belts, may thus be one of the most promising places to search for vestiges of early life on Earth.

References and Notes

1. W. E. Krumbein, C. A. Urzi, C. Gehrman, *Geomicrobiol. J.* **9**, 139 (1991).
2. K. A. Ross, R. V. Fisher, *Geology* **14**, 571 (1986).
3. I. H. Thorseth, H. Furnes, M. Heldal, *Geochim. Cosmochim. Acta* **56**, 845 (1992).
4. I. H. Thorseth, H. Furnes, O. Tumyr, *Chem. Geol.* **119**, 139 (1995).
5. H. Staudigel, R. A. Chastain, A. Yayanos, W. Bourcier, *Chem. Geol.* **126**, 147 (1995).
6. H. Staudigel *et al.*, *Earth Planet. Sci. Lett.* **164**, 233 (1998).
7. J. R. Rogers, P. C. Bennett, *Chem. Geol.* **203**, 91 (2004).
8. I. H. Thorseth, T. Torsvik, H. Furnes, K. Muehlenbachs, *Chem. Geol.* **126**, 137 (1995).
9. M. R. Fisk, S. J. Giovannoni, I. H. Thorseth, *Science* **281**, 978 (1998).
10. T. Torsvik, H. Furnes, K. Muehlenbachs, I. H. Thorseth, O. Tumyr, *Earth Planet. Sci. Lett.* **162**, 165 (1998).
11. H. Furnes, H. Staudigel, *Earth Planet. Sci. Lett.* **166**, 97 (1999).
12. H. Furnes *et al.*, *Geochem. Geophys. Geosyst.* **2**, 2000GC000150 (2001).
13. N. R. Banerjee, K. Muehlenbachs, *Geochem. Geophys. Geosyst.* **4**, 2002GC000470 (2003).
14. H. Furnes, K. Muehlenbachs, T. Torsvik, I. H. Thorseth, O. Tumyr, *Chem. Geol.* **173**, 313 (2001).
15. M. J. de Wit, R. A. Hart, J. Hart, *J. Afr. Earth Sci.* **6**, 681 (1987).
16. C. E. J. de Ronde, M. J. De Wit, *Tectonics* **13**, 983 (1994).
17. M. M. Tice, B. C. Bostick, D. R. Lowe, *Geology* **32**, 37 (2004).
18. The x-ray mapping was carried out using a JEOL JXA-8900R microprobe, with an accelerating voltage of 15 kV and probe current of 3×10^{-8} on an iridium-coated thin section.
19. Carbon stable-isotope analyses of disseminated carbonates in glassy and crystalline whole rocks were performed by pouring 100% phosphoric acid onto rock powders under vacuum and analyzing the exsolved CO_2 on a Finnigan MAT 252 mass spectrometer (*14*). The data are reported in the usual delta notation with respect to the Vienna Pee Dee belemnite standard.
20. K. Nakamura, Y. Kato, *Geochim. Cosmochim. Acta* **67** (suppl. 1), A326 (2003).
21. H. Furnes, K. Muehlenbachs, in *Ophiolites in Earth's History*, Y. Dilek, P. T. Robinson, Eds. (Special Publication 218, Geological Society of London, London, 2003), pp. 415–426.
22. M. Lopez- Martinez, D. York, J. A. Hanes, *Precambrian Res.* **57**, 91 (1992).
23. S. L. Kamo, D. W. Davis, *Tectonics* **13**, 167 (1994).
24. J. Dann, *S. Afr. J. Geol.* **103**, 47 (2000).
25. S. E. Hoffman, M. Wilson, D. S. Stakes, *Nature* **321**, 55 (1986).
26. M. J. de Wit, R. A. Hart, *Lithos* **30**, 309 (1993).
27. M. J. Russel, A. J. Hall, *J. Geol. Soc. London* **154**, 377 (1997).
28. B. Rasmussen, *Nature* **405**, 676 (2000).

29. Comments by A. Lepland, B. Robins, T. Chacko, and three anonymous reviewers improved the manuscript. We thank O. Levner for help with carbon isotope analyses and S. Matveev for help with the electron microprobe. Thanks to F. Daniels for hospitality and the Mpumalanga Parks Board for access during field work. Financial support to carry out this study was provided by the Norwegian Research Council, the National Sciences and Engineering Research Council of Canada, the U.S. National Science Foundation, the Agouron Institute, and the National Research Foundation of South Africa.

Supporting Online Material
www.sciencemag.org/cgi/content/full/304/5670/578/DC1
Fig. S1

21 January 2004; accepted 22 March 2004

The Fine-Scale Structure of Recombination Rate Variation in the Human Genome

Gilean A. T. McVean,[1*†] Simon R. Myers,[1*] Sarah Hunt,[2] Panos Deloukas,[2] David R. Bentley,[2] Peter Donnelly[1]

The nature and scale of recombination rate variation are largely unknown for most species. In humans, pedigree analysis has documented variation at the chromosomal level, and sperm studies have identified specific hotspots in which crossing-over events cluster. To address whether this picture is representative of the genome as a whole, we have developed and validated a method for estimating recombination rates from patterns of genetic variation. From extensive single-nucleotide polymorphism surveys in European and African populations, we find evidence for extreme local rate variation spanning four orders in magnitude, in which 50% of all recombination events take place in less than 10% of the sequence. We demonstrate that recombination hotspots are a ubiquitous feature of the human genome, occurring on average every 200 kilobases or less, but recombination occurs preferentially outside genes.

The nature and causes of recombination rate variation in the human genome are little known. Genetic maps estimated from pedigree studies have revealed chromosome-wide and sex-specific variation in the rate of recombination (*1*, *2*) but only have resolution above megabase scales. Analyses of recombination break points and, more recently, cross-over events in sperm have demonstrated the presence of recombination hotspots in a small number of genomic locations; the human leukocyte antigen (HLA) region (*3*, *4*), the minisatellite MS32 (*5*), the pseudoautosomal region (*6*), and the β-globin gene (*7*). Hotspots are a feature described in yeast and some prokaryotes (*8*) but not documented in other eukaryotes such as flies and worms. However, recent observations of a block-like structure to patterns of human linkage disequilibrium (*9*–*11*) and correlations in homozygosity (*12*) have led to speculation that most or all human recombination occurs at hotspots (*9*, *13*). But, in fact, it is not known how widespread hotspots are in the human genome, neither is the magnitude of rate differences, nor the physical scales over which this occurs, known.

An understanding of the genomic landscape of human recombination rate variation would facilitate the efficient design and analysis of disease association studies and greatly improve inferences from polymorphism data about selection and human demographic history. Fine-scale recombination rate estimates would also provide a new route to understanding the molecular mechanisms underlying human recombination. Current approaches cannot provide this information: Pedigree studies do not have the required resolution, whereas sperm analyses can only detect recombination rate variation in males and are impracticable for studies on chromosomal scales. Here, we present and validate a coalescent-based method for estimating recombination rate variation at kilobase scales from large surveys of single-nucleotide polymorphism (SNP) variation. With the advent of genome-wide diversity studies, such as the HapMap (*14*), this will allow the construction of the first fine-scale genetic map in humans.

Patterns of genetic diversity and LD are shaped by many factors (*15*), mutation, recombination, selection, population demography, and genetic drift. Typically, they display substantial stochastic variation. Extracting the signal of recombination rate variation from such data presents a challenging statistical problem (*16*). Our approach, based on an approximation to the coalescent, is motivated by recent developments in computationally intensive population genetics inference methods (*17*–*19*). Informally, we extend the composite likelihood approach of Hudson (*20*) to allow different recombination rates between each pair of SNPs and adopt a Bayesian implementation in which the prior distribution encourages short-range smoothness in estimated recombination rates and avoids over-fitting (fig. S1 and table S1).

[1]Department of Statistics, University of Oxford, Oxford OX1 3TG, UK. [2]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: mcvean@stats.ox.ac.uk

Downloaded from http://science.sciencemag.org/ on August 15, 2020

The method allows for rate estimation over various scales, from a finest scale of kilobases up to that of current genetic maps. It applies to phased or unphased genotype data, is largely insensitive to SNP ascertainment strategy, allows for repeat mutation, handles missing data, and is computationally practicable for genome-wide variation surveys (*21*). [See the supporting online material (SOM) for details.]

At a qualitative level, correlations have been observed over megabase scales between simple summaries of LD and the existing genetic map (*22*). It is less clear whether reliable fine-scale recombination rate estimates can be obtained from polymorphism data. We have validated our approach in three ways: by extensive simulation studies (see SOM) and by comparisons with independent estimates of recombination rates, both over large scales from the genetic map and over fine scales from sperm analysis.

Conventional genetic maps estimate recombination rates over megabase scales. Rate estimates at the same scales can be obtained from our population genetic method by summing local rates over equivalent intervals. We compared the genetic maps obtained by pedigree-based and population genetic methods for the long (q) arms of chromosomes 19 (*22*) and 22 (*23*) [from a European population (*24*)] (Fig. 1). There is strong agreement between the methods over the majority of the chromosomes, although some discrepancies are observed at the chromosome arm ends, particularly where marker density for the pedigree-based map is low.

Where recombination hotspots and coldspots have been previously detected in the HLA region (*3*), our novel population genetic approach obtains very similar estimates of the location and scale of recombination rate variation (Fig. 2). The population genetic analysis also provides information about recombination rate variation in females (unlike the sperm analysis). We conclude that there are no female-specific hotspots in the region and that the ones observed in males are also found in females (*25*). A further advantage of the population genetic approach is that the method can reliably estimate recombination rates in regions of low (but nonzero) recombination. For the 53-kb region between the *DNA3* and *DMB1* hotspots, we estimate the male recombination rate to be 0.08 centimorgan per megabase (or a sex-averaged rate of 0.19 cM/Mb) and confirm the presence of recombination by the detection (*26*) of at least seven historical recombination events. To estimate such a low rate accurately by counting cross-over events in sperm or pedigree analyses would require at least 250,000 informative meioses.

Conclusions about the fine-scale structure of recombination rate variation in the human genome drawn from our population genetics approach must be robust to variation in demographic history between populations, SNP spacing, SNP ascertainment, and differences among genomic regions. We have addressed the influence of such factors through simulation and the analysis of data sets that vary in population origin (and demographic influences), SNP density, SNP minor allele frequency distribution, and genomic location. Simulations show that the method is largely robust to deviations from the assumed neutral coalescent model, including population growth, population bottlenecks, and gene conversion (fig. S2 and table S1). Robustness to differences in population history is also confirmed empirically by the similarity across populations in both the details of estimated rates and the overall picture of recombination rate variation from homologous regions in a dense SNP survey of chromosome 20 (average SNP spacing of 2.3 kb) in European and African American populations (*27*) (Fig. 3 and fig. S3; the correlation between the natural log of the rate estimates between SNPs in the two populations is 0.75), and a genome-wide set of 74 regions (average SNP spacing of 5.8 kb) in European [Centre d'Etude du Polymorphism Humain (CEPH)] and African (Yoruban) populations (*10*) (Fig. 3C). The near-identical estimates of the pattern of rate variation obtained from different genomic regions with different SNP spacing (Fig. 3C) demonstrate that the patterns we observe are a general feature of the human genome and not an artifact of experimental design, nor of factors such as natural selection that affect diversity in a particular region. Similar results were also obtained by artificially thinning the chromosome 20 data set to an average SNP spacing of 4.6 kb (fig. S3). Because the data sets analyzed differ in how candidate SNPs were identified and in minor allele frequency distribution (important because it is an effect induced by SNP ascertainment and because simple measures of LD are influenced by allele frequency), we conclude that our approach is also robust to SNP ascertainment strategy, a conclusion confirmed by simulation [fig. S2 and (*28*)].

Figure 3a shows the estimated recombination rates in the chromosome 20 region for the European population, together with the pedigree-based estimate, the location of recombination hotspots with strong statistical support (see SOM) and the positions of known genes. Unlike the pedigree-based estimate, which is essentially uniform over the interval, we find evidence for unprecedented (and statistically significant – see

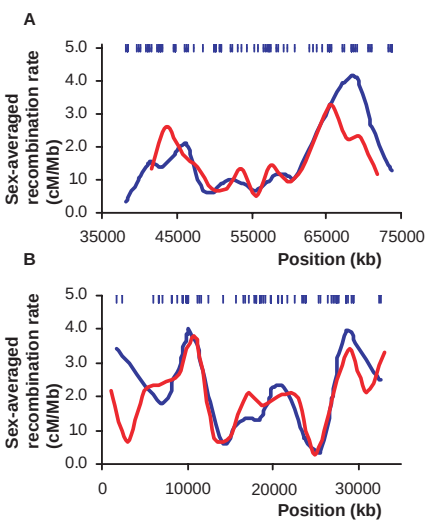

Fig. 1. Estimated recombination rate variation at a 2-Mb scale on the long arms of chromosomes 19 (**A**) and 22 (**B**). Estimates from pedigree data (*1*) (blue) and population genetic data, summed over 2-Mb intervals (red). Positions of markers in the pedigree-based map are indicated by blue dashes; SNP density for the population genetic maps is too high to represent at this resolution (1823 and 1504 SNPs for chromosomes 19 and 22, respectively). Genotype data from (*22*, *23*). To rescale population genetic estimates, N_e was estimated from a comparison of the genetic map distance across the entire chromosomal arm and the estimated population recombination rate for the same region.

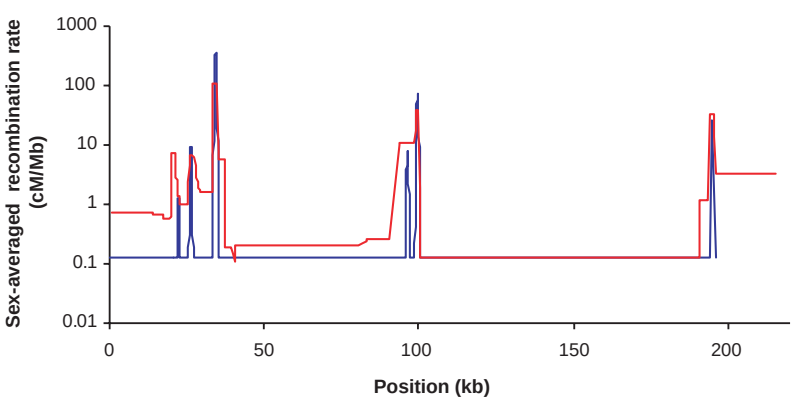

Fig. 2. Comparison between estimates of local recombination rates from population genetic data (red) and sperm analysis (blue) in the HLA region; data from (*3*). To convert the male crossing over rates to sex-averaged rates, we used the previous observation that the female crossing-over rate in this region is about four times that of males (*42*).

Downloaded from http://science.sciencemag.org/ on August 15, 2020

SOM) variation in the recombination rate, spanning four orders in magnitude (Fig. 3B). Our analysis establishes that recombination hotspots are a ubiquitous feature of the human genome, and occur on average every 200 kb or less. For example, in the European sample from chromosome 20, we find 48 positions where there is at least a fivefold increase in estimated local recombination rate and also strong statistical support ($P < 0.001$) for the presence of a hotspot (Figs. 3A and 4), which suggest the presence of at least 15,000 hotspots in the human genome. However, we also find more extended regions where the background recombination rate is relatively high. In addition to the very fine scale variation in recombination rate, an underlying fluctuation in the background rate can also be seen acting over a much larger scale. We also find that regions of low recombination tend to be more extensive than regions of high rate: The average length of regions where all SNP intervals fall in the bottom 10% of rates is 91 kb, compared with an average of 19 kb for the top 10% of rates.

The picture of recombination rate variation in chromosome 20 and the genome-wide regions is different from that previously reported in the HLA region (*3*). We can represent the nature of recombination rate variation both through the distribution of recombination rates (which is roughly exponential, Fig. 3B) and the proportion of recombination that occurs in a given fraction of the sequence (Fig. 3C; obtained by ordering SNP intervals by their estimated rate in descending value and plotting the cumulative genetic distance against the cumulative physical distance). For the chromosome 20 (*27*) and Gabriel *et al.* (*10*) genome-wide data from Europeans, we estimate that 50% of all recombination occurs in less than 10% of the sequence, whereas, in the HLA region, 80% of all recombination occurs in less than 10% of the sequence. In short, the pattern of short recombination hotspots separated by large recombination coldspots, as observed in the HLA region, does not appear to be typical of the genome as a whole. We detect slightly less recombination rate variation in the African American and Yoruban populations (50% of recombination occurs in 15% of the sequence). However, we also find from simulation that population bottlenecks, as have probably occurred in the history of European populations (*29*), increase the power to detect recombination rate variation (probably because they increase background LD and so make hotspots more prominent), which suggests that the estimate from the European population may be more accurate. For the chromosome 19 (*22*) and 22 (*23*) data sets, we detect less rate variation, a result of the considerably lower SNP density (fig. S4).

Fine-scale estimates of local recombination rate variation provide an opportunity to determine how recombination rate is influenced by genomic context. We find that recombination rates are lower in genic regions (defined by the beginning of the first exon of a gene to the end of the last) than in noncoding regions (ratio = 0.75, $P < 0.02$ from permutation). The triplet repeat gene *NCOA3* (start position at 46,769) shows a striking pattern of corresponding almost perfectly with a strong recombination coldspot (estimated sex-averaged rate = 0.01 cM/Mb). These results are in strong contrast with earlier reports of a positive correlation between gene density (or features associated with genes such as CpG islands) and recombination rate (as estimated from pedigree-based genetic maps) (*1*, *30*). The apparent contradiction may be explained if recombination hotspots are more likely to occur near genes rather than within them. Such a pattern may be expected if recombination typically occurs in regions of open chromatin (*31*), but there is selection against recom-

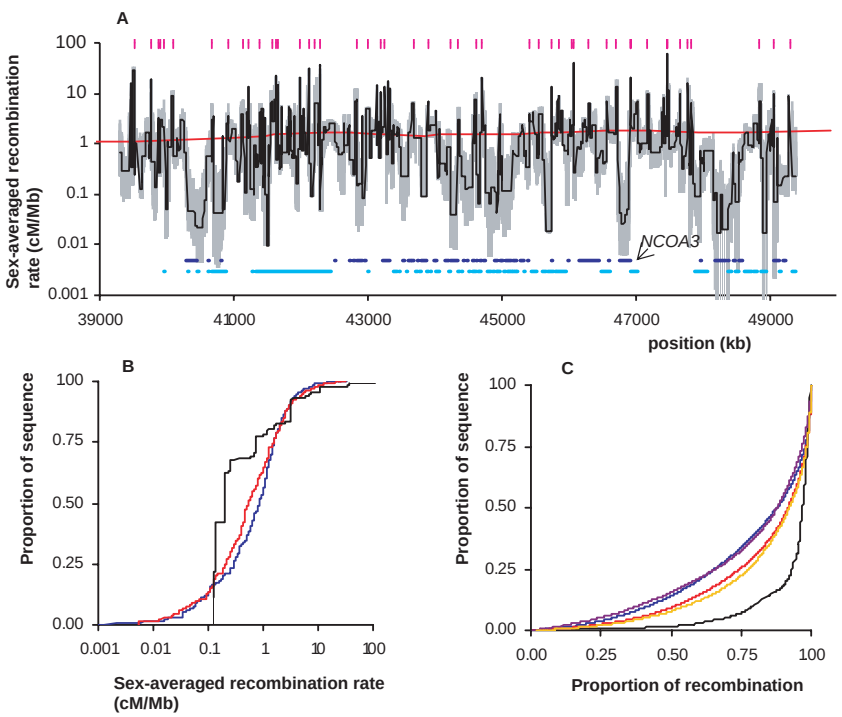

Fig. 3. Recombination rate variation. (**A**) Estimated recombination rate along a 10-Mb region of chromosome 20 in a European (UK Caucasian) sample (black line). Superimposed are the 2.5 and 97.5 percentiles of the sampling distribution for local rates (gray, see SOM for details); the position of recombination hotspots (estimated increase in rate by a factor of at least 5 over local background) with strong ($P < 0.001$) statistical support (vertical lines, see SOM); the recombination rate estimated from pedigrees (red line); and the location of genes on the plus (dark blue) and minus (light blue) strand (*43*). Note the remarkable lack of recombination around the triplet repeat *NCOA3* gene, start position 46,769 kb. (**B**) Cumulative distribution of recombination rates from the same sample, Europeans (red) and African Americans (blue), and for the HLA region (black). (**C**) Cumulative plot showing the estimated proportion of recombination occurring in a given fraction of sequence, where the SNP intervals have been ranked by decreasing recombination rate; colors as for (B). Data from (*27*). Also shown are estimates obtained from the genome-wide data of Gabriel *et al.* (*10*) for the CEPH (yellow) and African (purple) populations.

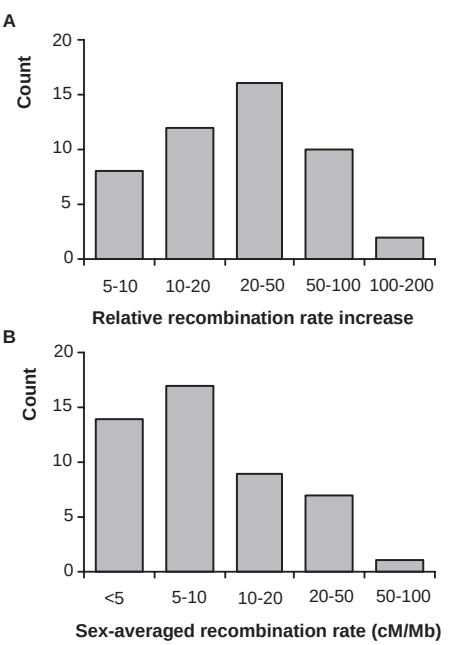

Fig. 4. Properties of recombination hotspots. (**A**) Distribution of estimated relative rate increase at the 48 recombination hotspots in the 10-Mb region of chromosome 20 for which there is strong statistical support ($P < 0.001$) in the European population. (**B**) The distribution of the maximum estimated absolute recombination rate (in cM/Mb) for the same 48 hotspots.

Downloaded from http://science.sciencemag.org/ on August 15, 2020

bination within exons, for example, if recombination is mutagenic (*32*, *33*).

The correlation between genes and recombination, although statistically significant, does not explain any appreciable amount of the variation in recombination rate. We have considered whether a variety of genomic features previously associated with recombination rate (*1*, *30*, *34*, *35*) [GC content, CpG dinucleotide frequency, poly(A)/poly(T) fraction, and the presence of $(AC)_n$ repeats] provide accurate predictors of local recombination rate. We find that no single factor, or combination, can explain more than a small fraction (less than 10%) of variation in recombination rate, when measured at the highest possible resolution (between adjacent SNPs). When estimates are obtained over increasing physical scales (by averaging over multiple SNP intervals), the correlation coefficient between GC content and recombination rate first increases (up to 50 kb), then decreases toward the value obtained when recombination rates are estimated from pedigree-based genetic maps (*35*). Rate estimates obtained over multiple SNP intervals have less variance than for individual intervals [GM (*36*)], hence, the pattern is indicative of a local relation between GC content and recombination rate. However, local GC content is only a poor predictor of local recombination rate. One possibility is that recombination directly influences local base composition, e.g., through biased gene conversion (*37*), but that recombination rates evolve much faster than local base composition (which is limited by the rate of mutation). Variation among individuals in recombination rate (*2*), polymorphisms that influence local rate (*38*), and differences in recombination hotspots between humans and other primates (*39*) all suggest that fine-scale recombination rate variation may change rapidly over evolutionary time.

Aside from inherent interest, our demonstration that recombination hotspots are common and widespread in the human genome has important implications for the interpretation of genetic maps. In particular, extrapolation of local recombination rates from pedigree-based maps would tend to overestimate the rate in most regions (see Fig. 3A). Similarly, if 50% of all recombination occurs in only 10% of the sequence (or 80% in 25%), variation in rate along existing genetic maps and between chromosomes must be largely due to differences in the density and intensity of recombination hotspots, rather than changes in the background rate (although the two may be correlated).

Fine-scale genetic maps have many applications in both medical and evolutionary biology. A fine-scale genetic map of the human genome will have a major influence on many of the key issues in the design and analysis of association mapping experiments, such as the choice of tagging SNPs, assessment of experimental power, and methods for fine-scale mapping. Fine-scale genetic maps in multiple species will also provide a powerful tool for answering fundamental questions in evolutionary biology, such as the relative influence of adaptive and purifying selection in molecular evolution (*40*) and what short-term benefits exist to sexual reproduction (*41*). Population-genetic data and the statistical methods described here enable the construction of fine-scale genetic maps in any species.

References and Notes

1. A. Kong *et al.*, *Nature Genet.* **31**, 241 (2002).
2. K. W. Broman, J. C. Murray, V. C. Sheffield, R. L. White, J. L. Weber, *Am. J. Hum. Genet.* **63**, 861 (1998).
3. A. J. Jeffreys, L. Kauppi, R. Neumann, *Nature Genet.* **29**, 217 (2001).
4. A. J. Jeffreys, A. Ritchie, R. Neumann, *Hum. Mol. Genet.* **9**, 725 (2000).
5. A. J. Jeffreys, J. Murray, R. Neumann, *Mol. Cell* **2**, 267 (1998).
6. S. Lien, J. Szyda, B. Schechinger, G. Rappold, N. Arnheim, *Am. J. Hum. Genet.* **66**, 557 (2000).
7. R. A. Smith, P. J. Ho, J. B. Clegg, J. R. Kidd, S. L. Thein, *Blood* **92**, 4415 (1998).
8. G. R. Smith, *Experientia* **50**, 234 (1994).
9. M. J. Daly, J. D. Rioux, S. F. Schaffner, T. J. Hudson, E. S. Lander, *Nature Genet.* **29**, 229 (2001).
10. S. B. Gabriel *et al.*, *Science* **296**, 2225 (2002).
11. N. Patil *et al.*, *Science* **294**, 1719 (2001).
12. D. E. Reich *et al.*, *Nature Genet.* **32**, 135 (2002).
13. D. B. Goldstein, *Nature Genet.* **29**, 109 (2001).
14. R. A. Gibbs *et al.*, *Nature* **426**, 789 (2003).
15. J. K. Pritchard, M. Przeworski, *Am. J. Hum. Genet.* **69**, 1 (2001).
16. M. P. Stumpf, G. A. McVean, *Nature Rev. Genet.* **4**, 959 (2003).
17. M. Stephens, P. Donnelly, *J. R. Stat. Soc. Ser. B* (*Stat. Method.*) **62**, 605 (2000).
18. P. Fearnhead, P. Donnelly, *Genetics* **159**, 1299 (2001).
19. N. Li, M. Stephens, *Genetics* **165**, 2213 (2003).
20. R. R. Hudson, *Genetics* **159**, 1805 (2001).
21. Analysis of 4337 SNPs genotyped in 96 individuals takes ~3 hours computer processing time on a 2.4 Ghz processor with 1 Gbyte random access memory (RAM).
22. M. S. Phillips *et al.*, *Nature Genet.* **33**, 382 (2003).
23. E. Dawson *et al.*, *Nature* **418**, 544 (2002).
24. For each chromosome arm, population genetic rate estimates (in $4N_er$ per kilobase) are rescaled by the genetic distance across the region as estimated from pedigree data to provide rate estimates in centimorgans per megabase.
25. If the recombination rate were constant in females, the sex-averaged recombination rate outside hotspots would be at least 1 cM/Mb, much higher than we estimate.
26. R. R. Hudson, N. L. Kaplan, *Genetics* **111**, 147 (1985).
27. X. Ke *et al.*, *Hum. Mol. Genet.* **13**, 577 (2004).
28. S. R. Myers, unpublished observations.
29. D. E. Reich *et al.*, *Nature* **411**, 199 (2001).
30. A. Eyre-Walker, *Proc. R. Soc. Lond. B. Biol. Sci.* **252**, 237 (1993).
31. A. Nicolas, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 87 (1998).
32. I. Hellmann, I. Ebersberger, S. E. Ptak, S. Paabo, M. Przeworski, *Am. J. Hum. Genet.* **72**, 1527 (2003).
33. D. A. Filatov, D. T. Gerrard, *Gene* **317**, 67 (2003).
34. A. Yu *et al.*, *Nature* **409**, 951 (2001).
35. S. M. Fullerton, C. A. Bernardo, A. G. Clark, *Mol. Biol. Evol.* **18**, 1139 (2001).
36. G. A. T. McVean, unpublished observations.
37. N. Galtier, G. Piganeau, D. Mouchiroud, L. Duret, *Genetics* **159**, 907 (2001).
38. A. J. Jeffreys, R. Neumann, *Nature Genet.* **31**, 267 (2002).
39. J. D. Wall, L. A. Frisse, R. R. Hudson, A. Di Rienzo, *Am. J. Hum. Genet.* **73**, 1330 (2003).
40. C. F. Aquadro, *Curr. Opin. Genet. Dev.* **7**, 835 (1997).
41. N. H. Barton, B. Charlesworth, *Science* **281**, 1986 (1998).
42. M. Cullen *et al.*, *Am. J. Hum. Genet.* **60**, 397 (1997).
43. ENSEMBL genes from Build 33 on www.ensembl.org/EnsMart

44. Thanks to G. Coop, D. Falush, J. Marchini, M. Przeworski, D. Reich, D. Sherratt, M. Stephens, C. Spencer, and three anonymous reviewers for discussion and comments on the manuscript. The work was supported by the Royal Society, the Nuffield Trust, the Wellcome Trust, NIH, and The SNP Consortium.

Supporting Online Material
www.sciencemag.org/cgi/content/full/304/5670/581/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S2
References and Notes

14 October 2003; accepted 29 March 2004

The Rise of the Rhizosolenid Diatoms

Jaap S. Sinninghe Damsté,[1]* Gerard Muyzer,[1,2] Ben Abbas,[1] Sebastiaan W. Rampen,[1] Guillaume Massé,[3] W. Guy Allard,[3] Simon T. Belt,[3] Jean-Michel Robert,[4] Steven J. Rowland,[3] J. Michael Moldowan,[5] Silvana M. Barbanti,[5,6] Frederick J. Fago,[5] Peter Denisevich,[5] Jeremy Dahl,[5] Luiz A. F. Trindade,[6] Stefan Schouten[1]

The 18S ribosomal DNA molecular phylogeny and lipid composition of over 120 marine diatoms showed that the capability to biosynthesize highly branched isoprenoid (HBI) alkenes is restricted to two specific phylogenetic clusters, which independently evolved in centric and pennate diatoms. The molecular record of C_{25} HBI chemical fossils in a large suite of well-dated marine sediments and petroleum revealed that the older cluster, composed of rhizosolenid diatoms, evolved 91.5 ± 1.5 million years ago (Upper Turonian), enabling an accurate dating of the pace of diatom evolution that is unprecedented. The rapid rise of the rhizosolenid diatoms probably resulted from a major reorganization of the nutrient budget in the mid-Cretaceous oceans, triggered by plate tectonics.

In modern oceans, diatoms are the most important group of phytoplankton, responsible for almost half of the marine primary productivity (*1*, *2*), but their ecological dominance occurred relatively recently in geological time (*3*). The fossil record of marine diatom

Downloaded from http://science.sciencemag.org/ on August 15, 2020

Science

The Fine-Scale Structure of Recombination Rate Variation in the Human Genome

Gilean A. T. McVean, Simon R. Myers, Sarah Hunt, Panos Deloukas, David R. Bentley and Peter Donnelly

Science **304** (5670), 581-584.
DOI: 10.1126/science.1092500

ARTICLE TOOLS http://science.sciencemag.org/content/304/5670/581

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2004/04/27/304.5670.581.DC1

REFERENCES This article cites 37 articles, 5 of which you can access for free
http://science.sciencemag.org/content/304/5670/581#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

American Association for the Advancement of Science

Downloaded from http://science.sciencemag.org/ on August 15, 2020