Published in final edited form as:
Hum Genet. 2013 January ; 132(1): . doi:10.1007/s00439-012-1223-x.

Investigation of genetic risk factors for chronic adult diseases for association with preterm birth

Nadia Falah[1,*], **Jude McElroy**[1,*], **Victoria Snegovskikh**[2], **Charles J. Lockwood**[2], **Errol Norwitz**[2], **Jeffey C. Murray**[3], **Edward Kuczynski**[4], **Ramkumar Menon**[5], **Kari Teramo**[6], **Louis J. Muglia**[1], and **Thomas Morgan**[1]

Nadia Falah: nadia.a.falah@vanderbilt.edu; Jude McElroy: jude.j.mcelroy@Vanderbilt.edu; Victoria Snegovskikh: Victoria.snegovskikh@yale.edu; Charles J. Lockwood: charles.lockwood@yale.edu; Errol Norwitz: errol.norwitz@yale.edu; Jeffey C. Murray: Jeff-murray@uiowa.edu; Edward Kuczynski: ekuczyns@central.UH.ED; Ramkumar Menon: ram.menon@utmb.edu; Kari Teramo: kari.teramo@hus.fi; Louis J. Muglia: louis.muglia@Vanderbilt.Edu; Thomas Morgan: thomas.morgan@vanderbilt.edu

[1]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN [2]Department of Obstetrics and Gynecology, Yale School of Medicine, New Haven, CT [3]Department of Pediatrics, University of Iowa [4]Biology and Biochemistry Department, University of Houston, TX [5]Department of Obstetrics and Gynecology, University of Texas, Galveston, TX [6]Department of Obstetrics and Gynecology, University of Helsinki, Helsinki, Finland

Abstract

Preterm birth (PTB) is the leading cause of infant mortality. PTB pathophysiology overlaps with those of adult cardiovascular, immune and metabolic disorders (CIMD), with mechanisms including inflammation, immunotolerance, thrombosis, and nutrient metabolism. Whereas many genetic factors for CIMD have been identified, progress in PTB has lagged. We hypothesized that highly validated genetic risk factors for CIMD may also be associated with PTB. We conducted case–control study of four female cohorts with spontaneous PTB (n = 673) versus term (n = 1119). Of 35 SNPs genotyped, there were 13 statistically significant associations (P<0.05), which were more than expected (binomial test; P = 0.02). In US White (307 cases/342 controls), the G allele of HLA-DQA1 (A/G) rs9272346 was protective for PTBin the initial discovery cohort (P = 0.02; OR = 0.65; 95 % CI 0.46, 0.94). This protective association replicated (P = 0.02; OR = 0.85; 95 % CI 0.75, 0.97) nominally in the Danish Cohort (883 cases, 959 controls), but lost significance upon multiple testing correction. We observed more statistically significant associations than expected, suggesting that chance is an unlikely explanation for one or more of the associations. Particularly, a protective association of the G allele of HLA-DQA1 was found in two independent cohorts, and in previous studies, this same allele was found to protect against type-1-diabetes (meta-analysis P value 5.52×10^{-219}). Previous investigations have implicated HLA phenotypic variation in recurrent fetal loss and in chronic chorioamnionitis. Given the limited sample size in his study, we suggest larger studies to further investigate possible HLA genetic involvement in PTB.

Background

Preterm birth (PTB), defined as live birth occurring between 20–37 weeks of gestation, complicates 12.2 % of pregnancies (Kochanek et al. 2012), contributes to over one-third of infant deaths in the United States annually (Mathews and MacDorman 2011), and is

Corresponding author: Thomas M Morgan, Thomas.morgan@vanderbilt.edu.
*Contributed equally

Conflict of interest The authors declare that they have no competing interests.

associated with serious medical complications including neurodevelopmental delay, retinopathy of prematurity, and chronic lung disease (Landry et al. 2011; Sharma et al. 2011). PTB is etiologically heterogeneous (Menon et al. 2011). Genetic influences on birth timing in humans are substantial (Chaudhari et al. 2008). However, association studies using candidates selected from suspected pathways have not detected robust susceptibility variants for PTB.

Genome-wide association studies (GWAS) are promising but will require a very large number of well-characterized subjects in order to overcome the challenge of multiple statistical comparisons. The pathophysiology of preterm birth is poorly understood, but the processes of inflammation, placental thrombosis, and nutrient metabolism have all been hypothesized to have a role (Park et al. 2005; Lusis 2000). Such processes may overlap substantially with the pathophysiology of adult cardiovascular, immune, and metabolic disorders (CIMD). Recent advances in GWAS have led to the identification of multiple well-validated risk factors for CIMD, while progress has lagged for PTB.

The existence of genetic interconnections between PTB and adult cardiovascular, immune, and metabolic disorders (CIMD) is strongly suggested by similarities in known pathophysiological pathways among the various disorders. Inflammation, hormonal stress-response signaling, nutrition/metabolism, connective tissue matrix elaboration, and thrombosis are all fundamental processes that underlie most CIMD and are critical in fetal growth and the timing of parturition (Park et al. 2005; Lusis 2000).

Given the similarity in fundamental pathophysiological processes between CIMD and PTB, there is an opportunity to test for pleiotropic effects of known genetic risk factors for CIMD that may also affect birth timing. We therefore hypothesized that highly validated genetic associations with CIMD may also be associated with PTB. The selection of single nucleotide polymorphisms (SNPs) with extraordinarily low P values (1×10^{-30}) for association with CIMD is proposed as an efficient strategy to reduce multiple comparisons testing while maximizing the prior probability that variants tested are true pathological markers. Given the practical limitations of assembling very large sample sizes in complex phenotypes such as PTB, there is a need for creative research strategies that seek to resolve the tension between the hypothesis-directed candidate gene approach and the relatively unbiased GWAS approach that requires sample sizes massive enough to accommodate a million or more genetic association tests. Accordingly, we tested 35 independent SNPs (Table 1) in 33 genes that were found to be strongly associated with CIMD.

Methods

Recruitment and Sample Collection

Study subjects were enrolled with approval by Institutional Review Boards at each participating institution. Informed consent was obtained from all participants. Mothers with preterm birth were included if the birth was spontaneous (non-iatrogenic), singleton, had no obvious precipitating stimulus (trauma, infection, drug use), and was fewer than 37 weeks (Yale University; New York University; University of Iowa) or 36 weeks (Centennial Hospital, Nashville, TN; Washington University) of completed gestation. Control mothers were included if they had delivered two or more children at 37 weeks or later spontaneously. Race/ethnicity was assigned by self-report. All specimens were linked with demographic (Plunkett et al. 2011) and medical data abstracted from maternal/neonatal records. DNA from blood or saliva was prepared by standard methods (DNA extraction from saliva samples collected with the Oragene® DNA self collection kit 2008).

Discovery cohorts

Unless otherwise specified below, births were known spontaneous (non-iatrogenic) singleton births, with no obvious precipitating stimulus (e.g., trauma, infection, drug use). Gestational age was determined by last menstrual period and corroborated by ultrasound dating. A total of 673 PTB case mothers (307 US White, 133 US African- American, 160 Finnish, 73 US Hispanic) and 1,119 term birth mother controls (342 US White, 285 US African- American, 200 Finnish, 292 US Hispanic) were recruited at the following sites.

Barnes Jewish Hospital in St. Louis, Missouri (Plunkett et al. 2011)—The participants in this study were collected between January 2004 and May 2007. White (N = 26) and African-American (N = 46) case participants were selected among mothers giving birth to infants < 36 weeks of gestation, whereas control participants were mothers having uncomplicated term births (57 African-American, 26 White).

New York University, NY and Yale-New Haven Hospital (Snegovskikh et al. 2011)—Consecutive Hispanic patients with preterm birth < 37 weeks were identified from the March of Dimes Perinatal Emphasis Research Initiative project (MOD #20-FY03-30) from Jan 1989 to June 2005. Cases (N = 73) were matched (4:1) with uncomplicated term deliveries (controls, N = 290).

University of Helsinki/University of Oulu, Finland (Plunkett et al. 2011)—166 White cases < 36 weeks and 199 term controls, all of Finnish origin. Non-elective PTB with either intact or premature rupture of fetal membranes was included among the SPTBs. Patients with risk factors for PTB were excluded (i.e., multiple gestation, polyhydramnios, septic infection or chronic disease of the mother, narcotic or alcohol abuse, accidents, and fetuses with congenital anomalies).

University of Iowa Children's Hospital—118 White mothers with PTB < 37 weeks; 78 White controls. The mothers' baby was admitted to the Neonatal Intensive Care Unit of University of Iowa Children's Hospital. The indication for the PTB was not available for most cases as the study uses a neonatal registry lacking this information (Ehn et al. 2007).

Nashville Perinatal Research Center—Cases (84 African-American/157 White) were mothers with PTB < 37 weeks, term controls (222 African-American/ 196 White). White and African-American subjects between the ages of 18 and 40 years with singleton live births were included. Cases (PTB) were defined as presence of regular uterine contractions (2 contractions/10 min) with documented cervical changes followed by delivery prior to 36 weeks gestation. Subjects with multiple gestations, pre-eclampsia, preterm premature rupture of membranes, placental previa, fetal anomalies, cerclage, gestational diabetes, pre-gestational diabetes, poly- and oligohydramnios, intrauterine growth restriction, and other complications such as surgeries during pregnancy were excluded. Controls consisted of women having delivery at term (≥ 37 weeks) with no medical or obstetrical complications during pregnancy (Menon et al. 2010).

Replication cohort

Danish National Birth Cohort—The database of Genotypes and Phenotypes (dbGaP) contains publicly available data from a genome-wide case/ control study using approximately 1,000 preterm mother pairs from the Danish National Birth Cohort (DNBC) with spontaneous onset of labor or preterm premature rupture of membranes (PPROM), along with 1,000 control mothers to whom a child was born at ~40 weeks' gestation. All were singleton live births, without congenital anomalies, and with no maternal conditions known to be associated with preterm delivery or often requiring early delivery of the baby

(placenta previa, placental abruption, hydramnios, isoimmunization, placental insufficiency, pre-eclampsia/eclampsia). The child's parents and all four grandparents were born in Denmark (except in 24 cases with one or two grandparents from other Nordic countries) (Murray 2010).

Molecular methods

DNA was extracted from saliva or from buffy coats or venous blood stored at −80 °C. All specimens were linked with demographic and medical data abstracted from maternal/neonatal records (Plunkett et al. 2011).

SNP genotyping

30 microgrms of DNA diluted to 5 ng/lL was loaded into 96-well plates and assayed by means of matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry through the Sequenom Technology/Genotyping Cores at Vanderbilt and Yale Universities. SNPs for which successful Sequenom designs could not be achieved were genotyped individually by TaqMan®. Initial genotyping of the Finnish cohort was performed using the Affymetrix® Genome-Wide Human SNP Array 6.0. Genotypes were called from cell intensity data by the birdseed v2 algorithm (using default thresholds) (Schillert and Ziegler 2012), implemented in Affymetrix® Genotyping Console 3.0. DNBC was genotyped on the Illumina 660Quad. The SNAP tool from the Broad Institute (Johnson et al. 2008) will be used to select appropriate proxy SNPs for the SNPs that are not directly genotyped on the array.

Statistical analysis

Sequenom design was feasible and genotypes were generated for 26 of independent 35 SNPs (36 SNPs total, with one pair, rs6457617 and rs6457620, in complete linkage disequilibrium —note that 35 "independent" SNPs will hereafter be referenced throughout the current report). Using TaqMan®, genotypes were generated for the ten SNPs which were not genotyped by Sequenom (rs6457620, rs9272346, rs9268645, rs12191877, rs12191877, rs2476601, rs6457617, rs9263739, rs3131379, rs6056, rs599839), seven of which are in the 6p21 HLA region containing repetitive DNA sequences. In preterm birth mother cohorts, eight variants were not genotyped by TaqMan® in Hispanic mothers (repository at Yale University). Data analysis was performed with PLINK, using Fisher exact tests. We excluded individuals based on genotyping quality (>95 % call rate) and possible cryptic relatedness and SNPs based on the following criteria: not in Hardy–Weinberg equilibrium in controls ($P<0.001$), > 95 % genotype call rate, minor allele frequency(MAF) < 0.05, duplicate probes.

Power and sample size considerations

Effect sizes estimated by odds ratios (OR) for CIMD phenotypes of SNPs included in this study ranged from 1.3 to 5.5. We assumed that a similar range of odds ratios might possibly pertain to PTB. In our analysis, each genetic variant was pre-specified based on published GWAS reports, and the frequencies of risk-associated variants was compared in cases and controls by using Fisher exact single-marker allelic association tests in PLINK (Purcell et al. 2007). The number of positive associations at the $P < 0.05$ level was also compared to the expected number by chance. Power estimates ranged from 0.34 (OR = 1.5) to 0.82 (OR = 2) in US. White mothers (307 cases, 342 controls), under the assumption of allelic Fisher exact model, baseline disease risk frequency of 0.1, and minor allele frequency of 0.1, with two-tailed P value of ≤ 0.05 (Power and Sample Size Calculation, William D. Dupont 1998). The largest clinical population in the present study was US White mothers; inclusion of African-Americans, Finns, and Hispanics supplemented power and permitted investigation of the

robustness of statistical associations across multiple populations. We considered statistical significance of the nominal P values of individual genetic associations, and also applied the Bonferroni correction for multiple comparisons (Lesack and Naugler 2011).

Results

Genotyping was successful in 99.4 % of maternal DNA samples. Genotyping efficiency was above 98 % in all our SNPs except rs12191877 (97 %) and rs599839 (90 %).

Genotype distributions of four SNPs in six study subpopulations violated HWE in controls ($P < 0.001$), as noted in Table 2 below. Thirteen SNPs were found to be in statistically significant association (P\0.05) with PTB (Table 2). By chance alone, one would expect an average of 6.6 statistically significant associations, given 4 study cohorts in which 35 independent SNPs were tested for association, with a total of 132 independent tests performed in the 4 clinical populations. Thus, 13 associations are more than expected by chance ($P = 0.015$; binomial test).

Positive associations were particularly frequent in US White mothers (307 cases/342 controls), in which group 6 of 35 SNPs met the nominal level of statistical significance, trending toward a surplus of positive associations ($P = 0.07$; binomial test). These six significant SNPs were: IL23R rs11465804, LPL rs12678919, CDKN2 rs4977574, HLA-DRB1 rs6457620 (and its proxy SNP, MHC rs6457617, linkage disequilibrium r2 = 1.0), HLA-DQA1 rs9272346, and HLA-DRA rs9268645, none of which violated HWE in the controls ($P < 0.001$). The most statistically significant association observed was with the G allele of rs9268645 HLA-DRA in US Whites ($P = 0.003$). However, this association (along with all others) fell short of the study-wide statistical significance threshold ($P = 0.0004$) after Bonferroni correction for 132 tests.

In order to augment our study power, we brought forward all of the 13 positive associations for potential validation in the Danish National Birth Cohort mothers (DNBC) (883 cases, 959 controls). Two of the 13 SNPs were not directly genotyped on DNBC's Illumina 660 Quad array, and we therefore selected proxy SNPs. Specifically, the minor allele A of HLA-DRA rs9268615 correlates with the minor allele G of HLA-DRA rs9268645 in the CEU population (r2 = 0.97). In addition, the minor allele A of the HLA-DQB1 rs1063355 correlates with the minor allele G of rs9272346 in the HLA-DQA1 contiguous MHC/HLA 6p21 locus in the CEU population ($r^2 = 0.97$). For 1 SNP that was not genotyped in the DNBC (PCKS9 rs11591147), no suitable proxy was available, leaving only 12 SNPs tested for association with PTB in the DNBC validation sample. Of these, only one association was validated in the DNBC (two-tailed $P < 0.05$). Specifically, the A allele of the rs1063355 in HLA-DQB1 had a protective association with preterm birth in the Danish subpopulation (OR = 0.85; 95 % CI 0.75–0.97; two-tailed $P = 0.016$). When testing a pre-specified genetic model for validation, a one-tailed test is appropriate, and 0.016/2 = 0.008. Given a Bonferroni correction for 12 independent tests, the corrected one-tailed P value would be 0.096. Thus, there was a statistically non-significant trend toward validation of the association between PTB in US Whites and rs9272346 in HLA-DQA1 (by proxy association with rs1063355 in HLA-DQB1 in the DNBC). Tables 3 and 4 show lower minor allele frequencies in cases compared with controls in both US White and DNBC populations (i.e., the minor allele is termed "protective" in that it appears in higher frequency among unaffected individuals). Specifically, in US Whites, the rs9272346 G allele frequency was 0.35 (Table 4). In the DNBC, the proxy SNP rs1063355 A allele frequency was also 0.41 (Table 5). By contrast, the minor allele frequencies of both SNPs were 0.45. For comparison, the CEU population (white mixed European descent population in Utah) minor

allele frequencies, of rs9272346 and rs1063355, are 0.41 and 0.43, respectively, which are similar to our findings.

Discussion

We observed more statistically significant associations between CIMD-related genetic risk variants and PTB than expected, suggesting that chance is an unlikely explanation for one or more of the associations. This was the main finding of our study, and it provides statistical support for our hypothesis that highly validated risk factors for CIMD may also alter the risk of PTB. However, we were unable to identify any particular association as being robust across all study samples. One association with PTB drew our attention more than the others, mainly because it was replicated in the Danish National Birth Cohort, and this was the protective association of SNPs in the HLA DQA1-DQB1 region of chromosome 6p21. Although this association was only nominally statistically significant in US Whites and in the DNBC, it is perhaps premature to dismiss it as being likely due to chance. Several other canonical causality criteria must be considered in addition to the overall Bonferroni-corrected statistical significance level. First, there is the apparent consistency of the association in the discovery cohort of White mothers with that observed in the replication cohort of Danish mothers, with similar odds ratio and overlapping confidence limits. Such consistency is tempered by the fact that the association was not present in three other study cohorts (African-Americans, Hispanics, Finns). Possible explanations for why a valid association may not be observed in all populations studied include: (1) inadequate power in relatively small cohorts; (2) population-specific genetic and/or clinical heterogeneity; (3) Winner's curse (in which the effect size is skewed upwards in the discovery cohort) (Kraft 2008). Another criterion to be considered is a particular SNPs known status as a marker of other diseases besides PTB. Pleiotropy is the classical genetic principle that a particular DNA sequence variant may exert an influence on multiple traits. Along those lines, the present study was largely motivated by observations that pleiotropy is emerging as a dominant theme in genome-wide association studies (Huang et al. 2011). The HLA DQA1-DQB1 SNPs included in this study are known disease markers. The G allele of HLADQA1 is an extraordinarily robust protective marker in type I diabetes (meta-analysis P value 5.52×10^{-219}) (Table 1). Thus, this SNP (perhaps in concert with many others in the MHC/HLA 6p21 locus) signifies important variation in function that has documented pathological importance. This raises the "pre-test probability," or suspicion that the SNP may be involved in other disease processes in addition to type I diabetes mellitus.

Another criterion to be considered is biological plausibility. HLA-DQA1 belongs to the HLA class II alpha chain paralogs. Class II molecules are expressed in antigen-presenting cells (e.g., B lymphocytes, dendritic cells, macrophages). The class II molecule is a heterodimer consisting of an alpha (DQA) and a beta chain (DQB), both anchored in the membrane. It plays a central role in the immune system by presenting peptides derived from extracellular proteins. One possible mechanism that immunological factors could lead to preterm birth has a correlation with those mechanisms related to transplant rejection. A recent study showed that maternal anti-HLA class I seropositivity in spontaneous preterm births is higher than that of normal term (OR = 5.90, 95 % CI 1.60–21.83) (Lee et al. 2011). Furthermore, previous investigations have implicated HLA phenotypic variation in recurrent fetal loss as well as in chronic chorioamnionitis, so there is at least a superficial biological plausibility for the association observed in the present study. However, no clear mechanism has been proven. The SNPs (rs1063355 in HLA-DQB1 and SNP rs9272346 in HLA-DQA1) are both noncoding SNPs. SNP rs9272346 is in the 50 region upstream from HLA-DQA1, and rs1063355 lies in the 30 untranslated region of HLADQB1 (Li et al. 2010). Thus, the effects of these particular SNPs on gene function in the context of broader haplotypes may be difficult to discern. Knowing only the status of one SNP (or its proxy) in each patient, we

were unable to reconstruct complex HLA haplotypes that were classically used in the study of genetic variation in human immunity.

The present study is not without important caveats. The sample sizes in each clinical group studied were modest. Given the special challenges of recruiting, phenotyping, and collecting DNA samples from pregnant women who labor prematurely, our study is similar in size to others in the field of obstetrical genetic research. Our study design was deliberately conceived to choose only a few dozen genetic markers, in order to maximize power with our existing clinical resources. We suggest larger, even more narrowly tailored, future studies to further investigate possible HLA genetic involvement in PTB.

The small sample size limited our ability to divide our cases and controls based on their birth weight being small or appropriate for gestational age. While initial findings regarding the Barker hypothesis, pertained primarily to birth weight, both low birth weight due to shortened gestation or intrauterine growth restriction exert similar increases in risk for adult disease (de Boo and Harding 2006).

We also acknowledge that the role of potential confounding factors was not necessarily fully controlled in the present study. The most prominent potential confounder in genetic studies is population stratification (i.e., differential ethnic admixture in cases and controls). We dealt with this potential threat to the validity of associations by performing separate analyses within each ethnic/racial subgroup. An alternative (or supplemental procedure) is to genotype large numbers of ethnically informative genetic markers throughout the genome for use as covariates to detect and attempt to correct for population stratification. The present study's size and scope were too limited to apply genomic control methods to it, but future studies should be designed to address potential confounding more fully (He et al. 2011). In addition, future studies should take into account the vast haplotype complexity of the MHC/HLA region. The present study was not designed with this issue in mind, but we believe that our results indicate the need for more detailed analysis of HLA genetic variation in relation to PTB.

Our study was confined to mothers, because we did not have DNA from sufficient numbers of infants to fully analyze the paternal contribution to PTB risk. However, our previous studies have indicated that the paternal contribution to PTB is only slight in comparison with the maternal genome (Plunkett et al. 2009). Even so, it would be of interest to genotype infants for the purpose of investigating the possible effects of maternal–fetal risk allele transmission. Global efforts are currently underway to amass the requisite sample size and standardized phenotyping for genome-wide association studies of PTB (Uzun et al. 2012; Murray 2010). We anticipate that initial results will soon be made available, and that further studies will be forthcoming.

Conclusions

Despite its limitations, our initial exploratory investigation for the genetic connection between CIMD and PTB identified at least one promising candidate SNP involved in immune system biology, and resulted in an apparent surplus of positive associations that supported our primary hypothesis. Genetic investigation of preterm birth is imperative, not only for developing predictive models that could focus attention on women most at risk, and potentially intervening early with antenatal steroids to enhance fetal lung development, but also for gaining further insights into the pathobiology of parturition.

Acknowledgments

Turner-Hazinski grant award; March of Dimes; Cara Sutcliffe and Rachel Wiseman in the DNA Resources Core at Vanderbilt University Medical Center for their assistance with genotyping. Funding support for the GWAS of Prematurity and its Complications study was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01HG004423). The GWAS of Prematurity and its Complications study is one of the genome-wide association studies funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by the NIH GEI (U01HG004438) and the NIH contract “High throughput genotyping for studying the genetic contributions to human disease” (HHSN268200782096C). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap through dbGaP accession number phs000103.v1.p1.

References

Chaudhari BP, Plunkett J, Ratajczak CK, Shen TT, DeFranco EA, Muglia LJ. The genetics of birth timing: insights into a fundamental component of human development. Clin Genet. 2008; 74(6): 493–501.10.1111/j.1399-0004.2008.01124.x [PubMed: 19037974]

de Boo HA, Harding JE. The developmental origins of adult disease (Barker) hypothesis. Aust N Z J Obstet Gynaecol. 2006; 46(1):4–14.10.1111/j.1479828X.2006.00506.x [PubMed: 16441686]

[Accessed 2 Feb 2012] DNA extraction from saliva samples collected with the Ora- gene®·DNA self collection kit. 2008. http://www.dnagenotek.com/pdf_files/568446855.pdf

Ehn NL, Cooper ME, Orr K, Shi M, Johnson MK, Caprau D, Dagle J, Steffen K, Johnson K, Marazita ML, Merrill D, Murray JC. Evaluation of fetal and maternal genetic variation in the progesterone receptor gene for contributions to preterm birth. Pediatr Res. 2007; 62(5):630–635.10.1203/PDR.0b013e3181567bfc [PubMed: 17805208]

[Accessed 14 Dec 2011] Genome-Wide Association Studies of Prematurity and Its Complications. 2010. http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000103.v1.p1

He H, Zhang X, Ding L, Baye TM, Kurowski BG, Martin LJ. Effect of population stratification analysis on false-positive rates for common and rare variants. BMC Proc. 2011; 5(Suppl 9):S116.10.1186/1753-6561-5-S9-S116 [PubMed: 22373282]

Huang J, Johnson AD, O’Donnell CJ. PRIMe: a method for characterization and evaluation of pleiotropic regions from multiple genome-wide association studies. Bioinformatics. 2011; 27(9): 1201–1206.10.1093/bioinformatics/btr116 [PubMed: 21398673]

Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O’Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008; 24(24): 2938–2939.10.1093/bioinformatics/btn564 [PubMed: 18974171]

Kochanek KD, Kirmeyer SE, Martin JA, Strobino DM, Guyer B. Annual summary of vital statistics: 2009. Pediatrics. 2012; 129(2):338–348.10.1542/peds.2011-3435 [PubMed: 22291121]

Kraft P. Curses—winner’s and otherwise—in genetic epidemiology. Epidemiology. 2008; 19(5):649–651. (discussion 657–648). 10.1097/EDE.0b013e318181b865 [PubMed: 18703928]

Landry JS, Chan T, Lands L, Menzies D. Long-term impact of bronchopulmonary dysplasia on pulmonary function. Can Respir J J Can Thorac Soc. 2011; 18(5):265–270.

Lee J, Romero R, Xu Y, Kim JS, Topping V, Yoo W, Kusanovic JP, Chaiworapongsa T, Hassan SS, Yoon BH, Kim CJ. A signature of maternal anti-fetal rejection in spontaneous preterm birth: chronic chorioamnionitis, anti-human leukocyte antigen antibodies, and C4d. PLoS One. 2011; 6(2):e16806.10.1371/journal.pone.0016806 [PubMed: 21326865]

Lesack K, Naugler C. An open-source software program for performing Bonferroni and related corrections for multiple comparisons. J Pathol Inform. 2011; 2:52.10.4103/2153-3539.91130 [PubMed: 22276243]

Li X, Howard TD, Zheng SL, Haselkorn T, Peters SP, Meyers DA, Bleecker ER. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. J Allergy Clin Immunol. 2010; 125(2):328–335. e311.10.1016/j.jaci.2009.11.018 [PubMed: 20159242]

Lusis AJ. Atherosclerosis. Nature. 2000; 407(6801):233–241.10.1038/35025203 [PubMed: 11001066]

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature. 2009; 461(7265):747–753.10.1038/nature08494 [PubMed: 19812666]

Mathews, TJ.; MacDorman, MF. Natl Vital Stat Rep. Vol. 59 . Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System Genome-Wide Association Studies of Prematurity and Its Complications; 2011. Infant mortality statistics from the 2007 period linked birth/infant death data set; p. 1-30.2010 http://www.ncbi.nlm.nih.gov/projects/gap/cgibin/study.cgi?study_id=phs000103.v1.p1 [Accessed 14 Dec 2011]

Menon R, Fortunato SJ, Edwards DR, Williams SM. Association of genetic variants, ethnicity and preterm birth with amniotic fluid cytokine concentrations. Ann Hum Genet. 2010; 74(2):165–183. [PubMed: 20369436]

Menon R, Torloni MR, Voltolini C, Torricelli M, Merialdi M, Betran AP, Widmer M, Allen T, Davydova I, Khodjaeva Z, Thorsen P, Kacerovsky M, Tambor V, Massinen T, Nace J, Arora C. Biomarkers of spontaneous preterm birth: an overview of the literature in the last four decades. Reprod Sci. 2011; 18(11):1046–1070.10.1177/1933719111415548 [PubMed: 22031189]

Park JS, Park CW, Lockwood CJ, Norwitz ER. Role of cytokines in preterm labor and birth. Minerva Ginecol. 2005; 57(4):349–366. [PubMed: 16170281]

Plunkett J, Feitosa MF, Trusgnich M, Wangler MF, Palomar L, Kistka ZA, DeFranco EA, Shen TT, Stormo AE, Puttonen H, Hallman M, Haataja R, Luukkonen A, Fellman V, Peltonen L, Palotie A, Daw EW, An P, Teramo K, Borecki I, Muglia LJ. Mother's genome or maternally-inherited genes acting in the fetus influence gestational age in familial preterm birth. Hum Hered. 2009; 68(3):209–219.10.1159/000224641 [PubMed: 19521103]

Plunkett J, Doniger S, Orabona G, Morgan T, Haataja R, Hallman M, Puttonen H, Menon R, Kuczynski E, Norwitz E, Snegovskikh V, Palotie A, Peltonen L, Fellman V, DeFranco EA, Chaudhari BP, McGregor TL, McElroy JJ, Oetjens MT, Teramo K, Borecki I, Fay J, Muglia L. An evolutionary genomic approach to identify genes involved in human birth timing. PLoS Genet. 2011; 7(4):e1001365.10.1371/journal.pgen.1001365 [PubMed: 21533219]

[Accessed 27 Aug 2012] Power and sample size calculation: a review and computer program. 1998. http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–575.10.1086/519795 [PubMed: 17701901]

Schillert A, Ziegler A. Genotype calling for the Affymetrix platform. Methods Mol Biol. 2012; 850:513–523.10.1007/978-1-61779-555-8_28 [PubMed: 22307717]

Sharma PK, Sankar MJ, Sapra S, Saxena R, Karthikeyan CV, Deorari A, Agarwal R, Paul V. Growth and neurosensory outcomes of preterm very low birth weight infants at 18 months of corrected age. Indian J Pediatr. 201110.1007/s12098-011-0442-9

Snegovskikh VV, Bhandari V, Wright JR, Tadesse S, Morgan T, Macneill C, Foyouzi N, Park JS, Wang Y, Norwitz ER. Surfactant protein-A (SP-A) selectively inhibits prostaglandin F2alpha (PGF2alpha) production in term decidua: implications for the onset of labor. J Clin Endocrinol Metab. 2011; 96(4):E624–E632.10.1210/jc.2010-1496 [PubMed: 21270323]

Uzun A, Laliberte A, Parker J, Andrew C, Winterrowd E, Sharma S, Istrail S, Padbury JF. dbPTB: a database for preterm birth. Database J Biol Databases Curation. 201210.1093/database/bar069

Table 1

GWAS-identified cardiovascular, inflammatory, and metabolic disease gene variant

We selected 35 SNPs located in 33 genes/regions related to cardiovascular disease (including myocardial infarction, hypertension, hemostasis/thrombosis), growth/metabolism (e.g., lipids, glucose metabolism, obesity, height/weight), or inflammatory/immunological disease, and associated with the specific phenotypes with a P-value cut off for association of $< 1.0 \times 10^{-30}$

Disease/Trait	Region	Reported Gene(s)	Strongest SNP-Risk Allele	Risk Allele Freq in Controls	P-value	EntrezGene Pathway
Rheumatoid arthritis	6p21.32	*HLA-DRB1*	rs6457620	0.5	4×10^{-186}	immune response
Type 1 diabetes	6p21.32	*HLA-DQA1***	rs9272346-G	0.61	5×10^{-134}	immune response
Type 1 diabetes	6p21.32	*HLA-DRA*	rs9268645	NR	1×10^{-100}	immune response
Psoriasis	6p21.33	*HLA-C***	rs12191877-T	0.15	1×10^{-100}	immune response
HDL cholesterol	16q13	*CETP***	rs1532624-C	0.57	9×10^{-94}	lipoprotein metabolism
Type 1 diabetes	1p13.2	*PTPN22***	rs2476601	NR	9×10^{-85}	T cell differentiation, T cell receptor signaling pathway
Rheumatoid arthritis	6p21.32	*MHC*	rs6457617-T	0.49	5×10^{-75}	immune response
Ulcerative colitis	6p21.33	*CCHCR1*, HLA region*	rs9263739-T	0.15	4×10^{-67}	cell differentiation and proliferation
Crohn's disease	1p31.3	*IL23R*	rs11465804-T	0.93	7×10^{-63}	inflammatory response, innate immune response
Triglycerides	11q23.3	*ZNF259*, APOC3, APOA4, APOA5*	rs964184-G	0.14	4×10^{-62}	cell proliferation
LDL cholesterol	19q13.32	*APOE, APOC1*, APOC4, APOC2*	rs4420638-G	0.2	1×10^{-60}	lipoprotein metabolism
Fasting plasma glucose	2q24.3	*G6PC2*	rs560887-C	0.7	1×10^{-57}	glycolysis, gluconeogenesis
Systemic lupus erythematosus	6p21.33	*MSH5*, HLA region*	rs3131379-A	0.1	2×10^{-52}	meiosis
LDL cholesterol	19p13.2	*LDLR*	rs6511720-T	0.1	2×10^{-51}	lipoprotein metabolism
Body mass index	16q12.2	*FTO***	rs9939609-A	0.41	4×10^{-51}	posttranslational modification of proteins, DNA repair, and fatty acid metabolism
Fasting plasma glucose	11q21	*MTNR1B*	rs10830963-G	0.28	3×10^{-50}	melatonin, circadian rhythm
Type 1 diabetes	11p15.5	*INS*	rs7111341	NR	4×10^{-48}	cell-cell signaling, glucose homeostasis
Soluble ICAM-1	19p13.2	*ICAM1*	rs1799969-G	0.12	4×10^{-47}	T cell activation
LDL cholesterol	1p32.3	*PCSK9*	rs11591147-T	0.01	2×10^{-44}	regulation of cholesterol homeostasis
Myocardial infarction (early onset)	9p21.3	*CDKN2A*, CDKN2B*	rs4977574-G	0.56	3×10^{-44}	cell cycle, induction of apoptosis

Disease/Trait	Region	Reported Gene(s)	Strongest SNP-Risk Allele	Risk Allele Freq in Controls	P-value	EntrezGene Pathway
C-reactive protein	1q23.2	*CRP*	rs7553007-A	NR	8×10^{-44}	negative regulation of lipid storage, inflammatory response
LDL cholesterol	1p13.3	*CELSR2*, PSRC1, SORT1*	rs12740374-T	0.21	2×10^{-42}	Wnt receptor signaling pathway
Type 2 diabetes	11p15.5	*KCNQ1*	rs2237892-C	0.61	2×10^{-42}	muscle contraction, negative regulation of insulin secretion
Systemic lupus erythematosus	2q32.3	*STAT4*	rs7574865-A	0.33	5×10^{-42}	T-cell differentiation, JAK-STAT cascade
Triglycerides	8p21.3	*LPL*	rs12678919-G	0.1	2×10^{-41}	triglyceride homeostasis
Fibrinogen	4q32.1	*FGB*, FGA, FGG***	rs6056-A	0.18	8×10^{-39}	platelet adhesion and activation
Weight	16q12.2	*FTO*	rs8050136-A	0.41	5×10^{-36}	posttranslational modification of proteins, DNA repair, and fatty acid metabolism
HDL cholesterol	15q22.1	*LIPC***	rs1532085-G	0.59	1×10^{-35}	triglycerid catabolic process, cholesterol homeostasis
Type 2 diabetes	10q25.2	*TCF7L2***	rs7903146-T	0.3	2×10^{-34}	Wnt receptor signaling pathway, glucose homeostasis
HDL cholesterol	8p21.3	*LPL*	rs12678919-G	0.1	2×10^{-34}	triglyceride homeostasis
Serum Matrix Metalloproteinase	11q22.2	*intronic*; MMP1/3 region*	rs495366-A	0.36	6×10^{-34}	degradation of extracellular collagens
LDL cholesterol	1p13.3	*CELSR2, PSRC1*-***	rs599839-G	0.19	1×10^{-33}	cell cycle, negative regulation of cell growth
Crohn's disease	2q37.1	*ATG16L1*	rs3828309-G	0.53	2×10^{-32}	autophagy, protein transport
Systemic lupus erythematosus	1q25.1	*TNFSF4*	rs2205960-A	0.27	3×10^{-32}	inflammatory response, T cell proliferation
Triglycerides	2p23.3	*GCKR***	rs780094-T	0.39	6×10^{-32}	glucose homeostasis
C-reactive protein	12q24.31	*HNF1A*	rs1183910-T	NR	1×10^{-30}	glucose homeostasis
Psoriasis	1q21.3	*intronic*, LCE3D, LCE3A*	rs4085613-A	0.43	7×10^{-30}	late cornified epithelium protein

Table 2

Summary of association analysis results

Fisher's exact P values of < 0.05 are highlighted in bold.

			US White Mothers (307 cases, 342 controls)		US African-American Mothers (133 cases, 285 controls)		Finnish Mothers (160 cases, 200 controls)		US Hispanic Mothers (73 cases, 292 controls)	
Locus	**SNP**	**MA**	**OR (95CI)**	**P-value**	**OR (95CI)**	**P-value**	**OR (95CI)**	**P-value**	**OR (95CI)**	**P-value**
MTNR1B	rs10830963	G	1.19(0.93,1.53)	0.18	0.87(0.51,1.51)	0.68	1.12(0.82,1.52)	0.48	0.68(0.42,1.12)	0.16
IL23R	rs11465804	G	0.61(0.39,0.95)	**0.035**	0.89(0.44,1.79)	0.86	0.94(0.47,1.89)	1	0.58(0.17,1.98)	0.6
PCSK9	rs11591147	T	0.49(0.17,1.42)	0.21	8.3(0.92,74.58)	**0.042**	0.83(0.35,1.96)	0.83	0.99(0.11,8.92)	1
HNF1A	rs1183910	T	1.09(0.86,1.40)	0.49	1.23(0.83,1.84)	0.3	0.94(0.69,1.28)	0.7	1.08(0.74,1.59)	0.7
HLA-C	rs12191877	T	0.83(0.50,1.40)	0.51	1.02(0.47,2.22)	1	0.91(0.57,1.48)	0.81	NA	NA
LPL	rs12678919	G	0.59(0.39,0.88)	**0.012**	1.23(0.77,1.98)	0.39	0.89(0.53,1.50)	0.69	0.88(0.4,1.93)	0.85
CELSR2, PSRC1, SORT1	rs12740374	T	0.91(0.70,1.19)	0.5	1.28(0.91,1.81)	0.18	1.10(0.77,1.55)	0.66	1.09(0.68,1.74)	0.72
LIPC	rs1532085	A	0.94(0.74,1.18)	0.6	0.87(0.65,1.17)	0.36	1.00(0.74,1.35)	1	1.10(0.76,1.60)	0.64
CETP	rs1532624	T	0.89(0.71,1.11)	0.32	0.88(0.58,1.33)	0.6	0.91(0.68,1.22)	0.55	1.01(0.69,1.47)	1
ICAM1	rs1799969	A	1.39(0.97,1.99)	0.08	0.73(0.26,2.06)	0.63	1.21(0.83,1.75)	0.34	0.91(0.56,1.48)	0.81
TNFSF4	rs2205960	T	1.11(0.84,1.47)	0.48	1.75(0.98,3.14)	0.06	1.03(0.74,1.45)	0.86	0.94(0.62,1.42)	0.83
KCNQ1	rs2237892[4]	T	1.24(0.67,2.28)	0.54	2.72(1.30,5.69)	**0.009**	0.73(0.36,1.48)	0.48	0.57(0.32,1.04)	0.08
PTPN22	rs2476601	T	0.75(0.39,1.44)	0.4	0.38(0.08,1.89)	0.31	0.92(0.59,1.44)	0.74	1.00(0.15,6.77)	1
MSH5, HLA region	rs3131379	A	0.90(0.52,1.55)	0.78	0.96(0.30,3.13)	1	0.91(0.54,1.55)	0.79	NA	NA
ATG16L1	rs3828309	C	1.17(0.94,1.47)	0.17	0.93(0.66,1.31)	0.73	1.11(0.82,1.49)	0.55	0.89(0.59,1.34)	0.61
intronic, LCE3D, LCE3A	rs4085613	A	1.02(0.81,1.29)	0.86	1.22(0.90,1.64)	0.22	1.06(0.78,1.43)	0.76	1.00(0.69,1.46)	1
APOE, APOC1, APOC4, APOC2	rs4420638[2]	G	1.11(0.82,1.50)	0.49	0.96(0.66,1.40)	0.85	0.82(0.59,1.14)	0.24	1.19(0.64,2.24)	0.62
intronic; MMP1/3 region	rs495366[4]	A	0.82(0.64,1.06)	0.14	0.96(0.70,1.30)	0.81	0.93(0.67,1.28)	0.68	0.94(0.65,1.36)	0.78
CDKN2A, CDKN2B	rs4977574	G	0.75(0.60,0.94)	**0.013**	1.03(0.71,1.48)	0.93	0.87(0.65,1.16)	0.37	1.28(0.89,1.86)	0.21
G6PC2	rs560887	A	0.89(0.69,1.15)	0.36	1.20(0.65,2.22)	0.63	1.11(0.81,1.52)	0.52	1.10(0.63,1.92)	0.77
CELSR2, PSRC1	rs599839	G	0.74(0.49,1.12)	0.17	1.50(0.73,3.10)	0.28	1.07(0.76,1.52)	0.72	NA	NA
FGB, FGA, FGG	rs6056	T	0.77(0.57,1.05)	0.1	0.48(0.25,0.92)	**0.024**	0.88(0.60,1.28)	0.56	0.84(0.46,1.54)	0.66
MHC	rs6457617	C	0.69(0.49,0.99)	**0.049**	1.11(0.65,1.91)	0.78	0.96(0.71,1.30)	0.82	NA	NA
HLA-DRB1	rs6457620	C	1.43(1.01,2.04)	**0.049**	1.11(0.65,1.91)	0.78	1.03(0.76,1.38)	0.88	NA	NA

			US White Mothers (307 cases, 342 controls)		US African-American Mothers (133 cases, 285 controls)		Finnish Mothers (160 cases, 200 controls)		US Hispanic Mothers (73 cases, 292 controls)	
Locus	**SNP**	**MA**	**OR (95CI)**	**P-value**	**OR (95CI)**	**P-value**	**OR (95CI)**	**P-value**	**OR (95CI)**	**P-value**
LDLR	rs6511720	T	0.97(0.68,1.39)	0.93	1.05(0.64,1.72)	0.9	1.26(0.76,2.10)	0.43	0.17(0.04,0.70)	**0.004**
INS	rs7111341	T	0.94(0.74,1.20)	0.66	0.75(0.55,1.02)	0.07	0.78(0.56,1.10)	0.17	1.01(0.65,1.57)	1
CRP	rs7553007	A	1.00(0.79,1.28)	1	1.02(0.71,1.47)	0.93	1.06(0.78,1.44)	0.76	0.85(0.57,1.25)	0.43
STAT4	rs7574865	T	1.12(0.86,1.47)	0.41	1.27(0.84,1.91)	0.28	0.87(0.62,1.23)	0.48	0.74(0.5,1.11)	0.17
GCKR	rs780094	A	1.11(0.88,1.40)	0.38	1.49(1.02,2.18)	**0.044**	1.05(0.77,1.44)	0.81	1.02(0.69,1.51)	0.92
TCF7L2	rs7903146	T	1.17(0.91,1.51)	0.23	0.96(0.68,1.35)	0.86	1.34(0.93,1.94)	0.11	1.44(0.93,2.25)	0.12
FTO	rs8050136	A	1.00(0.79,1.26)	1	1.12(0.83,1.52)	0.49	1.10(0.82,1.48)	0.55	1.05(0.69,1.6)	0.83
CCHCR1, HLA region	rs9263739	T	0.97(0.6,1.57)	0.9	0.90(0.45,1.78)	0.86	1.58(1.01,2.46)	**0.044**	NA	NA
HLA-DRA	rs9268645	G	1.75(1.22,2.50)	**0.003**	0.82(0.45,1.50)	0.55	0.83(0.61,1.14)	0.27	NA	NA
HLA-DQA1	rs9272346	G	0.65(0.46,0.94)	**0.022**	0.80(0.46,1.37)	0.49	1.18(0.88,1.58)	0.3	NA	NA
ZNF259, APOA1, APOC3, APOA4, APOA5	rs964184	G	1.29(0.94,1.79)	0.12	1.05(0.72,1.53)	0.85	1.18(0.78,1.79)	0.46	0.58(0.38,0.9)	**0.015**
FTO	rs9939609[1,2,4]	A	1.25(0.96,1.63)	0.11	1.07(0.75,1.51)	0.72	1.09(0.81,1.46)	0.6	1.31(0.7,2.43)	0.4

[1]US white controls are out of HWE (P<0.001)

[2]US African-American controls are out of HWE (P<0.001)

[3]Finnish controls are out of HWE (P<0.001)

[4]Hispanic controls are out of HWE (P<0.001)

SNP, single nucleotide polymorphism; MA, minor allele; OR, odds ratio; CI, confidence interval;

Table 3

Testing the significant SNPs for the association with preterm birth in the Danish National Birth Cohort

Danish National Birth Cohorts Mothers (882 cases, 958 controls)					
Locus	**SNP**	**MA**	**MAF**	**OR (95CI)**	**Fisher's -p-value**
IL23R	rs11465804	G	0.06	1.29(0.96,1.74)	0.1
GCKR	rs780094	A	0.35	0.98(0.86,1.12)	0.78
FGB, FGA, FGG	rs6056	T	0.19	0.11(0.92,1.27)	0.38
CCHCR1, HLA region	rs9263739	T	0.16	1.15(0.97,1.38)	0.12
HLA-DRA	rs9268615*	A	0.38	1.10(0.96,1.26)	0.16
HLA-DQB1	rs1063355**	A	0.41	0.85(0.75,0.97)	0.017
MHC	rs6457617	T	0.49	1.06(0.93,1.21)	0.37
LPL	rs12678919	G	0.1	0.99(0.80,1.22)	0.91
CDKN2A, CDKN2B	rs4977574	G	0.43	0.92(0.81,1.05)	0.22
KCNQ1	rs2237892	T	0.05	0.68(0.52,0.90)	0.008
ZNF259, APOA1, APOC3, APOA4, APOA5	rs964184	G	0.15	1.00(0.83,1.20)	1
LDLR	rs6511720	T	0.12	1.25(1.02,1.54)	0.038

* proxy for rs9268645

** rs1063355 was genotyped as proxy for rs9272346

MA, minor allele; MAF, minor allele frequency in controls

Table 4

The genotype frequencies for the rs9272346 HLA-DQA1 in US White Mothers.

HLA-DQA1	**Preterm labor < 37 weeks**		**Total**	**Fisher-P-value**
rs9272346	**Case**	**Control**		
GG	13(8.8%)	19(17.6%)	32	
GA	77(52.0%)	59(54.6%)	136	
AA	58(39.2%)	30(27.8%)	88	**0.022**
Total	148	108	256	
G freq	0.35	0.45		

Table 5

The genotype frequencies of the rs9272346 HLA-DQB1 in the Danish National Birth Cohort mothers.

	Preterm labor < 37 weeks		**Total**	**Fisher –P-Value**
rs1063355	**Case**	**Control**		
AA	146(16.67%)	192(20.25%)	338	
AC	432(49.32%)	472(49.79%)	904	
CC	298(34.02%)	284(29.96%)	582	**0.048**
Total	876	948	1824	
A freq	0.41	0.45		