A linkage between DNA markers on the X chromosome and male sexual orientation
Dean Hamer, Stella Hu, Victoria A. Magnuson, Nan Hu and Angela M.L. Pattatucci
Science. 261.5119 (July 16, 1993): p321.

Abstract:

The role of genetics in male sexual orientation was investigated by pedigree and linkage analyses on 114 families of homosexual men. Increased rates of same-sex orientation were found in the maternal uncles and male cousins of these subjects, but not in their fathers or paternal relatives, suggesting the possibility of sex-linked transmission in a portion of the population. DNA linkage analysis of a selected group of 40 families in which there were two gay brothers and no indication of nonmaternal transmission revealed a correlation between homosexual orientation and the inheritance of polymorphic markers on the X chromosome in approximately 64 percent of the sib-pairs tested. The linkage to markers on Xq28, the subtelomeric region of the long arm of the sex chromosome, had a multipoint lod score of 4.0 (P = [10.sup.-5]), indicating a statistical confidence level of more than 99 percent that at least one subtype of male sexual orientation is genetically influenced.

Full Text:

Human sexual orientation is variable. Although most people exhibit a heterosexual preference for members of the opposite sex, a significant minority display a homosexual orientation. This naturally occurring variation presents an opportunity to explore the mechanisms underlying human sexual development and differentiation.

The role of genetics in sexual orientation has been previously approached by twin, adoption, and nuclear family studies. From the rates of homosexuality observed in the monozygotic and dizygotic twins, ordinary siblings, and adoptive (adopted in) brothers and sisters of homosexual men (1, 2) and women (3, 4), overall heritabilities of 31 to 74 percent for males and 27 to 76 percent for females were estimated. However, the precise extent of genetic loading is unclear because systematic data on relatives raised apart (adopted out) are not available and because the number and nature of the putative inherited factors are unknown. The observation that male homosexuals usually have more gay brothers than gay sisters, whereas lesbians have more gay sisters than gay brothers, suggests that the factors responsible for this familial aggregation are at least partially distinct in men compared to women (3, 5).

Recent neuroanatomical studies have revealed differences between heterosexual and homosexual men in the structure of three regions of the brain; namely, the third interstitial nucleus of the anterior hypothalamus (6), the anterior commissure (7), and the suprachiasmatic nucleus (8). The role of gonadal steroids in the sexual differentiation of the mammalian brain is well established (9), but thus far the role of hormonal variations in normal human sexual development is unknown (10). Nonbiological sources of variation in human sexual expression have been under consideration in diverse disciplines including psychiatry, psychology, religion, history, and anthropology (11).

The goal of our work was to determine whether or not male sexual orientation is genetically influenced. We used the standard techniques of modern human genetics, namely pedigree analysis and family DNA linkage studies. Recent advances in human genome analysis, in particular the development of chromosomal genetic maps that are densely populated with highly polymorphic markers, make it feasible to apply such methods to complex traits, such as sexual orientation, even

if these traits are influenced by multiple genes or environmental or experiential factors, or some combination of these (12). Our data indicate a statistically significant correlation between the inheritance of genetic markers on chromosomal region Xq28 and sexual orientation in a selected group of homosexual males.

Characteristics of study participants. The subjects studied were self-acknowledged homosexual men and their relatives over age 18. The initial sample for pedigree analysis consisted of 76 index subjects who were recruited through the outpatient HIV clinic at the National Institutes of Health Clinical Center, the Whitman-Walker Clinic in Washington, D.C., and local homophile organizations. One or more relatives from 26 of these families also participated in the project (total n = 122). The sample for the sib-pair pedigree study consisted of 38 pairs of homosexual brothers, together with their parents or other relatives when available, who were recruited through advertisements in local and national homophile publications. Two additional families who were originally in the randomly ascertained pool were added to this group for the DNA linkage study (total n = 114). Subjects signed an Informed Consent, approved by the NCI Clinical Review Subpanel, prior to donating blood and completing an interview or questionnaire covering childhood gender identification, childhood and adolescent sexual development, adult sexual behavior, the Kinsey scales, handedness, alcohol and substance use, mental health history, medical genetics screen, HIV status, and demographics (13, 14). The participants were white non-Hispanic (92 percent), African American (4 percent), Hispanic (3 percent), and Asian (1 percent) and had an average educational level of 15.5 [+ or -] 2.4 (mean [+ or -] SD) years and an average age of 36 [+ or -] 9 (mean [+ or -] SD) years.

Sexual orientation was assessed by the Kinsey scales, which range from 0 for exclusive heterosexuality to 6 for exclusive homosexuality (13). Subjects rated themselves on four aspects of their sexuality: self-identification, attraction, fantasy, and behavior. Of the homosexual subjects, > 90 percent self-identified as either Kinsey 5 or 6 whereas > 90 percent of their nonhomosexual male relatives self-identified as either 0 or 1 (Fig. 1). The sexual attraction and fantasy scales gave even greater dispersions between the groups, with >95 percent of the participants either less than Kinsey 2 or more than Kinsey 4. Only the sexual behavior scale gave a small overlap between the two groups largely because of adolescent and early adult experiences. Therefore, for our study, it was appropriate to treat sexual orientation as a dimorphic rather than as a continuously variable trait. Similar bimodal distributions of Kinsey scores in males have been reported by others (1, 2).

The age of phenotypic expression of homosexuality was assessed by asking the subjects at what age they were first attracted to another male, when they acknowledged their sexual orientation to themselves, and when they acknowledged their orientation to others. Most of the subjects experienced their first same-sex attraction by age 10, which was prior to the average age of puberty at 12 years (Fig. 2). Self-acknowledgement occurred over a broad range of ages between 5 and 30 years, with the greatest increase occurring between years 11 and 19. The mean age for public acknowledgement was 21 years, which is similar to the average age for "coming out" reported by others (15). Since the average age of our subjects was 36 [+ or -] 9 years, we did not correct for age-dependent phenotypic expression in subsequent analyses.

Pedigree analysis. Traits that are genetically influenced aggregate in families and, in the case of dominant or sex-linked inheritance, are transmitted from one generation to the next. Family histories were collected from 114 homosexual male probands who were asked to rate their fathers, sons, brothers, uncles, and male cousins as either definitely homosexual (Kinsey 5 or 6, acknowledged to the proband or another family member) or not definitely known to be homosexual (heterosexual, bisexual, or unclear). The reliability of the probands' assessment of their family members' sexual

orientation was estimated by conducting interviews with 99 relatives of the index subjects. All (69/69) of the relatives identified as definitely homosexual verified the initial assessment, as did most (27/30) of the relatives considered to be nonhomosexual; the only possible discrepancies were one individual who considered himself to be asexual and two subjects who declined to answer all of the interview questions. Hence describing individuals as either homosexual or nonhomosexual, while undoubtedly overly simplistic, appears to represent a reliable categorization of the population under study.

On the basis of a separate study in which the uncles and male cousins of lesbians were interviewed (16), we estimated that the population prevalence of male homosexuality is 2 percent (14/717). Although this rate is lower than the popularly accepted figures of 4 to 10 percent for male homosexuality, probably due to the more stringent definition applied here, it was considered more accurate for this analysis since the sampling, interview format, and definition of homosexual orientation were identical to those used in the male study. Similarly low rates for the population incidence of homosexuality have been reported when recent sexual behavior was used as the criterion (17).

The pedigree analysis for the male relatives of the 76 randomly ascertained homosexual male probands indicated (Table 1) that the highest rate of homosexual orientation was in brothers, who had a 13.5 percent chance of being gay, representing a significant 6.7-fold increase over the estimated background rate of 2 percent P < 0.001). Among more distant relatives, only two groups had significantly higher rates of homosexual orientation than the population incidence, namely maternal uncles and the sons of maternal aunts. Both of these maternally related classes of relatives had rates of [is nearly equal] 7.5 percent, which were significantly higher than the background rate (P < 0.01). By contrast, fathers and all other types of paternally related relatives had rates that were lower or not significantly different from the background. Background rates of homosexuality were also observed in the female relatives of the homosexual male probands (except for sisters, who had a 5.4 percent rate versus a I percent background rate) and in the male relatives of lesbian probands (except brothers, who had a 4.7 percent rate) (16).

Table 1. Rates of homosexual orientation in the male relatives of homosexual male probands. The 76 random probands were ascertained without the investigator's knowledge of family history of sexual orientation. The 38 sib-pair probands were selectively ascertained because they had a homosexual brother and no indication of transmission through fathers or to females. The population frequency of male homosexuality was estimated from the data for the uncles and male cousins of lesbian probands (16). **P < 0.001 compared to population frequency. *P < 0.01 compared to population frequency.

Relationship	Homosexual total	Percent
Random probands (n = 76)		
Father	0/76	0
Son	0/6	0
Brother	14/104	13.5**
Maternal uncle	7/96	7.3*
Paternal uncle	2/119	1.7
Maternal cousin, aunt's son	4/52	7.7*
Maternal cousin, uncle's son	2/51	3.9
Paternal cousin, aunt's son	3/84	3.6
Paternal cousin, uncle's son	3/56	5.4
Sib-pair probands (n = 38)		

Maternal uncle	6/58	10.3**
Paternal uncle	1/66	1.5
Maternal cousin, aunt's son	8/62	12.9**
Maternal cousin, uncle's son	0/43	0
Paternal cousin, aunt's son	0/69	0
Paternal cousin, uncle's son	5/93	5.4
	Population frequency	
Uncles and cousins of	14/717	2.0

female probands

Although the observed rates of homosexual orientation in the maternally derived uncles and male cousins of gay men were higher than in female and paternally related male relatives, they were lower than would be expected for a simple Mendelian trait. Furthermore, there was a substantial number of families in which lesbians or paternally related gay men were present. This could be explained if some instances of homosexuality were male-limited and maternally inherited whereas others were either sporadic, not sex-limited, or not maternally transmitted. To test this, we recruited 38 families in which there were two homosexual brothers, no more than one lesbian relative, and no indication of direct father-to-son transmission of homosexuality (that is, neither the father nor son of a proband was gay). We hypothesized that this selected population of families would be enriched for the putative maternally transmitted genetic factor and therefore display further increases in the rates of homosexuality in maternally derived uncles and male cousins. Indeed, the rates of homosexuality in the relatives of these selected sib-pair probands were increased from 7.3 to 10.3 percent for maternal uncles and from 7.7 to 12.9 percent for the sons of maternal aunts (Table 1). By contrast, the rates of homosexuality in the other types of male relatives were unchanged or decreased compared to the initial study. The differences between the random and sib-pair populations were not significantly different ($P > 0.1$); however, the differences between all maternal relatives as compared to all nonmaternal relatives were significant within both the randomly ascertained group ($P < 0.05$) and the sib-pair group ($P < 0.001$).

Several examples of the apparent maternal transmission of male homosexual orientation are shown in Fig. 3. Families DH99002 and DH99017, which were randomly ascertained, are characterized by a single gay man in each of three maternally related generations. In family DH321, which was recruited as part of the sib-pair study, a pair of homosexual brothers have a maternally related gay nephew and uncle. Family DH210, which was ascertained as part of a separate study, contains seven homosexual males, all related through the sequential marriage of two sisters to the same husband in generation II. In several families, maternally related half-brothers or half-cousins shared a homosexual orientation (16). The striking feature of these multiplex pedigrees is the absence of transmission through the paternal line and the paucity of female homosexuals.

These results demonstrate increased rates of homosexual orientation not only in the brothers of gay men, as has been previously reported (1, 2), but also in maternal uncles and the sons of maternal aunts (18). Because uncles and cousins share inherited information with the index subjects, but are raised in different households by different parents, this observation favored an interpretation based on genetics rather than the rearing environment and suggested that linkage studies might be fruitful.

X chromosome linkage. One explanation for the maternal transmission of a male-limited trait is X chromosome linkage. Since males receive their single X chromosome exclusively from their mothers, any trait that is influenced by an X-linked gene will be preferentially passed through the mother's side of the family. DNA linkage analysis provides the means to distinguish X-linked inheritance from competing hypotheses such as maternal effects, imprinting, decreased

reproductive rates of expressing males, or differential knowledge concerning maternal versus paternal family members. If the X chromosome contains a gene that increases the probability of an individual's being homosexual, then genetically related gay men should share X chromosome markers close to that gene. If no such gene exists, then no statistically significant correlations between sexual orientation and X chromosome markers will be observed (19).

We performed the linkage analysis on the selected population of families described above in which there were two homosexual brothers. This sib-pair experimental design has several theoretical and practical benefits (20): (i) it is nonparametric and independent of gene penetrance and frequency; (ii) it is capable of detecting a single linked locus even if additional genes or environmental conditions are required to express the trait; (iii) it is more powerful to study siblings than more distant relatives for traits displaying limited familiality; (iv) "false negatives" (individuals who have or will have a homosexual orientation but choose to identify themselves as heterosexual) are irrelevant to the analysis because they are not studied; (v) "false positives" (individuals who have a heterosexual orientation but choose to identify themselves as homosexual) are expected to be rare; (vi) the sib-pair method is more stable to errors in genotyping and to mistakes or alterations in phenotype than are large pedigree methods; and (vii) it was more practical to obtain the cooperation of nuclear sib-pair families than of multigenerational families.

The sample for the linkage analysis consisted of 40 pairs of homosexual brothers (38 from the sib-pair pedigree study and 2 from the random sample) together with their mothers or other siblings if available. DNA was prepared from all available members of these families and typed for a series of 22 markers that span the X chromosome. Each sib-pair was scored as either concordant-by-descent (D) if the mother was known to be heterozygous and both sons inherited the same allele, concordant-by-state (S) if the mother was unavailable and both sons shared the same allele, discordant (-) if the two sons carried different alleles, or noninformative (n) if the mother was homozygous for the marker. For families in which DNA from the mother was not available, the data for the concordant-by-state pairs were corrected for the possibility that the mother was homozygous for the marker by taking into account the population frequency of the allele coinherited by the two sons (19, 20). Using a likelihood ratio test, we then calculated for each locus the probability ([z.sub.1]) of the brothers sharing the marker by-descent and the statistical significance (P) of deviations from the value of [z.sub.1] = 1/2 expected under the null hypothesis of no linkage.

The X chromosome markers used for linkage analysis were simple sequence repeats, variable number of tandem repeats, and restriction fragment length polymorphisms, all of which were detected by the polymerase chain reaction (PCR) (Table 2). Heterozygosities, which were determined by analyzing 62 to 150 independent X chromosomes from the sib-pair and related populations, ranged from 0.35 to 0.87. An example of genotype determination with a [(GT).sub.n] [GC(GT).sub.n]-repeat marker, DXS1108, is shown in Fig. 4. Despite the presence of shadow bands, the individual alleles were readily distinguishable, and concordant and discordant sib-pairs could be clearly differentiated. As expected for this X-specific marker, the alleles inherited by the sons were derived exclusively from the mother. By contrast, the marker DXYS154, which lies on the tip of Xq in a region of subtelomeric homology and genetic exchange between the X and Y chromosomes, displayed alleles contributed by both the father and the mother (21). As expected for this tightly sex-linked marker, almost all of the male siblings inherited the same Y chromosome allele from their fathers (22); therefore, only the contribution from the maternal X chromosome was considered in the analysis of this locus (23).

The linkage analysis included a statistical analysis of the pair-by-pair data (Tables 2 and 3) and multipoint mapping analysis of the X chromosome (Fig. 5). The main outcome was the detection of

linkage between homosexual orientation and markers in the distal portion of Xq28. Each of the five markers in this region gave values of [z.sub.1] > 0.8, and for the three most heterozygous loci the data were significant at P < 0.0003 (Table 2). The five terminal loci on Xq28 are clustered within 2.8 to 4.3 cM 21, 24), and within our collection of 40 families exhibited no unequivocal intramarker recombination events (25). Therefore, the entire distal region of Xq28 could be considered as a single extended locus with a haplotype heterozygosity of 0.99; this transformation of the data increases the power to detect linkage by decreasing the uncertainties due to nongenotyped and noninformative mothers. Of the 40 sib-pairs, 33 were concordant for all markers within this region, whereas 7 pairs were discordant at one or more loci (Table 3). This analysis gives a value of [z.sub.1] = 0.82 at a significance of P 1.2 x [10.sup.-5].

Evaluation of the data by multipoint mapping with the LINKMAP routine of the computer program LINKAGE 5.1 supported the linkage between homosexual orientation and distal Xq28. The model used for analysis was an X-linked, male-specific gene with a mutation rate of 0. The population frequency of the homosexuality-associated allele was assumed to be 0.02, and penetrances were set at 0 for all females, 0 for males lacking the trait-associated allele, and 0.5 for males having the trait-associated allele; heterosexual brothers were not included in the analysis. The peak multipoint lod score was 3.96 to 4.02 (Fig. 5), depending on whether compressed or full allele information was used; because the lod score is the logarithm to the base 10 of the odds ratio, this corresponds to an odds ratio of = 10,000:1. The apparent location of the peak was 8 cM distal of DXYS154. However, this is likely to be an overestimate due to the well-known phenomenon of biased recombination fraction estimation in the case of complex traits where the analysis model differs from the true model (19). Therefore, the data were reanalyzed under two alternative models (Fig. 5B). When the frequency of the trait-associated allele was increased from 0.02 to 0.1, as suggested by Risch and Giuffra (26), the peak lod score decreased to 3.9 and the distance from DXYS154 decreased to 5 cM. When the penetrance of the non-trait-associated allele was increased from 0 to 0.05, giving a substantial level of phenocopies, the peak lod score of 3.9 fell directly over DXYS154, and the lod scores throughout distal Xq28 were greater than 3.5. Given that DXYS154 lies within 1 Mb of the telomere (27), these latter models probably yield more accurate estimates of the locus position. More precise mapping will require more distal markers, a larger number of families, and additional information concerning the trait parameters.

There was no significant evidence for linkage between sexual orientation and loci lying outside of Xq28. Most of the markers on the remainder of the long arm, and all of the markers on the short arm, gave values of [z.sub.1] that were statistically indistinguishable from the null hypothesis (P > 0.05) (Table 1). Although there was a moderate excess of concordant pairs at the markers DXS456, DXS297, and DXS548 (0.002 [is less than or equal to] P [is less than or equal to] 0.02), it is unlikely that these loci play a significant role in sexual orientation because they were adjacent to markers that gave negative results. Furthermore, multipoint mapping gave lod score less than -2 throughout the region between the KAL locus at Xp22.3 and the DXS994 locus at Xq26 and around the fragile X locus at Xq27.3. However, a much larger sample would be required to stringently eliminate these regions from playing a role in sexual development in a small proportion of families.

Contribution of genetics to male sexual orientation. The proof for the involvement of genes in a human behavioral trait must ultimately consist of the chromosomal mapping of the loci and isolation of the relevant DNA sequences. Such molecular studies are essential to separate the role of inheritance from environmental, experiential, social, and cultural factors. DNA linkage studies of families in which the trait appears to be genetically segregating represent the first step in this approach.

We have now produced evidence that one form of male homosexuality is preferentially transmitted through the maternal side and is genetically linked to chromosomal region Xq28. In a selected population of families in which there were two homosexual brothers and no transmission through fathers or to females, 33 of 40 sib-pairs had coinherited genetic information in this subtelomeric region. Observing such an association by chance alone has a type I error rate of approximately 0.001 percent for a single tested region of the genome (haplotype P = 1.2 x [10.sup.-5]), and therefore an error rate of less than 0.03 percent for a collection of 22 independent markers (P = 22 x 1.2 x [10 .sup.-5] = 0.0003). Similarly, multipoint linkage mapping gave a peak lod score of 4.0, which is associated with an overall type I error level of 0.5 percent, even for a complete genome search (12, 19). Thus, both forms of analysis indicate that the linkage results are statistically significant at a confidence level of > 99 percent. As with all linkage studies, replication and confirmation of our results are essential. The observed excess coinheritance of Xq28 markers by homosexual brothers is not due to segregation distortion because normal, Mendelian segregation has been demonstrated for many different Xq28-linked traits and polymorphic markers (28). Rather, it appears that Xq28 contains a gene that contributes to homosexual orientation in males.

There were seven pairs of brothers who did not coinherit all of the Xq28 markers. Such discordant pairs could arise because of homozygosity of the mother at the sexual orientation-related locus, recombination between the locus and a marker gene, genetic heterogeneity, or nongenetic sources of variation in sexual orientation. We estimate that the last two categories comprise approximately 36 percent of the sib-pair population (29). At present, we can say nothing about the fraction of all instances of male homosexuality that are related or unrelated to the Xq28 candidate locus because of the selection for genetically loaded families that is imposed by linkage methods. We also have no information about the role, or lack thereof, of the Xq28 region in multiplex families containing multiple gay men or lesbians (or both) (29), nor about the presence or absence of the homosexuality-associated allele in brothers or other male relatives who identify as heterosexual. Given the overall complexity of human sexuality, it is not surprising that a single genetic locus does not account for all of the observed variability. Sib-pairs that are discordant at Xq28 should provide a useful resource for identifying additional genes or environmental, experiential, or cultural factors (or some combination of these) that influence the development of male sexual orientation.

Our experiments suggest that a locus (or loci) related to sexual orientation lies within approximately 4 million base pairs of DNA on the tip of the long arm of the X chromosome. Although this represents less than 0.2 percent of the human genome, it is large enough to contain several hundred genes. The fine mapping and eventual isolation of this locus will require either large numbers of sib-pairs, more extended families, or the complete DNA sequence of the region. Once a specific gene has been identified, we can find out where and when it is expressed and how it ultimately contributes to the development of both homosexual and heterosexual orientation. The Xq28 region is characterized by a high density of genetic loci (28) and contains both repeated DNA sequences (27) and a pseudoautosomal region of homology and genetic exchange between the X and Y chromosomes (21). Recombination between tandemly repeated sequences, or between active and inactive loci on the X and Y chromosomes, could generate DNA sequence variants at a high rate and thereby account for the genetic transmission of a trait that may reduce reproduction.

The subjects for our linkage study were males who self-identified as predominantly or exclusively homosexual within the context of modern American society; such studies could be broadened to include individuals who identify as bisexual or ambisexual. The role of the Xq28 candidate locus, and of other chromosomal regions, in female sexual orientation remains to be tested. Although nuclear family studies suggest that the overall heritability of sexual orientation is similar in men and women (2, 4), their pedigree segregation patterns appear to be distinct (16).

Our work represents an early application of molecular linkage methods to a normal variation in human behavior. As the human genome project proceeds, it is likely that many such correlations will be discovered. We believe that it would be fundamentally unethical to use such information to try to assess or alter a person's current or future sexual orientation, either heterosexual or homosexual, or other normal attributes of human behavior. Rather, scientists, educators, policy-makers, and the public should work together to ensure that such research is used to benefit all members of society.

REFERENCES AND NOTES [1.] R. C. Pillard and J. D. Weinrich, Arch. Gen. Psychiatry 43, 808 (1986). [2.] J. M. Bailey and R. C. Pillard, ibid. 48, 1089 (1991). [3.] J. M. Bailey and D. S. Benishay, Am. J. Psychiatry 150, 272 (1993). [4.] J. M. Bailey, R. C. Pillard, M. C. Neale, Y. Agyei, Arch. Gen. Psychiatry 50, 217 (1993). [5.] J. D. Weinrich, Sexual Landscapes (Scribner, New York, 1987); R. C. Pillard, in Homosexuality/ Heterosexuality: Concepts of Sexual Orientation, D. McWhirter, S. A. Sanders, J. M. Reinisch, Eds. (Oxford Univ. Press, London, 1990), pp. 88-100; J. M. Bailey, L. Willermaan, C. Parks, Arch. Sex. Behav. 20, 277 (1991). [6.] S. Levay, Science 253, 1034 (1991). [7.] L. S. Allen and R. A. Gorski, Proc. Natl. Acad. Sci. U. S.A. 89, 7199 (1992). [8.] D. F. Swaab and M. A. Hofman, Brain Res. 537, 141 (1990). [9.] R. A. Gorski, in Brain Endocrinology, M. Motta, Ed. (Raven, New York, 1991), pp. 71-104; B. S. McEwen et al., J Steroid Biochem. Mol. Biol. 39, 223 (1991). [10.] H. F. Meyer-Bahlburg, Clin. Endocrinol. Metab. 11, 681 (1982). [11.] R. Friedman, Psychoanal. Rev. 73, 483 (1986); C. A. Tripp, The Homosexual Matrix (Meridian, New York, 1987); R. A. Isay, Being Homosexual (Avon, New York, 1989); J. S. Spong, Living in Sin? (Harper and Row, San Francisco, 1988); R. Schow, W. Schow, M. Raynes, Eds., Peculiar People: Mormons and Same-Sex Orientation (Signature, Salt Lake City, 1991): A. L. Rouse, Homosexuals in History (Dorset, New York, 1977); M. Dubermaan, M. Vicinas, G. Chauncey, Jr., Eds., Hidden from History: Reclaiming the Gay and Lesbian Past (Meridian, New York, 1989); F. L. Whitam and R. M. Mathy, Male Homosexuality in Four Societies: Brazil, Guatemala, the Philippines, and the United States (Praeger, New York, 1986). [12.] E. S. Lander and D. Botstein, Proc. Natl. Acad. Sci. U.S.A. 83, 7353 (1986); E. S. Lander, in Genome Analysis, K. E. Davis, Ed. (IRL Press, New York, 1988), pp. 171-189. [13.] A. C. Kinsey and W. B. Pomeroy, Sexual Behavior in the Human Male (Sanders, Philadelphia, 1948). 14. K. Freund, R. Langevin, J. Satterberg, B. Steiner, Arch. Sex Behav. 6, 507 (1977); N. Buhrich J. M. Bailey, N. G. Martin, Behav. Genet. 21, 75 (1991). [15.] A. P. Bell and M. S. Weinberg, Homosexualities: A Study of Diversity among Men and Women (Simon and Schuster, New York, 1978). [16.] A. Pattatucci and D. Hamer, unpublished results. [17.] ACSF Investigators, Nature 360, 407 (1992); A. M. Johnson, J. Wadsworth, K. Wellings, S. Bradshaw, J. Field, ibid., P. 410. [18.] Increased rates of homosexuality in the maternal compared to paternal relatives of homosexual men can also be discerned in the data of G. W. Henry [Sex Variants: A Study of Homosexual Patterns (Hoeber, New York, 1941)], B. Zuger [Arch. Sex. Behav. 18, 155 (1989)], and Pillard and Weinrich (1). [19.] J. Ott, Analysis of Human Genetic Linkage (Johns Hopkins Univ. Press, Baltimore, 1991). [20.] B. K. Suarez, J. Rice, T. Reich, Ann. Hum. Genet. 42, 87 (1978); G. Thomson, Am. J. Hum. Genet. 39, 207 (1986); W. Blackwelder and R. Elston, Genet. Epidemiol. 2, 85 (1985). L. Goldin and E. Gershon, ibid. 5,35 (1988); N. Risch, Am. J. Hum. Genet. 46, 222 (1990); ibid., p. 242; D. T. Bishop and J. A. Williamson, ibid., p. 254. [21.] D. Freije, C. Helms, M. S. Watson, H. Donis-Keller, Science 258, 1784 (1992). [22.] The published male recombination rate between sex and DXYS154 is 4/195 = 2 percent (21). The observed rate in our sib-pair population was 1/36 - 3 percent. The single recombinant family was DH441, for which the genotypes (alleles) were mother = [6,6], son 1 = [2,61, and son 2 = [6,6]; therefore this family was noninformative for the maternal X chromosome. [23.] The maternal contribution for DXYS154 was calculated as follows. Families of type mother (alleles) = [1,2], and son 1 = son 2 = [1,x] or [2,xl, where x = any allele, were scored as concordant-by-descent. Families of type son 1 = son 2 = [l,l) were scored as concordant-by-state with an allele frequency of ([f.sub.1.sup.2] + [f.sub.2.sup.2])/(2[f.sub.1] + 2[f.sub.2] - [f.sub.1.sup.2] - [f.sub.2.sup.2] - 2[f.sub.1 f.sub.2.] Families of type son 1 = [1,2] and son

2 = [1,3] were scored as discordant. Families of type mother = [1,1] were noninformative. If the homosexuality-associated gene were in fact derived from the father rather than from the mother in some families, this treatment of the data would decrease rather than increase the evidence for linkage. [24.] J. Weissenbach et al., Nature 359, 794 (1992); CEPH database, version 6.0. [25.] There was one possible recombination event, in family Dhl 101, between the proximal Xq28 locus GABRA3, which lies approximately 2 cM centromeric of DXS52, and the distal Xq28 gene cluster (Table 3). However, this could not be confirmed because the mother was not available for genotyping. [26.] N. Risch and L. Giuffra, Hum. Hered. 42, 77 (1992). [27.] D. Freije and D. Schlessinger, Am. J. Hum. Genet. 51, 66 (1992). [28.] There appears to be no published evidence for segregation distortion for several well-studied Xq28-linked traits such as color blindness and G6PD deficiency [V. A. McKusick, Mendelian Inheritance in Man (Johns Hopkins Univ. Press, Baltimore, 1992)]. Analysis of the CEPH database (version 6.0) for the Xq28 marker loci F8C and DXS52 also showed no indication of segregation distortion in 33 informative families containing 142 sons. Analysis of the first two sons in each family as a single pair gave 18 concordant-by-state pairs versus 15 discordant pairs ([z.sub.1 = 0.55, 2ln L([z.sub.1]) = 0.3, P > > 0.05). Analysis of all n(n - 1)/2 pairs for each family (where n is the number of sons) gave 152 concordant-by-state pairs compared to 162 discordant pairs ([z.sub.1] [is less than or equal to] 0.5, P >> 0.05). [29.] The fraction ([alpha]) of a sib-pair population in which a trait is associated with excess coinheritance f an X-linked marker is minimally estimated by [alpha] = s[z.sub.1] -1, which takes into account the fact that 1/2 of all sib-pairs will coinherit the marker by chance alone. Therefore the fraction of the sib-pair population in which the trait is not linked to the gene is 1 - [alpha] = 2(1 - [z.sub.1]. For Xq28, [z.sub.1] = 33/40 = 0.82, giving 1 - [alpha] = 0.36 = 36 percent. The proportion of the entire population in which the trait is linked to a marker cannot be estimated without (i) information on the frequencies and penetrances of the linked trait allele d additional loci and (ii) the frequency of phenocopies, none of which are known; under the simple model of a single gene and a high rate of phenocopies, Xq28 could account for as little as 10 percent of total variance. The contribution of Xq28 to sib-pairs cannot be extrapolated to larger families without further information. We have observed considerably greater variability in sexual development and expression in families containing more than two gay brothers or multiple lesbians (or both), suggesting that ascertainment may be more complicated in these cases. [30.] Of the 40 families, there were 14 for which DNA from he mother was genotyped for all loci (DH10, 050, 040, 371, 1221, 070, 471, 060, 151, 170, 441, 391, 220, and 020), 1 family for which DNA from the mother was genotyped for some but not all loci (DH141), and 1 family for which DNA from a sister but not the mother was genotyped (DH31). For the remaining 24 families, DNA was available only from the two brothers. DNA was prepared from peripheral blood by SDS lysis and salt precipitation [D. Lahiri and J.L. Nurnberger, Jr., Nucleic Acids Res. 19, 5444 (1991)]. The PRC procedure followed standard conditions for each primer pair, with analysis by electrophoresis on 6 percent denaturing acrylamide, 8 percent acrylamide, or 1 percent agarose gels. Marker loci B, C, E, F, G, H, 1, K, and L were as described 24), marker loci U and V were as described (21). Marker loci A (primers KaL.PCR1.1/KaL.PCR1.2), J (primers XG30BL/XG30BR), M (primers VK23F/ VK23R), and P (primers RS46-CAL/RS46-CA2) were as described in the Genome Data Base (William H. Welch Medical Library, Johns Hopkins University). Marker locus D (DMD1) was described by J. P. Hugnot et al. [Nucleic Acids Res. 19, 3159 (1991)], locus 0 (FRAXAC2) was as described [R. I. Richards et al., J. Med. Genet. 28, 818 (1991)], locus Q (GABRA3) was described by A. A. Hicks et al. [Nucleic Acids Res, 19, 4016 (1991)), locus R (DXS52) was described by B. Richards et al. (ibid., p. 1944), locus S (G6PD) was described by B. Kurdi-haidar et al. [Am. J. Hum. Genet. 47, 1013 (1990)], and locus T (F8C) was described by V. L. Surin et al. [Nucleic Acids Res. 18, 3432 (1990)] The (CCG)n -repeat at locus N (FMR) was assayed with the use of primers 203/213 described by E. J. Kremer et al. [Science 252, 1711 (1991)] under PCR conditions described by S. Yu et al. [Am. J. Hum. Genet. 50, 968 (1992)].

[31.] The data were analyzed by the likelihood ratio method of N. Risch [Am. J. Hum, Genet. 46, 229 (1990)] as modified for X-linked markers in male sib-pairs. Let [z.sub.1] = the probability that a pair of brothers share a marker allele by-descent ([z.sub.1] is an unknown parameter that is estimated from the data), [Tau]([z.sub.1]) = the likelihood of the observed data at [z.sub.1], [Tau](1/2) = the likelihood ofthe observed data under the null hypothesis of [z.sub.1] = 1/2, and L([z.sub.1]) = [Tau]([z.sub.1])/[Tau](1/2). Then for the complete data set 2ln L([z.sub.1]) =

[MATHEMATICAL EXPRESSION OMITTED]

where G is the number of discordant pairs, D is the number of concordant-by-descent pairs, S is the number of concordant-by-state pairs, f, the population frequency of the allele shared by the jth pair of concordant-by-state brothers, and N = G + D + S = the number of informative pairs. This function was evaluated as a function of z, (between 0.5 and 1) to find the maximal value of 2ln L([z.sub.1]). The one-sided significance P was estimated by considering 2ln L([z.sub.1]) to be distributed as a chi-square at 1 degree of freedom. [32.] We thank all of the participants, without whose cooperation and interest this research would not have been possible; W. Gahl, L. Charnas, S. Schiesinger, and the staff at the Whitman-Walker, NIAID, and NIH Interinstitute Genetics clinics for their assistance; D. Freije and H. Donis-Keller for communicating results prior to publication; W. McBride, C. Amos, J. Eldridge, and the staff of the Biomedical Supercomputer Center for technical and statistical advice; and E, Gershon. L, Goldin, J. Nathans, W. Gahl. D. Goldman, L. Charnas, E. Lander, M. Boehnke, F. Collins, members of the Laboratory of Biochemistry, and the reviewers for comments and suggestions on the manuscript.

2 April 1993; accepted 17 June 1993