

Bibliography

- J. M. Akey, G. Zhang, K. Zhang, L. Jin, and M. D. Shriver. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res*, 12:1805–14, 2002.
- L. T. Amundadottir, P. Sulem, J. Gudmundsson, A. Helgason, A. Baker, B. A. Agnarsson, A. Sigurdsson, K. R. Benediktsdottir, J. B. Cazier, J. Sainz, M. Jakobsdottir, J. Kostic, D. N. Magnusdottir, S. Ghosh, K. Agnarsson, B. Birgisdottir, L. L. Roux, A. Olafsdottir, T. Blondal, M. Andresdottir, O. S. Gretarsdottir, J. T. Bergthorsson, D. Gudbjartsson, A. Gylfason, G. Thorleifsson, A. Manolescu, K. Kristjansson, G. Geirsson, H. Isaksson, J. Douglas, J. E. Johansson, K. Balter, F. Wiklund, J. E. Montie, X. Yu, B. K. Suarez, C. Ober, K. A. Cooney, H. Gronberg, W. J. Catalona, G. V. Einarsson, R. B. Barkardottir, J. R. Gulcher, A. Kong, U. Thorsteinsdottir, and K. Stefansson. A common variant associated with prostate cancer in European and African populations. *Nat Genet*, 38:652–8, 2006.
- S. E. Antonarakis and J. S. Beckmann. Mendelian disorders deserve more attention. *Nat Rev Genet*, 7:277–82, 2006.
- V. Bafna and V. Bansal. Inference about recombination from haplotype data: lower bounds and recombination hotspots. *J Comput Biol*, 13:501–21, 2006.
- D. Balding. The impact of low-cost and genome-wide resequencing on association studies. *Human Genomics*, 2:79–8, 2005.
- D. J. Balding. A tutorial on statistical methods for population association studies. *Nat Rev Genet*, 7:781–91, 2006.

- C. Bardel, V. Danjean, J. P. Hugot, P. Darlu, and E. Genin. On the use of haplotype phylogeny to detect disease susceptibility loci. *BMC Genet*, 6:24, 2005.
- I. Barroso, J. Luan, E. Wheeler, P. Whittaker, J. Wasson, E. Zeggini, M. N. Weedon, S. Hunt, M. Delgado, R. Venkatesh, T. M. Frayling, R. J. Neuman, R. Sherva, B. Glaser, M. Walker, G. Hitman, M. I. McCarthy, A. T. Hattersley, M. A. Permutt, N. J. Wareham, and P. Deloukas. Population-specific risk of type 2 diabetes (T2D) conferred by HNF4A P2 promoter variants: a cautionary tale for genome-wide association studies. *Diabetes*, Submitted, 2007.
- M. A. Beaumont. Detecting population expansion and decline using microsatellites. *Genetics*, 153:2013–29, 1999.
- T. Bersaglieri, P. C. Sabeti, N. Patterson, T. Vanderploeg, S. F. Schaffner, J. A. Drake, M. Rhodes, D. E. Reich, and J. N. Hirschhorn. Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet*, 74:1111–20, 2004.
- L. R. Cardon and J. I. Bell. Association study designs for complex diseases. *Nat Rev Genet*, 2:91–9, 2001.
- D. Carter, M. J. Minichiello, and R. Durbin. Imputing missing DNA values and discovering SNPs from low-coverage sequencing of multiple individuals. *In preparation*, 2007.
- M. Chiano and D. Clayton. Fine genetic mapping using haplotype analysis and the missing data problem. *Ann Hum Genet*, 62:55–60, 1998.
- G. A. Churchill and R. W. Doerge. Empirical threshold values for quantitative trait mapping. *Genetics*, 138:963–71, 1994.
- A. G. Clark. Inference of haplotypes from PCR-amplified samples of diploid populations. *Mol Biol Evol.*, 7:111–22, 1990.
- D. G. Clayton, N. M. Walker, D. J. Smyth, R. Pask, J. D. Cooper, L. M. Maier, L. J. Smink, A. C. Lam, N. R. Ovington, H. E. Stevens, S. Nutland, J. M. Howson, M. Faham, M. Moorhead, H. B. Jones, M. Falkowski, P. Hardenbol, T. D. Willis, and J. A. Todd.

- Population structure and differential bias and genomic control in a large-scale and case-control association study. *Nat Genet*, 37:1243–6, 2005.
- H. J. Cordell. Epistasis: what it means, what it doesn't mean and statistical methods to detect it in humans. *Hum Mol Genet*, 11:2463–8, 2002.
- H. J. Cordell. Estimation and testing of genotype and haplotype effects in case/control studies: comparison of weighted regression and multiple imputation procedures. *Genet Epidemiol*, 30:259–75, 2006.
- H. J. Cordell and D. G. Clayton. Genetic association studies. *Lancet*, 366:1121–31, 2005.
- D. C. Crawford, T. Bhangale, N. Li, G. Hellenthal, M. J. Rieder, D. A. Nickerson, and M. Stephens. Evidence for substantial fine-scale variation in recombination rates across the human genome. *Nat Genet*, 36:700–6, 2004.
- E. D. Crawford. Epidemiology of prostate cancer. *Urology*, 22:3–12, 2003.
- J. Y. Dai, I. Ruczinski, M. LeBlanc, and C. Kooperberg. Imputation methods to improve inference in SNP association studies. *Genet Epidemiol*, 30:690–702, 2006.
- M. J. Daly, J. D. Rioux, S. F. Schaffner, T. J. Hudson, and E. S. Lander. High-resolution haplotype structure in the human genome. *Nat Genet*, 29:229–32, 2001.
- Z. Dawy, B. Goebel, J. Hagenauer, C. Andreoli, T. Meitinger, and J. C. Mueller. Gene mapping and marker clustering using Shannon's mutual information. *IEEE/ACM Trans Comput Biol Bioinform*, 3:47–56, 2006.
- P. I. de Bakker, R. Yelensky, I. Pe'er, S. B. Gabriel, M. J. Daly, and D. Altshuler. Efficiency and power in genetic association studies. *Nat Genet*, 37:1217–23, 2005.
- B. Devlin and N. Risch. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics*, 29:311–22, 1995.

- F. Dudbridge and B. P. Koeleman. Efficient computation of significance levels for multiple associations in large studies of correlated data and including genomewide association studies. *Am J Hum Genet*, 75:424–35, 2004.
- C. Durrant, K. T. Zondervan, L. R. Cardon, S. Hunt, P. Deloukas, and A. P. Morris. Linkage disequilibrium mapping via cladistic analysis of single-nucleotide polymorphism haplotypes. *Am J Hum Genet*, 75:35–43, 2004.
- A. O. Edwards, R. Ritter, K. J. Abel, A. Manning, C. Panhuysen, and L. A. Farrer. Complement factor H polymorphism and age-related macular degeneration. *Science*, 308:421–4, 2005.
- D. M. Evans, L. R. Cardon, and A. P. Morris. Genotype prediction using a dense map of SNPs. *Genet Epidemiol*, 27:375–84, 2004.
- L. Excoffier and M. Slakin. Maximum likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol*, 12:921–7, 1995.
- S. Eyheramendy, J. Marchini, G. McVean, S. Myers, and P. Donnelly. A model-based approach to capture genetic variation for future association studies. *Genome Res*, 17:88–95, 2007.
- P. Fearnhead. SequenceLDhot: detecting recombination hotspots. *Bioinformatics*, 22:3061–6, 2006.
- P. Fearnhead and P. Donnelly. Estimating recombination rates from population genetic data. *Genetics*, 159:1299–318, 2001.
- P. Fearnhead, R. M. Harding, J. A. Schneider, S. Myers, and P. Donnelly. Application of coalescent methods to reveal fine-scale rate variation and recombination hotspots. *Genetics*, 167:2067–81, 2004.
- J. Felsenstein. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol*, 17:368–76, 1981.
- J. Felsenstein. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39:783–91, 1985.

- M. L. Freedman, C. A. Haiman, N. Patterson, G. J. McDonald, A. Tandon, A. Waliszewska, K. Penney, R. G. Steen, K. Ardlie, E. M. John, I. Oakley-Girvan, A. S. Whittemore, K. A. Cooney, S. A. Ingles, D. Altshuler, B. E. Henderson, and D. Reich. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proc Natl Acad Sci U S A*, 103:14068–73, 2006.
- S. B. Gabriel, S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S. N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E. S. Lander, M. J. Daly, and D. Altshuler. The structure of haplotype blocks in the human genome. *Science*, 296:2225–9, 2002.
- D. B. Goldstein, K. R. Ahmadi, M. E. Weale, and N. W. Wood. Genome scans and candidate gene approaches in the study of common diseases and variable drug responses. *Trends Genet*, 19:615–22, 2003.
- J. Graham and E. A. Thompson. Disequilibrium likelihoods for fine-scale mapping of a rare allele. *Am J Hum Genet*, 63:1517–30, 1998.
- R. C. Griffiths and P. Marjoram. Ancestral inference from samples of DNA sequences with recombination. *J Comput Biol*, 3:479–502, 1996.
- R. C. Griffiths and P. Marjoram. An ancestral recombination graph. In: *Donnelly P and Tavaré S (eds) and Progress in Population Genetics and Human Evolution and Springer Verlag*, pages 257–70, 1997.
- R. C. Griffiths and S. Tavaré. Ancestral inference in population genetics. *Statistical Science*, 9:307–19, 1994a.
- R. C. Griffiths and S. Tavaré. Sampling theory for neutral alleles in a varying environment. *Phil Trans R Soc Lond B*, 344:403–10, 1994b.
- R. C. Griffiths and S. Tavaré. Simulating probability distributions in the coalescent. *Theor Popn Biol*, 46:131–59, 1994c.

- J. Gudmundsson, P. Sulem, A. Manolescu, L. T. Amundadottir, D. Gudbjartsson, A. Helgason, T. Rafnar, J. T. Bergthorsson, B. A. Agnarsson, A. Baker, A. Sigurdsson, K. R. Benediktsdottir, M. Jakobsdottir, J. Xu, T. Blondal, J. Kostic, J. Sun, S. Ghosh, S. N. Stacey, M. Mouy, J. Saemundsdottir, V. M. Backman, K. Kristjansson, A. Tres, A. W. Partin, M. T. Albers-Akkers, J. G.-I. Marcos, P. C. Walsh, D. W. Swinkels, S. Navarrete, S. D. Isaacs, K. K. Aben, T. Graif, J. Cashy, M. Ruiz-Echarri, K. E. Wiley, B. K. Suarez, J. A. Witjes, M. Frigge, C. Ober, E. Jonsson, G. V. Einarsson, J. I. Mayordomo, L. A. Kiemeny, W. B. Isaacs, W. J. Catalona, R. B. Barkardottir, J. R. Gulcher, U. Thorsteinsdottir, A. Kong, and K. Stefansson. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet*, 39:631–7, 2007.
- K. L. Gunderson, F. J. Steemers, G. Lee, L. G. Mendoza, and M. S. Chee. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat Genet*, 37:549–54, 2005.
- D. Gusfield. Haplotyping as perfect phylogeny: conceptual framework and efficient solutions (extended abstract). *Proceedings of the 6th International Conference on Computational Molecular Biology*, page 166, 2002.
- D. Gusfield, D. Hickerson, and S. Eddhu. An efficiently-computed lower bound on the number of recombinations in phylogenetic networks: Theory and empirical study. *Discrete Applied Mathematics*, 155:806–30, 2007.
- C. A. Haiman, N. Patterson, M. L. Freedman, S. R. Myers, M. C. Pike, A. Waliszewska, J. Neubauer, A. Tandon, C. Schirmer, G. J. McDonald, S. C. Greenway, D. O. Stram, L. L. Marchand, L. N. Kolonel, M. Frasco, D. Wong, L. C. Pooler, K. Ardlie, I. Oakley-Girvan, A. S. Whittemore, K. A. Cooney, E. M. John, S. A. Ingles, D. Altshuler, B. E. Henderson, and D. Reich. Multiple regions within 8q24 independently affect risk for prostate cancer. *Nat Genet*, 39:638–44, 2007.
- J. L. Haines, M. A. Hauser, S. Schmidt, W. K. Scott, L. M. Olson, P. Gallins, K. L. Spencer, S. Y. Kwan, M. Nouredine, J. R. Gilbert, N. Schnetz-Boutaud, A. Agarwal, E. A. Postel,

- and M. A. Pericak-Vance. Complement factor H variant increases the risk of age-related macular degeneration. *Science*, 308:419–21, 2005.
- E. Halperin and E. Eskin. Haplotype reconstruction from genotype data using imperfect phylogeny. *Bioinformatics*, 20:1842–9, 2004.
- M. T. Hamblin, E. E. Thompson, and A. D. Rienzo. Complex signatures of natural selection at the Duffy blood group locus. *Am J Hum Genet*, 70:369–83, 2002.
- N. A. Hanchard, K. A. Rockett, C. Spencer, G. Coop, M. Pinder, M. Jallow, M. Kimber, G. McVean, R. Mott, and D. P. Kwiatkowski. Screening for recently selected alleles by analysis of human haplotype similarity. *Am J Hum Genet*, 78:153–9, 2006.
- D. A. Hinds, L. L. Stuve, G. B. Nilsen, E. Halperin, E. Eskin, D. G. Ballinger, K. A. Frazer, and D. R. Cox. Whole-genome patterns of common DNA variation in three human populations. *Science*, 18:1072–9, 2005.
- J. N. Hirschhorn and M. J. Daly. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet*, 6:95–108, 2005.
- J. N. Hirschhorn, K. Lohmueller, E. Byrne, and K. Hirschhorn. A comprehensive review of genetic association studies. *Genet Med*, 4:45–61, 2002.
- C. J. Hoggart, T. Clark, R. Lampariello, M. D. Iorio, J. Whittaker, and D. Balding. FREGENE: Software for simulating large genomic regions. *Technical Report and Department of Epidemiology and Public Health and Imperial College and University of London*, 2005.
- R. R. Hudson. Properties of a neutral allele model with intragenic recombination. *Theor Popul Biol*, 23:183–201, 1983.
- R. R. Hudson. Two-locus sampling distributions and their application. *Genetics*, 159:1805–17, 2001.
- R. R. Hudson. Generating samples under a wright-fisher neutral model of genetic variation. *Bioinformatics*, 18:337–8, 2002.

- R. R. Hudson and N. L. Kaplan. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, 111:147–64, 1985.
- J. P. Ioannidis, T. A. Trikalinos, E. E. Ntzani, and D. G. Contopoulos-Ioannidis. Genetic associations in large versus small studies: an empirical assessment. *Lancet*, 361:567–71, 2003.
- A. J. Jeffreys, L. Kauppi, and R. Neumann. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet*, 29:217–22, 2001.
- G. C. Johnson, L. Esposito, B. J. Barratt, A. N. Smith, J. Heward, G. D. Genova, H. Ueda, H. J. Cordell, I. A. Eaves, F. Dudbridge, R. C. Twells, F. Payne, W. Hughes, S. Nutland, H. Stevens, P. Carr, E. Tuomilehto-Wolf, J. Tuomilehto, S. C. Gough, D. G. Clayton, and J. A. Todd. Haplotype tagging for the identification of common disease genes. *Nat Genet*, 29:233–7, 2001.
- J. F. C. Kingman. On the genealogy of large populations. *J Appl Prob*, 19A:27–43, 1982.
- R. J. Klein, C. Zeiss, E. Y. Chew, J. Y. Tsai, R. S. Sackler, C. Haynes, A. K. Henning, J. P. SanGiovanni, S. M. Mane, S. T. Mayne, M. B. Bracken, F. L. Ferris, J. Ott, C. Barnstable, and J. Hoh. Complement factor H polymorphism in age-related macular degeneration. *Science*, 308:385–9, 2005.
- L. Kruglyak. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet*, 22:139–44, 1999.
- L. Kruglyak and D. A. Nickerson. Variation is the spice of life. *Nat Genet*, 27:234–6, 2001.
- M. K. Kuhner, J. Yamato, and J. Felsenstein. Maximum likelihood estimation of recombination rates from population data. *Genetics*, 156:1393–401, 2000.
- F. Larribe, S. Lessard, and N. J. Schork. Gene mapping via the ancestral recombination graph. *Theor Popul Biol*, 62:215–29, 2002.
- S. Lesage and C. C. Goodnow. Organ-specific autoimmune disease: a deficiency of tolerogenic stimulation. *J Exp Med*, 194:31–36, 2001.

- N. Li and M. Stephens. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, 165:2213–33, 2003.
- R. J. A. Little and D. B. Rubin. Statistical analysis with missing data. *John Wiley and Sons, Inc*, 1987.
- K. E. Lohmueller, C. L. Pearce, M. Pike, E. S. Lander, and J. N. Hirschhorn. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*, 33:177–82, 2003.
- J. Long, R. Williams, and M. Urbanek. An e-m algorithm and testing strategy for multiple-locus haplotypes. *Am J Hum Genet*, 56:799–810, 1995.
- R. Lyngsø, Y. S. Song, and J. Hein. Minimum recombination histories by branch and bound. *Proceedings of Workshop on Algorithms in Bioinformatics, Lecture Notes in Computer Science*, 3692:239–50, 2005.
- S. Macgregor and I. A. Khan. GAIA: An easy-to-use web-based application for interaction analysis of case-control data. *BMC Med Genet*, 7:34, 2006.
- J. Marchini, D. Cutler, N. Patterson, M. Stephens, E. Eskin, E. Halperin, S. Lin, Z. S. Qin, H. M. Munro, G. R. Abecasis, and P. Donnelly. A comparison of phasing algorithms for trios and unrelated individuals. *Am J Hum Genet*, 78:437–50, 2006.
- J. Marchini, B. Howie, S. Myers, G. McVean, and P. D. P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*, 39:906–13, 2007.
- H. Matsuzaki, S. Dong, H. Loi, X. Di, G. Liu, E. Hubbell, J. Law, T. Berntsen, M. Chadha, H. Hui, G. Yang, G. C. Kennedy, T. A. Webster, S. Cawley, P. S. Walsh, K. W. Jones, S. P. Fodor, and R. Mei. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nat Methods*, 1:109–11, 2004.
- G. A. McVean. A genealogical interpretation of linkage disequilibrium. *Genetics*, 162:987–91, 2002.

- G. A. McVean and N. J. Cardin. Approximating the coalescent with recombination. *Philos Trans R Soc Lond B Biol Sci*, 360:1387–93, 2005.
- G. A. T. McVean, P. Awadalla, and P. Fearnhead. A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics*, 160:1231–41, 2002.
- G. A. T. McVean, S. Myers, S. Hunt, P. Deloukas, D. R. Bentley, and P. Donnelly. The fine-scale structure of recombination rate variation in the human genome. *Science*, 304:581–584, 2004.
- F. K. Mensah, M. S. Gilthorpe, C. F. Davies, L. J. Keen, P. J. Adamson, E. Roman, G. J. Morgan, J. L. Bidwell, and G. R. Law. Haplotype uncertainty in association studies. *Genet Epidemiol*, 31:348–57, 2007.
- M. J. Minichiello and R. Durbin. Mapping trait loci by use of inferred ancestral recombination graphs. *Am J Hum Genet*, 79:910–22, 2006.
- J. Molitor, P. Marjoram, and D. Thomas. Fine-scale mapping of disease genes with multiple mutations via spatial clustering techniques. *Am J Hum Genet*, 73:1368–84, 2003.
- A. P. Morris, J. C. Whittaker, and D. J. Balding. Fine-scale mapping of disease loci via shattered coalescent modeling of genealogies. *Am J Hum Genet*, 70:686–707, 2002.
- A. P. Morris, J. C. Whittaker, and D. J. Balding. Little loss of information due to unknown phase for fine-scale linkage-disequilibrium mapping with single-nucleotide-polymorphism genotype data. *Am J Hum Genet*, 74:945–53, 2004.
- C. Myers, L. Bottolo, C. Freeman, G. McVean, and P. Donnelly. A fine-scale map of recombination rates and hotspots across the human genome. *Science*, 310:321–4, 2005.
- S. R. Myers and R. C. Griffiths. Bounds on the minimum number of recombination events in a sample history. *Genetics*, 163:375–94, 2003.
- S. A. Narod, A. Dupont, L. Cusan, P. Diamond, J.-L. Gomez, R. Suburu, and F. Labrie. The impact of family history on early detection of prostate cancer. *Nat Med*, 1:99–101, 1995.

- Nature Genetics Editorial. Freely associating. *Nat Genet*, 22:1–2, 1999.
- D. L. Nicolae. Testing untyped alleles (TUNA)-applications to genome-wide association studies. *Genet Epidemiol*, 30:718–27, 2006.
- R. Nielsen. Estimation of population parameters and recombination rates from single nucleotide polymorphisms. *Genetics*, 154:931–42, 2000.
- R. Nielsen, S. Williamson, Y. Kim, M. J. Hubisz, A. G. Clark, and C. Bustamante. Genomic scans for selective sweeps using SNP data. *Genome Res*, 15:1566–75, 2005.
- M. Nordborg and S. Tavaré. Linkage disequilibrium: what history has to tell us. *Trends Genet*, 18:83–90, 2002.
- L. J. Palmer and L. R. Cardon. Shaking the tree: mapping complex disease genes with linkage disequilibrium. *Lancet*, 366:1223–34, 2005.
- P. Paschou, M. W. Mahoney, A. Javed, J. R. Kidd, A. J. Pakstis, S. Gu, K. K. Kidd, and P. Drineas. Intra- and interpopulation genotype reconstruction from tagging SNPs. *Genome Res*, 17:96–107, 2007.
- N. Patil, A. J. Berno, D. A. Hinds, W. A. Barrett, J. M. Doshi, C. R. Hacker, C. R. Kautzer, D. H. Lee, C. Marjoribanks, D. P. McDonough, B. T. Nguyen, M. C. Norris, J. B. Sheehan, N. Shen, D. Stern, R. P. Stokowski, D. J. Thomas, M. O. Trulson, K. R. Vyas, K. A. Frazer, S. P. Fodor, and D. R. Cox. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science*, 23:1719–23, 2001.
- N. Patterson, N. Hattangadi, B. Lane, K. E. Lohmueller, D. A. Hafler, J. R. Oksenberg, S. L. Hauser, M. W. Smith, S. J. O’Brien, D. Altshuler, M. J. Daly, and D. Reich. Methods for high-density admixture mapping of disease genes. *Am J Hum Genet*, 74:979–1000, 2004.
- A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*, 38:904–9, 2006.

- J. K. Pritchard and M. Przeworski. Linkage disequilibrium in humans. *Am J Hum Genet*, 69:1–14, 2001.
- J. K. Pritchard, M. Stephens, and P. Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155:945–59, 2000.
- S. Purcell, S. S. Cherny, and P. C. Sham. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, 19:149–50, 2003.
- B. Rannala and J. P. Reeve. High-resolution multipoint linkage-disequilibrium mapping in the context of a human genome sequence. *Am J Hum Genet*, 69:159–78, 2001.
- F. A. Reed and S. A. Tishkoff. Positive selection can create false hotspots of recombination. *Genetics*, 172:2011–4, 2006.
- J. D. Rioux, M. J. Daly, M. S. Silverberg, K. Lindblad, H. Steinhart, Z. Cohen, T. Delmonte, K. Kocher, K. Miller, S. Guschwan, E. J. Kulbokas, S. O’Leary, E. Winchester, K. Dewar, T. Green, V. Stone, C. Chow, A. Cohen, D. Langelier, G. Lapointe, D. Gaudet, J. Faith, N. Branco, S. B. Bull, R. S. McLeod, A. M. Griffiths, A. Bitton, G. R. Greenberg, E. S. Lander, K. A. Siminovitch, and T. J. Hudson. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet*, 29:223–8, 2001.
- N. Risch and K. Merikangas. The future of genetic studies of complex human diseases. *Science*, 13:1516–7, 1996.
- N. J. Risch. Searching for genetic determinants in the new millennium. *Nature*, 405:847–56, 2000.
- D. L. Rohde, S. Olson, and J. T. Chang. Modelling the recent common ancestry of all living humans. *Nature*, 431:562–6, 2004.
- S. Romeo, L. A. Pennacchio, Y. Fu, E. Boerwinkle, A. Tybjaerg-Hansen, H. H. Hobbs, and J. C. Cohen. Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. *Nat Genet*, 39:513–16, 2007.

- D. B. Rubin. Multiple imputation for nonresponse in surveys. *John Wiley and Sons*, 1987.
- W. L. Ruzzo and M. Tompa. A linear time algorithm for finding all maximal scoring subsequences. *Proc Int Conf Intell Syst Mol Biol*, pages 234–41, 1999.
- P. C. Sabeti, D. E. Reich, J. M. Higgins, H. Z. Levine, D. J. Richter, S. F. Schaffner, S. B. Gabriel, J. V. Platko, N. J. Patterson, G. J. McDonald, H. C. Ackerman, S. J. Campbell, D. Altshuler, R. Cooper, D. Kwiatkowski, R. Ward, and E. S. Lander. Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419:832–7, 2002.
- R. Sachidanandam, D. Weissman, S. C. Schmidt, J. M. Kakol, L. D. Stein, G. Marth, S. Sherry, J. C. Mullikin, B. J. Mortimore, D. L. Willey, S. E. Hunt, C. G. Cole, P. C. Coggill, C. M. Rice, Z. Ning, J. Rogers, D. R. Bentley, P. Y. Kwok, E. R. Mardis, R. T. Yeh, B. Schultz, L. Cook, R. Davenport, M. Dante, L. Fulton, L. Hillier, R. H. Waterston, J. D. McPherson, B. Gilman, S. Schaffner, W. J. V. Etten, D. Reich, J. Higgins, M. J. Daly, B. Blumenstiel, J. Baldwin, N. Stange-Thomann, M. C. Zody, L. Linton, E. S. Lander, D. Altshuler, and I. S. M. W. Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409:928–33, 2001.
- P. D. Sasieni. From genotypes to genes: Doubling the sample size. *Biometrics*, 53:1253–61, 1997.
- D. J. Schaid and S. J. Jacobsen. Biased tests of association: comparisons of allele frequencies when departing from hardy-weinberg proportions. *Am J Epidemiol*, 149:706–11, 1999.
- P. Scheet and M. Stephens. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet*, 78:629–44, 2006.
- B. Servin and M. Stephens. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genetics*, preprint, 2007.
- M. W. Smith, N. Patterson, J. A. Lautenberger, A. L. Truelove, G. J. McDonald, A. Waliszewska, B. D. Kessing, M. J. Malasky, C. Scafe, E. Le, P. L. D. Jager, A. A. Mignault,

- Z. Yi, G. D. The, M. Essex, J. L. Sankale, J. H. Moore, K. Poku, J. P. Phair, J. J. Goedert, D. Vlahov, S. M. Williams, S. A. Tishkoff, C. A. Winkler, F. D. L. Vega, T. Woodage, J. J. Sninsky, D. A. Hafler, D. Altshuler, D. A. Gilbert, S. J. O'Brien, and D. Reich. A high-density admixture map for disease gene discovery in african americans. *Am J Hum Genet*, 74:1001–13, 2004.
- Y. S. Song and J. Hein. On the minimum number of recombination events in the evolutionary history of DNA sequences. *J Math Biol*, 48:160–86, 2004.
- Y. S. Song and J. Hein. Constructing minimal ancestral recombination graphs. *J Comput Biol*, 12:147–69, 2005.
- Y. S. Song, R. Lyngsø, and J. Hein. Counting all possible ancestral configurations of sample sequences in population genetics. *IEEE/ACM Trans Comput Biol Bioinform*, 3:239–251, 2006.
- O. W. Souverein, A. H. Zwinderman, and M. W. Tanck. Multiple imputation of missing genotype data for unrelated individuals. *Ann Hum Genet*, 70:372–81, 2006.
- M. Stephens and P. Donnelly. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet*, 73:1162–69, 2003.
- M. Stephens and P. Scheet. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet*, 76:449–62, 2005.
- M. Stephens, N. J. Smith, and P. Donnelly. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*, 68:978–89, 2001.
- S. Tavaré, D. J. Balding, R. C. Griffiths, and P. Donnelly. Inferring coalescence times from DNA sequence data. *Genetics*, 145:505–18, 1997.
- A. R. Templeton. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping or DNA sequencing. V. Analysis of case/control sampling designs: Alzheimer's disease and the apoprotein E locus. *Genetics*, 140:403–9, 1995.

- A. R. Templeton, E. Boerwinkle, and C. F. Sing. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in drosophila. *Genetics*, 117:343–51, 1987.
- A. R. Templeton, K. A. Crandall, and C. F. Sing. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, 132:619–33, 1992.
- A. R. Templeton, T. Maxwell, D. Posada, J. H. Stengard, E. Boerwinkle, and C. F. Sing. Tree scanning: a method for using haplotype trees in phenotype/genotype association studies. *Genetics*, 169:441–53, 2005.
- A. R. Templeton and C. F. Sing. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, 134:659–69, 1993.
- A. R. Templeton, C. F. Sing, A. Kessler, and S. Humphries. A cladistic analysis of phenotype associations with haplotypes inferred from restriction endonuclease mapping. II. The analysis of natural populations. *Genetics*, 120:1145–54, 1988.
- J. D. Terwilliger and T. Hiekkalinna. An utter refutation of the fundamental theorem of the HapMap. *Eur J Hum Genet*, 14:426–37, 2006.
- The International HapMap Consortium. A haplotype map of the human genome. *Nature*, 437:1299–320, 2005.
- The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447:661–78, 2007.
- H. Ueda, J. M. Howson, L. Esposito, J. Heward, H. Snook, G. Chamberlain, D. B. R. K. M. Hunter, A. N. Smith, G. D. Genova, M. H. Herr, I. Dahlman, F. Payne, D. Smyth, C. Lowe, R. C. Twells, S. Howlett, B. Healy, S. Nutland, H. E. Rance, V. Everett, L. J. Smink, A. C. Lam, H. J. Cordell, N. M. Walker, C. Bordin, J. Hulme, C. Motzo, F. Cucca, J. F. Hess,

- M. L. Metzker, J. Rogers, S. Gregory, A. Allahabadia, R. Nithiyanthan, E. Tuomilehto-Wolf, J. Tuomilehto, P. Bingley, K. M. Gillespie, D. E. Undlien, K. S. Ronningen, C. Guja, C. Ionescu-Tirgoviste, D. A. Savage, A. P. Maxwell, D. J. Carson, C. C. Patterson, J. A. Franklyn, D. G. Clayton, L. B. Peterson, L. S. Wicker, J. A. Todd, and S. C. Gough. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. *Nature*, 423:506–11, 2003.
- B. Vaidya and S. Pearce. The emerging role of the CTLA-4 gene in autoimmune endocrinopathies. *Eur J Endocrinol*, 150:619–26, 2004.
- B. F. Voight, S. Kudravalli, X. Wen, and J. K. Pritchard. A map of recent positive selection in the human genome. *PLoS Biol*, 7:e72, 2006.
- E. R. Waldron, J. C. Whittaker, and D. J. Balding. Fine mapping of disease genes via haplotype clustering. *Genet Epidemiol*, 30:170–9, 2006.
- L. Wang, K. Zhang, and L. Zhang. Perfect phylogenetic networks with recombination. *J Comput Biol*, 8:69–78, 2001.
- W. Y. Wang, B. J. Barratt, D. G. Clayton, and J. A. Todd. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet*, 6:109–18, 2005.
- I. J. Wilson and D. J. Balding. Genealogical inference from microsatellite data. *Genetics*, 150:499–510, 1998.
- C. Wiuf. Inference on recombination and block structure using unphased data. *Genetics*, 166:537–45, 2004.
- S. Wright. Evolution in Mendelian populations. *Genetics*, 16:97–159, 1931.
- M. Yeager, N. Orr, R. B. Hayes, K. B. Jacobs, P. Kraft, S. Wacholder, M. J. Minichiello, P. Fearnhead, K. Yu, N. Chatterjee, Z. Wang, R. Welch, B. J. Staats, E. E. Calle, H. S. Feigelson, M. J. Thun, C. Rodriguez, D. Albanes, J. Virtamo, S. Weinstein, F. R. Schumacher, E. Giovannucci, W. C. Willett, G. Cancel-Tassin, O. Cussenot, A. Valeri, G. L. Andriole, E. P. Gelmann, M. Tucker, D. S. Gerhard, J. R. Fraumeni, R. Hoover, D. J.

-
- Hunter, S. J. Chanock, and G. Thomas. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet*, 39:645–649, 2007.
- K. Zhang, M. Deng, T. Chen, M. S. Waterman, and F. Sun. A dynamic programming algorithm for haplotype block partitioning. *Proc Natl Acad Sci U S A*, 28:7335–9, 2002.
- S. Zollner and J. K. Pritchard. Coalescent-based association mapping and fine mapping of complex trait loci. *Genetics*, 169:1071–92, 2005.
- G. Zsurka, Y. Kraytsberg, T. Kudina, C. Kornblum, C. E. Elger, K. Khrapko, and W. S. Kunz. Recombination of mitochondrial DNA in skeletal muscle of individuals with multiple mitochondrial DNA heteroplasmy. *Nat Genet*, 37:873–7, 2005.