

DR TING FAN LEUNG (Orcid ID : 0000-0002-6469-1926)

DR GARY WK WONG (Orcid ID : 0000-0001-5939-812X)

Article type : Original

Cadherin-related Family Member 3 Gene Impacts Childhood Asthma in Chinese Children

Ting Fan Leung^{1,2}, Man Fung Tang^{1,2}, Agnes Sze Yin Leung¹, Alice Pik Shan Kong³, Tak Chi Liu¹, Renee Wan Yi Chan^{1,2}, Ronald Ching Wan Ma³, Hing Yee Sy¹, Juliana Chung Ngor Chan³ and Gary Wing Kin Wong¹

Department of ¹Paediatrics and ³Medicine and Therapeutics, and ²Hong Kong Hub of Paediatric Excellence, The Chinese University of Hong Kong, Hong Kong

Running head: *CDHR3* and childhood wheezing and lung function

Word counts: 3,045 (manuscript) and 259 (abstract)

Correspondence:

Ting F. Leung, Room 84043, Lui Che Woo Clinical Sciences Building, Prince of Wales Hospital, Shatin, N.T., Hong Kong

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/PAI.13138](#)

This article is protected by copyright. All rights reserved

31 Tel: (852) 3505 2981
32 Fax: (852) 2636 0020
33 E-mail: tfleung@cuhk.edu.hk

34 **Abstract**

35 **Background** A missense variant (rs6967330) of the gene encoding cadherin-related family
36 member 3 (*CDHR3*) was associated with recurrent severe exacerbations in preschoolers. However,
37 there was limited data on its relationship with preschool lung function and school-age asthma.
38 This study replicated the association between polymorphic markers at the region of *CDHR3*
39 around rs6967330 and wheezing phenotypes in two independent cohorts of Chinese children.

40 **Methods** Ten tagging SNPs located 10-kb around rs6967330 were selected by HaploView
41 5.0 based on 1000 Genomes database for Southern Han Chinese. Their associations with wheezing
42 and lung function were examined in 1341 Chinese preschool children while those for asthma
43 phenotypes were examined in an independent group of 2079 school-age children. Genotypic and
44 haplotypic associations were analyzed by multivariate regression, and generalized multifactor
45 dimensionality reduction was used to examine epistatic interactions for wheezing traits.

46 **Results** The mean (SD) age of preschool cohort was 4.7 (1.0) years. Rs6967330 was
47 associated with current wheeze (odds ratio [OR] 1.63, 95% confidence interval [CI] 1.09-2.43)
48 and its severity (OR 1.64, 95% CI 1.10-2.44) among preschool children. This SNP was also
49 associated with school-age asthma (OR 1.32, 95% CI 1.04-1.69). The minor allele of rs408223
50 was associated with lower FEV_{0.5} (β =-2.411, P =0.004) and FEV_{0.5}/FVC (β =-1.292, P =0.015).
51 Lower spirometric indices were also associated with minor allele of rs140154310. GAC haplotype
52 from rs4730125, rs6967330 and rs408223 was associated with preschool current wheeze and
53 school-age asthma. Epistatic interaction was found between unrelated *CDHR3* SNPs for FEV_{0.5}
54 among preschoolers.

55 **Conclusion** *CDHR3* is a candidate gene for early-life wheezing, school-age asthma and lung
56 function in Chinese children.

57
58 **Keywords:** *CDHR3*; rhinovirus C receptor; incentive spirometry; preschoolers; wheeze

63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94

Introduction

Asthma is characterized by airway hyperresponsiveness accompanied by episodic and reversible airflow limitation and respiratory symptoms, which results from complex interactions between many genetic and environmental factors.¹ A number of candidate genes were reported for asthma diagnosis and subphenotypes², but none of them targeted asthma exacerbations as the main outcome. A genome-wide association study (GWAS) revealed significant association between a single-nucleotide polymorphism (SNP; rs6967330) of the gene encoding cadherin-related family member 3 (*CDHR3*) and recurrent severe exacerbations in Danish preschoolers.³ *CDHR3* is a member of a cadherin family of transmembrane proteins that is highly expressed in the human lung and airway epithelium.^{4,5} This top variant is a nonsynonymous coding SNP that causes an amino acid change from cysteine to tyrosine at position 529. Subsequently, experimental evidence identified *CDHR3* as a receptor for human rhinovirus-C (HRV-C) that mediates entry of this virus into host cell.⁶ Transfection of cells with the risk variant of rs6967330 resulted in 10-fold increases

in HRV-C binding and progeny yields compared with wild type variant, suggesting the GWAS association with this top SNP of *CDHR3* might be due to the increased susceptibility to HRV-C infections. Early-life respiratory virus infections, predominantly those caused by HRV, are a well-known risk factor for subsequent childhood wheezing and asthma development.⁷⁻⁹ However, the roles of *CDHR3* variants in early-life wheezing and school-age asthma as well as preschool lung function measured objectively by spirometry remain unclear. This study aimed to replicate the relationship between *CDHR3* variants and early-life wheezing and forced expiratory indices in Chinese preschoolers and further examined the relationship between *CDHR3* and asthma phenotypes in an independent cohort of Chinese school-age children. This study also examined any epistatic interaction between unrelated SNPs of *CDHR3* for different wheezing phenotypes.

Patients and Methods

Study populations

Two independent cohorts were recruited to investigate the relationship between *CDHR3* variants and different childhood wheezing and lung function phenotypes. Our preschool cohort consisted of unrelated Chinese children aged 2-7 years from local nurseries and kindergartens who contributed to our spirometric references for preschoolers.^{10,11} Their history of wheezing and allergy status was ascertained using validated Chinese International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire.¹² The severity of wheezing was defined according to the frequency of current wheeze. Subjects in our school-age cohort were recruited from both pediatric clinics and local schools.¹³⁻¹⁵ Cases in this cohort were unrelated Chinese children aged 5-18 years who had physician-diagnosed asthma according to American Thoracic Society criteria.¹ Controls consisted of children who attended our clinics for minor non-respiratory and non-inflammatory illnesses or from the above school-based community studies for allergic diseases, who did not report any history of asthma, eczema or allergic rhinitis by our ISAAC questionnaire. All subjects were free from respiratory tract infection for at least four weeks as reported by self or parent before study. The Clinical Research Ethics Committee of our University approved these studies, and subjects and/or their parents gave informed written consent.

Assessment of atopy

127 Plasma total IgE concentrations were measured by microparticle immunoassay (IMx analyzer,
128 Abbott Laboratories, Abbott Park, IL). Allergen-specific IgE to *Dermatophagoides pteronyssinus*,
129 cat, dog, cockroach and mixed moulds were measured by fluorescent enzyme immunoassay
130 (AutoCAP; Phadia AB, Uppsala, Sweden), with levels ≥ 0.35 kIU/l being positive.¹⁷⁻¹⁹ Skin prick
131 tests (SPTs) were performed with diluent (negative control), histamine (10 mg/ml; positive control)
132 and standardized extracts of the same panel of five locally prevalent allergens (ALK Abelló,
133 Round Rock, TX).²⁰ Atopy was defined as subjects with at least one positive SPT or
134 allergen-specific IgE result.

135

136 **Spirometric measurements**

137 Preschool children provided valid readings of forced expiratory volume in 0.5-second (FEV_{0.5})
138 and forced vital capacity (FVC) by incentive spirometry (Master Screen, Jaeger GmbH, Würzburg,
139 Germany) according to international guideline.² Asthmatics in the school-age cohort performed
140 pre-bronchodilator spirometry (COMPACT II; Vitalograph, Buckingham, UK). Results were
141 compared with age-specific local references.^{10,22}

142

143 **SNP selection and genotyping**

144 Tagging SNPs were selected by the tagging algorithm³ implemented in HaploView 5.0 (Daly Lab,
145 Cambridge, MA) based on 1000 Genomes database search for Southern Han Chinese (CHS) at
146 pairwise $r^2 \geq 0.8$ for linkage disequilibrium (LD) for all SNPs located 5-kb both upstream and
147 downstream from the top GWAS SNP rs6967330³ with minor allele frequencies (MAFs) ≥ 0.01 .
148 Rs6967330 was forced to be included in the tagging process. Table S1 summarizes the details of
149 our selected SNPs. Genomic DNA was extracted from buccal cells in preschool children and
150 peripheral blood leucocytes in school-age children. All SNPs were genotyped by TaqMan
151 genotyping assays (Applied Biosystems, Foster City, CA) on QuantStudio 12K Flex real-time
152 PCR system (Applied Biosystems).

153

154 **Statistical analysis**

155 Subjects' clinical and laboratory characteristics were compared by student t-test for numerical data
156 and Pearson χ^2 or Fisher exact test for categorical variables. Allele frequencies for all SNPs were
157 estimated by the gene-counting method, and Hardy-Weinberg equilibrium (HWE) was examined
158 by χ^2 exact test. The associations between SNPs and clinical phenotypes were analyzed by

multivariable regression (see online supplement). All comparisons were made two-tailed using SPSS v.21 (IBM Corp., Armonk, NY). False discovery rate (FDR) was applied separately to each phenotype to correct for multiple testing. The method of predictive margins was used to assess the probability of falling into a corresponding cluster of phenotype according to the allele of each significant SNP. Pairwise linkage disequilibrium (LD) coefficient was calculated for each SNP pair by Haploview 5.0 (Daly Lab). The haplotypes were determined by Haploview 5.0, and their associations with wheezing traits were tested by haplo.stats 1.7.9 on R v.3.5.0 (<http://www.r-project.org/>).⁴ The haplotypic associations with dichotomous and continuous variables were analyzed by multivariate logistic and linear regression respectively. Generalized multifactor dimensionality reduction (GMDR, v.0.9; <http://www.healthsystem.virginia.edu/internet/addiction-genomics/software/>) was used to test for interactions among *CDHR3* SNPs from different LD blocks in our preschool cohort (see online supplement). The best model was chosen based on cross-validation consistency (CVC) $\geq 8/10$ and testing accuracy (TA) score ≥ 0.55 . Statistical significance was derived empirically from 5000 permutations before selecting the best performing models. Pearson χ^2 , standard analysis of variance and student t-tests were then used to compare the wheezing phenotypes and spirometric indices among different groups assigned by GMDR.

***In silico* analysis**

In silico analysis was performed to evaluate potential biological functions of seven intronic SNPs genotyped (rs3887998, rs140154310, rs73195657, rs146004234, rs4730125, rs73195665, and rs408223) on *CDHR3* mRNA splicing using Human Splicing Finder (HSF) (v. 3.1) (<http://www.umd.be/HSF/>).

Results

Study populations

The preschool cohort consisted of 1341 subjects whose mean (SD) age was 4.7 (1.0) years. Table S2 summarizes their demographic and clinical characteristics. The prevalence of asthma and wheezing ever in this cohort was 5.3% and 16.9%, respectively. The school-age cohort consisted of 892 asthmatics and 1187 non-allergic controls (Table S3), and the former group had higher

191 plasma total IgE level (mean [SD] 2.61 [0.69] vs 1.86 [0.69]; $P < 0.0001$). Atopy was detected in
192 75.3% of cases and 38.0% of controls ($P < 0.0001$).
193

194 **Genotypic associations for asthma traits**

195 Rs448025 and rs543085868, being monomorphic (Table S1), were excluded from analyses. The
196 genotyping efficiency for *CDHR3* SNPs was $\geq 95\%$ in both cohorts, and all of them followed
197 HWE (see footnote in Table S1). Table 1 summarizes the genetic associations between eight SNPs
198 and wheezing phenotypes in preschool children. Rs6967330, the top GWAS SNP, was associated
199 with current wheeze among preschoolers (odds ratio [OR] 1.63 and 95% confidence interval [CI]
200 1.09-2.43; $P=0.017$). This SNP was also associated with severity of current wheeze (OR 1.64 and
201 95% CI 1.10-2.44; $P=0.016$). Another tagging SNP rs140154310 was also associated with current
202 wheeze (OR 2.20 and 95% CI 1.16-4.19; $P=0.016$) and its severity (OR 2.17 and 95% CI
203 1.15-4.12; $P=0.018$).

204 In our school-age cohort (Table 2), we found significant association between asthma
205 diagnosis and rs6967330 under both additive model (OR 1.32 and 95% CI 1.04-1.69; $P=0.025$)
206 and dominant model (OR 1.34 and 95% CI 1.03-1.73; $P=0.027$). Among asthmatics who received
207 inhaled corticosteroid (ICS) treatment (i.e. more severe disease; Table S4), asthma was also
208 associated with this same SNP under additive model (OR 1.52 and 95% CI 1.05-2.20; $P=0.027$)
209 and dominant model (OR 1.60 and 95% CI 1.09-2.35; $P=0.016$). Another tagging SNP, rs4730125,
210 was strongly associated with asthma diagnosis among these ICS-treated patients under additive
211 model (OR 1.34 and 95% CI 1.08-1.67; $P=0.008$) and dominant model (OR 1.66 and 95% CI
212 1.21-2.27; $P=0.002$).

213 When analyzed by the method of predictive margins, carriers of the minor alleles at
214 rs6967330, rs140154310 and rs4730125 were associated with higher probability to have current
215 wheeze, asthma, and/or asthma requiring ICS treatment (Fig. 1). On average, carrying one more
216 minor allele of the SNPs rs6967330, rs140154310 and rs4730125 would increase 7% of the
217 chance developing current wheeze, asthma, and/or asthma requiring ICS treatment. None of the
218 *CDHR3* SNPs was associated with atopy and plasma total IgE concentration among our
219 schoolchildren.
220

221 **Genotypic association for spirometric indices**

222 Table 3 summarizes the genotypic associations for spirometric indices in our preschool cohort.

The minor allele of rs408223 was associated with lower FEV_{0.5} ($\beta=-2.411$, $P=0.004$) and FEV_{0.5}/FVC ($\beta=-1.292$, $P=0.015$). Lower spirometric indices were also associated with the minor allele of rs140154310. Besides these two SNPs, the minor allele of rs73195657 was associated with higher FVC ($\beta=2.733$, $P=0.030$). None of these *CDHR3* SNPs was associated with spirometric indices in our school-age cohort.

Haplotypic associations for wheezing and lung function phenotypes

Fig. S1 describes LD pattern for *CDHR3* SNPs, which revealed two haplotype blocks. In our preschool cohort (Table S5), current wheeze was associated with an at-risk haplotype GAC formed by rs4730125, rs6967330 and rs408223 (OR 1.54 and 95% CI 0.99-2.39; $P=0.048$). Among school-age children, the same at-risk haplotype GAC was also associated with asthma (OR 1.39 and 95% CI 1.08-1.80; $P=0.012$) and asthma requiring ICS treatment (OR 1.63 and 95% CI 1.07-2.48; $P=0.023$) (Table S5). Furthermore, ICS-treated asthma was associated with TGC haplotype formed by rs4730125, rs6967330 and rs408223 (OR 1.41 and 95% CI 1.12-1.79; $P=0.004$). Haplotypic association for school-age asthma requiring ICS treatment remained significant following the removal of rs6967330 (Table S6).

Tables S7 summarizes the haplotypic associations for spirometric parameters in the preschool cohort. The GGG haplotype from rs4730125, rs6967330 and rs408223 was associated with reduced FEV_{0.5} ($\beta=-1.993$, $P=0.027$) and FEV_{0.5}/FVC ($\beta=-1.464$, $P=0.010$). FVC was also associated with AC haplotype from rs3887998 and rs73195657 ($\beta=2.619$, $P=0.042$). Following the removal of rs6967330, the 2-locus haplotype GG remained to be significantly associated with FEV_{0.5} and FEV_{0.5}/FVC (Table S8). Although none of the SNPs was associated with lung function of school-age children, we found significant haplotypic associations between GGG haplotype from rs4730125, rs6967330 and rs408223 and FEV₁ ($\beta=-4.308$, $P=0.039$) and between AT haplotype and FEV₁ ($\beta=4.210$, $P=0.006$) and log₁₀-transformed FVC ($\beta=0.019$, $P=0.009$) (Table 4).

Interactions among unrelated *CDHR3* SNPs

GMDR analyses detected a significant two-locus interaction model between *CDHR3*_rs140154310 and *CDHR3*_rs408223, not in LD according to Fig. S1, for FEV_{0.5} among preschool children (TA>55% and $P<0.05$; Fig. 2). FEV_{0.5} were lower in high-risk than low-risk preschool children (mean: 89.13 vs 92.55; $P<0.0001$). There was no epistatic interaction for wheezing phenotypes.

255

256 ***In silico* analysis of intronic SNPs**

257 The prediction of the consequence of intronic SNPs showed that some of these SNPs occurred in
258 acceptor splicing sites while others were located in donor splicing sites with change of HSF scores
259 (Table S9-S15). Only rs146004234 and rs408223 showed potential alteration of splicing:
260 rs146004234 created activation of an intronic cryptic acceptor site while rs408223 created
261 activation of an intronic cryptic donor site.

262

263

264 **Discussion**

265 To the best of our knowledge, this study replicating rs6967330 and nearby SNPs of *CDHR3* is the
266 first in a Chinese population which not only studied asthma and early-life wheezing but also
267 analyzed additional relevant phenotypes such as spirometric parameters in preschool and
268 school-age children. We detected significant associations between *CDHR3* SNPs and current
269 wheeze, severity of wheezing and forced expiratory indices among preschool children. These
270 findings were replicated at single SNP and haplotype levels in our school-age cohort. This study
271 also identified epistatic interactions between two unrelated *CDHR3* SNPs for FEV_{0.5} among
272 preschoolers.

273 Rs6967330 of *CDHR3* was a risk factor for severe asthma exacerbations in Caucasian
274 preschool children.³ Bønnelykke et al. collected nasal samples during acute infections in two
275 groups of Caucasian children.²⁵ The risk allele A of rs6967330 was associated with increased risk
276 of respiratory illnesses related specifically to HRV-C (incidence risk ratio 1.51). Besides, birth
277 cohorts reported associations between early-life HRV infection and school-age asthma^{7-9,26} and
278 between HRV infection and asthma exacerbations²⁷⁻²⁹. Very recent data found that rs6967330
279 influenced *CDHR3* expression, HRV-C binding and replication as well as ciliogenesis and
280 Forkhead Box J1 expression in airway epithelial cells.³⁰ Consistent with such findings, current
281 wheeze and its severity among our preschoolers were consistently associated with rs6967330 with
282 its minor allele showing the same direction of association with increased risk (Table 1). Moreover,
283 the minor allele of this SNP was associated with asthma diagnosis, in particular among patients
284 requiring ICS treatment, in our school-age cohort (Tables 2 and S4). A case-control association
285 study of 3366 Japanese adults found that A allele of rs6967330 was associated with early-onset
286 asthma especially among the atopic individuals.³¹ These findings supported the relationship

287 between *CDHR3* and occurrence of childhood wheezing illnesses. Our results extend the current
288 knowledge to potential shared genetic risk for HRV infection and early-life wheezing and
289 childhood asthma development.

290 None of *CDHR3* SNPs other than rs6967330 was previously reported to be associated with
291 asthma traits.^{3,32} This study attempted to replicate the previous findings for rs6967330 together
292 with seven additional tagging SNPs over the region spanning 5-kb upstream and downstream from
293 this top SNP. Besides this coding SNP rs6967330, we identified significant associations between
294 rs140154310 and occurrence and severity of current wheeze in preschool children (Table 1) and
295 between rs4730125 and school-age asthmatics receiving ICS treatment (Table 2). This study also
296 examined *CDHR3* haplotypes due to more powerful multi-locus analysis to uncover
297 disease-causing variants for complex human diseases. Notably, the same at-risk haplotype GAC
298 formed by rs4730125, rs6967330 and rs408223 was associated with stronger ORs for current
299 wheeze in preschoolers and asthma in school-age children (Table S5) whereas the two-locus
300 haplotype TC after removing rs6967330 was still associated with school-age asthma requiring ICS
301 treatment (Table S6). Some of these genotypic and haplotypic associations survived FDR
302 adjustment for multiple testing (Table S16). These novel observations highlight the need to
303 determine multiple polymorphic markers over an extended region in asthma genetic studies.

304 We further examined the relationship between *CDHR3* and spirometric indices. The minor
305 alleles of rs140154310 and rs408223 were risk factors for lower indices in preschool children. By
306 GMDR, preschoolers with high-risk genotypes defined by these SNPs had significantly lower
307 FEV_{0.5} (Fig. 2). Besides, GGG haplotype formed by rs4730125, rs6967330 and rs408223 of
308 *CDHR3* was consistently associated with lower FEV_{0.5} and FEV_{0.5}/FVC among preschool children
309 (Table S7). The haplotypic associations for FEV_{0.5} and FEV_{0.5}/FVC were even stronger after
310 removal of the key SNP rs6967330 (Table S8). *CDHR3* haplotypes were also associated with
311 spirometric indices among schoolchildren with asthma (Table 4). These findings supported a
312 possible role of *CDHR3* in influencing lung function. As wheeze and asthma exacerbations were
313 risk factors of reduced lung function³³⁻³⁵, our finding for the functional rs6967330 would be
314 expected to enhance HRV-C infection and its associated wheezing whereby lowering subjects'
315 lung function. Further studies are needed to replicate these associations and to delineate the
316 relevant mechanisms.

317 Besides rs6967330, all other seven tagging SNPs are located in introns of *CDHR3* within
318 5-kb upstream and downstream from rs6967330. An increasing number of disease-related intronic

319 mutations were reported to alter splicing.^{36,37} *In silico* analysis of our seven intronic tagging SNPs
320 (Table S9-15) showed that some occurred in acceptor splicing sites while others occurred in donor
321 splicing sites. Only rs146004234 and rs408223 showed potential alteration of splicing (Table S12).
322 In the mutant, branching is directed to a new site. For rs146004234, the minor A allele abolishes
323 SF2/ASF (IgM-BRCA1) motif and SF2/ASF motif bound to splicing enhancer which creates a
324 new high score of silencer motif from Sironi et al. (63.22) bound to splicing silencer. On the other
325 hand, rs408223 created activation of an intronic cryptic donor site (Table S15). The intronic
326 variations of rs408223 did not show difference for branch point between mutant and reference
327 sequences. The mutant G allele abolishes SC35 motif and SF2/ASF (IgM-BRCA1) motif bound to
328 splicing enhancer, which creates a new high score of silencer motif from Sironi et al. (61.02)
329 bound to splicing silencer. Our additional analysis revealed all SNPs except rs10258293 that are
330 tagged to our tagging SNPs are located in intron of *CDHR3* (data not shown). Rs10258293
331 encodes for a synonymous variant in exon. As all our tagged and tagging SNPs are intronic or
332 synonymous, it is unlikely that they would lead to any functional effect on *CDHR3*.

333 We found lower MAFs of 0.07 for the top SNP rs6967330 in our school-age controls than the
334 published value of 0.17 among Danish controls.³ Assuming asthma prevalence to be 10%, the
335 sample size of our school-age cohort has respective power of 10% and 52% to detect OR 3.0 with
336 95% confidence for Chinese and Caucasian children. Therefore, this study might be underpowered
337 to detect any genotypic association with asthma traits. We attempted to minimize this problem by
338 capturing more tagging SNPs on this region and exploring associations at the haplotypic level as
339 well as for any SNP-SNP and gene-gene interaction. The cross-sectional nature of this study
340 meant that we did not record the occurrence of HRV infection in our subjects, which precluded us
341 from analyzing any HRV-*CDHR3* interaction for asthma traits in preschool and school-age
342 children. However, we extended the knowledge of *CDHR3* to its potential effects on spirometric
343 indices. These novel findings highlight the importance of *CDHR3* as a target gene for preventing
344 wheeze and asthma and for modulating childhood lung function. In the literature of asthma
345 genetics, 17q21 locus was widely replicated across different ethnic groups.⁵⁻⁸ Our group also
346 reported multiple SNPs on 17q21, especially around gasdermin B (*GSDMB*), to be associated with
347 asthma, atopy and spirometric indices in Chinese children and adults.^{17,38} It would be interesting to
348 examine any epistatic interaction between SNPs of these important asthma genes in different
349 populations.

350 In conclusion, this study identified genotypic and haplotypic associations between *CDHR3*

351 and early-life wheezing and childhood asthma. This candidate gene was also associated with
352 preschool lung function. We also showed epistatic interaction between unrelated *CDHR3* SNPs for
353 forced expiratory index among preschoolers. These results need replication in larger prospective
354 studies.

355

356 **Funding**

357 This work was supported by Research Grants Council General Research Fund (470909), Health
358 and Health Services Research Fund (06070261), Direct Grants for Research (4054292 and
359 4054367) of The Chinese University of Hong Kong, and Hong Kong Institute of Allergy Research
360 Grant 2017.

361

362 **Author contributions**

363 TFL obtained research funding, designed this study, carried out the analyses and drafted this
364 manuscript. MFT collected clinical data, performed genotyping experiments and analysed study
365 data. ASYL, KCCC, APSK, RCWM, JCNC and GWKW recruited patients and collected clinical
366 data. TCL and HYS helped with patient recruitment and spirometric measurement. WYC
367 supervised genotyping and data analyses. All authors approved the submitted version of this
368 manuscript.

369

370 **Acknowledgments**

371 We thanked children and their families for participating in this study, and Brenda Li, Air Chan,
372 Jack Szeto, Lawrence Cheung, Christy Leung, Easton Chan and Jonathan Leung for helping with
373 the field works in both cohorts.

374

375 **Conflicts of interest**

376 All authors declared no competing interests.

377 **References**

- 378 1. Ober C. Perspectives on the past decade of asthma genetics. *J Allergy Clin Immunol* 2005;
379 116:274-8.
- 380 2. Leung TF, Ko FW, Sy HY, Tsui SK, Wong GW. Differences in asthma genetics between
381 Chinese and other populations. *J Allergy Clin Immunol* 2014; 133:42-8.
- 382 3. Bønnelykke K, Sleiman P, Nielsen K, Kreiner-Moller E, Mercader JM, Belgrave D, et al. A

- genome-wide association study identifies *CDHR3* as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet* 2014; 46:51-5.
4. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 2005; 21:650-9.
5. Ross AJ, Dailey LA, Brighton LE, Devlin RB. Transcriptional profiling of mucociliary differentiation in human airway epithelial cells. *Am J Respir Cell Mol Biol* 2007; 37:169-85.
6. Bochkov YA, Watters K, Ashraf S, Griggs TF, Devries MK, Jackson DJ, et al. Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication. *Proc Natl Acad Sci USA* 2015; 112:5485-90.
7. Lemanske RF, Jr., Jackson DJ, Gangnon RE, Evans MD, Li Z, Shult PA, et al. Rhinovirus illnesses during infancy predict subsequent childhood wheezing. *J Allergy Clin Immunol* 2005; 116:571-7.
8. Jamieson KC, Warner SM, Leigh R, Proud D. Rhinovirus in the pathogenesis and clinical course of asthma. *Chest* 2015; 148:1508-16.
9. Kusel MM, de Klerk NH, Keadze T, Vohma V, Holt PG, Johnston SL, et al. Early-life respiratory viral infections, atopic sensitization, and risk of subsequent development of persistent asthma. *J Allergy Clin Immunol* 2007; 119:1105-10.
10. Leung TF, Liu TC, Mak KK, Su X, Sy HY, Li AM, et al. Reference standards for forced expiratory indices in Chinese preschool children. *Pediatr Pulmonol* 2013; 48:1119-26.
11. Leung TF, Yung E, Wong YS, Lam CW, Wong GW. Parent-reported adverse food reactions in Hong Kong Chinese pre-schoolers: epidemiology, clinical spectrum and risk factors. *Pediatr Allergy Immunol* 2009; 20:339-46.
12. Wong GW, Leung TF, Ko FW, Lee KK, Lam P, Hui DS, et al. Declining asthma prevalence in Hong Kong Chinese schoolchildren. *Clin Exp Allergy* 2004; 34:1550-5.
13. Wong GW, Mahesh PA, Ogorodova L, Leung TF, Fedorova O, Holla AD, et al. The EuroPrevall-INCO surveys on the prevalence of food allergies in children from China, India and Russia: the study methodology. *Allergy* 2010; 65:385-90.
14. Ozaki R, Qiao Q, Wong GW, Chan MH, So WY, Tong PC, et al. Overweight, family history of diabetes and attending schools of lower academic grading are independent predictors for metabolic syndrome in Hong Kong Chinese adolescents. *Arch Dis Child* 2007; 92:224-8.
15. Kong AP, Choi KC, Li AM, Hui SS, Chan MH, Wing YK, et al. Association between

physical activity and cardiovascular risk in Chinese youth independent of age and pubertal stage. BMC Public Health 2010; 10:303.

16. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. Am Rev Respir Dis 1987; 136:225-44.

17. Leung TF, Sy HY, Wong GW, Ng MC, Chan IH, Li CY, et al. Asthma and atopy are associated with chromosome 17q21 markers in Chinese children. Allergy 2009; 64:621-8.

18. Leung TF, Kong AP, Chan IH, Choi KC, Ho CS, Chan MH, et al. Association between obesity and atopy in Chinese schoolchildren. Int Arch Allergy Immunol 2009; 149:133-40.

19. Leung TF, Liu EK, Tang NL, Ko FW, Li CY, Lam CW, et al. Nitric oxide synthase polymorphisms and asthma phenotypes in Chinese children. Clin Exp Allergy 2005; 35:1288-94.

20. Leung TF, Wong YS, Chan IH, Yung E, Sy HY, Lam CW, et al. Domestic exposure to aeroallergens in Hong Kong families with asthmatic children. Pediatr Pulmonol 2011; 46:632-9.

21. Beydon N, Davis SD, Lombardi E, Allen JL, Arets HG, Aurora P, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. Am J Respir Crit Care Med 2007; 175:1304-45.

22. Ip MS, Karlberg EM, Karlberg JP, Luk KD, Leong JC. Lung function reference values in Chinese children and adolescents in Hong Kong. I. Spirometric values and comparison with other populations. Am J Respir Crit Care Med 2000; 162:424-9.

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

24. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002; 70:425-34.

25. Bønnelykke K, Coleman AT, Evans MD, Thorsen J, Waage J, Vissing NH, et al. Cadherin-related family member 3 genetics and rhinovirus C respiratory illnesses. Am J Respir Crit Care Med 2018; 197:589-94.

26. Kotaniemi-Syrjänen A, Vainionpää R, Reijonen TM, Waris M, Korhonen K, Korppi M. Rhinovirus-induced wheezing in infancy - the first sign of childhood asthma? J Allergy Clin Immunol 2003; 111:66-71.

27. Miller EK, Edwards KM, Weinberg GA, Iwane MK, Griffin MR, Hall CB, et al. A novel

- group of rhinoviruses is associated with asthma hospitalizations. *J Allergy Clin Immunol* 2009; 123:98-104.
28. Leung TF, To MY, Yeung AC, Wong YS, Wong GW, Chan PK. Multiplex molecular detection of respiratory pathogens in children with asthma exacerbation. *Chest* 2010; 137:348-54.
29. Mak RK, Tse LY, Lam WY, Wong GW, Chan PK, Leung TF. Clinical spectrum of human rhinovirus infections in hospitalized Hong Kong children. *Pediatr Infect Dis J* 2011; 30:749-53.
30. Basnet S, Bochkov YA, Brockman-Schneider RA, Kuipers I, Aesif SW, Jackson DJ, et al. *CDHR3* asthma-risk genotype affects susceptibility of airway epithelium to rhinovirus C infections. *Am J Respir Cell Mol Biol* 2019 Mar 27. doi: 10.1165/rcmb.2018-0220OC.
31. Kanazawa J, Masuko H, Yatagai Y, Sakamoto T, Yamada H, Kaneko Y, et al. Genetic association of the functional *CDHR3* genotype with early-onset adult asthma in Japanese populations. *Allergol Int* 2017; 66:563-7.
32. Kanazawa J, Masuko H, Yatagai Y, Sakamoto T, Yamada H, Kaneko Y, et al. Genetic association of the functional *CDHR3* genotype with early-onset adult asthma in Japanese populations. *Allergol Int* 2017; 66:563-7.
33. Kelly WJ, Hudson I, Raven J, Phelan PD, Pain MC, Olinsky A. Childhood asthma and adult lung function. *Am Rev Respir Dis* 1988; 138:26-30.
34. Kjellman B, Hesselmar B. Prognosis of asthma in children: a cohort study into adulthood. *Acta Paediatr* 1994; 83:854-61.
35. Belgrave DC, Granell R, Turner SW, Curtin JA, Buchan IE, Le Souef PN, et al. Lung function trajectories from pre-school age to adulthood and their associations with early life factors: a retrospective analysis of three population-based birth cohort studies. *Lancet Respir Med* 2018; 6:526-34.
36. Baralle D, Baralle M. Splicing in action: assessing disease causing sequence changes. *J Med Genet* 2005; 42:737-48.
37. Buratti E, Baralle M, Baralle FE (2006) Defective splicing, disease and therapy: searching for master checkpoints in exon definition. *Nucleic Acids Res* 34: 3494–3510.
38. Tang MF, Sy HY, Kong AP, Ko FW, Wang SS, Liu TC, et al. Genetic effects of multiple asthma loci identified by genomewide association studies on asthma and spirometric indices. *Pediatr Allergy Immunol* 2016; 27:185-94.

Table 1 Associations between tagging *CDHR3* SNPs and different wheezing phenotypes in the preschool cohort

SNPs (A/a)	OR (95% CI), <i>P</i> value				
	Current wheeze (n=142) vs No current wheeze (n=1199)	Wheeze ever (n=227) vs Controls* (n=749)	Current wheeze (n=142) vs Controls* (n=749)	Recurrent wheeze (>3 episodes, n=39) vs Controls (n=749)	Frequency of current wheeze [†]
rs3887998 (G/A)	1.027 (0.768-1.375), 0.855	1.120 (0.873-1.437), 0.372	1.058 (0.782-1.432), 0.713	1.190 (0.703-2.015), 0.517	1.030 (0.771-1.377), 0.839
rs140154310 (C/T)	2.204 (1.161-4.185), 0.016	1.355 (0.721-2.548), 0.345	1.915 (0.977-3.754), 0.058	1.579 (0.459-5.428), 0.469	2.171 (1.145-4.116), 0.018
rs73195657 (T/C)	0.523 (0.251-1.089), 0.083	1.102 (0.667-1.822), 0.704	0.572 (0.267-1.225), 0.150	0.519 (0.122-2.209), 0.375	0.520 (0.250-1.084), 0.081
rs146004234 (G/A)	0.726 (0.287-1.838), 0.500	0.623 (0.287-1.353), 0.232	0.616 (0.239-1.586), 0.316	0.459 (0.062-3.395), 0.445	0.720 (0.284-1.828), 0.490
rs4730125 (G/T)	0.846 (0.649-1.102), 0.215	0.940 (0.751-1.176), 0.585	0.838 (0.637-1.101), 0.205	0.687 (0.409-1.155), 0.156	0.842 (0.646-1.097), 0.204
rs6967330 (G/A)	1.628 (1.089-2.432), 0.017	1.390 (0.962-2.008), 0.079	1.570 (1.032-2.389), 0.035	1.649 (0.814-3.343), 0.165	1.637 (1.097-2.440), 0.016
rs73195665 (G/A)	1.297 (0.720-2.338), 0.386	1.318 (0.771-2.251), 0.313	1.368 (0.729-2.568), 0.330	1.589 (0.540-4.672), 0.400	1.292 (0.717-2.326), 0.393
rs408223 (C/G)	1.070 (0.739-1.551), 0.719	1.122 (0.814-1.548), 0.482	1.168 (0.796-1.715), 0.427	1.371 (0.725-2.593), 0.332	1.076 (0.744-1.557), 0.698

Analyzed by logistic regression, fitting age and gender as covariates for additive model.

† Ordinal regression, fitting age and gender as covariates for additive model. The coding for frequency of current wheeze is 0 = never, 2 = 1-3 episodes, 8 = 4-12 episodes, 14 = >12 episodes.

* Controls are defined as those without asthma, eczema and wheeze ever.

Table 2 Associations between asthma diagnosis and tagging SNPs of *CDHR3* in school-age cohort

Genotypes	All asthmatics		All controls		Pearson χ^2 / Fisher's Exact* <i>P</i> -value	Additive model		Dominant model		Recessive model	
	n = 892		n = 1187			OR (95% CI)	<i>P</i> -value	OR (95% CI)	<i>P</i> -value	OR (95% CI)	<i>P</i> -value
	n	%	n	%							
rs3887998											
GG	528	59.3	679	57.3	0.587	0.928 (0.797-1.080)	0.334	0.885 (0.734-1.067)	0.200	1.039 (0.706-1.530)	0.847
GA	308	34.6	436	36.8							
AA	55	6.2	71	6.0							
rs140154310											
CC	860	96.4	1128	95.0	0.172*	0.786 (0.496-1.247)	0.307	0.795 (0.498-1.268)	0.335	ND	ND
CT	32	3.6	58	4.9							
TT	0	0	1	0.1							
rs73195657											
TT	792	88.8	1055	89.0	0.884*	0.955 (0.727-1.255)	0.744	0.979 (0.731-1.310)	0.884	0.555 (0.150-2.054)	0.378
TC	96	10.8	124	10.5							
CC	4	0.4	7	0.6							
rs146004234											
GG	852	95.5	1120	94.4	0.188*	0.821 (0.543-1.242)	0.350	0.795 (0.522-1.212)	0.287	ND	ND
GA	39	4.4	67	5.6							
AA	1	0.1	0	0							

rs4730125												
GG	333	37.8	474	40.1	0.399	1.070 (0.934-1.226)	0.329	1.143 (0.945-1.382)	0.168	0.997 (0.761-1.304)	0.980	
GT	432	49.0	544	46.0								
TT	116	13.2	164	13.9								
rs6967330												
GG	741	83.1	1034	87.1	0.032*	1.322 (1.035-1.687)	0.025	1.338 (1.034-1.733)	0.027	1.602 (0.470-5.467)	0.452	
GA	145	16.3	148	12.5								
AA	6	0.7	5	0.4								
rs73195665												
GG	817	91.6	1111	93.6	0.189*	1.272 (0.904-1.788)	0.167	1.294 (0.910-1.840)	0.151	0.903 (0.078-10.449)	0.935	
GA	74	8.3	74	6.2								
AA	1	0.1	2	0.2								
rs408223												
CC	711	81.0	918	78.3	0.318	0.849 (0.682-1.056)	0.142	0.830 (0.658-1.045)	0.113	1.068 (0.374-3.052)	0.902	
CG	160	18.2	245	20.9								
GG	7	0.8	10	0.9								

Analyzed by logistic regression, fitting age and gender as covariates.

479 **Table 3** Associations between tagging SNPs of *CDHR3* and spirometric indices in preschool
 480 cohort

SNP (A/a)	β , <i>P</i> -value		
	FEV _{0.5}	FVC	FEV _{0.5} /FVC
rs3887998 (G/A)	0.053, 0.935	0.277, 0.665	0.109, 0.789
rs140154310 (C/T)	-3.766, 0.039	-3.739, 0.037	-0.342, 0.765
rs73195657 (T/C)	0.539, 0.675	2.733, 0.030	-1.082, 0.179
rs146004234 (G/A)	0.849, 0.653	-0.055, 0.976	-0.086, 0.942
rs4730125 (G/T)	0.688, 0.236	1.007, 0.076	-0.275, 0.449
rs6967330 (G/A)	0.585, 0.572	0.404, 0.691	0.336, 0.605
rs73195665 (G/A)	-1.105, 0.440	-0.635, 0.823	-0.569, 0.525
rs408223 (C/G)	-2.411, 0.004	-1.193, 0.152	-1.292, 0.015

481 *P*-values significant after FDR correction were highlighted in italic.

482 Analyzed by linear regression, fitting age, gender, current wheeze as covariates for additive
 483 model.

484

485

Table 4 Relationship between *CDHR3* haplotypes and spirometric indices measured among asthmatics in the school-age cohort

Spirometric index	Combination of <i>CDHR3</i> SNPs	Haplotype	Frequency	β	$P^{\dagger}$
FEV ₁	rs3887998_rs73195657	GT	0.767	Reference	
		AT	0.173	4.210	0.006
		AC	0.055	-0.318	0.902
	rs4730125_rs6967330_rs408223	GGC	0.436	Reference	
		TGC	0.394	-1.990	0.116
		GGG	0.086	-4.308	0.039
		GAC	0.079	-2.282	0.302
LogFVC	rs3887998_rs73195657	GT	0.767	Reference	
		AT	0.172	0.019	0.009
		AC	0.056	-0.003	0.791
	rs4730125_rs6967330_rs408223	GGC	0.435	Reference	
		TGC	0.395	-0.006	0.335
		GGG	0.087	-0.012	0.210
		GAC	0.081	-0.008	0.434

486

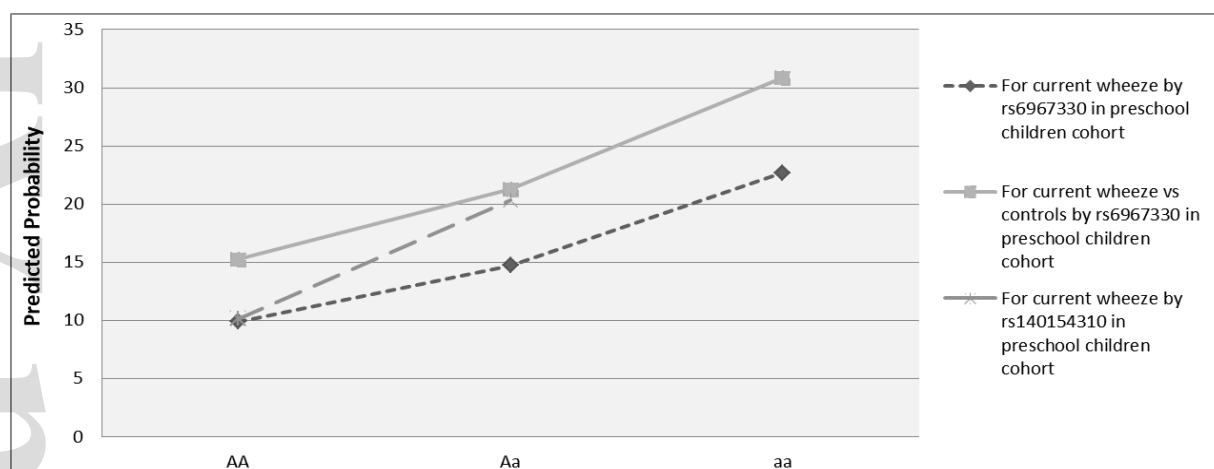
LogFVC, log₁₀-transformed FVC.

487

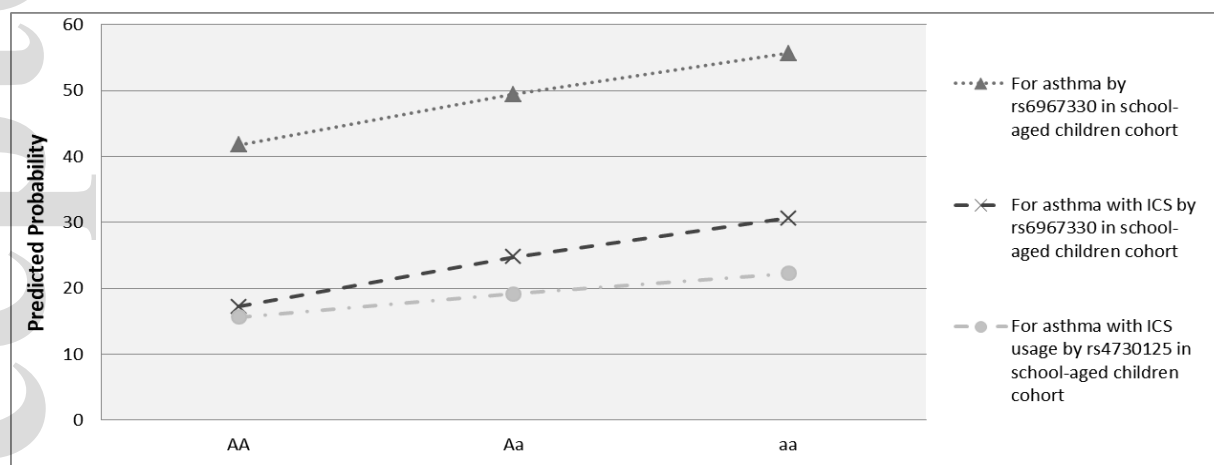
[†] Adjusted for age, sex and ICS use as covariates.

489 **Figure 1** Fitted predicted probability curves of falling into corresponding phenotype
 490 (wheeze/asthma) according to different genotypes of *CDHR3* SNPs: rs6967330 (C/T);
 491 rs140154310 (C/T); rs4730125 (G/T) among participants in (A) preschool cohort and (B)
 492 school-age cohort. There were no subjects homozygous for the minor allele (T) of the SNP
 493 rs140154310. $P < 0.0001$, except between CT and TT genotypes of rs6967330 for asthma and
 494 asthma with ICS use. AA, homozygous major genotype; Aa, heterozygous genotype; aa,
 495 homozygous minor genotype

496
 497 (A)



498
 499
 500 (B)

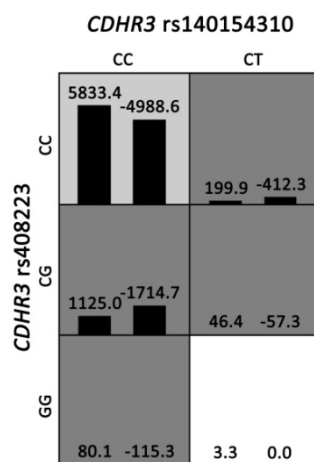


501

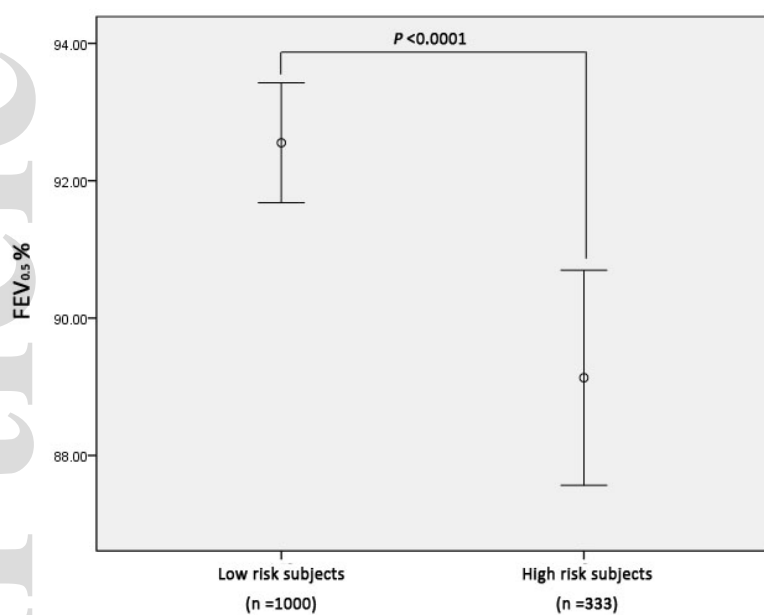
502

Figure 2 (a) Best two-locus model for FEV_{0.5} in the preschool cohort. High-risk genotypes are in dark grey, low-risk genotypes are in light grey, and white cells indicate unclassified genotypes due to small numbers of subjects. Left and right bars represent children without and with current wheeze respectively. The figures above the bars in each cell are the sum of scores for the corresponding group of individuals. The heights of the bars are proportional to the sum of scores in each group. (b) Distribution of FEV_{0.5} in preschoolers with high-risk and low-risk genotypes defined by 2-locus model between *CDHR3*_rs140154310 and *CDHR3*_rs408223 ($P < 0.0001$ by student t-test).

(a)



(b)



ONLINE SUPPLEMENT

Cadherin-related Family Member 3 Gene Impacts Childhood Asthma in Chinese Children

Ting Fan Leung^{1,2}, Man Fung Tang¹, Agnes Sze Yin Leung¹, Alice Pik Shan Kong³, Tak Chi Liu¹, Renee Wan Yi Chan^{1,2}, Ronald Ching Wan Ma³, Hing Yee Sy¹, Juliana Chung Ngor Chan³ and Gary Wing Kin Wong¹

Department of ¹Paediatrics and ³Medicine and Therapeutics, and ²Hong Kong Hub of Paediatric Excellence, The Chinese University of Hong Kong, Hong Kong

METHODS

Statistical analysis

Subjects' clinical and laboratory characteristics were compared by student t-test for numerical data and Pearson χ^2 or Fisher exact test for categorical variables. Allele frequencies for all SNPs were estimated by the gene-counting method, and Hardy-Weinberg equilibrium (HWE) was examined by χ^2 exact test. The associations between SNPs and categorical and quantitative phenotypes were analyzed by logistic and linear regression respectively, adjusting for age, gender and wheezing ever as covariates for the preschool cohort, and adjusting for age, gender, disease status and inhaled corticosteroid (ICS) treatment as covariates in the school-age cohort. The genotypic associations with the frequency of preschool wheeze in the past year was analysed by ordinal regression where the categories were coded as 0 (none), 2 (1-3 episodes), 8 (4-12 episodes) and 14 (>12 episodes). All these comparisons were made two-tailed using SPSS v.21 (IBM Corp., Armonk, NY). The nominal significance threshold was set at $P < 0.05$, which was adjusted for multiple testing using the false discovery rate (FDR; see Table S16). The method of predictive margins was used to assess the probability of falling into a corresponding cluster of phenotype according to the allele of each significant SNP. Pairwise linkage disequilibrium (LD) coefficient was calculated for each SNP pair by Haploview 5.0 (Daly Lab, Cambridge, MA, USA). The haplotypes were determined by Haploview 5.0, and their associations with wheezing traits were tested by haplo.stats 1.7.9 on R v.3.5.0 (<http://www.r-project.org/>).^{S1} The haplotypic associations with dichotomous and continuous variables were analyzed by multivariable logistic and linear regression respectively.

Generalized multifactor dimensionality reduction (GMDR, v. 0.9; <http://www.healthsystem.virginia.edu/internet/addiction-genomics/software/>) was used to test

for interactions among unrelated *CDHR3* SNPs in different LD blocks (Fig. S1) in our preschool children cohort.^{S2} GMDR is an extension of MDR^{S3} that allows adjustment for discrete and quantitative covariates and is applicable to both continuous and dichotomous phenotypes. It assigned subjects as high-risk and low-risk according to score-based statistics. The epistatic interactions among SNPs were evaluated using 10-fold cross-validation. The statistical significance of different interaction models was determined by comparing the average prediction error from the observed data with the distribution of average prediction errors under the null hypothesis (rejected at $P < 0.05$) for absence of association derived empirically from 5000 permutations prior to selecting the best performing model. The best model was then chosen based on cross-validation consistency (CVC) $\geq 8/10$ and testing accuracy (TA) score ≥ 0.55 . Pearson's χ^2 , standard analysis of variance and Student's t-tests were used to compare wheezing phenotypes and spirometric indices among different groups assigned by GMDR.

***In silico* analysis**

In silico analysis was performed to evaluate potential biological functions of seven intronic SNPs genotyped (rs3887998, rs140154310, rs73195657, rs146004234, rs4730125, rs73195665, and rs408223) on *CDHR3* mRNA splicing. These intronic SNPs were analyzed by Human Splicing Finder (HSF) (v. 3.1) (<http://www.umd.be/HSF/>), which integrated all available matrices to identify exonic and intronic motifs: exonic splicing enhancer (ESE) motifs from HSF (ESE-HSFs) for serine/arginine-rich (SR) proteins (SRp40, SC35, SF2/ASF, SF2/ASF IgM/BRCA1, and SRp55), RESCUE ESE hexamers (RESCUE-ESE), putative 8-mer ESEs (PESEs) and putative 8-mer exonic splicing silencers (PESSs), exon identity and intron-identity elements (EIEs and IIEs), heterogeneous nuclear ribonucleoprotein (hnRNP)-binding motifs, and Fas exonic splicing silencers (ESS).

References

- S1. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002; 70:425-34.
- S2. Lou XY, Chen GB, Yan L, Ma JZ, Zhu J, Elston RC, et al. A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am J Hum Genet* 2007; 80:1125-37.
- S3. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bioinformatics* 2003; 19:376-82.

Table S1 The details and minor allele frequencies of our selected tagging *CDHR3* SNPs

SNP	Position	Cytogenetic band	Nature of SNP	Major allele	Minor allele	MAF of CHS from 1000 genomes	MAF in preschool cohort	MAF in school-age cohort	
								Case	Control
rs3887998	105653867	7q22.3	Intron	G	A	0.21	0.227	0.235	0.244
rs140154310	105653880	7q22.3	Intron	C	T	0.033	0.024	0.018	0.025
rs73195657	105654609	7q22.3	Intron	T	C	0.071	0.049	0.058	0.058
rs146004234	105656263	7q22.3	Intron	G	A	0.024	0.022	0.023	0.028
rs4730125	105656840	7q22.3	Intron	G	T	0.371	0.374	0.377	0.369
rs6967330	105658451	7q22.3	Missense (C529Y)	G	A	0.071	0.074	0.088	0.067
rs73195665	105659249	7q22.3	Intron	G	A	0.024	0.038	0.043	0.033
rs448025	105660496	7q22.3	Intron	T	G	0.024	0	0	0
rs408223	105662024	7q22.3	Intron	C	G	0.129	0.113	0.099	0.113
rs543085868	105662156	7q22.3	Intron	A	C	0.019	0	0	0

MAF, minor allele frequency; ND, not done.

All genotyped SNPs of all subjects in two cohorts were in Hardy-Weinberg equilibrium with $P > 0.05$.

Table S2 Demographic, environmental and clinical factors of subjects in the preschool cohort

Characteristic	<i>n</i> = 1341
Age, years	4.7 ± 1.0
Male (%)	740 (55.2)
Doctor-diagnosed asthma	71 (5.3)
Doctor-diagnosed hyperactive airway	434 (32.4)
Doctor-diagnosed allergic rhinitis	384 (28.7)
Doctor-diagnosed eczema	461 (34.4)
Doctor-diagnosed food allergy	53 (4.0)
Tobacco smoker in household	553 (41.2)
Maternal smoking during pregnancy	54 (4.0)
Breastfeeding ever	686 (51.2)
Current pet keeping	89 (6.7)
Wheezing ever	227 (16.9)
Current wheeze (in past 12 months)	142 (10.6)
Frequency wheeze within past 12 months	
No wheeze	1199 (89.4)
1-3 episodes	103 (7.7)
4-12 episodes	31 (2.3)
>12 episodes	8 (0.6)
Spirometric variables	
FEV _{0.5} (L)	0.753 ± 0.192
FEV _{0.5} , % predicted	91.7 ± 14.3
FVC (L)	1.061 ± 0.279
FVC, % predicted	102.2 ± 14.4
FEV _{0.5} /FVC, % predicted	71.6 ± 9.1

FEV_{0.5}, forced expiratory volume in first 0.5 second; FVC, forced vital capacity; ND, not done.

Results expressed in mean ± standard deviation or number (percentage).

Table S3 Clinical characteristics of patients and controls in the school-age cohort

Characteristic	Asthma (<i>n</i> = 892)	Control (<i>n</i> = 1187)
Age, years	11.0 ± 4.1*	13.7 ± 4.5
Male (%)	550 (61.7)*	525 (44.2)
Log ₁₀ -transformed plasma total IgE level	2.61 ± 0.69*	1.86 ± 0.69
Increased plasma total IgE level (%)**	616 (78.4)*	265 (40.3)
Atopy*** (%)	572 (75.3)*	248 (38.0)
Spirometric variables		
FEV ₁ , % predicted	92.6 ± 17.5	ND
FVC, % predicted	103.5 ± 20.3	ND
FEV ₁ /FVC, % predicted	82.2 ± 10.5	ND
FEV ₁ reversibility, %	7.2 ± 11.1	ND

FEV₁, forced expiratory volume in first one second; FVC, forced vital capacity; ND, not done.

Results expressed in mean ± standard deviation or number (percentage).

* $P < 0.001$.

** Compared with in-house upper limits of normal: 120 kIU/l for 5-6 years, 160 kIU/l for 6-7 years and 180 kIU/l for ≥8 years.

*** Defined as having one or more positive SPT or RAST to aeroallergens.

GG	216	91.1	1111	93.6							
GA	21	8.9	74	6.2	0.292*	1.328	0.289	1.369	0.251	ND	ND
AA	0	0	2	0.2		(0.786-2.242)		(0.801-2.341)			
rs408223											
CC	185	80.4	918	78.3							
CG	42	18.3	245	20.9	0.488*	0.892	0.519	0.830	0.390	1.694	0.460
GG	3	1.3	10	0.9		(0.630-1.262)		(0.586-1.232)		(0.419-6.854)	

P-values significant after FDR correction were highlighted in italic.

Analyzed by binary logistic regression, fitting age and gender as covariate.

* Analyzed by Fisher exact test.

† Heterozygous vs homozygous major.

Table S5 Relationship between *CDHR3* haplotypes and different wheezing phenotypes

Phenotype	Combination of <i>CDHR3</i> SNPs	Haplotype	Frequency		OR	95% CI	<i>P</i> [†]
			With phenotype	Control			
Current wheeze in preschoolers	rs3887998_rs73195657	GT	0.763	0.778	Reference		
		AT	0.208	0.176	1.171	0.855-1.606	0.326
		AC	0.029	0.045	0.627	0.291-1.349	0.233
	rs4730125_rs6967330_rs408223	GGC	0.434	0.447	Reference		
		TGC	0.337	0.378	0.902	0.669-1.216	0.499
		GGG	0.120	0.103	1.144	0.759-1.724	0.520
		GAC	0.109	0.071	1.535	0.987-2.389	0.048
School-age asthma	rs3887998_rs73195657	GT	0.763	0.755	Reference		
		AT	0.178	0.187	0.935	0.790-1.107	0.436
		AC	0.056	0.057	0.915	0.692-1.210	0.533
	rs4730125_rs6967330_rs408223	GGC	0.438	0.457	Reference		
		TGC	0.378	0.365	1.098	0.950-1.269	0.207
		GGG	0.098	0.110	0.922	0.730-1.166	0.498
		GAC	0.086	0.064	1.392	1.075-1.801	0.012
School-age asthma receiving ICS treatment	rs3887998_rs73195657	GT	0.760	0.755	Reference		
		AT	0.185	0.187	0.967	0.736-1.269	0.807
		AC	0.052	0.057	0.892	0.567-1.402	0.621
	rs4730125_rs6967330_rs408223	GGC	0.389	0.457	Reference		
		TGC	0.410	0.365	1.414	1.115-1.793	0.004
		GGG	0.102	0.110	1.067	0.729-1.564	0.738
		GAC	0.085	0.064	1.631	1.072-2.482	0.023

[†] Adjusted for age and sex as covariates.

Table S6 Relationship between *CDHR3* haplotypes formed by rs4730125 and rs408223 and different wheezing phenotypes

Phenotype	Combination of <i>CDHR3</i> SNPs	Haplotype	Frequency		OR	95% CI	<i>P</i> [†]
			With phenotype	Control			
Current wheeze in preschoolers	rs4730125_rs408223	GC	0.542	0.518	Reference		
		TC	0.338	0.378	0.846	0.635-1.126	0.251
		GG	0.120	0.103	1.085	0.725-1.624	0.692
School-age asthma	rs4730125_rs408223	GC	0.524	0.521	Reference		
		TC	0.377	0.366	1.056	0.918-1.216	0.446
		GG	0.099	0.110	0.894	0.711-1.124	0.337
School-age asthma receiving ICS treatment	rs4730125_rs408223	GC	0.473	0.521	Reference		
		TC	0.422	0.366	1.361	1.087-1.703	0.007
		GG	0.105	0.110	1.049	0.728-1.512	0.796

[†] Adjusted for age and sex as covariates.

Table S7 Relationship between *CDHR3* haplotypes and forced expiratory indices measured by incentive spirometry among preschoolers

Spirometric Index	Combination of <i>CDHR3</i> SNPs	Haplotype	Frequency	β	$P^\dagger$
FEV _{0.5}	rs3887998_rs73195657	GT	0.772	Reference	
		AT	0.179	-0.102	0.886
		AC	0.049	0.576	0.659
		GGC	0.438	Reference	
	rs4730125_rs6967330_rs408223	TGC	0.374	0.389	0.531
		GGG	0.112	-1.993	0.027
		GAC	0.073	0.988	0.363
FVC	rs3887998_rs73195657	GT	0.772	Reference	
		AT	0.179	-0.259	0.711
		AC	0.049	2.619	0.042
		GGC	0.439	Reference	
	rs4730125_rs6967330_rs408223	TGC	0.374	0.945	0.124
		GGG	0.112	-1.502	0.393
		GAC	0.073	0.768	0.472
FEV _{0.5} /FVC	rs3887998_rs73195657	GT	0.772	Reference	
		AT	0.179	0.325	0.466
		AC	0.049	-0.907	0.267
		GGC	0.439	Reference	
	rs4730125_rs6967330_rs408223	TGC	0.374	-0.524	0.179
		GGG	0.112	-1.464	0.010
		GAC	0.073	0.151	0.824

[†] Adjusted for age, sex and current wheeze as covariates.

Table S8 Relationship between *CDHR3* haplotypes formed by rs4730125 and rs408223 and forced expiratory indices among preschoolers

Spirometric Index	Combination of <i>CDHR3</i> SNPs	Haplotype	Frequency	β	$P^\dagger$
FEV _{0.5}	rs4730125_rs408223	GC	0.512	Reference	
		TC	0.374	0.233	0.700
		GG	0.113	-2.310	0.009
FVC	rs4730125_rs408223	GC	0.512	Reference	
		TC	0.374	0.823	0.167
		GG	0.113	-0.838	0.335
FEV _{0.5} /FVC	rs4730125_rs408223	GC	0.512	Reference	
		TC	0.374	-0.548	0.148
		GG	0.113	-1.528	0.006

[†] Adjusted for age, sex and current wheeze as covariates.

Table S9 The prediction of the consequence of rs3887998 mutation

cDNA position	Splice site type	Motif	New splice site	Wild type	Mutant	If cryptic site use, exon length variation	Variation (%)
(A) HSF matrices							
+387	Donor	tgggtaaat	TGGataaat	81	54.16	NA	Site broken (-33.14)
(B) Branch Points							
No difference between mutant and reference sequences was found with this matrice.							

cDNA position	Linked SR protein	Reference motif (value 0-100)	Linked SR protein	Mutant motif (value 0-100)	Variation
(C) ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins					
+387	SRp55	tgggta (74.12)			Site broken -100

cDNA position	Sironi motif reference	Reference silencer (value 0-100)	Sironi mutant motif	Mutant silencer (value 0-100)	Variation
(D) Silencer motifs from Sironi et al.					
+383	Motif 2 - [T/G]G[T/A]GGGG	tgtctgg (68.11)	Motif 2 - [T/G]G[T/A]GGGG	tgtctgg (68.11)	0 %
+385	Motif 2 - [T/G]G[T/A]GGGG	tctgggt (71.20)			Site broken -4.34

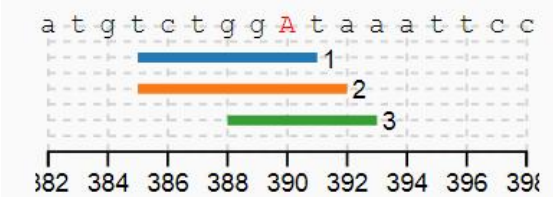
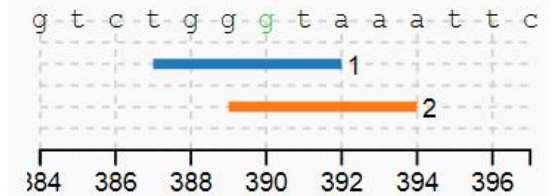
Predicted signal	Prediction algorithm	cDNA position	Interpretation
(E) Interpreted data			
ESS Site broken	1 - Sironi et al. - Motif 2		Alteration of an intronic ESS site. Probably no impact on splicing.
	2 - IIEs from Zhang et al.		
	3 - Fas-ESS hexamers		
New ESE Site	1 - EIEs from Zhang et al.		Creation of an intronic ESE site. Probably no impact on splicing.
	2 - HSF Matrices - 9G8		

Table S10 The prediction of the consequence of rs140154310 mutation

cDNA position	Splice site type	Motif	New splice site	Wild type	Mutant	If cryptic site use,	Variation (%)
(A) HSF Matrices							
+396	Acceptor	tccacaccgtagaa	tccacactgtagAA	78.55	79.45	NA	+1.15
(B) Branch Points							
No difference between mutant and reference sequences was found with this matrice.							

cDNA position	Linked SR protein	Reference motif (value 0-100)	Linked SR protein	Mutant motif (value 0-100)	Variation
(C) ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins					
+397	SRp40	ccacacc (90.90)	SRp40	ccacacT (81.74)	-10.08 %
+398	SF2/ASF (IgM-BRCA1)	cacaccg (81.77)	SF2/ASF (IgM-BRCA1)	cacacTg (82.54)	+0.94 %
+398	SF2/ASF (IgM-BRCA1)	cacaccg (81.77)	SF2/ASF	cacacTg (83.40)	+2 %
+398	SF2/ASF	cacaccg (80.78)	SF2/ASF (IgM-BRCA1)	cacacTg (82.54)	+2.18 %
+398	SF2/ASF	cacaccg (80.78)	SF2/ASF	cacacTg (83.40)	+3.24 %
+400	SF2/ASF (IgM-BRCA1)	caccgta (81.08)	SF2/ASF (IgM-BRCA1)	cacTgta (71.46)	-11.86 %
+400	SF2/ASF	caccgta (78.33)	SF2/ASF (IgM-BRCA1)	cacTgta (71.46)	-8.77 %
+403			SRp55	Tgtaga (74.69)	New site

cDNA position	Sironi motif reference	Reference silencer (value 0-100)	Sironi mutant motif	Mutant silencer (value 0-100)	Variation
(D) Silencer motifs from Sironi et al.					
No difference between mutant and reference sequences was found with this matrice.					

Predicted signal	Prediction algorithm	cDNA position	Interpretation
(E) Interpreted Data			
New ESE Site	1 - ESR Sequences from Goren et al.		Creation of an intronic ESE site. Probably no impact on splicing.
	2 - PESE Octamers from Zhang & Chasin		
	3 - ESE-Finder - SRp55		

cDNA Position	Linked SR protein	Reference Motif (value 0-100)	Linked SR protein	Mutant Motif (value 0-100)	Variation
(C) ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins					
-954	SC35	agctcctg (87.83)	SC35	agctccCg (83.90)	-4.48 %
-952	SRp40	ctcctgg (80.12)	SF2/ASF (IgM-BRCA1)	ctccCgg (88.00)	+9.84 %
-952	SRp40	ctcctgg (80.12)	SF2/ASF	ctccCgg (83.46)	+4.17 %
-952	SF2/ASF (IgM-BRCA1)	ctcctgg (74.08)	SF2/ASF (IgM-BRCA1)	ctccCgg (88.00)	+18.8 %
-952	SF2/ASF (IgM-BRCA1)	ctcctgg (74.08)	SF2/ASF	ctccCgg (83.46)	+12.67 %

cDNA Position	Sironi Motif Reference	Reference silencer (value 0-100)	Sironi Mutant motif	Mutant silencer (value 0-100)	Variation
(D) Silencer motifs from Sironi et al.					
-953	Motif 3 - TCTCCCAA	gctcctgg (63.35)	Motif 3 - TCTCCCAA	gctccCgg (80.54)	+27.13 %

Predicted signal	Prediction algorithm	cDNA Position	Interpretation
(E) Interpreted Data			
ESS Site broken	1 - Fas-ESS hexamers		Alteration of an intronic ESS site. Probably no impact on splicing.
	2 - IIEs from Zhang et al.		
	3 - ESR Sequences from Goren et al.		

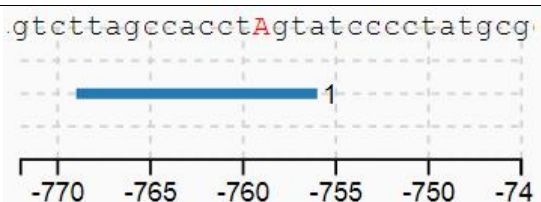
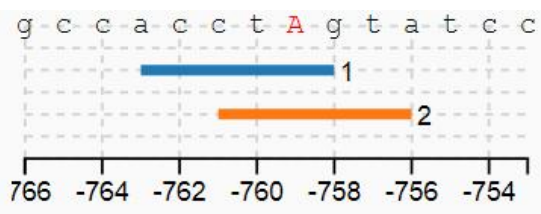
Table S12 The prediction of the consequence of rs146004234 mutation

cDNA position	Splice site type	Motif	New splice site	Wild type	Mutant	If cryptic site use, exon length variation	Variation (%)
(A) HSF Matrices							
-770	Acceptor	ttagccacctggta	ttagccacctagTA	47.7	76.65	+758	New site +60.69
-762	Donor	ctggatatcc	CTAgtatcc	72.26	61.68	+750	Site broken -14.64

cDNA position	Branch point motif	CV for reference sequence (0-100)	CV for mutant sequence (0-100)	Variation
(B) Branch Points				
-765	cacctGg	39.92	69.55	New site

cDNA position	Linked SR protein	Reference motif (value 0-100)	Linked SR protein	Mutant motif (value 0-100)	Variation
(C) ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins					
-765	SF2/ASF (IgM-BRCA1)	cacctgg (77.38)			Site broken -100
-765	SF2/ASF	cacctgg (75.31)			Site broken -100

cDNA position	Sironi motif reference	Reference silencer (value 0-100)	Sironi mutant motif	Mutant silencer (value 0-100)	Variation
(D) Silencer motifs from Sironi et al.					
-766			Motif 3 - TCTCCCAA	ccacctAg (63.22)	New site -30.51

Predicted signal	Prediction algorithm	cDNA Position	Interpretation
(E) Interpreted Data			
New Acceptor Site	1 - HSF Matrices		Activation of an intronic cryptic acceptor site. Potential alteration of splicing.
ESS Site broken	1 - ESR Sequences from Goren et al. 2 - IIEs from Zhang et al.		Alteration of an intronic ESS site. Probably no impact on splicing.

Predicted signal	Prediction algorithm	cDNA Position	Interpretation
(E)			
ESS Site broken	1 - Sironi et al. - Motif 1		Alteration of an intronic ESS site. Probably no impact on splicing.
	2 - HSF Matrices - hnRNP A1		

Table S14 The prediction of the consequence of rs73195665 mutation

cDNA position	Splice site type	Motif	New splice site	Wild type	Mutant	If cryptic site use, exon length variation	Variation (%)
(A) HSF Matrices							
+733	Acceptor	cagatgggtagat	cagatgggatagAT	66.76	66.83	NA	+0.1
+738	Donor	gggtagat	GGGatagat	72.88	46.04	NA	Site broken -36.83
(B) Branch Points							
No difference between mutant and reference sequences was found with this matrice.							
(C) ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins							
No Enhancer motif found with this matrice							

cDNA position	Sironi motif reference	Reference silencer (value 0-100)	Sironi mutant motif	Mutant silencer (value 0-100)	Variation
(D) Silencer motifs from Sironi et al.					
+734	Motif 2 - [T/G]G[T/A]GGGG	agatggg (82.32)	Motif 2 - [T/G]G[T/A]GGGG	agatggg (82.32)	0 %
+735	Motif 2 - [T/G]G[T/A]GGGG	gatgggg (80.18)	Motif 2 - [T/G]G[T/A]GGGG	gatgggA (69.35)	-13.51 %
+741	Motif 1 - CTAGAGGT	gtagatct (64.69)	Motif 1 - CTAGAGGT	Atagatct (67.32)	+4.06 %

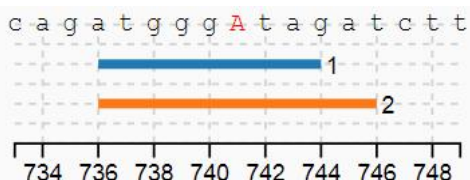
Predicted signal	Prediction algorithm	cDNA Position	Interpretation
(E) Interpreted Data			
ESS Site broken	1 - Fas-ESS hexamers		Alteration of an intronic ESS site. Probably no impact on splicing.
	2 - IIEs from Zhang et al.		

Table S15 The prediction of the consequence of rs408223 mutation

cDNA position	Splice site type	Motif	New splice site	Wild type	Mutant	If cryptic site use, exon length variation	Variation (%)
(A) HSF Matrices							
-683	Acceptor	tctctgccccaggc	tctctgccccagGG	92	92.24	+671	+0.26
-673	Acceptor	aggctccctgagct	agggtccctgagCT	72.1	70.19	+661	-2.65
-673	Donor	aggctccct	AGGgtccct	41.51	68.35	+661	New site +64.66
(B) Branch Points							
No difference between mutant and reference sequences was found with this matrice.							

cDNA position	Linked SR protein	Reference motif (value 0-100)	Linked SR protein	Mutant motif (value 0-100)	Variation
(C) ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins					
-676	SF2/ASF (IgM-BRCA1)	cccaggc (80.00)	SF2/ASF (IgM-BRCA1)	cccaggG (87.62)	+9.52 %
-676	SF2/ASF (IgM-BRCA1)	cccaggc (80.00)	SF2/ASF	cccaggG (82.18)	+2.72 %
-676	SF2/ASF	cccaggc (73.62)	SF2/ASF (IgM-BRCA1)	cccaggG (87.62)	+19.02 %
-676	SF2/ASF	cccaggc (73.62)	SF2/ASF	cccaggG (82.18)	+11.63 %
-672	SC35	ggctccct (80.76)			Site broken -100
-671	SC35	gctccctg (84.02)	SC35	gGtccctg (96.50)	+14.85 %
-670	SF2/ASF (IgM-BRCA1)	ctccctg (76.08)			Site broken -100

cDNA position	Sironi motif reference	Reference silencer (value 0-100)	Sironi mutant motif	Mutant silencer (value 0-100)	Variation
(D) Silencer motifs from Sironi et al.					
-675			Motif 1 - CTAGAGGT	ccaggGtc (61.02)	New site -15.12
-671	Motif 3 - TCTCCCAA	gctccctg (84.74)	Motif 3 - TCTCCCAA	gGtccctg (70.41)	-16.91 %

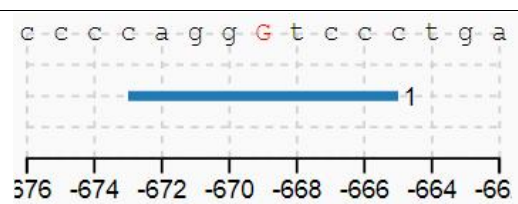
Predicted signal	Prediction algorithm	cDNA Position	Interpretation
(E) Interpreted Data			
New Donor Site	1 - HSF Matrices		Activation of an intronic cryptic donor site. Potential alteration of splicing.

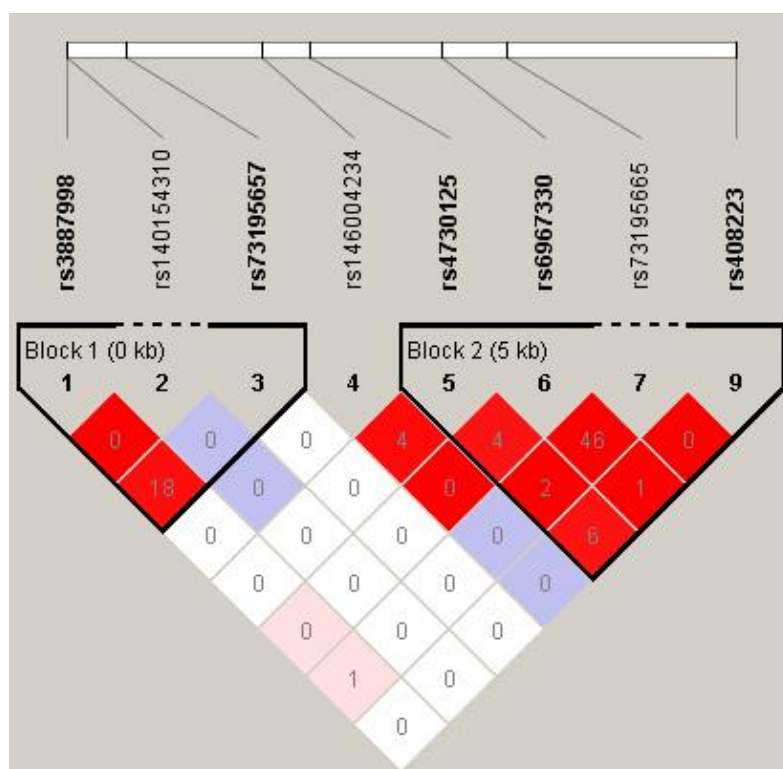
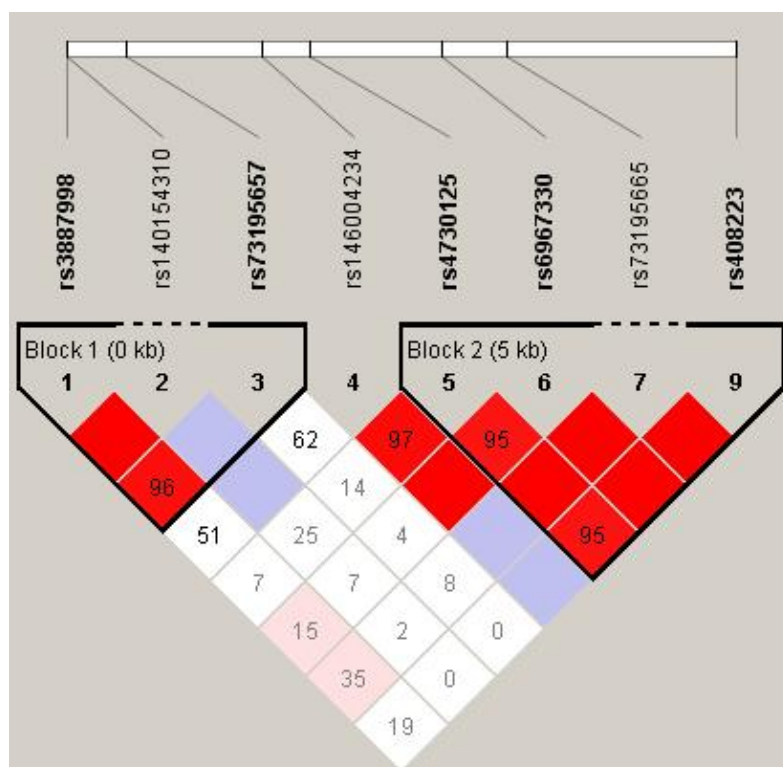
Table S16 Correction of multiple comparisons by false discovery rate (FDR) at a level of 0.05

Comparisons on Current wheeze vs No current wheeze in preschool cohort			Comparisons on Wheeze ever vs Controls in preschool cohort			Comparisons on Current wheeze vs Controls in preschool cohort			Comparisons on Recurrent wheeze vs Controls in preschool cohort		
Observed p values (p_i)	Rank (i)	FDR thresholds ($0.05 \times i/8$)	Observed p values (p_i)	Rank (i)	FDR thresholds ($0.05 \times i/22$)	Observed p values (p_i)	Rank (i)	FDR thresholds ($0.05 \times i/22$)	Observed p values (p_i)	Rank (i)	FDR thresholds ($0.05 \times i/22$)
0.016	1	0.00625	0.079	1	0.00625	0.035	1	0.00625	0.035	1	0.00625
0.017	2	0.0125	0.232	2	0.0125	0.058	2	0.0125	0.077	2	0.0125
0.083	3	0.01875	0.313	3	0.01875	0.15	3	0.01875	0.083	3	0.01875
0.215	4	0.025	0.345	4	0.025	0.205	4	0.025	0.116	4	0.025
0.386	5	0.03125	0.372	5	0.03125	0.316	5	0.03125	0.144	5	0.03125
0.500	6	0.0375	0.482	6	0.0375	0.33	6	0.0375	0.145	6	0.0375
0.719	7	0.04375	0.585	7	0.04375	0.427	7	0.04375	0.158	7	0.04375
0.855	8	0.05	0.704	8	0.05	0.713	8	0.05	0.183	8	0.05

Comparisons on Frequency of current wheeze in preschool cohort			Comparisons on Asthma diagnosis in school-aged cohort			Comparisons on asthma diagnosis with ICS use in school-aged cohort			Comparisons on FEV _{0.5} in preschool cohort		
Observed p values (p_i)	Rank (i)	FDR thresholds (0.05 x i/8)	Observed p values (p_i)	Rank (i)	FDR thresholds (0.05 x i/22)	Observed p values (p_i)	Rank (i)	FDR thresholds (0.05 x i/22)	Observed p values (p_i)	Rank (i)	FDR thresholds (0.05 x i/22)
0.016	1	0.00625	0.025	1	0.00227	0.002	1	0.00263	0.004	1	0.00625
0.018	2	0.0125	0.027	2	0.00455	0.008	2	0.00526	0.039	2	0.0125
0.081	3	0.01875	0.113	3	0.00682	0.016	3	0.00789	0.236	3	0.01875
0.204	4	0.025	0.142	4	0.00909	0.027	4	0.01053	0.44	4	0.025
0.393	5	0.03125	0.151	5	0.01136	0.143	5	0.01316	0.572	5	0.03125
0.49	6	0.0375	0.167	6	0.01364	0.21	6	0.01579	0.653	6	0.0375
0.698	7	0.04375	0.168	7	0.01591	0.216	7	0.01842	0.675	7	0.04375
0.839	8	0.05	0.2	8	0.01818	0.251	8	0.02105	0.935	8	0.05
			0.287	9	0.02045	0.289	9	0.02368			
			0.307	10	0.02273	0.39	10	0.02632			
			0.329	11	0.025	0.407	11	0.02895			
			0.334	12	0.02727	0.46	12	0.03158			
			0.335	13	0.02955	0.519	13	0.03421			
			0.35	14	0.03182	0.661	14	0.03684			
			0.378	15	0.03409	0.664	15	0.03947			
			0.452	16	0.03636	0.67	16	0.04211			
			0.744	17	0.03864	0.74	17	0.04474			
			0.847	18	0.04091	0.779	18	0.04737			
			0.884	19	0.04318	0.843	19	0.05			
			0.902	20	0.04545						
			0.935	21	0.04773						
			0.98	22	0.05						

Comparisons on FVC in preschool cohort			Comparisons on FEV _{0.5} /FVC in preschool cohort		
Observed <i>p</i> values (<i>p_i</i>)	Rank (i)	FDR thresholds (0.05 x i/8)	Observed <i>p</i> values (<i>p_i</i>)	Rank (i)	FDR thresholds (0.05 x i/22)
0.03	1	0.00625	0.015	1	0.00625
0.037	2	0.0125	0.179	2	0.0125
0.076	3	0.01875	0.449	3	0.01875
0.152	4	0.025	0.525	4	0.025
0.665	5	0.03125	0.605	5	0.03125
0.691	6	0.0375	0.765	6	0.0375
0.823	7	0.04375	0.789	7	0.04375
0.976	8	0.05	0.942	8	0.05

Figure S1 Haploview analysis of the LD patterns for *CDHR3* SNPs, with the numbers inside the boxes being the D' (upper) and r^2 (lower) values for different SNP pairs



Compliance with the STREGA Guideline

Item	Item number	Answer
Title and Abstract	1	Stated in manuscript P.1-2
Introduction		
Background rationale	2	Stated in manuscript lines 54-72
Objectives	3	Stated in manuscript lines 72-77
Methods		
Study design	4	Stated in manuscript lines 32-38
Setting	5	Stated in manuscript lines 83-95
Participants	6	Stated in manuscript lines 83-95
Variables	7	Stated in manuscript lines 81- 125; For the details of ISAAC, please refer to reference 12
Data sources/measurement	8	DNA was extracted from buccal swab samples in preschool cohort and from peripheral blood samples in school-age cohort with QIAamp DNA Mini Kit (Qiagen) according to standard protocol. For details, please refer to references 10, 11 and 13-15
Bias	9	
Study size	10	
Quantitative variables	11	
Statistical methods	12	Stated in details in online supplement P.2-3
Results		
Participants	13	Stated in manuscript lines 158-164
Descriptive data	14	Stated in Tables S1- S3
Outcome data	15	Stated in Table S1
Main results	16	Stated in manuscript P.7-8
Other analyses	17	Stated in manuscript P.9-10
Discussion		
Key results	18	Stated in manuscript P.10-12
Limitations	19	Stated in manuscript P.13
Interpretation	20	Stated in manuscript P.13
Generalizability	21	
Other information		
Funding	22	Stated in manuscript lines 323-326