGAAAATCTGGC	CCAGATGCCACAGTTCACTA	202	sY 137	DYF48S1	AGGGTAAAATGCATAAGCCC	TTATCCTGCTGCTGGGTAAG	312
sY 47*	DXYS9	ATCTGTGAAGCACCCACTGT	CCTGTTGCAAATGACAAATG	174	sY 138	DYF49S1	CACATGAAGCACTGGAACTG	AGGGCCTGAGTCTCCAGG	170
sY 48*	DXYS12	GAGGCAATCACAAATTCTGG	TGGCTGAATGATAACCCATT	329	sY 139	DYS227	TTCAGAGGAATCATGTGGGT	AATGTTTCATCACCATTATCCC	120
sY 49	DXYS123	ATACAGCTTTATTTCTTGCTTTTC	CGTGTTTCAAAACGCTGTAG	104	sY 140	DYS228	ACAAGTCCTCAAACACACTGG	CTCCATGCTTGCTTTTTCTC	107
sY 50	DXYS124	TTAAAGGCATTGGTGGAAAA	TACTTTCCCTGTTTGTGCCA	169	sY 141	DYS229	GCAGTTCCATTGTTTGCTTC	GCAGCATAATAGCTATACAGTATGG	290
sY 51	DXYS128	AAGCACGCCTACCTTCACCT	CTTTGCTAGGTAAGACCCACAA	300	sY 142	DYS230	AGCTTCTATTCGAGGGCTTC	CTCTCTGCAATCCCTGACAT	196
sY 52*†	DXYS4	CTCAAAATTTCCGTGTGCTT	CGCCAAAACAGGAGATTAGA	292	sY 143	DYS231	GCAGGATGAGAAGCAGGTAG	CCGTGTGCTGGAGACTAATC	311
sY 53	DYF56S1	CAGACACAGCAGGACTAGCA	GCAATTTCAATAGGAAGCCA	231	sY 144	DYF50S1	TCATCTGCCACCATCAACAT	ACGTGTTTCTACACCTGCCC	143
sY 54	DYF57S1	CCAGAGATTTATGGCAGGTG	GGCTGAATCTCCAGCAGTAGT	222	sY 145	DYF51S1	CAACACAAAAACACTCATATACTCG	TTGAGAATAATTGTATGTTACGGG	160
sY 55*†	DYF67S1	CTTTGATTCCAGGCATGAAC	CTCAACACAAAACCACCCAT	256	sY 146	DYF52S1	ACAAAAATGTGGCTCAGGGA	AAATAGTGTGCCCACCCAAA	173
sY 56	DYF58S1	GACTGCCAGCCTCATAAAAA	TCCAGAAGGCATGTTAGGAA	251	sY 147	DYS232	TTTCTCGTTTGATGATCCTAG	TTAATATGAGAATGAGAACAGATGT	100
sY 57	DYS257	GAACTTGTCGGGAGGCAAT	TGATACACTTCCTCCTTTAGTGG	288	sY 148	DYS233	AAATGAAAAAAGATACGAAACTCG	GAATCCCACCCAAGAATCTG	202
sY 58	DYS258	TTCCCAAGTGTCATCCTGTT	AACTACCCCAAATCGGTCTC	144	sY 149*	DYS1	TGTCACACTGCCCTAATCCT	TGGTCATGACAAAAGACGAA	132
sY 59	DYF59S1	AAATCTGTACATTCCTAACAGCG	TGCAAAGGATGGATTTTTGT	267	sY 150	DYS235	GGGAGAGTCACATCACTTGG	TTGAATTATCTGCCTGAGTGC	158
sY 60*	DYZ4	GGTTTGGGATCCTATTATTGAGTG	TTTTTCCTTAGTATCTAGGCCAATG	311	sY 151**	KAL-Y	AAATCTGTAGTCTCATATCAATCTG	TTACTTGATTTAGCAATAAAAAGG	183
sY 61	DYF60S1	GTCTCTCCTGGAAAGAAAACG	AACAAAACAAAACAGTCACTGG	221	sY 152	DYS236	AAGACAGTCTGCCATGTTTCA	ACAGGAGGGTACTTAGCAGT	125
sY 62	DYF61S1	TGGCTCAAAGGTTTGGAATA	GCTATAATGGCTCAAACTACTGG	151	sY 153	DYS237	GCATCCTCATTTTATGTCCA	CAACCCAAAAGCACTGAGTA	139
sY 63‖	DYF65S2	AATGTGCCCACACAGAGTTC	TGGAAGACCAGGATTTCATG	625§	sY 154	DYS238	TTTGCACCAGGATTAAGTGA	TTTTTTCAGATAAACTTTCAGTGG	245
sY 64	DYF62S1	CAACTGCACTCTCAGAAGAGC	TCACAAACATTCAGCGGATA	101	sY 155	DYF53S1	ATTTTGCCTTGCATTGCTAG	TTTTTAAGCCTGTGACCTGG	349
sY 65	DYS260	ACTAAAACACCATTAGAAACAAAGG	CTGAGCAACATAGTGACCCC	309	sY 156#	DYS239	AGGAACTGGCAGGATTAGCC	ATGTCAGGGTTTCCTTTGCC	950
sY 66¶	DYS261	CTGGACTGCACAAAACAACA	AGAATATGGTGGGTGGGACT	293	sY 157	DYS240	CTTAGGAAAAAGTGAAGCCG	CCTGCTGTCAGCAAGATACA	285
sY 67	DYS262	ACCGTGCCTGGCTATTATTT	ACAGCATCAAGAGCACATGA	301	sY 158	DYS241	CTCAGAAGTCCTCCTAATAGTTCC	ACAGTGGTTTGTAGCGGGTA	231
sY 68	DYS263	CCCACCCACTTCAGTATGAA	AGGCTGACAGACAAGTCCAC	267	sY 159*††	DYZ2	TACATGTTATGTGCTATGCC	CACATTATATAATATGTATGTTGTC	550§
sY 69	DYS264	GGAACAGCATCTTGCTCTGT	ACTATGGGAGACCAAGGCTC	234	sY 160*	DYZ1	TACGGGTCTCGAATGGAATA	TCATTGCATTCCTTTCCATT	236
sY 70*	AMGL	AGCTTGGTTCTATCCCATCC	ACATTTGTCAGCAGCTTGTG	367	sY 161	DYS243	GCTAGTCCAATTTCTTGCCA	GCCTAGCAAAGAGTCAGTGC	316
sY 71	DYS265	CCATCTGGCTCAATGGTTAG	CTGAAGGTGGCCATTTCTTA	122	sY 162	DXYS94	GGAGGATACAGTGCTTGCAT	TGTTTGTGTCAGACTATCACAGC	157
sY 72	DYS266	CTTGTGTGACATTCCCTCCA	ATGTTTGTGGGTCATTCAGG	109	sY 163	DYS244	CTGGGTATTCCCAGATAAAA	CTGCTGAAAGGAAATAATTAATT	98
sY 73*	DXYS1	GAATGTGGAATTGCCTTTTG	AGCCCAAATTAGCTAGGTGG	325	sY 164‖	DYF65S1	AATGTGCCCACACAGAGTTC	TGGAAGACCAGGATTTCATG	690§
sY 74	DYS267	TTTTAGAGTCATTGGCCAGG	CTCTGAAAAAAGGCAGCAG	132	sY 165	DYS246	CGTCTCAAAAACAAACAAACAA	TGTTTCCACATCTGCCATATT	207
sY 75*	DYF27	TTCAAGAAGAGGGTCACACA	TTCTTTCCTGGCATTTATGC	275	sY 166	DYS247	GAACTCCAATCATTCCCTGA	TTGGCTCTACTTTTCCCCTT	115
sY 76	DYS268	ATGTACAGGGAGAGACAGATAATAA	TCTGTTATACCAACCATCACTACC	115	sY 167	DYS248	GAAGGCATTAAGACATAAAGCA	ACTTTTGATATGAAAACAAACTTCA	117
sY 77	DYS269	TTCAAACTCCACAACTTCACTG	GTTTTCATTGCTGTCTGTGAGA	322	sY 168	DYF54S1	CATTGGTAGAAATATGTAACAGAGG	TTGCCCACATATGGGATTAA	175
sY 78*	DYZ3	TCCTTTTCCACAATAGACGTCA	GGAAGTATCTTCCCTTAAAAGCTATG	170	sY 169	DYS249	AGCCTGTAGACACATTAGCTCA	TTTATTTCTGAAGAAATTACAATGG	194
sY 79	DYS270	TTCCATTTGATGTATTCCCG	GATTGGACTGGAATGGAATG	216	sY 170	DXYS95	GGGAGGAGTATGTCGTTCTC	CCAGATGGATTCACAGCATA	134
sY 80¶#	DYF46S2	CATTTGCTCAGTGGGAGTCA	ACCTTTCAGTGAGGGTGGAA	1075§	sY 171	DXYS96	CACATTACTGGTGGGAGGAG	CATCATCGCCCTACAGTACC	312
sY 81	DYS271	AGGCACTGGTCAGAATGAAG	AATGGAAAATACAGCTCCCC	209	sY 172	DXYS97	TAGTGGCCATCAGGACAATT	CGCTGTTTCTGCATTTTTTT	109
sY 82	DYS272	ATCCTGCCCTTCTGAATCTC	CAGTGTCCACTGATGGATGA	264	sY 173	DYF55S1	AACATCACTTTAAGTGTTCTGCC	CAGCACCCAGATGATGTTAC	304
sY 83*	DYS11	CTTGAATCAAAGAAGGCCCT	CAATTTGGTTTGGCTGACAT	275	sY 174	DXYS98	ACCGGAATACAGCATGAAGA	AGTTGGCTTTGATGGCCTC	83
sY 84	DYS273	AGAAGGGTCTGAAAGCAGGT	GCCTACTACCTGGAGGCTTC	326	sY 175	DXYS99	TGACACCATGTACCCAAATG	GGGTTCACAGGTTGTGTTTC	109
sY 85#	DYS274	TGGCAATTTGCCTATGAAGT	ACAGGCTATTTGACTGGCAG	369	sY 176	DXYS100	GAAGTTTCGAGCACCATTGT	ACTCGCCTGGAAGTATAGGC	218
sY 86*	DYS148	GTGACACACAGACTATGCTTC	ACACACAGAGGGACAACCCT	320	sY 177	DXYS101	TGACTGTTGTGATATTTTTCTGG	TCCTGCACCCATAAATCAGT	267
sY 87	DYS275	TCTGTTGCTTGAAAAGAGGG	ACTGCAGGAAGAATCAGCTG	252	sY 178	DXYS102	TACTAAGAGCCAAAATTCCCA	TTCTGACAGCAGGAGACGTG	89
sY 88	DYS276	TTGTAATCCAAATACATGGGC	CACCCAGCCATTTGTTTTAC	123	sY 179	DXYS103	ACCGGAGTCCTGGTCTCTTG	TTCCTCAGGCGTCTTCCC	97
sY 89	DYS277	CTGTTGCAGGAATTAAGGGA	TATGGTGAAGGTCATGAGACA	203	sY 180	DXYS104	TCACACTGTTGTCTTAACCTGC	TTCTCAAGAACCAGGAACCA	105
sY 90	DYS278	CAGTGCCCCATAACACTTTC	ATGGTAATACAGCAGCTCGC	176	sY 181	DXYS105	TTTCTGGAAAATTCATTTTTCG	TTAGCTTGATTGTTATGTTAATTCG	203
sY 91	DYS136	CCCAAAAGTGACCACTGACA	ATCTTTCCAAAGGGAAATTGG	101	sY 182**	KAL-Y	TCAGAAGTGAAACCCTGTATG	GCATGTGACTCAAAGTATAAGC	125

* See (44). † Scored using heteroduplexes. ‡ Anneal at 65°C. § Approximate size. ‖ See (46). ¶ Also a distinct X locus. # See (47). ** See (48). †† See (49).

breakpoints, 19 bound Y intervals known to contain Xp-homologous sequences (*33*). In one instance, the translocation breakpoint has been sequenced and found to result from homologous recombination (*34*). Thus, a mix of X-Y sequence similarity and meiotic proximity may account for the observed preponderance of X;Y translocations in our data set (*35*).

Other types of recombination events may have generated deleted chromosomes elsewhere on the map. The numerous blocks of Y-specific repeats distributed about the chromosome could serve as substrates in intrachromosomal recombination events. For example, isodicentric chromosomes could be generated by sister chromatid exchange within a repeat array or between two separate arrays, provided that recombination occurred between inverted repeats on the same arm of the chromosome (*36*). Cytologically, the two halves of an isodicentric chromosome appear identical, as though a fragment of the chromosome containing an entire arm, the centromere, and a portion of the other arm duplicated and fused at the point of breakage. Of the 97 abnormal Y chromosomes, 14 are known to be isodicentric (Fig. 2). The breakpoints of 12 of these 14 cluster in three regions of Yq and one region of Yp—all regions rich in Y-specific repeats, which is consistent with this hypothesis (*11, 37*).

Interstitial deletions could result from recombination between direct repeats on the same arm of the chromosome. The abnormal Y in WHT1165 may have been produced by such a process. It has a breakpoint adjacent to the centromere and is missing all euchromatic sequences from Yq, but the chromosome retains heterochromatic repeats from distal Yq that, although reduced in copy number, are sufficiently abundant to be detected by in situ hybridization (*38*). A degenerate pentameric repeat array of sequences like that of satellite-3 lies adjacent to the centromeric breakpoint, and these sequences are similar to the *DYZ1* repeats found in the Yq heterochromatin (*11*). Homologous recombination between similar pericentric and distal Yq sequences may have created the interstitial deletion seen in WHT1165.

Two assumptions underlie the construction of the deletion map: (i) that the order of the 43 intervals is the same for each of the 97 chromosomes studied, and (ii) that each deleted chromosome has sustained a single break with loss of all DNA on one side and retention of all DNA on the other. To the extent that either of these assumptions is not valid, chromosomes with more than one breakpoint will appear on the map. Only 5 of the 97 chromosomes show evidence of more than one breakpoint, an indication that the assumptions, are, to a large extent, valid (*39*). Several explanations are possible for the five exceptional chromosomes. First, a chromosome may have sustained multiple breaks during a complex rearrangement. The ring Y of WHT1344 and the well-characterized deletion of WHT1013 (*40*) may be of this type. Second, a paternal Y chromosome with a preexisting interstitial deletion may have sustained a terminal deletion. Third, Y chromosomes with inversions or which are otherwise structurally variant may be present in the data set. Structural polymorphism has been hypothesized for the Y as a consequence of its restricted meiotic exchange and postulated reduced content of genes (*16, 41*). Let us suppose that one of the 97 chromosomes is a structural variant in which the order of several intervals is inverted, and that a single breakpoint falls within the inverted region. To fit data from this chromosome to a map based on the other 96 chromosomes, it would be necessary to hypothesize an interstitial deletion as well as a terminal deletion, two more breakpoints than actually present. The exceptional chromosome of WHT715 may be of this type. It displays three breakpoints and was previously hypothesized to derive from a structural variant with a paracentric inversion of Yp (*42*). If the order of intervals 4A-3C was inverted in the Y chromosome of the father, then a single break near interval 4A could account for the observed data. Such an inversion would juxtapose Y-specific repeats present in interval 3C with very similar repeats near the alphoid array in 4B (*11*). The observation that this chromosome carries an infrequent allele for a restriction fragment polymorphism in interval 6 is further evidence that it may be a rare variant (*43*). The overall paucity of apparent interstitial deletions in Fig. 2 suggests that gross structural polymorphism of the euchromatic Y chromosome is quite limited, at least among the largely Caucasian population tested here, in contrast to the findings of some other investigators (*41*). With the deletion and YAC maps as guides, comparison of the structure of the Y within human populations and among primates should provide further insights into the degree of structural polymorphism, the mechanisms of chromosome rearrangement, and, possibly, the evolution of the human species.

REFERENCES AND NOTES

1. L. Tiepolo and O. Zuffardi, *Hum. Genet.* **34**, 119 (1976).
2. P. A. Jacobs and A. Ross, *Nature* **210**, 352 (1966); E. M. Buhler, *Hum. Genet.* **55**, 145 (1980); R. M. Davis, *J. Med. Genet.* **18**, 161 (1981).
3. G. Vergnaud *et al.*, *Am. Hum. Genet.* **38**, 109 (1986).
4. N. A. Affara *et al.*, *Nucleic Acids Res.* **14**, 5353 (1986); U. Muller *et al.*, *ibid.*, p. 6489; C. J. Oosthuizen *et al.*, *Hum. Genet.* **85**, 205 (1990); Y. Nakahori *et al.*, *Genomics* **9**, 765 (1991); M. Kotecki, J. Jaruzelska, M. Skowronska, P. Fichna, *Hum. Genet.* **87**, 234 (1991); B. Bardoni *et al.*, *Genomics* **11**, 443 (1991); K. Ma *et al.*, *Hum. Mol. Genet.* **1**, 29 (1992).
5. A. H. Sinclair *et al.*, *Nature* **346**, 240 (1990).
6. E. M. C. Fisher *et al.*, *Cell* **63**, 1205 (1990).
7. E. Simpson *et al.*, *Nature* **326**, 876 (1987).
8. M. A. Cantrell *et al.*, *Genomics* **13**, 1255 (1992).
9. R. K. Saiki *et al.*, *Science* **230**, 1350 (1985); R. K. Saiki *et al.*, *ibid.* **239**, 487 (1988).
10. M. Olson, L. Hood, C. Cantor, D. Botstein, *ibid.* **245**, 1434 (1989).
11. S. Foote, D. Vollrath, A. Hilton, D. C. Page, *ibid.* **258**, 60 (1992).
12. E. D. Green and M. V. Olson, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1213 (1990); E. D. Green *et al.*, *Genomics* **11**, 548 (1991).
13. J. C. Fuscoe, L. M. Clark, D. M. Van, *Cytogenet. Cell. Genet.* **43**, 79 (1986).
14. Fragments were isolated either by PCR with primers bracketing the cloning site of the Charon 21A phage vector or by conventional subcloning into plasmid vectors and preparation of double-stranded templates by alkaline lysis. The templates generated by PCR were resolved on low–melting-point agarose gels, and the bands were excised, melted at 65°C, cooled to 37°C, and digested with Agarase (Sigma) for 30 minutes; the digestion products were used for sequencing. From 400 to 500 bp of sequence were obtained from template ends with the Applied Biosystems dye-primer cycle-sequencing protocol and an ABI 373A machine.
15. Sequences were compared to a consensus Alu element and a collection of L1 and transposable human elements (THE) with FastN [W. R. Pearson and D. J. Lipman, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2444 (1988)]. Sequences with initial match scores of 100 or more to any of these repeats were generally not pursued further. Sequences with scores below 100 were used to search Genbank by the BLASTN algorithm [S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990)] to detect similarities to less abundant interspersed repeats. Duplication was avoided because each new sequence was matched against all previous sequences.
16. B. Charlesworth, *Science* **251**, 1030 (1991). Alternatively, the data may reflect some systematic bias in our selection protocol. For example, all sequences were obtained from Hind III subclones.
17. S. Lincoln, M. Daly, E. Lander, personal communication. Adjustable program parameters were as follows: primer T_m, optimal 58°C, range 54° to 63°C; primer length, optimal 20, range 18 to 25; GC content 20 to 80 percent; product length 100 to 400 bp; salt concentration 50 mM. The repeat file contained THE, alphoid, and consensus Alu sequences. In a few cases, working primers were generated from sequences subsequently found to contain interspersed repeats.
18. Reactions were performed in a 10- or 20-μl volume in 1.5 mM $MgCl_2$, 5.0 mM NH_4Cl, 10 mM tris (pH 8.2 at 25°C), 100 mM KCl, 100 μM dNTP's, with 5 units of Taq DNA polymerase per 100 μl of reaction volume, 50 to 100 ng of human genomic DNA per 10 μl of reaction, and each primer at 1.0 μM. The cycling protocol was 1 minute at 94°C, 1 minute at 61°C, and 1 minute at 72°C for 30 cycles, except where otherwise specified. Most products were resolved on 8 percent polyacrylamide gels in 0.5X tris-borate EDTA buffer [J. Sambrook, E. F. Fritsch, T. Maniatis, Eds., *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989, p. 6.71)], but some were resolved on 4 percent agarose gels.
19. D. C. Page, M. E. Harper, J. Love, D. Botstein, *Nature* **311**, 119 (1984).
20. H. J. Cooke, W. A. R. Brown, G. A. Rappold, *ibid.*, p. 259.
21. ———, *ibid.* **317**, 687 (1985); M. C. Simmler *et al.*, *ibid.*, p. 692; M. Koenig, J. P. Moisan, R. Heilig, J.

L. Mandel, *Nucleic Acids Res.* **13**, 5485 (1985).
22. C. Bishop, G. Guellaen, D. Geldwerth, M. Fellous, J. Weissenbach, *J. Mol. Biol.* **173**, 403 (1984).
23. M. Marcus, R. Tantravahi, V. G. Dev, D. A. Miller, O. J. Miller, ***Nature*** **262**, 63 (1976).
24. DNA's from somatic cell hybrids 937 and 1006 were purchased from Bios Corp. (New Haven, CT).
25. C. J. Brown and H. F. Willard, *Am. J. Hum. Genet.* **46**, 273 (1990).
26. The hybrid lines were X3000.11 [R. L. Nussbaum, S. D. Airhart, D. H. Ledbetter, *Am. J. Med. Genet.* **23**, 457 (1986)], GM10063 and GM10095 from the Coriell Institute Repository, Camden, NJ, and one derived from patient 16 described by A. Ballabio *et al.* [*Proc. Natl. Acad. Sci. U.S.A.* **86**, 10001 (1989)].
27. Because the DNA samples were collected over a long period, much of the initial characterization was done by DNA hybridization methods. Many of the 300 individuals tested appeared either to lack Y DNA or to carry an intact Y chromosome. Some of the remaining 150 individuals were not extensively studied because the breakpoints fell among Yq heterochromatic repeats or in the Yp-Xq21 region, which is difficult to map from data obtained with the PCR assays described.
28. The correspondence between hybridization probes used to obtain data in Fig. 2 (and STS's generated from them): pDP307 (sY20), pDP522b (sY21), p47a (sY24), p41a (sY27), lambda Y215 (sY28), p17 (sY35), p7a (sY36), p16 (sY37), pDP61 (sY40), pDP1045 (sY43), lambda Y103 (sY45), pDP1057 (sY47), St25/2 (sY48), pDP1040 (sY51), p1 (sY52), and pY431-HinfA (sY159). Details of the restriction enzymes used and sizes of Y-specific restriction fragments scored are those described in (*44*) or in D. C. Page *et al.*, in preparation.
29. Of the 27 loci, four (sY27, sY37, sY45, and sY52) were scored by hybridization and PCR on the DNA's of up to 15 individuals. The two methods yielded identical results on the samples tested.
30. D. C. Page, *Development* **101** (suppl.), 151 (1987); A. Ballabio *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 10001 (1989); L. Alvesalo and A. de la Chapelle, *Ann. Hum. Genet.* **54**, 49 (1981); M. Andersson *et al.*, *Hum. Genet.* **79**, 2 (1988).
31. N. E. Morton, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7474 (1991).
32. F. Rouyer, M.-C. Simmler, D. C. Page, J. Weissenbach, *Cell* **51**, 417 (1987).
33. The 19 breakpoints bound one of five intervals (containing Xp-homologous STS's): 3D (sY66), 4A (sY70), 5E (sY182), 5F (sY151), and 5I (sY99).
34. P. H. Yen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8944 (1991).
35. The apparent frequencies of various chromosomal anomalies may be affected by the method of ascertainment of individuals harboring those anomalies and by their viability.
36. A possible role for sister chromatid exchange in the generation of isochromosomes has been suggested [C. B. Sharp, H. M. Bedford, H. F. Willard, *Hum. Genet.* **85**, 330 (1990)].
37. The 12 isodicentric chromosomes with breaks near repeats are FF, WHT2171, WHT2251, RS, WHT1820, DM, JC, WHT1318, WHT1434, WHT2064, WHT2240, and WHT1301.
38. J. Skare *et al.*, *Am. J. Med. Genet.* **36**, 394 (1990).
39. We assume that all chromosomes require one functional centromere and a telomere at each end for stable propagation. The ends of 71 of the 97 chromosomes are accounted for because they are either translocation products, isochromosomes, ring chromosomes, or the result of interstitial deletion. Lacking assays for Yqter, we cannot yet determine whether any of the 26 remaining Yq chromosomes are truly terminally deleted or are, instead, interstitially deleted.
40. D. C. Page, E. M. C. Fisher, B. McGillivray, L. G. Brown, *Nature* **346**, 279 (1990).
41. M. A. Ferguson-Smith, N. A. Affara, R. E. Magenis, *Development* **101** (suppl.), 41 (1987).
42. D. C. Page, *Cold Spring Harbor Symp. Quant. Biol.* **51**, 229 (1986).
43. C. M. Disteche *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7841 (1986).
44. J. Weissenbach and P. N. Goodfellow, *Cytogenet. Cell. Genet.* **58**, 967 (1991).
45. Of the 182 STS's, 14 derive from YAC ends as described (*11*). All primer sequences and PCR conditions have been deposited with the Genome Data Base.
46. J. Silver, A. Rabson, T. Bryan, R. Willey, M. A. Martin, *Mol. Cell. Biol.* **7**, 1559 (1987).
47. P. Vogt *et al.*, *Hum. Genet.* **86**, 341 (1991). The correspondence between STS's and DNA clones is as follows: sY112 (pHBS65), sY113 (pYD14), sY114 (pYBaH34), sY156 (pYHP52), sY85 (pYHP35).
48. J. A. Grifo *et al.*, *J. Am. Med. Assoc.* **268**, 727 (1992).
49. S. Guioli *et al.*, *Nature Genet.* **1**, 337 (1992); B. Incerti *et al.*, *ibid.*, in preparation.
50. N. A. Affara *et al.*, *Nucleic Acids Res.* **15**, 7325 (1987).
51. M. Munke *et al.*, *Hum. Genet.* **80**, 219 (1988).
52. C. M. Disteche *et al.*, *ibid.* **74**, 372 (1986).
53. E. Goulmy *et al.*, *Immunogenetics* **17**, 523 (1983).
54. A. J. O'Reilly *et al.*, in preparation.
55. D. S. Borgaonkar, in *The Y Chromosome, Part B: Clinical Aspects of Y Chromosome Abnormalities*, A. A. Sandberg, Ed. (Liss, New York, 1985), pp. 15–26.
56. N. Blagowidow, D. C. Page, D. Huff, M. T. Mennuti, *Am. J. Med. Genet.* **34**, 159 (1989).
57. We thank R. Alagappan, K. Bieker, D. Pettay, and J. Pollack for contributions to the work reported here; W. Gilbert and M. P. Reeve for help with sequence analysis; A. de la Chapelle for encouragement and contributions to this work; P. de Jong for one of the flow-sorted phage libraries and V. Nowotny for the composition of his PCR buffer; N. Affara, R. Bernstein, H. Bode, J. Bodurtha, M. Bofinger, W. R. Breg, P. Chandler, A. de la Chapelle, C. Disteche, A. Donnenfeld, B. Emanuel, M. Feingold, M. Ferguson-Smith, S. Gollin, M. Greally, R. Greenstein, E. Goulmy, S. Hahm, M. J. Hajianpour, K. Harper, J. Hersh, D. Hoar, L. Holmes, J. Hoo, I. A. Hughes, P. Jacobs, M. Jamehdor, L. Jenkins, R. McChane, P. McDonough, B. McGillivray, E. Magenis, M. Mennuti, W. Miller, A. Milunsky, P. Murphy, A. Newlin, F. Novoa, M. Phelan, J. Priest, S. Patil, I. Rosenthal, C. Rudlin, D. Salomon, L. Shapiro, S. Shekhter-Levin, M. H. K. Shokeir, J. Shulkin, J. Siegel, E. Simpson, J. Skare, U. Surti, M. Steele, P. Townes, M. Verma, J. Vigorita, N. Wang, L. Weiss, J. Weissenbach, S. Wenger, and D. Wilson for human DNA, blood, or cell line samples; J. Weissenbach, P. Vogt, K. Smith, and H. Donis-Keller for plasmids; and A. Ballabio, R. Nussbaum, and H. Willard for somatic cell hybrids. Supported by grants from the National Institutes of Health, the Searle Scholars Program/Chicago Community Trust, and the Harvard-MIT Health Sciences and Technology Johnson & Johnson Research Fund; a Smith Kline Beecham fellowship of the Life Sciences Research Foundation (D.V.); and a Markey Foundation fellowship (S.F.).

20 July 1992; accepted 20 August 1992