

cation. The presence of multiple binding sites results in a high avidity of the polymerase array for single-stranded RNA, retaining the intermediates of RNA amplification while maintaining a low local affinity at each polymerase to allow the local template and primer movements necessary for processive RNA elongation. RNA recombination, which occurs by template switching during RNA replication (36), is responsible for much of the diversity of RNA viruses. The high frequency of switching between RNA templates might result from the presence of multiple templates and nascent strands on a shared polymerase lattice.

References and Notes

1. P. R. Cook, *Science* **284**, 1790 (1999).
2. E. Wimmer, C. U. T. Hellen, X. Cao, *Annu. Rev. Genet.* **27**, 353 (1993).
3. J. H. Strauss, E. G. Strauss, *Microbiol. Rev.* **58**, 491 (1994).
4. T. J. Chambers, C. S. Hahn, R. Galler, C. M. Rice, *Annu. Rev. Microbiol.* **44**, 649 (1990).
5. T. A. Van Dyke, J. B. Flanagan, *J. Virol.* **35**, 732 (1980).
6. J. Lama, A. V. Paul, K. S. Harris, E. Wimmer, *J. Biol. Chem.* **269**, 66 (1994).
7. D. A. Hope, S. E. Diamond, K. Kirkegaard, *J. Virol.* **71**, 9490 (1997).
8. J. M. Lyle et al., *J. Biol. Chem.* **277**, 16324 (2002).
9. J. Pata, S. C. Schultz, K. Kirkegaard, *RNA* **1**, 466 (1995).
10. M. T. Beckman, K. Kirkegaard, *J. Biol. Chem.* **273**, 6724 (1998).
11. W. Xiang, A. Cuconati, D. Hope, K. Kirkegaard, E. Wimmer, *J. Virol.* **72**, 6732 (1998).
12. J. L. Hansen, A. M. Long, S. C. Schultz, *Structure* **5**, 1109 (1997).
13. S. E. Diamond, K. Kirkegaard, *J. Virol.* **68**, 863 (1994).
14. S. D. Hobson et al., *EMBO J.* **20**, 1153 (2001).
15. S. D. Hobson, thesis, University of Colorado, Boulder (2000).
16. J. M. Lyle, E. Bullitt, K. Bienz, K. Kirkegaard, unpublished data.
17. H. Flyvbjerg, E. Jobs, S. Leibler, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5975 (1996).
18. S. J. Plotch, O. Palant, Y. Gluzman, *J. Virol.* **63**, 216 (1989).
19. D. A. Suhy, T. H. Giddings Jr., K. Kirkegaard, *J. Virol.* **74**, 8953 (2000).
20. R. C. Rust et al., *J. Virol.* **75**, 9808 (2001).
21. A. Schlegel, T. H. Giddings Jr., M. S. Ladinsky, K. Kirkegaard, *J. Virol.* **70**, 6576 (1996).
22. S. Dales, H. J. Eggers, I. Tamm, G. E. Palade, *Virology* **26**, 379 (1965).
23. K. Bienz, D. Egger, L. Pasamontes, *Virology* **160**, 220 (1987).
24. K. Bienz, D. Egger, T. Pfister, M. Troxler, *J. Virol.* **66**, 2740 (1992).
25. D. Egger, L. Pasamontes, R. Bolten, V. Boyko, K. Bienz, *J. Virol.* **70**, 8675 (1996).
26. M. W. Klymkowsky, *Nature Cell Biol.* **1**, 121 (1999).
27. J. R. Doedens, L. A. Maynell, M. W. Klymkowski, K. Kirkegaard, *Arch. Virol.* **9**, 159 (1994).
28. A. V. Paul et al., *Virology* **272**, 72 (2000).
29. D. Chretien, S. D. Fuller, E. Karsenti, *J. Cell Biol.* **129**, 1311 (1995).
30. J. Löwe, L. A. Amos, *EMBO J.* **18**, 2364 (1999).
31. A. Bravo, M. Salas, *EMBO J.* **17**, 6096 (1998).
32. S. Winston, N. Sueoka, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2834 (1980).
33. C. Weigel et al., *Mol. Microbiol.* **34**, 53 (1999).
34. G. L. Nelsestuen, *Chem. Phys. Lipids* **101**, 37 (1999).
35. A. R. Hill Jr., C. Bohler, L. E. Orgel, *Origins Life Evol. Biosph.* **28**, 235 (1998).
36. T. C. Jarvis, K. Kirkegaard, *EMBO J.* **11**, 3135 (1992).
37. Swiss PDB Viewer is described in N. Guex, M. C. Peitsch, *Electrophoresis* **18**, 2714 (1997) and is available at www.expasy.ch/spdbv. POV-RAY is available at www.povray.org.
38. Materials and methods are available as supporting

material on Science Online at www.sciencemag.org/cgi/content/full/296/5576/2218/DC1.

39. We thank S. M. Crowder, K. C. Garcia, D. Egger, I. R. Lehman, E. Mocarski Jr., and P. Sarnow for critical reading of the manuscript; D. Egger for experimental contributions; and S. Schultz and R. Striker for experimental suggestions. J.M.L. is recipient of a predoctoral fellowship from the Howard Hughes Medical Institute. Funded by the Hutchison Foundation for Translational Research, Eli Lilly, Inc., NIH grant AI-42119, and Swiss National Science Foundation grant 1-055397.98.

5 February 2002; accepted 13 May 2002

Covariation of Synaptonemal Complex Length and Mammalian Meiotic Exchange Rates

Audrey Lynn,^{1*} Kara E. Koehler,^{1*} LuAnn Judis,¹ Ernest R. Chan,¹ Jonathan P. Cherry,¹ Stuart Schwartz,^{1,2} Allen Seftel,^{3,4} Patricia A. Hunt,¹ Terry J. Hassold^{1†}

Analysis of recombination between loci (linkage analysis) has been a cornerstone of human genetic research, enabling investigators to localize and, ultimately, identify genetic loci. However, despite these efforts little is known about patterns of meiotic exchange in human germ cells or the mechanisms that control these patterns. Using recently developed immunofluorescence methodology to examine exchanges in human spermatocytes, we have identified remarkable variation in the rate of recombination within and among individuals. Subsequent analyses indicate that, in humans and mice, this variation is linked to differences in the length of the synaptonemal complex. Thus, at least in mammals, a physical structure, the synaptonemal complex, reflects genetic rather than physical distance.

Virtually all human genetic linkage studies have examined individual chromosomes or chromosome segments. Consequently, little is known about the overall number and location of meiotic exchanges in individual germ cells. Only one systematic linkage analysis of genome-wide levels of recombination in humans has been published. Broman and colleagues (1, 2) analyzed the inheritance of short tandem repeat polymorphisms in eight of the CEPH (Centre d'Etude du Polymorphisme Humain) reference families, examining all detectable recombination events per meiosis. This approach provides a useful tool for studying human recombination but has at least two limitations. First, it requires well-characterized, three-generation (or deeper) families. Hence, without acquisition of additional families, analysis is effectively limited to the few hundred meioses available from the CEPH registry. Second, the approach relies on analysis of transmitted haploid products instead of cells undergoing meiosis; consequently, only one-half of all exchanges can be detected (for example, after a single ex-

change, only two of the four chromatids are recombinant).

Recent cytological studies suggest that, by using antibodies against the DNA mismatch repair protein MLH1 to analyze meiosis I spermatocytes and oocytes (3), it may be possible to overcome these limitations. Specifically, studies analyzing the localization of MLH1 foci on synaptonemal complexes (SCs) in mouse (4) and human (5) spermatocytes suggested that these foci identify the sites of meiotic exchanges. However, as these analyses were based on small numbers of cells—45 spermatocytes from three mice (4) and 46 spermatocytes from a single human (5)—it was not possible to examine intra- and interindividual variation, nor was it possible to determine whether recombination varied with intrinsic or extrinsic factors (for example, the age of the individual).

To address these issues directly, we analyzed pachytene-stage cells from 14 control males (Fig. 1, table S1) (6); first, we asked whether the number and location of MLH1 foci conformed to expectations for a molecule that marks the sites of exchange. Details of these initial analyses are provided in supporting online text. Briefly, observations on 1384 cells from the 14 individuals yielded an overall mean of 49.1 ± 4.8 foci per cell and a range of 34 to 66 foci per cell, which is remarkably similar to data from CEPH males (fig. S1); estimates of chromosome-specific and total autosomal male maps were consis-

¹Department of Genetics, Case Western Reserve University, ²The Center for Human Genetics, University Hospitals of Cleveland, ³Department of Urology, Case Western Reserve University and University Hospitals of Cleveland, ⁴Cleveland VA Medical Center, 10900 Euclid Avenue, Cleveland, OH 44106, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: tjh6@po.cwru.edu

REPORTS

tent with previous observations (table S2); MLH1 foci were preferentially distally located, as expected (fig. S2A); and the foci displayed cross-over interference (fig. S2B, table S3), a well-known property of meiosis. Thus, our observations provide strong evidence that MLH1 foci do "mark" the sites of meiotic exchange in human males.

Next, we were interested in assessing the range of values in our study population. Almost nothing is known about variation in recombination in humans, and the available information is contradictory. Cytological studies of diakinesis-stage spermatocytes (7) suggest significant individual variation in mean chiasma frequencies, and studies of recombination with single sperm polymerase chain reaction assays (8) indicate differences over short genetic intervals. However, genome scans of CEPH families (1) could not identify significant differences among males in the overall number of meiotic exchanges. Thus, the extent of variation in the number and location of exchanges in the human male is not clear.

We compared the number of foci per cell among the 14 individuals (Fig. 2A, table S1), and we observed surprising variability both within and among individuals. For most individuals, we observed cells with as few as 40 and as many as 60 MLH1 foci; thus, in individuals with apparently normal spermatogenesis, some cells had only two-thirds the number of exchanges of other cells. Similarly, results of a permutation test with an analysis of variance-like F statistic (9) demonstrated significant interindividual variation ($P < 0.001$). The mean number of MLH1 foci per cell ranged from 46.2 ± 3.3 to 52.8 ± 4.8 , a difference of nearly 15% (Fig. 2A). This was not attributable to patient status, stage of pachytene, or age (Fig. 2, B, C, and D).

Although the basis for these differences is unclear, we were intrigued by the possibility that variation in SC formation and/or maintenance might play a role. Early in our study, we noted a striking difference in the SC lengths of chromosomes 21 and 22. Although these chromosomes have similar physical sizes (21q is 39 Mb, 22q is 43 Mb) (10, 11), their genetic lengths are quite different [for males, 54 centimorgans (cM) for 21q, and 70 cM for 22q] (12). This led us to ask if SC length reflects genetic or physical length of a chromosome. To address that question, we selected two sets of chromosome pairs: one set with similar physical sizes but different genetic lengths (chromosomes 21 and 22, described above) and one set with similar genetic lengths but different physical sizes (chromosome 16, 98 Mb and 106 cM; chromosome 19, 67 Mb and 104 cM) (12, 13). In both instances, SC length was correlated with genetic rather than physical length (14),

which suggests that the SC measures genetic distance.

To determine whether this holds for the entire genome, we analyzed cells from 7 of 14 study participants, comparing total autosomal SC length (determined by summing the lengths of all autosomes) with the number of autosomal MLH1 foci (Fig. 3A). Trend lines for the individuals had somewhat different slopes, which suggests that the interindividual variation we observed in the overall

number of MLH1 foci (Fig. 2A) was also reflected in the relationship between recombination rate and SC length. Nevertheless, for all seven individuals there was a simple, striking linear relationship between the number of MLH1 foci per cell and the total autosomal SC length.

To determine whether this relationship extends to other mammals, we conducted similar analyses on mouse meiocytes. We examined the relationship between total SC length

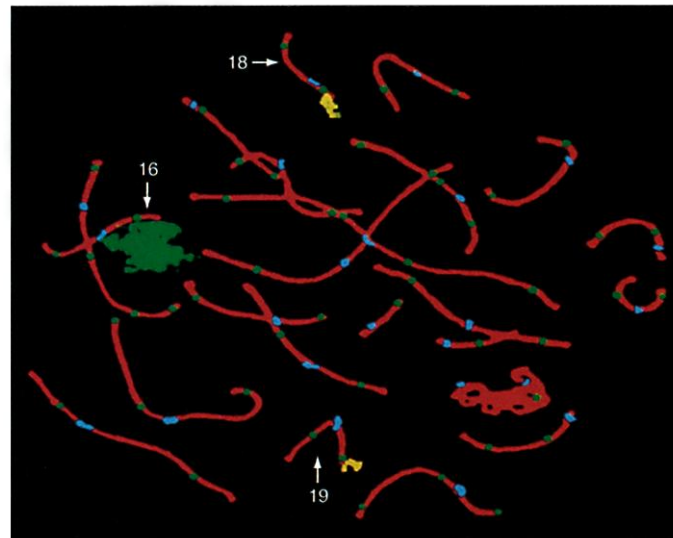


Fig. 1. Combined immunofluorescence/FISH analysis of a human pachytene spermatocyte. The cell was treated with antibodies against SCP3 (to visualize synaptonemal complexes; red), and MLH1 (to identify meiotic exchanges; green), and with CREST antisera (to detect centromeric regions; blue). Chromosome-specific FISH probes were used to identify chromosomes 16 (green), 18 (yellow, upper), and 19 (yellow, lower).

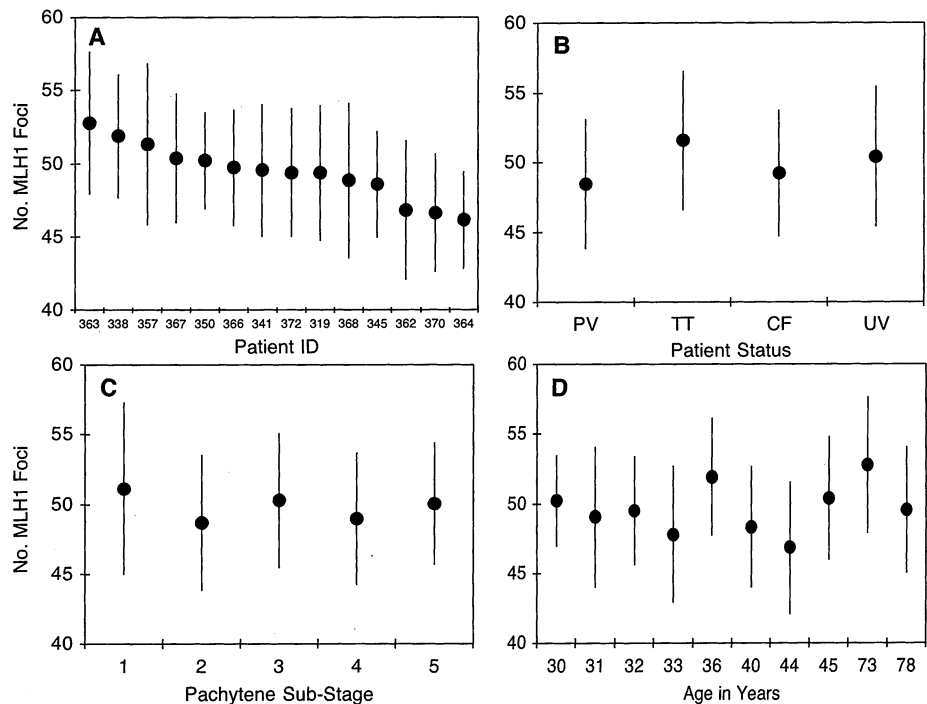


Fig. 2. Mean ($\pm$ SD) number of MLH1 foci per cell is plotted for the 14 individuals, with data categorized by individual (A), patient status (B), pachytene substage (C), and patient age (D). (A) Highly significant interindividual variation was observed among the 14 individuals, with mean numbers of foci ranging from 46.2 to 52.8. (B to D) There were no obvious reasons for the differences, as patient status (B), stage of pachytene [scored according to Solari (26)] (C), and age (D) had no apparent effect. CF, cystic fibrosis; PV, previous vasectomy; TT, testicular tumor; UV, unilateral varicocele.

and number of MLH1 foci per cell for males and females of three inbred strains. In previous studies of male mice, we identified strain-specific differences in the mean number of MLH1 foci per cell (15). Thus, in this analysis we asked whether the strain differences were correlated with variation in total SC length (as for humans) and whether the variation occurred because, on average, some strains had longer total SC lengths than others. The results are summarized in Fig. 3B. Although mice showed less variation than humans, the overall trends were similar; for each strain and for both sexes, the number of MLH1 foci was positively correlated with total SC length. Further, strain differences in MLH1 foci could be explained by differences in total SC length; indeed, the relationship between SC length and the number of foci was a virtual constant in males: 7.1-, 6.9-, and 6.9- μm SC length per MLH focus for CAST/Ei, C57BL/6, and SPRET/Ei, respectively.

These results provide evidence that a physical structure, the synaptonemal complex, "measures" genetic distance, at least in mammals. Previous studies reported a rough correlation between SC length and cross-over frequency (16); and in yeast, mutations in the SC-associated locus *Zip1* affect both cross-over frequency and interference (17). However, this study directly demonstrates that cross-over frequency and SC length co-vary; that is, interindividual variation in recombination is linked to variation in SC length. Conceptually, we can think of at least two ways this might occur. First, if the SC is central to the initiation/regulation of recombination (for example, as in *Drosophila*) (18), the number of recombination "nodules" (19) localizing to the SC might be affected by the length of the structure—longer SCs could promote more nodules and, ultimately, more chiasmata. Alternatively, if the SC forms in

response to earlier double-strand break-associated events (for example, as in *Saccharomyces cerevisiae*) (20), SC length might vary with the number and location of double-strand breaks. The SC would still measure genetic distance, but it would not be the source of variation in exchange frequencies. In humans, the exact temporal relationship between recombination proteins and SC components is not clear. However, if one assumes a sequence of events similar to that in mice (21), recombination likely precedes synapsis in both species. Thus, we suggest that allelic variation in loci encoding recombination machinery proteins (such as SPO11, MRE11, RAD51, and DMC1) may mediate differences in SC length and exchange frequency.

Regardless of the validity of this model, it seems likely that the relationship between exchange frequency and SC length is restricted to a subset of exchanges. That is, in humans 39 autosomal arms are available for recombination (excluding the five acrocentric short arms, where exchanges seldom occur) and, in virtually all cells we examined, each arm had at least one MLH1 focus. Thus, as in other organisms, there is likely a requirement for at least one exchange per arm. This suggests that the effect of SC length on exchange frequency is restricted to those "optional" exchanges in excess of 39; that is, in humans the SC adds an average of 7 to 14 exchanges, depending on the individual, and in mice it adds 2 to 6 exchanges depending on strain background. Mapping studies with mouse strains that have low and high levels of genome-wide recombination may reveal the loci responsible for the variation and thus provide candidate loci for analyses of human males.

Note added in proof: In an upcoming article, Tease *et al.* (22) analyze oocytes from a single human female fetus. They found that the mean number (70.3) and location of MLH1 foci fit the expectations predicted by female genotype data (1, 12). Additionally, for individual chromosomes, they noted that increasing SC length was accompanied by an increased number of MLH1 foci, although no formal analyses were conducted.

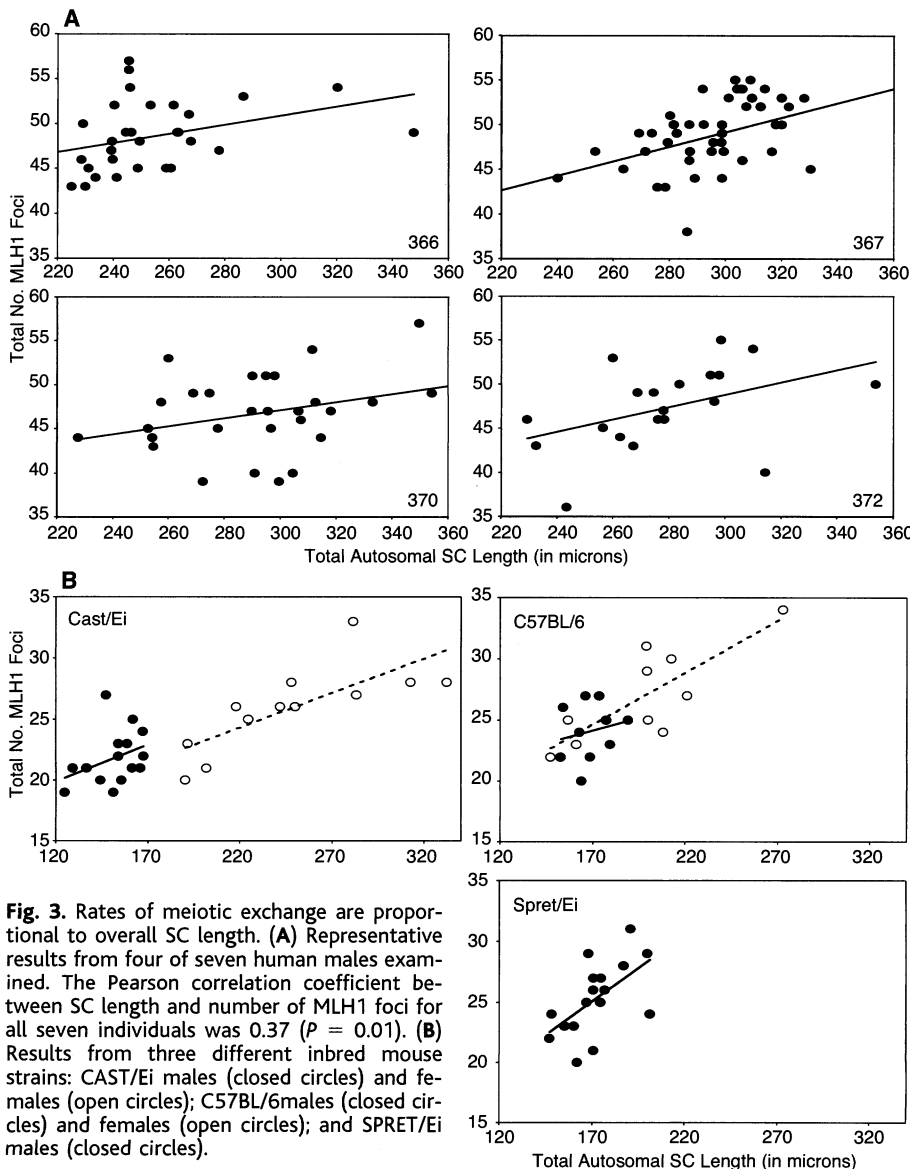


Fig. 3. Rates of meiotic exchange are proportional to overall SC length. (A) Representative results from four of seven human males examined. The Pearson correlation coefficient between SC length and number of MLH1 foci for all seven individuals was 0.37 ($P = 0.01$). (B) Results from three different inbred mouse strains: CAST/Ei males (closed circles) and females (open circles); C57BL/6 males (closed circles) and females (open circles); and SPRET/Ei males (closed circles).

References and Notes

1. K. W. Broman *et al.*, *Am. J. Hum. Genet.* **63**, 861 (1998).
2. K. W. Broman, J. L. Weber, *Am. J. Hum. Genet.* **66**, 1911 (2000).
3. S. M. Baker *et al.*, *Nature Genet.* **13**, 336 (1996).
4. L. K. Anderson, A. Reeves, L. M. Webb, T. Ashley, *Genetics* **151**, 1569 (1999).
5. A. L. Barlow, M. A. Hulten, *Eur. J. Hum. Genet.* **6**, 350 (1998).
6. The optimal study population would consist of volunteers of known fertility, but difficulties acquiring testicular material make this impractical. Instead, we analyzed males who were seeing a urologist for reasons unrelated to meiotic defects. Specifically, we obtained testicular biopsies from men who were being examined for obstructive azoospermia because of a previous vasectomy ($n = 6$), cystic fibrosis ($n = 3$), unilateral

The Structure of Haplotype Blocks in the Human Genome

Stacey B. Gabriel,¹ Stephen F. Schaffner,¹ Huy Nguyen,¹
 Jamie M. Moore,¹ Jessica Roy,¹ Brendan Blumenstiel,¹
 John Higgins,¹ Matthew DeFelice,¹ Amy Lochner,¹
 Maura Faggart,¹ Shau Neen Liu-Cordero,^{1,2} Charles Rotimi,³
 Adebawale Adeyemo,⁴ Richard Cooper,⁵ Ryk Ward,⁶
 Eric S. Lander,^{1,2} Mark J. Daly,¹ David Altshuler^{1,7*}

Haplotype-based methods offer a powerful approach to disease gene mapping, based on the association between causal mutations and the ancestral haplotypes on which they arose. As part of The SNP Consortium Allele Frequency Projects, we characterized haplotype patterns across 51 autosomal regions (spanning 13 megabases of the human genome) in samples from Africa, Europe, and Asia. We show that the human genome can be parsed objectively into haplotype blocks: sizable regions over which there is little evidence for historical recombination and within which only a few common haplotypes are observed. The boundaries of blocks and specific haplotypes they contain are highly correlated across populations. We demonstrate that such haplotype frameworks provide substantial statistical power in association studies of common genetic variation across each region. Our results provide a foundation for the construction of a haplotype map of the human genome, facilitating comprehensive genetic association studies of human disease.

Variation in the human genome sequence plays a powerful but poorly understood role in the etiology of common medical conditions. Because the vast majority of heterozygosity in the human population is attributable to common variants and because the evolutionary history of common human diseases (which determined the allele spectrum for causal alleles) is not yet known, one promising approach is to comprehensively test common genetic variation for association to medical conditions (1–3). This approach is increasingly practical because 4 million (4, 5) of the estimated 10 million (6) common single nucleotide polymorphisms (SNPs) are already known.

In designing and interpreting association studies of genotype and phenotype, it is necessary to understand the structure of haplotypes in the human genome. Haplotypes are the particular combinations of alleles observed in a pop-

ulation. When a new mutation arises, it does so on a specific chromosomal haplotype. The association between each mutant allele and its ancestral haplotype is disrupted only by mutation and recombination in subsequent generations. Thus, it should be possible to track each variant allele in the population by identifying (through the use of anonymous genetic markers) the particular ancestral segment on which it arose. Haplotype methods have contributed to the identification of genes for Mendelian diseases (7–9) and, recently, disorders that are both common and complex in inheritance (10–12). However, the general properties of haplotypes in the human genome have remained unclear.

Many studies have examined allelic associations [also termed “linkage disequilibrium” (LD)] across one or a few gene regions. These studies have generally concluded that linkage disequilibrium is extremely variable within and among loci and populations [reviewed in (13–15)]. Recently, examination of a higher density of markers over contiguous regions (16–18) suggested a surprisingly simple pattern: blocks of variable length over which only a few common haplotypes are observed punctuated by sites at which recombination could be inferred in the history of the sample. In one segment of the major histocompatibility complex (MHC) on chromosome 6, it has been directly demonstrated that “hotspots” of meiotic recombination coincided with boundaries between such blocks (17). These studies suggested a model for human haplotype structure but left many questions unanswered. First, how much of the hu-

varicocele ($n = 3$), or after surgery for a testicular tumor ($n = 2$). In all cases, histological examination indicated normal spermatogenesis. We analyzed karyotypes and conducted STS-based assays for Yq microdeletions on eight individuals and detected no abnormalities. Seminiferous tubules were prepared for immunolocalization and fluorescence in situ hybridization (FISH) as described (5). We conducted standard immunostaining (23) with CREST antiserum to detect kinetochores and antibodies against MLH1 and either of two SC proteins, SCP1 and SCP3. Pachytene cells were identified on a Zeiss epifluorescence microscope, and images were captured with an Applied Imaging Quips Pathvision System. For subsequent FISH, we used paint probes (Vysis) to identify chromosomes 1, 16, 21, and 22. Cells with previously captured images were relocated on the microscope, and new images were captured. We attempted to analyze 50 cells per individual, scoring the number of MLH1 foci per autosomal bivalent and the total number of foci per autosomal complement. The XY bivalent was excluded, as it desynapses before the autosomes. In virtually all cells analyzed, at least one MLH1 focus was present on each autosome. Thus, special mechanisms to segregate achiasmate chromosomes, a feature of meiosis in organisms such as *Drosophila* (24), are unlikely to be an important component of human male meiosis.

7. D. A. Laurie, M. A. Hulten, *Ann. Hum. Genet.* **49**, 189 (1985).
8. J. Yu et al., *Am. J. Hum. Genet.* **59**, 1186 (1996).
9. B. F. J. Manly, *Randomization, Bootstrap and Monte Carlo Methods in Biology* (Chapman and Hall/CRC, New York, 1997).
10. M. Hattori et al., *Nature* **405**, 311 (2000).
11. I. Dunham et al., *Nature* **402**, 489 (1999).
12. A. Lynn, C. Kashuk, A. Chakravarti, unpublished data.
13. N. E. Morton, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7474 (1991).
14. MicroMeasure 3.3 (25) was used to measure SC lengths of individual chromosomes. Because of cell-cell variation, we compared only SCs measured in the same cell. SCs were considered equal if their lengths (in micrometers) were within 10% of each other; they were considered unequal if the difference exceeded 10%. For SCs of chromosomes 21 and 22, $22 > 21$ in 50 cells and $21 > 22$ in 1 cell. For SCs of chromosomes 16 and 19, $16 > 19$ in 9 cells, $16 = 19$ in 35 cells, and $19 > 16$ in 8 cells.
15. K. E. Koehler, J. P. Cherry, A. Lynn, P. A. Hunt, T. J. Hassold, unpublished data.
16. S. Stack, *J. Cell Sci.* **71**, 159 (1984).
17. M. Sym, G. S. Roeder, *Cell* **79**, 283 (1994).
18. S. L. Page, R. S. Hawley, *Genes Dev.* **15**, 3130 (2001).
19. A. T. C. Carpenter, in *Genetic Recombination*, R. Kucherlapati, G. R. Smith, Eds. (American Society for Microbiology, Washington, DC, 1988), pp. 529–548.
20. S. Agarwal, G. S. Roeder, *Cell* **102**, 245 (2000).
21. S. K. Mahadevaiah et al., *Nature Genet.* **27**, 271 (2001).
22. C. Tease et al., *Am. J. Hum. Genet.*, in press.
23. L. M. Woods et al., *J. Cell Biol.* **145**, 1395 (1999).
24. R. S. Hawley, W. E. Theurkauf, *Trends Genet.* **9**, 310 (1993).
25. A. Reeves, J. Tear, MicroMeasure for Windows, version 3.3. (2000). <http://www.colostate.edu/Depts/Biology/MicroMeasure>.
26. A. Solari, *Chromosoma* **81**, 315 (1980).
27. We thank T. Ashley, H. Willard, and G. Matera for helpful suggestions. This work was funded by National Institutes of Health (NIH) grants HD21341 (to T.J.H.) and HD37502 (to P.A.H.). A.L. is supported by NIH training grant HD07518, and K.E.K. is supported by an American Cancer Society fellowship.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1071220/DC1
 SOM Text

Figs. S1 and S2
 Tables S1 to S3

25 February 2002; accepted 17 May 2002

Published online 6 June 2002;

10.1126/science.1071220

Include this information when citing this paper.

¹Whitehead/MIT Center for Genome Research, Cambridge, MA 02139, USA. ²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA. ³National Human Genome Center, Howard University, Washington, DC 20059, USA. ⁴Department of Pediatrics, College of Medicine, University of Ibadan, Ibadan, Nigeria. ⁵Department of Preventive Medicine and Epidemiology, Loyola University Medical School, Maywood, IL 60143, USA. ⁶Institute of Biological Anthropology, University of Oxford, Oxford, England OX2 6QS. ⁷Departments of Genetics and Medicine, Harvard Medical School; Department of Molecular Biology and Diabetes Unit, Massachusetts General Hospital, Boston, MA 02114, USA.

*To whom correspondence should be addressed. E-mail: altshuler@molbio.mgh.harvard.edu