npg

ARTICLE

Genome-wide genetic investigation of serological measures of common infections

Rohina Rubicz*,[1], Robert Yolken[2], Eugene Drigalenko[3], Melanie A Carless[3], Thomas D Dyer[3], Jack Kent Jr[3], Joanne E Curran[3], Matthew P Johnson[3], Shelley A Cole[3], Sharon P Fowler[4], Rector Arya[5], Sobha Puppala[3], Laura Almasy[3], Eric K Moses[6], Ellen Kraig[7], Ravindranath Duggirala[3], John Blangero[3], Charles T Leach[8] and Harald HH Göring[3]

Populations and individuals differ in susceptibility to infections because of a number of factors, including host genetic variation. We previously demonstrated that differences in antibody titer, which reflect infection history, are significantly heritable. Here we attempt to identify the genetic factors influencing variation in these serological phenotypes. Blood samples from > 1300 Mexican Americans were quantified for IgG antibody level against 12 common infections, selected on the basis of their reported role in cardiovascular disease risk: *Chlamydia pneumoniae*; *Helicobacter pylori*; *Toxoplasma gondii*; cytomegalovirus; herpes simplex I virus; herpes simplex II virus; human herpesvirus 6 (HHV6); human herpesvirus 8 (HHV8); varicella zoster virus; hepatitis A virus (HAV); influenza A virus; and influenza B virus. Pathogen-specific quantitative antibody levels were analyzed, as were three measures of pathogen burden. Genome-wide linkage and joint linkage and association analyses were performed using ~ 1 million SNPs. Significant linkage (lod scores > 3.0) was obtained for HHV6 (on chromosome 7), HHV8 (on chromosome 6), and HAV (on chromosome 13). SNP rs4812712 on chromosome 20 was significantly associated with *C. pneumoniae* ($P = 5.3 \times 10^{-8}$). However, no genome-wide significant loci were obtained for the other investigated antibodies. We conclude that it is possible to localize host genetic factors influencing some of these antibody traits, but that further larger-scale investigations will be required to elucidate the genetic mechanisms contributing to variation in antibody levels.
European Journal of Human Genetics (2015) **23**, 1544–1548; doi:10.1038/ejhg.2015.24; published online 11 March 2015

INTRODUCTION

Individuals exhibit variation in their response, including potential disease outcome, upon exposure to infectious pathogens. This is likely in part because of underlying host genetic differences that influence the ability of pathogens to attach/invade and replicate on/inside the human host (or regulate behavior that alters exposure). There is evidence that host genetic factors influence antibody responses to both vaccines (eg varicella and rubella)[1,2] and to naturally occurring infections such as malaria, Chagas, and Epstein–Barr virus.[3–5] We previously demonstrated that differences in antibody titer, which reflect infection history, are significantly heritable for a number of common infectious pathogens.[6] Here we attempt to localize the genetic factors influencing serological phenotypes of these pathogens using genome-wide statistical gene-mapping approaches. We had selected 12 pathogens for examination because of reports linking them to risk for cardiovascular disease. These include the following: two bacterial pathogens: *Chlamydia pneumoniae* and *Helicobacter pylori*; one protozoan: *Toxoplasma gondii*; six herpes viruses: cytomegalovirus (CMV), herpes simplex type 1 virus (HSV-1), herpes simplex type 2 virus (HSV-2), human herpesvirus 6 (HHV-6), human herpesvirus 8 (HHV-8), and varicella zoster virus (VZV); and three

other viruses: hepatitis A virus (HAV), influenza A virus, and influenza B virus. We also examined three measures of pathogen burden (ie, the number of pathogens to which an individual has been exposed) – all herpes viruses, all viruses, and all pathogens – in order to determine whether there were underlying genetic variants shared among these traits. To our knowledge, published genome-wide investigations are not currently available for most of the infectious antibody traits examined here, with the exception of anti-CMV antibodies.[7]

MATERIALS AND METHODS

Sample collection

Participants consisted of 1932 members of extended Mexican–American families in the San Antonio Family Study (SAFS), including 771 men and 1161 women, most of whom were initially recruited during the years 1991–1995. On average, each phenotyped individual had 40 other phenotyped family members in this study, with up to six generations represented and the largest pedigree consisting of 133 phenotyped members. Participant ages ranged from 16 to 94 years, and the average age was 41 years. The SAFS combines the San Antonio Family Heart Study (SAFHS)[8] and the San Antonio Family Diabetes Gallbladder Study (SAFDGS).[9,10] The combined study is now focused on the identification of genetic risk factors for many common diseases and their quantitative correlates, beyond the initial respective foci on cardiovascular

[1]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [2]Stanley Division of Developmental Neurovirology, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [3]Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA; [4]Division of Clinical Epidemiology, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; [5]Division of Endocrinology and Diabetes, Department of Pediatrics, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; [6]Center for Genetic Epidemiology and Biostatistics, The University of Western Australia, Perth, WA, Australia; [7]Department of Cellular and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; [8]Department of Pediatrics, Section of Infectious Diseases, Baylor College of Medicine/Children's Hospital of San Antonio, San Antonio, TX, USA
*Correspondence: Dr R Rubicz, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., Mailstop: M4-A402, P.O Box 19042, Seattle, 98109 WA, USA. Tel: 206 667 6904; Fax: 206 667 2717; E-mail: rrubicz@fredhutch.org
Received 30 April 2014; revised 9 December 2014; accepted 27 January 2015; published online 11 March 2015

disease and diabetes. Individuals in the SAFHS were randomly ascertained (without regard for disease status), whereas SAFDGS participants were initially recruited on the basis of a single diabetic proband in each pedigree. Although the SAFDGS is weakly enriched for diabetes, the rate of diabetes (and other major diseases including heart disease and obesity) is similar in the two component studies, indicating that the overall study may be viewed (and analyzed) as a randomly ascertained sample. The study and protocols were approved by the Institutional Review Board of the University of Texas Health Science Center at San Antonio, and all study participants signed informed consent statements.

Serology
Fasting blood samples were collected in EDTA vacutainers from study participants during the initial recruitment (1991–1995). Plasma was aliquoted and frozen as previously described,[11] and stored at −80 °C. Samples were thawed just before their use in determining IgG antibody levels, using commercially available ELISA assays for the following infectious agents: *C. pneumoniae* (Bioclone Australia, Marrickville, NSW, Australia); *H. pylori* and CMV (Inverness Medical Professional Diagnostics, Palatine, IL, USA); *T. gondii,* VZV, and influenza A and B (IBL America, Minneapolis, MN, USA); HSV-1 and HSV-2 (Focus Diagnostics Inc, Philadelphia, PA, USA); HAV (Bio-Rad Laboratories, Redmond, WA, USA); and HHV-6 (Advanced Biotechnologies, Rockville, MD, USA). These antibody titer data, along with information on age and sex, are available in Supplementary Table 1.

Statistical genetic analysis
The quantitative IgG antibody level traits were analyzed for each pathogen. Statistical analyses were performed using a variance components (VC) method with the SOLAR software computer package (Texas Biomedical Research Institute, San Antonio, TX, USA).[12] Because of the sensitivity of VC analyses to extreme values (outliers or 'thick tails' generating high kurtosis), the quantitative antibody level traits (optical density values) were inverse-normalized (by rank) before analysis. In addition, the sum of seropositive reactions to the pathogens examined here was used as a measure of pathogen burden for (1) all herpes viruses, (2) all viruses, and (3) all pathogens. We have previously published gene-mapping results on antibodies directed against herpes virus Epstein–Barr virus (EBV),[13] and these results are not repeated in this paper. However, the EBV antibodies were included when calculating the pathogen burden trait. The influence of shared environmental factors was modeled using a 'household' variance component to account for shared residency at the time of the blood draw.[14]

Given that the SAFS contains extended families, several genome-wide analyses (including linkage and association) were performed using SOLAR[12] in order to take full advantage of the information available for this sample. To identify genomic regions containing variants that may influence the antibody traits, genome-wide linkage analysis was run based on multipoint identity-by-descent estimations that were calculated with LOKI[15,16] using 28,388 SNPs with very low linkage disequilibrium between them. In addition, genome-wide joint linkage and association analysis was performed using 944,565 SNPs. This was done by including a random-effects linkage component (a QTL effect) and a fixed-effect allelic component (an additive allele dosage effect). SNP genotypes were generated using several versions of Illumina's SNP genotyping BeadChip microarrays (HumanHap550v3, HumanExon510Sv1, Human1Mv1, and Human1M-Duov3 (Illumina, Inc., San Diego, CA, USA)), and before analysis underwent rigorous quality control measures as described previously.[13] A VC-based random-effects linkage model was used for the linkage component, and the association component was implemented as an additive measured genotype model. For our sample, we used the effective number of SNPs (calculated from linkage disequilibrium) to estimate that $P \leq 1.3 \times 10^{-7}$ is the genome-wide significance threshold corresponding to $\alpha = 0.05$. To account for population substructure, which may produce spurious association results,[17] we ran principal components analysis using R princomp (R development team) on the SNPs to identify differential ethnic contributions of ancestral populations to study participants. This was done in the founders of pedigrees only, and the offspring were then assigned the average values of their parents for each principal component in order to avoid the risk that principal components may

tag potentially important genetic differences between pedigrees *per se.* The top five principal components were included in all statistical analyses as additional covariates. Variant locations were based on the hg19 reference sequence.

RESULTS

Antibody levels and heritability estimates
We measured IgG antibody levels against 12 infectious pathogens in Mexican–American families from San Antonio, TX, USA. These pathogens have previously been suggested to have a role in cardio-vascular disease risk. The seroprevalence rates for these pathogens ranged from 10% for *T. gondii* to 92% for VZV (Table 1), and are described in greater detail in the study by Rubicz et al.[18] Heritability estimates were significant for all pathogens, ranging from 0.08 for HSV-2 to 0.37 for HHV-8 (Table 1 and Rubicz et al.).[6] Three measures of pathogen burden (ie, the number of seropositive reactions to the investigated pathogens) were generated – namely, for all herpes viruses, all viruses, and all pathogens. These three burden traits were also significantly heritable, with heritability estimates between 0.16 and 0.24.

Genome-wide linkage results
Genome-wide linkage analysis, using identity-by-descent allele sharing probabilities computed from >28 000 SNPs in low linkage disequilibrium, was conducted in order to locate any major loci that might be influencing the infectious pathogen IgG antibody traits. The genome-wide linkage plots are provided in Supplementary Figure 1. Significant LOD scores (LOD score ≥ 3.0) were obtained for HHV-6 on chromosome 7q36.3 (LOD 3.62), for HHV-8 at 6q14.1 (3.80) located outside the human leukocyte antigen region, and for HAV at 13q34 (3.03) (Table 2), but not for the other pathogens investigated. Additional suggestive LOD scores (≥2.0) were obtained for eight pathogens and for all three measures of pathogen burden.

Genome-wide joint linkage and association results
We next performed genome-wide joint linkage and association analysis using nearly 1 million SNPs in order to potentially increase our power (in particular for common variants) and refine our search for responsible loci (Manhattan-style genome-wide association plots

Table 1 Estimated rates of seroprevalence and heritability for pathogens examined in this study

Pathogen	% Seropositive	Heritability (P-value)
C. pneumoniae[a]	76.2	$0.32 \ (2.7 \times 10^{-18})$
H. pylori[a]	54.9	$0.36 \ (2.8 \times 10^{-24})$
T. gondii[a]	10.0	$0.23 \ (1.3 \times 10^{-10})$
CMV[a]	58.8	$0.35 \ (1.3 \times 10^{-20})$
HSV-1[a]	82.0	$0.22 \ (3.9 \times 10^{-10})$
HSV-2[a]	23.8	$0.08 \ (0.01)$
HHV-6[a]	76.8	$0.28 \ (2.4 \times 10^{-14})$
HHV-8[b]	28.2	$0.37 \ (8.1 \times 10^{-16})$
VZV[b]	92.1	$0.25 \ (2.5 \times 10^{-9})$
HAV[b]	75.6	$0.34 \ (5.7 \times 10^{-17})$
Influenza A[b]	76.3	$0.27 \ (4.5 \times 10^{-9})$
Influenza B[b]	55.0	$0.23 \ (3.0 \times 10^{-7})$
All herpes viruses[b,c]	—	$0.18 \ (6.6 \times 10^{-5})$
All viruses[b,c]	—	$0.16 \ (6.1 \times 10^{-4})$
All pathogens[b,c]	—	$0.24 \ (1.3 \times 10^{-6})$

[a]Phenotype data available for 1932 participants.
[b]Phenotype data available for 1345 participants.
[c]Including IgG level to EBV, as previously described in Rubicz et al.[13] (seroprevalence rate = 58.2%, heritability = 0.39 ($P = 7.0 \times 10^{-19}$), for 1932 participants).

are given in Supplementary Figure 2). Note that the joint linkage and association test involves two degrees-of-freedom (the linkage variance component and the measured genotype covariate), and hence, although these analyses use more available information, it is not guaranteed that joint linkage and association analysis actually increases power. For our sample, consisting of large multigenerational families, we estimated $P \leq 1.3 \times 10^{-7}$ to be the genome-wide significance threshold (at $\alpha = 0.05$) (see Methods). Genome-wide significant results ($P = 5.34 \times 10^{-8}$) were obtained for *C. pneumoniae* for SNP rs4812712 (Chr.20.hg19:g.42104939 A > C) on chromosome 20 (Table 3). A quantile–quantile plot of observed *versus* expected P-values for genome-wide association results for *C. pneumoniae* is provided in Supplementary Figure 3. The genomic inflation factor[19] for this trait ($\lambda = 1.016$) does not deviate greatly from 1, indicating no obvious inflation of significance levels. No other pathogen yielded genome-wide significant results. There are suggestive results (at $P \leq 1.3 \times 10^{-6}$) for *C. pneumoniae* on chromosome 11, for CMV on chromosome 14, for HHV-8 on chromosome 6, for influenza A on chromosomes 15 and 19, and for all herpes viruses on chromosome 11.

DISCUSSION

The heterogeneity observed among individuals and populations in infection status to various pathogens and subsequent disease progression may be influenced by social and environmental factors including population density, hygiene, nutritional status, and stress, as well as the genetic composition of both the pathogen and the host. Here we conducted a genome-wide investigation into host genetic variation influencing IgG antibody response to 12 common infectious pathogens in extended Mexican-American families. Most of the pathogens examined in this study are common, and many of them are thought to be transmitted relatively easily, primarily through person-to-person contact or through respiratory secretions. Individuals are likely to be exposed to many of these pathogens multiple times throughout their lifetime. Exceptions to these are HSV-2, which is sexually transmitted, and *T. gondii*, which is contracted mainly through eating undercooked meat or contact with cat feces. This is likely part of the explanation for the fairly low seroprevalences of these two pathogens (24% and 10%, respectively), which are the lowest among the examined infectious agents. If a given individual indeed is likely to have been exposed to most of these microbes, then the absence of antibodies against these pathogens should also be informative about host genetics (ie, an individual was likely exposed, but either did not produce antibodies or produced only low levels of antibodies). We therefore included both seropositive and seronegative individuals in this investigation. Our contention that seronegative individuals do provide host genetic information is supported by our previous demonstration that these antibody level traits are substantially heritable when one includes seronegative individuals in the analysis, with the smallest heritability estimate obtained for the sexually transmitted HSV-2 (in which case seronegative individuals may well have not been exposed at all and thus would not provide any information on host genetics).[6] It should be noted that antibody level does not necessarily correlate with

Table 2 Significant (LOD $\geq$ 3.0) and suggestive (LOD $\geq$ 2.0) multipoint linkage peaks for IgG antibody level traits

Pathogen	Chromosome	cM location	LOD score
C. pneumoniae	11	22	2.07
	11	73	2.77
T. gondii	5	187	2.02
	8	67	2.06
	10	52	2.30
CMV	14	13	2.25
	17	18	2.63
HSV1	15	102	2.09
	16	44	2.30
	19	59	2.23
HHV6	**7**	**188**	**3.62**
	8	121	2.10
	17	1	2.32
	19	109	2.28
HHV8	**6**	**92**	**3.80**
	20	15	2.2
VZV	12	88	2.62
	19	55	2.07
HAV	11	24	2.37
	13	**123**	**3.03**
Influenza A	1	77	2.26
	10	69	2.03
	15	34	2.49
	21	27	2.64
All herpes viruses	11	7	2.15
All viruses	8	139	2.99
All pathogens	8	150	2.52
	11	7	2.36

Bold = significant at LOD $\geq$ 3.0.

Table 3 Significant and suggestive results of genome-wide joint linkage and association analysis for IgG antibody measurements against infectious pathogens

Trait	SNP	SNP location[a]	Nearest gene	Minor allele	Minor allele frequency[b]	P-value (β_{SNP})[c]
C. pneumoniae	rs7122209	Chr.11.hg19:g.71766174 T > C	*NUMA1*	C	0.003	7.7×10^{-7} (−1.98)
	rs4812712	**Chr.20.hg19:g.42104939 A > C**	***SRSF6/L3MBTL1***	**A**	**0.403**	**5.3×10^{-8} (0.18)**
Human herpesvirus 8	rs2123993	Chr.6.hg19:g.81428336 T > C	*FAM46A*	C	0.398	5.7×10^{-7} (−0.16)
	rs985737	Chr.6.hg19:g.81428748 T > G	*FAM46A*	G	0.330	7.5×10^{-7} (0.16)
	rs9352866	Chr.6.hg19:g.81481783C > T	*FAM46A*	T	0.397	1.5×10^{-7} (−0.17)
Influenza A	rs16957709	Chr.15.hg19:g.43702948 T > C	*TP53BP1*	C	0.006	5.4×10^{-7} (1.04)
	rs6508941	Chr.19.hg19:g.41015965C > T	*SPTBN4*	T	0.217	7.9×10^{-7} (0.26)
Cytomegalovirus	rs12889813	Chr.14.hg19:g.24306537C > T	*LOC728667*	T	0.152	8.2×10^{-7} (−0.20)
All herpes viruses	rs4638315	Chr.11.hg19:g.69037823G > A	*MYEOV*	A	0.290	5.6×10^{-7} (0.22)

Bold = genome-wide significant at $P \leq 1.3 \times 10^{-7}$; significance threshold was estimated on the basis of our sample of Mexican-American participants. The threshold for suggestive evidence for association used here is $P \leq 1.3 \times 10^{-6}$.
[a]Based on hg19.
[b]Minor allele frequency was estimated for our study sample using maximum likelihood methods and taking family relationships into account.
[c]The direction of effect of the regression coefficient is relative to the minor SNP allele.

resistance against a particular pathogen. Furthermore, antibody levels are likely also influenced by the timing (how long since last exposure) and/or frequency of pathogen exposure, dose of pathogen, and efficiency of the immune system in dealing with the pathogen. Such factors are expected to add noise to antibody levels, which – similar to the manifold complexities underlying most complex traits – would reduce power to detect genetic loci influencing these traits and require large sample sizes to achieve genome-wide significance. This is in agreement with our observation here that it is difficult, but not impossible, to identify genomic regions harboring genetic variants influencing these types of serological traits. The IgG antibody traits examined here are naturally occurring for the most part, as vaccines were not available against the majority of these pathogens at the time of blood sample collection (1991–1995) or, when available, were not in common use in the study population. For transcripts significantly correlated with anti-C. pneumoniae antibody level and other antibody level traits, please see Supplementary Appendix 1

In our genome-wide analyses, our top association finding in this study was for the C. pneumoniae antibody level trait with SNP rs4812712:A > C on chromosome 20. It should be noted that there is no evidence that IgG antibodies against C. pneumonia confer resistance to infection and/or disease. The nearest genes to rs4812712: A > C are SRSF6 and L3MBTL1. The SRSF6 gene encodes a protein that is involved in mRNA splicing and its overexpression is suggested to contribute to the development of lung and colon cancers.[20] C. pneumoniae seropositivity may be a risk factor for lung cancer, with higher anti-C. pneumoniae antibody titer correlated with cancer risk.[21] Although the mechanism by which infection with C. pneumoniae may increase the risk of lung cancer is not known, one possible explanation is that it may induce irregular apoptosis in tissues.[22] The human SRSF6 protein has also been linked to infection with HIV-1, reportedly regulating HIV-1 mRNA processing and possibly being involved in nuclear export of spliced mRNAs.[23,24] L3MBTL1, on the other hand, is a tumor suppressor gene that appears to suppress genes and microRNAs that re-activate during tumor growth and whose expression is associated with decreased risk of death due to breast cancer.[25]

We found suggestive evidence of association for anti-CMV antibody level on chromosome 14, with our top SNP located nearest to the pseudogene LOC728667. However, the nearby DHRS4 gene may be of interest in this trait as it is involved in retinol metabolism. Retinol is known to influence both immune function and ocular health, and infection with CMV has a role in the development of certain ocular diseases, including retinitis.[26] We compared our association results with those of Kuparinen et al.,[7] who also conducted genome-wide association analysis for the anti-CMV IgG antibody level. Although they did not find evidence for a major genetic locus, they did identify AGBL1 as potentially influencing this trait, based on suggestive evidence of association of five SNPs in this gene. The AGBL1 gene may influence infection with CMV by means of its potential role in processing tubulin, given that CMV capsids are reported to utilize the microtubule network in nuclear targeting of epithelial cells,[27] or alternatively by evading lysosomal fusion through microtubule network disaggregation during macrophage replication, thus allowing the virus to persist in these cells.[28] Our results did not, however, provide further supporting evidence, possibly in part owing to population-level differences in our sample (which consisted of Mexican Americans, versus Finnish individuals in their study).

We investigated whether it is possible to localize genetic factors influencing pathogen burden traits, including all herpes viruses, all viruses, and all pathogens. These traits were of interest for the study because the number of seropositive reactions against infectious pathogens has previously been identified as a risk factor for chronic disease including atherosclerosis.[29–31] Although no major loci were significantly associated with any of these pathogen burden measures, there was suggestive evidence for association and/or linkage on chromosomes 8 and 11, including a genetic variant in the oncogene MYEOV for all herpes viruses (various herpes viruses are shown to contribute to the development of cancers, eg, EBV to Burkitt's lymphoma and Hodgkin's lymphoma).[32,33]

In summary, this study identified a variant on chromosome 20 that is significantly associated with anti-C. pneumoniae IgG. However, no other antibody level traits obtained genome-wide significance levels of association. This may be in part due to the heterogeneity of factors that contribute to variation in antibody levels, including differences in exposure, its frequency, and its timing relative to blood draw, which would reduce power to detect underlying genetic loci influencing these traits. Further investigation may offer additional insight into the genetic factors and biological pathways involved in antibody titer levels, and possibly the relationship of antibody levels to resistance or susceptibility to infection and disease progression, potentially identifying novel methods of prevention and/or treatment. We hope that our data will be useful for larger-scale meta-analytical association analyses on many of these antibody traits.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

1 Klein NP, Fireman B, Enright A, Ray P, Black S, Dekker CL: A role for genetics in the immune response to the varicella vaccine. Pediatr Infect Dis J 2007; 26: 300–305.
2 Jacobson RM, Ovsyannikova IG, Poland GA: Genetic basis for variation of vaccine response: our studies with rubella vaccine. Paediatr Child Health 2009; 19: S156–S159.
3 Duah NO, Weiss HA, Jepson A, Tetteh KK, Whittle HC, Conway DJ: Heritability of antibody isotype and subclass responses to Plasmodium falciparum antigens. PLoS One 2009; 4: e7381.
4 Williams-Blangero S, Vandeberg JL, Blangero J, Teixeira AR: Genetic epidemiology of seropositivity for Trypanosoma cruzi infection in rural Goias, Brazil. Am J Trop Med Hyg 1997; 57: 538–543.
5 Besson C, Amiel C, Le-Pendeven C et al: Strong correlations of anti-viral capsid antigen antibody levels in first-degree relatives from families with Epstein-Barr virus-related lymphomas. J Infect Dis 2009; 199: 1121–1127.
6 Rubicz R, Leach CT, Kraig E et al: Genetic factors influence serological measures of common infections. Hum Hered 2011; 72: 133–141.
7 Kuparinen T, Seppala I, Jylhava J et al: Genome-wide association study does not reveal major genetic determinants for anti-cytomegalovirus antibody response. Genes Immun 2012; 13: 184–190.
8 Mitchell BD, Kammerer CM, Blangero J et al: Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. Circulation 1996; 94: 2159–2170.
9 Hunt KJ, Lehman DM, Arya R et al: Genome-wide linkage analyses of type 2 diabetes in Mexican Americans: the San Antonio Family Diabetes/Gallbladder Study. Diabetes 2005; 54: 2655–2662.

10 Puppala S, Dodd GD, Fowler S et al: A genomewide search finds major susceptibility loci for gallbladder disease on chromosome 1 in Mexican Americans. Am J Hum Genet 2006; 78: 377–392.

11 Cheng ML, Woodford SC, Hilburn JL, VandeBerg JL: A novel system for storage of sera frozen in small aliquots. J Biochem Biophys Methods 1986; 13: 47–51.

12 Almasy L, Blangero J: Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 1998; 62: 1198–1211.

13 Rubicz R, Yolken R, Drigalenko E et al: A genome-wide integrative genomic study localizes genetic factors influencing antibodies against Epstein-Barr virus nuclear antigen 1 (EBNA-1). PLoS Genet 2013; 9: e1003147.

14 Spence MA, Westlake J, Lange K: Estimation of the variance components for dermal ridge count. Ann Hum Genet 1997; 41: 111–115.

15 Heath SC: Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. Am J Hum Genet 1997; 61: 748–760.

16 Heath SC, Snow GL, Thompson EA, Tseng C, Wijsman EM: MCMC segregation and linkage analysis. Genet Epidemiol 1997; 14: 1011–1016.

17 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006; 38: 904–909.

18 Rubicz R, Leach CT, Kraig E et al: Seroprevalence of 13 common pathogens in a rapidly growing U.S. minority population: Mexican Americans from San Antonio, TX. BMC Res Notes 2011; 4: 433.

19 Devlin B, Roeder K: Genomic control for association studies. Biometrics 1999; 55: 997–1004.

20 Cohen-Eliav M, Golan-Gerstl R, Siegfried Z et al: The splicing factor SRSF6 is amplified and is an oncoprotein in lung and colon cancers. J Pathol 2013; 229: 630–639.

21 Zhan P, Suo LJ, Qian Q et al: Chlamydia pneumoniae infection and lung cancer risk: a meta-analysis. Eur J Cancer 2011; 47: 742–747.

22 Fan T, Lu H, Hu H et al: Inhibition of apoptosis in chlamydia-infected cells: blockade of mitochondrial cytochrome c release and caspase activation. J Exp Med 1998; 187: 487–496.

23 Swanson CM, Sherer NM, Malim MH: SRp40 and SRp55 promote the translation of unspliced human immunodeficiency virus type 1 RNA. J Virol 2010; 84: 6748–6759.

24 Tranell A, Tingsborg S, Fenyo EM, Schwartz S: Inhibition of splicing by serine-arginine rich protein 55 (SRp55) causes the appearance of partially spliced HIV-1 mRNAs in the cytoplasm. Virus Res 2011; 157: 82–91.

25 Zeng H, Irwin ML, Lu L et al: Physical activity and breast cancer survival: an epigenetic link through reduced methylation of a tumor suppressor gene L3MBTL1. Breast Cancer Res Treat 2012; 133: 127–135.

26 Newman H, Gooding C: Viral ocular manifestations: a broad overview. Rev Med Virol 2013; 23: 281–294.

27 Ogawa-Goto K, Tanaka K, Gibson W et al: Microtubule network facilitates nuclear targeting of human cytomegalovirus capsid. J Virol 2003; 77: 8541–8547.

28 Fish KN, Britt W, Nelson JA: A novel mechanism for persistence of human cytomegalovirus in macrophages. J Virol 1996; 70: 1855–1862.

29 Espinola-Klein C, Rupprecht HJ, Blankenberg S et al: Impact of infectious burden on extent and long-term prognosis of atherosclerosis. Circulation 2002; 105: 15–21.

30 Zhu J, Quyyumi AA, Norman JE et al: Effects of total pathogen burden on coronary artery disease risk and C-reactive protein levels. Am J Cardiol 2000; 85: 140–146.

31 Epstein SE, Zhu J, Burnett MS, Zhou YF, Vercellotti G, Hajjar D: Infection and atherosclerosis: potential roles of pathogen burden and molecular mimicry. Arterioscler Thromb Vasc Biol 2000; 20: 1417–1420.

32 de-The G, Geser A, Day NE et al: Epidemiological evidence for causal relationship between Epstein-Barr virus and Burkitt's lymphoma from Ugandan prospective study. Nature 1978; 274: 756–761.

33 Henle W, Henle G, Ho HC et al: Antibodies to Epstein-Barr virus in nasopharyngeal carcinoma, other head and neck neoplasms, and control groups. J Natl Cancer Inst 1970; 44: 225–231.