Genome Analysis and the Human X Chromosome

Jean-Louis Mandel,* Anthony P. Monaco, David L. Nelson, David Schlessinger, Huntington Willard

A unified genetic, physical, and functional map of the human X chromosome is being built through a concerted, international effort. About 40 percent of the 160 million base pairs of the X chromosome DNA have been cloned in overlapping, ordered contigs derived from yeast artificial chromosomes. This rapid progress toward a physical map is accelerating the identification of inherited disease genes, 26 of which are already cloned and more than 50 others regionally localized by linkage analysis. This article summarizes the mapping strategies now used and the impact of genome research on the understanding of X chromosome inactivation and X-linked diseases.

The X chromosome is the most extensively studied of all human chromosomes, a result of the wide interest in X-linked diseases and in the phenomenon of X chromosome inactivation. Both of these features are related to the different dosage of X in males and females.

Because males have only a single X, recessive diseases tend to be revealed; this accounts both for the large numbers of X-linked diseases (*1*) and for their characteristic inheritance pattern. The need to understand these diseases better and to provide diagnostic tools for genetic counselling in affected families was a major impetus for mapping studies. As a result of this driving force, the X chromosome was the first to have a genetic map based on restriction fragment length polymorphisms (RFLPs) and systematic approaches to physical coverage were undertaken that have been expanded as part of the Human Genome Initiative. Genes for two X-linked diseases (chronic granulomatous disease and Duchenne muscular dystrophy) were the first to be isolated by a mapping approach in the absence of functional information about the gene, by what is now called positional cloning (*2*). The isolation of the Duchenne muscular dystrophy gene (*DMD*) was an enormous undertaking; it is by far the largest gene known in any organism (2.4 Mb), contains introns as large as 100 to 200 kb, and may take 24 hours to be transcribed from end to end. More recently, the study of two other diseases (the fragile X mental retardation syndrome and spinobulbar muscular atrophy) uncovered a new and unexpected mutation mechanism, the expansion of trinucleotide repeats (*3*).

The second fascinating aspect of the human X chromosome (shared by its mammalian homologs) is the process of stable inactivation of one of the two Xs in females. Mapping approaches have led recently to the isolation of the probably non-protein-coding gene *XIST* which is expressed only from inactive X chromosomes and is likely to be important in inactivation (*4*).

The authors are presently the Genome Data Base editors for the X chromosome. J.-L. Mandel is at the Laboratoire de Genetique Moleculaire des Eucaryotes du CNRS, INSERM, 67085 Strasbourg Cedex France. A. P. Monaco is at The Human Genetics Laboratory, Imperial Cancer Research Fund, John Radcliffe Hospital, Headington, Oxford OX3 9DU. D. L. Nelson is at the Institute for Molecular Genetics, Baylor College of Medicine, Houston, TX 77030. D. Schlessinger is in the Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110. H. Willard is in the Department of Genetics, Case Western Reserve University, School of Medicine, Cleveland, OH 44106–4955.

*To whom communications should be addressed.

Developing a Physical Map of the X Chromosome

Physical maps of large chromosomal regions are defined by a series of DNA markers, preferably at closely and evenly distributed intervals. Such maps can be developed without cloning most of the chromosomal DNA—for example, by localizing sites for rare-cutting restriction enzymes (*5*). It is an advantage, however, to have cloned DNA in order to hunt for genes, study gene expression, or examine chromosome replication. Consequently, physical mappers usually aim to reassemble chromosome equivalents from purified DNA.

As the molecule of DNA that makes up the X chromosome is much too large to be handled intact (~160 Mb), it has been broken into cloned fragments that must be arranged by overlaps to achieve long-range contiguity. The collections of clones can be of various types, but it is intuitively obvious that larger clones make physical mapping easier. The development of yeast artificial chromosomes (YACs) as a cloning vector has promoted rapid progress in mapping in recent years (*6*). YACs provide isolated fragments of human DNA of up to a megabase or more, so that a few hundred would be sufficient to extend across a chromosome.

Two alternative (but often complementary) methods may be used to organize DNA fragments into a mapped region. In the first approach, which has been the primary one used for the X, chromosome-specific probes are used to screen YAC libraries for cognate clones. Many such probes have been characterized as a result of ongoing mapping activities related to the localization and identification of disease genes (*7*) and have been genetically or cytogenetically assigned to regions of the chromosome (*8*). Probes defining genetically mapped, polymorphic loci have been used to find corresponding larger YAC clones and provide markers that format the physical map. Historically, most probes have been obtained as clones (in plasmids, phage, or cosmids) derived from flow-sorted chromosomes (*9*) or genomic libraries constructed from somatic cell hybrids. More recently, additional methods have been introduced to recover region-specific probes; they include the polymerase chain reaction (PCR)–based amplification of microdissected fragments of individual chromosomes and amplification of segments flanked by human-specific interspersed, repetitive sequences present in hybrid cells and YACs (Alu or LINE sequences) (*10*). In the second approach, clones for all or part of a chromosome are systematically analyzed by fingerprinting techniques [for instance by sizing restriction fragments (*11*) or studying fragments that contain certain repetitive sequences (*12*)] and overlaps between clones are detected by computer analysis.

YAC libraries used to build the X physical map can be divided in two classes: total genomic libraries (*13*) or X chromosome–specific libraries constructed from appropriate somatic hybrids. The targeted libraries most intensively studied are for Xq24 to q28 (*14*) and pter to q27.3 (*15*). Chromosome-specific libraries have a smaller number of clones which favors screening with probes. Currently, the chromosome-specific libraries appear to contain a smaller percentage of chimeric clones (clones that contain segments from two regions of the genome that became artifactually linked) than total genome libraries. Although chimeric clones pose problems for contig building, constructing a chromosome-specific library with long DNA inserts and adequate representation is a tedious and difficult task. Some total libraries are of special value for the X map as they are based on cells from patients with four or five Xs, and have a 2- to 2.5-fold enrichment in X-specific clones. Other libraries have been useful because they have a large average insert size (*13*). YAC

Table 1. Positional cloning of X-linked diseases. T, D, and Y indicate whether translocations (T), large deletions (D), or YACs (Y) were crucial for localization and cloning of the gene.

Year cloned	Disease	Method	Reference
1986	Chronic granulomatous disease	D	*(2)*
1986	Duchenne muscular dystrophy	D, T	*(2)*
1990	Choroideremia	T, D	*(89)*
1991	Fragile X syndrome	Fragile site, Y	*(69)*
1991	Kallmann syndrome	D, Y	*(90)*
1991	Lowe syndrome	T, Y	*(91)*
1992	Norrie disease	D, Y	*(92)*

libraries that are accessible to many investigators are considered to be "reference libraries," as databases of characterized clones allow integration of mapping information from various sources (*9*, *16*, *17*).

Screening YAC libraries with specific probes can be performed either by hybridization of the radioactively labeled probe to a matrix array of the clones (*16*) or by PCR with a pair of oligonucleotides that produce a small (60 to 1000 bp) product characteristic of that unique site in the genome. Such a site is called a sequence tagged site (STS) (*17*). The hybridization approach is the most direct for probes that have not yet been sequenced. However, it may be subject to false positive or negative screening artifacts. Filters with a high density of ordered arrays have been made for the Imperial Cancer Research Fund (ICRF) library, which facilitates hybridization screening (*16*). Most total genome YAC libraries have now been formatted in pools of clones for efficient PCR screening (*17*). In fact the PCR-based method that relies on STSs has become a feature of the U.S. Human Genome Initiative. The goal of this program is the development of a map as a group of contigs of an average 2 Mb, aligned and oriented with a minimum of gaps and with an STS every 100 to 200 kb. In STS content mapping, overlapping YACs are fitted together by their common content of STSs (*18*).

Prior to being completely linked in a set of YACs, DNA segments can be compartmentalized by localizing them to mapping intervals of specified regions of a chromosome. Intervals can be defined through use of hybrid cells containing segments of the chromosome, by deletions or translocation breakpoints found in patients with X-linked disorders (*19*), by in situ hybridization of probes to chromosomes in metaphase (*20*), or by radiation hybrid mapping (*21*). The preassignment of clones to intervals helps to reduce the effort required to infer overlaps among them.

Interval mapping is especially valuable in providing early localization of otherwise random clones (or in targeting probes to regions of particular interest). Many groups have concentrated on targeted mapping of regions around some disease loci or the X inactivation center in Xq13, and this has led to the isolation of many of the disease genes listed in Table 1. However, such approaches are unlikely to give coverage for the whole X, and strategies for long-range mapping are thus indispensible.

Long-Range Physical Mapping

The first attempt at long-range mapping was directed towards the Xq24–q28 region by means of a targeted YAC library (*14*). The library was organized as a matrix array of 800 clones, and the strategy was based almost entirely on hybridization of probes (a large number of these were available for Xq26–q28 from groups mapping the fragile X syndrome or diseases in Xq28). For closing gaps, additional PCR screening of general libraries and isolation of YAC ends for walking steps were performed (*22*). For the Xq24–q25 region, where few mapped probes were available, in situ hybridization of YACs provided useful information (*23*). As a result, the distal half of the long arm of the X chromosome is now organized in 21 contigs, the largest of 8.5 Mb around *HPRT* and *F9* (*8*, *22*). As these contigs cover a total of 45 Mb, the remaining gaps are expected to be small. In Xq26–q28, where genetic and somatic cell hybrid maps had been extensively developed, contigs could be easily ordered and oriented. In contrast, in Xq24–q25, most contigs have not been ordered completely.

Two major efforts are now under way to complete the physical coverage of the X map. An STS content mapping approach is used by David Schlessinger and colleagues, starting from seed YACs (which contain selected genes or anonymous STSs or are taken at random from the Xpter–q27.3–specific library). PCR primer pairs are developed for the ends of their human DNA inserts to screen for overlapping clones that extend the map locally, until maps of subregions meet. A strategy based on ligation-mediated PCR followed by automatic sequencing, allows efficient development of end STSs (about 30 per week) (*24*). The STS-based approach avoids the necessity of accumulating large numbers of probes from many laboratories to test batteries of clones. Furthermore, integration of a functional map of expressed genes with the physical map is possible (as cloned genes assigned to the X are sequenced and easily converted to STSs). The genetic map can also be incorporated as it is increasingly based on highly polymorphic microsatellites (small, tandem, di- or trinucleotide repeats, usually of CA) which are already in STS format. Such integration also provides material rapidly to groups interested in a particular region. The most extensive example of pure STS content mapping thus far has been provided by the studies of Page and coworkers (*25*) who have assembled 30 Mb of the Y chromosome into two contigs, including regions of the Y short arm homologous to extensive regions of the X chromosome (about 2 Mb in the pseudoautosomal region and 10 to 20 Mb in Xq21) (*7*, *22*). An alternative hybridization-based strategy (used by Hans Lehrach, Anthony Monaco, and their colleagues) is based on the high density filters of the ICRF library (*16*); it can be used with single copy or complex, region-specific probes (*16*). This has been very successful for the Xp21.3–q21 region due to the availability of large numbers of very well mapped probes from the region (*26*). Large contigs have already been developed around the 2.4-Mb dystrophin gene, which allowed mapping of its 79 exons and reconstruction of the complete gene (except for exon 60) in a 2.6-Mb YAC (*27*).

Somatic Cell Genetics

The mapping of the X chromosome and isolation of disease genes has been advanced by application of somatic cell genetic approaches that take advantage of two aspects of X chromosome genetics: the availability of X-linked selectable genetic markers that ensure retention of portions of the X in human/rodent somatic cell hybrids and the availability of a large number of structurally abnormal X chromosomes from patients with cytogenetic abnormalities. Unlike mapping of most autosomes in somatic cell hybrids, which relies in many instances on the fortuitous retention of the relevant chromosome, mapping to the X has been greatly facilitated by both positive and negative selection for the hypoxanthine-guanine phosphoribosyltransferase (HPRT) locus, which maps to Xq26. When normal, $HPRT^+$ human cells are fused with $HPRT^-$-deficient rodent cells, there is direct selection for retention of the human HPRT locus (and, therefore, the X chromosome) in the resulting somatic cell hybrid (*28*). More recently, a second X-linked selectable locus, the ubiquitin-activating

enzyme E1 (UBE1) locus in Xp11.2, has been used to complement a temperature-sensitive mouse mutant (*29*). Thus, selection for growth of human/mouse hybrids at the nonpermissive temperature results in retention of the short arm portion of the X chromosome.

Because of the functional hemizygosity of the X chromosome, many translocations between X and an autosome as well as other structural abnormalities (such as deletions, duplications, and isochromosomes) are detected clinically. By means of the selectable markers described above, such rearranged chromosomes have been isolated in somatic cell hybrids and have provided a rich resource for interval mapping, especially in the pericentromeric region and in the middle and distal long arm regions (*19*). In some instances, the density of available breakpoints is such that resolution of physical mapping is in the range of a few megabases (*8*), allowing the unambiguous ordering of markers that could not be definitively ordered by other techniques.

Newer somatic cell genetic techniques have been reported that should increase the resolution of physical mapping based solely on naturally occurring abnormal chromosomes. These approaches depend on introduction of experimentally induced chromosome breaks, the subsequent isolation or "rescue" of broken chromosomes in somatic cell hybrids, and the use of such hybrids for mapping DNA markers. The original development of this approach was by Goss and Harris (*30*), who used low-dosage irradiation to introduce breaks between genetic markers. While these original hybrids found some use among physical mappers (*30, 31*), they have been effectively replaced by hybrids retaining even smaller DNA fragments generated by irradiation, chromosome transfer, and fragmentation (*21*) or by simultaneous positive and negative counterselection of multiple selectable X-linked markers (*32*). An elegant application of this approach has utilized telomere-directed chromosome breakage, in which human telomeric DNA has been introduced into a human/rodent somatic cell hybrid carrying the X chromosome. In a proportion of clones, the cloned telomeric DNA induced formation of a new telomere interstitially on the X, resulting in terminal deletion of the chromosome (*33*). A panel of such terminally deleted chromosomes would be enormously useful for physical mapping and would complement existing panels of hybrids based on clinical material.

The Genetic Map

The X chromosome was the first one for which an RFLP-based map was produced, by linkage analysis in a defined panel of normal families with large sibships (*34*). This strategy became the standard for building a genetic map of the human genome as an international collaboration under the auspices of the Centre d'Étude du Polymorphisme Humain (CEPH) in Paris. In 1985 the map contained 20 RFLP markers, with an estimated genetic length of 185 centimorgans (cM) (*34*). The current length is estimated at 200 cM (*8*). The map provided an extremely useful starting tool for genetic mapping of a rapidly increasing number of X-linked diseases. In turn, the interest in such diseases led to a steady improvement of the X genetic map through the building of regional maps of increased density and accuracy (*35*). Mapping based on the early RFLP array was not optimal, since most RFLP markers were only diallelic and not very informative and few probes were available that were highly polymorphic, multiallelic markers (*36*). This has changed with the use of microsatellites (*37*), which can be rapidly tested by PCR; about 70 such markers have already been characterized and are being placed on the genetic map (*38*). At present the CEPH database contains genotypic information on about 160 markers (~130 loci). The largest systematic effort over the whole chromosome resulted in the characterization and genetic mapping of 55 RFLPs and a preliminary map including 111 markers over a distance of 200 cM was presented by Fain and Barker at the third X chromosome workshop in April 1992 (*39*). This map contained only three gaps greater than 10 cM. The map presented in this issue of *Science* is a composite of the Fain and Barker map and of published regional maps; it includes many of the most widely used markers.

A striking feature of the genetic map is the tenfold difference in recombination frequency between males and females for the 2.5-Mb pseudoautosomal region shared between the tips of the short arms of the X and Y chromosomes. Polymorphic markers in this region show a gradient of sex linkage from the pseudoautosomal boundary (*PABX*), which marks the beginning of X-specific sequences, to markers near the telomere (DXYS14) that show 50% recombination in males between the X and Y chromosomes. A 50% recombination frequency indicates that one crossover occurs at each meiosis in this region, which is probably the most highly recombinogenic in the human genome (*40*). Recombination in this region between the two female X chromosomes is at a frequency of only 5 to 6%, a value close to that predicted from the physical length (*40*). The pseudoautosomal region is particularly rich in highly polymorphic minisatellite markers (which have longer repeat units than microsatellites) and it remains to be seen whether these tandem repeat sequences are involved in the high male recombination rate. On autosomes, minisatellites appear also to cluster near telomeres, in regions with high recombination rates (*41*). Increased recombination was also found within the dystrophin gene (*DMD*) (*42*) and in regions flanking the fragile X site in Xq27.3 (*FRAXA*), although the fragile locus itself is not a hot spot of recombination (*43*). The pericentromeric region shows less recombination than expected from its size (*8, 35*) perhaps reflecting a centromere effect on recombination frequency.

Mapping X-Linked Diseases

At least 16 X-linked disease genes have been cloned on the basis of prior knowledge of the defective protein (*8*). However, for most of the remaining diseases, the biochemical defect was unknown or very uncertain and it was thus necessary to use mapping strategies to identify the corresponding genes. Genes for seven X-linked diseases have been isolated by positional cloning (Table 1), including those for Duchenne muscular dystrophy and the fragile X mental retardation syndrome, the two most common, severe, X-linked diseases. In several other cases, mapping studies were useful in selecting an independently cloned "candidate" gene (a gene coding for a protein whose function or localization fits with the disease manifestations), which could then be validated by finding gene-specific mutations in patients. Linkage studies in affected families are being actively pursued in many laboratories for more than 50 diseases (*7*). Increasingly, microsatellite markers are used for more efficient and precise mapping, and meanwhile provide diagnostic assays for genetic counseling.

Fortunately, X-linked diseases have features that facilitate positional cloning. Chromosomal assignment, which can be an arduous task for rare autosomal diseases, is obvious because of the inheritance pattern. For about ten diseases, rare affected females have been found with balanced X-autosome translocations. In these patients, the normal X chromosome is generally inactive, and the translocated X active, because of a selection process operating in early embryogenesis against cells carrying an inactive translocated X. Translocations that have a breakpoint within a gene will lead to expression of the corresponding disease, as the uninterrupted copy on the normal X is inactive. Such translocations have provided precise localization for the relevant disease genes that could be confirmed by linkage analysis in affected families and have been instrumental in the cloning of several genes (Table 1). This approach is now being used for anhidrotic ectodermal dys-

plasia (*EDA*) and Menkes disease (*MNK*) (*44*).

A puzzling case is represented by incontinentia pigmenti, a rare disorder with developmental neuroectodermal abnormalities. It affects only females and is assumed to be an early lethal in affected male fetuses. Several females with different de novo X-autosome translocations have been described, with breakpoints in Xp11.1 or Xp11.21. Some of these breakpoints appear to be separated by at least 1250 kb (*45*), suggesting that the corresponding gene, *IP1*, may be very large. Such a putative large gene should be a target for other types of mutations. However a familial form of the disease has been localized to a different region by genetic mapping, Xq28 (defining locus *IP2*) (*46*).

Males have only one X chromosome, and it was thus surprising to find rare male patients who survive with deletions (often detected cytogenetically) encompassing several megabases of DNA. In most cases the lack of function for genes in the deleted region results in a contiguous gene syndrome. In other words, several diseases are associated in a single patient, allowing very accurate mapping of the relevant genes if a series of overlapping deletions are available for analysis (*47*). This was first observed in the case of the BB deletion encompassing part of *DMD* and genes for chronic granulomatous disease (*CYBB*), McLeod syndrome (*XK*), and retinitis pigmentosa-3 (*RP3*) (*48*). The BB deletion was instrumental for the cloning in 1986 of *DMD* and *CYBB*, and now is being used to clone *XK* and *RP3* (*49*). Large deletions have been found in the Xp22.3, Xp21.2, Xq21.2, Xq25, and Xq27 regions and are illustrated in the chart in this issue of *Science*. Very likely such regions have relatively low gene densities, or the deletions would be lethal (*50*). In the Xp21.2 region this can be accounted for in part by the huge size of *DMD*, but more generally it illustrates the great heterogeneity in gene density throughout the genome, with gene-poor regions (in general AT-rich, Giemsa dark bands) alternating with gene-rich ones [such as the distal part of Xq28 where many genes are clustered (*7*, *51*)].

For X-linked diseases that severely decrease reproductive fitness in affected males, unrelated patients can be expected to carry independent mutations (this is not the case for autosomal recessive diseases), and a proportion of these may be small deletions. When a disease locus has been mapped within about a megabase, it can thus be efficient to screen a panel of unrelated patients for such deletions with probes from the region, as this may pinpoint the exact location of the gene. This strategy was first used for Duchenne muscular dystrophy (*52*).

The high proportion of families whose X-linked diseases are a result of new mutations impairs, however, precise localization by linkage analysis. Even with a very dense genetic map composed of highly polymorphic markers, most rare diseases will not be mapped to intervals smaller than 2 to 5 cM (on the order of 2 to 5 Mb and containing up to 100 genes) (*53*). For some milder or late onset diseases that can be transmitted through many generations, precise localization may be facilitated by the study of specific association (linkage disequilibrium) between the disease and alleles at marker loci in homogeneous populations susceptible of showing a founder effect for the disease (*54*). In the more general case, and in the absence of translocations or deletions large enough to be easily detected, isolation of genes for rare X-linked diseases will be difficult, even if YAC contigs are available for the region.

An even more challenging problem is posed by the heterogeneous causes of X-linked mental retardation. The fragile X syndrome, which affects 1 in 1500 males, may account for only 30 to 50% of X-linked mental retardation. Rare syndromes have been described in which mental retardation is accompanied by specific clinical features, but in many families such features are absent or nonspecific (*55*). In the latter cases, one cannot pool the linkage results obtained in different families, as they may not have a defect in the same gene and this severely impairs genetic mapping. Studies of some large families have indeed indicated that genes at a number of different positions along the X can cause similar "nonspecific" mental retardation (*55*). Contiguous gene syndromes in patients with large deletions (such as in Xp22.3 and Xq21.3) may help to isolate some of these mental retardation genes (*47*, *56*).

Another medically important, but very difficult mapping problem is posed by Rett syndrome. This severe neurological disorder affects 1 in 10,000 live-born females and is thought (but not proven) to be an X-linked dominant trait with lethality in males. As affected females do not reproduce, the extremely rare families with two affected half-sisters (from different fathers) are thought to be due to maternal germinal mosaïcism. These are being now used for exclusion mapping of the putative X-linked mutations (*57*). In exclusion mapping, the gene may lie only in those regions where the two affected daughters share the same maternal alleles. Unfortunately, X-autosome translocations found in two females with a Rett phenotype gave inconsistent mapping information.

As the Human Genome Initiative advances, the production of a catalog of genes along the X chromosome, and the characterization of tissue-specific patterns of expression, may provide a general route to the selection of candidate genes. It will then be critical to screen these candidates for mutations at the DNA or RNA level (*58*).

Mutations in X-Linked Disease Genes

The spectrum of mutations and incidence of severe X-linked diseases is due to the constant input of new mutations and directly reflects the mutational sensitivity of the gene, with little influence of population genetic parameters such as selection, drift, or founder effect. This is because in cases of severe disease, affected males do not reproduce. Therefore, at each generation the number of mutations decreases by one-third (since males have one-third of the X chromosomes in the population) and the particular mutation becomes extinct after a few generations. This does not apply to mutations with mild or no effect on reproductive fitness such as those responsible for glucose-6-phosphate dehydrogenase (G6PD) deficiency, color blindness, or some cases of mild hemophilia A or B. Systematic analyses of mutations in unrelated patients provide information on the mechanisms of mutations in different regions of the genome, including parameters such as parental origin of mutations or the effect of paternal age. Such studies are often of great importance for diagnostic applications and will benefit from the complete sequencing of the genes and of their neighboring regions. It is already clear that there is a striking difference in deletion frequency for various diseases. In X-linked ichthyosis, 80 to 90% of the mutations are large deletions encompassing the entire gene. Many of these are 1.9-Mb deletions that result from unequal recombination between flanking low-copy repetitive elements (*59*). Duchenne muscular dystrophy is another disease with a high frequency of partial deletions (60 to 70%) and a significant level of partial duplications (6 to 7%) (*60*). In part, this may be due to the huge target size of the dystrophin gene (2.4 Mb), but there is a clear clustering of deletion breakpoints in two regions, which poses the problem of the mechanism of this selective instability (*61*). For most other diseases analyzed thoroughly (notably hemophilia A and B, as well as ornithine transcarbamylase and HPRT deficiencies) the frequency of deletions or other rearrangements detectable by Southern blot is on the order of 5 to 15%.

Databases of point mutations have been established for hemophilia A and B (*62*, *63*). Because of the smaller size of the coagulation factor IX gene (which has been totally sequenced), hemophilia B has been more thoroughly analyzed, with about 400

mutations reported (62). A change from a CpG dinucleotide to TpG or CpA accounts for 33% of these mutations, a 20-fold increase in mutation rate with respect to the average of other nucleotide changes (64) and very similar observations were made for hemophilia A (63, 65). In the latter case, thorough analysis of all exons by PCR and denaturing gradient gel electrophoresis allowed the identification of 89% of the mutations in patients with mild to moderate hemophilia A (mostly missense mutations), but only 53% of the mutations in patients with a severe form, who lack factor VIII activity (65). Thus many null mutations may occur in introns or flanking sequences, (regions that have not yet been sequenced, as the gene is very large). Two hemophilia A mutations are due to insertion of a LINE1 interspersed repeat element. This LINE element was used to identify the parental sequence, which encodes a functioning reverse transcriptase (66). A case of hemophilia B resulting from de novo insertion of an Alu repeat has also been reported (67). These observations indicate that transposable elements can cause mutations in humans, although it is a rare occurrence. For G6PD, only missense point mutations have been found, with the exception of one single codon deletion, and it is likely that mutations which totally abolish G6PD activity are early lethals (68).

Expansion of a trinucleotide repeat sequence is a newly discovered disease-causing mechanism (3) originally described as the sole type of mutation in two X-linked diseases. The fragile X mental retardation syndrome is caused by an astonishingly unstable expansion of a CGG repeat in a 5′ exon of the gene *FMR-1*, which is correlated in patients with abnormal methylation of the adjacent CpG island (69). In spinobulbar muscular atrophy, the mutation is a more moderate expansion of a CAG repeat in the NH_2-terminal coding region of the androgen receptor gene (*AR*) (70) while other heterogeneous mutations in *AR* result in the completely different phenotype of testicular feminization. The identification of the mutation has had a profound impact on both the understanding of the peculiar inheritance of the fragile X syndrome, and on its diagnosis (69). Trinucleotide expansion has been found more recently in myotonic dystrophy, an autosomal dominant disease (3).

Mapping of Genes Controlling or Escaping X Inactivation

The stable inactivation of one of the two X chromosomes in females is a fascinating phenomenon as it can spread over 100 Mb of DNA, turning off several thousands of genes. It is thought to occur in three steps: initiation early in embryogenesis at a site called the X inactivation center, propagation along the length of the chromosome, and stabilization of the inactive state of individual loci (71, 72). (Stabilization is a term used to indicate that once the chromosome is inactivated, all descendants of that chromosome will be inactivated.) Genome mapping has already revealed some important and surprising features of this process.

The study of structurally abnormal X chromosomes in humans and mouse gave strong evidence for the existence of a cis acting locus, the X inactivation center (*XIC*), necessary for initiation of inactivation (71, 72). This locus was mapped recently in both species to a region between the ectodermal dysplasia locus (*EDA*, or *Tabby* in mouse) and the phosphoglycerate kinase gene (*PGK*) (73) a region of 2.5 Mb in humans now largely covered by a YAC contig (26). A gene cloned by serendipity was found to map to this region that is transcribed only from inactive X chromosomes (4, 74). It is also expressed in males from the single X that is inactivated during spermatogenesis. The portion characterized does not appear to code for a protein, and extensive alternative splicing occurs in humans but not in mouse (4, 72, 74). The localization and expression of this gene, named *XIST* (for X inactive–specific transcripts), suggests that it is involved in initiation or spreading of inactivation. Recently, studies of two deletions in females showing nonrandom inactivation have suggested that a region near the *FRAXA-IDS* loci (in Xq27.3–proximal q28 where proximal refers to the side closest to the centromere) may be involved in inactivation of distal loci (75). A second surprise was the discovery that several genes along the X appear to escape X inactivation (72). Although this feature was expected for genes in the pseudoautosomal region (which exist in two copies in males), it was also found for at least three genes in three other regions (*ZFX*, *UBE1*, and *RPS4X* in Xp22.1, Xp11.23, and Xq13.1, respectively). These genes are subject to inactivation in mouse. Abnormal dosage of such genes may be involved in the embryonic lethality and clinical phenotype of X0 females (Turner syndrome) in humans, whereas X0 mice are practically normal (76).

Stabilization of inactivation is very well correlated with DNA methylation (at the cytosine of a CpG dinucleotide), which occurs at CpG-rich regions (CpG islands) near the 5′ end of many genes (77). This feature is useful for mapping by pulsed-field gel electrophoresis (as the restriction enzymes used are sensitive to methylation), and as a way to identify expressed genes in cloned regions (5, 51). Analysis of such differential DNA methylation near some polymorphic X loci is also used for the study of clonality in tumors and of biased inactivation in female carriers of some X-linked diseases (78).

Of Mice, Humans, and Kangaroos

It was postulated originally by Ohno (79) that genes which are X-linked in one mammalian species should be X-linked in all others. The rationale for this was that as these genes function at a single dose, while those on autosomes function at double dose, exchange of a large chromosome segment from autosomes to X would result in impaired dosage and would be expected to be lethal. The mouse X chromosome is being actively mapped, and because of the conservation of protein-coding DNA sequences between mouse and human, the same genes or conserved probes can be mapped in parallel in the two species (80). Indeed, almost all genes tested that are X-linked in humans have their counterpart on the mouse X chromosome and five large regions can be defined that appear to have conserved gene content and order between the human and mouse Xs (80). Thus, relatively few intrachromosomal rearrangements have occurred on X since divergence of the two species. The correspondence established between the maps of the two chromosomes has allowed the validation of mouse mutants as homologous models for human diseases with similar phenotypes, based on their equivalent map positions, and may be useful for positional cloning of such loci (44, 81). For example, the *Hyp*, *Ta*, and *Mo* mutants correspond to the genes for hypophosphatemic rickets (*HYP*), ectodermal dysplasia (*EDA*), and Menkes disease (*MNK*), respectively in humans.

An interesting exception to the conservation of genes is the apparent absence in mouse of sequences homologous to the steroid sulfatase, Kallmann syndrome (*KAL*), *MIC2*, and *GS1* genes, four genes that are closely located in Xp22.3 in humans (72). Deletion of this region in males results in a relatively mild phenotype, and it may have been similarly deleted in a mouse ancestor. Another surprise was the finding that all genes tested that are located on the short arm of the human X are autosomal in marsupials and monotremes. This is at variance with Ohno's hypothesis, and suggests that the short arm was of autosomal origin and was added to the X chromosome in eutherian mammals (82).

The Future of the X Chromosome

Overall, about 40% of the X has now been recovered in YAC contigs (8). It can be forecast that the majority of the chromo-

some will be represented in cloned DNA of known location within the next 18 to 24 months. Although the physical map of the chromosome will approach completion, much additional effort will be required to close gaps and validate both the map and the clones representing it. In many respects, this will be a continual process, as alternate cloning vehicles and methods are developed. Assembly of such large regions in cloned DNA also permits the initiation of more refined structural studies, including the placement of repetitive sequences, the mapping of expressed genes, and evolutionary comparisons of human DNA to other species. For detailed analysis, YACs can be subcloned into phage or cosmid vectors or used as probes to recover cosmids from high-density arrays of genomic or chromosome-specific libraries (*8, 16, 83*). These will promote vertical integration from contigs toward sequence, and will provide reagents for studying disease-associated rearrangements.

For functional analysis, a number of techniques that rely on the cloned contigs are increasingly in use to identify expressed sequences. Whole YAC screening of cDNA libraries, cDNA selection strategies involving hybridization of PCR-amplified cDNA inserts to DNA from the contig, and exon amplification procedures show significant promise (*84*). The isolation of a significant fraction of the genes present in distal Xq28, a region particularly rich in disease loci has been recently reported (*85*). In a complementary approach, systematic sequencing of cDNA clones provide expressed sequence tags, or ESTs (*86*) that may be placed on the YAC contigs, thus providing accurate localization of the corresponding gene. The requirement of searching through entire genomes to find those ESTs in a defined region of interest may limit its present utility for positional cloning projects, and the development of efficient strategies to assign large numbers of such ESTs to YAC contigs is still needed. In addition to providing candidates for genes conferring phenotypes, such studies will begin to answer basic questions regarding the density and arrangement of genes in human chromatin. The isolation of disease genes will be accelerated and, as shown by those loci recently identified, will likely provide significant new insights into basic biochemical, genetic, and developmental processes in addition to their impact on the diagnosis, treatment, and prevention of the diseases.

The use of probes or STSs sufficiently conserved during evolution may permit researchers to recover equivalent regions of many genomes across phyla (*87*). Thus, the preservation of the fossil record in genome sequences from a variety of organisms can provide a method of formatting and comparing maps for all of them. Many evolutionary events will likely be revealed from the reconstructed maps.

Until now, genomic sequencing on the X has been limited to a small number of genes (*8*). With the advent of computer algorithms capable of identifying candidate-coding sequences in genomic DNA (*88*), large-scale sequencing projects become more attractive. They will probably first involve the disease-rich region in Xq28 or the deletion-prone regions in the enormous gene *DMD* whose large introns may contain unsuspected genes. Sequencing may also be used in the study of recombination, for instance in the pseudoautosomal region, or in the analysis of control elements involved in the various phases of X inactivation. The yeast and *Caenorhabditis elegans* large-scale sequencing projects have uncovered a higher than expected gene density, and this should further encourage similar projects on the X chromosome.

REFERENCES AND NOTES

1. V. A. McKusick, *Mendelian Inheritance in Man* (Johns Hopkins Univ. Press, Baltimore/London, ed. 10, 1992).
2. F. S. Collins, *Nat. Genet.* **1**, 3 (1992); Duchenne muscular dystrophy: A. P. Monaco *et al.*, *Nature* **323**, 646 (1986); chronic granulomatous disease: B. Royer-Pokora *et al.*, *ibid.* **322**, 32 (1986).
3. C. T. Caskey, A. Pizzuti, Y. H. Fu, R. G. Fenwick, D. L. Nelson, *Science* **256**, 784 (1992); R. I. Richards and G. Sutherland, *Nat. Genet.* **1**, 7 (1992).
4. C. J. Brown *et al.*, *Nature* **349**, 38 (1991); C. J. Brown *et al.*, *Cell*, in press; N. Brockdorff *et al.*, *ibid.*, in press.
5. D. P. Barlow and H. Lehrach, *Trends Genet.* **3**, 167 (1987); on the X, such maps have been developed over 10 Mb for the Xp22.3 and Xq28 regions; C. Petit, J. Levilliers, J. Weissenbach, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3680 (1990); A. Poustka *et al.*, *ibid.* **88**, 8302 (1991).
6. D. T. Burke, G. F. Carle, M. V. Olson, *Science* **236**, 806 (1987).
7. K. E. Davies, J. L. Mandel, A. P. Monaco, R. L. Nussbaum, H. F. Willard, *Cytogenet. Cell Genet.* **58**, 853 (1992).
8. The wall chart enclosed with this issue of *Science* has a visual display of the state of X chromosome mapping. Additional copies are available directly from *Science*.
9. An X-specific cosmid library gridded at high density has been made widely available for screening; it is especially useful for identification and characterization of genes, and to search for highly polymorphic genetic markers; D. Nizetic *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 3233 (1991).
10. H. J. Ludecke, G. Senger, U. Claussen, B. Horsthemke, *Nature* **338**, 348 (1989); M. C. Hirst *et al.*, *Genomics* **10**, 243 (1991); S. Hadano *et al.*, *ibid.* **11**, 364 (1991); D. L. Nelson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6686 (1989); C. G. Cole, P. N. Goodfellow, M. Bobrow, D. R. Bentley, *Genomics* **10**, 816 (1991).
11. A. Coulson, J. Sulston, S. Brenner, J. Karn, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7821 (1986); M. V. Olson *et al.*, *ibid.*, p. 7826.
12. G. Cangiano, H. Ameer, R. Waterston, A. Lavolpe, *Nucleic Acids Res.* **18**, 5077 (1990); I. Zucchi and D. Schlessinger, *Genomics* **12**, 264 (1992); this approach appears applicable to physical mapping of the whole genome starting from a total human YAC library if fingerprinting and especially computer analysis are scaled accordingly; D. Cohen *et al.*, Cold Spring Harbor Meeting on Genome Mapping and Sequencing, 32 (abstr.) (1992).
13. B. H. Brownstein *et al.*, *Science* **244**, 1348 (1989); R. Anand, J. H. Riley, R. Butler, J. C. Smith, A. F. Markham, *Nucleic Acids Res.* **18**, 1951 (1990); H. M. Albertsen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4256 (1990); Z. Larin, A. P. Monaco, H. Lehrach, *ibid.* **88**, 4123 (1991).
14. F. E. Abidi, M. Wada, R. D. Little, D. Schlessinger, *Genomics* **7**, 363 (1990); M. Wada *et al.*, *Am. J. Hum. Genet.* **46**, 95 (1990).
15. J. T. Lee, A. Murgia, D. M. Sosnoski, I. M. Olivos, R. L. Nussbaum, *Genomics* **12**, 526 (1992).
16. H. Lehrach *et al.*, in *Genome Analysis*, K. E. Davies and S. M. Tilghman, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1990), vol. 1, p. 39; A. P. Monaco *et al.*, *Nucleic Acids Res.* **19**, 3315 (1991); M. T. Ross *et al.*, *Nat. Genet.* **1**, 284 (1992).
17. M. V. Olson, L. Hood, C. Cantor, D. Botstein, *Science* **245**, 1434 (1989); E. D. Green and M. V. Olson, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1213 (1990).
18. E. Green and P. Green, *PCR Methods Appl.* **1**, 77 (1991).
19. F. P. Cremers *et al.*, *Am. J. Hum. Genet.* **43**, 452 (1988); G. K. Suthers *et al.*, *ibid.* **47**, 187 (1990); R. G. Lafreniere *et al.*, *Genomics* **11**, 352 (1991); D. L. Nelson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6157 (1991); F. Rousseau *et al.*, *Am. J. Hum. Genet.* **48**, 108 (1991).
20. P. Lichter *et al.*, *Science* **247**, 64 (1990).
21. D. R. Cox, M. Burmeister, E. R. Price, S. Kim, R. M. Myers, *ibid.* **250**, 245 (1990); F. Benham *et al.*, *Genomics* **4**, 509 (1989).
22. R. D. Little, G. Pilia, S. Johnson, M. D'Urso, D. Schlessinger, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 177 (1992); D. Schlessinger *et al.*, *Genomics* **11**, 783 (1991).
23. V. Montanaro *et al.*, *Am. J. Hum. Genet.* **48**, 183 (1991).
24. P. R. Mueller and B. Wold, *Science* **246**, 780 (1989); M. J. Palazzolo, S. A. Sqwyer, C. H. Martin, D. A. Smoller, D. L. Hartl, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8034 (1991); J. Kere *et al.*, *Genomics*, in press.
25. D. Vollrath *et al.*, *Science* **258**, 52 (1992); S. Foote *et al.*, *ibid.*, p. 60.
26. R. G. Lafreniere *et al.*, Cold Spring Harbor Meeting on Genome Mapping and Sequencing, 63 (abstr.) (1992); M. Ho *et al.*, *Cytogenet. Cell. Genet.* **58**, 2067 (1991); A. P. Monaco *et al.*, unpublished data.
27. A. P. Monaco, A. P. Walker, I. Millwood, Z. Larin, H. Lehrach, *Genomics* **12**, 465 (1992); A. J. Coffey *et al.*, *ibid.*, p. 474; R. G. Roberts, A. J. Coffey, M. Bobrow, D. R. Bentley, *ibid.* **13**, 942 (1992); J. T. Den Dunnen *et al.*, *Hum. Mol. Genet.* **1**, 19 (1992).
28. J. W. Littlefield, *Science* **145**, 709 (1964).
29. C. J. Brown, V. E. Powers, D. M. Monroe, R. Sheinin, H. F. Willard, *Somatic Cell Mol. Genet.* **15**, 173 (1989). As the UBE1 gene escapes inactivation, it can serve as selectable marker for an inactive X chromosome.
30. S. Goss and H. Harris, *Nature* **255**, 680 (1975).
31. H. F. Willard, S. J. Goss, M. T. Holmes, D. L. Munroe, *Hum. Genet.* **71**, 138 (1985).
32. C. J. Brown, A. Vivona, S. Parikh, M. Bedford, H. F. Willard, *Cytogenet. Cell Genet.* **51**, 970 (1989); S. T. Warren *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3856 (1990).
33. C. Farr, J. Fantes, P. Goodfellow, H. Cooke, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7006 (1991).
34. D. Drayna and R. White, *Science* **230**, 753 (1985); D. Drayna *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2836 (1984).
35. D. F. Barker *et al.*, *Hum. Genet.* **88**, 189 (1991); G. K. Suthers *et al.*, *Genomics* **9**, 37 (1991); T. Alitalo *et al.*, *Hum. Genet.* **86**, 599 (1991); M. M. Mahtani *et al.*, *Genomics* **10**, 849 (1991).
36. The only ones available before 1990 were DXS52 in Xq28, DXS255 in Xp11, and a series of minisatellites in the pseudoautosomal region at the distal end of Xp22.3.

37. J. L. Weber and P. E. May, *Am. J. Hum. Genet.* **44**, 388 (1989); M. Litt and J. A. Luty, *ibid.*, p. 397.
38. T. H. M. Huang, R. W. Cottingham, D. H. Ledbetter, H. Y. Zoghbi, *Genomics* **13**, 375 (1992); M. P. Coleman *et al.*, *ibid.* **11**, 991 (1991); A. K. Gedeon, K. Holman, R. I. Richards, J. C. Mulley, *Am. J. Med. Genet.* **43**, 255 (1992).
39. P. Fain and D. Barker, Third X chromosome workshop, Almalfi (1992).
40. C. Petit, J. Levilliers, J. Weissenbach, *EMBO J.* **7**, 2369 (1988); F. Rouyer *et al.*, *Nature* **319**, 291 (1986).
41. N. J. Royle, R. E. Clarkson, Z. Wong, A. J. Jeffreys, *Genomics* **3**, 325 (1988).
42. S. Abbs, R. G. Roberts, C. R. Mathew, D. R. Bentley, M. Bobrow, *ibid.* **7**, 602 (1990); C. Oudet, A. Hanauer, P. Clemens, T. Caskey, J. L. Mandel, *Hum. Mol. Genet.*, in press.
43. R. I. Richards *et al.*, *J. Med. Genet.* **28**, 818 (1991).
44. J. Zonana *et al.*, *Am. J. Hum. Genet.*, in press; V. Verga *et al.*, *ibid.* **48**, 1133 (1991); Z. Tumer *et al.*, *Hum. Genet.* **88**, 668 (1992).
45. J. L. Gorski *et al.*, *Am. J. Hum. Genet.* **48**, 53 (1991).
46. A. Sefiani *et al.*, *Hum. Genet.* **86**, 297 (1991).
47. A. Ballabio, *Curr. Op. Genet. Dev.* **1**, 25 (1991); contiguous gene syndromes are also known for some autosomal regions, but in this case patients are monosomic for the deleted region, and express mostly dominant traits.
48. U. Francke *et al.*, *Am. J. Hum. Genet.* **37**, 250 (1985).
49. M. F. Ho *et al.*, *ibid.* **50**, 317 (1992); M. A. Musarella *et al.*, *Genomics* **11**, 263 (1991).
50. An average density of about 15 to 30 genes per megabase can be estimated assuming the human genome has 50 to 100 thousand genes in total.
51. E. Maestrini *et al.*, *Genomics* **8**, 664 (1990); E. Maestrini *et al.*, *Hum. Mol. Genet.* **1**, 275 (1992); S. Saccone, A. DeSario, G. DellaValle, G. Bernard, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4913 (1992).
52. A. P. Monaco *et al.*, *Nature* **316**, 842 (1985).
53. To map within an interval of about 2 cM, it is necessary to study at least 100 informative siblings, born from demonstrated carrier mothers, and whose status at the disease locus is fully known (affected or normal males and those females for which carrier status can be determined). Furthermore, germinal mosaicism for new mutations appears to be rather frequent and, if unsuspected, can lead to errors in interpretation of linkage analysis [E. Bakker *et al.*, *J. Med. Genet.* **26**, 553 (1989); B. Arveiler *et al.*, *Am. J. Hum. Genet.* **46**, 906 (1990)].
54. K. G. Kupke, M. B. Graeber, U. Müller, *Am. J. Hum. Genet.* **50**, 808 (1992); E. M. Sinkala *et al.*, *Hum. Genet.* **84**, 66 (1989).
55. G. Neri, P. Chiurazzi, F. Arena, H. A. Lubs, I. A. Glass, *Am. J. Med. Genet.* **43**, 373 (1992); B. Kerr, G. Turner, J. Mulley, A. Gedeon, M. Partington, *J. Med. Genet.* **28**, 378 (1992).
56. I. Bach; D. Robinson, N. Thomas, H. H. Ropers, F. P. M. Cremers, *Hum. Genet.*, in press.
57. K. A. Ellison *et al.*, *Am. J. Hum. Genet.* **50**, 278 (1992); N. Archidiacono *et al.*, *Hum. Genet.* **86**, 604 (1991).
58. R. G. Roberts, M. Bobrow, D. R. Bentley, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2331 (1992).
59. P. H. Yen *et al.*, *Cell* **61**, 603 (1990); A. Ballabio, B. Bardoni, S. Guioli, E. Basler, G. Camerino, *Genomics* **8**, 263 (1990).
60. M. Koenig *et al.*, *Am. J. Hum. Genet.* **45**, 489 (1989); X. Hu *et al.*, *ibid.* **46**, 682 (1990); J. T. Den Dunnen *et al.*, *ibid.* **45**, 835 (1989).
61. L. A. J. Blonden *et al.*, *Genomics* **10**, 631 (1991); the deletion-prone regions may also be hot spots of meiotic recombination (*42*).
62. F. Gianelli *et al.*, *Nucleic Acids Res.* **19**, 2193 (1991).
63. E. G. D. Tuddenham *et al.*, *ibid.*, p. 4821.
64. C. D. K. Bottema *et al.*, *Am. J. Hum. Genet.* **49**, 839 (1991).
65. M. Higuchi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7405 and 8307 (1991).
66. B. A. Dombroski, S. L. Mathias, E. Nanthakamur, A. F. Scott, H. H. Kazazian, *Science* **254**, 1805 (1991).
67. D. Vidaud *et al.*, *Eur. J. Hum. Genet.*, in press.
68. T. Vulliamy, P. Mason, L. Luzzatto, *Trends Genet.* **8**, 138 (1992); in contrast with the other cited diseases, several G6PD mutations have spread in human populations, as they conferred a selective advantage in regions endemic for malaria.
69. E. J. Kremer *et al.*, *Science* **252**, 1711 (1991); I. Oberlé *et al.*, *ibid.*, p. 1097; F. Rousseau *et al.*, *N. Engl. J. Med.* **325**, 1673 (1991); Y. H. Fu *et al.*, *Cell* **67**, 1047 (1991); A. J. Verkerk *et al.*, *ibid.* **65**, 905 (1991).
70. A. R. La Spada, E. M. Wilson, D. B. Lubahn, A. E. Harding, K. H. Fischbeck, *Nature* **352**, 77 (1991).
71. M. F. Lyon, *Am. J. Hum. Genet.* **42**, 8 (1988).
72. A. Ballabio and H. F. Willard, *Curr. Opin. Genet. Dev.* **2**, 439 (1992).
73. S. Rastan and S. D. M. Brown, *Genet. Res.* **56**, 99 (1990); B. M. Cattanach *et al.*, *Cytogenet. Cell Genet.* **56**, 137 (1991); C. J. Brown *et al.*, *Nature* **349**, 82 (1991).
74. G. Borsani *et al.*, *Nature* **351**, 325 (1991); N. Brockdorff *et al.*, *ibid.*, p. 329.
75. M. Schmidt, *Am. J. Med. Genet.* **43**, 279 (1992); J. T. R. Clarke *et al.*, *Am. J. Hum. Genet.* **51**, 316 (1992).
76. A. Ashworth, S. Rastan, R. Lovell-Badge, G. Kay, *Nature* **351**, 406 (1991).
77. B. R. Migeon, *Genet. Res.* **56**, 91 (1990).
78. Y. Boyd and N. J. Fraser, *Genomics* **7**, 182 (1990); R. W. Hendriks, Z. Y. Chen, H. Hinds, R. K. B. Schuurman, I. W. Craig, *Hum. Mol. Genet.* **1**, 187 (1992).
79. S. Ohno, *Sex Chromosomes and Sex Linked Genes* (Springer, Berlin, 1967).
80. S. Brown, P. Avner, G. E. Herman, *Mamm. Genome*, in press.
81. N. Brockdorff *et al.*, *ibid.* **1**, 152 (1991).
82. J. M. Watson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11256 (1991).
83. S. Baxendale, G. P. Bates, M. E. MacDonald, J. F. Gusella, H. Lehrach, *Nucleic Acids Res.* **19**, 6651 (1991).
84. P. Elvin *et al.*, *ibid.* **18**, 3913 (1990); G. M. Duyk, S. Kim, R. M. Myers, D. R. Cox, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8995 (1990); M. Lovett, J. Kere, L. M. Hinton, *ibid.* **88**, 9628 (1991); S. Parimoo, S. R. Patanjali, H. Shukla, D. D. Chaplin, S. M. Weissman, *ibid.*, p. 9623; A. J. Buckler *et al.*, *ibid.*, p. 4005.
85. B. Korn *et al.*, *Hum. Mol. Genet.* **1**, 235 (1991).
86. M. D. Adams *et al.*, *Science* **252**, 1651 (1991).
87. R. Mazzarella *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3681 (1992).
88. E. C. Uberbacher and R. J. Mural, *ibid.* **88**, 11261 (1991).
89. F. P. Cremers, D. J. van de Pol, L. P. van Kerkhoff, B. Wieringa, H. H. Ropers, *Nature* **347**, 347 (1990).
90. B. Franco *et al.*, *ibid.* **353**, 529 (1991); R. Legouis *et al.*, *Cell* **67**, 423 (1991).
91. O. Attree *et al.*, *Nature* **358**, 239 (1992).
92. W. Berger *et al.*, *Nat. Genet.* **1**, 199 (1992); Z. Y. Chen *et al.*, *ibid.*, p. 204.