Published in final edited form as:

Genet Epidemiol. 2010 January ; 34(1): 78–91. doi:10.1002/gepi.20437.

Correcting “winner's curse” in odds ratios from genomewide association findings for major complex human diseases

Hua Zhong^{1,2} and Ross L. Prentice³

¹ Rosetta Inpharmatics, LLC, a Merck & Co., Inc., 401 Terry Avenue N., Seattle, WA 98109, USA

³ Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

Abstract

Genome-wide association studies (GWAS) provide an important approach for identifying common genetic variants that predispose to human disease. However, odds ratio (OR) estimates for the reported findings from GWAS discovery data are typically affected by a bias away from the null sometimes referred to the “winner's curse”. Also standard confidence intervals (CIs) may have far from the desired coverage rates. We applied a bias reduction method to GWAS findings from several major complex human diseases, including breast cancer, colorectal cancer, lung cancer, prostate cancer, type I diabetes and type II diabetes. We found the simple bias correction procedure allows one to estimate bias-adjusted ORs that have substantial consistency with ORs from subsequent replication studies; and that corresponding selection-adjusted CIs appear to help quantify the uncertainty of the findings. Selection-adjusted ORs and CIs can provide a reliable summary of GWAS data, and can help to choose single nucleotide polymorphisms (SNPs) for subsequent validation studies.

Keywords

bias correction; odds ratio; confidence interval; genomewide association study; complex human disease

Introduction

Genome-wide association studies (GWAS) provide a powerful method for identifying disease susceptibility loci for common diseases, offering the promise of novel targets for therapeutic intervention that act on the root cause of disease [Risch and Merikangas, 1996]. Within the last few years, many GWAS have been conducted and produced findings that generally support the hypothesis that the heritable risk of several major complex human diseases is attributable in part by common, low-risk variants [Easton et al., 2007; The Wellcome Trust Case Control Consortium (WTCCC), 2007; Hunter et al., 2007; COGENT Study, 2008; Eeles et al., 2008; Zeggini et al., 2008; Amos et al., 2008]. A typical GWAS calls for the use of high-throughput platforms to genotype a very large number of single nucleotide polymorphism (SNP) markers located throughout the human genome, in a set of cases and controls. Statistical tools are applied that compare the frequencies of SNP alleles between diseased and non-diseased individuals in a study cohort. Typically, only a very

² Correspondence regarding this manuscript should be addressed to: Dr. Hua Zhong, Department of Genetics, Rosetta Inpharmatics, LLC, 401 Terry Ave North, Seattle, WA 98109, zhh@u.washington.edu, T: 206-734-8886.

Web Resources: The R program used for the method is available from <http://students.washington.edu/zhh/>.

small fraction of SNPs are plausibly related to disease risk. Odds ratios (ORs) may be reported, to display the association strengths for SNPs selected for reporting. A significant SNP-disease risk association may be declared following a single stage study, in which case corresponding reliable OR estimates and confidence intervals (CIs) are needed to assess the magnitude of association and to decide upon next research steps. These could, for example, involve an intensive study of close-by SNPs or neighboring genes. Such a declaration may also take place following an early, intermediate, or late stage of a multistage design. Even the design of the next stage of a multistage design could be influenced by OR estimation, for example, if the sample size calculation for a subsequent stage is based on an overestimated OR, replication stages are likely to be underpowered and more likely to fail [Zhong and Prentice, 2008].

The selection, for further evaluation or reporting, of only SNPs that meet a statistical significance criterion affects the probability density of the ORs for these selected SNPs, thus potentially causing bias in the OR estimates. This is an example of the “regression to the mean” or “winner's curse” effect [Capen et al., 1971]. The magnitude of bias depends on various factors, including the power of the study. The effect sizes of individual variants, the need for stringent thresholds for establishing statistical significance, and financial constraints on numbers of variants that can be followed up inevitably constrain study power, which may result in profound bias in GWAS findings [COGENT Study, 2008]. Therefore, correcting the bias that attends standard OR estimators is particularly relevant in this context. The standard unadjusted CIs centering around the uncorrected point estimate may have far from the desired coverage rate for the selected ORs. This is a principal reason that many reported associations with common variants have not successfully replicated, and that the estimated effects from replication studies tend to be smaller than those reported in the initial discovery phase.

To address the over-estimation and inaccurate CIs, several methods have been proposed [Garner, 2007; Sun and Bull, 2005; Zollner and Pritchard, 2007; Yu et al., 2007; Zhong and Prentice, 2008]. We recently developed methods to correct the winner's curse and adjust CIs for ORs from GWAS, where SNPs are selected based on extreme P-values, either by a pre-specified threshold or rank-based selection [Zhong and Prentice, 2008]. The correction procedures only use the reported OR estimates and CIs, and do not require the raw individual-level genotype data, and hence can be conveniently applied to published findings from GWAS that have clearly specified selection criteria. In this paper, we applied conditional maximum likelihood estimation and the quantile based CI procedures [Zhong and Prentice, 2008] to estimate bias-corrected ORs and selection-adjusted CIs from several published GWAS on complex human diseases. These studies include GWAS on breast cancer [Easton et al., 2007; Hunter et al., 2007], lung cancer [Amos et al., 2008], prostate cancer [Eeles et al., 2008], colorectal cancer [COGENT Study, 2008], type I diabetes [The WTCCC, 2007] and type II diabetes [Zeggini et al., 2008]. Table 1 describes the design, sample size, genotyped SNP number, and SNP selection criterion for these studies, and the corresponding p-value cutoff choice used in the correction procedure. These diseases were selected based on the availability of one or more high dimensional GWAS with SNP selection criteria that were specified or can be inferred, and the availability of replication study results. We then compared the uncorrected OR estimates, the bias-adjusted estimates, and the replication estimates, and their corresponding CIs. A likelihood ratio based test was used to statistically compare the bias-adjusted estimates and the replication estimates. If the test failed to detect statistical significant difference between them, a combined estimated was given using the bias-adjusted estimate, a procedure that makes appropriate use of the data from both the discovery and the replication stages. We found that the simple bias correction procedure generally produces OR point estimates that are closer to estimates from

independent replication studies. Furthermore, the selection adjusted CIs derived from the initial GWAS appears to be useful for predicting whether SNPs will successfully replicate.

Methods

We recently developed several bias correction methods where SNPs are selected based on extreme P-values from the trend test [Zhong and Prentice, 2008]. Here the per allele bias-adjusted OR estimates were calculated using conditional maximum likelihood and corresponding CIs were calculated using a quantile-based method [Zhong and Prentice, 2008].

Some studies use a χ^2_2 test that allows different OR estimates for the heterozygote and homozygote for the minor allele. Some refinement of our published methods for bias correction of per allele odds ratios [Zhong and Prentice, 2008] is needed to simultaneously bias-correct odds ratios for persons who are heterozygous and homozygous for the minor SNP allele. To do so, the same conditional likelihood approach can be applied, as was shown in a 2008 Department of Biostatistics doctoral dissertation by Hua Zhong. A brief description of these bias-corrected odds ratio estimates and confidence interval methods is given in the appendix. The likelihood ratio test to compare the bias-corrected estimate to the replication study estimate is given also in the appendix.

Results

Lung Cancer GWAS

Recent GWAS have demonstrated that 15q25.1 variation influences lung cancer risk [Amos et al., 2008; Hung et al., 2008; Thorgeirsson et al., 2008]. For example, Amos et al. [2008] conducted a multistage GWAS, where they analyzed 315,450 tagging SNPs in 1,154 current and former (ever) smoking cases of European ancestry and 1,137 frequency-matched, ever-smoking controls from Houston in the discovery phase. They evaluated the ten most highly significant SNPs in an additional 2,724 cases and 3,694 controls (711 cases and 632 controls from Texas and 2,013 cases and 3,062 controls from the UK) [Amos et al., 2008]. In applying our selection-adjusted procedure we defined the correction cutoff as the largest p-value among those for the ten selected SNPs, 4.5×10^{-5} . As shown in table 2, all corrected ORs imply a weaker effect than do the uncorrected ORs. None of the bias-adjusted ORs have a statistically significant difference from the replication OR estimates. Of the 10 SNPs, 8 bias-adjusted ORs are closer to the replication estimates, the exceptions being rs8034191 and rs1051730, which showed very extreme p-values in the replication stage. Among the 10 SNPs, 8 failed to show significance in the replication cohort. Interestingly, their selection-adjusted CIs all include 1, which is consistent with the findings from the subsequent replication studies.

Colorectal Cancer GWAS

Recent GWAS have identified colorectal cancer susceptibility loci mapping to 8q23.3 [Tomlinson et al., 2008], 8q24 [Zanke et al., 2007], 10p14 [Tomlinson et al., 2008], 11q23 [Tenesa et al., 2008], 15q13 [Jaeger et al., 2008] and 18q21 [Tenesa et al., 2008; Broderick et al., 2007]. The COGENT team recently conducted a meta-analysis of data from two cohorts (London and Edinburgh) [COGENT Study, 2008]. Of the 23 SNPs associated with colorectal cancer risk at $P < 10^{-5}$, 14 map to regions that have been previously replicated. They therefore followed up the nine remaining SNPs using eight independent case control series [COGENT Study, 2008]. The bias-adjusted results for the nine SNPs are shown in table 3. Again, the bias-adjusted OR estimates and CI have good consistency with the replication based estimates and CIs. All corrected ORs have weaker effect than the

uncorrected ORs. Except rs4951039, the remaining bias-adjusted ORs are not statistically different from the replication OR estimates. Of the 9 SNPs, 8 bias-adjusted ORs are closer to the replication estimates. The selection adjusted CIs of 7 SNPs are consistent with those from the replication stage.

Prostate Cancer GWAS

Two GWAS demonstrated 8q24 variation to be associated with prostate cancer risk [Freedman et al., 2006; Amundadottir et al., 2006; Yeager et al., 2007]. Eeles et al. [2008] conducted a GWAS to identify common alleles associated with prostate cancer. In the discovery stage, they studied 569,243 SNPs on 1,854 prostate cancer cases and 1,894 controls recruited through national studies in UK. Of the 53 SNPs significant at the $P < 10^{-6}$ level, 20 SNPs were on 8q24, and six were on chromosome 17, regions previously shown to harbor loci associated with prostate cancer susceptibility [Eeles et al., 2008]. The remaining 27 SNPs were located in eight genomic regions. They evaluated 11 SNPs in a replication stage comprising 3,268 prostate cancer cases and 3,366 controls from studies in UK and Australia [Eeles et al., 2008]. After bias correction, we find that although the corrected OR estimates are smaller than the uncorrected OR, most of them are still larger than the OR estimates from the replication study. Of the 11 SNPs, 10 selection-adjusted CIs are consistent with those from the replication estimates (Table 4). The difference observed in the point estimates could be explained by the enriched nature of the cases and controls in stage 1. In the initial study, cases were selected as being were diagnosed through clinical symptoms rather than through routine screening by prostate-specific antigen (PSA) in order to maximize the proportion of cases that cause morbidity and mortality. Also these cases are genetically 'enriched' by including men diagnosed by age 60 years or having a family history of prostate cancer, as such individuals are thought to be more likely to carry susceptibility alleles, thereby increasing statistical power [Eeles et al., 2008]. These enriched genetic cases may have comparatively stronger genetic risk compared to controls, leading to higher ORs than in the replication study.

Breast Cancer GWAS

Recent GWAS studies have identified several new risk alleles for breast cancer [Easton et al., 2007; Hunter et al., 2007; Stacey et al., 2008; Gold et al., 2008]. Easton et al. [2007] conducted a two-stage GWAS in 4,398 breast cancer cases and 4,316 controls. They genotyped 227,876 SNPs in the first stage and selected 12,711 SNPs, approximately 5% of those typed in stage 1, to genotype in the second stage containing 3,990 invasive breast cancer cases and 3,916 controls. In the replication stage, they tested 30 SNPs that are either significant from the combined stage 1 and 2 data or are identified as candidate SNPs from other sources in 21,860 cases and 22,578 controls from 22 cohorts [Easton et al., 2007]. Nine SNPs showed associations in replication study with $P < 0.05$ with effects in the same direction as in stages 1 and 2. Under the assumption that the vast majority of 227,876 SNPs are from the null distribution, the 227,876 p-values of the difference in genotype frequency between cases and controls follow a distribution that similar to a uniform distribution between 0 and 1. Therefore, selecting the top 5% SNPs among all the SNPs is approximately equivalent to selecting the SNPs with $P < 0.05$. So the p-value cutoff point for stage 1 is set to 0.05. At stage 2, we approximately set the p-value cutoff point as $30/12,711 = 2.3 \times 10^{-3}$. Using the above two p-value cutoffs, we applied the bias correction method to the 24 SNPs (Table 5). The OR estimates from the replication study were less than the OR estimates from the initial discovery scan, showing some selection bias in the unadjusted estimates. The bias-adjusted OR estimates were weaker than the unadjusted estimates, therefore closer to the replication OR estimates. The selection-adjusted CIs were wider than the unadjusted CIs, and more consistent with those from the replication studies. All the selection adjusted CIs exclude 1, which is consistent with the replication studies. Again, the difference observed

between the bias-corrected ORs and the replication study ORs could possibly be explained by the enriched nature of the cases in stage 1. In stage 1, cases were women through clinical genetics centers diagnosed with invasive breast cancer under the age of 60 years who had a family history score of at least 2, where the score was computed as the total number of first-degree relatives plus half the number of second-degree relatives affected with breast cancer [Easton et al., 2007].

As part of the National Cancer Institute Cancer Genetic Markers of Susceptibility (CGEMS) Project, Hunter et al. [2007] conducted a GWAS of breast cancer by genotyping 528,173 SNPs in 1,145 postmenopausal women of European ancestry with invasive breast cancer and 1,142 controls. The GWAS identified several genomic locations as potentially associated with breast cancer. Of 528,173 SNPs tested, they reported six having the most significant P values [Hunter et al., 2007]. They attempted to replicate the initial associations in the GWAS in an additional independent 1,776 affected individuals and 2,072 controls. We performed the bias adjustment using cutoff 1.5×10^{-5} , the largest observed p-value among the top six SNPs (Table 5). Rs1219648 and rs2420946 were successfully replicated and the bias-corrected ORs are very close to those from the replication study. Rs10510126 failed to show significance in the replication study. Although the corrected OR estimate is smaller than the uncorrected OR, it still overestimates the OR compared to the replication study. All the three SNPs selected from the χ^2 test failed to replicate in the follow-up studies. The corrected OR estimates are smaller than the uncorrected ORs for all the three SNPs. The selection adjusted CIs include 1 for both the heterozygote and the homozygote for all three SNPs. These CIs are consistent with the CIs from the replication studies.

To identify further loci at which common variants are associated with breast cancer risk, a further 925 SNPs that showed evidence for association in the first two stages of Easton's cohort (combined $P_{\text{trend}} < 0.014$) were attempted to genotype in a third stage, 3,878 cases and 3,928 controls from three studies corresponding to stage 2 of CGEMS [Ahmed et al. 2009]. After combination of these data with the original GWAS data, three SNPs had P values $< 10^{-5}$. These SNPs were then evaluated on 36,141 controls and 33,134 cases. Strong evidence was found for two SNPs, rs4973768 and rs6504950. We performed the bias adjustment using p-value cutoffs of 0.014 and 10^{-5} . For both SNPs, the bias-corrected estimates show good concordance with the replication based estimates (Table 5).

Type I Diabetes GWAS

The WTCCC [2007] identified six chromosomal regions: 12q24, 12q13, 16p13, 18p11, 12p13 and 4q27 associated with T1D at $P < 5 \times 10^{-7}$. Todd et al. [2007] genotyped 11 SNPs that had shown association with $P < 1.64 \times 10^{-5}$ from 11 chromosome regions not previously associated with T1D. They genotyped samples from 4,000 affected individuals and 5,000 controls that were independent of the WTCCC study [Todd et al. 2007]. We performed the bias correction methods on the 12 selected SNPs imposing the selection criteria $P < 1.64 \times 10^{-5}$ (Table 6). Six SNPs showed evidence for associations in the follow up study with $P < 1.82 \times 10^{-6}$ and effects in the same direction as in WTCCC. Their corrected OR estimates are similar or slightly less than the uncorrected OR estimates. The selection adjusted CIs for these SNPs are similar to the unadjusted CIs. Three SNPs showed evidence for associations in the follow-up study with p-values between 0.02 and 0.05. Their corrected OR estimates are less than the uncorrected OR estimates. The selection adjusted CIs are wider than the unadjusted CIs. However, the selection adjusted CIs for these SNPs exclude 1 as well. The corrected point estimates and CIs are more consistent with those from the replication studies. Three SNPs failed to show significance in the follow up study, though they had p-values from 4.24×10^{-6} to 1.64×10^{-5} in the GWAS. Their corrected OR estimates are much smaller than the uncorrected OR and closer to the replication OR

estimates. The selection adjusted CIs for these SNPs are much wider than the unadjusted CIs and all include 1.

Type II Diabetes GWAS

GWAS have identified multiple loci that modestly, but reproducibly, influence risk of T2D [Diabetes Genetics Initiative, 2007; The WTCCC, 2007; Scott et al., 2007]. A meta-analysis has been recently conducted on data from three independent [Diabetes Genetics Initiative, 2007; The WTCCC, 2007; Scott et al. 2007] cohorts [Zeggini et al., 2008]. They selected 69 SNPs for replication in stage 2, which modeled 22426 additional samples, based on the statistical significance. Of these SNPs, 11 showed $P < 0.005$ in stage 2 alone and $P < 10^{-5}$ from the combined stage 1 and stage 2 data. Then they further genotyped these 11 SNPs in a replication stage that evaluated 57,366 additional samples. We approximated the p-value cutoff point by $69/2, 202, 892 = 3.1 \times 10^{-5}$ at stage 1; and at 10^{-5} at stage II. The bias-adjusted results for the 11 SNPs are shown in table 7. Again, the bias-adjusted OR estimates and CIs have good consistency with the replication based estimates and CIs.

Discussion

In this paper, we applied bias-adjusted point estimates and confidence interval procedures [Zhong and Prentice, 2008] to published GWAS findings. It is widely recognized that estimates of the genetic effect based on high-dimensional association studies tend to be upwardly biased [Garner, 2007; Zollner and Pritchard, 2007; Yu et al., 2007]. However as demonstrated by several examples here and by extensive simulation studies [Zhong and Prentice, 2008], a simple bias reduction procedure, with careful selection of correction p-value cutoff, allows one to estimate bias-adjusted ORs that have fairly good consistency with the replication based OR estimates. More importantly, the selection adjusted CIs help quantify the uncertainty of the findings from the discovery stage, which provide insights as to whether the significant findings will replicate in independent cohorts. Therefore, this simple bias correction procedure evidently enables one to obtain more reliable OR estimates and CIs from the discovery stage of a GWAS, before embarking on expensive replication studies. When there is not a clearly pre-specified p-value selection threshold, the cutoff can be usefully approximated by the maximum p-value among all the selected SNPs or by the ratio of the number of selected SNPs to the total number of SNPs as in some of our applications. With either specified or approximated thresholds, the bias-corrected estimates show reasonable concordance with the subsequent replication-based estimates.

Some bias-adjusted estimates were still more extreme than the estimates from the independent replication studies. In those instances, the selection adjusted CIs typically excluded the null hypotheses value when the SNPs failed to replicate. Residual difference between bias-corrected ORs and replication ORs may be due to the ascertainment differences between the initial sample and the replication sample or to the approximation we chose for the p-value selection used to choose SNPs for further evaluation, in some of the studies examined. For example, the initial sample might be genetically 'enriched' by including cases with a family history of the study disease to increase statistical power. These enriched genetic cases may have a comparatively more extreme ORs compared to those from populations used for subsequent replication studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work was partially supported by grant CA 53996 and contract HHSN268200764314C from the National Institute of Health.

References

- Ahmed S, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet.* 2009; 41:585–590. [PubMed: 19330027]
- Amos CI, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet.* 2008; 40:616–622. [PubMed: 18385676]
- Amundadottir LT, et al. A common variant associated with prostate cancer in European and African populations. *Nat Genet.* 2006; 38:652–658. [PubMed: 16682969]
- Broderick P, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat Genet.* 2007; 39:1315–1317. [PubMed: 17934461]
- Capen EC, Clapp RV, Campbell WM. Competitive bidding in high-risk situations. *Journal of Petroleum Technology.* 1971; 23:641–653.
- COGENT Study. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet.* 2008; 40:1426–1435. [PubMed: 19011631]
- Diabetes Genetics Initiative. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science.* 2007; 316:1331–1336. [PubMed: 17463246]
- Easton DF, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature.* 2007; 447:1087–1093. [PubMed: 17529967]
- Eeles RA, et al. Multiple newly identified loci associated with prostate cancer susceptibility. *Nat Genet.* 2008; 40:316–321. [PubMed: 18264097]
- Freedman ML, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proc Natl Acad Sci USA.* 2006; 103:14068–14073. [PubMed: 16945910]
- Garner C. Upward bias in odds ratio estimates from genome-wide association studies. *Genet Epidemiol.* 2007; 31:288–295. [PubMed: 17266119]
- Gold B, et al. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc Natl Acad Sci USA.* 2008; 105:4340–4345. [PubMed: 18326623]
- Hung RJ, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.* 2008; 452:633–637. [PubMed: 18385738]
- Hunter DJ, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet.* 2007; 39:870–874. [PubMed: 17529973]
- Jaeger E, et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat Genet.* 2008; 40:26–28. [PubMed: 18084292]
- Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science.* 1996; 273:1516–1517. [PubMed: 8801636]
- Scott LJ, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science.* 2007; 316:1341–1345. [PubMed: 17463248]
- Stacey SN, et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet.* 2008; 40:703–706. [PubMed: 18438407]
- Sun L, Bull S. Reduction of selection bias in genomewide studies by resampling. *Genet Epidemiol.* 2005; 28:352–367. [PubMed: 15761913]
- Tenesa A, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet.* 2008; 40:631–637. [PubMed: 18372901]
- The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; 447:661–678. [PubMed: 17554300]
- Thorgerirsson TE, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.* 2008; 452:638–642. [PubMed: 18385739]

- Todd JA, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet.* 2007; 39:857–864. [PubMed: 17554260]
- Tomlinson IP, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet.* 2008; 40:623–630. [PubMed: 18372905]
- Yeager M, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet.* 2007; 39:645–649. [PubMed: 17401363]
- Yu K, Chatterjee N, et al. Flexible Design for Following up Positive Findings. *Am J Hum Genet.* 2007; 81:540–551. [PubMed: 17701899]
- Zanke BW, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet.* 2007; 39:989–994. [PubMed: 17618283]
- Zeggini E, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet.* 2008; 40:638–645. [PubMed: 18372903]
- Zhong H, Prentice RL. Bias-reduced estimators and confidence intervals for odds ratios in genome-wide association studies. *Biostatistics.* 2008; 9:621–634. [PubMed: 18310059]
- Zollner S, Pritchard J. Overcoming the winner's curse: estimating penetrance parameters from case-control data. *Am J Hum Genet.* 2007; 80:605–615. [PubMed: 17357068]

Overview of Study Design

Table 1

Disease	Type	Study	Design	Ncase/Ncontrol	N _{SNP}	GWAS Selection Criteria	Correction Cutoff
Lung Cancer	Discovery	UTMD	One Stage	1,154/1,137	315,450	Top ranked 10 ¹	$p^I = 4.5 \times 10^{-5}$
	Replication	Texas&UK		2,724/3,694	10		
Colorectal Cancer	Meta-analysis	London&Edinburgh	One Stage	4878/4914	38,710	$p^I < 5.0 \times 10^{-5}$	$p^I = 5.0 \times 10^{-5}$
	Replication	4 cohorts		13,114/14,304	7		
T1D	GWAS	WTCCC	One Stage	1,924/2,938	393,143	$p^I < 10^{-5}$	$p^I = 1.64 \times 10^{-5}$
	Replication	Todd		4,000/5,000	11		
Prostate Cancer	GWAS	Eeles	One Stage	1,854/1,894	541,129	$p^I < 10^{-6}$	$p^I = 10^{-6}$
	Replication	UK&Australia		3,268/3,366	11		
Breast Cancer	GWAS	Easton	Two Stage	Stage I: 390/364	Stage I 205,586	Top ranked 5% ¹ (N=12,711)	$p^I = 0.05$
				Stage II: 3,990/3,916	Stage II 10,405	Top ranked and candidate gene ²	
	Replication	22 cohorts		21,860/22,578	30		$p^I_{combine} = 2.3 \times 10^{-3}$
	GWAS	CGEMS	One Stage	1,145/1,142	528,173	Top ranked 6 ³	
	Replication	3 cohorts		1,776/2,072	8		
T2D	Meta-analysis	Diagram	Two Stage	Stage I: 4,549/5,579	Stage I: 2,202,892	Top ranked 69 ¹	$p^I = 3.1 \times 10^{-5}$
	Replication	10 cohorts		Stage II: 10,037/12,389	Stage II: 69	$p^I_{combine} < 10^{-5}$	
				14,157/43,209	11		$p^I_{combine} = 10^{-5}$

¹ P-values are computed from 1 degree of freedom association trend tests;

² P-values are computed from 1 degree of freedom association tests, combining stages 1 and 2;

³ P-values are either from 1 degree of freedom association tests or a score test with two degrees of freedom.

Table 2
Summary odds ratios and p-values for the SNPs showing association with Lung Cancer

rsID	Gene	Allele ^d	Chr	Position ^b	Trend p-value			Per Allele OR (95% CI)			
					Stages 1&2	Replication	Unadjusted	Adjusted	Replication	P ^c _{het}	Combined
rs2808630	CRP	C/T	1	156493941	1.6 × 10 ⁻⁵	3.9 × 10 ⁻¹	0.76 (0.67-0.86)	0.88 (0.68-1.01)	0.93 (0.79-1.10)	1.00	0.93 (0.80-1.01)
rs7626795	IL1RAP	A/G	3	191833163	1.9 × 10 ⁻⁵	4.5 × 10 ⁻¹	1.46 (1.23-1.74)	1.17 (0.98-1.70)	1.05 (0.93-1.19)	1.00	1.05 (0.99-1.19)
rs2202507	GYP A	A/C	4	145615286	8.7 × 10 ⁻⁶	5.6 × 10 ⁻¹	1.30 (1.16-1.46)	1.20 (0.99-1.45)	0.98 (0.91-1.06)	0.18	1.01 (0.98-1.07)
rs11099666	ARHGAP10	A/G	4	148991033	4.5 × 10 ⁻⁵	6.5 × 10 ⁻¹	0.65 (0.53-0.80)	0.94 (0.56-1.04)	0.97 (0.85-1.11)	0.87	0.96 (0.85-1.02)
rs1481847	MSC	C/T	8	72944049	1.2 × 10 ⁻⁵	7.2 × 10 ⁻¹	1.30 (1.15-1.46)	1.15 (0.99-1.45)	1.01 (0.94-1.10)	0.64	1.02 (0.99-1.10)
rs855974	EMX2	C/T	10	119436858	1.4 × 10 ⁻⁵	9.1 × 10 ⁻¹	0.74 (0.65-0.85)	0.83 (0.65-1.01)	1.00 (0.92-1.10)	0.37	0.98 (0.91-1.02)
rs8034191	LOC123688	C/T	15	76593078	1.9 × 10 ⁻⁵	2.2 × 10 ⁻¹⁴	1.30 (1.15-1.47)	1.05 (0.99-1.44)	1.33 (1.24-1.43)	0.13	1.31 (1.22-1.40)
rs1051730	CHR3	C/T	15	76681394	1.1 × 10 ⁻⁵	9.1 × 10 ⁻¹³	1.31 (1.16-1.48)	1.16 (1.00-1.46)	1.32 (1.22-1.43)	0.28	1.30 (1.21-1.40)
rs12956651	CBLN2	A/G	18	68265435	1.1 × 10 ⁻⁵	8.7 × 10 ⁻¹	0.59 (0.47-0.75)	0.71 (0.48-1.02)	0.99 (0.85-1.14)	0.41	0.96 (0.84-1.02)
rs6069045	DOK5	A/C	20	52949270	8.6 × 10 ⁻⁶	7.5 × 10 ⁻¹	0.75 (0.66-0.85)	0.82 (0.67-1.01)	1.01 (0.93-1.10)	0.25	0.98 (0.92-1.02)

^aMajor/minor allele;

^bFrom NCBI build 139;

^csignificance level (p-value) for testing equality of bias-adjusted and replication odds ratios

Summary odds ratios and p-values for the SNPs showing association with Colorectal Cancer

Table 3

rsID	Gene	Allele ^a	Chr	Position ^b	Trend p-value		Per Allele OR (95% CI)				
					Stages 1&2	Replication	Unadjusted	Adjusted	Replication	P ^c _{het}	Combined
rs10411210	RHPN2	C/T	19	38224140	2.0 × 10 ⁻⁷	6.9 × 10 ⁻⁴	0.79 (0.72-0.86)	0.81 (0.72-0.95)	0.90 (0.85-0.96)	0.24	0.89 (0.84-0.94)
rs961253		C/A	20	6352281	7.8 × 10 ⁻⁷	3.4 × 10 ⁻⁵	1.13 (1.08-1.19)	1.10 (1.00-1.18)	1.11 (1.06-1.17)	0.87	1.11 (1.06-1.15)
rs355527		G/A	20	6336068	7.8 × 10 ⁻⁷	3.4 × 10 ⁻⁵	1.13 (1.08-1.19)	1.10 (1.00-1.18)	1.11 (1.06-1.17)	0.87	1.11 (1.06-1.15)
rs9929218	CDH1	G/A	16	67378447	1.1 × 10 ⁻⁶	1.5 × 10 ⁻⁴	0.88 (0.84-0.93)	0.91 (0.84-1.00)	0.93 (0.90-0.97)	0.71	0.93 (0.90-0.96)
rs4444235	BMP4	T/C	14	53480669	5.6 × 10 ⁻⁶	1.8 × 10 ⁻⁴	1.12 (1.07-1.18)	1.03 (0.99-1.17)	1.10 (1.05-1.16)	0.42	1.09 (1.04-1.14)
rs1862748	CDH1	C/T	16	67390444	8.5 × 10 ⁻⁷	1.5 × 10 ⁻⁴	0.88 (0.84-0.93)	0.91 (0.84-1.00)	0.93 (0.90-0.97)	0.64	0.93 (0.90-0.96)
rs4951291		G/A	1	202273161	6.6 × 10 ⁻⁶	5.7 × 10 ⁻¹	0.85 (0.79-0.91)	0.97 (0.80-1.01)	1.02 (0.95-1.09)	0.35	0.99 (0.95-1.01)
rs7259371	RHPN2	G/A	19	38226481	3.4 × 10 ⁻⁶	2.1 × 10 ⁻³	0.86 (0.81-0.92)	0.93 (0.81-1.01)	0.91 (0.86-0.97)	0.84	0.91 (0.86-0.97)
rs4951039		A/G	1	202273220	6.6 × 10 ⁻⁶	5.2 × 10 ⁻²	0.85 (0.79-0.91)	0.97 (0.80-1.01)	1.09 (1.00-1.19)	0.03	0.99 (0.96-1.01)

^aMajor/minor allele;

^bFrom NCBI build 139;

^csignificance level (p-value) for testing equality of bias-adjusted and replication odds ratios.

Table 4
Summary odds ratios and p-values for the SNPs showing association with Prostate Cancer

rsID	Gene	Allele ^d	Chr	Position ^b	Trend p-value		Per Allele OR (95% CI)				
					Stages 1&2	Replication	Unadjusted	Adjusted	Replication	P ^c _{het}	Combined
rs2660753		C/T	3	87193364	9.5 × 10 ⁻⁸	1.8 × 10 ⁻³	1.52 (1.30-1.77)	1.30 (1.01-1.71)	1.18 (1.06-1.31)	0.69	1.19 (1.07-1.31)
rs9364554	SLC22A3	C/T	6	160753654	9.3 × 10 ⁻⁷	4.8 × 10 ⁻⁵	1.28 (1.16-1.41)	1.02 (1.01-1.31)	1.17 (1.08-1.26)	0.29	1.16 (1.07-1.25)
rs6465657	LMTK2	T/C	7	97654263	1.2 × 10 ⁻⁸	1.0 × 10 ⁻³	1.30 (1.19-1.43)	1.24 (1.02-1.41)	1.12 (1.05-1.20)	0.33	1.13 (1.06-1.21)
rs7931342		G/T	11	68751073	2.4 × 10 ⁻⁷	1.1 × 10 ⁻⁶	0.79 (0.72-0.86)	0.90 (0.74-1.01)	0.84 (0.79-0.90)	0.61	0.84 (0.79-0.90)
rs902774		C/T	12	91560171	2.0 × 10 ⁻⁷	4.9 × 10 ⁻¹	1.39 (1.23-1.57)	1.21 (0.99-1.52)	1.03 (0.94-1.14)	0.45	1.04 (0.99-1.14)
rs2659056		G/T	19	56027755	1.2 × 10 ⁻⁷	4.2 × 10 ⁻¹	1.33 (1.20-1.49)	1.10 (0.99-1.42)	0.97 (0.89-1.05)	0.25	1.01 (0.99-1.07)
rs5945619	NUDT10/11	T/C	X	51258412	2.2 × 10 ⁻⁸	8.0 × 10 ⁻⁴	1.46 (1.28-1.66)	1.39 (1.03-1.64)	1.19 (1.07-1.31)	0.33	1.25 (1.06-1.45)
rs7920517	MSMB	A/G	10	51202627	7.2 × 10 ⁻¹³	9.3 × 10 ⁻⁹	1.39 (1.27-1.53)	1.39 (1.24-1.52)	1.22 (1.14-1.31)	0.05	1.26 (1.18-1.34)
rs10993994	MSMB	C/T	10	51219502	8.0 × 10 ⁻²⁴	1.5 × 10 ⁻¹⁰	1.62 (1.47-1.78)	1.62 (1.47-1.78)	1.25 (1.17-1.34)	1.60 × 10 ⁻⁵	1.25 (1.17-1.34)
rs2735839	KLK2/3	G/A	19	56056435	2.4 × 10 ⁻²⁰	2.0 × 10 ⁻⁴	0.56 (0.50-0.64)	0.56 (0.49-0.63)	0.83 (0.75-0.91)	1.42 × 10 ⁻⁶	0.83 (0.75-0.91)
rs266849		A/G	19	56040902	1.0 × 10 ⁻¹⁶	2.3 × 10 ⁻¹	0.62 (0.55-0.69)	0.62 (0.55-0.70)	0.95 (0.87-1.03)	1.02 × 10 ⁻⁷	0.95 (0.87-1.03)

^aMajor/minor allele;

^bFrom NCBI build 139;

^csignificance level (p-value) for testing equality of bias-adjusted and replication odds ratios.

Table 5

Summary odds ratios and p-values for the SNPs showing association with Breast Cancer

rsID	Gene	Allele ^a	Chr	Position ^b	Trend p-value			Per Allele OR (95% CI)			P ^c _{het}	Combined
					Stages 1&2	Replication	Unadjusted	Adjusted	Replication			
										Breast Cancer (Easton)		
rs2981582	FGFR2	G/A	10q	123342307	7.7 × 10 ⁻¹⁵	4.0 × 10 ⁻⁶²	1.28 (1.20-1.36)	1.27 (1.19-1.35)	1.26 (1.23-1.30)	0.95	1.26 (1.23-1.30)	
rs889312	MAP3K1	A/C	5q	56067641	3.6 × 10 ⁻⁵	2.7 × 10 ⁻¹⁵	1.15 (1.07-1.23)	1.12 (1.01-1.21)	1.13 (1.10-1.16)	0.81	1.13 (1.10-1.16)	
rs12443621	TNRC9 LOC643714	A/G	16q	51105538	2.0 × 10 ⁻⁷	1.0 × 10 ⁻¹³	1.17 (1.10-1.24)	1.16 (1.08-1.23)	1.11 (1.08-1.14)	0.28	1.12 (1.09-1.15)	
rs8051542	TNRC9 LOC643714	C/T	16q	51091668	5.5 × 10 ⁻⁶	3.9 × 10 ⁻⁸	1.15 (1.08-1.23)	1.13 (1.04-1.21)	1.09 (1.06-1.13)	0.41	1.10 (1.06-1.13)	
rs2107425	H19	C/T	11p	1977651	8.9 × 10 ⁻⁶	1.4 × 10 ⁻²	0.86 (0.81-0.92)	0.88 (0.82-0.99)	0.96 (0.94-0.99)	0.08	0.96 (0.93-0.99)	
rs13281615		T/C	8q	128424800	8.1 × 10 ⁻⁸	6.2 × 10 ⁻⁷	1.18 (1.11-1.26)	1.17 (1.09-1.25)	1.08 (1.05-1.11)	0.05	1.09 (1.06-1.13)	
rs3817198		T/C	8q	128424800	2.0 × 10 ⁻⁵	1.5 × 10 ⁻⁵	1.15 (1.07-1.22)	1.12 (1.01-1.21)	1.07 (1.04-1.11)	0.37	1.08 (1.04-1.11)	
rs981782		A/C	5p	45321475	7.5 × 10 ⁻⁵	3.2 × 10 ⁻³	0.89 (0.84-0.94)	0.92 (0.85-0.99)	0.96 (0.93-0.99)	0.40	0.96 (0.93-0.98)	
rs4666451		G/A	2p	19150424	1.8 × 10 ⁻⁶	3.6 × 10 ⁻²	0.86 (0.81-0.92)	0.88 (0.82-0.95)	0.97 (0.94-0.99)	0.03	0.96 (0.93-0.99)	
Breast Cancer (CGEMS): Trend Test												
rs10510126		C/T	10q	124992475	7.1 × 10 ⁻⁷	6.4 × 10 ⁻¹	0.62 (0.51 - 0.75)	0.67 (0.52 - 0.99)	0.97 (0.84-1.11)	0.20	0.95 (0.83-1.08)	
rs1219648	FGFR2	A/G	10q	123336180	3.2 × 10 ⁻⁶	1.0 × 10 ⁻⁵	1.32 (1.17 - 1.49)	1.22 (0.99 - 1.45)	1.23 (1.12-1.35)	0.48	1.23 (1.11-1.33)	
rs2420946	FGFR2	C/T	10q	123341314	3.5 × 10 ⁻⁶	2.6 × 10 ⁻⁵	1.32 (1.17 - 1.49)	1.22 (0.99 - 1.45)	1.22 (1.11-1.34)	0.51	1.22 (1.11-1.32)	
Breast Cancer (CGEMS): 2-DF Test												
rs12505080		C/T	4p	37171906	8.1 × 10 ⁻⁶	1.4 × 10 ⁻¹	1.22 ^c (1.02-1.45)	1.09 ^c (0.96-1.37)	1.10 ^c (0.96-1.25)	1.00	1.10 ^c (0.96-1.25)	

rsID	Gene	Allele ^a	Chr	Position ^b	Trend p-value			Per Allele OR (95% CI)		
					Stages 1&2	Replication	Unadjusted	Adjusted	Replication	P ^c _{het}
rs17157903	RELN	C/T	7q	103221987	8.8 × 10 ⁻⁶	6.4 × 10 ⁻¹	0.51 ^d (0.35-0.73)	0.75 ^d (0.41-1.03)	0.85 ^d (0.65-1.12)	0.67 (0.63-1.12)
							1.60 ^c (1.31-1.95)	1.22 ^c (0.99-1.79)	1.01 ^c (0.87-1.18)	1.01 ^c (0.87-1.18)
rs7696175	TLR1/6	C/T	4p	38643552	1.5 × 10 ⁻⁶	9.6 × 10 ⁻¹	0.77 ^d (0.42-1.41)	0.90 ^d (0.50-1.35)	0.81 ^d (0.51-1.29)	0.89 (0.51-1.29)
							1.39 ^c (1.15-1.68)	1.27 ^c (1.00-1.62)	1.00 ^c (0.86-1.15)	1.00 ^c (0.86-1.15)
							0.86 ^d (0.67-1.09)	0.90 ^d (0.70-1.10)	1.02 ^d (0.85-1.23)	1.02 ^d (0.85-1.23)
Breast Cancer (Easton & CGEMS)										
rs4973768	SLC4A7, NEK10	C/T	3p24	27391017	3.7 × 10 ⁻⁶	1.4 × 10 ⁻¹⁸	1.11 (1.06-1.16)	1.05 (1.00-1.15)	1.11 (1.08-1.13)	1.11 (1.08-1.13)
rs6504950	COX11	G/A	17q23	50411470	3.4 × 10 ⁻⁶	1.0 × 10 ⁻⁴	0.89 (0.85-0.94)	0.95 (0.86-1.00)	0.95 (0.92-0.97)	0.95 (0.93-0.97)

^aMajor/minor allele;

^bFrom NCBI build 139;

^codds ratio for heterozygote;

^dodds ratio for homozygote;

^esignificance level (p-value) for testing equality of bias-adjusted and replication odds ratios.

Table 6

Summary odds ratios and p-values for the SNPs showing association with Type I Diabetes

rsID	Gene	Allele ^d	Chr	Position ^b	Trend p-value			Per Allele OR (95% CI)			P ^c _{het}	Combined
					Stages 1&2	Replication		Unadjusted	Adjusted	Replication		
rs6679677	PHTF1 PTPN22	C/A	1p13	114105331	8.03 × 10 ⁻²⁴	confirmed previously	1.89 (1.67-2.13)	1.89 (1.67-2.13)	confirmed previously			
rs3741208	INS	C/T	11p15	2126350	2.28 × 10 ⁻⁷	confirmed previously	1.25 (1.15-1.35)	1.24 (1.09-1.35)	confirmed previously			
rs2292239	ERBB3	C/A	12q13	54768447	1.49 × 10 ⁻⁹	1.89 × 10 ⁻¹⁴	1.30 (1.20-1.42)	1.30 (1.17-1.41)	1.28 (1.20-1.36)	0.82	1.29 (1.22-1.35)	
rs2542151	PTPN2	A/C	18p11	12769947	8.40 × 10 ⁻⁸	3.36 × 10 ⁻¹⁰	1.33 (1.20-1.49)	1.29 (1.05-1.47)	1.29 (1.19-1.40)	0.98	1.29 (1.20-1.38)	
rs12708716	KIAA0350	A/G	16p13	11087374	1.28 × 10 ⁻⁸	7.07 × 10 ⁻⁹	0.77 (0.70-0.84)	0.78 (0.70-0.89)	0.83 (0.78-0.89)	0.34	0.82 (0.77-0.87)	
rs17696736	C12orf30	A/G	12q24	110971201	7.27 × 10 ⁻¹⁴	1.82 × 10 ⁻⁶	1.37 (1.27-1.49)	1.37 (1.26-1.48)	1.16 (1.09-1.23)	1.00 × 10 ⁻³	1.23 (1.16-1.29)	
rs9653442	AFF3 LOC150577	A/G	2q11	100191799	4.78 × 10 ⁻⁶	0.02	1.21 (1.11-1.32)	1.02 (0.99-1.26)	1.07 (1.01-1.14)	0.65	1.07 (1.01-1.13)	
rs17388568	Tenr-IL2-IL21	G/A	4q27	123548812	6.35 × 10 ⁻⁷	0.02	1.27 (1.15-1.39)	1.22 (1.01-1.38)	1.08 (1.01-1.15)	0.24	1.09 (1.02-1.16)	
rs7722135	Q8WY63	G/A	5q14	86330425	4.24 × 10 ⁻⁶	0.05	0.79 (0.71-0.87)	0.89 (0.73-1.00)	0.92 (0.86-1.00)	0.82	0.92 (0.85-0.99)	
rs2666236	NRP1	G/A	10p11	33458878	1.05 × 10 ⁻⁵	0.13	1.21 (1.11-1.31)	1.09 (0.99-1.28)	1.05 (0.99-1.12)	0.80	1.05 (1.00-1.12)	
rs6546909	DQX1	T/A	2p13	74599830	8.53 × 10 ⁻⁶	0.71	1.31 (1.16-1.47)	1.10 (0.99-1.42)	1.02 (0.93-1.11)	0.69	1.03 (0.99-1.12)	
rs12061474	PIK3C2B	G/A	1q32	202655937	1.64 × 10 ⁻⁵	0.93	0.75 (0.65-0.85)	1.00 (0.72-1.02)	1.00 (0.91-1.10)	0.64	0.98 (0.91-1.02)	

^aMajor/minor allele;^bFrom NCBI build 139;^csignificance level (p-value) for testing equality of bias-adjusted and replication odds ratios.

Table 7
Summary odds ratios and p-values for the SNPs showing association with Type II Diabetes

rsID	Gene	Allele ^a	Chr	Position ^b	Trend p-value			Per Allele OR (95% CI)			
					Stages 1&2	Replication	Unadjusted	Adjusted	Replication	P ^c _{het}	Combined
rs864745	JAZF1	C/T	7	27953796	4.9 × 10 ⁻⁹	1.3 × 10 ⁻⁷	1.10 (1.06-1.13)	1.08 (1.03-1.12)	1.10 (1.06-1.15)	0.50	1.09 (1.06-1.12)
rs12779790	CDC123 CAMK1D	A/G	10	12368016	1.4 × 10 ⁻⁸	1.5 × 10 ⁻⁴	1.12 (1.08-1.16)	1.09 (1.04-1.15)	1.09 (1.04-1.14)	0.91	1.09 (1.05-1.13)
rs7961581	TSPAN8 LGR5	T/C	12	69949369	1.4 × 10 ⁻⁶	4.3 × 10 ⁻⁵	1.09 (1.05-1.13)	1.05 (1.00-1.11)	1.09 (1.04-1.13)	0.31	1.08 (1.04-1.11)
rs758597	THADA	C/T	2	43644474	2.3 × 10 ⁻⁸	9.2 × 10 ⁻⁵	1.17 (1.11-1.24)	1.13 (1.05-1.21)	1.12 (1.05-1.20)	0.80	1.13 (1.07-1.18)
rs4607103	ADAMTS9	T/C	3	64686944	1.2 × 10 ⁻⁷	3.5 × 10 ⁻³	1.11 (1.07-1.15)	1.08 (1.01-1.13)	1.06 (1.01-1.11)	0.65	1.07 (1.03-1.11)
rs10923931	NOTCH2	G/T	1	120230001	4.3 × 10 ⁻⁷	1.9 × 10 ⁻³	1.14 (1.08-1.2)	1.09 (1.00-1.17)	1.11 (1.05-1.18)	0.73	1.1 (1.05-1.15)
rs1153188	DCD	T/A	12	53385263	7.7 × 10 ⁻⁷	8.8 × 10 ⁻³	1.09 (1.06-1.13)	1.06 (1.00-1.11)	1.06 (1.02-1.10)	0.91	1.06 (1.02-1.09)
rs17036101	SYN2 PPARG	A/G	3	12252845	3.8 × 10 ⁻⁷	1.2 × 10 ⁻²	1.19 (1.11-1.27)	1.12 (1.00-1.23)	1.11 (1.02-1.2)	0.88	1.11 (1.04-1.19)
rs2641348	ADAM30	A/G	1	120149926	4.0 × 10 ⁻⁵	7.8 × 10 ⁻³	1.11 (1.06-1.17)	1.01 (0.99-1.01)	1.09 (1.03-1.16)	0.12	1.07 (1.01-1.12)
rs9472138	VEGFA	C/T	6	43919740	1.1 × 10 ⁻⁵	9.5 × 10 ⁻²	1.09 (1.05-1.13)	1.01 (0.99-1.1)	1.03 (1.00-1.07)	0.68	1.03 (1.00-1.06)
rs10490072	BCL11A	C/T	2	60581582	1.6 × 10 ⁻⁷	6.5 × 10 ⁻¹	1.11 (1.07-1.15)	1.08 (1.01-1.13)	1 (0.97-1.04)	0.06	1.01 (1.00-1.05)

^aMajor/minor allele;

^bFrom NCBI build 139;

^csignificance level (p-value) for testing equality of bias-adjusted and replication odds ratios