Comprehensive Testing of Positionally Cloned Asthma Genes in Two Populations

Craig P. Hersh, Benjamin A. Raby, Manuel E. Soto-Quirós, Amy J. Murphy, Lydiana Avila, Jessica Lasky-Su, Jody S. Sylvia, Barbara J. Klanderman, Christoph Lange, Scott T. Weiss, Juan C. Celedón

Online Data Supplement

Methods

Childhood Asthma Management Program (CAMP) Genetics Ancillary Study

CAMP is a multi-center North American clinical trial designed to investigate the long-term effects of inhaled anti-inflammatory medications in children with mild to moderate asthma (1, 2). Of the 1,041 children originally enrolled, 968 children and 1,518 parents contributed DNA samples for an ancillary study of the genetic determinants of asthma. Sufficient quantities of DNA were available for all family members for 458 of the 474 nuclear families of self-reported non-Hispanic white ancestry studied previously (3). Thirty-two of these 458 families had more than one asthmatic offspring, including thirty families with two asthmatic offspring and two families with three, resulting in a total of 497 asthmatic children available for analysis. A diagnosis of asthma was based on methacholine hyperresponsiveness ($PC_{20} \leq 12.5$ mg/ml) and one or more of the following criteria for at least six months in the year prior to recruitment: (i) asthma symptoms at least two times per week; (ii) at least two usages per week of an inhaled bronchodilator; and (iii) use of daily asthma medication. Airway responsiveness was assessed prior to treatment randomization by methacholine challenge with the Wright nebulizer tidal breathing technique (1). Serum total immunoglobulin E (IgE) was measured by radioimmunosorbent assays from blood samples collected during the CAMP screening sessions. The Institutional Review Board of the Brigham and Women's Hospital (BWH), as well as those of the other CAMP study centers, approved this study. Informed assent and consent were obtained from the study participants and their parents to collect DNA for genetic studies.

Costa Rica Study

Schoolchildren aged 6-14 years with asthma were recruited through 95 schools in the Central Valley of Costa Rica, as previously described (4). The Central Valley is a relatively genetically isolated population (5), with extensive genealogical records that can be used to trace ancestry back to approximately 4000 founding individuals in the 1697 census (6). Inclusion criteria included physician-diagnosed asthma, at least two respiratory symptoms (cough, wheeze or dyspnea) or a history of asthma attacks in the previous year, and a high probability of having at least at least six great-grandparents born in the Central Valley of Costa Rica (as determined by our study genealogist on the basis of the paternal and maternal last names of each of the child's parents). The latter criterion was required to increase the likelihood that children would be descendants of the founder population of the Central Valley (7); the birthplace of great-grandparents of participating children was later confirmed by our genealogy team by review of notary and/or church records. Of the participating children, 91% had confirmation of having at least six great-grand parents born in the Central Valley (lack of confirmation was mostly due to adoption or birth out of wedlock). Parents of probands were invited to participate. Parents provided written informed consent for themselves and for their children, who also gave written assent. The study was approved by the Institutional Review Boards of BWH and the Hospital Nacional de Niños, San José, Costa Rica.

Phenotyping protocols have been previously described (4). For each proband, a parent completed a study questionnaire, which was a modified version of that used by the Collaborative Study on the Genetics of Asthma (CSGA)(8) that had been translated into Spanish. Spirometry was performed according to American Thoracic Society standards

(9, 10). Methacholine challenge testing utilized a modified version of the Chatham protocol (11, 12). Serum total IgE levels were determined by the UniCAP 250 system (Pharmacia & Upjohn, Kalamazoo, MI), with samples measured in duplicate (13).

SNP Selection and Genotyping

Using genotype data from European-Americans (CEU) in the International HapMap project (14), we applied a linkage disequilibrium (LD) tagging algorithm to capture common variation (r^2>0.8, minor allele frequency>0.1) across the genes studied (15). In addition, we included SNPs that were associated with asthma or asthma-related phenotypes in the original reports. Results for sixteen SNPs in ADAM33 have been previously published for CAMP (3). Additional ADAM33 SNPs were genotyped in CAMP to complete the LD-tagging of the locus.

SNPs were genotyped in a highly multiplexed allele-specific hybridization assay with the Illumina Bead Station 500G (San Diego, CA). Mendelian transmission was tested with PedCheck (16), and inconsistent SNPs and individuals were removed from analysis. The list of SNPs successfully genotyped in both cohorts is shown in Table E1 (online only).

Statistical Analysis

Pairwise LD was expressed as both D' and r^2, calculated using Haploview (17). To ensure phenotype comparability to the CAMP enrollment criteria, a strict definition of asthma was used in the Costa Rica study, which included methacholine hyperresponsiveness ($PD_{20} \leq 8.58$ μmoles) or bronchodilator responsiveness plus the

recruitment criteria above. In addition to asthma, two intermediate phenotypes common to both cohorts were analyzed: (i) airways hyperresponsiveness (AHR), measured as $\log_{10}$-transformed dose-response slope to methacholine (18); and (ii) $\log_{10}$-transformed total serum IgE levels (19). Haplotype blocks were defined using the Gabriel algorithm (20), and haplotype-tagging SNPs were identified in Haploview. Family-based association testing of single SNPs and haplotypes was performed using PBAT software (21) under additive genetic models. The family-based association test is a generalization of the transmission disequilibrium test that allows for the analysis of quantitative traits, tests of different genetic models, and missing parental genotypes (22). Analyses of the two quantitative traits were adjusted for age and gender; AHR analysis was additionally adjusted for height. No covariates were included in the asthma analyses (23). Power for replication of both affection status and the quantitative traits was estimated in PBAT. Additional statistical analyses were performed in SAS version 9.1 (Cary, NC) and in R. Since these genes have all been associated with asthma phenotypes in prior studies, we used a p-value<0.05 in both samples to define statistical significance in the setting of multiple tests, instead of an adjusted p-value (24). If we consider a gene-based multiple comparisons correction (given the relatively tight LD within each gene), a global α level of 0.05 still would be preserved after correction for the number of tests conducted (0.0025 * 5 genes * 3 phenotypes = 0.0375).

References

E1. The Childhood Asthma Management Program (CAMP): design, rationale, and methods. Childhood Asthma Management Program Research Group. *Control Clin Trials* 1999;20:91-120.

E2. Long-term effects of budesonide or nedocromil in children with asthma. The Childhood Asthma Management Program Research Group. *N Engl J Med* 2000;343:1054-63.

E3. Raby BA, Silverman EK, Kwiatkowski DJ, Lange C, Lazarus R, Weiss ST. ADAM33 polymorphisms and phenotype associations in childhood asthma. *J Allergy Clin Immunol* 2004;113:1071-8.

E4. Hunninghake GM, Soto-Quiros ME, Avila L, Ly NP, Liang C, Sylvia JS, Klanderman BJ, Silverman EK, Celedon JC. Sensitization to Ascaris lumbricoides and severity of childhood asthma in Costa Rica. *J Allergy Clin Immunol* 2007;119:654-61.

E5. Service S, DeYoung J, Karayiorgou M, Roos JL, Pretorious H, Bedoya G, Ospina J, Ruiz-Linares A, Macedo A, Palha JA, et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat Genet* 2006;38:556-60.

E6. Melendez C. Conquistadores y Pobladores: Origenes Historico-Sociales de los Costarricenses.San Jose, Costa Rica: Editorial Universidad Estatal a Distancia;1982.

E7. Escamilla MA, Spesny M, Reus VI, Gallegos A, Meza L, Molina J, Sandkuijl LA, Fournier E, Leon PE, Smith LB, et al. Use of linkage disequilibrium approaches to map genes for bipolar disorder in the Costa Rican population. *Am J Med Genet* 1996;67:244-53.

E8. Blumenthal MN, Banks-Schlegel S, Bleecker ER, Marsh DG, Ober C. Collaborative studies on the genetics of asthma--National Heart, Lung and Blood Institute. *Clin Exp Allergy* 1995;25 Suppl 2:29-32.

E9. American Thoracic Society. Standardization of Spirometry, 1994 Update. *Am J Respir Crit Care Med* 1995;152:1107-36.

E10. Hersh CP, Soto-Quiros ME, Avila L, Lake SL, Liang C, Fournier E, Spesny M, Sylvia JS, Lazarus R, Hudson T, et al. Genome-wide linkage analysis of pulmonary function in families of children with asthma in Costa Rica. *Thorax* 2007;62:224-230.

E11. Chatham M, Bleecker ER, Norman P, Smith PL, Mason P. A screening test for airways reactivity. An abbreviated methacholine inhalation challenge. *Chest* 1982;82:15-8.

E12. Celedon JC, Soto-Quiros ME, Avila L, Lake SL, Liang C, Fournier E, Spesny M, Hersh CP, Sylvia JS, Hudson TJ, et al. Significant linkage to airway responsiveness on chromosome 12q24 in families of children with asthma in Costa Rica. *Hum Genet* 2007;120:691-699.

E13. Raby BA, Soto-Quiros ME, Avila L, Lake SL, Murphy A, Liang C, Fournier E, Spesny M, Sylvia JS, Verner A, et al. Sex-specific linkage to total serum immunoglobulin E in families of children with asthma in Costa Rica. *Hum Mol Genet* 2007;16:243-53.

E14. The International HapMap Project. *Nature* 2003;426:789-96.

E15. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet* 2005;37:1217-23.

E16. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998;63:259-66.

E17. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-5.

E18. O'Connor G, Sparrow D, Taylor D, Segal M, Weiss S. Analysis of dose-response curves to methacholine. An approach suitable for population studies. *Am Rev Respir Dis* 1987;136:1412-7.

E19. Barbee RA, Halonen M, Lebowitz M, Burrows B. Distribution of IgE in a community population sample: correlations with age, sex, and allergen skin test reactivity. *J Allergy Clin Immunol* 1981;68:106-11.

E20. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, et al. The structure of haplotype blocks in the human genome. *Science* 2002;296:2225-9.

E21. Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM. PBAT: tools for family-based association studies. *Am J Hum Genet* 2004;74:367-9.

E22. Rabinowitz D, Laird N. A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum Hered* 2000;50:211-23.

E23. Poisson LM, Rybicki BA, Coon SW, Barnholtz-Sloan JS, Chase GA. Susceptibility scoring in family-based association testing. *BMC Genet* 2003;4 Suppl 1:S49.

E24. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007;445:881-5.

Figure Legends

Figure E1: Detailed linkage disequilibrium (LD) plots in GPR154 (NPSR1) in the (A) Costa Rica and (B) CAMP studies. LD is measured as D', with darker red colors indicating higher values. Haplotype blocks are defined using the confidence interval method (20). Online only.

Table E1: SNPs successfully genotyped in Costa Rica and CAMP studies. Online only.

Gene	SNP	Location*
ADAM33	rs2787094	chr20:3597161
	rs677044	chr20:3597431
	rs628965	chr20:3597713
	rs630712	chr20:3598066
	rs2280090	chr20:3598205
	rs597980	chr20:3599165
	rs44707	chr20:3599226
	rs615436	chr20:3600835
	rs3918395	chr20:3601149
	rs2280093	chr20:3601692
	rs2485700	chr20:3602993
	rs487377	chr20:3606931
DPP10	rs10208402	chr2:114920147
	rs6737251	chr2:114929905
	rs975453	chr2:115783214
	rs2420815	chr2:115794672
	rs7589983	chr2:115843613
	rs1980189	chr2:115867543
	rs12711825	chr2:115884302
	rs843385	chr2:115893809
	rs1521079	chr2:115899494
	rs1448461	chr2:115945009
	rs982174	chr2:115971413
	rs2008031	chr2:116027070
	rs7576583	chr2:116056043
	rs4516432	chr2:116098127
	rs4538204	chr2:116109896
	rs12105131	chr2:116112288
	rs4241136	chr2:116129881
	rs7575094	chr2:116161946
	rs2901392	chr2:116173021
	rs2421277	chr2:116186655
	rs2901396	chr2:116233972
	rs2034257	chr2:116249176
	rs10496510	chr2:116297183
	rs2421343	chr2:116306653
	rs272071	chr2:116322805
GPR154	rs2609234	chr7:34461665
	rs1006392	chr7:34464182
	rs714588	chr7:34466466
	rs898070	chr7:34472105
	rs1379928	chr7:34474529
	rs963218	chr7:34485009
	rs2609215	chr7:34489007
	rs10274146	chr7:34492887
	rs323917	chr7:34514883
	rs323920	chr7:34520257

	rs324377	chr7:34529215
	rs6961859	chr7:34533448
	rs324381	chr7:34541351
	rs182718	chr7:34543239
	rs1419789	chr7:34543803
	rs324957	chr7:34574612
	rs324960	chr7:34575631
	rs10278663	chr7:34581711
	rs740347	chr7:34585542
	rs324981	chr7:34591353
	rs1860370	chr7:34637678
	rs727162	chr7:34647278
	rs10250709	chr7:34660065
	rs6972158	chr7:34662422
	rs10238983	chr7:34672813
	rs2040854	chr7:34678046
	rs1833090	chr7:34681026
HLA-G	rs1736920	chr6:29905152
	rs2517892	chr6:29909453
SETDB2/	rs7987940	chr13:48906647
PHF11/	rs7982297	chr13:48925271
RCBTB1	rs7338755	chr13:48927977
	rs9316454	chr13:48940485
	rs7998427	chr13:48948621
	rs2057413	chr13:48955098
	rs11619265	chr13:48955634
	rs4941643	chr13:48966388
	rs3794381	chr13:48971564
	rs2031532	chr13:48978848
	rs2247119	chr13:48985143
	rs9568232	chr13:48987845
	rs2274276	chr13:48993954
	rs7332573	chr13:48996377
	rs7333668	chr13:48999527
	rs3829366	chr13:49001412
	rs3186013	chr13:49005905
	rs1925742	chr13:49007585
	rs7982555	chr13:49009511
	rs2407697	chr13:49018436
	rs3751383	chr13:49021650
	rs3751381	chr13:49021833
	rs942871	chr13:49022622
	rs9568246	chr13:49023623
	rs2274278	chr13:49024383
	rs6561542	chr13:49025521
	rs2038881	chr13:49033178
	rs11148151	chr13:49042614
	rs7981396	chr13:49049927
	rs1359541	chr13:49056940
	rs1325659	chr13:49059043

rs1409015	chr13:49062564

*May 2004 human genome reference sequence (NCBI Build 35).

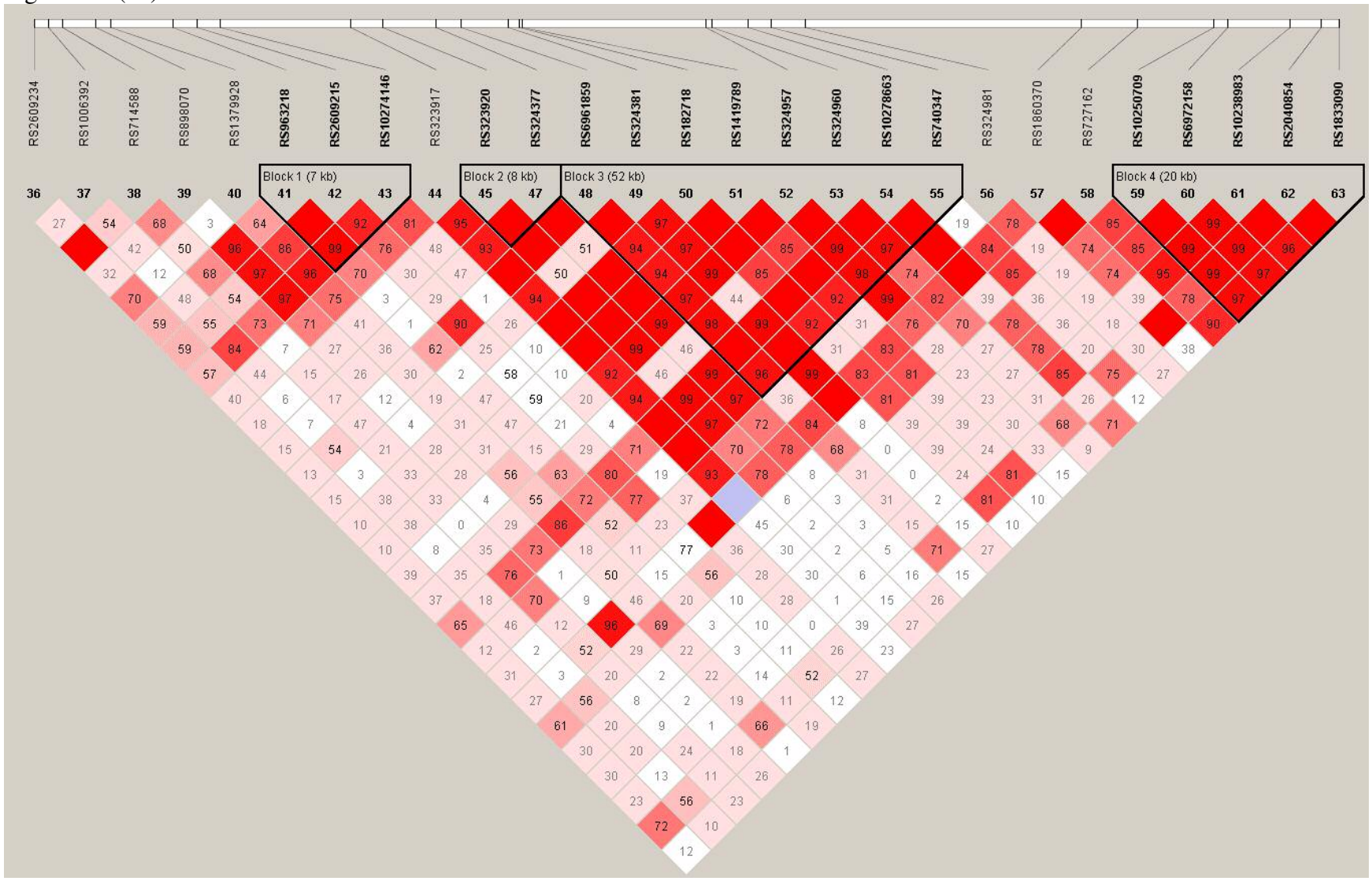
Table E2: Genetic associations with total Immunoglobulin E levels in Costa Rica and CAMP studies. SNPs with p-value ≤0.05 in either cohort are listed. No SNPs in ADAM33, DPP10, or HLAG (see Table E1 for list) were significant in either cohort. Online only.

Gene	SNP	Costa Rica			CAMP		
		MAF*	Families†	P-value	MAF*	Families†	P-value
GPR154	rs1860370	0.04	60	0.04	0.08	129	0.09
PHF11	rs9568232	0.14	166	0.9	0.08	122	0.03

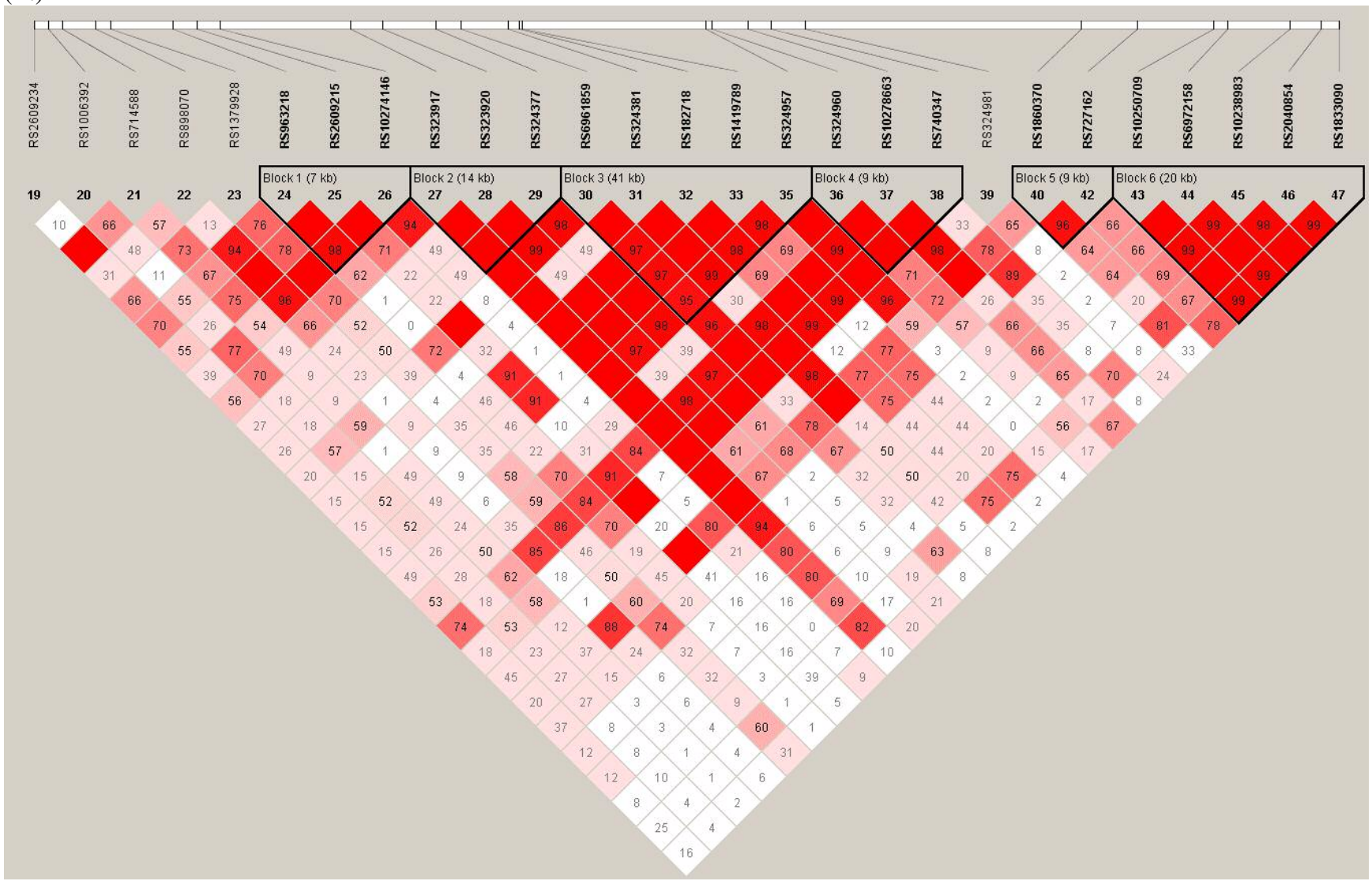
*MAF=minor allele frequency.

†Number of informative families.

Figure E1: (A.)


RS2609234
RS1006392
RS714588
RS898070
RS1379928
RS963218
RS2609215
RS10274146
RS323917
RS323920
RS324377
RS6961859
RS324381
RS182718
RS1419789
RS324957
RS324960
RS10278663
RS740347
RS324981
RS1860370
RS727162
RS10250709
RS6972158
RS10238983
RS2040854
RS1833090
Block 1 (7 kb)
Block 2 (8 kb)
Block 3 (52 kb)
Block 4 (20 kb)
36 37 38 39 40 41 42 43 44 45 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63

(B.)


RS2609234
RS1006392
RS714588
RS898070
RS1379928
RS963218
RS2609215
RS10274146
RS323917
RS323920
RS324377
RS6961859
RS324381
RS182718
RS1419789
RS324957
RS324960
RS10278663
RS740347
RS324981
RS1860370
RS727162
RS10250709
RS6972158
RS10238983
RS2040854
RS1833090
Block 1 (7 kb)
Block 2 (14 kb)
Block 3 (41 kb)
Block 4 (9 kb)
Block 5 (9 kb)
Block 6 (20 kb)
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 35 36 37 38 39 40 42 43 44 45 46 47