ERJ Express. Published on July 2, 2009 as doi: 10.1183/09031936.00018909

Interaction of T-cell and antigen presenting-cell co-stimulatory genes in childhood IgE

Renske W.B. Bottema[1], Dirkje S. Postma[1], Naomi E. Reijmerink[1,2], Carel Thijs[3], Foekje F. Stelma[4], Henriette A. Smit[5], Constant P. van Schayck[6], Bert Brunekreef[7], Gerard H. Koppelman[8], Marjan Kerkhof[9]

[1]Department of Pulmonology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[2]Department of Pediatrics, Beatrix Children's Hospital, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[3]Department of Epidemiology, Care and Public Health Research Institute and Nutrition and Toxicology Research Institute Maastricht, Maastricht University, Maastricht, The Netherlands
[4]Department of Medical Microbiology, University Hospital Maastricht, Maastricht, The Netherlands
[5]Institute for Public Health and the Environment, Bilthoven, The Netherlands
[6]Department of General Practice and Care and Public Health Research Institute, Maastricht University, Maastricht, The Netherlands
[7]Institute for Risk Assessment Sciences, University of Utrecht, and Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
[8]Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children's Hospital, University Medical Center Groningen, Groningen, The Netherlands
[9]Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Corresponding author (including requests for reprints)**:**
Prof. Dr. D.S. Postma
Department of Pulmonology
University Medical Center Groningen
P.O. Box 196
9700 AD Groningen
The Netherlands
Telephone: +31 50 3613532, Fax: +31 50 3619320, e-mail: d.s.postma@int.umcg.nl

Short title: co-stimulatory genes and childhood serum IgE (max 45 characters)

Word count: 2959 (max 3000)

Abstract

It is likely that multiple genes contribute to IgE production. Co-stimulatory molecules are crucial for the cross-talk between antigen presenting cells and T-lymphocytes which drives the IgE response. We evaluated gene-gene interactions of haplotype tagging polymorphisms in a pathway of 24 co-stimulatory genes in relation to serum IgE levels. We assessed this at ages 1-2 years and 6-8 years in 3,062 Dutch children from a pooled data set of three birth cohorts PIAMA, PREVASC and KOALA. Single- and multilocus associations with serum IgE levels (3^{rd} vs. 1^{st} tertile) were evaluated by Chi^2-tests and multidimensionality-reduction method (MDR) in co-stimulatory genes *VTCN1*, *TNFRSF4*, *TNFRSF18*, *TNFRSF14*, *TNFSF18*, *TNFSF4*, *CD28*, *CTLA4*, *ICOS*, *PDCD1*, *BTLA*, *CD80*, *CD86*, *HLA-G*, *CD274*, *PDCD1LG2*, *CD276*, *LILRA4*, *LILRB1*, *LILRB2*, *LILRB4*, *CD40*, *ICOSLG*, and *CD40LG*. We found multiple statistically significant single $^{(S)}$ and multilocus $^{(M)}$ associations for the genes $VTCN1^{SM}$, $TNFSF18^{SM}$, $TNFSF4^{S}$, $CD28^{S}$, $CTLA4^{M}$, $ICOS^{S}$, $BTLA^{M}$, $CD80^{M}$, $CD86^{SM}$, $CD274^{SM}$, $PDCD1LG2^{M}$, $LILRA4^{SM}$, $LILRB4^{M}$, and $CD40^{SM}$ with serum IgE. Two-locus interactions of *CD86* with *VTCN1* and *CD274* with *LILRA4* were confirmed by logistic regression. In conclusion, serum IgE levels are regulated by multiple gene-gene interaction effects in the co-stimulatory pathway. We suggest using research strategies that model multiple gene-gene interactions in genetic studies. (200 words)

Key words: birth cohort, co-stimulation, genetic interaction, IgE, MDR

Introduction

Atopic diseases asthma, hay fever and atopic dermatitis are complex in origin, *i.e.* development of these diseases results from an interplay between genetic variants and environmental factors. It is increasingly clear that genes exert only a small or even no effect on the development of atopy on their own, and that they may act synergistically [(1)]. To identify which genes contribute to the development of atopic disease by association analyses, a research strategy thus should be applied that is capable to combine single nucleotide polymorphisms in multiple genes. A promising method designed for studying multifactorial diseases by computational algorithms is the Multifactor Dimensionality Reduction (MDR) method [(2)].

A core feature of atopy is the presence of increased total serum IgE. A key control point for IgE synthesis and regulation is the necessity for co-stimulation to activate T-lymphocytes. T-lymphocytes recognize antigens by T-cell receptor binding to the antigen presenting MHC class II molecules on antigen presenting cells. Subsequent cross-talk through co-stimulatory molecules integrating various positive and negative signals is necessary for differentiation and activation of the T-lymphocytes. Activated T-lymphocytes can direct an immune response and when directed towards a Th2 response, they induce B-lymphocytes to produce IgE. Due to their pivotal role in fine tuning T-lymphocyte differentiation and activation, genes encoding co-stimulatory receptors and ligands are promising candidate genes for atopic disease. Indeed, results of previous single candidate gene studies have suggested that polymorphisms in co-stimulatory genes, *i.e. CD86* [(3)], *CD40* [(4)], *CTLA4* [(5)], *HLA-G* [(6)], *ICOS* [(7)] and *PDCD1* [(8)] may influence susceptibility to atopic disease.

We hypothesized that polymorphisms in genes encoding co-stimulatory receptors and ligands influence susceptibility to elevated IgE levels. We evaluated haplotype tagging SNPs of co-stimulatory pathway genes in relation to serum IgE levels in a combined dataset of 3,062 children participating in three Dutch birth cohorts (Allergenic study [(9)]) at ages 1-2 years and 6-8 years. In

addition to studying main effects of these genes, we investigated two-, three- and four-locus gene-gene interaction with respect to serum IgE levels using MDR.

Subjects and methods

Study populations

The Allergenic study includes three prospective Dutch birth cohorts of similar design, *i.e.* PIAMA [10], PREVASC [11;12] and KOALA [13]. Genetic studies were approved by local medical ethics committees of participating institutes. All parents provided written informed consent.

IgE measurements

Total IgE levels were determined in capillary or venous blood collected at age 1 and 8 years in PIAMA, age 1, 2, and 6 in PREVASC, and age 1 and 2 in KOALA (Sanquin Research, Amsterdam). Total IgE levels were measured by radioimmunoassay as described previously [14-16] and expressed as international units per milliliter (1 IU representing 2.4 ng of IgE). Total IgE measurements were clustered at age 1-2 years and 6-8 years and analysed in tertiles, see online repository.

SNP selection and genotyping

Haplotype tagging SNPs were selected from the HapMap database [17] or from the Innate Immunity web site [18] depending on the largest number of SNPs with a minor allele frequency > 0.1 available in each database. Additionally, the biomedical literature was screened for SNPs within the candidate genes known to have functional impact or to be associated with asthma or atopy. Genomic DNA was extracted from buccal swabs or blood using standard methods [19]. DNA was amplified by using REPLI-g UltraFast technology (Qiagen™). Genotyping was performed by Competitive Allele-Specific PCR using KASPar™ genotyping chemistry, performed under contract by K-Biosciences. Quality of genotype data was guaranteed by standards of K-Biosciences and verified by comparing the genotyping results in genomic versus amplified DNA in a subset of children.

Statistical methods

All SNPs were analysed for Hardy-Weinberg equilibrium (HWE) using χ^2 statistics ($p>0.01$). We used χ^2 tests to analyse whether genotypes in this pathway were associated with elevated serum IgE levels at 1-2 years and at 6-8 years (highest vs. lowest tertile) by using a co-dominant model. For SNPs with a $p<0.10$, AIC (Akaike Information Criterion) was evaluated to determine the best fitting genetic model (additive, dominant or recessive). Odds ratios (ORs) and 95% confidence intervals (CI) were calculated by logistic regression analysis. SNPs located on the X-chromosome were analyzed in boys and girls separately. Calculations were performed using SPSS 14.0 statistical software and considered significant if $p<0.05$ (two sided).

Haplotypes were constructed from the SNPs available in each gene and frequency distributions among cases and controls were estimated by the expectation-maximization algorithm. Differences in these frequency distributions were evaluated by a log-likelihood ratio test (in house software). Since the haplotype tagging SNP selection uses multi-marker predictors to capture all information of the gene (*i.e.* aggressive tagging), we also analysed SNPs that were captured by multiple SNPs by constructing haplotypes.

Gene-gene interactions were analysed using Multifactor Dimensionality Reduction (MDR) (version 1.0.0). The MDR approach has been described previously [20], see online repository. The significance of the average prediction error was calculated using MDR permutation test and a p-value <0.05 was considered significant. Logistic regression analyses were performed to confirm significant 2-way interaction results from MDR analyses, if the interaction term was significant (in a multiplicative model) the best fitting genetic model e.g. dominant or recessive was analysed.

Results

Study population

3,062 children were genotyped and 2,927 Dutch children were selected for genetic analyses (table 1). Children who were not from Dutch origin (5.7%) were excluded from further analyses because inclusion of non-Caucasians may result in spurious genetic effects due to population stratification. Cut-off values determined by tertiles of serum IgE (online supplement table E1) identified 503 cases and 541 controls at age 1-2 years, and 307 cases and 308 controls at age 6-8 years.

SNP selection and genotyping

145 SNPs of 24 genes relevant to co-stimulation, *VTCN1* (also named *B7-H4*/ *B7x*), *TNFRSF4* (*OX40*), *TNFRSF18* (*GITR*), *TNFRSF14* (*HVEM*), *TNFSF18* (*GITRL*), *TNFSF4* (*OX40L*), *CD28*, *CTLA4*, *ICOS*, *PDCD1* (*PD-1*), *BTLA*, *CD80* (*B7-1*), *CD86* (*B7-2*), *HLA-G*, *CD274* (*B7-H1*/*PD-L1*), *PDCD1LG2* (*PD2L*), *CD276* (*B7-H3*), *LILRA4* (*ILT7*), *LILRB1* (*ILT2*), *LILRB2* (*ILT4*), *LILRB4* (*ILT3*), *CD40*, *ICOSLG*, and *CD40LG* (*CD40L*), were selected for genotyping (figure 1 and table E2 in the Supplementary Data). Five SNPs failed amplification; one SNP was monomorphic (rs7602383, *ICOS*). Genotypes for three SNPs, rs1181390 (*CD28*), rs9848900 (*CD86*), and rs7565639 (*PDCD1*) deviated from Hardy-Weinberg equilibrium in controls (at 1-2 years and/or at 6-8 years of age) and were not considered for further analyses leaving 136 SNPs for SNP and haplotype analysis. MDR analyses were performed with all 136 SNPs and a selection of 54 SNPs, after exclusion of 80 SNPs that were in LD (D'>0.8) with one or more SNPs (Supplementary Data table E2).

Single SNP analysis

A total of 8 SNPs in 4 genes were significantly associated with serum IgE at age 1-2 years at either the allele or genotype level, and 13 SNPs in 7 genes at age 6-8 years ($p<0.05$, table 2). SNPs in *CD40* and *LILRA4* were associated with IgE in both age groups, but with different SNPs, *i.e.*

rs3746821 and rs3745419 at age 1-2 years and rs3765459 and rs2241384 at age 6-8. Interestingly, the associated SNPs in both genes showed high LD (D'=0.8; and r^2=0.03 for both SNP combinations). None of the single SNP associations remained significant when corrected for multiple comparisons using false discovery rate (data not shown).

Haplotype analysis

We constructed haplotypes combining all tagging SNPs for each gene and haplotypes that captured non-genotyped SNPs (presented in supplemental data table E3). One *CD86* haplotype was significantly more prevalent in cases than controls at age 1-2 years (10 vs. 6%, p=0.01). This *CD86* haplotype comprehends the haplotype that was constructed of rs2681415 and rs2681411 to tag the non-genotyped SNPs rs9872438 and rs2681408. The combined minor alleles of these SNPs were significantly more prevalent in cases compared to controls at age 1-2 years (12 vs. 7%, p=0.0002). Another haplotype, consisting of two minor alleles for *ICOSLG* rs2070561 and rs3746963, was less prevalent among cases compared to controls at age 6-8 years (16 vs. 22%, p=0.04).

MDR analyses

Table 3 shows the best multilocus models for elevated serum IgE detected by MDR when considering one, two, three, and four loci in all co-stimulatory genes. At age 1-2 years, the prediction error of each model was statistically significant (p=0.03, 0.02, 0.02, and 0.02 respectively) based on 1,000 permutations. At 6-8 years, the 2-, 3-, and 4-locus models showed statistical significance (p=0.04, 0.02, and 0.01 respectively). At both ages multiple synergistic interactions between SNPs were observed, as indicated by the red and orange colour in the dendrograms (Figure 2).

The gene-gene interaction effects of the best 2-locus models were confirmed by conventional logistic regression analysis. At age 1-2 years, the best 2-locus model, *i.e.* rs10804556 (*CD86*) and

rs12030415 (*VTCN1*), showed a dendrogram without synergistic effect (figure 2a). The information gain (entropy based) calculated for this pair of SNPs indicated redundancy, which may be interpreted as the two SNPs acting in parallel redundant ways to increase IgE. Logistic regression revealed that individuals with minor alleles of each SNP had an increased risk to develop an elevated serum IgE level, and this risk did not further increase in individuals having the minor alleles of both SNPs. A borderline statistically significant interaction existed between the SNPs in logistic regression analysis (p-value for interaction 0.058, figure 3a). At age 6-8 years, the dendrogram showing the relation between rs4143815 (*CD274*) and rs2241384 (*LILRA4*) indicated synergistic interaction (figure 2b). Logistic regression revealed that individuals carrying one or two minor alleles of either SNP had a decreased risk to develop elevated serum IgE. In contrast, this decreased risk was not observed in individuals having one or two minor alleles of both SNPs, as indicated by a statistically significant interaction (p-value for interaction 0.004, figure 3b).

Integrating analytical strategies

Table 4 shows a summary of both the single SNP, the multilocus, and the haplotype associations of the co-stimulatory pathway with IgE at age 1-2 and 6-8 years. When considering associations at a gene level, 3 genes, *i.e. CD86*, *CD274,* and *LILRA4* showed significant associations with serum IgE at both ages 1-2 years and at 6-8 years. Three genes, *i.e. BTLA*, *CTLA4* and *VTCN1* showed association at 1-2 years of age, but not at 6-8 years. Another 5 genes, *i.e. CD276*, *ICOS*, *LILRB4*, *TNFSF4*, and *TNFSF18* showed association at 6-8 years, but not at 1-2 years.

Interestingly, 12 SNPs, in the genes *BTLA*, *CD40*, *CD80, CD86, CD274*, *CD276*, *CTLA4*, *LILRA4*, *LILRB4*, and *VTCN1* associated in multilocus models and did not associate with IgE in the single SNP or haplotype analyses. Thus, multilocus analyses by MDR identified gene variations that associated with elevated serum IgE without having a main effect.

Discussion

This study evaluated single SNPs, haplotypes, and multilocus associations of haplotype tagging SNPs in a pathway of co-stimulatory genes and their association with the predisposition to elevated serum IgE levels at ages 1-2 and 6-8 years. We expand current knowledge by showing that within a biological pathway multiple gene combinations contribute to serum IgE levels. We found multiple multilocus associations which showed statistical significance. One, two, three, and four loci models were found to better predict the presence of elevated serum IgE than would be expected by chance. Some polymorphisms were not significantly associated when tested in a single SNP analysis, yet they were significantly associated in the multilocus models. This stresses the importance of applying research strategies that model multiple interactions in genetic association studies.

It has been well established that multiple genes are involved in the predisposition to elevated serum IgE [1;21-23]. In comparison to other studies that used the MDR approach in atopic disease, we found many statistically significant multilocus models [21-23]. This is most likely because we evaluated genes in a biological pathway, whereas previous studies selected genes based on earlier reported associations with atopic disease.

Some of the evaluated genes have previously been described to associate with atopic phenotypes in single gene association studies, *i.e. CD86* [3], *CD40* [4], *CTLA4* [5;24], *HLA-G* [6], *ICOS* [7] and *PDCD1* [8] yet other genes were found not to be associated with atopy such as *CD28* [5], and *LILRB4* [25]. A summary of these association studies is presented in the Supplementary Data table E4. We here confirm main effects of the genes *CD86*, *CD40*, and *ICOS* with respect to serum IgE levels. In addition, we show that *LILRB4*, which did not appear to have a main effect on IgE levels in single gene association studies [5;25], does affect IgE levels in interaction with other gene polymorphisms. Although we did not find main effects for *CTLA4*, *HLA-G* and *PDCD1*, our results are not in contradiction with previous studies, because the associated atopic phenotypes were different, *e.g.* specific IgE to grass pollen (*PDCD1*), or asthma and airway

hyperresponsiveness (*HLA-G*), or the studies evaluated different subgroups of individuals, *e.g.* females (*CTLA-4*), or an adult asthma population (*CTLA-4*).

IgE production is known to rise during childhood and the influence of certain genes may be age-specific [26]. We therefore tested associations with IgE at two different age groups. The single and multilocus models were different between the age groups 1-2 and 6-8 years, suggesting that genes in this pathway have indeed age-specific effects on IgE development. However in our unbiased approach by applying MDR analysis, only the best models for each data set are given by MDR. Since multiple loci in this pathway may be associated with IgE development, these loci may compete with each other causing different results in each age group. Therefore we conclude that MDR is not suitable for comparison between data sets in different age groups.

Of the interactions identified, MDR indicated that the type of interaction in most of the models was synergistic. The negative gene-gene interactions in the logistic regression of the 2-locus models suggest that polymorphisms of these genes counteract each other's effect. This seems to be biologically plausible, because by counteracting effects of small genetic variations, the immune system would prevent itself from derailment by small genetic changes. Further biological interpretation of multilocus models is rather speculative, but current data provide a first suggestion that activation of allergen-specific T-lymphocyte responses can take place by modifying co-stimulatory signals. Thus the genes under study are likely important homeostatic regulators of T-lymphocyte activation and subsequent IgE production.

Our results support the hypothesis that multiple gene-gene interactions are involved in IgE regulation by fine-tuning of lymphocyte responses. Activation of T-lymphocytes requires, besides TCR-MHCII/peptide complex recognition, additional secondary signals provided by co-stimulatory molecules expressed on antigen presenting cells (APCs). The interaction between CD28 on T-lymphocytes and its two ligands B7-1 (CD80) and B7-2 (CD86) on APCs is considered to be the master co-stimulatory pathway for optimal T-cell responses [27]. CD86 is constitutively expressed on APCs at low levels and rapidly upregulated upon stimulation, whereas

CD80 is inducible and expressed later than CD86. In contrast to the stimulatory signals provided by CD28, interaction of CD80 or CD86 with the CD28 homolog CTLA-4 induces signals that down-regulate T-cell activation. CTLA-4 is constitutively expressed only on $CD4^{+}CD25^{+}$ regulatory T-lymphocytes and is induced on activated T-lymphocytes and CTLA-4 signalling plays an important role in regulating the intensity of allergic disease [(28)].

Strengths of our study are its large sample size and prospective follow-up, which enabled us to evaluate the influence of co-stimulatory pathway polymorphisms in two age groups that represent different stages of the developing immune system in early childhood and primary school age. We have previously published the successful identification of important genetic mechanisms in the development of childhood atopy in our Allergenic cohort [(9)], thus showing it's high potential for genetic studies. Furthermore, the selection of haplotype tagging SNPs has made it very unlikely that we have missed important signals from genes.

To appreciate our results we should also consider some potential limitations to our study. First, environmental influences were not considered in our analyses. Several studies have shown that environmental influences can be of great importance in the development of atopy and we recommend considering these in future research. As a result of recruitment strategies, our study represents a selected population with a relatively high number of children with atopic parents and our results may not be fully representative of the general population. Secondly, the proportion of the cohorts that participated in each age group was variable, *i.e.* in age group 1-2 years all cohorts contributed to the IgE measurements whereas at age 6-8 only IgE measurements of PREVASC and PIAMA could be evaluated. As a result, age groups may have been subject to different selection effects or variable environmental exposures and may therefore not be completely comparable. It is therefore important to note that several genes were associated with IgE at 1-2 as well as 6-8 years, internally replicating our results.

None of the single SNP associations remained significant when corrected for multiple comparisons using false discovery rate. However, due to linkage disequilibrium between the

evaluated single SNPs, the statistical tests performed are not completely independent and correction may be overly conservative. We tested the best models obtained from MDR analysis for significance based on 1000 permutations. To assess if our results in the pooled cohorts at age 1-2 and 6-8 years are valid, we investigated whether we could replicate the observed two-way interaction in our separate cohorts. Indeed we found a similar and significant interaction in 2 different cohorts at age 1-2 years in logistic regression analyses, p-values for interaction being 0.008 and 0.03 for the PREVASC and KOALA cohort respectively. We were unable to replicate the borderline significant interaction we observed in the full cohort at age 6-8 years in the two separate cohorts. This interaction was observed in the PIAMA (n=437) but not the PREVASC cohort that contained lower numbers of individuals (n=130).

The SNPs found to be associated in this study were based on haplotype selection, hence their functional role is not clear. Furthermore, MDR is a new technique that can be considered as an unbiased data-mining approach. Thus, this study can be viewed as hypothesis generating. The selected important genes in the co-stimulatory pathway can now guide replication studies and functional analyses. This may ultimately lead to novel targets for early prevention of atopy development.

In conclusion, serum IgE levels are regulated by multiple gene-gene interaction effects of many genes in the co-stimulatory pathway. The genetic interactions we observed occur in a biologically plausible way. Our results implicate that investigation of genetic contribution to complex traits will not be possible without analytical approaches that consider effects of multiple interacting loci in one gene as well as in multiple genes.

Acknowledgements

The authors would like to thank the children and parents of the PIAMA, PREVASC and KOALA study for their participation. In addition we acknowledge the field workers, secretaries and scientific collaborators dedicated to the PIAMA, PREVASC and KOALA cohorts, Marcel Bruinenberg for his advice on DNA isolation, processing and genotyping, Ilja Nolte for her advice on haplotype analyses, and Antoon van Oosterhout for his advice on gene selection. We thank Jason H. Moore for his helpful comments to our questions with respect to MDR analyses.

This study was financially supported by ZonMW grant number 912-03-031. G.H. Koppelman is supported by a Zon-Mw VENI grant, number 91656091.

322 **Web Resources**

HapMap database

URL: http://www.hapmap.org; data release 19. phase II. October 2005

Innate Immunity web site. Innate Immunity PGA. NHLBI Program in Genomic Applications

URL: http://www.innateimmunity.net/data/homology; October 2005

323 K-Biosciences. Cambridge. UK

324 URL: http://www.kbiosciences.co.uk

References

(1) Kabesch M. Candidate gene association studies and evidence for gene-by-gene interactions. Immunol Allergy Clin North Am 2005; 25(4):681-708.

(2) Blakey JD. Looking for a bit of co-action? Thorax 2007; 62(3):196-7.

(3) Corydon TJ, Haagerup A, Jensen TG, Binderup HG, Petersen MS, Kaltoft K et al. A functional CD86 polymorphism associated with asthma and related allergic disorders. J Med Genet 2007; 44(8):509-15.

(4) Park JH, Chang HS, Park CS, Jang AS, Park BL, Rhim TY et al. Association analysis of CD40 polymorphisms with asthma and the level of serum total IgE. Am J Respir Crit Care Med 2007; 175(8):775-82.

(5) Howard TD, Postma DS, Hawkins GA, Koppelman GH, Zheng SL, Wysong AK et al. Fine mapping of an IgE-controlling gene on chromosome 2q: Analysis of CTLA4 and CD28. J Allergy Clin Immunol 2002; 110(5):743-51.

(6) Ober C, Tsalenko A, Parry R, Cox NJ. A second-generation genomewide screen for asthma-susceptibility alleles in a founder population. Am J Hum Genet 2000; 67(5):1154-62.

(7) Shilling RA, Pinto JM, Decker DC, Schneider DH, Bandukwala HS, Schneider JR et al. Cutting edge: polymorphisms in the ICOS promoter region are associated with allergic sensitization and Th2 cytokine production. J Immunol 2005; 175(4):2061-5.

(8) James ES, Harney S, Wordsworth BP, Cookson WO, Davis SJ, Moffatt MF. PDCD1: a tissue-specific susceptibility locus for inherited inflammatory disorders. Genes Immun 2005; 6(5):430-7.

(9) Bottema RW, Reijmerink NE, Kerkhof M, Koppelman GH, Stelma FF, Gerritsen J, Thijs C, Brunekreef B, van Schayck CP, Postma DS. IL13, CD14, pet and tobacco smoke influence atopy in 3 Dutch cohorts; The Allergenic study. Eur Respir J. 2008 Apr 16; Epub ahead of print.

(10) Brunekreef B, Smit J, de JJ, Neijens H, Gerritsen J, Postma D et al. The prevention and incidence of asthma and mite allergy (PIAMA) birth cohort study: design and first results. Pediatr Allergy Immunol 2002; 13 Suppl 15:55-60.

(11) Schonberger HJ, Dompeling E, Knottnerus JA, Maas T, Muris JW, van WC et al. The PREVASC study: the clinical effect of a multifaceted educational intervention to prevent childhood asthma. Eur Respir J 2005; 25(4):660-70.

(12) Kuiper S, Maas T, van Schayck CP, Muris JW, Schonberger HJ, Dompeling E et al. The primary prevention of asthma in children study: design of a multifaceted prevention program. Pediatr Allergy Immunol 2005; 16(4):321-31.

(13) Kummeling I, Thijs C, Penders J, Snijders BE, Stelma F, Reimerink J et al. Etiology of atopy in infancy: the KOALA Birth Cohort Study. Pediatr Allergy Immunol 2005; 16(8):679-84.

(14) Aalberse RC, Koshte V, Clemens JG. Immunoglobulin E antibodies that crossreact with vegetable foods, pollen, and Hymenoptera venom. J Allergy Clin Immunol 1981; 68(5):356-64.

(15) Akkerdaas JH, Wensing M, Asero R, Fernandez RM, Knulst AC, Bolhaar S et al. IgE binding to pepsin-digested food extracts. Int Arch Allergy Immunol 2005; 138(3):203-8.

(16) Stallman PJ, Aalberse RC. Estimation of basophil-bound IgE by quantitative immunofluorescence microscopy. Int Arch Allergy Appl Immunol 1977; 54(1):9-18.

(17) A haplotype map of the human genome. Nature 2005; 437(7063):1299-320.

(18) Lazarus R, Vercelli D, Palmer LJ, Klimecki WJ, Silverman EK, Richter B et al. Single nucleotide polymorphisms in innate immunity genes: abundant variation and potential role in complex human disease. Immunol Rev 2002; 190:9-25.

(19) Sambrook J, Russell D. Molecular Cloning. 3rd Edition. Preparation of plasmid DNA by lysis with SDS. 2007.

(20) Ritchie MD, Hahn LW, Moore JH. Power of multifactor dimensionality reduction for detecting gene-gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity. Genet Epidemiol. 2003 Feb;24(2):150-7.

(21) Chan IH, Leung TF, Tang NL, Li CY, Sung YM, Wong GW et al. Gene-gene interactions for asthma and plasma total IgE concentration in Chinese children. J Allergy Clin Immunol 2006; 117(1):127-33.

(22) Park HW, Shin ES, Lee JE, Kwon HS, Chun E, Kim SS et al. Multilocus analysis of atopy in Korean children using multifactor-dimensionality reduction. Thorax 2007; 62(3):265-9.

(23) Leung TF, Chan IH, Wong GW, Li CY, Tang NL, Yung E et al. Association between candidate genes and lung function growth in Chinese asthmatic children. Clin Exp Allergy 2007; 37(10):1480-6.

(24) Sohn MH, Kim SH, Song TW, Kim KW, Kim ES, Park HS et al. Cytotoxic T lymphocyte-associated antigen-4 gene polymorphisms confer susceptibility to atopic asthma in Korean children. Pediatr Pulmonol 2007; 42(6):542-7.

(25) Heinzmann A, Blattmann S, Forster J, Kuehr J, Deichmann KA. Common polymorphisms and alternative splicing in the ILT3 gene are not associated with atopy. Eur J Immunogenet 2000; 27(3):121-7.

(26) Bottema RW, Reijmerink NE, Koppelman GH, Kerkhof M, Postma DS. Phenotype definition, age, and gender in the genetics of asthma and atopy. Immunol Allergy Clin North Am 2005; 25(4):621-39.

(27) Harding, F. A., J. G. McArthur, J. A. Gross, D. H. Raulet, J. P. Allison. 1992. CD28-mediated signalling co-stimulates murine T cells and prevents induction of anergy in T-cell clones. *Nature* 356: 607-609.

(28) van Wijk, F., S. Hoeks, S. Nierkens, S. J. Koppelman, P. van Kooten, L. Boon, L. M. Knippels, R. Pieters. 2005. CTLA-4 signaling regulates the intensity of hypersensitivity responses to food antigens, but is not decisive in the induction of sensitization. *J. Immunol.* 174: 174-179.

(29) Beier KC, Kallinich T, Hamelmann E.T-cell co-stimulatory molecules: novel targets for the treatment of allergic airway disease. Eur Respir J. 2007 Aug;30(2):383-90. Review.

(30) Kallinich T, Beier KC, Wahn U, Stock P, Hamelmann E. T-cell co-stimulatory molecules: their role in allergic immune reactions. Eur Respir J. 2007 Jun;29(6):1246-55. Review.

405 **Table 1.** Characterictics of participating children in the Allergenic birth cohort.

Characteristics	PIAMA	PREVASC	KOALA	P[a]
Participants in genetic study (number)	1,037	374	1,651	-[b]
Ethnicity (% Dutch origin)	95.1	95.7	95.2	-
Boys (%)	51.2	49.2	50.6	0.80
Total serum IgE				
1 year (IU/ml) [c]	7.1 (2.0-17.0) N=369	8.6 (3.5-19.4) N=226	6.0 (2.6-12.5) N=699	**0.002**
2 years (IU/ml) [c]	n.a.[d]	11.7 (4.2-28.7) N=358	12.0 (3.7-38.0) N=704	0.80
6 years (IU/ml) [c]	n.a.	22.5 (7.6-67.0) N=218	n.a.	-
8 years (IU/ml) [c]	64.9 (23.0-240.0) N=748	n.a.	n.a.	-

406 [a] P = p-value based for comparison between cohorts by Chi-square test or analysis of variance
407 where appropriate; [b] - = not tested; [c] geometric mean (interquartile range); [d] n.a. = not available.

Table 2. SNPs significantly associated with increased serum IgE, at allele and genotype level ($p<0.05$).

		Association at allele level 1-2 years					**Association at genotype level 1-2 years**			
Gene	rs number	MAF[a] (controls / cases)	Allele	OR[b]	(95% CI)[c]	P[d]	Genotype	OR[b]	(95% CI)[c]	P[e]
CD86	rs10804556	0.18 / 0.24	G	1.46	1.14-1.87	**0.003**	A:A	1.00		**0.002**
							G:A	1.56	1.19-2.04	
							G:G	1.79	0.97-3.28	
CD86	rs2681415	0.10 / 0.15	G	1.35	1.00-1.82	**0.008**	A:A	1.00		**0.004**
							G:A	1.67	1.23-2.27	
							G:G	1.34	0.54-3.33	
VTCN1	rs10047089	0.45 / 0.50	A	1.25	1.05-1.49	**0.01**	G:G	1.00		**0.05**
							G:A	1.29	0.96-1.72	
							A:A	1.53	1.08-2.17	
VTCN1	rs12030415	0.24 / 0.29	A	1.31	1.07-1.60	**0.05**	G:G	1.00		**0.003**
							G:A	1.05	0.81-1.37	
							A:A	2.46	1.47-4.12	
CD86	rs1915087	0.32 / 0.36	C	1.32	1.07-1.62	**0.01**	T:T	1.00		**0.04**
							C:T/C:C	1.50	1.02-2.20	
CD86	rs11717893	0.27 / 0.24	-	-	-	-	T:T	1.00		**0.04**
			-	-	-	-	C:T	0.98	0.76-1.27	
			-	-	-	-	C:C	0.50	0.29-0.86	
CD40	rs3746821	0.11 / 0.09	-	-	-	-	G:G/G:T	1.00		**0.04**
			-	-	-	-	T:T	0.26	0.07-0.93	
LILRA4	rs3745419	0.17 / 0.19	-	-	-	-	T:T/A:T	1.00		**0.04**
			-	-	-	-	A:A	2.31	1.04-5.17	
		Association at allele level 6-8 years					**Association at genotype level 6-8 years**			
Gene	rs number	MAF[a] (controls / cases)	Allele	OR[b]	(95% CI)[c]	P[d]	Genotype	OR[b]	(95% CI)[c]	P[e]
CD274	rs2297136	0.42 / 0.50	T	1.35	1.07-1.69	**0.01**	C:C	1.00		**0.03**
							C:T	1.25	0.86-1.83	
							T:T	1.89	1.19-3.02	
TNFSF18	rs2236876	0.23 / 0.28	T	1.32	1.02-1.72	**0.04**	C:C	1.00		**0.05**
							C:T	1.52	1.09-2.13	
							T:T	1.34	0.66-2.71	
LILRA4	rs17836364	0.15 / 0.20	A	1.37	1.01-1.84	**0.04**	G:G	1.00		**0.04**
							G:A/A:A	1.44	1.02-2.04	
ICOS	rs4521021	0.19 / 0.25	C	1.34	1.02-1.76	**0.04**	T:T	1.00		**0.02**
							C:T/C:C	1.5	1.08-2.08	
TNFSF18	rs975074	0.51 / 0.45	G	0.79	0.62-1.00	**0.05**	T:T	1.00		**0.04**
							G:T/G:G	0.66	0.45-0.98	
CD28	rs1181390	0.37 / 0.32	-	-	-	-	C:C	1.00		**0.04**
			-	-	-	-	C:A	1.15	0.80-1.66	
			-	-	-	-	A:A	0.63	0.40-0.99	
CD274	rs10975123	0.19 / 0.16	-	-	-	-	C:C/C:T	1.00		**0.02**
			-	-	-	-	T:T	0.26	0.08-0.78	
TNFSF4	rs11811856	0.26 / 0.30	-	-	-	-	C:C	1.00		**0.05**
			-	-	-	-	C:G	0.97	0.69-1.36	
			-	-	-	-	G:G	2.19	1.14-4.22	
CD40	rs3765459	0.21 / 0.25	-	-	-	-	G:G/G:A	1.00		**0.02**
			-	-	-	-	A:A	2.53	1.19-5.39	
TNFSF18	rs723858	0.19 / 0.23	-	-	-	-	A:A	1.00		**0.05**
			-	-	-	-	A:T/T:T	1.39	1.00-1.94	
LILRA4	rs2241384	0.22 / 0.18	-	-	-	-	C:C	1.00		**0.04**
			-	-	-	-	C:T/T:T	0.71	0.50-0.99	
CD274	rs2297137	0.26 / 0.21	T	0.75	0.58-0.99	**0.04**	-	-	-	-
CD274	rs4143815	0.31 / 0.25	C	0.77	0.59-0.99	**0.04**	-	-	-	-

[a] MAF = minor allele frequency; [b] OR = odds ratio calculated by logistic regression; [c] 95% confidence interval; [d] P-value determined by Chi-square test (1 df); [e] P-value determined by logistic regression

412 **Table 3.** Results of MDR analysis evaluating all co-stimulatory genes.

Loci[a]	Best model		Mean CV[b] consistency	Mean prediction error (%)	P[c]
	Gene(s)	SNP(s)			
1-2 years					
1	*CD86*	rs10804556	9.4	**45.3**	**0.03**
2	*CD86 and VTCN1*	rs10804556 rs12030415	6.4	**43.4**	**0.02**
3	*VTCN1 BTLA*	rs9288953 rs7023227			
	CD274	rs11805655	4.6	**42.8**	**0.02**
4	*VTCN1 CD40*	rs9288953 rs745307			
	CD86 CTLA4	rs2332096 rs231806	1.2	**39.9**	**0.006**
6-8 years					
1	*ICOS*	rs4521021	5.4	45.2	0.09
2	*CD274 LILRA4*	rs4143815 rs2241384	2.6	42.3	**0.04**
3	*TNFSF18 LILRA4*	rs975074 rs2241385			
	LILRB4	rs3745871	5.6	**39.8**	**0.02**
4	*CD276 CD80*	rs11072430 rs610902			
	CD80 CD86	rs7648642 rs4308217	1.4	**38.1**	**0.006**

413 [a] Number of loci considered; [b] CV=cross-validation; [c] Significance of prediction error (empirical p-
414 value based on 1000 permutations).

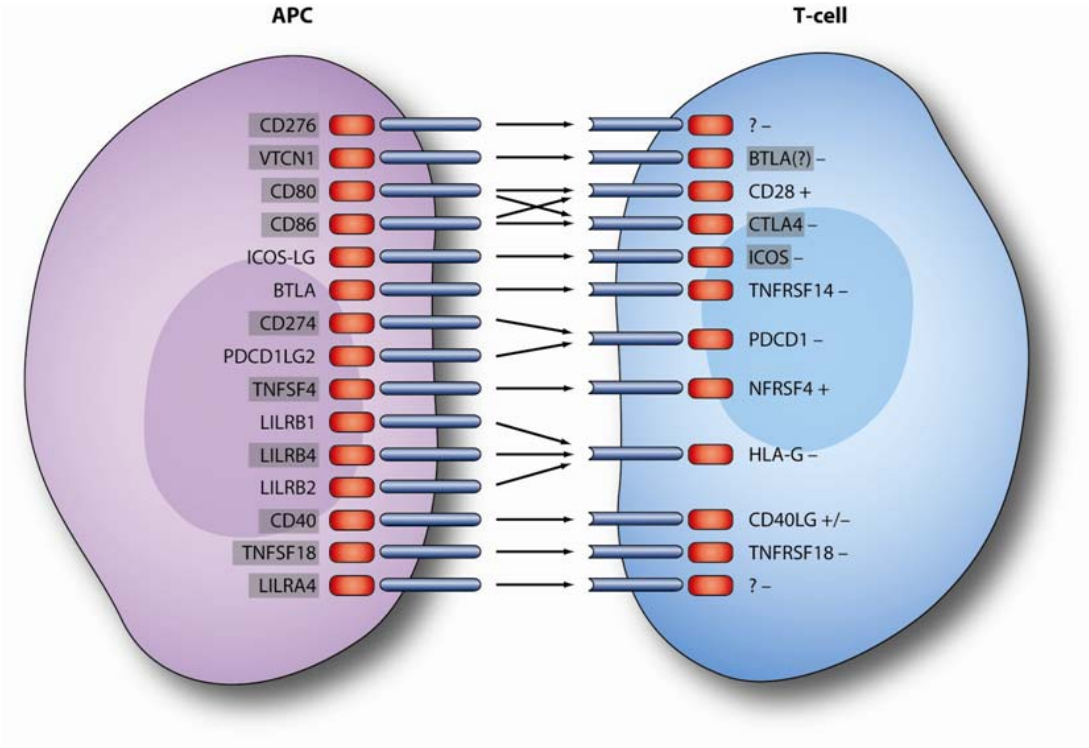
Table 4. Summary of single and multilocus associations (p-values) in the co-stimulatory pathway at ages 1-2 years and 6-8 years.

Gene	Rs number	Association with IgE at 1-2 years						Association with IgE at 6-8 years					
		Allele	Genotype	Association by MDR				Allele	Genotype	Association by MDR			
				1 locus	2 locus	3 locus	4 locus			1 locus	2 locus	3 locus	4 locus
BTLA	rs7023227					**0.02**							
CD28	rs1181390		**0.04**										
CD40	rs3746821		**0.03**										
CD40	rs3765459								**0.02**				
CD40	rs745307						**0.006**						
CD80	rs610902												**0.006**
CD80	rs7648642												**0.006**
CD86	rs10804556	**0.003**	**0.002**	**0.03**	**0.02**								
CD86	rs2681415	**0.008**	**0.003**										
CD86	rs1915087	**0.01**	**0.04**										
CD86	rs11717893		**0.04**										
CD86	rs2332096						**0.006**						
CD86	rs4308217												**0.006**
CD274	rs10975123								**0.02**				
CD274	rs2297137							**0.04**					
CD274	rs4143815							**0.04**			**0.04**		
CD274	rs2297136							**0.01**	**0.03**				
CD274	rs11805655					**0.02**							
CD276	rs11072430												**0.006**
CTLA4	rs231806						**0.006**						
TNFSF18	rs2236876							**0.04**	**0.05**				
TNFSF18	rs975074							**0.05**	**0.04**			**0.02**	
TNFSF18	rs723858								**0.05**				
ICOS	rs4521021							**0.04**	**0.02**				
LILRA4	rs17836364							**0.04**	**0.04**				
LILRA4	rs3745419		**0.03**										
LILRA4	rs2241384								**0.04**		**0.04**		
LILRA4	rs2241385											**0.02**	
LILRB4	rs3745871											**0.02**	
TNFSF4	rs11811856								**0.04**				
VTCN1	rs10047089	**0.01**	**0.05**										
VTCN1	rs12030415	**0.05**	**0.002**		**0.02**								
VTCN1	rs9288953					**0.02**	**0.006**						

418 **Figure legends**

419 **Figure 1.** Hypothetical scheme of co-stimulatory receptor and ligand pairs evaluated in this study.
420 Genes that show association with serum IgE are boxed. APC=antigen presenting cell; *BTLA* = B
421 and T lymphocyte attenuator; *CD40LG* = CD40 antigen ligand; *CTLA4* = cytotoxic T-lymphocyte-
422 associated antigen 4; *TNFSF18* = tumor necrosis factor (ligand) superfamily, member 18; *HLA-G*
423 = HLA-G histocompatibility antigen, class I, G; *ICOS* = inducible T-cell co-stimulator; *ICOSLG* =
424 inducible T-cell co-stimulator ligand; *LILRA4* = leukocyte immunoglobulin-like receptor,
425 subfamily A (with TM domain), member 4; *LILRB1* = leukocyte immunoglobulin-like receptor,
426 subfamily B (with TM and ITIM domains), member 1; *LILRB2* = leukocyte immunoglobulin-like
427 receptor, subfamily B (with TM and ITIM domains), member 2; *LILRB4* = leukocyte
428 immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 4; *PDCD1*=
429 programmed cell death 1; *PDCD1LG2* = programmed cell death 1 ligand 2; *TNFRSF14* = tumor
430 necrosis factor receptor superfamily, member 14; *TNFRSF18* = tumor necrosis factor receptor
431 superfamily, member 18; *TNFRSF4* = tumor necrosis factor receptor superfamily, member 4;
432 *TNFSF4* = tumor necrosis factor (ligand) superfamily, member 4; *VTCN1* = V-set domain
433 containing T cell activation inhibitor 1; ? = receptor unknown; - = inhibitory signalling effect; +
434 positive signalling effect; Information adapted from [29-30].

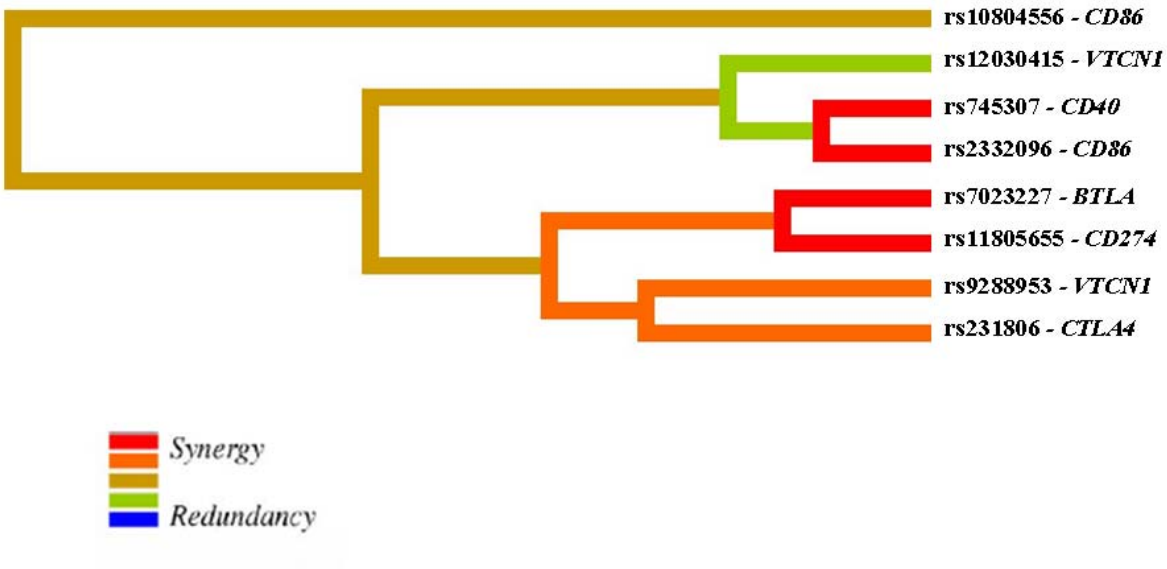
Figure 1



435
436

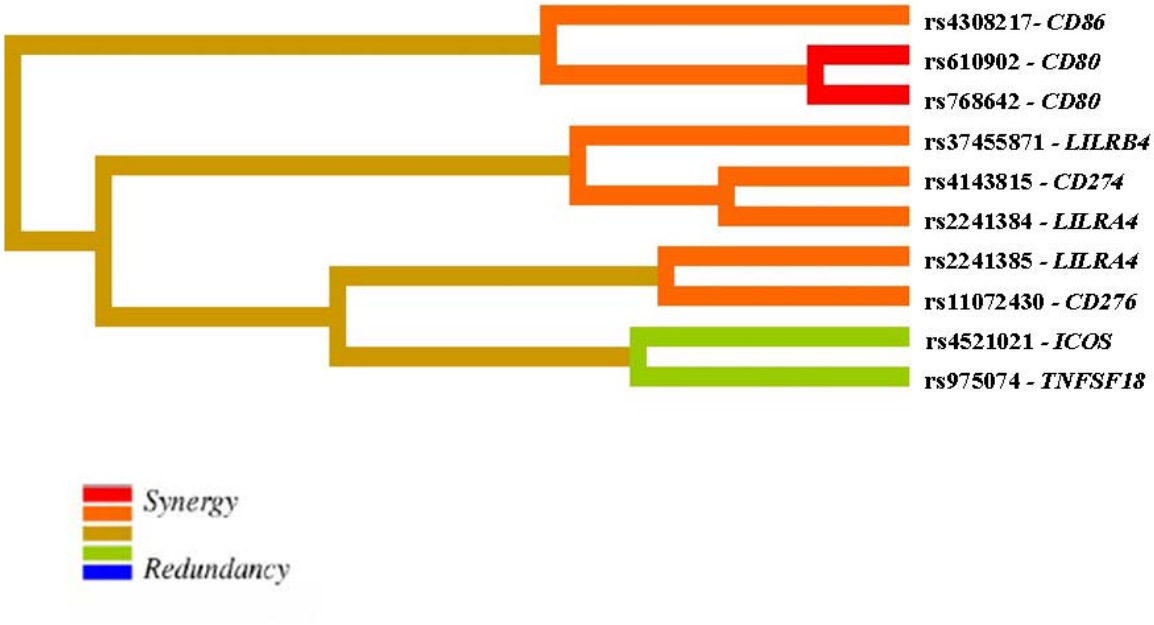
437 **Figure 2.** Interaction dendrogram visualizing the information gain (entropy based) associated with
438 considering SNP interactions in the co-stimulatory pathway at (a) 1-2 years and (b) 6-8 years of
439 age. The color of the line connecting pairs of SNPs indicates the degree of synergy (positive
440 information gain i.e. the SNPs combined provide a different (stronger or weaker) effect or
441 redundancy (negative information gain). The shorter the line the stronger the interaction.

Figure 2 (a)



442

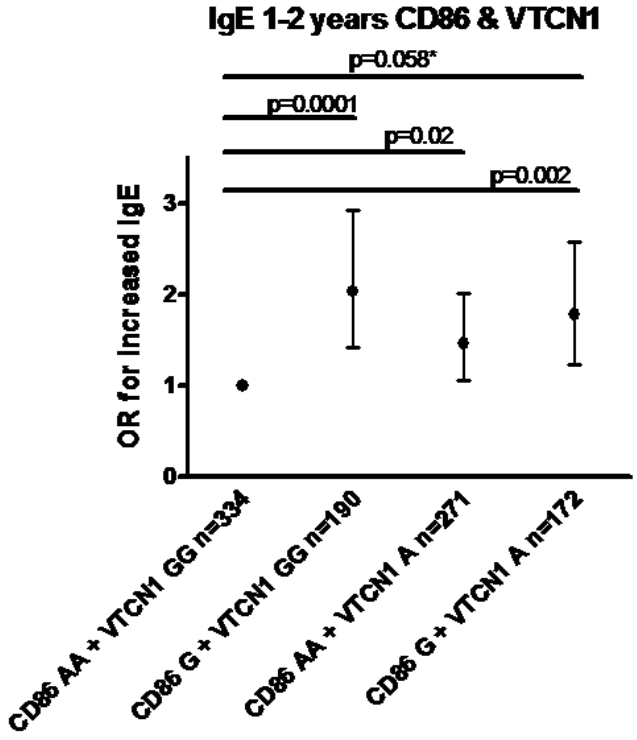
Figure 2 (b)



443
444

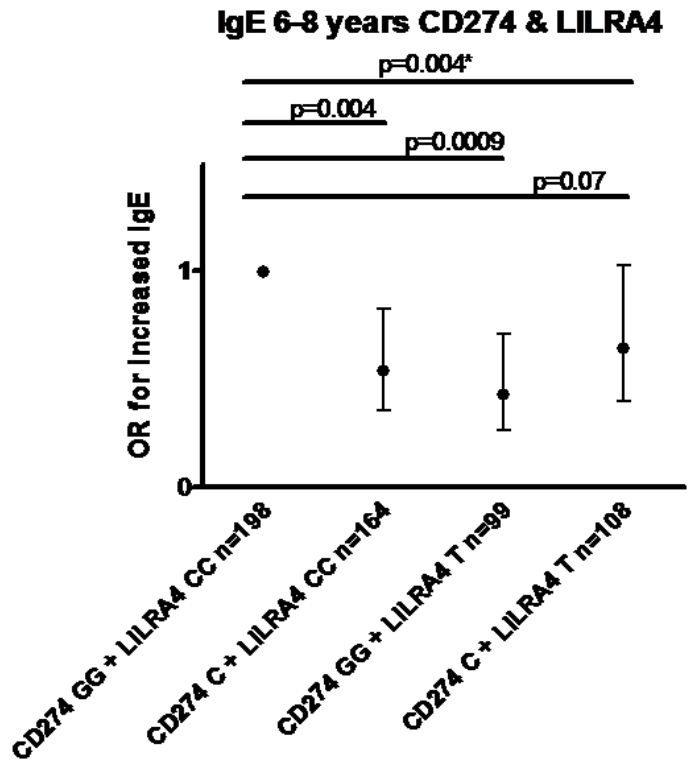
445 **Figure 3.** Two locus associations between co-stimulatory molecules by logistic regression.
446 (a) best 2 locus model at 1-2 years; and (b) best 2 locus model at 6-8 years; OR= odds ratio and
447 95% confidence interval; * p-value for interaction.

Figure 3 (a)



448

Figure 3 (b)



449