

Host genetic and epigenetic factors in toxoplasmosis

Sarra E Jamieson^{1,2}, Heather Cordell^{2,3}, Eskild Petersen⁴, Rima McLeod^{5,6}, Ruth E Gilbert⁷,
Jenefer M Blackwell^{1,2/+}

¹Telethon Institute for Child Health Research, Centre for Child Health Research, The University of Western Australia, PO Box 855, West Perth, Western Australia 6872, Australia ²Cambridge Institute for Medical Research and Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, UK ³Institute of Human Genetics, Newcastle University, International Centre for Life, Central Parkway, Newcastle upon Tyne, UK ⁴Department of Infectious Diseases, Aarhus University Hospital, Aarhus, Denmark ⁵Departments of Ophthalmology, Medicine, Pediatrics, Committees on Immunology, Molecular Medicine and Genetics, University of Chicago, Illinois, USA ⁶Michael Reese Hospital and Medical Center, Chicago, Illinois, USA ⁷Centre for Paediatric Epidemiology and Biostatistics, Institute of Child Health, University College London, London, UK

Analysing human genetic variation provides a powerful tool in understanding risk factors for disease. Toxoplasma gondii acquired by the mother can be transmitted to the fetus. Infants with the most severe clinical signs in brain and eye are those infected early in pregnancy when fetal immunity is least well developed. Genetic analysis could provide unique insight into events in utero that are otherwise difficult to determine. We tested the hypothesis that propensity for T. gondii to cause eye disease is associated with genes previously implicated in congenital or juvenile onset ocular disease. Using mother-child pairs from Europe (EMSCOT) and child/parent trios from North America (NCCCTS), we demonstrated that ocular and brain disease in congenital toxoplasmosis associate with polymorphisms in ABCA4 encoding ATP-binding cassette transporter, subfamily A, member 4 previously associated with juvenile onset retinal dystrophies including Stargardt's disease. Polymorphisms at COL2A1 encoding type II collagen, previously associated with Stickler syndrome, associated only with ocular disease in congenital toxoplasmosis. Experimental studies showed that both ABCA4 and COL2A1 show isoform-specific epigenetic modifications consistent with imprinting, which provided an explanation for the patterns of inheritance observed. These genetic and epigenetic risk factors provide unique insight into molecular pathways in the pathogenesis of disease.

Key words: congenital infection - toxoplasmosis - genetics - epigenetics

Following the human genome project, analyzing human genetic variation has emerged as a powerful tool in understanding both direct genetic and modifiable environmental (Davey Smith et al. 2008, Ebrahim & Davey Smith 2008) risk factors for disease. These approaches could be applied to analysis of clinical outcomes for both congenital and acquired toxoplasmosis. We recently used a genetic approach to look at congenital toxoplasmosis (Jamieson et al. 2008). As is well known, *Toxoplasma gondii* is a ubiquitous protozoan parasitic infection that, if acquired for the first time during pregnancy, can be transmitted to the fetus. At birth, infants infected *in utero* may have intracranial calcification, hydrocephalus and ocular disease broadly defined as retinochoroiditis or inflammation of the retina and choroid with associated vitritis (Gilbert & Gras 2003, Gras et al. 2005, McLeod et al. 2006). New ocular lesions can occur at any age after birth, in untreated and some treated children. Whilst severity of disease is influenced by trimester in which infection is acquired by the mother (Dunn et al. 1999, Remington et al. 2005), other factors including genetic predisposition may contribute. For example, previous

studies suggest that genes affecting immune response, including HLA (Mack et al. 1999), influence clinical outcome in the child. However, since infants who have the most severe clinical signs in the brain and eye are those infected early in pregnancy (Dunn et al. 1999, Remington et al. 2005) when fetal immunity is least well developed, we considered whether genes that encode molecules that play a role in developmental processes in the eye or the brain could provide another avenue to identify candidate genes that contribute to clinical phenotype observed in the child. This could provide unique insight into events *in utero* that determine the clinical outcome of infection.

The eye disorder gene hypothesis?

One hypothesis we pursued (Jamieson et al. 2008) was whether the propensity for *T. gondii* to cause eye disease was associated with genes previously implicated in congenital or juvenile onset ocular disease. Two genes were studied, *ABCA4* encoding ATP-binding cassette transporter subfamily A member 4 associated with juvenile onset retinal dystrophies including Stargardt's disease (Ducroq et al. 2002, Koenekoop 2003), and *COL2A1* encoding type II collagen associated with Stickler syndrome (Rose et al. 2005), in which there is congenital abnormal vitreous and lattice retinal degeneration. The full results of these studies are presented elsewhere (Jamieson et al. 2008). All of the data provided here are from this paper, which is therefore not further referenced in the results section below. Here we summarize the salient

Financial support: Main funders for the published work on which this article is based were Guide Dogs for the Blind Association, UK; NIH, USA.

+ Corresponding author: jmb37@cam.ac.uk

Received 10 October 2008

Accepted 17 February 2009

findings and examine the implications of this research to the application of genetics as a tool to understand the multiple pathologies associated with toxoplasmosis.

Cohorts for genetic studies

Two unique cohorts (Table I) were available to test this novel candidate gene hypothesis. These cohorts are from the European Multicentre Cohort Study on Congenital Toxoplasmosis (EMSCOT), which recruited prospectively for mothers with primary *T. gondii* in pregnancy (Gilbert & Gras 2003, Gras et al. 2005), and from the National Collaborative Chicago-based Congenital Toxoplasmosis Study (NCCCTS) (McLeod et al. 2006), in North America, to which infants and children with congenital infection with *T. gondii* are referred.

The EMSCOT cohort - For the EMSCOT cohort, DNA was successfully obtained from 457 mother-child pairs with confirmed infection during pregnancy. Transmission of infection during pregnancy was confirmed in 232 (51%) infants; 225 (49%) infants remained uninfected. All infants were monitored for clinical signs until at least three years of age and many have been followed for > 10 years. Of the 232 infected infants, 79 (34%) had clinical signs (referred to as affected) of congenital toxoplasmosis: 53 (67%) with ocular lesions (retinochoroidal lesions), 45 (57%) with brain lesions (hydrocephalus or intracranial calcifications detected on ultrasound examination of the brain), 19 (24%) of these infants had both eye and brain lesions. One hundred fifty three (66%) infected infants had no clinical signs of disease up to a minimum follow-up period of three years of age. Two-thirds of mother-child pairs reported grandparents' countries of birth. Of these, 95% were Caucasian and 5% of African origin.

The NCCCTS cohort - For the second independently ascertained NCCCTS cohort, DNA was successfully obtained from 149 children with confirmed congenital infection (69% Caucasian, 15% Hispanic, 8% Asian or Pacific Islander, 3% African American, 0.7% Native American, 4.7% mixed race) plus available parents. At birth or time of diagnosis, 92 (62%) infected children had brain calcifications with/without hydrocephalus and retinal lesions, 21 (14%) had retinal lesions only, 11 (7%) had brain calcifications with/without hydrocephalus

only and 25 (17%) infected children were without these clinical findings. Only the 124 children with confirmed clinical findings in eye and/or brain were included in the allelic association analysis for this cohort. This provided data on 124 trios classified as affected (i.e. eye or brain lesions or both), 113 trios classified as eye disease (with or without brain disease) and 103 trios classified as brain disease (with or without eye disease).

Results of genetic analyses

Seven single nucleotide polymorphisms (SNPs) at *ABCA4* and seven SNPs at *COL2A1* (Jamieson et al. 2008) were genotyped in both cohorts.

Genetic associations with clinical outcome

For the EMSCOT cohort, nominally significant (i.e. p values without correction for multiple testing) allelic associations were observed for SNPs at *ABCA4* and at *COL2A1* when affected children (i.e. children with retinal or brain disease or both) in the EMSCOT cohort were compared with infected but unaffected children. Large effect sizes (odds ratios or OR 2.9-6.8) were observed for these associations, particularly in children or homozygous for the disease allele (Table II). Significance at both genes was improved when the analysis was enriched for children with the eye lesion phenotype (with/without brain lesions, i.e. leaving out children with brain disease only) compared to unaffected children, but not when the analysis was similarly enriched for children with brain lesions (with/without eye lesions, leaving out children with eye disease only) compared to unaffected children.

In the NCCCTS cohort, significant allelic associations were observed at *ABCA4* and *COL2A1* under a dominant model of inheritance (Table III). The use of case/parent trios and transmission disequilibrium testing in FBAT controlled for ethnic admixture in this cohort. For *COL2A1*, significance improved when the analysis was enriched for the eye lesion phenotype (with/without brain lesions), with the associated markers all within one strong haplotype block.

Overall, the evidence from the children with eye and brain signs associated with congenitally acquired toxoplasmosis in these two cohorts was for association with eye disease and these two previously recognized eye disorder genes, *ABCA4* and *COL2A1*.

TABLE I
Comparison of the EMSCOT and USA study samples

EMSCOT		NCCCTS	
457 mother-child pairs; mothers screened and confirmed infected during pregnancy		149 child-parent trios; no screening during pregnancy	
232 children infected <i>in utero</i>		All children inferred infected <i>in utero</i>	
79 children with ocular or brain clinical signs	153 children born without clinical signs	124 children with ocular or brain clinical signs	25 children with other clinical signs
34 with ocular signs only		21 with ocular signs only	
26 with brain signs only		11 with brain signs only	
19 with both		92 with both	
225 children not infected <i>in utero</i>			

TABLE II
Logistic regression analysis for genotype-wise associations between *ABCA4* and *COL2A1* and risk of congenital toxoplasmosis in the EMSCOT mother-child pairs

SNP Database Identity	Genotype	OR disease child			OR disease mother ^a		
		OR	95% CI	p	OR	95% CI	p
<i>ABCA4</i> .rs2297633	GG	6.81	1.46-31.7	0.0145	8.84	1.88-41.6	0.0058
	GT	4.74	0.98-23.0	0.0534	3.19	0.65-15.8	0.1542
<i>ABCA4</i> .rs1761375	GG	1.98	0.69-5.61	0.1998	8.92	1.94-41.1	0.0050
	GA	1.43	0.48-4.28	0.5252	3.03	0.63-14.6	0.1660
<i>ABCA4</i> .rs3112831	CC	2.86	1.02-7.93	0.0439	3.12	1.01-9.60	0.0472
	CT	2.13	0.73-6.22	0.1670	1.62	0.51-5.21	0.4165
<i>ABCA4</i> .rs952499 ^a	TT	2.20	0.96-5.09	0.0634	3.66	1.50-8.89	0.0042
	TC	1.50	0.67-3.37	0.3203	1.61	0.71-3.61	0.2526
	GA	3.63	1.29-10.21	0.0144	3.98	1.45-10.97	0.0074
<i>COL2A1</i> .rs2276455	GG	3.63	1.29-10.21	0.0144	3.98	1.45-10.97	0.0074
	GA	3.13	1.12-8.77	0.0292	1.67	0.63-4.41	0.3038

a: OR for mothers infected during pregnancy having an infected child with clinical signs; CI: 95% confidence intervals; OR: odds ratio; p: probability, nominal $p \leq 0.05$ is shown in bold; SNP: single nucleotide polymorphism. All analyses for the EMSCOT study were adjusted for country of birth, trimester of seroconversion and country of origin of grandparents as a surrogate for ethnicity. Results are for genotype associations comparing affected children against unaffected children from infected mothers and mothers of affected children against mothers of unaffected children. Adapted from Jamieson et al. (2008).

TABLE III

FBAT analysis under dominant model of inheritance for associations between SNPs at *ABCA4* or *COL2A1* and congenital toxoplasmosis in the NCCCTS child-parent trios

Gene/SNP	Allele ^a	Z score	p
<i>ABCA4</i> /rs952499 ^a	C	- 2.255	0.024
<i>COL2A1</i> /rs6823	G	+ 1.949	0.051
<i>COL2A1</i> /rs2070739	T	+ 2.556	0.011
<i>COL2A1</i> /rs2276454 ^b	A	- 2.682	0.007
<i>COL2A1</i> /rs1635544 ^c	C	- 2.269	0.023
<i>COL2A1</i> /rs3803183	T	+ 2.449	0.014

a: OR for heterozygote compared to homozygous disease allele is 0.37 (95% confidence intervals: 0.14-0.92; $p = 0.032$); b: OR for carrying disease allele is 2.45 (1.22-4.94; $p = 0.012$); c: OR 2.6 (1.25-5.39; $p = 0.010$). A positive Z score indicates association with disease; a negative Z score indicates the non-associated allele; only Z scores for $p < 0.1$ are shown. The use of trios in the NCCCTS cohort was robust to ethnic admixture. Adapted from Jamieson et al. (2008).

Influence of mother's genotype on clinical outcome

To determining whether mother's genotype had an effect on disease outcome in the child, logistic regression analysis was carried out comparing mothers of affected children (eye and/or brain disease) with mothers of infected unaffected children for the EMSCOT cohort. This proved interesting in that (i) most of the associations were more significant in comparisons of mothers than in comparisons of infants (Table II), again with large effect sizes (OR 3.1-8.9), (ii) additional *ABCA4* and *COL2A1* were associated with mothers of affected infants compared to mothers of unaffected infants and (iii) all of the *ABCA4* SNPs but none of the *COL2A1* SNPs

were associated with mothers of infants with brain disease suggesting a role for *ABCA4* in brain pathology. Statistically, it appeared that genetic effects of these two loci on disease in the child were being diluted out in making direct comparisons of the affected versus unaffected children relative to the evidence for association when comparing the mothers.

Unusual patterns of inheritance observed when comparing genotype-wise associations in children versus mothers suggested that there may be a direct effect of mother's genotype, or that there may be parent-of-origin effects (imprinting), i.e. that for the child it is the origin of, and not just the combination of, alleles that is important in determining disease risk. Using a log-linear method previously designed to evaluate maternal genotype and/or parent-of-origin effects in case-parent trios (Weinberg et al. 1998), evidence for imprinting given mother's and child's genotypes was obtained at *ABCA4* ($p = 0.033$) and *COL2A1* ($p = 0.05$) in the NCCCTS trios. For EMSCOT we adapted the method for use with mother-child pairs and found evidence for effects of maternal genotype given child's genotype and imprinting for two SNPs at *COL2A1* ($p = 0.0088$ and $p = 6.87 \times 10^{-5}$) and evidence for imprinting taking account of both mother's and child's genotype for one *COL2A1* SNP ($p = 0.0025$). Given the statistical limitations of small sample size and power in the modeling analysis and the fact that a direct effect of mother's genotype seemed unlikely biologically, we looked for experimental evidence of epigenetic effects or imprinting for *ABCA4* and *COL2A1*.

Evidence for imprinting

To obtain experimental evidence for imprinting, an anonymised EBV B cell bank was screened (i) to find individuals heterozygous for exonic SNPs at both loci and (ii) to determine mono-allelic expression in elicited transcripts obtained from RNA from these EBV cell

lines. EBV lines were identified that were heterozygous for SNP rs3112831 in exon 10 of *ABCA4* and for SNP rs3737548 in exon 7 of *COL2A1*. For the known functional exon 10-containing isoform at *ABCA4*, four of the five EBV lines that were heterozygous for genomic DNA showed monoallelic expression in cDNA for the exon 10 rs3112831 SNP (Fig. 1A). The most likely explanation for this was polymorphic imprinting of the paternally-derived allele. Monoallelic expression was also observed (Fig. 1B) for a SNP in exon 7 in cDNA specifically amplified for the IIB short form of *COL2A1*, but not for cDNA amplified for the IIA long form, consistent with isoform-specific imprinting in this case with the maternally-derived allele silenced.

DISCUSSION

In our published study (Jamieson et al. 2008) we examined the specific genetic hypothesis that polymorphisms in two genes known to be associated with ocular disorders, *ABCA4* and *COL2A1*, are associated with eye disease caused by congenital toxoplasmosis. These associations were replicated across two independently ascertained cohorts, providing further confidence in this result. How these findings are related to the direct functional roles of *ABCA4* and *COL2A1* are discussed in more detail in the original paper. The observation that both loci were influenced by epigenetic modification of gene expression was also novel and exciting. The study provided new insight into processes that occur in early in embryonic development when it is not easy to determine what is happening when the fetus is first infected with a parasite such as *T. gondii*. The specific genetic associa-

tions observed have seeded follow-up studies to address how the parasite triggers the development of ocular disease during embryogenesis and fetal development. What is known about the chronological expression and localization of *COL2A1* and *ABCA4* during development? How does the parasite influence this? Since *T. gondii* is a potent trigger for, and direct regulator of, the NFκB signaling pathway (Molestina & Sinai 2005a, b), could the parasite trigger disease via polymorphic NFκB binding in the promoter regions of these two genes? Does the parasite directly interfere with methylation and/or histone acetylation patterns of host DNA, thereby directly affecting epigenetic regulation of gene expression? Studies in progress are designed to test some of these hypotheses.

In this initial genetic study of congenital toxoplasmosis we took a specific hypothesis-driven candidate gene approach. Such an approach could also be applied, for example, to developmental genes previously associated with other clinical outcomes of congenital toxoplasmosis such as hydrocephalus. Alternatively, given a sufficiently large sample size, further studies could take a non-hypothesis-driven genome-wide approach to identify critical developmental pathways and pathogenic processes associated with congenital disease. Since other congenital infections, e.g. cytomegalovirus, often elicit similar pathologies, it would also be of interest to test both this hypothesis-driven candidate gene analysis to determine whether the same or different genes influence specific pathologies (such as eye or brain lesions) and/or genome-wide approaches across cohorts of different congenital and perinatally-acquired infectious diseases. Understanding more about clinical outcomes associated with *T. gondii* infection acquired in childhood could also benefit from the application of the tools of modern post-genomic genetic and epigenetic analysis.

One major problem with all candidate gene studies for infectious diseases reported to date is that they were under-powered (Burgner et al. 2006). Until recently, this was a general problem in genetic analysis of complex disease, along with issues relating to study design and population history (Cordell & Clayton 2005, Palmer & Cardon 2005). Fig. 2 compares power to detect association at OR = 1.5 or OR = 2, given different risk allele frequencies, p values and sample sizes. This shows that 500 trios or case-control pairs have little power to detect association for small effect sizes (OR = 1.5). Even with 1,000 trios or case-control pairs, power is limited for risk alleles with frequency < 0.2 for an effect size (OR) 1.5, although low frequency (e.g., 0.10) risk alleles with larger effect sizes (OR > 2) may be detected. This means that the studies we can perform on congenital toxoplasmosis using the current EMSCOT and NCCCTS resources are only powered to detect large effect sizes (OR > 2). To carry out further candidate gene studies of congenital or acquired toxoplasmosis with confidence, the research community needs to work together to develop much larger cohorts for primary and replication studies. This applies also to the application of SNP-chip based genome-wide association studies which have proved to be quite successful in rapidly increasing the number of loci that have been positively associated with complex diseases. For example, the Consortium WTCCC study

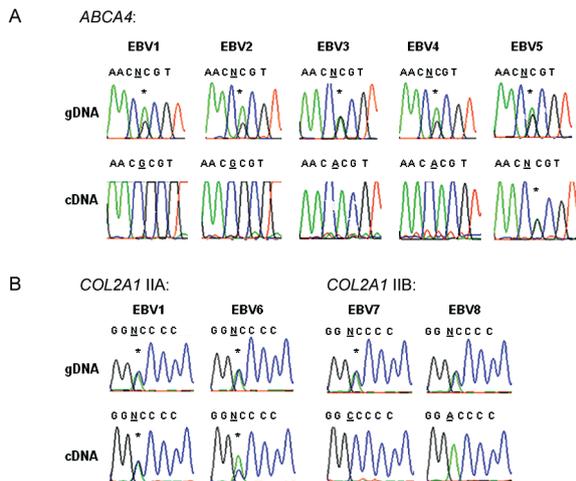


Fig. 1: experimental evidence for monoallelic expression at *ABCA4* and *COL2A1*. A and B show sequence analysis of genomic DNA (gDNA) and cDNA in EBV lines heterozygous for exonic SNPs. A: EBV lines heterozygous for *ABCA4* rs3112831 in gDNA; lines EBV1 to EBV4 homozygous (i.e. monoallelic) in cDNA specific for the exon 10-containing isoform. Line EBV5 is heterozygous, indicating that monoallelic silencing is polymorphic; B: EBV lines show monoallelic expression for *COL2A1* SNP rs3737548 in PCR products specific for isoform IIB, but not IIA. Positions of SNPs (*) indicated by N where heterozygous, with the bp underlined for monoallelic expression in cDNA. Adapted from Jamieson et al. (2008).

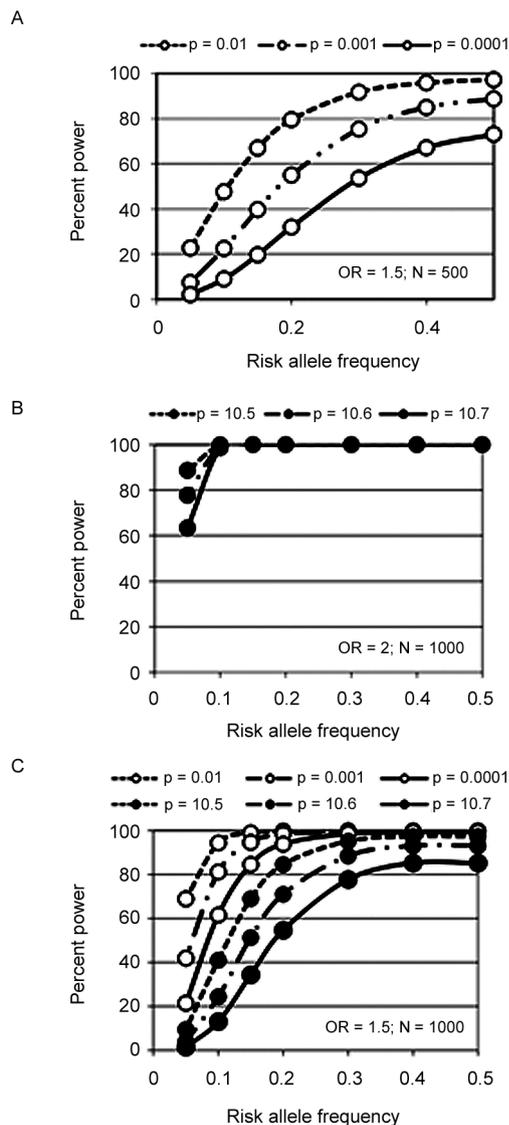


Fig. 2: percent power (Y-axis) of $N = 500$ (A) or $N = 1000$ (B, C) trios of case-control pairs to detect allelic association for a risk allele with effect size (OR) 1.5 (A, C) or 2 (B), given different risk allele frequencies (X-axis) and p values (10^{-2} - 10^{-7} as indicated on key for different lines on each graph). Greater robustness to type one error, i.e. lower thresholds for p values, is required for GWAS.

(2007) of 14,000 cases of seven common diseases and 3,000 shared controls has itself identified 24 independent association signals at $p < 5 \times 10^{-7}$, nine of which were in Crohn's disease, three in rheumatoid arthritis, seven in type 1 diabetes and three in type 2 diabetes. The increased problem of multiple testing associated with genotyping in excess of 500,000 SNPs in large numbers of cases and controls necessitated this stringent threshold p value $< 5 \times 10^{-7}$ to achieve genome-wide significance. However, across all diseases, a large number of further signals, including 58 loci with single point p values between 10^{-5} - 5×10^{-7} , were identified which are likely to yield additional susceptibility loci. A number of papers providing validation of the original WTCCC data

have already been published (Cooper et al. 2007, Parkes et al. 2007, Plenge et al. 2007, Samani et al. 2007, Thomson et al. 2007, Todd et al. 2007, Zeggini et al. 2007, Beckly et al. 2008), including one on type 1 diabetes that increases the number of loci with compelling evidence of association from six to at least 10 (Todd et al. 2007). The WTCCC study is but one of an increasing number of published GWAS for complex diseases (Buch et al. 2007, Burton et al. 2007, Cupples et al. 2007, Duggan et al. 2007, Easton et al. 2007, Florez et al. 2007, Gudmundsson et al. 2007, Hampe et al. 2007, Kathiresan et al. 2007, Raelson et al. 2007, Roberts et al. 2007, Scuteri et al. 2007, Sladek et al. 2007, Sulem et al. 2007, Tomlinson et al. 2007, Zanke et al. 2007, Cronin et al. 2008). As pointed out in the recent News Feature on "Genetics by Numbers" in Nature (Baker 2008), identification of the initial SNP association on a primary GWAS is only the first step on the path to validating and identifying the etiological genes and associated mechanisms of disease susceptibility. Some papers have demonstrated at least two independent etiological variants at one locus (6q23) associated with rheumatoid arthritis, while regions like the HLA complex require intensive fine mapping. For example, following the WTCCC type 1 diabetes study, a combined total of 1,729 polymorphisms across HLA were genotyped in > 6000 cases and controls across cohorts and statistical methods of recursive partitioning and regression applied to pinpoint disease susceptibility to the MHC class I genes HLA-B and HLA-A [risk ratios > 1.5 ; p (combined) = 2.01×10^{-19} and 2.35×10^{-13} , respectively] in addition to the established associations of the MHC class II HLA-DQB1 and HLA-DRB1 genes (Nejentsev et al. 2007). This demonstrates the intensive mapping that must follow any primary GWAS to validate the SNP associations and identify the genes associated with disease and the statistical power that can be achieved by genotyping large numbers of cases across multiple cohorts.

The application of GWAS technology to the study of complex disease has also stimulated development of genetic statistical tools, with recent papers highlighting the use of localized haplotype cluster analysis to improve detection of disease associated variants (Browning & Brownin 2008), better methods for correcting for population stratification by identifying hidden population structures (Li & Yu 2008), robust methods to allow primary screening and replication using the same data set in family-based association testing (Van Steen et al. 2005), a new multipoint method for GWAS by imputation of genotypes based on knowledge of tag-SNPs (Marchini et al. 2007) and the effect of local LD patterns around the primary analysis signal(s) and tagging efficacy of typed markers in determining a suitable density of tag-SNPs for fine-mapping (Wiltshire et al. 2008). All of these have implications for study design, which given the global distribution of toxoplasmosis, need to be carefully planned. Recent papers discuss many of the issues that affect modern genome-wide approaches to genetic association studies using case-control and trio designs (Teo 2008, Teo et al. 2008a, b, 2009a, b). Case-control approaches are more efficient in terms of the amount of genotyping required, but will not permit detection of maternal genotype effect, maternal-foetal interactions

or parent-of-origin (imprinting) effects (Buyske 2008). Family-based studies (minimally mother-child pairs, or child/parent trios) make it possible to distinguish between these different mechanisms (Weinberg et al. 1998, Cordell et al. 2004, Hsieh et al. 2006) and are also valuable in controlling for mixed ethnicity or ethnic admixture in the population under study. Our specific study of congenital toxoplasmosis has demonstrated how important these effects can be. It is also of note that, despite the exciting results of highly-powered genome-wide case-control association studies, the genes identified to date generally fail to account for all of the attributable genetic risk. This suggests that many genes are being missed by failure to take these other models and the huge role that epigenetic effects are likely to play, into account.

In summary, we believe that the application of a genetic approach has provided novel insight into disease mechanisms in congenital toxoplasmosis. We hope this will seed new collaborative consortia that will allow application of this technology with greater efficiency and with adequate statistical power to determine associations of relevance to our understanding of the underlying mechanisms of disease.

ACKNOWLEDGMENTS

To the patients, their families and their physicians, for their participation in EMSCOT and the NCCCTS.

REFERENCES

- Baker M 2008. Genome studies: genetics by numbers. *Nature* 451: 516-518.
- Beckly JB, Hancock L, Geremia A, Cummings JR, Morris A, Cooney R, Pathan S, Guo C, Jewell DP 2008. Two-stage candidate gene study of chromosome 3p demonstrates an association between nonsynonymous variants in the MST1R gene and Crohn's disease. *Inflamm Bowel Dis* 14: 500-507.
- Browning BL, Browning SR 2008. Haplotypic analysis of Wellcome Trust Case Control Consortium data. *Hum Genet* 123: 273-280.
- Buch S, Schafmayer C, Volzke H, Becker C, Franke A, von Eller-Eberstein H, Kluck C, Bassmann I, Brosch M, Lammert F, Miquel JF, Nervi F, Wittig M, Roskopf D, Timm B, Holl C, Seeger M, ElSharawy A, Lu T, Egberts J, Fandrich F, Folsch UR, Krawczak M, Schreiber S, Nurnberg P, Tepel J, Hampe J 2007. A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease. *Nat Genet* 39: 995-999.
- Burgner D, Jamieson SE, Blackwell JM 2006. Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better? *Lancet Infect Dis* 6: 653-663.
- Buyske S 2008. Maternal genotype effects can alias case genotype effects in case-control studies. *Eur J Hum Genet* 16: 784-785.
- Consortium WTCC 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447: 661-678.
- Cooper JD, Smyth DJ, Bailey R, Payne F, Downes K, Godfrey LM, Masters J, Zeitels LR, Vella A, Walker NM, Todd JA 2007. The candidate genes TAF5L, TCF7, PDCD1, IL6 and ICAM1 cannot be excluded from having effects in type 1 diabetes. *BMC Med Genet* 8: 71.
- Cordell HJ, Barratt BJ, Clayton DG 2004. Case/pseudocontrol analysis in genetic association studies: a unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genet Epidemiol* 26: 167-185.
- Cordell HJ, Clayton DG 2005. Genetic association studies. *Lancet* 366: 1121-1131.
- Cronin S, Berger S, Ding J, Schymick JC, Washecka N, Hernandez DG, Greenway MJ, Bradley DG, Traynor BJ, Hardiman O 2008. A genome-wide association study of sporadic ALS in a homogenous Irish population. *Hum Mol Genet* 17: 768-774.
- Cupples LA, Arruda HT, Benjamin EJ, D'Agostino RB, Sr., Demissie S, DeStefano AL, Dupuis J, Falls KM, Fox CS, Gottlieb DJ, Govindaraju DR, Guo CY, Heard-Costa NL, Hwang SJ, Kathiresan S, Kiel DP, Laramie JM, Larson MG, Levy D, Liu CY, Lunetta KL, Mailman MD, Manning AK, Meigs JB, Murabito JM, Newton-Cheh C, O'Connor GT, O'Donnell CJ, Pandey M, Seshadri S, Vasan RS, Wang ZY, Wilk JB, Wolf PA, Yang Q, Atwood LD 2007. The Framingham Heart Study 100K SNP genome-wide association study resource: overview of 17 phenotype working group reports. *BMC Med Genet* 8 (Suppl. 1): S1.
- Davey Smith G, Timpson N, Ebrahim S 2008. Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization. *Ann Med* 40: 524-541.
- Ducroq D, Rozet JM, Gerber S, Perrault I, Barbet D, Hanein S, Hakiki S, Dufier JL, Munnich A, Hamel C, Kaplan J 2002. The ABCA4 gene in autosomal recessive cone-rod dystrophies. *Am J Hum Genet* 71: 1480-1482.
- Duggan D, Zheng SL, Knowlton M, Benitez D, Dimitrov L, Wiklund F, Robbins C, Isaacs SD, Cheng Y, Li G, Sun J, Chang BL, Marovich L, Wiley KE, Balter K, Stattin P, Adami HO, Gielzak M, Yan G, Sauvageot J, Liu W, Kim JW, Bleecker ER, Meyers DA, Trock BJ, Partin AW, Walsh PC, Isaacs WB, Gronberg H, Xu J, Carpten JD 2007. Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. *J Natl Cancer Inst* 99: 1836-1844.
- Dunn D, Wallon M, Peyron F, Petersen E, Peckham C, Gilbert R 1999. Mother-to-child transmission of toxoplasmosis: risk estimates for clinical counselling. *Lancet* 353: 1829-1833.
- Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struwing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O, Johnson N, Seal S, Stratton MR, Rahman N, Chenevix-Trench G, Bojesen SE, Nordestgaard BG, Axelsson CK, Garcia-Closas M, Brinton L, Chanock S, Lissowska J, Peplonska B, Nevanlinna H, Fagerholm R, Eerola H, Kang D, Yoo KY, Noh DY, Ahn SH, Hunter DJ, Hankinson SE, Cox DG, Hall P, Wedren S, Liu J, Low YL, Bogdanova N, Schurmann P, Dork T, Tollenaar RA, Jacobi CE, Devilee P, Klijn JG, Sigurdson AJ, Doody MM, Alexander BH, Zhang J, Cox A, Brock IW, MacPherson G, Reed MW, Couch FJ, Goode EL, Olson JE, Meijers-Heijboer H, van den Ouweland A, Uitterlinden A, Rivadeneira F, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Hopper JL, McCredie M, Southey M, Giles GG, Schroen C, Justenhoven C, Brauch H, Hamann U, Ko YM, Spurdle AB, Beesley J, Chen X, Mannermaa A, Kosma VM, Kataja V, Hatakainen J, Day NE, Cox DR, Ponder BA 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447: 1087-1093.
- Ebrahim S, Davey Smith G 2008. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Hum Genet* 123: 15-33.
- Florez JC, Manning AK, Dupuis J, McAteer J, Irenze K, Gianniny L, Mirel DB, Fox CS, Cupples LA, Meigs JB 2007. A 100K genome-wide association scan for diabetes and related traits in the Framingham Heart Study: replication and integration with other genome-wide datasets. *Diabetes* 56: 3063-3074.

- Gilbert R, Gras L 2003. Effect of timing and type of treatment on the risk of mother to child transmission of *Toxoplasma gondii*. *Bjog* 110: 112-120.
- Gras L, Wallon M, Pollak A, Cortina-Borja M, Evengard B, Hayde M, Petersen E, Gilbert R 2005. Association between prenatal treatment and clinical manifestations of congenital toxoplasmosis in infancy: a cohort study in 13 European centres. *Acta Paediatr* 94: 1721-1731.
- Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, Rafnar T, Bergthorsson JT, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Xu J, Blondal T, Kostic J, Sun J, Ghosh S, Stacey SN, Mouy M, Saemundsdottir J, Backman VM, Kristjansson K, Tres A, Partin AW, Albers-Akkers MT, Godino-Ivan Marcos J, Walsh PC, Swinkels DW, Navarrete S, Isaacs SD, Aben KK, Graif T, Cashy J, Ruiz-Echarri M, Wiley KE, Suarez BK, Witjes JA, Frigge M, Ober C, Jonsson E, Einarsson GV, Mayordomo JI, Kiemeny LA, Isaacs WB, Catalona WJ, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K 2007. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* 39: 631-637.
- Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, Albrecht M, Mayr G, De La Vega FM, Briggs J, Gunther S, Prescott NJ, Onnie CM, Hasler R, Sipos B, Folsch UR, Lengauer T, Platzer M, Mathew CG, Krawczak M, Schreiber S 2007. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 39: 207-211.
- Hsieh HJ, Palmer CG, Harney S, Newton JL, Wordsworth P, Brown MA, Sinsheimer JS 2006. The v-MFG test: investigating maternal, offspring and maternal-fetal genetic incompatibility effects on disease and viability. *Genet Epidemiol* 30: 333-347.
- Jamieson SE, de Roubaix LA, Cortina-Borja M, Tan HK, Mui EJ, Cordell HJ, Kirisits MJ, Miller EN, Peacock CS, Hargrave AC, Coyne JJ, Boyer K, Bessieres MH, Buffolano W, Ferret N, Franck J, Kieffer F, Meier P, Nowakowska DE, Paul M, Peyron F, Stray-Pedersen B, Prusa AR, Thulliez P, Wallon M, Petersen E, McLeod R, Gilbert RE, Blackwell JM 2008. Genetic and epigenetic factors at *COL2A1* and *ABCA4* influence clinical outcome in congenital toxoplasmosis. *PLoS ONE* 3: e2285.
- Kathiresan S, Manning AK, Demissie S, D'Agostino RB, Surti A, Guiducci C, Giannini L, Burt NP, Melander O, Orho-Melander M, Arnett DK, Peloso GM, Ordovas JM, Cupples LA 2007. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med Genet* 8 (Suppl. 1): S17.
- Koenekoop RK 2003. The gene for Stargardt disease, *ABCA4*, is a major retinal gene: a mini-review. *Ophthalmic Genet* 24: 75-80.
- Li Q, Yu K 2008. Improved correction for population stratification in genome-wide association studies by identifying hidden population structures. *Genet Epidemiol* 32: 215-226.
- Mack DG, Johnson JJ, Roberts F, Roberts CW, Estes RG, David C, Grumet FC, McLeod R 1999. HLA-class II genes modify outcome of *Toxoplasma gondii* infection. *Int J Parasitol* 29: 1351-1358.
- Marchini J, Howie B, Myers S, McVean G, Donnelly P 2007. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39: 906-913.
- McLeod R, Boyer K, Karrison T, Kasza K, Swisher C, Roizen N, Jalbrzikowski J, Remington J, Heydemann P, Noble AG, Mets M, Holfels E, Withers S, Latkany P, Meier P 2006. Outcome of treatment for congenital toxoplasmosis, 1981-2004: the National Collaborative Chicago-Based, Congenital Toxoplasmosis Study. *Clin Infect Dis* 42: 1383-1394.
- Molestina RE, Sinai AP 2005a. Detection of a novel parasite kinase activity at the *Toxoplasma gondii* parasitophorous vacuole membrane capable of phosphorylating host I κ B α . *Cell Microbiol* 7: 351-362.
- Molestina RE, Sinai AP 2005b. Host and parasite-derived IKK activities direct distinct temporal phases of NF- κ B activation and target gene expression following *Toxoplasma gondii* infection. *J Cell Sci* 118: 5785-5796.
- Nejentsev S, Howson JM, Walker NM, Szeszko J, Field SF, Stevens HE, Reynolds P, Hardy M, King E, Masters J, Hulme J, Maier LM, Smyth D, Bailey R, Cooper JD, Ribas G, Campbell RD, Clayton DG, Todd JA 2007. Localization of type 1 diabetes susceptibility to the MHC class I genes *HLA-B* and *HLA-A*. *Nature* 450: 887-892.
- Palmer LJ, Cardon LR 2005. Shaking the tree: mapping complex disease genes with linkage disequilibrium. *Lancet* 366: 1223-1234.
- Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, Roberts RG, Nimmo ER, Cummings FR, Soars D, Drummond H, Lees CW, Khawaja SA, Bagnall R, Burke DA, Todhunter CE, Ahmad T, Onnie CM, McArdle W, Strachan D, Bethel G, Bryan C, Lewis CM, Deloukas P, Forbes A, Sanderson J, Jewell DP, Satsangi J, Mansfield JC, Cardon L, Mathew CG 2007. Sequence variants in the autophagy gene *IRGM* and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 39: 830-832.
- Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burt NP, Blumenstiel B, DeFelice M, Parkin M, Barry R, Winslow W, Healy C, Graham RR, Neale BM, Izmailova E, Roubenoff R, Parker AN, Glass R, Karlson EW, Maher N, Haffner DA, Lee DM, Seldin MF, Remmers EF, Lee AT, Padyukov L, Alfredsson L, Cobyln J, Weinblatt ME, Gabriel SB, Purcell S, Klareskog L, Gregersen PK, Shadick NA, Daly MJ, Altshuler D 2007. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nat Genet* 39: 1477-1482.
- Raelson JV, Little RD, Ruether A, Fournier H, Paquin B, Van Eerdewegh P, Bradley WE, Croteau P, Nguyen-Huu Q, Segal J, Debrus S, Allard R, Rosenstiel P, Franke A, Jacobs G, Nikolaus S, Vidal JM, Szego P, Laplante N, Clark HF, Paulussen RJ, Hooper JW, Keith TP, Belouchi A, Schreiber S 2007. Genome-wide association study for Crohn's disease in the Quebec Founder Population identifies multiple validated disease loci. *Proc Natl Acad Sci USA* 104: 14747-14752.
- Remington JS, McLeod R, Thullie P, Desmonts G 2005. Toxoplasmosis. In JS Remington, C Baker, E Wilson, JO Klein, *Infectious diseases of the fetus and newborn infant*, 6th ed., WB Saunders, Philadelphia, p. 947-1091.
- Roberts R, Stewart AF, Wells GA, Williams KA, Kavaslar N, McPherson R 2007. Identifying genes for coronary artery disease: an idea whose time has come. *Can J Cardiol* 23 (Suppl. A): 7A-15A.
- Rose PS, Levy HP, Liberfarb RM, Davis J, Szymko-Bennett Y, Rubin BI, Tsilou E, Griffith AJ, Francomano CA 2005. Stickler syndrome: clinical characteristics and diagnostic criteria. *Am J Med Genet A* 138: 199-207.
- Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, Konig IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H 2007. Genomewide association analysis of coronary artery disease. *N Engl J Med* 357: 443-453.
- Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orru M, Usala G, Dei M, Lai S, Maschio A, Busonero F, Mulas A, Ehret GB, Fink AA, Weder AB, Cooper RS, Galan P, Chakravarti A, Schlessinger D, Cao A, Lakatta E, Abecasis GR 2007. Genome-wide association scan shows genetic variants in the *FTO* gene are associated with obesity-related traits. *PLoS Genet* 3: e115.
- Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P,

- Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881-885.
- Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, Jakobsdottir M, Steinberg S, Palsson S, Jonasson F, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsson KR, Aben KK, Kiemeny LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K 2007. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 39: 1443-1452.
- Teo YY 2008. Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. *Curr Opin Lipidol* 19: 133-143.
- Teo YY, Fry AE, Sanjoaquin MA, Pederson B, Small KS, Rockett KA, Kwiatkowski DP, Clark TG 2009a. Assessing genuine parents-offspring trios for genetic association studies. *Hum Hered* 67: 26-37.
- Teo YY, Inouye M, Small KS, Fry AE, Potter SC, Dunstan SJ, Seielstad M, Barroso I, Wareham NJ, Rockett KA, Kwiatkowski DP, Deloukas P 2008a. Whole genome-amplified DNA: insights and imputation. *Nat Methods* 5: 279-280.
- Teo YY, Small KS, Clark TG, Kwiatkowski DP 2008b. Perturbation analysis: a simple method for filtering SNPs with erroneous genotyping in genome-wide association studies. *Ann Hum Genet* 72: 368-374.
- Teo YY, Small KS, Fry AE, Wu Y, Kwiatkowski DP, Clark TG 2009b. Power consequences of linkage disequilibrium variation between populations. *Genet Epidemiol Epub* 33: 128-135.
- Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, Donn R, Symmons D, Hider S, Bruce IN, Wilson AG, Marinou I, Morgan A, Emery P, Carter A, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Strachan D, Worthington J 2007. Rheumatoid arthritis association at 6q23. *Nat Genet* 39: 1431-1433.
- Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, Bailey R, Nejentsev S, Field SF, Payne F, Lowe CE, Szeszko JS, Hafler JP, Zeitel S, Yang JH, Vella A, Nutland S, Stevens HE, Schuilenburg H, Coleman G, Maisuria M, Meadows W, Smink LJ, Healy B, Burren OS, Lam AA, Ovington NR, Allen J, Adlem E, Leung HT, Wallace C, Howson JM, Guja C, Ionescu-Tirgoviste C, Simmonds MJ, Heward JM, Gough SC, Dunger DB, Wicker LS, Clayton DG 2007. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 39: 857-864.
- Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, Penegar S, Chandler I, Gorman M, Wood W, Barclay E, Lubbe S, Martin L, Sellick G, Jaeger E, Hubner R, Wild R, Rowan A, Fielding S, Howarth K, Silver A, Atkin W, Muir K, Logan R, Kerr D, Johnstone E, Sieber O, Gray R, Thomas H, Peto J, Cazier JB, Houlston R 2007. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet* 39: 984-988.
- Van Steen K, McQueen MB, Herbert A, Raby B, Lyon H, Demeo DL, Murphy A, Su J, Datta S, Rosenow C, Christman M, Silverman EK, Laird NM, Weiss ST, Lange C 2005. Genomic screening and replication using the same data set in family-based association testing. *Nat Genet* 37: 683-691.
- Weinberg CR, Wilcox AJ, Lie RT 1998. A log-linear approach to case-parent-triad data: assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. *Am J Hum Genet* 62: 969-978.
- Wellcome Trust Case Control Consortium; Australo-Anglo-American Spondylitis Consortium (TASC), Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, Todd JA, Donnelly P, Barrett JC, Davison D, Easton D, Evans DM, Leung HT, Marchini JL, Morris AP, Spencer CC, Tobin MD, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan MC, Owen MJ, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop TD, Iles MM, Maqbool A, Yuldasheva N, Hall AS, Braund PS, Dixon RJ, Mangino M, Stevens S, Thompson JR, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Matthew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP, Webster J, Brown MJ, Lathrop MG, Connell J, Dominiczak A, Marcano CA, Burke B, Dobson R, Gungadoo J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue M, Caulfield M, Farrall M, Barton A; Biologics in RA Genetics and Genomics Study Syndicate (BRAGGS) Steering Committee, Bruce IN, Donovan H, Eyre S, Gilbert PD, Hilder SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DP, Thomson W, Worthington J, Dunger DB, Widmer B, Frayling TM, Freathy RM, Lango H, Perry JR, Shields BM, Weedon MN, Hattersley AT, Hitman GA, Walker M, Elliott KS, Groves CJ, Lindgren CM, Rayner NW, Timpson NJ, Zeggini E, Newport M, Sirugo G, Lyons E, Vannberg F, Hill AV, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Franklyn JA, Heward JM, Simmonds MJ, Gough SC, Seal S; Breast Cancer Susceptibility Collaboration (UK), Stratton MR, Rahman N, Ban M, Goris A, Sawcer SJ, Compston A, Conway D, Jallow M, Newport M, Sirugo G, Rockett KA, Bumpstead SJ, Chaney A, Downes K, Ghorri MJ, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widdon C, Withers D, Cardin NJ, Davison D, Ferreira T, Pereira-Gale J, Hallgrimsdottir IB, Howie BN, Su Z, Teo YY, Vukcevic D, Bentley D, Brown MA, Compston A, Farrall M, Hall AS, Hattersley AT, Hill AV, Parkes M, Pembrey M, Stratton MR, Mitchell SL, Newby PR, Brand OJ, Carr-Smith J, Pearce SH, McGinnis R, Keniry A, Deloukas P, Reveille JD, Zhou X, Sims AM, Dowling A, Taylor J, Doan T, Davis JC, Savage L, Ward MM, Leach TL, Weisman MH, Brown M. 2007. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 39: 1329-1337.
- Wiltshire S, Morris AP, Zeggini E 2008. Examining the statistical properties of fine-scale mapping in large-scale association studies. *Genet Epidemiol* 32: 204-214.
- Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A, Farrington SM, Prendergast J, Olschwang S, Chiang T, Crowley E, Ferretti V, Laflamme P, Sundararajan S, Roumy S, Olivier JF, Robidoux F, Sladek R, Montpetit A, Campbell P, Bezieau S, O'Shea AM, Zogopoulos G, Cotterchio M, Newcomb P, McLaughlin J, Younghusband B, Green R, Green J, Porteous ME, Campbell H, Blanche H, Sahbatou M, Tubacher E, Bonaiti-Pellie C, Buecher B, Riboli E, Kury S, Chanock SJ, Potter J, Thomas G, Gallinger S, Hudson TJ, Dunlop MG 2007. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 39: 989-994.
- Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS, McCarthy MI, Hattersley AT 2007. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316: 1336-1341.