A Case-only Genome-Wide Association Study on Gene-Sex Interaction in Allergic Rhinitis

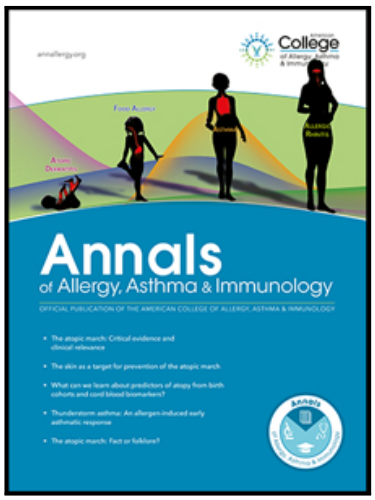
Mohammadnejad, Afsaneh; Brasch-Andersen, Charlotte; Li, Weilong; Haagerup, Annette; Baumbach, Jan; Tan, Qihua

Published in:
Annals of Allergy, Asthma, & Immunology

DOI:
10.1016/j.anai.2018.07.001

Publication date:
2018

Document version:
Accepted manuscript



A Case-only Genome-Wide Association Study on Gene-Sex Interaction in Allergic Rhinitis

Afsaneh Mohammadnejad , Charlotte Brasch-Andersen , Weilong Li , Annette Haagerup , Jan Baumbach , Qihua Tan

PII: S1081-1206(18)30533-7
DOI: 10.1016/j.anai.2018.07.001
Reference: ANAI 2613

To appear in: *Annals of Allergy, Asthma Immunology*

Received date: 8 March 2018
Revised date: 18 June 2018
Accepted date: 1 July 2018

Please cite this article as: Afsaneh Mohammadnejad , Charlotte Brasch-Andersen , Weilong Li , Annette Haagerup , Jan Baumbach , Qihua Tan , A Case-only Genome-Wide Association Study on Gene-Sex Interaction in Allergic Rhinitis, *Annals of Allergy, Asthma Immunology* (2018), doi: 10.1016/j.anai.2018.07.001

A Case-only Genome-Wide Association Study on Gene-Sex Interaction in Allergic Rhinitis

Afsaneh Mohammadnejad[1,2,#], Charlotte Brasch-Andersen[3,4,#], Weilong Li[1], Annette Haagerup[5], Jan Baumbach[2,6], Qihua Tan[1,3,*]

1. Unit of Epidemiology and Biostatistics, Department of Public Health, University of Southern Denmark, Denmark
2. Department of Mathematics and Computer Science, University of Southern Denmark, Denmark
3. Unit of Human Genetics, Department of Clinical Research, University of Southern Denmark, Denmark
4. Department of Clinical Genetics, Odense University Hospital, Denmark
5. NIDO | danmark, Gødstrup Hospital, and Human Genetics, Institute of Clinical Biomedicine, Aarhus University, Denmark
6. Chair of Experimental Bioinformatics, TUM School of Life Sciences, Technical University of Munich, Munich, Germany

# Equal contributions

*Addresses for correspondence:

Qihua Tan, MD, PhD

Epidemiology and Biostatistics, Department of Public Health, University of Southern Denmark

J. B. Winsløws Vej 9B, DK-5000, Odense C, Denmark

Tel.: +45 65503536 Fax: +45 65411911

E-mail: qtan@health.sdu.dk

Conflict of Interest: none

Funding Source: This work was supported by the Novo Nordisk Foundation grant R121-A10459 awarded to Dr. Charlotte Brasch Andersen, the Region of Southern Denmark Research Grant #13/26014 awarded to Professor Qihua Tan. The work of JB was supported by JB's Villum Young Investigator grant nr. 13154.

Abbreviations: none

Allergic rhinitis (AR) is a condition with a significant impact on patient's quality of life. Sex differences in the prevalence and clinical presentation of rhinitis have been reported in multiple studies. Additionally, it has comorbidities with asthma, atopic dermatitis, rhinosinusitis, otitis media, anosmia, nasal polyps and lower airway infection. As a complex phenotype, AR is affected by both genetic and environmental factors. Recent literature has focused on difference in the genomic architecture of asthma for the two sexes by considering gene-sex interaction, whereas AR has been neglected. The cost-effective case-only design is an efficient approach for detecting interactions, compared to the case-control design. In addition, the hypothesis-free genome-wide association study (GWAS) is a promising approach to discovering novel phenotype-associated variants. Application of the case-only design to GWAS should help with detecting interaction effects efficiently.

We conducted a case-only GWAS to detect gene-sex interaction in AR in 434 Danish patients, among which 243 siblings and 191 unrelated individuals, 184 males and 250 females (eTable 1). Sample collection and diagnostic criterions have been described in detail elsewhere.[1] SNP genotyping was conducted using the Affymetrix Genome-Wide Human SNP array 5.0 containing nearly 500,000 SNPs. Before statistical analysis, SNPs with a minor allele frequency < 5% and call rate < 90% were removed resulting in 363,536 genotyped SNPs. Based on the observed SNP genotypes and linkage disequilibrium (LD) information from the 1,000 Genomes

Project and International HapMap Project, SNP imputation was performed using IMPUTE2 software package with The Genome Reference Consortium Human build 37 as reference. After imputation, SNPs with the measured information metric (Info score) < 0.5 were removed. The final number of genotyped and imputed SNPs included in the analysis was 1,348,667. The case-only GWAS was carried out by fitting a logistic regression model of sex on SNP genotype, assuming additive genetic effect using the *geepack* package in R which accounts for sibling correlation. Our analysis showed no genome-wide significant SNP with p-value < 5×10^{-8} but 12 suggestive SNPs with p-value < 1×10^{-5}. These SNPs are positioned on chromosomes 2, 4, 5, 7 and 16 including six imputed and six genotyped SNPs (eFigure 1). A cluster of suggestive SNPs on chromosome 5 were plotted in eFigure 2 with top SNP rs566750 ($\beta = 0.67$, p-value 2.8×10^{-6}) at position 134,513,184 base pair (bp) mapped to C5orf66 gene.

Next, we performed a group test of SNPs within a gene to boost the power using VEGAS2. The gene-based test considers LD between SNPs estimated from the 1000 Genome Project and assigns SNPs to genes based on hg19 genomic location, and then uses a simulation approach for gene-based testing. The analysis found 1194 genes with p-value < 0.05. The list of top 20 genes ordered by p-values are presented in Table 1, with the top four genes *TWF1* (p-value = 5.0×10^{-5}), *PUS7L* (p-value = 7.2×10^{-5}), *C5orf66* (p-value = 7.2×10^{-5}) and *IRAK4* (p-value = 7.3×10^{-5}). Note that *C5orf66* located on 5q31.1 was identified by both SNP and gene-based analyses. *C5orf66* is a protein coding gene and has not yet been reported to associate with atopy. In Table 1, *IRAK4* on 12q12 hosts SNP rs4251459 ($\beta = 1.06$, p-value = 1.28×10^{-5}). Zhang and colleagues[2] found two SNPs (rs4251431 and rs6582484) in *IRAK4* displaying significant association with AR in males, while in our result, rs4251459 appears to increase the risk of AR in females compared to males. The *GALNT14* gene (Table 1) has been associated

with otitis media (OM) in one GWAS.[3] AR has been frequently reported to associate with OM in different populations.[4] Our result provides further evidence that the two diseases (AR and OM) could be genetically connected. Chromosome 1p22.2 also showed an important association to AR because it harbors three genes *GBP4, GBP7* and *GBP2* (Table 1), each plays a role in the innate immune functions.[5] Since the innate immune system is indispensable for autoimmunity and AR is an autoimmune disease, the detected genes are sensible although they are subject to further evaluations. From chromosome 2, *CUL3* hosting top SNP rs11688397 is a protein-coding gene related to the innate immune system. This gene is near *DOCK10* on 2q36.2 which was identified as being associated with asthma.[6] In addition, the expression of *TFAP2B* (Table 1) has been shown to distinguish controlled and therapy-resistant childhood asthma.[7]

The summary statistics from the gene-based test were then used for pathway-based analysis to cluster genes into pathways using VEGAS2 which calculated pathway-based test and empirical p-values to obtain the significance of each detected pathway. The top 10 pathways are shown in eTable 2. The top pathway "Trophectodermal cell differentiation" comprises eight genes, among which *CDH1* is known to relate with airway remodeling and lung function in asthma, as SNPs within this gene are associated with epithelial E-cadherin expression.[8] Moreover, dysfunction in epithelial cell morphogenesis, a sub-pathway of cell morphogenesis, leads to loss of differentiation, reduced junctional integrity, and impaired innate defense which predates atopy and development of allergic disease.[9] Children with asthma were likely to have higher triglyceride level than those without asthma. (remove citation 5 and add new citation here, then re-order citations)

Finally, we applied HaploReg v4 to investigate the significance of non-coding regions related to AR in the regulatory units outside the coding regions. It used the list of query SNPs with p-values < 10^{-4} from GWAS and calculated the coverage of strong enhancers for each cell type and then compared with background set using a binomial test statistic. At the end, enrichment was reported based on p-values < 0.05. Among the summary results, one enhancer, "Esophagus" was identified with a binomial p-value of 0.02. Eosinophilic esophagitis (EE) is a chronic allergy, causing inflammation of the esophagus. Interestingly, it has been reported that there is a similar histopathology between EE and atopic diseases and most of the patients with EE have allergy disorders.[10]

Overall, our case-only GWAS revealed evidence of gene-sex interaction in AR either in support of previously published results or presenting novel findings that could be considered as a reference for future verifications.

Acknowledgements

We thank professor Jørgen Vestbo from Unit of Lung Medicine at the Department of Clinical Research, University of Southern Denmark for his contribution in patients collection for this study.

REFERENCE

1. Kruse LV, Nyegaard M, Christensen U, Møller-Larsen S, Haagerup A, Deleuran M, et al. A genome-wide search for linkage to allergic rhinitis in Danish sib-pair families. Eur J Hum Genet.

2012 Sep;20(9):965-72. doi: 10.1038/ejhg.2012.46.

2. Zhang Y, Lin X, Desrosiers M, Zhang W, Meng N, Zhao L, et al. Association pattern of interleukin-1 receptor-associated kinase-4 gene polymorphisms with allergic rhinitis in a han Chinese population. PLoS One, 2011;6(6): e21769. doi:10.1371/journal.pone.0021769.

3. Rye, M. S, Warrington, N. M, Scaman, E S.H, Vijayasekaran S, Coates H.L, Anderson D, et al. Genome-wide association study to identify the genetic determinants of otitis media susceptibility in childhood. PLoS ONE, e48215 (2012).

4. Kreiner-Møller E, Chawes BLK, Caye-Thomasen P, Bønnelykke K, Bisgaard H. Allergic Rhinitis Is Associated With Otitis Media With Effusion: A Birth Cohort Study. Clin Exp Allergy. 2012; 42(11):1615–1620.

5. Gerrit J.K.Praefckea. Regulation of innate immune functions by guanylate-binding proteins. INT J MED MICROBIOL, 2018; 308(1): 237-245.

6. Romina A. Ortiz and Kathleen C. Barnes. Genetics of allergic diseases. Immunol Allergy Clin North Am. 2015; 35(1):144.

7. Helena Persson, Andrew T. Kwon, Jordan A. Ramilowski, Gilad Silberberg, Cilla Söderhäll, Christina Orsmark-Pietras, et al. Transcriptome analysis of controlled and therapy-resistant childhood asthma reveals distinct gene expression profiles. J Allergy Clin Immunol 2015; 136:638-48.

8. D. Ierodiakonou, D.S. Postma, G.H. Koppelman, H.M. Boezen, J. Gerritsen, N. ten Hacken, et al. E-cadherin gene polymorphisms in asthma patients using inhaled corticosteroids. Eur Respir J, 2012; 38:1044–1052.

9. Schleimer RP, Berdnikovs S. Etiology of epithelial barrier dysfunction in patients with type 2 inflammatory diseases. J Allergy Clin Immunol. 2017 Jun;139(6):1752-1761.

10. Ram G, Lee J, Ott M, Brown-Whitehorn TF, Cianferoni A, Shuker M, et al. Seasonal exacerbation of esophageal eosinophilia in children with eosinophilic esophagitis and allergic rhinitis. Ann Allergy Asthma Immunol. 2015;115(3):224-228.e1. doi: 10.1016/j.anai.2015.07.004.

Table 1: Description and test statistics of top 20 genes from gene-based analysis using VEGAS2.

Chr	Gene	nSNPs	Start	Stop	Test	P-value	Top SNP	Top SNP P-value
12	TWF1	72	44137525	44250178	583.49	5.00e-05	rs4251459	1.28e-05
12	PUS7L	80	44072409	44202620	726.80	7.20e-05	rs4251459	1.28e-05
5	C5orf66	198	134318969	134730370	653.91	7.20e-05	rs566750	2.81e-06
12	IRAK4	83	44102746	44233346	719.80	7.30e-05	rs4251459	1.28e-05
2	GALNT14	238	31083330	31411592	1154.93	1.04e-04	rs2028677	3.87e-04

5	LOC644936	43	79544916	79646297	302.89	1.12e-04	rs11744364	1.87e-04
6	TFAP2B	83	50736438	50865326	566.26	1.39e-04	rs7772880	5.09e-04
12	LINC00615	69	91261799	91392446	441.28	1.87e-04	rs1920765	7.15e-04
1	GBP4	27	89596830	89714633	271.77	2.08e-04	rs604593	2.77e-04
1	GBP7	60	89547433	89691723	566.53	2.20e-04	rs2026043	2.65e-04
16	LOC100506172	123	73370703	73505295	512.21	2.24e-04	rs727014	1.53e-05
2	CUL3	138	225284866	225500114	1085.94	2.42e-04	rs11688397	2.23e-04
5	SPZ1	50	79565789	79667660	303.84	2.64e-04	rs11744364	1.87e-04
14	CGRRF1	38	54926586	55055334	233.90	3.10e-04	rs4143720	3.65e-05
5	CRSP8P	51	79596423	79697785	268.44	3.68e-04	rs11744364	1.87e-04
2	CXCR1	33	218977567	219081716	237.10	5.13e-04	rs1008562	6.28e-04
1	GBP2	59	89521815	89641842	474.76	5.65e-04	rs2026043	2.65e-04
11	EI24	71	125389282	125504584	413.15	5.74e-04	rs515735	7.79e-04
18	SIGLEC15	63	43355544	43472521	389.03	6.12e-04	rs8085782	1.60e-04
6	TFAP2D	65	50631256	50790746	579.81	6.41e-04	rs7772880	5.09e-04