Am. J. Hum. Genet. 73:5–16, 2003

REVIEW ARTICLE
Hot and Cold Spots of Recombination in the Human Genome: the Reason We Should Find Them and How This Can Be Achieved

Norman Arnheim, Peter Calabrese, and Magnus Nordborg

Molecular and Computational Biology Program, University of Southern California, Los Angeles

Introduction

There is currently a great deal of interest in how the recombination rate varies over the human genome. The primary reason for this is the excitement about using linkage disequilibrium (LD) to map and identify human disease genes (reviewed in Ardlie et al. 2002; Nordborg and Tavaré 2002; Cardon and Abecasis 2003). LD mapping (also known as "association mapping") differs from traditional linkage mapping methods in that marker-disease associations are sought in populations of unrelated individuals. Such associations reflect the long evolutionary history of the chromosomal region (or the entire genome, in case of genomewide screens), involving recombination, mutation, demographic events, and (perhaps most importantly) random reproductive success. This is in sharp contrast with the associations utilized in linkage mapping, which simply reflect recombination and segregation of DNA markers and phenotypes in a small number of meioses in pedigrees or crosses. The dependence of LD on an unknown history has two important consequences for association mapping. First, although LD is expected to decay with increasing recombination, the relationship is expected to be extremely noisy (reviewed in Nordborg and Tavaré 2002). Thus, although recent studies have identified considerable variation in LD across the genome (i.e., "haplotype blocks," see Daly et al. 2001; Patil et al. 2001; Dawson et al. 2002; Gabriel et al. 2002; Phillips et al. 2003), it is by no means obvious to what extent this reflects variation in the underlying recombination rate (Reich et al. 2002; Wang et al. 2002; Innan et al. 2003; Phillips et al. 2003). It is clear that nonuniformity of recombination rates (hot spots and cold spots) along a chromosome *could* give rise to blocks of LD, but it does not follow that variation in recombination can explain all blocks of LD, nor does it follow that all blocks of LD are caused by variation in recombination.

A second consequence of a long evolutionary history (and a correspondingly large number of meioses) is that LD decays along the chromosome over genetic distances far shorter than those used to find linkage associations when utilizing traditional mapping methods. In most studies, LD levels change over physical distances from as little as 10 kb to perhaps as much as 100 kb (Daly et al. 2001; Gabriel et al. 2002; Phillips et al. 2003). If we extrapolate from a commonly accepted genome average recombination rate of ~1 cM/Mb, this means LD levels can vary over genetic distances ranging from 0.01 to 0.1 cM. (In this review, we will use the recombination fraction and the unit of genetic distance [cM] interchangeably, since the likelihood of a double crossover event in such small intervals is negligible.) Recombination-rate estimates in existing genetic maps are very inaccurate for intervals <1 cM (Kong et al. 2002; Weber 2002). This makes it impossible to determine the extent to which variation in LD is determined by variation in recombination. In what follows, we focus in particular on experimental methods for measuring recombination over distances <1 cM and discuss their strengths and limitations. Before turning to this, we address briefly why knowing more about the pattern of recombination is both important and interesting.

The Importance of Recombination Patterns

It can be argued that, from the point of view of LD mapping, it does not matter why there is LD, as long as it is there. The past is the past. This is a strong argument, but it may not be completely valid. For one thing, the cause of LD may have practical implications for marker choice in association studies. The human HapMap project (see "Electronic-Database Information" section) proposes to select markers on the basis of the pattern of LD in three samples from three populations. These markers are then to be used to map disease genes in samples from other populations, a strategy that may be more likely to work if variation in LD is due mostly to variation in the rate of recombination and not due to chance or demographic events. (We write "may" because it is entirely

Received April 9, 2003; accepted for publication April 10, 2003; electronically published May 22, 2003.

Address for correspondence and reprints: Dr. Norman Arnheim, Molecular and Computational Biology Program, University of Southern California, 835 West 37th Street, SHS 172, Los Angeles, CA 90089-1340. E-mail arnheim@usc.edu

© 2003 by The American Society of Human Genetics. All rights reserved.
0002-9297/2003/7301-0003$15.00

possible that recombination rates in some regions vary considerably among human populations.) This is an issue that will probably have to be determined by looking at patterns of LD in various samples.

Estimates of the recombination rate are necessary, however, if we wish to model the evolution of a chromosomal region on a fine scale (from hundreds to a few kb). Most algorithms for fine-scale mapping make assumptions about the local recombination rate (Liu et al. 2001; Morris et al. 2002; Niu et al. 2002). The same is true for many other kinds of evolutionary inference methods. There is, for example, great interest in determining whether selection has been acting on particular loci and in inferring human demographic history. These kinds of inference are notoriously sensitive to assumptions about local recombination rates (Wall 1999; Wall and Przeworski 2000; Wall et al. 2002).

Finally, variation in the recombination rate is interesting in and of itself. Recombination and gene conversion in humans and other mammals are poorly understood, and knowledge of their genomic pattern of variation may help us better understand these phenomena at a mechanistic level.

Characteristics of Genetic Maps: Resolution and Accuracy

The resolution of a genetic map determines how precisely we know the position at which a recombination event takes place, and it is ultimately limited by the density of polymorphisms. We cannot define the position of an event to less than the distance between the closest flanking markers. However, given that two human genomes differ, on average, once every 1–2 kb (Sachidanandam et al. 2001), marker density itself will not be limiting for many regions of the human genome.

The accuracy of any estimate of a recombination fraction over a particular interval is a function of the number of meiotic events studied. For example, if an interval has a true recombination fraction of 0.01 (1 cM), estimating recombination within this segment using a sample size of 200 meioses will leave significant doubt about the true genetic distance. On the basis of this sample size, the true genetic distance will be estimated to lie between 0.1 and 3.6 cM, a 36-fold range (95% CI). (In this review, CIs are computed by assuming the observed number of recombinants is the expected number, given the recombination rate, and then calculating the standard binomial CI.) Fundamental statistical principles show that accuracy goes up only with the square root of the number of observations, so achieving high levels of accuracy can be laborious and expensive. In the example just given, a sample size of 5,000 meioses would significantly increase the accuracy and narrow the estimate of the recombination fraction to 0.7–1.3 cM (95% CI).

There is a trade-off between resolution and accuracy. Given a constant number of meioses, increasing resolution decreases accuracy, and vice versa. Consider again the same 1-cM region as above, but with nine additional markers placed to divide it equally into 10 segments. In this case, crossovers will be located with greater resolution (100 kb). However, the accuracy of the measured genetic distance in any one of these smaller intervals must be lower, because it is based on fewer crossover events (assuming a random distribution of crossovers). For each 0.1-cM region, the accuracy of the recombination-fraction estimate reduces to 0–1.84 cM for 200 meioses and to 0.03–0.23 cM for 5,000 meioses.

The recent mapping study by Kong et al. (2002) analyzed the largest number of meioses to date (Weber 2002). They typed 5,136 microsatellite polymorphisms, using 1,257 meioses as part of an effort to map disease genes. If the markers they used were uniformly distributed over the genome, the average interval size would be 584 kb. Given the same number of meioses, the 95% CI for an estimate of genetic distance over intervals this length is 0.2–1.1 cM. These upper and lower bounds of genetic distance vary approximately sixfold, and the interval size is significantly larger than that needed to inform the overall relationship between LD and recombination. If the authors had instead achieved 100-kb resolution, the accuracy of the estimate of the genetic distance of each interval would actually have been reduced (95% CI 0.002–0.44 cM). Accurate estimates of the recombination fraction at the resolution required to study the relationship between LD and recombination (physical distances far <1 Mb) demand many more meioses.

Recombination Hot Spots

It has been known for decades that many regions of the human genome have higher (or lower) recombination fractions than would be expected on the basis of the genome average recombination rate (1 cM/Mb). These regions have all been called “hot spots” (or “cold spots”). A corollary to the definition of a hot spot (or cold spot) is that comparable cold spots (or hot spots) must also exist to maintain the genome average recombination rate. The definition of a hot spot (or cold spot) can also be relative to local recombination rates. A segment of DNA that undergoes recombination at the genome average rate can also be considered a hot spot if it is imbedded in a very “cold” region of the genome.

The term “hot spot” (or “cold spot”) has been used to describe chromosomal segments that vary considerably in size (from hundreds of base pairs to tens of megabases). Hot spots (and cold spots) of recombination have been observed in all experimental organisms that have been studied in any detail (reviewed in Lichten and Goldman 1995; Petes 2001; Nachman 2002).

Hot spots of recombination in humans and other mam-

mals have been discovered inadvertently during linkage studies or inferred from patterns of LD. The recombination hot spots in the human pseudoautosomal region (~2.6 Mb [Rouyer et al. 1986]), the DMD locus (60 kb [Grimm et al. 1989]), the phosphoglucomutase gene (58 kb [Yip et al. 1999]), the subtelomeric region of chromosome 16 (3 kb [Badge et al. 2000]), and the mouse major histocompatibility complex (<1 kb [Kobori et al. 1986; Steinmetz et al. 1986; Shiroishi et al. 1991]) are just some examples of the former. The hot spots close to the human β-globin gene (12 kb [Antonarakis et al. 1982; Chakravarti et al. 1984]) and insulin locus (20 kb [Lebo et al. 1983; Chakravarti et al. 1986]) were the first examples of the latter.

Note that the definition of a chromosome region as a hot spot does not define a priori how recombination events are distributed within that region. It is possible that crossover initiation sites of equal strength are distributed randomly throughout the genome but that the rate at which crossovers are initiated at these sites is higher throughout some chromosome regions than in others of the same size. Alternatively, discrete recombination initiation sites of equal strength might be distributed unevenly throughout the genome, creating some chromosome segments with a higher density of such "initiators." Of course, hot spot activity could also be observed, given both a random distribution of initiators and a nonrandom distribution of their strength across chromosomes.

Discovering Recombination Hot Spots by Use of Pedigree Analysis

In family studies, each offspring can provide information on the product of one paternal and one maternal meiotic event. A limited number of microsatellite and variable number of tandem repeats (VNTR) DNA polymorphisms and meioses were used in constructing genetic maps in the past. The result was (on average) barely sufficient to allow accurate estimates of the recombination fraction between markers several Mb apart. This was not an issue, since the primary objective of these mapping studies was getting the marker order right. The ordered maps of DNA polymorphisms that came from such maps (Donis-Keller et al. 1987; Buetow et al. 1994; Murray et al. 1994; Dib et al. 1996; Robinson 1996; Broman et al. 1998; Mahtani and Willard 1998; Yu et al. 2001; Kong et al. 2002) contributed greatly to the cloning of disease genes (Botstein et al. 1980).

In spite of the low-resolution nature of the above mapping studies, pedigree analysis has accurately defined numerous large chromosomal regions with cM:Mb ratios that vary by up to 100-fold (Yu et al. 2001). On the other hand, quite dramatic hot spots (even recombination rates thousands of times greater than the genome average) might be completely missed in low-resolution analyses. Imagine a linkage study showing that a 1-Mb interval between two markers undergoes recombination at approximately two times the genome average rate (note that only a study using far more meioses than is typical would have the statistical power to document excess recombination of this magnitude). A twofold increase in cM:Mb ratio could be explained if recombination initiators are randomly positioned throughout this interval and the initiation events are two times more likely than the genome average rate. On the other hand, 999,000 bp recombining at the genome average rate and a 1-kb segment recombining at a rate 1,000 times greater is also consistent with the data. In this case, 999,000 bp would contribute 0.999 cM of genetic distance. The remaining 1,000 bp would contribute 1 cM (1,000 cM/Mb × .001 Mb). Could we distinguish between these two explanations? The answer is "perhaps."

Table 1 shows the number of meioses necessary to reject (with 75% power) the genome average recombination rate for an interval as both the width of the interval studied and the intensity (multiples of the genome average) of the single hot spot varies. The hot spot (assumed to be 1 kb in width), is considered to be completely contained in the interval. In the case of the 1-Mb region containing a 1-kb hot spot with 1,000 times the genome average recombination rate, detection could be achieved with ~1,000 meioses. However, a hot spot with only twice the genome average recombination rate would not be detected, even if the number of meioses were 100,000.

By linearly adjusting the recombination rate, one may also use the table to consider hot spots of different sizes or intervals with more than one hot spot. For example, a 0.5-kb hot spot with a recombination rate 20 times the genome average corresponds to the "10 × " column in the table, or two 1-kb hot spots in the same interval, each with recombination rate 50 times the genome average, correspond to the "100 × " column.

The computations in table 1 are for a single interval

Table 1

The Number of Meioses Necessary to Reject the Genome Average Recombination Rate of an Interval with 75% Power

Interval Width (kb)	Hot-Spot Recombination Rate			
	2 ×	10 ×	100 ×	500 ×
1	>100,000	39,500	1,500	500
10	>100,000	>100,000	4,000	1,000
20	>100,000	>100,000	4,500	1,000
100	>100,000	>100,000	6,500	1,000
1,000	>100,000	>100,000	10,500	1,000

Note.—The hot spot is assumed to be localized to a 1-kb portion of the interval. The columns correspond to the hot-spot recombination rate given as multiples of the genome average rate.

with a single hot spot, but we are interested in testing many intervals. If one were to test, for example, 100 intervals, a multiple-tests correction must be applied to evaluate whether any of them are significantly greater than the genome average. In this case, the number of required meioses listed in table 1 must increase by approximately twofold (preserving 75% power; exact numbers not shown).

To detect a hot spot with width ~1 kb, it is best to resolve the location of the recombination events as narrowly as possible, despite the loss of accuracy in estimating the recombination rate. The reason is that the signal from a narrow hot spot will be diluted in a large interval.

As an example of how resolution and sample size are interrelated, imagine a 1,000-kb region that we can consider either as one interval or divided into 100 10-kb intervals. If the hot spot has width 1 kb and recombination rate 100 times the genome average, detecting it requires 10,500 meioses in the one interval scenario, whereas the 100-interval scenario, even after application of the multiple-tests correction, requires only 7,000 meioses.

In addition to testing whether a hot spot exists within a particular interval, one could also test whether the recombination rate is homogeneous across an interval. It is possible that a region would have multiple hot spots (some perhaps too weak to be detected by current methods) and multiple cold spots. A χ^2 analysis, using the number of events observed in each interval, would be appropriate to test for significance.

Sperm Typing Can Improve the Resolution and Accuracy of Recombination Estimates

In 1988, a PCR method was developed that made it possible to genotype single diploid and haploid cells at the DNA level (Li et al. 1988). This led to a new approach for studying crossing over called "single-sperm typing" (Cui et al. 1989). Sperm typing can now be applied to either single gametes or total-sperm DNA. Studying recombination between X-chromosomes is obviously not feasible using male gametes.

Single-Sperm Typing

A single human sperm cell represents one of the four products of a single meiotic event at which ~50 meiotic exchanges (MLH1 foci) occur among the 23 pairs of chromosomes (Lynn et al. 2002). A semen sample contains hundreds of millions of sperm and thus represents a large sample of the meiotic crossover events that occurred in the donor prior to sperm collection. Because a large number of meiotic products from a single individual can be studied, this technique has the capability of estimating recombination with high accuracy.

The principle of single-sperm typing is simple (figs. 1*A* and 1*B*). Individual gametes are isolated manually (Li et al. 1988; Lien et al. 1993) or by fluorescence-activated cell sorting (Cui et al. 1989). In its simplest form, PCR is performed on each lysed sperm, by use of primer pairs flanking two informative polymorphic markers that bound the interval of interest. By analyzing the genotype at each locus in many individual sperm, the phase of the two markers (AI, aI or Ai, ai) in the donor can be deduced and individual sperm categorized as having a recombinant or a nonrecombinant haplo-

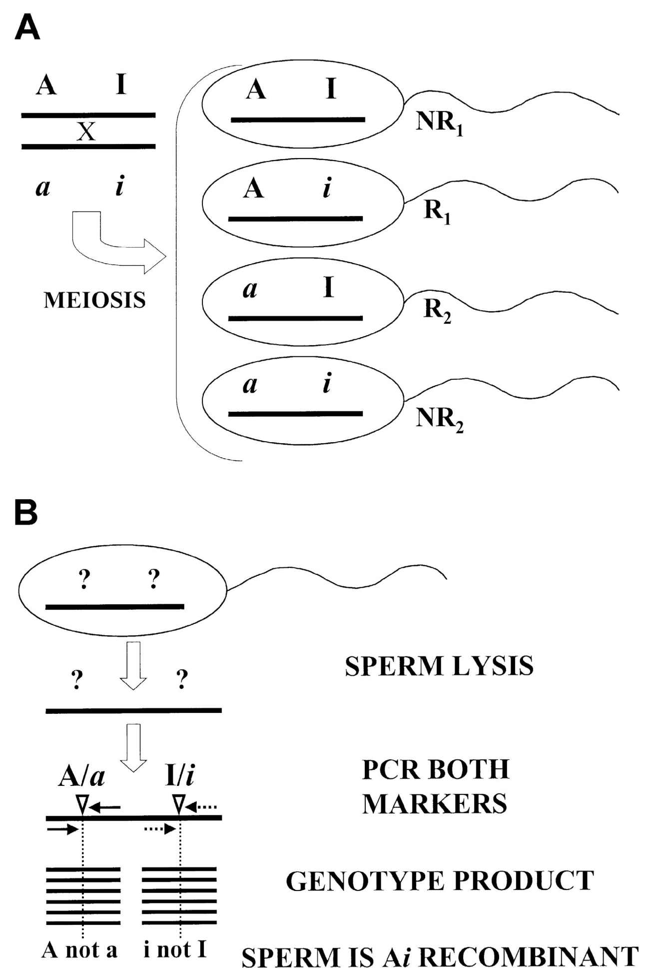

Figure 1 Basic principle of single-sperm typing. *A,* The sperm donor is heterozygous at both the A/*a* and I/*i* loci. The four possible sperm haplotypes with respect to these two loci are shown (NR = nonrecombinant; R = recombinant). *B,* Sperm lysis followed by simultaneous PCR, using primers specific for the A/*a* and I/*i* loci. Separate aliquots of the PCR product can be typed for the A/*a* or I/*i* markers, thereby determining the haplotype that, in this case, is A*i*.

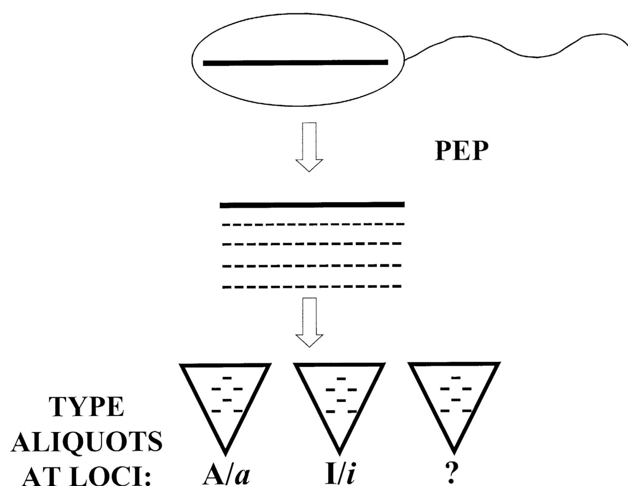

Figure 2 Basic principle of PEP. To make multiple copies of a single-sperm genome, individual sperm are lysed, and a collection of all possible 15-nt-long primers is annealed to the DNA and extended multiple times (usually 50–60) by use of *Taq* DNA polymerase. The dashed lines show the primer extension products after several rounds and indicate that multiple copies of each genomic sequence can be produced. Aliquots of the product are placed in tubes (*inverted triangles*) and used for genotyping.

type. Significantly more than two loci (in one case, 17 and, in the other, 32) have been amplified from a single sperm (Lien et al. 2000; Pramanik and Li 2002). The single-sperm typing method, including technical details and possible errors (Li et al. 1991; Cullen and Carrington 2001; Lien et al. 2002), has been reviewed elsewhere.

Single-sperm typing made it possible for the first time to measure human recombination fractions in a single individual (male), because the large number of sperm available from a single donor enables a high level of accuracy to be achieved. Family studies, on the other hand, necessarily generate a population average recombination fraction (male, female, or both sexes) for any interval, because data from different families must be pooled to obtain a sample size for even a minimal level of accuracy.

Single-sperm typing has been used to ask very specific questions about human meiotic recombination that are difficult, if not impossible, to study using pedigree analysis. This includes meiotic drive (Williams et al. 1993; Leeflang et al. 1996; Takiyama et al. 1997, 1999; Grewal et al. 1999; Girardet et al. 2000; Crouau-Roy and Clayton 2002), the effect of specific chromosomal abnormalities on crossing over (Brown et al. 1998; Shi et al. 2001), variation in crossover frequency among different individuals (Yu et al. 1996; Lien et al. 2000; Cullen et al. 2002), and the effect of age on recombination (Shi et al. 2002). However, most relevant to this review are those single-sperm typing studies aimed at characterization of chromosome regions with unusually high recombination activity (Day et al. 1994; Hubert et al. 1994; Schmitt et al. 1994; Lien et al. 2000; Cullen et al. 2002; Schneider et al. 2002; Shi et al. 2002).

Locating Recombination Hot Spots by Single-Sperm Typing

A useful and economical method for locating recombination hot spots makes use of a simple bisection strategy (Hubert et al. 1994). Individual sperm (fig. 2) are first subject to whole-genome amplification (WGA) by multiple rounds of primer extension, using random sequence 15-base oligonucleotides (Zhang et al. 1992). The goal is to produce as many copies of the sperm's genome as possible so as to allow genotyping at any chosen genetic locus. Primer extension preamplification (PEP) can produce an average of 30–60 copies of a single haploid genome with a high probability that a small aliquot of the total PEP product will have at least one whole genome available for typing a marker (Zhang et al. 1992). A new approach to WGA (Dean et al. 2002) may make it possible to produce tens of thousands of copies of a single sperm's genome.

Two aliquots of the PEP product are taken from each sperm and are used to type a pair of markers flanking any interval of interest. If this interval is 1 Mb in size, for example, then 1% of the sperm are expected to have a recombinant haplotype for these two markers.

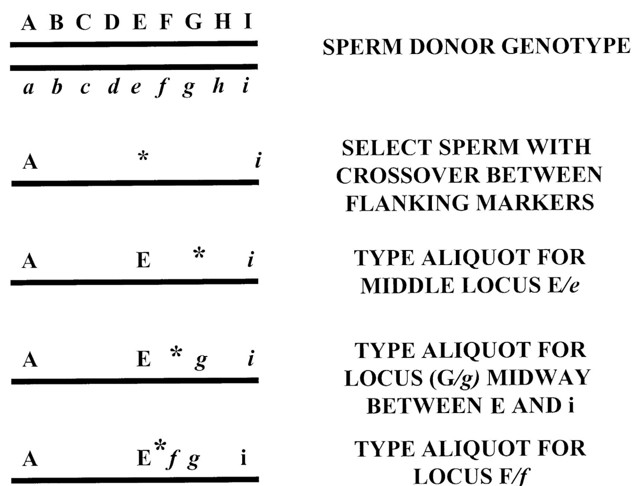

Figure 3 Single-sperm bisection strategy. The genotype of the sperm donor at nine loci (A/*a* through I/*i*) is shown at the top. After PEP, sperm are identified that have a crossover between the A/*a* and I/*i* loci. Successive PEP aliquots are taken for additional typings and, at each step, the marker in the middle of the smallest interval (*) in which the crossover could have occurred is examined until the desired resolution is achieved.

Once the recombinant sperm are identified, more aliquots of each sperm's PEP product are taken to type additional markers within the 1-Mb region (fig. 3). To achieve a resolution of 100 kb, for example, the genotype of nine additional markers from each sperm must be evaluated. However, nine additional PEP aliquots from each sperm are not required. In this example, the bisection strategy requires a maximum of only four typings to locate the 100-kb subinterval in which the crossover event occurred. In general, if there are n interior markers, a maximum of $1 + [\log_2 n]$ typings are required to specify the adjacent markers of the recombination event. Although PEP may produce only 30–60 copies of the original sperm genome, it does not necessarily result in an inability to achieve the highest possible resolution. Consider the extreme case where a 10-Mb interval is to be analyzed. The expected fraction of recombinant sperm is 10%. Even if informative markers were spaced every 1 kb (10,000 total markers), the bisection strategy would require only a maximum of 14 additional typings of each sperm (besides the two originally needed to define the sperm as recombinant in the 10-Mb interval) to identify the 1-kb interval in which crossing over took place. Although very high resolution can be achieved with the bisection strategy, the accuracy, of course, will depend on how many single sperm are examined for each interval.

The largest single-sperm typing study to date (Cullen et al. 2002) measured recombination events across a 3.3-Mb interval in the human major histocompatibility complex (MHC), known previously through classical linkage studies to harbor recombination hot spots (reviewed in Carrington 1999). Haplotypes were obtained for ~20,000 single sperm from 12 different individuals. Forty-eight different polymorphic simple tandem repeat markers were used to define 30 intervals. A total of 325 recombinants could be localized to within these intervals, 8 of which were <100 kb in length.

Six of the 30 intervals were identified as having hotspot activity. Three of them showed highly statistically significant increases in cM:Mb ratios (range 2.5-fold–5.2-fold) greater than that observed for the whole HLA. Note that the fold increase in cM:Mb ratio was not calculated relative to the genome as a whole but to the total rate over the 3.3-kb MHC (0.49 cM/Mb). It is intriguing that the authors found that donors varied in the pattern of recombination within the MHC. Comparison of MHC-identical sib pairs with nonidentical donor pairs suggested that variation in recombination pattern might be related to the donor's particular MHC type.

Recombination Detection Using Total-Sperm DNA

An alternative sperm-typing approach uses total-sperm DNA and allele-specific PCR to study recombination. The method has been applied to both equal and unequal crossing over and gene conversion (Hogstrand and Bohme 1994; Tusie-Luna and White 1995; Jeffreys et al. 1998*a,* 2001; Han et al. 2000; Guillon and de Massy 2002). Using allele-specific PCR (reviewed in Ugozzoli and Wallace 1991), recombinant molecules in total-sperm DNA are selectively amplified. Because recombinants are identified only if they can be amplified, the largest interval in which recombination can be measured is restricted to the largest PCR product that can be produced. Practically speaking, this limits the application to intervals ~5 kb in length. This is due not only to the general difficulty of long PCR using genomic DNA but also to the fact that the recombinant target makes up a very small proportion of the total-sperm DNA. The recombination fraction in an interval 5 kb long is expected to be on the order of 5×10^{-5}, given the genome average recombination rate.

The PCR assay is designed to detect one of the two reciprocal crossover products in preference to the other crossover product and the two nonrecombinants. Four allele-specific primers form the basis for two rounds of fully nested PCR (see fig. 4) that provide the specificity required. Thus, sperm donors informative for two SNPs on each side of every studied interval must be found.

Once an assay has been developed, appropriate dilutions of total-sperm DNA are made so that no more than one recombinant molecule is present in any aliquot. The difficulty of selectively amplifying a recombinant using allele-specific PCR increases with the number of nonrecombinants that are also present. How large a number this can be without losing the specificity of the PCR must be determined empirically for each interval. To date, this number varies from as few as 1,000 to as many as 66,000 nonrecombinant genomes. At least two factors contribute to this variation. The greater the length of the DNA segment being amplified, the longer the required PCR extension time, and this is negatively correlated with allele specificity. A second factor is that allele-specific amplification is also affected by exactly which nucleotide bases are present at the polymorphic site (Kwok et al. 1990; Huang et al. 1992).

Once a single recombinant is amplified, the site of recombination can be further localized by using the product to determine the genotype at any informative polymorphisms existing in the region between the two markers that were used for the second round of PCR. A target that is initially 7 kb will be reduced (on average) to 5 kb at the end of the second round. Assuming that the sperm donor is informative for markers within the

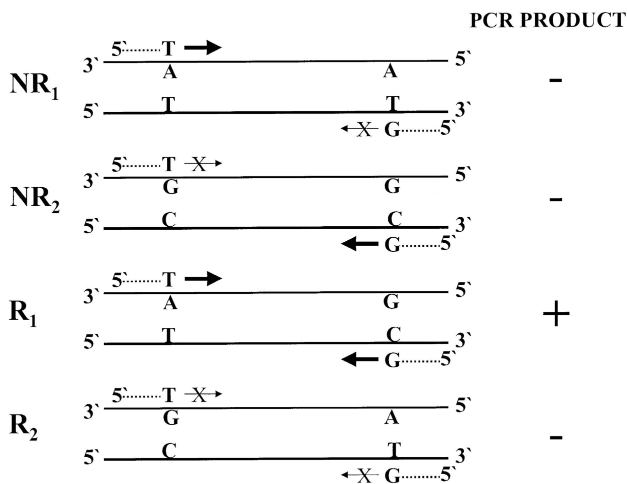

Figure 4 Principle of identification of crossover molecules by use of allele-specific PCR. The four possible products of meiosis, including both nonrecombinants (NR_1 and NR_2) and recombinants (R_1 and R_2), are shown. For each meiotic product, both strands of the double-stranded DNA molecule are shown. The interval being studied is bounded on each side by an SNP (two are required on each side, but, for simplicity, only one is shown; see text). For each SNP, one allele contains an A/T base pair and the other allele a G/C base pair. The four haplotypes, with respect to these two SNPs, are also shown. To select for the R_1 haplotype (A/T on the left and G/C on the right), PCR primers specific for A/T and G/C sites are used for PCR. The primer on the left of the interval contains a T at the 3′ end, whereas the primer on the right contains a G at its 3′ end. At each SNP, the primer is extended efficiently (*thick arrows*) on two of the four haplotypes but inefficiently (*thin arrows with an X*) on the remaining two haplotypes, because of a T opposite G base-pair mismatch. As a result, only R_1 molecules can be efficiently amplified with the two primers.

5-kb region, the positions of crossovers can be determined at higher resolution. Certain regions of the genome with higher densities of markers make it more likely that such resolution can be obtained. Similarly, regions with a low marker density may be analyzed only at lower resolution.

The accuracy of results using this method can be very high, since the number of sperm genomes that can be examined in one experiment is enormous. If each PCR aliquot queries 10,000 molecules, 100 reactions will provide a sample size of 1 million genomes. However, not all allele-specific PCR reactions perform with the same stringency. If 1,000 sperm genomes per aliquot were the upper limit for detecting a single recombinant, 1,000 aliquots must be examined to acquire the same number of recombinants for this small interval. At the genome average rate (0.001 cM/kb), a total of 50 recombinants will be expected in a 5-kb interval, if a million genomes are analyzed. This sample size will allow one to be 95% sure that the true recombination fraction is 0.0037–0.0066 cM, an approximately twofold range.

The effort needed for studying any particular interval depends on both the allele specificity of the PCR that can be achieved (determining the number of aliquots that must be tested) and the recombination fraction (determining the number of recombinants). Like family analysis and single-sperm typing, the number of recombinants needed depends on the accuracy required for testing the specific hypothesis under investigation.

Recombination Analysis in the Future

The Human Genome Project acquired vast amounts of data on the basis of the premise that exceptional insights into biology and biomedicine would be forthcoming. The benefits of this decision are already accruing. However, it would be difficult to argue for a complete analysis of the high-resolution pattern of recombination throughout all human chromosomes. Instead, a significantly more detailed understanding of the relationship between LD and recombination could be achieved by high-resolution studies on judiciously sampled chromosome regions ranging from a few kilobases to as many as 100 kilobases. The emphasis here needs to be on the term "sampled." Among relevant human sperm-typing studies to date (Day et al. 1994; Hubert et al. 1994; Schmitt et al. 1994; Tusie-Luna and White 1995; Jeffreys et al. 1998*a,* 2001; Buard et al. 2000; Han et al. 2000; Lien et al. 2000; Cullen et al. 2002; May et al. 2002; Schneider et al. 2002; Shi et al. 2002), most have been devoted to the analysis of chromosomal regions known or suspected, on the basis of classical genetic analysis (the β-δ globin intergenic region, the human MHC, and the PAR1 pseudoautosomal arm), to contain a hot spot. Together, all studied regions make up <~0.6% of the human genome. Whether these regions are representative is unknown. As discussed above, choosing regions to perform high-resolution, sperm-typing analysis on the basis of information from classical linkage studies may be misleading, since significant hot spots may go unrecognized when recombination is analyzed at low resolution. Clearly, high-resolution recombination analysis of considerable accuracy should be performed, in an unbiased fashion, on many more chromosomal segments. With respect to association studies, for example, targets could include a large sample of regions containing genes.

Understanding the forces responsible for establishing the observed patterns of high and low LD over both large and small DNA segments is necessary for understanding human genetic variation. Direct measurements of recombination at the appropriate level of resolution and accuracy will help to achieve this goal, if well-defined questions are chosen.

Advantages and Disadvantages of Sperm-Typing Methods

Each of the two sperm-typing technologies has its strengths and weaknesses for an unbiased recombination survey. Single-sperm analysis can provide accurate estimates of the recombination fraction and can pinpoint the position of those hot spots that have a sufficient impact on recombination to be detected. As shown in table 1, no more than 6,500 sperm would have to be studied to show that a 100-kb (or smaller) interval contained a hot spot with a recombination rate 100 (or more) times the genome average.

The PCR selection method can provide highly accurate estimates of the recombination fraction, but only over intervals a few kilobases in length. This method is technically demanding, because of the stringent prerequisite for having two informative polymorphisms flanking each side of every interval and of the need to achieve stringent allele specificity at each polymorphic site during amplification. If the allele specificity is inadequate, "false" recombinants can be generated, thereby raising the background and lowering the accuracy. Gene regions with an unusually high density of polymorphism—such as the MHC or duplicated and diverged repeats, in the case of studies on unequal crossing over (often 1–2 differences per 100 bp in the latter case)—provide exceptionally advantageous opportunities for finding candidate markers for the PCR selection method (Hogstrand and Bohme 1994; Tusie-Luna and White 1995; Jeffreys et al. 1998*b,* 2001; Han et al. 2000). It will undoubtedly be more difficult for other chromosomal regions. The possibility that recombination itself introduces mutations into the DNA (Strathern et al. 1995; Holbeck and Strathern 1997; Rattray et al. 2001) and thus may be correlated with polymorphism density is also a possible confounding factor in obtaining an unbiased estimate of recombination patterns.

Both sperm-typing methods require that every interval to be studied be bounded by informative markers. How informative any one sperm donor is for any particular region depends on many factors. However, it is clear that a large number of sperm samples will be needed to make sure that individuals can be found who are informative for any desired chromosome segment. Multiple individuals need to be studied for each interval to obtain the relevant (with regard to LD) population-average recombination rate, since variation in recombination rates among individuals is known (Yu et al. 1996; Han et al. 2000; Lien et al. 2000; Cullen et al. 2002).

Is There a Possible Technological Fix?

Given the present state of the art, localization of crossovers in intervals between ~100 kb and ~5 kb in length would require that a large sample size of single sperm be studied. Alternatively, using total-sperm DNA, large intervals could be analyzed in many separate experiments in which each experiment examines a single ~5-kb-long subinterval.

A hybrid approach between the two sperm-typing methods may be one option to increase throughput and reduce labor. This approach, which is based on a recent technology (Mitra and Church 1999; Mitra et al. 2003), would eliminate single-sperm isolation by allowing single molecules from total-sperm DNA to be amplified in situ on microscope slides at high density without the use of allele-specific PCR methods. The resulting PCR colonies ("polonies" [Mitra and Church 1999]) that are derived from one starting template molecule could then be easily haplotyped by conventional methods and with no need to develop highly selective allele-specific PCR protocols for each polymorphic marker. Of course, the initial (non–allele-specific) single-molecule PCR must be performed under conditions minimizing misincorporation mutations that would lead to the observation of "false recombinants." By coupling automated slide scanning and imaging technologies, the effort required to study recombination over a large number of 5-kb (or perhaps even larger [Mitra et al. 2003]) intervals would be dramatically reduced. Another possibility would depend upon improvements in allele-specific amplification. If an enzyme could be developed that would discriminate with equal efficiency between a correct match and a mismatch, regardless of the exact nature of the bases involved in mismatch formation, it would significantly reduce the time and labor needed to develop robust allele-specific amplification systems for every new interval studied, as compared with the current total-sperm DNA method.

It would be possible, at least in theory, to study recombination, range 100–10 kb, by means of pedigree analysis. An advantage would be that both male and female recombination could be analyzed. However, it would appear to be totally impracticable in the foreseeable future. The cost would be prodigious, considering the numbers of families and offspring that would have to be genotyped to accumulate the vast number of informative meioses required to permit adequate accuracy down to resolutions approaching 10 kb.

Possible Limitations of Sperm-Typing Studies

The haplotypes found in the human population are the products of recombination events during human evolution that occurred in individuals of both sexes. Sperm typing measures only male recombination. However, over evolutionary periods, every autosome spends half its time in each sex. Given that recombination, on average, is only ~1.5-fold–1.7-fold greater in females than males

(Yu et al. 2001; Kong et al. 2002) and that haplotypes have evolved over tens of thousands of generations, making conclusions about haplotype formation by studying only male crossing over may be valid. Of course, this statement is true only to the extent that the greater recombination rate in females holds true everywhere in the genome. However, substantial sex-dependent regional variation is known (reviewed in Robinson 1996). For any one segment, the bigger the difference between the sexes, the more caution that will be needed when interpreting the significance of sperm-typing data. In the mouse, for example, a 341-bp MHC recombination hot spot has been shown to function only in females (Shiroishi et al. 1990, 1991, 1995; Isobe et al. 2002).

Recombination data relevant to this question, and at the resolution required to inform questions about LD, is virtually absent in humans. At the present time, there is no way to isolate female gametes in numbers that would directly allow comparison with high-resolution sperm-typing data, although, in the mouse and cow, single-oocyte and polar-body typing has been used to measure crossing over, albeit at low resolution (Cui et al. 1992; Da et al. 1995).

In addition to crossing over and recurrent mutation, new haplotypes could also arise because of gene-conversion events. Modeling studies support this idea (Ardlie et al. 2001; Frisse et al. 2001; Przeworski and Wall 2001). The direct detection of gene-conversion events, whether associated or unassociated with crossing over, is possible in fungi where the four products of a single meiotic event are found in a single ascus. Given the present state of the art, however, gene-conversion events in humans are observed virtually only in the context of reciprocal exchanges studied at very high resolution (for examples, see Hogstrand and Bohme [1994, 1999]; Tusie-Luna and White [1995]; Jeffreys et al. [1998*a*]; Han et al. [2000]; Guillon and de Massy [2002]; Jeffreys and Neumann [2002]; Tayebi et al. [2003]), although a few attempts have been made by use of crossover-independent methods of analysis (reviewed in Martinsohn et al. 1999). A direct assessment of the significance of gene conversion unassociated with crossing over to LD will require new experimental approaches.

Electronic-Database Information

The URL for data presented herein is as follows:

HapMap, http://www.genome.gov/page.cfm?pageID=10001688

References

Antonarakis SE, Boehm CD, Giardina PJ, Kazazian HH Jr (1982) Nonrandom association of polymorphic restriction sites in the β-globin gene cluster. Proc Natl Acad Sci USA 79:137–141

Ardlie K, Liu-Cordero SN, Eberle MA, Daly M, Barrett J, Winchester E, Lander ES, Kruglyak L (2001) Lower-than-expected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion. Am J Hum Genet 69:582–589

Ardlie KG, Kruglyak L, Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 3:299–309

Badge RM, Yardley J, Jeffreys AJ, Armour JA (2000) Crossover breakpoint mapping identifies a subtelomeric hotspot for male meiotic recombination. Hum Mol Genet 9:1239–1244

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 63:861–869

Brown GM, Leversha M, Hulten M, Ferguson-Smith MA, Affara NA, Furlong RA (1998) Genetic analysis of meiotic recombination in humans by use of sperm typing: reduced recombination within a heterozygous paracentric inversion of chromosome 9q32-q34.3. Am J Hum Genet 62:1484–1492

Buard J, Shone AC, Jeffreys AJ (2000) Meiotic recombination and flanking marker exchange at the highly unstable human minisatellite CEB1 (*D2S90*). Am J Hum Genet 67:333–344

Buetow KH, Weber JL, Ludwigsen S, Scherpbier-Heddema T, Duyk GM, Sheffield VC, Wang Z, Murray JC (1994) Integrated human genome-wide maps constructed using the CEPH reference panel. Nat Genet 6:391–393

Cardon LR, Abecasis GR (2003) Using haplotype blocks to map human complex trait loci. Trends Genet 19:135–140

Carrington M (1999) Recombination within the human MHC. Immunol Rev 167:245–256

Chakravarti A, Buetow KH, Antonarakis SE, Waber PG, Boehm CD, Kazazian HH (1984) Nonuniform recombination within the human β-globin gene cluster. Am J Hum Genet 36:1239–1258

Chakravarti A, Elbein SC, Permutt MA (1986) Evidence for increased recombination near the human insulin gene: implication for disease association studies. Proc Natl Acad Sci USA 83:1045–1049

Crouau-Roy B, Clayton J (2002) Testing two possible mechanisms for gametic association in human: frozen haplotypes and segregation distortion. Genetica 114:17–23

Cui X, Gerwin J, Navidi W, Li H, Kuehn M, Arnheim N (1992) Gene-centromere linkage mapping by PCR analysis of individual oocytes. Genomics 13:713–717

Cui XF, Li HH, Goradia TM, Lange K, Kazazian HH Jr, Galas D, Arnheim N (1989) Single-sperm typing: determination of genetic distance between the G γ-globin and parathyroid hormone loci by using the polymerase chain reaction and allele-specific oligomers. Proc Natl Acad Sci USA 86:9389–9393

Cullen M, Carrington M (2001) Single-sperm typing: a rapid

alternative to family-based linkage analysis. In: Carrington M, Hoelzel AR (eds) Molecular epidemiology: the practical approach series. Oxford University Press, Oxford, United Kingdom, pp 145–179

Cullen M, Perfetto SP, Klitz W, Nelson G, Carrington M (2002) High-resolution patterns of meiotic recombination across the human major histocompatibility complex. Am J Hum Genet 71:759–776

Da Y, Jarrell VL, Wang T, Fernando RL, Wheeler MB, Lewin HA (1995) Multilocus analysis for gene-centromere mapping using first polar bodies and secondary oocytes. Genetics 139: 1091–1097

Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES (2001) High-resolution haplotype structure in the human genome. Nat Genet 29:229–232

Dawson E, Abecasis GR, Bumpstead S, Chen Y, Hunt S, Beare DM, Pabial J, et al (2002) A first-generation linkage disequilibrium map of human chromosome 22. Nature 418: 544–548

Day CE, Schmitt K, Robinson MA (1994) Frequent recombination in the human T-cell receptor β gene complex. Immunogenetics 39:335–342

Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M, Lasken RS (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci USA 99:5261–5266

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154

Donis-Keller H, Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith TP, et al (1987) A genetic linkage map of the human genome. Cell 51:319–337

Frisse L, Hudson RR, Bartoszewicz A, Wall JD, Donfack J, Di Rienzo A (2001) Gene conversion and different population histories may explain the contrast between polymorphism and linkage disequilibrium levels. Am J Hum Genet 69:831–843

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296: 2225–2229

Girardet A, McPeek MS, Leeflang EP, Munier F, Arnheim N, Claustres M, Pellestor F (2000) Meiotic segregation analysis of RB1 alleles in retinoblastoma pedigrees by use of single-sperm typing. Am J Hum Genet 66:167–175

Grewal RP, Cancel G, Leeflang EP, Durr A, McPeek MS, Draghinas D, Yao X, Stevanin G, Alnot MO, Brice A, Arnheim N (1999) French Machado-Joseph disease patients do not exhibit gametic segregation distortion: a sperm typing analysis. Hum Mol Genet 8:1779–1784

Grimm T, Muller B, Dreier M, Kind E, Bettecken T, Meng G, Muller CR (1989) Hot spot of recombination within DXS164 in the Duchenne muscular dystrophy gene. Am J Hum Genet 45:368–372

Guillon H, de Massy B (2002) An initiation site for meiotic crossing-over and gene conversion in the mouse. Nat Genet 32:296–299

Han LL, Keller MP, Navidi W, Chance PF, Arnheim N (2000) Unequal exchange at the Charcot-Marie-Tooth disease type 1A recombination hot-spot is not elevated above the genome average rate. Hum Mol Genet 9:1881–1869

Hogstrand K, Bohme J (1994) A determination of the frequency of gene conversion in unmanipulated mouse sperm. Proc Natl Acad Sci USA 91:9921–9925

——— (1999) Gene conversion can create new MHC alleles. Immunol Rev 167:305–317

Holbeck SL, Strathern JN (1997) A role for REV3 in mutagenesis during double-strand break repair in *Saccharomyces cerevisiae.* Genetics 147:1017–1024

Huang MM, Arnheim N, Goodman MF (1992) Extension of base mispairs by *Taq* DNA polymerase: implications for single nucleotide discrimination in PCR. Nucleic Acids Res 20:4567–4573

Hubert R, MacDonald M, Gusella J, Arnheim N (1994) High resolution localization of recombination hot spots using sperm typing. Nat Genet 7:420–424

Innan H, Padhukasahasram B, Nordborg M. The pattern of polymorphism on human chromosome 21. Genome Res (in press)

Isobe T, Yoshino M, Mizuno K, Lindahl KF, Koide T, Gaudieri S, Gojobori T, Shiroishi T (2002) Molecular characterization of the Pb recombination hotspot in the mouse major histocompatibility complex class II region. Genomics 80:229–235

Jeffreys AJ, Kauppi L, Neumann R (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet 29:217–222

Jeffreys AJ, Murray J, Neumann R (1998*a*) High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. Mol Cell 2:267–273

Jeffreys AJ, Neil DL, Neumann R (1998*b*) Repeat instability at human minisatellites arising from meiotic recombination. Embo J 17:4147–4157

Jeffreys AJ, Neumann R (2002) Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. Nat Genet 31:267–271

Kobori JA, Strauss E, Minard K, Hood L (1986) Molecular analysis of the hotspot of recombination in the murine major histocompatibility complex. Science 234:173–179

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high-resolution recombination map of the human genome. Nat Genet 31:241–247

Kwok S, Kellogg DE, McKinney N, Spasic D, Goda L, Levenson C, Sninsky JJ (1990) Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res 18:999–1005

Lebo RV, Chakravarti A, Buetow KH, Cheung MC, Cann H, Cordell B, Goodman H (1983) Recombination within and between the human insulin and β-globin gene loci. Proc Natl Acad Sci USA 80:4808–4812

Leeflang EP, McPeek MS, Arnheim N (1996) Analysis of meiotic segregation, using single-sperm typing: meiotic drive at the myotonic dystrophy locus. Am J Hum Genet 59:896–904

Li H, Cui X, Arnheim N (1991) Analysis of DNA sequence variation in single cells. Methods 2:49–59

Li HH, Gyllensten UB, Cui XF, Saiki RK, Erlich HA, Arnheim

N (1988) Amplification and analysis of DNA sequences in single human sperm and diploid cells. Nature 335:414–417

Lichten M, Goldman AS (1995) Meiotic recombination hotspots. Annu Rev Genet 29:423–444

Lien S, Kaminski S, Alestrom P, Rogne S (1993) A simple and powerful method for linkage analysis by amplification of DNA from single sperm cells. Genomics 16:41–44

Lien S, Szyda J, Arnheim N (2002) Single-sperm typing. In: Haines JL (ed) Current protocols in human genetics. John Wiley & Sons, New York, Unit 1.6

Lien S, Szyda J, Schechinger B, Rappold G, Arnheim N (2000) Evidence for heterogeneity in recombination in the human pseudoautosomal region: high resolution analysis by sperm typing and radiation-hybrid mapping. Am J Hum Genet 66: 557–566

Liu JS, Sabatti C, Teng J, Keats BJ, Risch N (2001) Bayesian analysis of haplotypes for linkage disequilibrium mapping. Genome Res 11:1716–1724

Lynn A, Koehler KE, Judis L, Chan ER, Cherry JP, Schwartz S, Seftel A, Hunt PA, Hassold TJ (2002) Covariation of synaptonemal complex length and mammalian meiotic exchange rates. Science 296:2222–2225

Mahtani MM, Willard HF (1998) Physical and genetic mapping of the human X chromosome centromere: repression of recombination. Genome Res 8:100–110

Martinsohn JT, Sousa AB, Guethlein LA, Howard JC (1999) The gene conversion hypothesis of MHC evolution: a review. Immunogenetics 50:168–200

May CA, Shone AC, Kalaydjieva L, Sajantila A, Jeffreys AJ (2002) Crossover clustering and rapid decay of linkage disequilibrium in the Xp/Yp pseudoautosomal gene SHOX. Nat Genet 31:272–275

Mitra RD, Butty VL, Shendure J, Williams BR, Housman DE, Church GM (2003) Digital genotyping and haplotyping with polymerase colonies. Proc Natl Acad Sci USA 100:5926–5931

Mitra RD, Church GM (1999) In situ localized amplification and contact replication of many individual DNA molecules. Nucleic Acids Res 27:e34

Morris AP, Whittaker JC, Balding DJ (2002) Fine-scale mapping of disease loci via shattered coalescent modeling of genealogies. Am J Hum Genet 70:686–707

Murray JC, Buetow KH, Weber JL, Ludwigsen S, Scherpbier-Heddema T, Manion F, Quillen J, et al (1994) A comprehensive human linkage map with centimorgan density. Cooperative Human Linkage Center (CHLC). Science 265: 2049–2054

Nachman MW (2002) Variation in recombination rate across the genome: evidence and implications. Curr Opin Genet Dev 12:657–663

Niu T, Qin ZS, Xu X, Liu JS (2002) Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. Am J Hum Genet 70:157–169

Nordborg M, Tavaré S (2002) Linkage disequilibrium: what history has to tell us. Trends Genet 18:83–90

Patil N, Berno AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR, Kautzer CR, Lee DH, Marjoribanks C, McDonough DP, Nguyen BT, Norris MC, Sheehan JB, Shen N, Stern D, Stokowski RP, Thomas DJ, Trulson MO, Vyas KR, Frazer KA, Fodor SP, Cox DR (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science 294:1719–1723

Petes TD (2001) Meiotic recombination hot spots and cold spots. Nat Rev Genet 2:360–369

Phillips MS, Lawrence R, Sachidanandam R, Morris AP, Balding DJ, Donaldson MA, Studebaker JF, et al (2003) Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. Nat Genet 33:382–387

Pramanik S, Li H (2002) Direct detection of insertion/deletion polymorphisms in an autosomal region by analyzing high-density markers in individual spermatozoa. Am J Hum Genet 71:1342–1352

Przeworski M, Wall JD (2001) Why is there so little intragenic linkage disequilibrium in humans? Genet Res 77:143–151

Rattray AJ, McGill CB, Shafer BK, Strathern JN (2001) Fidelity of mitotic double-strand-break repair in *Saccharomyces cerevisiae*: a role for SAE2/COM1. Genetics 158:109–122

Reich DE, Schaffner SF, Daly MJ, McVean G, Mullikin JC, Higgins JM, Richter DJ, Lander ES, Altshuler D (2002) Human genome sequence variation and the influence of gene history, mutation and recombination. Nat Genet 32:135–142

Robinson WP (1996) The extent, mechanism, and consequences of genetic variation, for recombination rate. Am J Hum Genet 59:1175–1183

Rouyer F, Simmler MC, Johnsson C, Vergnaud G, Cooke HJ, Weissenbach J (1986) A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes. Nature 319:291–295

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, et al (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933

Schmitt K, Lazzeroni LC, Foote S, Vollrath D, Fisher EM, Goradia TM, Lange K, Page DC, Arnheim N (1994) Multipoint linkage map of the human pseudoautosomal region, based on single-sperm typing: do double crossovers occur during male meiosis? Am J Hum Genet 55:423–430

Schneider JA, Peto TE, Boone RA, Boyce AJ, Clegg JB (2002) Direct measurement of the male recombination fraction in the human β-globin hot spot. Hum Mol Genet 11:207–215

Shi Q, Spriggs E, Field LL, Ko E, Barclay L, Martin RH (2001) Single sperm typing demonstrates that reduced recombination is associated with the production of aneuploid 24,XY human sperm. Am J Med Genet 99:34–38

Shi Q, Spriggs E, Field LL, Rademaker A, Ko E, Barclay L, Martin RH (2002) Absence of age effect on meiotic recombination between human X and Y chromosomes. Am J Hum Genet 71:254–261

Shiroishi T, Hanzawa N, Sagai T, Ishiura M, Gojobori T, Steinmetz M, Moriwaki K (1990) Recombinational hotspot specific to female meiosis in the mouse major histocompatibility complex. Immunogenetics 31:79–88

Shiroishi T, Koide T, Yoshino M, Sagai T, Moriwaki K (1995) Hotspots of homologous recombination in mouse meiosis. Adv Biophys 31:119–132

Shiroishi T, Sagai T, Hanzawa N, Gotoh H, Moriwaki K (1991) Genetic control of sex-dependent meiotic recombination in the major histocompatibility complex of the mouse. Embo J 10:681–686

Steinmetz M, Stephan D, Fischer Lindahl K (1986) Gene organization and recombinational hotspots in the murine major histocompatibility complex. Cell 44:895–904

Strathern JN, Shafer BK, McGill CB (1995) DNA synthesis er-

rors associated with double-strand-break repair. Genetics 140: 965–972

Takiyama Y, Sakoe K, Amaike M, Soutome M, Ogawa T, Nakano I, Nishizawa M (1999) Single sperm analysis of the CAG repeats in the gene for dentatorubral-pallidoluysian atrophy (DRPLA): the instability of the CAG repeats in the DRPLA gene is prominent among the CAG repeat diseases. Hum Mol Genet 8:453–457

Takiyama Y, Sakoe K, Soutome M, Namekawa M, Ogawa T, Nakano I, Igarashi S, Oyake M, Tanaka H, Tsuji S, Nishizawa M (1997) Single sperm analysis of the CAG repeats in the gene for Machado-Joseph disease (MJD1): evidence for non-Mendelian transmission of the MJD1 gene and for the effect of the intragenic CGG/GGG polymorphism on the intergenerational instability. Hum Mol Genet 6:1063–1068

Tayebi N, Stubblefield BK, Park JK, Orvisky E, Walker JM, LaMarca ME, Sidransky E (2003) Reciprocal and nonreciprocal recombination at the glucocerebrosidase gene region: implications for complexity in Gaucher disease. Am J Hum Genet 72:519–534

Tusie-Luna MT, White PC (1995) Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase gene) and CYP21P involve different mechanisms. Proc Natl Acad Sci USA 92:10796–10800

Ugozzoli L, Wallace RB (1991) Allele-specific polymerase chain reaction. Methods 2:42–48

Wall JD (1999) Recombination and the power of statistical tests of neutrality. Genet Res 74:65–79

Wall JD, Andolfatto P, Przeworski M (2002) Testing models of selection and demography in *Drosophila simulans*. Genetics 162:203–216

Wall JD, Przeworski M (2000) When did the human population size start increasing? Genetics 155:1865–1874

Wang N, Akey JM, Zhang K, Chakraborty R, Jin L (2002) Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. Am J Hum Genet 71:1227–1234

Weber JL (2002) The Iceland map. Nat Genet 31:225–226

Williams C, Davies D, Williamson R (1993) Segregation of δ F508 and normal CFTR alleles in human sperm. Hum Mol Genet 2:445–448

Yip SP, Lovegrove JU, Rana NA, Hopkinson DA, Whitehouse DB (1999) Mapping recombination hotspots in human phosphoglucomutase (PGM1). Hum Mol Genet 8:1699–1706

Yu A, Zhao C, Fan Y, Jang W, Mungall AJ, Deloukas P, Olsen A, Doggett NA, Ghebranious N, Broman KW, Weber JL (2001) Comparison of human genetic and sequence-based physical maps. Nature 409:951–953

Yu J, Lazzeroni L, Qin J, Huang MM, Navidi W, Erlich H, Arnheim N (1996) Individual variation in recombination among human males. Am J Hum Genet 59:1186–1192

Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, Arnheim N (1992) Whole genome amplification from a single cell: implications for genetic analysis. Proc Natl Acad Sci USA 89: 5847–5851