

# Polymorphisms in the airway epithelium related genes CDHR3 and EMSY are associated with asthma susceptibility

**Miao-miao Zhang**

Sichuan University

**Guo Chen**

Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital

**Yu Wang**

Sichuan University

**Shou quan Wu**

Sichuan University

**Andrew J Sandford**

The University of British Columbia and St. Paul's Hospital

**Jian-Qing He** (✉ [jianqing\\_he@scu.edu.cn](mailto:jianqing_he@scu.edu.cn))

Sichuan University West China Hospital <https://orcid.org/0000-0003-0557-3089>

---

## Research article

**Keywords:** CDHR3, EMSY, asthma, polymorphism, susceptibility

**Posted Date:** October 22nd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-31705/v3>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on November 19th, 2020. See the published version at <https://doi.org/10.1186/s12890-020-01334-0>.

# Abstract

**Background:** As a main line of defense of the respiratory tract, the airway epithelium plays an important role in the pathogenesis of asthma. *CDHR3* and *EMSY* were reported to be expressed in the human airway epithelium. Although previous genome-wide association studies found that the two genes were associated with asthma susceptibility, similar observations have not been made in the Chinese Han population.

**Methods:** A total of 300 asthma patients and 418 healthy controls unrelated Chinese Han individuals were enrolled. Tag-single nucleotide polymorphisms (Tag-SNPs) were genotyped and the associations between SNPs and asthma risk were analyzed by binary logistic regression analysis.

**Results:** After adjusting for confounding factors, the A allele of rs3847076 in *CDHR3* was associated with increased susceptibility to asthma (OR = 1.407, 95% CI: 1.030-1.923). For the *EMSY* gene, the T alleles of both rs2508746 and rs12278256 were related with decreased susceptibility to asthma (additive model: OR = 0.718, 95% CI: 0.536-0.961; OR = 0.558, 95% CI: 0.332-0.937, respectively). In addition, the GG genotype of rs1892953 showed an association with increased asthma risk under the recessive model (OR = 1.667, 95% CI: 1.104-2.518) and the GATCTGAGT haplotype in *EMSY* was associated with reduced asthma risk (P = 0.037).

**Conclusions:** This study identified novel associations of rs3847076 in *CDHR3*, as well as rs1892953, rs2508746 and rs12278256 in *EMSY* with adult asthma susceptibility in the Chinese Han population. Our observations suggest that *CDHR3* and *EMSY* may play important roles in the pathogenesis of asthma in Chinese individuals. Further study with larger sample size is needed.

## Background

Asthma is a chronic airway inflammatory disease that affects populations throughout the world. A World Health Organization report [1] predicted that the number of asthma patients would increase to 400 million by 2025 and 250,000 patients may die from this disease each year. A recent survey indicated that the prevalence of asthma among individuals aged >14 years was 1.24% and there are approximately 30 million asthmatic patients in China[2]. The pathogenesis of asthma is still incompletely understood but it is known that genetic factors play a significant part in asthma susceptibility. The heritability of asthma was estimated to be 60% to 70% in an Australian twin study[3]. Genetic factors contributed to 90% of the variance in the susceptibility to asthma in a 5-year-old twin pair study [4].

As the first barrier between the human body and the environment, the airway epithelium has an important role in regulating the inflammation, immunity and tissue repair in the pathogenesis of asthma [5]. One genome-wide association study (GWAS) of a Danish population identified Cadherin related family member 3 (*CDHR3*), which is highly expressed in human airway epithelium, as a susceptibility locus for childhood asthma with severe exacerbations[6]. A GWAS in 2017 demonstrated that Chromosome 11 open reading frame 30 (*C11orf30*), also called *EMSY* or BRCA2-interacting transcriptional repressor,

another gene expressed in airway epithelium[7], was a risk locus for food allergy in a Canadian population [8] and this gene has been shown to be involved in the epigenetic regulation of gene expression [9]. However, there have been few studies of these two genes in Chinese asthmatics. Therefore, this study aimed to investigate the association of common variants in *CDHR3* and *EMSY* with adult asthma in the Chinese population.

## Methods

### Study population

The asthmatic cases were diagnosed by at least three respiratory physicians from the West China Hospital according to the criteria of the Global Strategy for Asthma Management and Prevention [10]. The healthy controls were collected from the physical examination center in the same hospital. Subjects were excluded if any of the following conditions were present: chronic obstructive pulmonary disease, diabetes, tumors, any immune disease, and immune deficiency. Use of hormones or immunosuppressive drugs was also exclusion criteria. Cases and controls were unrelated Chinese Han individuals. After signing the informed consent, 3 ml of venous blood were drawn from every subject and stored in a -80°C freezer. All blood specimens were collected from September 2013 to September 2016. The study was approved by the ethical committee of the West China Hospital of Sichuan University (Protocol No. 23).

### SNP Selection and Genotyping

Single nucleotide polymorphisms (SNPs) with minor allele frequency (MAF)  $\geq 0.05$  and  $r^2 \geq 0.64$ , located in the region 3000 base pairs upstream to 300 base pairs downstream of *CDHR3* were downloaded from the Han Chinese in Beijing database of the Genome Variation Server 147 (<http://gvs.gs.washington.edu/GVS/>), which is an online resource based on dbSNP. The final selections were 23 tag-SNPs including rs3887998, rs12155008, rs41267, rs3892893, rs10270308, rs34426483, rs193795, rs2526978, rs381188, rs10241452, rs3847076, rs11981655, rs10808147, rs193806, rs2528883, rs41269, rs2526979, rs2526976, rs41262, rs41266, rs6967330, rs41270 and rs448024 (Table S1). The selection of SNPs in *EMSY* was based on the tag-SNP strategy and literature review[11-14]. The tag-SNP selection strategy was the same as above except for  $r^2 \geq 0.80$ . The 17 SNPs selected were rs3753051, rs7125744, rs7926009, rs4945087, rs2508740, rs1939469, rs7115331, rs1044265, rs12278256, rs2513513, rs2508755, rs2155219, rs2513525, rs2508746, rs1892953, rs7130588 and rs10899234 (Table S2).

Genomic DNA was extracted from the blood samples using a genomic DNA

purification kit (Axygen Scientific Inc, Union City, CA, USA). SNPs were genotyped by Genesky Bio-Tech Co., Ltd (<http://geneskybiotech.com/index.html>) using the SNPscan<sup>TM</sup> multiplex SNP genotyping technique based on double ligation and multiplex fluorescence polymerase Chain Reaction (PCR) [15]. As a quality control measure, 5% of random samples were repeated genotyped with a concordance rate of 100%.

## Data Analyses

Statistical tests were performed using the Statistical Package for the Social Sciences (SPSS, SPSS Inc., Chicago, IL, USA), version 21.0. A  $p$  value  $<0.05$  was considered to be statistically significant. Continuous and categorical variables were analyzed by student's t-test and chi-square test, respectively. Genotype distributions under additive, dominant and recessive models were calculated by binary logistic regression analysis. Hardy-Weinberg equilibrium (HWE) among the controls was tested using plink software. Haploview and SHEsis software (<http://analysis.bio-x.cn>) were combined to perform linkage disequilibrium (LD) and haplotype analysis. Statistically significant SNPs were predicted by the software RegulomeDB (<http://www.regulomedb.org/>) and Haploreg v4 (<http://compbio.mit.edu/HaploReg>). An Excel spreadsheet set up by Tomas Andersson was applied to calculate the biological interaction on an additive scale[16]. And there are three measures: RERI, the relative excess risk due to interaction; AP, the attributable proportion due to interaction; and S, the synergy index. RERI and S have similar meaning, and AP means the proportion of all cases that can be attributed to the interaction of the two factors. If there is no biological interaction, RERI and AP are equal to 0 and S is equal to 1. The interaction between these significant SNPs and smoking (smoking status=1, non-smoking status=0), sex (male=1, female=0) and body mass index (BMI, BMI $\geq$ 24=1, BMI $<$ 24=0) was calculated.

## Results

### Subject characteristics

A total of 300 asthma patients and 418 healthy controls were enrolled. The average ages of asthma patients and controls were  $43.6\pm 13.48$  and  $44.09\pm 13.75$  years, respectively. No significant differences in sex, body mass index (BMI) and smoking history were observed between case and control groups (Table 1). Late-onset asthma (age of asthma onset  $\geq 18$  years) accounted for 74.3% in the case group. Most asthma individuals were outpatients (88.67%), and we could only get half of the patients' reports of eosinophil count, total serum immunoglobulin E (IgE), pulmonary function test and provocation or relaxation test. The other half of the patients' relevant tests were done in other hospitals, but we couldn't acquire. 58.33% of the patients adopted the step 4 treatment plan according to Global Strategy for Asthma Management and Prevention (2018 update)[17], 12.67% adopted step 5, 3.33% used step 3 and the other patients' treatment information was lost.

### Association analyses between *CDHR3*, *EMSY* SNPs and asthma susceptibility

The characteristics of the selected SNPs are listed in Table S1 and S2. Rs10899234 in *EMSY* and rs6967330 in *CDHR3* were excluded due to their deviation from HWE in the control subjects ( $P<0.05$ ). The genotyping assays failed for rs12155008, rs41270 and rs448024 in *CDHR3*.

After adjusting for confounding factors including age, sex, BMI and smoking history, four SNPs were found to be associated with asthma susceptibility (Table 2 and Figure S1). The A allele of rs3847076 in *CDHR3* was associated with increased susceptibility to asthma under the additive model ( $P = 0.032$ , OR =

1.407, 95% CI: 1.030-1.923). For *EMSY*, both the TC/TT genotype and T allele of rs2508746 were associated with decreased risk of asthma (dominant model:  $P = 0.019$ , OR = 0.660, 95% CI: 0.465-0.935; additive model:  $P = 0.026$ , OR = 0.718, 95% CI: 0.536-0.961). The TG/TT genotype and T allele of rs12278256 were associated with reduced asthma risk (dominant model:  $P = 0.033$ , OR = 0.563, 95% CI: 0.332-0.953; additive model:  $P = 0.027$ , OR = 0.558, 95% CI: 0.332-0.937). Finally, the GG genotype of rs1892953 showed an association with increased asthma risk under the recessive model ( $P = 0.015$ , OR = 1.667, 95% CI: 1.104-2.518). After excluding people who were lack of smoking or BMI information, we used the online software SNPStats (<https://snpstats.net/>) for statistical analysis again and the results (shown in the Table S3) were similar to Table 1. However, it should be reminded that some significant associations maybe were expected just by chance.

Stratified analysis results by gender, smoking status, BMI status and onset age of asthma were shown in Table 3. The cut-off point of adult BMI in China is different from other countries, as  $18.5 \leq \text{BMI} < 24 \text{kg/m}^2$  meaning normal weight range and  $\text{BMI} \geq 24 \text{kg/m}^2$  meaning overweight or obese[18]. Allele A of rs3847076 was associated with increased susceptibility to asthma in male subgroup, smoking subgroup,  $\text{BMI} < 24 \text{kg/m}^2$  subgroup and late onset asthma subgroup ( $P=0.023$ , OR=1.869;  $P=0.009$ , OR=2.168;  $P=0.005$ , OR=1.835 and  $P=0.023$ , OR=1.457, respectively). Similarly, rs2508746 TC+TT was related with decreased asthma susceptibility in the non-smoking subgroup, non-overweight subgroup, and late-onset asthma subgroup in dominant model ( $P=0.014$ , OR=0.618;  $P=0.027$ , OR=0.612 and  $P=0.016$ , OR=0.637, respectively). Meanwhile, rs1892953 GG shown increased risk of asthma in the female subgroup, non-smoking subgroup, non-overweight subgroup, and late onset asthma subgroup in recessive model ( $P=0.038$ , OR=1.738;  $P=0.04$ , OR=1.615;  $P=0.017$ , OR=1.910 and  $P=0.017$ , OR=1.680, respectively). Rs12278256 T was still associated with decreased asthma susceptibility in female subgroups, non-smoking subgroups, and non-overweight subgroups in additive model ( $P=0.032$ , OR=0.465;  $P=0.02$ , OR=0.508 and  $P=0.028$ , OR=0.481, respectively). The interaction between these four SNPs and smoking, sex and BMI were shown in Table S4. We got significant interaction between rs3847076 and rs1892953 and smoking, sex and BMI, while no interaction was found between rs12278256 and these clinical phenotypes. Meanwhile, significant interaction could also be observed between rs2508746 and either gender or BMI.

We further explored the relationship between eosinophil count, total serum IgE, pulmonary function test of asthma patients and gene variants. Eosinophil count was higher in asthma patients with genotype CC of rs3847076 comparing to individuals with genotype CA (Table S5). Total IgE was related with four variants of *CDHR3* and one variant of *EMSY* (Table S6). Both FEV1% predicted and FEV1/FVC% were significant different in nine SNP genotypes, including rs2508746 and rs1892953. Higher FEV1/FVC% was also seen in genotype GG of rs12278256 (Table S7). Due to the small number of samples, further verification research is needed.

## Haplotype and LD analysis

The LD between SNPs of *CDHR3* and *EMSY* was low and those SNPs were divided into eight haplotype blocks with Haploview software (Figures 1 and 2). Only the haplotype consisting of GATCTGAGT in block 1 of *EMSY* was associated with decreased risk of asthma ( $P = 0.037$ , OR = 0.615, 95% CI: 0.388-0.975) (Table 4).

### Functional prediction results

Four statistically significant SNPs were predicted using the software RegulomeDB and Haploreg v4 (Table S8). Rs144934374 is strongly linked to rs12278256 and its RegulomeDB scores is lower than that of rs12278256, suggesting that it may be the functional site represented by rs12278256. Acting as promoter histone marks or enhancer histone marks, or affecting DNase is suggested to be associated with chromatin status, and binding proteins or altering regulatory motifs in CHIP-Seq suggest that transcription levels may be affected. It seems that these four SNPs may have certain effects on chromatin status and transcription level. Rs1892953 appears as an expression quantitative trait loci (eQTL) SNP in thyroid tissue [13].

## Discussion

In this group of Chinese Han adults, the relationship between two airway epithelial-related genes *EMSY* and *CDHR3* and risk of asthma were investigated, and four polymorphisms related to asthma susceptibility were obtained, which were rs3847076 of *CDHR3* and rs2508746, rs1892953 and rs12278256 of *EMSY*. A further subgroup analysis of these four variants revealed that their association with asthma was present in different subgroups.

*CDHR3*, located on chromosome 7, is specifically expressed in ciliated airway epithelial cells which are the targets of Rhinovirus C (RV-C) infection, and its expression was positively associated with RV-C binding, replication and entry into the host cells [19, 20]. There are only a few studies describing the relationship between *CDHR3* polymorphisms and asthma, and the results were inconsistent in different populations. The A allele of rs6967330 in *CDHR3* increased the risk of wheezing illnesses and hospitalizations for childhood asthma in a Danish study[6]. Rs17152490, in LD with rs6967330 was reported to affect asthma risk through *cis*-regulation of its gene expression in cells from human bronchial epithelial biopsy[21]. However, rs6967330 was only related to early-onset asthma in a Japanese population [22] and no association between rs6967330 and asthma was found in Chinese children[23]. In the present study, rs6967330 was not in HWE and our data suggest that rs3847076 may increase the risk of asthma in adults, which were inconsistent with the previous studies. The potential reasons for this discrepancy are as follows: Firstly, the susceptibility to asthma may differ in different populations, and secondly, late-onset asthma patients accounted for the majority of the case group in this study, in contrast to the above Japanese study which reported the positive relationship between rs6967330 and early-onset asthma in children. A future study of different asthma phenotypes would be beneficial to the accurate prevention and treatment of asthma.

Peripheral blood eosinophil were one of the main inflammatory cells involved in asthma and other allergic diseases [24]. Meta-analysis showed that the level of eosinophil in peripheral blood could better reflect the inflammatory status of eosinophil in airway[25], predicted the trend of long-term decline of lung function [26] and the risk of asthma attack in adults and children[27]. And more research is needed to determine whether rs3847076 genotypes of *CDHR3* are related to the number of eosinophils. Some studies have shown that the serum total IgE level is related to the severity and control of asthma[28]. The relationship between *CDHR3* variants and total IgE needed further investigated.

*EMSY*, located on chromosome 11q13.5, is expressed in the human airway epithelium and encoded by the EMSY protein. GWAS studies showed that *EMSY* was involved in allergic diseases including atopic dermatitis and food allergy[29, 30]. Several SNPs, rs7130588, rs10899234, rs6592657, as well as SNPs rs2508746 and rs1892953 that we studied were associated with total serum IgE levels in non-Hispanic Caucasian asthmatic patients [11]. In an eQTL analysis, Li et al [21] reported that rs2508740, rs2513525, rs4300410 (in complete LD with rs7926009), rs10793169 (in complete LD with rs7926009), rs2513513 and rs4245443 were significantly correlated with mRNA expression levels of *EMSY* in human bronchial alveolar lavage. Another GWAS study reported that rs7130588 in *EMSY* was associated with asthma[31]. A meta-analysis demonstrated that rs2155219 in *EMSY* increased the risk of allergic sensitization[12]. In the present study, three SNPs (rs2508746, rs1892953 and rs12278256) were related to asthma susceptibility in the Chinese Han population, of which rs12278256 has not been reported in previous studies. As a variant located in the upstream region of *EMSY*, rs12278256 might affect the regulatory motifs and chromatin status of this gene and further study is needed to verify this hypothesis. Based on our results, rs2508746, rs1892953 and rs12278256 genotypes were associated with level of FEV1% predicted and/or FEV1/FVC%, which also suggested that gene *EMSY* was likely related with lung function.

Studies in the twin population have shown that susceptibility to asthma can be attributed to genetic factors [3, 4]. Although current genome-wide association studies have identified numerous polymorphisms associated with asthma susceptibility, the odds ratio (OR) is around 1.2, and only a small percentage of asthma prevalence can be contributed to them. Some experts have proposed to study the interaction between genes and environment[32, 33]. It is well known that environmental factors such as smoking and obesity are susceptibility factors for asthma, but the specific mechanism is not clear. A number of studies have shown that smoking is associated with increased risk of asthma, reduced efficacy of inhaled corticosteroids treatment, acute exacerbations, and airway remodeling in asthma[34-38]. Mechanisms of asthma in the obese may include mechanical factors and inflammatory immunity [39]. Studies have shown that the SNPs at 17q21.2 is associated with BMI levels in asthmatic patients[40]. Functional prediction suggests that the alternate A allele of rs3847076 decrease the effect on motif TCF4 relative to the reference C allele, according to the library [41].

Recently, genetic studies have detected a lot of susceptibility genes for asthma. This study was the first attempt to investigate the association between *CDHR3* and *EMSY* and adult asthma susceptibility in the Chinese Han population. We found rs3847076 in *CDHR3*, rs2508746, rs1892953 and rs12278256 in

*EMSY* were associated with the risk of adult asthma. However, there were some limitations to this study. Adjustment was not performed to correct the results for multiple testing, due to the weak effect of each single polymorphism on asthma susceptibility. In addition, the allergic phenotypes of the asthma patients were not clear and serum IgE levels were not analyzed in the study. Lastly, *CDHR3* is a huge gene spanning over 159kb and the strategy of tag-SNPs selection with  $r^2 > 0.64$  in this study may have missed some SNPs associated with the disease.

## Conclusions

In conclusion, this study is the first to identify that the airway epithelium related genes *EMSY* and *CDHR3* were associated with adult asthma susceptibility in the Chinese Han population. The *CDHR3* rs3847076 allele A and *EMSY* rs1892953 genotype GG may increase the risk of asthma. The *EMSY* rs2508746 and rs12278256 allele T may decrease asthma risk. A population with a larger sample size is needed for further exploration of the association.

## Abbreviations

Tag-SNPs Tag-single nucleotide polymorphisms

GWAS Genome-Wide Association Study

*CDHR3* Cadherin related family member 3

*C11orf30* Chromosome 11 open reading frame 30

SNPs Single Nucleotide Polymorphisms

MAF Minor Allele Frequency

PCR Polymerase Chain Reaction

SPSS Statistical Package for the Social Sciences

HWE Hardy-Weinberg equilibrium

LD Linkage Disequilibrium

BMI Body Mass Index

RV-C Rhinovirus C

IgE Immunoglobulin E

eQTL expression Quantitative Trait Loci

## Declarations

### Ethics approval and consent to participate

All protocols for this study were reviewed and approved by the Institutional Review Board of the West China Hospital of Sichuan University (Protocol No. 23). Written informed consent was obtained from all the study participants.

### Consent for publication

Not applicable.

### Availability of data and materials

Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

### Competing interests

The Authors declare that there is no conflict of interest.

### Funding

The design of the study, sample collection, Genomic DNA extraction, SNP selection and genotyping were supported by the National Natural Science Foundation of China [Grant No. 81370121] and the Health and Family Planning Commission of Sichuan Province Project [Grant No. 16PJ413]. The analysis and interpretation of data and manuscript writing were supported by the Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital Project [Grant No. 2017QN11] and the Sichuan Provincial Cadre Health Research Project [Grant No. 2018-211].

### Authors' Contributions

Conceptualization, JQH; Data and Formal analysis, YW and SQW; Project administration, MMZ, Guo Chen and JQH; Supervision, JQH; Writing – original draft, MMZ and GC; Writing – review & editing, AJS. All authors read and approved the final manuscript.

### Acknowledgements

We thank everyone who provided blood samples and consent for genetic analysis. And we thank all of the clinicians, nurses and study coordinators for their contributions to the work.

### Contributor Information

Miao-miao Zhang, Email: zhangmiaomiao0522@126.com

Guo Chen, Email: [5503850@163.com](mailto:5503850@163.com)

Yu Wang, Email: [390628840@qq.com](mailto:390628840@qq.com)

Shou-Quan Wu, Email: [76201313@qq.com](mailto:76201313@qq.com)

Andrew J Sandford, Email: [Andrew.Sandford@hli.ubc.ca](mailto:Andrew.Sandford@hli.ubc.ca)

Jian-Qing He, Email: [jianqing\\_he@scu.edu.cn](mailto:jianqing_he@scu.edu.cn) ; [jianqhe@gmail.com](mailto:jianqhe@gmail.com)

## References

1. WHO. Global surveillance, prevention and control of chronic respiratory diseases: WHO; 2007 [[http://www.who.int/respiratory/publications/global\\_surveillance/en/](http://www.who.int/respiratory/publications/global_surveillance/en/) (accessed April 23, 2018)]. 2007.
2. Lin J, Wang W, Chen P, et al. Prevalence and risk factors of asthma in mainland China: The CARE study. *Respir Med*. 2018;137:48-54.
3. Duffy DL, Martin NG, Battistutta D, et al. Genetics of asthma and hay fever in Australian twins. *Am Rev Respir Dis*. 1990;142(6 Pt 1):1351-8.
4. van Beijsterveldt CE, Boomsma DI. Genetics of parentally reported asthma, eczema and rhinitis in 5-year-old twins. *Eur Respir J*. 2007;29(3):516-21.
5. Lambrecht BN, Hammad H. The airway epithelium in asthma. *Nat Med*. 2012;18(5):684-92.
6. Bonnelykke K, Sleiman P, Nielsen K, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet*. 2014;46(1):51-5.
7. Fagerberg L, Hallström BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*. 2014;13(2):397-406.
8. Asai Y, Eslami A, van Ginkel CD, et al. Genome-wide association study and meta-analysis in multiple populations identifies new loci for peanut allergy and establishes C11orf30/EMSY as a genetic risk factor for food allergy. *J Allergy Clin Immunol*. 2018;141(3):991-1001.
9. Varier RA, Carrillo de Santa Pau E, van der Groep P, et al. Recruitment of the Mammalian Histone-modifying EMSY Complex to Target Genes Is Regulated by ZNF131. *J Biol Chem*. 2016;291(14):7313-24.
10. Reddel HK, Bateman ED, Becker A, et al. A summary of the new GINA strategy: a roadmap to asthma control. *Eur Respir J*. 2015;46(3):622-39.
11. Li X, Ampleford EJ, Howard TD, et al. The C11orf30-LRRC32 region is associated with total serum IgE levels in asthmatic patients. *J Allergy Clin Immunol*. 2012;129(2):575-8, 8 e1-9.

12. Bonnelykke K, Matheson MC, Pers TH, et al. Meta-analysis of genome-wide association studies identifies ten loci influencing allergic sensitization. *Nat Genet.* 2013;45(8):902-6.
13. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015;348(6235):648-60.
14. Weidinger S, Willis-Owen SA, Kamatani Y, et al. A genome-wide association study of atopic dermatitis identifies loci with overlapping effects on asthma and psoriasis. *Hum Mol Genet.* 2013;22(23):4841-56.
15. Du W, Cheng J, Ding H, et al. A rapid method for simultaneous multi-gene mutation screening in children with nonsyndromic hearing loss. *Genomics.* 2014;104(4):264-70.
16. Andersson T, Alfredsson L, Källberg H, et al. Calculating measures of biological interaction. *Eur J Epidemiol.* 2005;20(7):575-9.
17. Global Initiative for asthma. Global Strategy for Asthma Management and Prevention, 2018. Available from: [www.ginasthma.org](http://www.ginasthma.org).
18. Department of Disease Control, people's Republic of China. Guidelines for Prevention and Control of overweight and Obesity in Chinese Adults. 2003.
19. Bochkov YA, Watters K, Ashraf S, et al. Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication. *Proc Natl Acad Sci U S A.* 2015;112(17):5485-90.
20. Griggs TF, Bochkov YA, Basnet S, et al. Rhinovirus C targets ciliated airway epithelial cells. *Respir Res.* 2017;18(1):84.
21. Li X, Hastie AT, Hawkins GA, et al. eQTL of bronchial epithelial cells and bronchial alveolar lavage deciphers GWAS-identified asthma genes. *Allergy.* 2015;70(10):1309-18.
22. Kanazawa J, Masuko H, Yatagai Y, et al. Genetic association of the functional CDHR3 genotype with early-onset adult asthma in Japanese populations. *Allergol Int.* 2017;66(4):563-7.
23. Chen J, Zhang J, Hu H, et al. Polymorphisms of RAD50, IL33 and IL1RL1 are associated with atopic asthma in Chinese population. *Tissue Antigens.* 2015;86(6):443-7.
24. Matucci A, Vultaggio A, Maggi E, et al. Is IgE or eosinophils the key player in allergic asthma pathogenesis? Are we asking the right question? *Respir Res.* 2018;19(1):113.
25. Korevaar DA, Westerhof GA, Wang J, et al. Diagnostic accuracy of minimally invasive markers for detection of airway eosinophilia in asthma: a systematic review and meta-analysis. *The Lancet Respiratory medicine.* 2015;3(4):290-300.
26. Hancox RJ, Pavord ID, Sears MR. Associations between blood eosinophils and decline in lung function among adults with and without asthma. *Eur Respir J.* 2018;51(4).
27. Price DB, Rigazio A, Campbell JD, et al. Blood eosinophil count and prospective annual asthma disease burden: a UK cohort study. *The Lancet Respiratory medicine.* 2015;3(11):849-58.
28. Tanaka A, Jinno M, Hirai K, et al. Longitudinal increase in total IgE levels in patients with adult asthma: an association with poor asthma control. *Respir Res.* 2014;15:144.

29. Marenholz I, Grosche S, Kalb B, et al. Genome-wide association study identifies the SERPINB gene cluster as a susceptibility locus for food allergy. *Nat Commun.* 2017;8(1):1056.
30. Esparza-Gordillo J, Weidinger S, Folster-Holst R, et al. A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat Genet.* 2009;41(5):596-601.
31. Ferreira MA, Matheson MC, Duffy DL, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet.* 2011;378(9795):1006-14.
32. Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med.* 2010;363(13):1211-21.
33. Ober C. Asthma Genetics in the Post-GWAS Era. *Annals of the American Thoracic Society.* 2016;13 Suppl 1:S85-90.
34. Nakamura K, Nagata C, Fujii K, et al. Cigarette smoking and the adult onset of bronchial asthma in Japanese men and women. *Ann Allergy Asthma Immunol.* 2009;102(4):288-93.
35. Shimoda T, Obase Y, Kishikawa R, et al. Influence of cigarette smoking on airway inflammation and inhaled corticosteroid treatment in patients with asthma. *Allergy Asthma Proc.* 2016;37(4):50-8.
36. Heijink I, van Oosterhout A, Kliphuis N, et al. Oxidant-induced corticosteroid unresponsiveness in human bronchial epithelial cells. *Thorax.* 2014;69(1):5-13.
37. Silverman RA, Boudreaux ED, Woodruff PG, et al. Cigarette smoking among asthmatic adults presenting to 64 emergency departments. *Chest.* 2003;123(5):1472-9.
38. Fattahi F, Hylkema MN, Melgert BN, et al. Smoking and nonsmoking asthma: differences in clinical outcome and pathogenesis. *Expert Rev Respir Med.* 2011;5(1):93-105.
39. Dixon AE, Holguin F, Sood A, et al. An official American Thoracic Society Workshop report: obesity and asthma. *Proc Am Thorac Soc.* 2010;7(5):325-35.
40. Wang L, Murk W, DeWan AT. Genome-Wide Gene by Environment Interaction Analysis Identifies Common SNPs at 17q21.2 that Are Associated with Increased Body Mass Index Only among Asthmatics. *PLoS One.* 2015;10(12):e0144114.
41. Kheradpour P, Kellis M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res.* 2014;42(5):2976-87.

## Tables

Table 1 Characteristics of cases and controls

Characteristic	Control(n%)	Case (n%)	P value
Gender			0.876
Male	162(38.76%)	118(39.33%)	
Female	256(61.24%)	182(60.67%)	
Age (mean±SD,years)	44.09 ± 13.75	43.6 ± 13.48	0.64
Smoking status			
Current and ex-smokers	55(13.16%)	49(16.33%)	0.179
Non-smoking	207(49.52%)	247(82.33%)	
Smoking status unclear	156(37.32%)	4(1.33%)	
BMI(mean±SD)	22.94 ± 3.34	23.11 ± 3.28	0.517
BMI<24	227(54.31%)	197(65.67%)	
BMI≥24	121(29.67%)	103(34.33%)	
Types of patients			
Emergency patients or inpatients		34(11.33%)	
Outpatients		266(88.67%)	
Asthma onset time			
Early-onset asthma(<18 years old)		42(14.00%)	
Late-onset asthma(≥18 years old)		223(74.33%)	
Onset time unclear		35(11.67%)	
Eosinophil count		171(57.00%)	
Total IgE		139(46.33%)	
Asthma with pulmonary function test		174(58.00%)	
FEV1% predicted (mean±SD)		83.61 ± 19.97	
FEV1/FVC%(mean±SD)		72.37 ± 13.63	
Provocation test or relaxation test		164(54.67%)	
Positive provocation test or relaxation test		127(42.33%)	
Treatment scheme			
Step 3 treatment		10(3.33%)	
Step 4 treatment		175(58.33%)	
Step 5 treatment		38(12.67%)	

Values are means  $\pm$  standard deviation (SD) and absolute numbers (percentages). BMI, body mass index; Early-onset asthma, age of asthma onset <18years; Late-onset asthma, age of asthma onset  $\geq$ 18years; FEV1, forced expiratory volume in one second; FVC, forced vital capacity.

Table 2 The four SNPs associated with asthma

Genes	SNPs	Genetic models	Genotypes	Control n(%)	Case n(%)	<i>P</i> *	OR 95%CI*
<i>CDHR3</i>	rs3847076	Dom	CC	285(68.2)	185(61.7)	0.081	1.378(0.962-1.973)
			CA+AA	133(31.8)	115(38.3)		
		Rec	CC+CA	408(97.6)	285(95.0)	0.060	2.689(0.958-7.545)
			AA	10(2.4)	15(5.0)		
		Add	CC/CA/AA			<b>0.032*</b>	<b>1.407(1.030-1.923)*</b>
<i>EMSY</i>	rs2508746	Dom	CC	244(58.4)	197(65.7)	<b>0.019*</b>	<b>0.660(0.465-0.935)*</b>
			TC+TT	174(41.6)	103(34.3)		
		Rec	CC+TC	396(94.7)	288(96.0)	0.445	0.733(0.331-1.626)
			TT	22(5.3)	12(4.0)		
		Add	CC/TC/TT			<b>0.026*</b>	<b>0.718(0.536-0.961)*</b>
<i>EMSY</i>	rs1892953	Dom	AA	115(27.5)	76(25.3)	0.647	1.094(0.745-1.605)
			GA+GG	303(72.5)	224(74.7)		
		Rec	AA+GA	319(76.3)	219(73.0)	<b>0.015*</b>	<b>1.667(1.104-2.518)*</b>
			GG	99(23.7)	81(27.0)		
		Add	AA/GA/GG			0.081	1.240(0.974-1.579)
<i>EMSY</i>	rs12278256	Dom	GG	357(85.4)	272(90.7)	<b>0.033*</b>	<b>0.563(0.332-0.953)*</b>
			TG+TT	61(14.6)	28(9.3)		
		Rec	GG+TG	417(99.8)	300(100)	1	-
			TT	1(0.2)	0(0)		
		Add	GG/TG/TT			<b>0.027*</b>	<b>0.558(0.332-0.937)*</b>

\* Adjusted for sex, age, body mass index and smoking history with logistic regression,  $P < 0.05$ . Add: additive model; Dom: dominant model; Rec: recessive model.

Table 3 Results of stratification analysis based on gender, smoking status, BMI status, and onset age of asthma

SNPs	Genetic models	Stratified by gender		Stratified by smoking status		Stratified by BMI status		Stratified by onset age of asthma					
		<i>P</i>	OR 95%CI	<i>P</i>	OR 95%CI	<i>P</i>	OR 95%CI	<i>P</i>	OR 95%CI				
rs3847076	Dom	male	0.048*	1.834(1.005-3.347)	smoking	0.018*	2.252(1.149-4.413)	BMI $\geq$ 24	0.004*	1.925(1.233-3.007)	late onset	0.063	1.428(0.981-2.077)
	Rec		0.115	5.656(0.654-48.882)		0.09	4.222(0.799-22.320)		0.12	2.835(0.761-10.559)	asthma	0.049*	2.861(1.006-8.134)
	Add		0.023*	1.869(1.091-3.202)		0.009*	2.168(1.212-3.872)		0.005*	1.835(1.234-2.726)		0.023*	1.457(1.054-2.013)
rs2508746	Dom	female	-	-	non-smoking	0.014*	0.618(0.420-0.908)	BMI $\geq$ 24	0.027*	0.612(0.396-0.946)	late onset	0.016*	0.637(0.441-0.919)
	Rec		-	-		0.498	0.737(0.304-1.782)		0.862	0.920(0.361-2.347)	asthma	0.419	0.706(0.304-1.641)
	Add		-	-		0.022*	0.685(0.495-0.947)		0.06	0.710(0.496-1.015)		0.021*	0.696(0.511-0.948)
rs1892953	Dom	female	0.548	1.159(0.717-1.873)	non-smoking	0.456	0.174(0.770-1.790)	BMI $\geq$ 24	0.702	1.097(0.682-1.766)	late onset	0.692	1.084(0.726-1.620)
	Rec		0.038*	1.738(1.031-2.927)		0.04*	1.615(1.021-2.553)		0.017*	1.910(1.123-3.250)	asthma	0.017*	1.680(1.095-2.578)
	Add		0.108	1.282(0.947-1.737)		0.091	1.259(0.964-1.644)		0.096	1.297(0.955-1.761)		0.094	1.241(0.964-1.599)
rs12278256	Dom	female	0.037*	0.468(0.229-0.955)	non-smoking	0.023*	0.512(0.287-0.913)	BMI $\geq$ 24	0.033*	0.485(0.249-0.944)	late onset	-	-
	Rec		1			1			1		asthma	-	-
	Add		0.032*	0.465(0.231-		0.02*	0.508(0.288-		0.028*	0.481(0.250-		-	-

0.936)

0.897)

0.923)

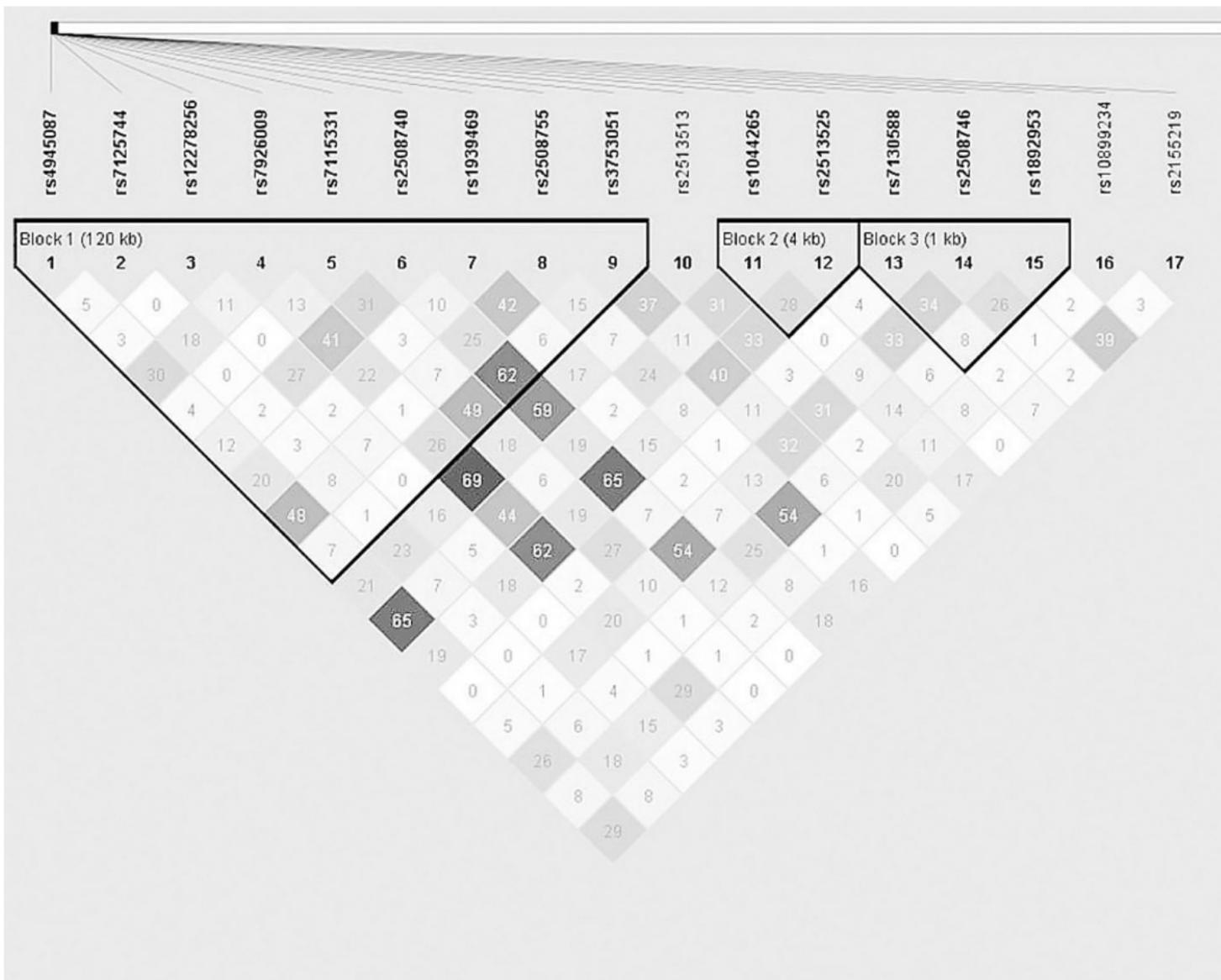
\* Adjusted for sex, age, body mass index and smoking history with logistic regression,  $P < 0.05$ . Add: additive model; Dom: dominant model; Rec: recessive model.

Table 4 The association between *EMSY* haplotypes in block 1 and asthma susceptibility

Haplotype	Case N (%)	Control N (%)	Chi <sup>2</sup>	Pearson's p	OR (95% CI)
AAGTTAAAT	207.00(0.345)	285.58(0.342)	0	0.999	1.000(0.801-1.248)
GAGCGGAGC	43.00(0.072)	57.59(0.069)	<b>0.023</b>	<b>0.878</b>	<b>1.033(0.685-1.556)</b>
GAGCTAAAT	44.00(0.073)	58.04(0.069)	0.054	0.815	1.050(0.699-1.577)
GAGCTGAGC	35.00(0.058)	52.18(0.062)	0.134	0.714	0.921(0.592-1.432)
GAGTTAGGT	178.00(0.297)	230.22(0.275)	0.589	0.443	1.095(0.868-1.382)
GATCTGAGT	28.00(0.047)	61.00(0.073)	4.346	0.037*	0.615(0.388-0.975)*
GGGCTAAAT	60.00(0.100)	76.32(0.091)	0.246	0.62	1.094(0.766-1.562)
Global result	600	836	4.912565	0.555	

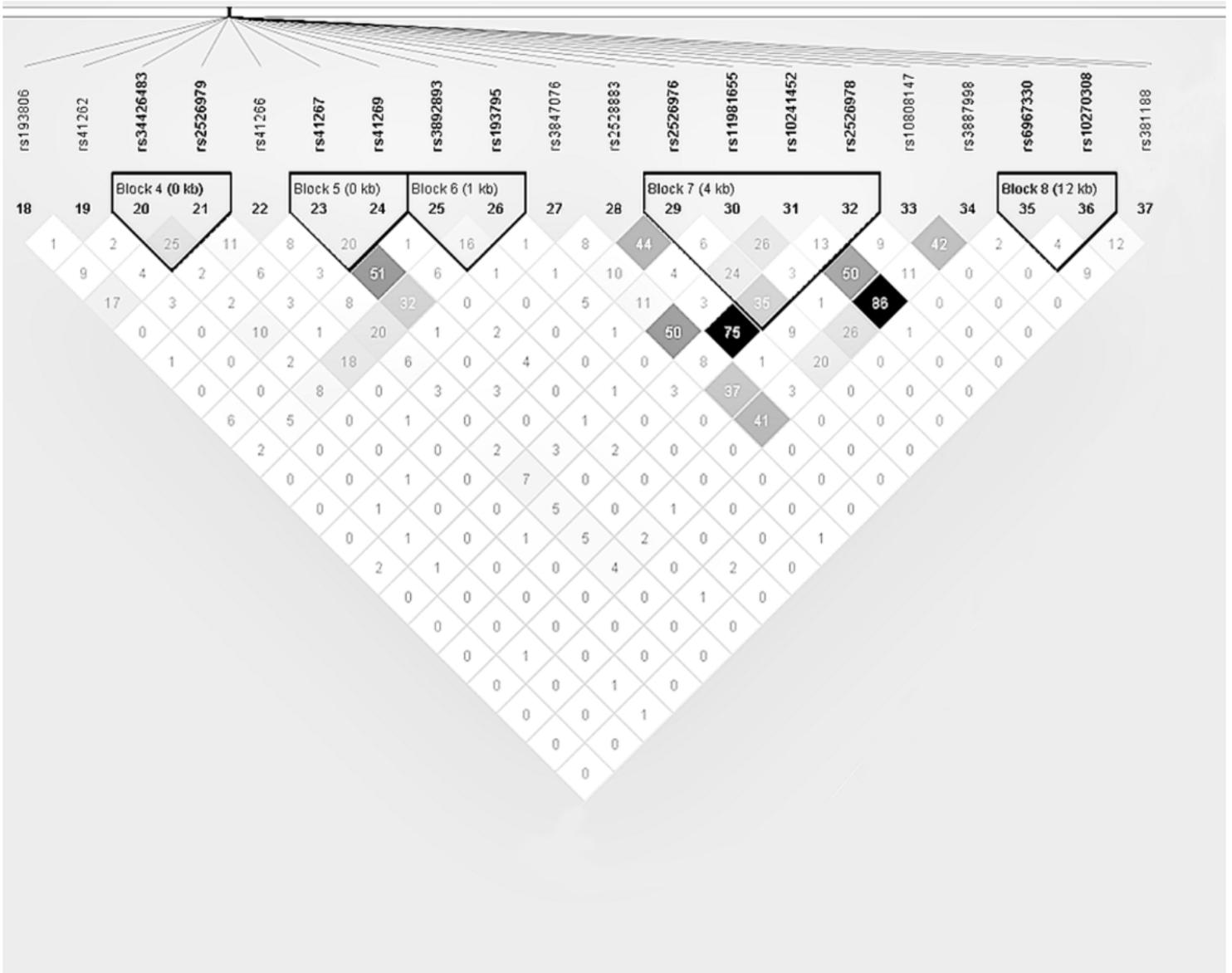
For each haplotype, alleles were arranged in order of rs4945087, rs7125744, rs12278256, rs7926009, rs7115331, rs2508740, rs1939469, rs2508755 rs3753051.

## Figures



**Figure 1**

Analysis of linkage disequilibrium of 17 SNPs in EMSY. Note: Each square represents the linkage disequilibrium of two corresponding SNPs, which is displayed as  $r^2 \times 100$ . The larger the darkness of the square, the larger the value of  $r^2 \times 100$ .



**Figure 2**

Analysis of linkage disequilibrium of tag-SNPs in CDHR3 Note: Each square represents the linkage disequilibrium of two corresponding SNPs, which is displayed as  $r^2 \times 100$ . The larger the darkness of the square, the larger the value of  $r^2 \times 100$ .

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterialsR2.doc](#)