Repository - Unmarked Text

1 **ONLINE REPOSITORY**

2

3 **Identification of *ATPAF1* as a novel candidate gene for asthma in children**

4

5 Eric M. Schauberger[1], Susan L. Ewart[1], Syed H. Arshad[2,3], Marianne Huebner[1], Wilfried
6 Karmaus[4], John W. Holloway[3], Karen H. Friderici[1], Julie T. Ziegler[5], Hongmei Zhang[4], Mathew
7 J. Rose-Zerilli[3], Sheila J. Barton[3], Stephen T. Holgate[3], Jeffrey R. Kilpatrick[6], John B. Harley[6,7],
8 Stephane Lajoie-Kadoch[8], Isaac T. W. Harley[8], Qutayba Hamid[9], Ramesh J. Kurukulaaratchy[2,3],
9 Max A. Seibold[10], Pedro C. Avila[11], William Rodriguez-Citron[12], Jose R. Rodriguez-Santana[13],
10 Donglei Hu[14], Christopher Gignoux[14], Isabelle Romieu[15], Stephanie J. London[16], Esteban G.
11 Burchard[14], Carl D. Langefeld[5], and Marsha Wills-Karp[8]

12

13 *[1]Michigan State University, East Lansing, MI, USA*

14 *[2]The David Hide Asthma and Allergy Research Centre, Isle of Wight, UK*

15 *[3]University of Southampton School of Medicine, University of Southampton, Southampton, UK*

16 *[4]University of South Carolina, Columbia, SC, USA*

17 *[5]Wake Forest School of Medicine, Winston-Salem, NC, USA*

18 *[6]JKA Genomics, Oklahoma City, OK, USA*

19 *[7]Oklahoma Medical Research Foundation, Oklahoma City, OK, USA*

20 *[8]Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA*

21 *[9]Meakins-Christie Laboratory, McGill University, Montreal, Quebec, Canada*

22 *[10]National Jewish Health, Denver, CO, USA*

23 *[11]Northwestern University Feinberg School of Medicine, Chicago, IL, USA*

24 [12]*Veterans Affairs Medical Center, San Juan, PR*

25 [13]*Centro de Neumologia Pediatrica, CSP, San Juan, PR*

26 [14]*University of California, San Francisco, CA, USA*

27 [15]*National Institute of Public Health, Cuernevaca, MX*

28 [16]*National Institute of Environmental Health Sciences, National Institutes of Health, Department*

29 *of Health and Human Services, Research Triangle Park, NC, USA*

30

31 **Corresponding author:** Dr. Susan Ewart, ewart@cvm.msu.edu , G100 Veterinary Medical

32 Center, Michigan State University, East Lansing, MI, USA; Tel: (517) 884-2005; Fax: (517)

33 432-2310

34

35 **METHODS**

36

37 ***Study subjects***

38 **Isle of Wight birth cohort**.[E1] The primary population was a cohort of children (n=1,456) born
39 and enrolled between January 1, 1989 and February 28, 1990, on the Isle of Wight, UK. Children
40 were subsequently assessed at ages 1, 2, 4, and 10 years, with asthma diagnosis at each
41 assessment based on a minimum criteria of physician diagnosed asthma plus wheeze in the
42 previous 12 months, using a validated questionnaire.[E2] A subset of 112 children with asthma at
43 age 10 years and a control group of 165 children without asthma or wheeze at any age was
44 selected for genome-wide screening.

45

46 **Wessex families.**[E3] Caucasian families (n=341) containing at least two biological siblings aged
47 5-21 years with a current physician's diagnosis of asthma and who were taking asthma
48 medication on a regular basis were recruited from Wessex, UK. Asthma diagnosis was based on
49 validated health survey questionnaires completed by each family member.[E4]

50

51 **Childhood Asthma Management Program (CAMP) project and Childhood Asthma**
52 **Research and Education (CARE) network case-parent trios**. The publically available
53 SHARP data used (study accession: phs000166.v2.p1) were the case-parent trios from the CARE
54 network[E5, E6] and the CAMP project.[E5, E7, E8] The CAMP study cases were children aged 5-12
55 years who had chronic asthma with medication use and either asthma symptoms at least twice
56 per week or at least two usages per week of an inhaled bronchodilator. The CARE study children
57 were aged 6-17 years and had a positive asthma predicted index based on at least three

58 exacerbations of wheezing during the previous twelve months. The CARE population did not
59 include the PEAK study of infants at risk.
60

61 **Genetics of Asthma in Latino Americans (GALA) case-parent trios.**[E9] The GALA Study
62 includes children (probands) and their biological parents recruited from schools, clinics, and
63 hospitals that cared for Latino patients at four sites: San Francisco Bay Area, New York City,
64 Puerto Rico, and Mexico City. In all health care centers, medical records were reviewed to
65 identify patients with physician-diagnosed mild or moderate-severe asthma based on medical
66 billing records (ICD 9 codes). Patients were contacted to participate in the study if approved by
67 their primary physician. Patients were assessed by interviews and questionnaires (available in
68 English and Spanish) administered by bilingual and bicultural physicians specializing in asthma.
69 Children were included in the study if they were between the ages of 8-40 with physician
70 diagnosed mild to moderate-severe asthma and had experienced two or more symptoms
71 (including wheezing, coughing and/or shortness of breath) in the previous two years at the time
72 of recruitment. Trios were enrolled if both parents and 4 sets of grandparents of the proband self-
73 identified as either Puerto Rican (395 trios) or Mexican (298 trios) ethnicity.
74

75 **Mexico Childhood Asthma Study case-parent trios.**[E10-E12] Asthmatic children aged 5-17 years
76 and their parents were recruited from a population of patients at a referral clinic in central
77 Mexico City, MX (n=492 trios). The diagnosis of asthma was based on clinical symptoms and
78 response to treatment by pediatric allergists at a major referral hospital. The severity of asthma
79 was rated by a pediatric allergist according to symptoms in the Global Initiative on Asthma
80 schema as mild (intermittent or persistent), moderate, or severe.[E13]

81

82 **Consortia controls**. Additional consortia controls were acquired from 1) the Wellcome Trust
83 Case Control Consortium (WTCCC) (n=3,004), which derives from control subjects[E14] from the
84 1958 British Birth Cohort and from blood donors recruited by the three national UK Blood
85 Services, and 2) the Genetics of Systemic Lupus Erythematosus (SLEGEN) (n=3,471)
86 containing females from UK and USA origins.[E15] The asthma status of the consortia controls was
87 unknown.

88

89 **Asthmatic subjects examined for gene expression.**[E16, E17]. Eight patients with severe atopic
90 asthma (mean age: 41, FEV_1<80% predicted, on oral or inhaled steroids), five with mild asthma
91 (mean age: 40, FEV_1>80% predicted, on bronchodilators only, all nonatopic except one), and
92 four control subjects (mean age: 43, normal FEV_1 and PC_{20}, all nonatopic) were recruited for
93 bronchial biopsy and subsequent gene expression studies. All subjects were Caucasian
94 nonsmokers from the Montreal area of Quebec, Canada.

95

96 ***Genotyping and quality control***

97 **Isle of Wight pooled samples.** DNA samples assessed by NanoDrop spectrophotometry
98 (Wilmington, DE, USA) to have A_{260}/A_{280} range of 1.65-2.0 and A_{260}/A_{230} range of 1.0-2.2
99 qualified for inclusion in a pool. DNA samples were separated on 0.8% agarose gels to confirm
100 lack of degradation or RNA contamination. Equimolar amounts of DNA from individuals were
101 combined for a total of 250 ng/pool. DNA pools were digested with XbaI or HindIII enzyme,
102 adapter ligated, and PCR amplified. Then samples were separated on 4% agarose gels to ensure

DNA fragmentation in the 100-300 bp range. PCR yields (>1,200 ng/μl accepted) were compared between microarray chips to ensure uniformity and PCR products were separated on 2% agarose gels to ensure the proper range of product was amplified. Genotyping was conducted using the Affymetrix 100K GeneChip Mapping Array and GeneChip Genotyping software[E18] (v.4.0, Affymetrix, Inc.) was used for relative quality control assessment, detection rates (>98.5%), and allele distributions (<10% difference between pools). Hybridization intensity comparisons of the case and control pools were used to identify significant allele frequency differences for each SNP (Table E1).

Isle of Wight individual samples. To further investigate the top region associated with asthma (within 1p33-p32.31) tag SNPs in were chosen based on HapMap[E19] data (CEU Population, Release 21a/Phase II Jan 07, B35 data set) using Tagger[E20] and SNPbrowser software[E21] (Applied Biosystems) with a minor allele frequency threshold of 10% and pairwise R^2 threshold of 100%. SNPs with genotype and HapMap data available from the pooled sample analysis were also included. The selection list was scrutinized and haplotype groups were developed based on CEU phase data.[E19] A final list of tag SNPs was chosen from each group based on predicted assay design scores for the SNP beadarray. Individual sample (n=277) genotyping of SNPs was done by custom Illumina GoldenGate assays included a ThermoElectron KingFisher96 automated magnetic bead wash. Haploview[E22] was used to examine Hardy-Weinberg equilibrium, minor allele frequencies, and linkage disequilibrium. The genotype success rate for each SNP was 98.9% and the overall call rate was 99.7%. Five samples failed completely and were not reported here. Two samples assayed in duplicate as technical replicates had complete concordance. Eight samples from a different study genotyped here had a 99.8% call rate.

126

127 **Wessex samples.** DNA samples from the Wessex population were genotyped for selected SNPs
128 using KASPar competitive allele-specific PCR (Kbiosciences, Herts, UK). Replication sample
129 genotyping assays[E23] were designed and tested with a random panel of DNAs to ensure SNP
130 polymorphism. The genotype success rate for each SNP was ≥97.4% and the overall call rate
131 was 91%. Quality control measures included negative results for water controls, intra-plate
132 testing of a known DNA, and clear, distinct result clusters.

133

134 **CAMP and CARE samples.**[E5] CAMP and CARE population DNA samples were genotyped
135 using the Affymetrix 6.0 GeneChip Array. Quality control measures for markers included
136 comparing minor allele frequencies of the Caucasian subjects to the HapMap CEU samples, call
137 rates (>95%), Hardy-Weinberg equilibrium within Caucasian subjects ($p<10^{-6}$), and Mendelian
138 errors (<5). In addition, examined was the proportion of missing data for each plate across all
139 markers and the frequency differences across each plate using stringently quality controlled
140 markers. Based on these criteria 83.5% of markers in the CAMP study and 81.3% of markers in
141 the CARE study were retained. Quality control measures for samples included evaluation of
142 missingness, Mendelian inconsistencies, F inbreeding statistic, X chromosome heterozygosity
143 rates to check gender, allele sharing analysis, and population structure analysis. Meeting these
144 criteria were 96.8% of CAMP samples and 97.2% of CARE samples.

145

146 **GALA samples.** Genotyping was performed on the Affymetrix 6.0 GeneChip Array containing
147 > 900K SNPs prior to QC measures. Subjects were filtered based on 95% call rates, complete
148 trios, and Mendelian errors. Markers were filtered based on 95% call rates, Hardy-Weinberg

149 equilibrium p-values >10^{-6}, <1% Mendelian inconsistencies, unambiguous mapping to the
150 human reference genome, and no evidence for previous plate effects. The total number of trios
151 passing QC was 538 (n=1,614), and the total number of markers passing QC was 729,685. In
152 addition to direct genotyping, genotypes at some SNPs were imputed in the GALA subjects with
153 imputation performed using MACH[E24] using the phased CEU, YRI, and ASN phase 2, release 21
154 consensus HapMap genotypes as a reference. Allele frequencies in transmitted and untransmitted
155 chromosomes were estimated summing the dosages of the parents of complete trios, subtracting
156 the dosages of the children, and then taking the average (i.e. untransmitted allele frequency).
157 Family-based association statistics were calculated using allelic dosages assuming an additive
158 model in PBAT version 6.4.0.[E25]
159
160 **Mexico Childhood Asthma Study samples.**[E12] Genotyping was performed using the Illumina
161 HumanHap 550v3 BeadChip and genotypes determined by the Illumina BeadStudio Genotyping
162 Module. Successful genotype call rates in 1,491 subjects exceeded 95% with an average call rate
163 of 99.7%. Three trios were excluded because of a low call rate of 1 family member. Quality
164 control analyses for the 561,466 SNPs in the GWAS scan were conducted by using PLINK.[E26]
165 Sequential SNP exclusions were made due to poor chromosomal mapping (n=173), missing rate
166 of greater than 5% (n=4125), minor allele frequency of less than 1% (n=16,949), a Hardy-
167 Weinberg equilibrium p-value of less than 1×10^{-10} (n=557), Mendelian errors in more than 2
168 families (n=4,945), and heterozygous genotype calls for chromosome X SNPs in more than 1
169 male subject (n=380). SNPs with 1 or more discordant genotypes across 14 HapMap replicate
170 samples identified by using the Genotyping Library and Utilities application[E27] were also
171 excluded (n=921). Quality control at the subject level included inspection for unusual autosomal

172 homozygosity, inconsistent sex between genotype and collected phenotype data, and subject
173 relatedness. There were 492 complete case-parent trios in the final data set.
174

175 **WTCCC samples.** [E14] Genome-wide association study of 14,000 cases of seven common
176 diseases and 3,000 shared controls). Genotyping was conducted using the Affymetrix GeneChip
177 500K array. A genotype calling algorithm, CIAMO, was developed and applied to the whole
178 project. Successful samples were those with a call rate of 93% at p=0.33 for each array, over
179 90% concordance for the 50 SNPs that are common to the two arrays, both arrays agreed on
180 gender, and showed over 70% identity to the Sequenom genotypes supplied by WTCCC.
181

182 **SLEGEN samples.**[E15] Genotyping was conducted using the Illumina HapMap300 array. The
183 265,648 (84%) SNPs that met the following criteria were used in the data analysis: 1) no
184 statistically significant differences in the proportions of missing genotype data between cases and
185 controls ($p>0.05$), 2) overall <5% missing genotype data, 3) Hardy-Weinberg equilibrium in
186 controls $p>0.01$, and in cases $p>0.0001$, and 4) allele frequencies of controls statistically
187 consistent with expectations for ethnicity matched HapMap samples ($p>0.01$). Genotype data
188 were only used from samples with a call rate greater than 90% (all samples), from SNPs with a
189 call frequency greater than 90% (removed 27 SNPs), and an Illumina GenTrain score greater
190 than 0.7 (removed 152 SNPs). Samples identity was verified in 91 SNPs that had been
191 previously genotyped on 42% of the samples. At least one sample previously genotyped was
192 randomly placed on each array and used to track samples throughout the genotyping process.
193

194 ***Statistical analyses***

195 **Pooled samples.** Allele frequencies for Affymetrix data were determined by a structured
196 analysis for pooled samples[E28] from Isle of Wight subjects. Z^2 p-values were used to rank SNPs.
197 An autoclustering algorithm was used on all SNPs. Clusters of SNPs were manually inspected
198 when they had low call rate ($<98.5\%$), low clustering score (<0.6), or significant departure from
199 Hardy-Weinberg equilibrium ($p<0.05$). A total of 38 of the 96 clusters were edited to adjust the
200 autoclustering algorithm. A cluster analysis of Z^2 statistics was performed. Cluster parameters of
201 size, SNP numbers, and level of significance were considered with preference given to small
202 clusters of high significance. A subset of these clusters was assayed by individual genotyping.
203

204 **Haplotypes.** Odds ratios for haplotypes for Isle of Wight and Wessex data were computed using
205 regression models. Conditional probabilities of subjects' compatible haplotype pairs were used
206 given the observed marker phenotypes and an estimate of haplotype frequencies.[E29] Haploview
207 was used to generate linkage disequilibrium plots for the populations.[E22]
208

209 **Consortia controls.** The use of consortia controls has the benefit of mitigating the concern of
210 lack of statistical power while remaining efficient.[E8, E15] Power analysis performed using
211 simulations with and without the inclusion of 3,000 external controls, and assuming a
212 misclassification rate of 15%, demonstrated testing powers increased at least 20%. Consortia
213 control data from SLEGEN and WTCCC were independently combined with Isle of Wight control
214 data. Several SNPs in common with the Isle of Wight SNPs had been directly genotyped in the
215 SLEGEN and WTCCC control populations. In addition, a high degree of linkage disequilibrium
216 in this region allowed imputation to infer additional genotypes using BimBam software[E30] and
217 CEU Hapmap build 36. Results of analysis from the variants that were directly genotyped were

218 found to be similar to those of imputed genotypes (for example, rs2218189 had a p-value=0.00010
219 from direct genotyping and p=0.00011 from genotype imputation). Hardy Weinberg equilibrium
220 was calculated for both WTCCC and SLEGEN. SNPGWA[E15] software with an additive model
221 using a Cochran-Armitage trend test was used to determine statistical significance in combined Isle
222 of Wight and consortia data.
223

224 ***ATPAF1 expression in biopsied bronchial tissue***

225 Gene expression studies were conducted on bronchial biopsy samples collected from 13
226 asthmatic and 4 control subjects. Endoscopic biopsies were obtained using 3.5 mm cup forceps
227 from the lower lobe of the right lung and stored in the Tissue Bank (MCI/Meakins-Christie
228 Tissue Bank, McGill University, Montreal, Quebec, Canada).[E16, E17] Tissue RNA was extracted
229 using the RNeasy micro kit extraction columns (Qiagen, Valencia, CA). Total RNA was reverse
230 transcribed with oligo(dT) (Amersham Pharmacia Biotech, Pittsburgh, PA) and Superscript II
231 (Invitrogen, Carlsbad, CA) in the presence of RNAguard (Amersham Pharmacia Biotech,
232 Pittsburgh, PA). *ATPAF1* levels were normalized using ribosomal protein S9 gene expression.
233 Primers spanned at least one intron. Primer sequences: ATPAF1.sense 5'-
234 AAGTGGAGTTCAGTACCTGTCCA-3'; ATPAF1.antisense 5'-
235 GGCTCAGTCCTGTCCAACA-3'; S9.sense 5'-TGCTGACGCTTGATGAGAAG-3';
236 S9.antisense 5'-CGCAGAGAGAAGTCGATGTG-3'. No template controls and standard curves
237 were run for each gene. The standard curves contained four standards ranging from 5K to 50M
238 copies for ribosomal protein S9 and 1 to 10K copies for *ATPAF1*. Test samples were run in
239 singlet (due to paucity of sample amounts) and all test samples measured within the standard
240 curve values. Expression data were analyzed with Kruskal-Wallis followed by Dunn's multiple

241 comparison tests using GraphPad Prism (v.4.0, GraphPad Software, San Diego, CA,

242 http://www.graphpad.com).

243

244 **REFERENCES FOR ONLINE REPOSITORY**

245 E1. Kurukulaaratchy RJ, Fenn MH, Waterhouse LM, Matthews SM, Holgate ST, Arshad
246 SH. Characterization of wheezing phenotypes in the first 10 years of life. Clin Exper Allergy
247 2003; 33:573-578.
248 E2. The international study of asthma and allergies in childhood (ISAAC) steering committee.
249 Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and
250 atopic eczema: ISAAC. Lancet 1998; 351:1225-1232.
251 E3. Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, Torrey D,
252 Pandit S, McKenny J, Braunschweiger K, et al. Association of the ADAM33 gene with
253 asthma and bronchial hyperresponsiveness. Nature 2002; 418:426-430.
254 E4. Beghe B, Barton S, Rorke S, Peng Q, Sayers I, Gaunt T, et al. Polymorphisms in the
255 interleukin-4 and interleukin-4 receptor alpha chain genes confer susceptibility to asthma and
256 atopy in a Caucasian population. Clin Exper Allergy 2003; 33:1111-1117.
257 E5. The database of Genotypes and Phenotypes (dbGaP), National Center for Biotechnology
258 Information, U. S. National Library of Medicine. From http://www.ncbi.nlm.nih.gov/gap
259 E6. Guilbert TW, Morgan WJ, Krawiec M, Lemanske RF, Jr., Sorkness C, Szefler SJ, et al. The
260 Prevention of Early Asthma in Kids study: design, rationale and methods for the Childhood
261 Asthma Research and Education network. Control Clin Trials 2004; 25:286-310.
262 E7. The Childhood Asthma Management Program Research Group. Long-term effects of
263 budesonide or nedocromil in children with asthma. N Engl J Med 2000; 343:1054-63.
264 E8. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al.
265 Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. Am J
266 Hum Genet 2009; 84:581-93.
267 E9. Burchard EG, Avila PC, Nazario S, Casal J, Torres A, Rodriguez-Santana JR, et al. Lower
268 bronchodilator responsiveness in Puerto Rican than in Mexican subjects with asthma. Am J
269 Respir Crit Care Med 2004; 169:386-392.
270 E10. David GL, Romieu I, Sienra-Monge JJ, Collins WJ, Ramirez-Aguilar M, del Rio-Navarro
271 BE, et al. Nicotinamide adenine dinucleotide (phosphate) reduced:quinine oxidoreductoase and
272 glutathione S-transferase M1 polymorphisms and childhood asthma. Am J Respir Crit Care Med
273 2003; 168:1199-204.
274 E11. Hancock DB, Romieu I, Shi M, Sienra-Monge JJ, Wu H, Chiu GY, et al. Genome-wide
275 association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in
276 Mexican children. PLoS Genet 2009; 5:e1000623.
277 E12. Wu H, Romieu I, Shi M, Hancock DB, Li H, Sienra-Monge JJ, et al. Evaluation of
278 candidate genes in a genome-wide association study of childhood asthma in Mexicans. J Allergy
279 Clin Immunol 2010; 125:321-7.
280 E13. National Heart, Lung and Blood Institute (NHLBI). Global Initiative for Asthma (GINA).
281 Bethesda: National Institutes of Health; 1998.
282 E14. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000
283 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447:661-78.
284 E15. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, et al.
285 Genome-wide association scan in women with systemic lupus erythematosus identifies
286 susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 2008; 40:204-10.

287 E16. Foley SC, Mogas AK, Olivenstein R, Fiset PO, Chakir J, Bourbeau J, et al. Increased
288 expression of ADAM33 and ADAM8 with disease progression in asthma. Journal of Allergy and
289 Clinical Immunology 2007; 119:863-71.
290 E17. Shannon J, Ernst P, Yamauchi Y, Olivenstein R, Lemiere C, Foley S, et al. Differences in
291 airway cytokine profile in severe asthma compared to moderate asthma. Chest 2008; 133(2):420-
292 06.
293 E18. Affymetrix. Genechip® genotyping analysis software (gtype). 2007. Available
294 from: http://www.affymetrix.com/products/software/specific/gtype.affx.
295 E19. Gibbs RA, Belmont JW, Hardenbol P, Willis TD, Yu FL, Yang HM, Ch'ang LY, Huang
296 W, Liu B, Shen Y, et al. The International HapMap project. Nature 2003; 426:789-796.
297 E20. de Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and
298 power in genetic association studies. Nature Genet 2005; 37:1217-1223.
299 E21. De La Vega FM, Isaac HI, Scafe CR. A tool for selecting SNPs for association
300 studies based on observed linkage disequilibrium patterns. Pac Symp Biocomput
301 2006: 487-498.
302 E22. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and
303 haplotype maps. Bioinformatics 2005; 21:263-5.
304 E23. KBiosciences. December 2008. Available from: http://www.kbioscience.co.uk.
305 E24. Li Y, Abecasis GR. Mach 1.0: Rapid haplotype reconstruction and missing genotype
306 inference. Am J Hum Genet 2006; S79, 2290.
307 E25. Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM. PBAT: tools for family-based
308 association studies. Am J Hum Genet 2004; 74, 367-369.
309 E26. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A
310 tool set for whole-genome association and population-based linkage analyses. American Journal
311 of Human Genetics 2007; 81:559-75.
312 E27. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, et al. Multiple loci
313 identified in a genome-wide association study of prostate cancer, Nat Genet 2008; 40:310-315.
314 E28. Yang HC, Pan CC, Lu RCY, Fann CSJ. New adjustment factors and sample size
315 calculation in a DNA-pooling experiment with preferential amplification. Genetics 2005;
316 169:399-410.
317 E29. Lake SL, Lyon H, Tantisira K, Silverman EK, Weiss ST, Laird NM, Schaid DJ. Estimation
318 and tests of haplotype-environment interaction when linkage phase is ambiguous. Human
319 Heredity 2003; 55:56-65.
320 E30. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population
321 genotype data: applications to inferring missing genotypes and haplotypic phase. Am J Hum
322 Genet 2006; 78:629-44.
323
324
325

326 **FIGURES**

327

328 Figure E1. Manhattan plot of pooled DNA samples from the Isle of Wight population. False
329 discovery rates at 1% and 5% levels are indicated.

330

331

332

333 Figure E2. Linkage disequilibrium plots of the asthma-associated region within chromosome
334 1p33-p32.31 created using Haploview with HapMap release 28 data from a) CEU, b) ASW, and
335 c) MEX populations. Shown within the blocks are D' values. Color scheme: LOD≥2 is shades of
336 pink/bright red; LOD<2 and D'=1 is blue; LOD<2 and D'=1 is white.

337

338

339

TABLES

Table E1. Allele frequencies of pooled DNA samples from Isle of Wight children for chromosome 1p33-p32.31 SNPs

SNP	Gene Symbol	Function	Chr 1 Position (1p33)[a]	Asthma case pools allele frequencies					Control pools allele frequencies				
				Pool 1	Pool 2	Pool 3	Pool 4	Pool 5	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
				(n=)22	22	22	23	23	30	30	30	37	37
rs10489769	NSUN4	intron	46579290	0.6420	0.7993	0.7197	0.7024	0.7838	0.7270	0.7063	0.7518	0.7207	0.5701
rs10489770	NSUN4	intron	46580184	0.8257	0.7575	0.7486	0.7392	0.7932	0.7171	0.7689	0.7657	0.7935	0.8664
rs952947			46775476	0.4367	0.2237	0.2047	0.1818	0.2178	0.3001	0.2314	0.1345	0.2302	0.3294
rs6429606	MKNK1	intron	46835487	0.2234	0.3420	0.1412	0.3280	0.1262	0.2157	0.2963	0.2675	0.3562	0.2460
rs2181412 *			46869234	0	0	0	0	0	0	0	0	0	0
rs1273237	ATPAF1	intron	46881838	0.0448	0.0424	0.0598	0.0602	0.0703	0.0587	0.0511	0.0544	0.05281	0.0428
rs1933932 *	ATPAF1	intron	46881915	0	0	0	0	0	0	0	0	0	0
rs631840	ATPAF1	intron	46890531	0.9200	0.9814	0.8965	0.9383	0.9190	0.9673	0.8722	0.8923	0.9619	1
rs620913	ATPAF1	intron	46890654	0.7079	0.5701	0.6649	0.5749	0.5804	0.5336	0.6313	0.5064	0.4793	0.5575
rs2289447	ATPAF1	intron	46890755	0.9687	0.8312	0.9200	0.9572	0.7297	0.7406	0.5777	0.5698	0.7613	0.6743
rs1150068	ATPAF1	intron	46891505	0.8052	0.7078	0.7063	0.7954	0.7390	0.6758	0.5765	0.6872	0.5871	0.6484
rs1048380	KIAA0494	3' UTR	46915125	0.7311	0.6892	0.8251	0.8247	0.7328	0.5972	0.6946	0.6976	0.6471	0.4893
rs2275380	KIAA0494	intron	46920315	0.6085	0.3971	0.3791	0.3331	0.4137	0.4742	0.5673	0.4467	0.5134	0.6321
rs1150064	KIAA0494	intron	46920631	0.2289	0.2785	0.2795	0.2225	0.3083	0.3510	0.3950	0.3940	0.3588	0.3820
rs1440487	KIAA0494	intron	46939662	0.5558	0.8685	0.7766	0.7421	0.7529	0.7595	0.6417	0.8989	0.6548	0.7247
rs1440486	KIAA0494	intron	46939825	0.0978	0.1980	0.1797	0.1516	0.2480	0.2223	0.4916	0.3963	0.1678	0.4513
rs720413			46992410	0.4851	0.5188	0.4154	0.4916	0.4403	0.3361	0.5404	0.3456	0.3830	0.4244
rs10493124 *			47002322	0	0	0	0	0	0	0	0	0	0
rs2405335	CYP4B1	intron	47044772	0.2694	0.2939	0.2428	0.1290	0.4212	0.2575	0.2504	0.2810	0.2048	0.2174
rs10493125			47098257	0.2501	0.0897	0.1155	0.1670	0.1197	0.0618	0.1615	0.1982	0.2231	0.1577
rs1002378	CYP4Z1	intron	47317899	0.4859	0.4587	0.4463	0.5505	0.3939	0.4959	0.4471	0.4640	0.4226	0.4388
rs2405340	CYP4Z1	intron	47322106	0.5046	0.5179	0.4573	0.6875	0.5523	0.2817	0.3415	0.5722	0.5372	0.5261
rs1343294	CYP4A22	intron	47377339	0.8217	0.7462	0.8115	0.8748	0.5377	0.6459	0.7217	0.7635	0.8552	0.8008

[a]Position information from HapMap Data Rel 28 PhaseII+III, August 10, on NCBI B36 assembly, dbSNP b126

*SNPs monomorphic in HapMap CEU population; n = number of DNA samples per pool

Table E2. Alleles, frequencies, and odds ratios in the Isle of Wight

	Isle of Wight[a]					
SNP	**Minor Allele**	**Control MAF**	**Asthma MAF**	**Odds Ratio**	**Lower CI**	**Upper CI**
rs1258000	G	0.296	0.215	0.520	0.330	0.820
rs2289447	T	0.255	0.168	0.470	0.290	0.780
rs620431	A	0.283	0.179	0.490	0.310	0.780
rs1150068	C	0.269	0.173	0.450	0.270	0.730
rs1048380	T	0.265	0.168	0.440	0.270	0.730
rs2275380	G	0.509	0.421	0.780	0.550	1.110
rs1150064	T	0.269	0.173	0.460	0.280	0.750
rs1440487	T	0.233	0.243	1.320	0.880	1.980
rs1440486	A	0.265	0.173	0.460	0.280	0.750
rs2218189	G	0.269	0.175	0.460	0.280	0.750
rs6670495	A	0.250	0.154	0.490	0.290	0.810

[a]Data are from individually genotyped samples
MAF=Minor allele frequency
CI=Confidence interval

Table E3. Genetic associations for asthma in the Isle of Wight plus consortia control populations

			Primary population		**Primary plus consortia control populations**	
			IOW pooled GWAS	**Individual IOW samples**	**IOW individual + SLEGEN**	**IOW individual + WTCCC**
		Genotyping platform	**Affymetrix 100K GeneChip array**	**Custom Illumina Goldengate**	**Illumina HumanHap300**	**Affymetrix 500K GeneChip array**
		Model	**Allelic**	**Additive**	**Additive**	**Additive**
		Software	**LatteThunder**	**SNPGWA**	**SNPGWA**	**SNPGWA**
		Statistical test	**Z^2 p-value[a]**	**Cochran Armitage p-value**	**Cochran-Armitage p-value[b]**	**Cochran-Armitage p-value[b]**
SNP	**SLEGEN MAF**	**WTCCC MAF**				
rs1258000	0.296	0.296	-	0.0282	0.0024 (0.0024)	(0.2090)
rs2289447	0.255	0.255	2.20×10^{-8}	0.0156	(0.0001)	(0.0070)
rs620431	0.283	0.281	-	0.0091	(0.0004)	(0.0104)
rs1150068	0.269	0.270	0.0034	0.0065	(0.0002)	(0.0068)
rs1048380	0.265	0.267	0.0006	0.0044	(0.0001)	(0.0067)
rs2275380	0.509	0.492	0.0124	0.0612	0.0171 (0.01708)	0.0587 (0.0573)
rs1150064	0.269	0.267	0.0053	0.0084	(8.79×10^{-5})	(0.0066)
rs1440487	0.233	0.237	0.4309	0.7177	0.2167 (0.1878)	0.7623 (0.7025)
rs1440486	0.265	0.265	2.26×10^{-5}	0.0095	(0.0002)	(0.0105)
rs2218189	0.269	0.270	-	0.0060	0.0001 (0.0001)	(0.0077)
rs6670495	0.250	0.252	-	0.0060	(7.93×10^{-6})	(0.0009)

[a]False discovery rate cut-off for α=0.05 is Z^2 p-value=2.27×10^{-5}

[b]p-values from imputed data indicated in parentheses

Table E4. Results of sliding-window haplotype association in CAMP and CARE family studies

Race[a]	**1-marker**		**2-marker**			**3-marker**			
	SNP	**p-value**	**SNP 1**	**SNP 2**	**p-value**	**SNP 1**	**SNP 2**	**SNP 3**	**p-value**
Caucasian	rs1933932	0.08326	rs1933932	rs2486161	0.04505	rs1933932	rs2486161	rs12087698	0.06802
	rs12087698	0.3173	rs12087698	rs629412	0.04214	rs12087698	rs629412	rs11211334	0.05825
	rs629412	0.03389	rs629412	rs11211334	0.04265	rs629412	rs11211334	rs631368	0.07071
	rs682000	1	rs682000	rs654509	0.01141	rs682000	rs654509	rs12094663	0.3342
	rs654509	0.01141	rs654509	rs12094663	0.3342	rs654509	rs12094663	rs1025806	0.5159
African American	rs12048954	0.04953	rs12048954	rs7412469	0.1993	rs7412469	rs2275380	rs1890473	0.2376
	rs7412469	0.4458	rs7412469	rs2275380	0.2712	rs2275380	rs1890473	rs6665021	0.03613
	rs3766217	0.4458	rs3766217	rs6671124	0.02204	rs6671124	rs1440487	rs12026027	0.02562
	rs6671124	0.006485	rs6671124	rs1440487	0.02482	rs1440487	rs12026027	rs11801744	0.3445
	rs11582403	0.2568	rs11582403	rs614078	0.03207	rs614078	rs720413	rs6662321	0.5853
Other ethnicity	rs11582403	0.03251	rs11582403	rs614078	0.08669	rs11582403	rs614078	rs720413	0.09493
	rs614078	0.8575	rs614078	rs720413	0.3376	rs614078	rs720413	rs6662321	0.01219
	rs720413	0.715	rs720413	rs6662321	0.01195	rs720413	rs6662321	rs17102513	0.02249
	rs6662321	0.0009111	rs6662321	rs17102513	0.0008954	rs6662321	rs17102513	rs11211355	0.0008954

[a]No significant results were found in the Hispanic families

75 SNPs through the region were tested; only SNPs with an asthma-associated haplotype are shown

Table E5. Asthma severity in the Isle of Wight population

Asthma Severity Classifications[a]	**Severity Scores**	**Number (n) of individuals**
No treatment	0	9
Mild intermittent: bronchodilator only	1	28
Mild persistent: one regular prophylactic medication	2	41
Moderate persistent: multiple prophylactic medications and/or multiple steroid courses	3	23
Severe: multiple asthma emergency attendance or admissions	4	11

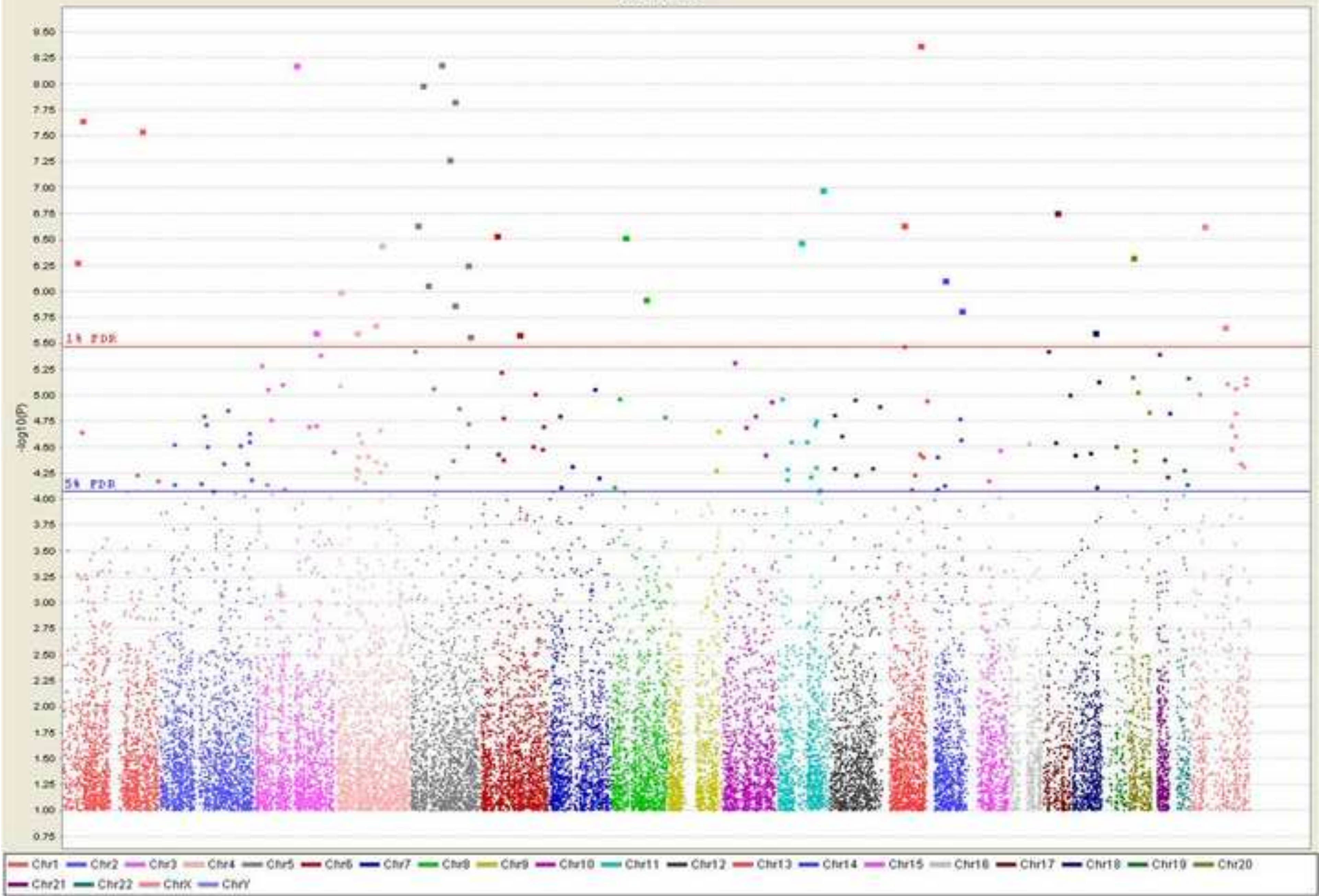
[a]Classifications based on the Global Initiative for Asthma (GINA) 2007 report

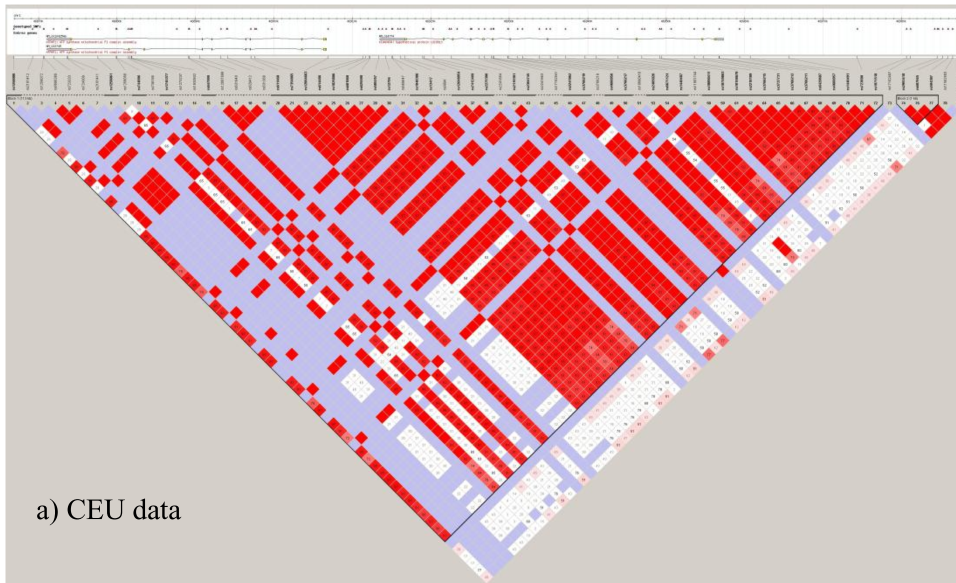
Table E6. Alleles and frequencies in the replication populations

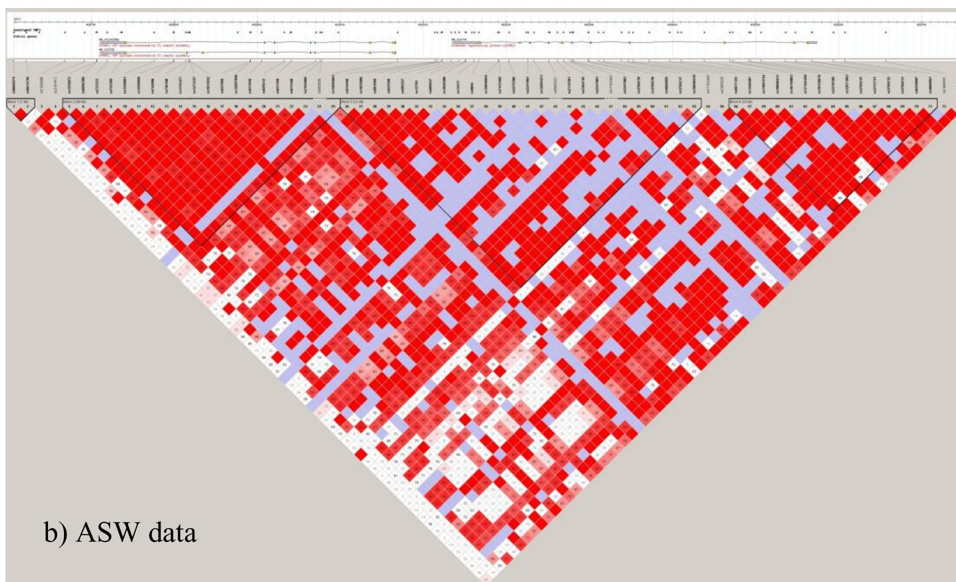
	Wessex		**CAMP + CARE**								**GALA**[a]			**Mexico Childhood Asthma Study**	
	Caucasian		**Caucasian**		**African American**		**Hispanic**		**Other ethnicity**			**PR**	**Mexican**	**Mexican**	
SNP	**Minor Allele**	**MAF**	**Minor allele**	**Case MAF**	**Minor allele**	**Case MAF**	**Minor allele**	**Case MAF**	**Minor allele**	**Case MAF**	**Minor Allele**	**MAF**	**MAF**	**Minor Allele**	**MAF**
rs1258000	G	0.285												G	0.427
rs2289447	T	0.244									T	(0.269)	(0.291)		
rs620431	A	0.257													
rs1150068	C	0.259									C	(0.484)	(0.498)		
rs629412			G	0.010	G	0.345	G	0.063	G	0.163					
rs654509			A	0.008	A	0.024	A	0.006	A	0.028					
rs601060			G	0.213	G	0.193	G	0.121	G	0.152	G	0.161	0.129		
rs1048380	T	0.253									G	(0.326)	(0.357)		
rs12048954			C	0.538	C	0.188	C	0.495	C	0.386	T	0.341	0.363		
rs2275380	A	0.487	A	0.544	A	0.244	A	0.500	A	0.413	A	0.346	0.373	A	0.428
rs1150064	T	0.258									T	(0.474)	(0.474)		
rs6665021			G	0.002	G	0.012	G	0	G	0.022	G	(0.002)	(0.002)		
rs4660956			T	0.218	T	0.052	T	0.121	T	0.098	T	0.142	0.118		
rs6671124			C	0.243	C	0.568	C	0.379	C	0.446	C	0.489	0.499		
rs1440487	T	0.253	T	0.213	T	0.184	T	0.121	T	0.141	T	0.158	0.127	T	0.080
rs1440486	A	0.251									A	(0.323)	(0.359)		
rs10749863			G	0.211	G	0.170	G	0.121	G	0.141	G	0.158	0.127		
rs2218189	G	0.257									A	(0.419)	(0.440)	G	0.428
rs6670495	A	0.24									A	(0.458)	(0.467)		
rs11582403			G	0.084	G	0.227	G	0.084	G	0.109	G	0.121	0.099		
rs6662321			T	0.056	T	0.301	T	0.053	T	0.076	T	(0.101)	(0.088)		

[a]Data in parentheses were imputed
Minor alleles reported are per NCBI Entrez SNP convention

Repository - Unmarked E Figure No. 1




a) CEU data


b) ASW data

c) MEX data