

Next Generation Exome Sequencing of paediatric Asthma Identifies Rare and Novel Variants in Candidate Genes

Neda Bogari (✉ nmbogari@uqu.edu.sa)

Umm Al-Qura University College of Medicine <https://orcid.org/0000-0002-7071-0492>

Amr A Amin

University of Umm alqura

Husni H Rayes

maternity and child hospital

Ahmed Abdelmotelb

Tanta University

Mohiuddin M Taher

Umm Al-Qura University College of Medicine

Faisal A. Al-Allaf

Umm Al-Qura University College of Medicine

Abdellatif Bouazzaoui

Umm Al-Qura University

Luke O'Gorman

University of Southampton

John W Holloway

University of Southampton

Research article

Keywords: Asthma, next generation sequence, exome, variants

Posted Date: June 2nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-31295/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Multiple genes have been implicated to have a role in asthma predisposition by association studies. Paediatric patients often manifest more extensive disease and a particularly severe disease course. It is likely that genetic predisposition could play a more substantial role in this group. This study aims to identify the spectrum of rare and novel variation in known paediatric asthma susceptibility genes using whole exome sequencing-analysis in nine individual cases of childhood onset allergic asthma. Data were processed through an analytical pipeline to align sequence reads, conduct quality checks, identify and annotate variants where patient sequence differed from the reference sequence.

Results: DNA samples from the nine children with a history of bronchial asthma diagnosis underwent targeted exome capture and sequencing. For each patient, the entire complement of rare variation within strongly associated candidate genes was catalogued. The analysis showed 21 variants in the subjects, 13 had been previously identified and 8 were novel. Also, amongst of which, nineteen were non-synonymous and 2 were nonsense. With regard to the novel variants, the 2 non-synonymous variants in the *PRKG1* gene (*PRKG1*: p.C519W and *PRKG1*: p.G520W) were presented in 4 cases, and a non-synonymous variant in the *MAVS* gene (*MAVS*: p.A45V) was identified in 3 cases. The variants we found in this study will enrich the variants spectrum and build up the database in the Saudi population. Novel eight variants were identified in the study which provides more evidence in the genetic susceptibility in asthma among Saudi children.

Conclusion: Screening a cohort of Saudi children for molecular identification of polymorphisms associated with allergic asthmatic response. These, together with the clinical phenotypes, revealed genetic determinants for paediatric asthma and also we compared to the similar previous reports.

Providing a genetic screening map for the molecular genetic determinants of allergic disease in Saudi children, with the goal of reducing the impact of chronic diseases on the health and the economy. We believe that the advanced specified statistical filtration/annotation programs used in this study succeeded to release such results in preliminary study, exploring the genetic map of that disease in Saudi children.

Background

Asthma and other allergic diseases, including allergic rhinitis, eczema and food allergy, cause a substantial burden of disease in childhood. Although a rapid increase in asthma and allergies has been identified over the latter part of the 20th century, the reasons for this are still unknown [1]. Recent changes in environmental factors and their interactions with genetic profiles have been suggested as major factors responsible for the increase in asthma and allergic diseases [2].

Asthma, a chronic inflammatory respiratory condition characterized by hyper-responsive airways and reversible airflow obstruction is a substantial public health problem that affects nearly 155 million individuals worldwide with the prevalence of current asthma is higher in children compared than adults [3–5]. Although environmental factors are important, there are strong genetic predispositions for the development of allergic diseases. It has been reported that there are more than 100 candidate genes in every chromosome which are identified to have a linkage with asthma and the strength of association of these single nucleotide polymorphisms (SNPs) with asthma varies in different parts of the world [6–8]. A better knowledge of asthma susceptibility will hold promise for a better understanding of the pathology, diagnosis, prevention, treatment and management of this increasingly frequent disease.

Next generation sequencing (NGS) technology, in particular exome sequencing, currently represents the most powerful and cost effective approach to identifying variation in the human genome and has already been shown to uncover important disease-causing variation missed by GWAS studies [9–16]. Previous studies have implicated rare variants in asthma and asthma-related traits for a number of relevant genes [17, 18], suggesting that rare variants may indeed play an important role in asthma susceptibility.

Chromosome 17q21 was the first asthma susceptibility locus discovered by genome wide association studies (GWAS) [13, 17, 19–22]. However, none of the genes within the locus had previously been implicated in asthma pathogenesis. SNPs in 17q21 showing highly significant associations with childhood asthma correlate with the expression of *ORMDL3* transcripts, suggesting *ORMDL3* was a plausible asthma candidate gene in the locus [1]. Later, allele-specific gene expression was also observed for other genes in the 17q21 locus [23].

At least twelve GWAS of asthma have been conducted and have yielded numerous associations, with the most significant (and in most cases replicated) associations occurring in or near the following genes: *ORMDL3* [1, 2], *PDE4D* [4], *HLADRB1* [2], *HLA-DQ* [2, 6], *RAD50-IL13* [6], *DENND1B* [7], *TLE4* [8], *SMAD3* [2], *IL1RL1* [5], *IL18R1* [2], *IL33* [2], *IL2RB* [2], *RORA* [2], and *SLC22A5* [2]. These findings have greatly expanded our understanding of the disease, having identified several novel genetic loci that had never previously been implicated in the pathogenesis of asthma (e.g. *ORMDL3*, *RAD50*, *DENND1B*, and *TLE4*). Despite these successes, no definitive causal variants have been identified in any of these genes. It is asserted that the associated variants are in LD with the causal variants in these genes, but more effort must be made to identify causal variants so that the biology of these genes in the aetiology of asthma can be better understood.

Despite the success of GWAS in identifying the associated genetic variants for complex diseases and more than 1000 studies in the past few years were conducted to identify the genetic complexity of many immune diseases including asthma, this approach still could -in part - explain the heritability of asthma regarding the clinical prediction of phenotypic heritability and immunological pathways [24].

This study aims to reveal the genetic determinants for paediatric asthma in Saudi Arabia using whole exome sequencing technique, reviewing the results with similar international studies.

Methods

Recruitment of paediatric asthma cohort of patients

Seventy-nine children included in this study were initially selected from the paediatric allergy and immunology clinic, maternity children hospital Makkah, KSA between Jan 2014 and Oct 2015. All children had been diagnosed with asthma after clinical evaluation in addition to both physiological assessment including pulmonary function tests (PFT), and the use of the International study of asthma and allergies in childhood (ISAAC) questionnaire (modified for the population) (Note that PFT was not considered for young age who was not able to perform the test properly). They were aged between 5 and 14 years at the time of recruitment. Written informed consent was obtained from the attending parents of all the children. In the initial recruitment interview, clinical data and venous blood samples (3 ml of whole blood for CBC and DNA extraction and 3 ml for plasma separation) were collected.

Additional comprehensive clinical data were extracted from their medical records with their consent. For each patient, the information gathered included gender, dates of birth and initial diagnosis, laboratory investigations, physiological assessment, disease history, parents' history for any allergic and/or other autoimmune diseases, medication history (use of steroids, immunomodulators, biological therapies), and history of potential allergen such as carpet, plants and/or animal exposure. Only four non-asthmatic children were recruited (three male (75%) and one female), sequenced in the study as a control group.

Selection of samples for whole exome analysis

Asthmatic patients were recruited based on the Global Initiative for Asthma (GINA) guidelines: A history of respiratory symptoms such as shortness of breath, chest tightness, wheezing and coughing that varies in intensity over time, as well as variable expiratory airflow limitation. Nine patient samples were selected to represent the asthma cohort for exome sequencing based on change to final diagnosis (Table 1).

Table 1
Summary of selected patients phenotypes and characteristics

ID	Diagnosis	Gender	Age at diagnosis (year)	Blood eosinophil (%)	PFT FEV1% Predicted	Treatment
1	BA	M	5.1	0.2	89%	β-agonists, and steroids
2	BA	F	5.9	0.22	90%	β-agonists, and steroids
3	BA/Obese	M	8	0.26	89%	β-agonists, and steroids
4	BA	M	6	0.18	92%	β-agonists, and steroids
5	BA	M	5.5	0.41	83%	β-agonists, and steroids
6	BA	M	6.5	0.29	87%	β-agonists, and steroids
7	BA	M	6	7.57%	NA	β-agonists,, anti-histaminic
8	BA (uncontrolled)	M	7.5	10.6%	NA	β-agonists, and steroids
9	BA	M	5	0.32	87%	β-agonists, and steroids

Table legend: Phenotypic data of the patients (in addition to clinical and physiological assessment data including pulmonary function test, blood eosinophil percent, and the treatment used. PFT not performed for both samples 7 & 8)
 *BA: Bronchial asthma
 *BA (unc): Bronchial asthma (with uncontrolled symptoms)
 *PFT: Pulmonary function test
 FEV1 -FULL FORM.

DNA extraction

Genomic DNA was extracted from EDTA anti coagulated peripheral venous blood samples using a spin-column method using the Applied Biosystem DNA extraction kit. All Genomic DNA was quantified using a Nano-Drop 2000 spectrophotometer (Thermo Scientific).

Whole exome sequencing

Whole exome sequencing was performed for nine patients and four controls DNA samples using Ion Torrent technology (Thermo Fisher Scientific). Libraries were prepared following the protocol described in Ion AmpliSeq Exome RDY library preparation user guide (Publication number MAN0010084) using Ion AmpliSeq Library Kit Plus (Cat. No. 4488990). One hundred ng of genomic DNA was used in library preparation with Ion AmpliSeq Exome RDY Kit 1 × 8 (Cat. No. A38262) then the libraries were purified using a magnetic plate and subsequently quantified using Ion Library TaqMan quantitation kit (Part number 4468802) and real-time PCR instrument 7500 Fast (Life Technologies, USA). Each sample was assigned a distinct barcode using an IonXpress barcode adapters 1–16 kit (Cat. No. 4471250). Barcoded libraries were diluted to 100 pM, pooled three sample's libraries and subjected to template preparation on Ion Sphere Particles (ISP) using the Ion OneTouch 2 System and the Ion PI Template OT2 Hi-Q Kit. Each templated ISP was loaded on the Ion PI v3 chip and sequenced on the Ion Proton instrument (Life Technologies, USA) using Ion PI Hi-Q sequencing Kit. Base calling and alignment were performed on a Torrent Suite v4.2 Server.

Informatic analysis

Filtering low quality reads and removal of adapters were followed by alignment against the human reference genome (hg19 build). GATK v3.7 [25] was used to call SNPs and short indels in a multisample VCF file. GATK v3.7 variant quality score recalibration (VQSR) was applied to the dataset. Annotation was performed using AnnoVar [26] with a database of RefSeq transcripts [27], the Exome Aggregation Consortium (ExAC) [28] and pathogenicity scores including Sort Intolerant From Tolerant (SIFT) [29], PhyloP, Phast Cons[30], Genomic Evolutionary Rate Profiling (GERP++) [31] and Combined Annotation Dependent Depletion (CADD) [32].

Verify BamID v1.1.13 software was used to estimate possible contamination and a 'free mix' value threshold of 0.02 was applied [33]. Gene coverage was determined using SAM tools v1.3.1 [34] and BED tools v2.17.0 [35]. Variants with GATK quality score recalibration (VQSR) tranche > 99% were excluded. Synonymous variants and variants which were found in ExAC (all populations) > 5% were also excluded from the analysis. Variants were prioritized by damaging pathogenicity scores (SIFT < 0.05, CADD Phred > 15[36]and GERP++ > 2).

Selection of a panel of known asthma target genes

All studies in the GWAS catalogue with "asthma" as a phenotype or keyword till 28 March 2017 were reviewed, and studies identifying asthma susceptibility variants that included paediatric subjects were identified and the significant loci retrieved. Following this, 16 additional genes from four studies were added following manual review of the literature. For a list of studies contributing to the candidate gene list (see additional table 3). In a total twenty loci encompassing 131 potential candidate genes amongst which 110 genes were covered in this analysis.

Results

Verify BamIDf reemix values for eight samples were below the 0.02 threshold and did not indicate substantial levels of contamination. One sample (subject 9) showed evidence for contamination and was excluded from further analysis. The minimum mean read depth was 47.7X and the minimum coverage at 10X depth was 86.6% across the eight samples (additional Table 2). Order of Tables 1,2,3.

Table 2
Twenty one shortlisted variants following the application of filter criteria for 8 Saudi Arabian paediatric asthma patients.

chr	Position	REF	ALT	Gene	Consequence	Amino Acid	rsID	Novel	1-sift	CADD Phred	gERP	Sample IDs	
chr1	152488110	G	A	CRCT1	nonsynonymous SNV	p.R84K	.	.	yes	.	22.8	4.73	H
chr1	158908886	C	T	PYHIN1	nonsynonymous SNV	p.S143F	rs147827524	0.0107	.	0.007	23.1	1.27	
chr1	180651520	G	C	XPR1	nonsynonymous SNV	p.A32P	.	.	yes	0.013	30	5.71	
chr2	218683301	C	T	TNS1	nonsynonymous SNV	p.D1127N	rs184451758	0.0006	.	0.013	25.7	4.74	
chr4	38799539	T	A	TLR1	nonsynonymous SNV	p.H305L	rs3923647	0.0296	.	0.002	21.1	4.79	Het
chr4	119952527	C	T	SYNPO2	nonsynonymous SNV	p.P866L	rs745975778	2.47E-05	.	0.042	29.8	5.76	Het
chr4	119978796	C	T	SYNPO2	nonsynonymous SNV	p.R1134W	rs61529635	0.0057	.	0	26.2	3.99	
chr6	32946010	G	T	BRD2	nonsynonymous SNV	p.K562N	.	.	yes	0.128	22.8	2.29	
chr6	32946011	C	A	BRD2	nonsynonymous SNV	p.H563N	.	.	yes	0.304	20.5	5.24	
chr6	165703388	A	T	C6orf118	nonsynonymous SNV	p.I430N	rs202221904	4.97E-05	.	0.037	22.8	1.26	
chr8	118184784	G	A	SLC30A8	nonsynonymous SNV	p.R325Q	rs16889462	0.0163	.	0.099	19.95	-2.29	
chr9	35704426	C	T	TLN1	nonsynonymous SNV	p.A1984T	rs35642290	0.0033	.	1	20.9	5	Het
chr10	53822301	A	G	PRKG1	nonsynonymous SNV	p.N267S	rs34997494	0.0445	.	0.053	23.3	5.76	H
chr10	54041969	T	G	PRKG1	nonsynonymous SNV	p.C519W	.	.	yes	0	29.1	5.49	Het
chr10	54041970	G	T	PRKG1	nonsynonymous SNV	p.G520W	.	.	yes	0	34	5.49	Het
chr12	80645886	G	A	OTOGL	nonsynonymous SNV	p.R388Q	rs747186265	3.81E-05	.	0.064	35	5.95	
chr19	7267725	C	A	INSR	nonsynonymous SNV	p.G95W	.	.	yes	0	33	5.19	
chr19	7267726	A	C	INSR	stopgain	p.Y94*	.	.	yes	.	32	-8.95	
chr19	35434582	G	T	ZNF30	nonsynonymous SNV	p.G239W	rs142299823	0.0214	.	0.004	26.1	2.12	Het
chr19	58213743	G	A	ZNF154	stopgain	p.R192*	rs74939505	0.0088	.	.	24.9	-2.67	
chr20	3842992	C	T	MAVS	nonsynonymous SNV	p.A45V	rs779234123	8.27E-06	.	0	26.9	3.22	

Twenty one shortlisted variants following the application of filter criteria for 8 Saudi Arabian paediatric asthma patients. Chrom, chromosome; Location hg19; REF, reference allele; ALT, alternative allele; Gene, gene symbol; Exonic. Func, consequence of the variant; Amino acid, Amino acid change; dbSNP144, rs; Alternate allele frequency from ExAC database (all populations); Novel; 1-sift, sort intolerant from tolerant; CADD Phred, Combined Annotation Dependent De; genomic evolutionary rate profiling; Samples IDs, "het" indicates a heterozygous genotype. Note : Sample number 9 that primarily selected was removed after

In total of 8 subjects, 21 variants were identified following filtering for potential functionality, amongst which 13 had previously been identified in dbSNP and 8 were novel (Table 2). Nineteen of the 21 variants were non-synonymous and 2 were nonsense (stop gain). Amongst the 13 non-synonymous variants in dbSNP, 2 variants have been previously studied. The rs3923647 was reported to be associated with increased production of the Th1 cytokines, IFN-γ and IL-2, following BCG vaccination [37]. The variant rs16889462 was identified and encoded for SLC30A8 gene reported for many functions in T2D pathophysiology involving the lowering of T2D risk in case of reduced SLC30A8 gene-activity [38]. With regard to the novel variants, the 2 non-synonymous variants in the PRKG1 gene (54041969 and 5404970) were present in 4 cases (1, 3, 4 and 5), and a non-synonymous variant in the MAVS gene (3842992) was identified in 3 cases (4, 5, and 9).

Discussion

This report describes the results of a pilot study of exome sequencing in Saudi children diagnosed with Asthma. Exome sequencing of eight children in this study has identified 21 potentially deleterious SNPs in known asthma genes; among them eight novel variants and 13 variants previously deposited in dbSNP.

In this study, the genetic association was not repeated in the selected samples because this study was not designed as family-based study. Our study was presented to define - for the first time in KSA- the genetic variants that are mostly associated with the asthma development in paediatric age in Makkah region through a large asthma GWAS cohort study, recognizing the possible genetic causes that implicated in asthma and perhaps also suggesting related biological pathways that play a role in the pathogenesis of asthma. Using advanced filtration and annotation programs, the total resulted variants were reduced to only 28 candidate genetic-variants on 10 loci were associated with childhood-onset disease.

On chromosome 1, three genetic variants were identified in three different samples (cases 3,4,7) with two novel variants. One was reported previously, rs147827524 encoded for pyrin and HIN domain family member 1 gene (PYHIN1) that has been accounted for HIN200 proteins which are primarily nuclear proteins involved in transcriptional regulation of genes important for cell cycle control, differentiation, and apoptosis in addition to a surprisingly large proportion of asthma risk in people of African descent[39]. The two novel SNPs discovered, rs152488110 and rs180651520 encoded for both cysteine-rich C-terminal 1 gene (CRCT1) and xenotropic and polytropic retrovirus receptor 1 gene (XPR1), respectively. The CRCT1 SNP is located near the end of the exon region in the gene and among many non-synonymous reported SNPs. Hence, we supposed that it will be of non-significant effect on the protein function. However, the second variant, XPR1 has already been mentioned to play a role in modulating human airway smooth muscle (ASM) contraction, cell growth, and pro-inflammatory cytokine production that promote broncho-constriction, airway inflammation, and remodelling in asthma [40]. On chromosome 2, only one variant was identified that was reported before, rs184451758 encoded for tensin 1 gene (TNS-1), the gene that has been accounted essentially for myofibroblast differentiation and extracellular matrix formation. The polymorphism in that gene is reported to be significantly associated with FULL FORM COPD risk[41]. *On chromosome 4*, three polymorphic variants were identified, one that previously reported, rs3923647, that encoded to Toll-like receptors1 (TLR1) polymorphisms that seems to play a role in susceptibility to asthma, atopic eczema, and allergic rhinitis [42]. The other two polymorphisms identified rs745975778 and rs61529635 were reported for the same gene, synaptopodin 2(SYNPO2) that was reported to be associated with total serum IgE in asthmatics in an independent GWAS, suggesting roles for this gene in asthma[43]. **On chromosome 6**, two closed sequentially novel-variants were identified, rs32946010 and rs32946011 encoded for the same transcription factor element gene, bromodomain (BRD2) located on exon 8 (of 12 exons for that gene) where its protein has been shown to modulate transcription, in particular, in cell cycle-induced transcriptional activation. It has been reported recently that BRD2 protein inhibition attenuates neutrophil-dominant allergic airway disease in mice models [44]. **On chromosome 8**: one non-synonymous polymorphic variant was identified, rs16889462 that was reported previously encoding one of the zinc efflux transporters, solute carrier family 30 member 8 (SLC30A8) which has been classified as one of the major components for providing zinc to insulin maturation and/or storage processes in insulin-secreting pancreatic β -cells. Five genome-wide association studies (GWAS) identified SLC30A8 polymorphism rs13266634 among Asian and European but not African populations[45]. **On chromosome 9**: only one reported polymorphism was identified, rs35642290 that encoded for the cytoskeletal protein, talin 1 (TLN-1) which is considered as one of the genes that might be associated with total IgE in asthmatics[46]. **On chromosome 10**: Surprisingly, three variants in the same gene were identified including one previously reported SNP and two sequential novel ones, rs54041969 and rs54041970 that identified in 50% of samples (cases 1,3,4, and 5) located on exon 14 (of 18 exons reported for that gene). All three variants encode the same protein kinase c GMP-dependent 1 gene (PRKG-1). All PRKG protein- isoforms act as key mediators of the nitric oxide/cGMP signalling pathway and are important components of many signal transduction processes in diverse cell types [47]. It was reported that asthma is typically associated with high levels of exhaled nitric oxide (NO) which reduce the normal levels of S-nitrosothiols, which act as a bronchodilator in the airway [47].

On chromosome 12: one previously reported SNP, rs747186265 was identified – in only one sample (case number 6)-that encoded for the otogelin like gene/protein, which has been accounted to be expressed in the inner ear of vertebrates with the highest level of expression seen at the embryonic stage. No significant ear complications were recorded for that child in our study. We suppose that this variant does not significantly involve in the pathophysiology of asthma development. Also, we suppose that this SNP does not affect the protein structure and function even it is classified as non-synonymous, and/or we recommend performing further studies for that particular SNP effects. *On chromosome 19*: four polymorphisms were identified in our study including two previously reported SNPs, rs142299823 and rs74939505 encoded for both genes of zinc finger protein family, ZNF30 and ZNF154, respectively which involved in the process of DNA binding transcription factor activity., Additionally two novel variants were identified both in two children in the analysis (cases 4 and 5), non-synonymous (rs7267725) and stop-gain (rs7267726). Both variants encoded for the insulin receptor (INSR) gene region that was suggestively associated with asthma risk [48]. **On chromosome 20**: only one non-synonymous previously reported variant, rs779234123 was identified - in three samples (cases, 4,5,8) -, that encoded for mitochondrial antiviral-signalling gene (MAVS), which expresses the protein required for protein kinase activity which is essential for gene expression. Impaired antiviral interferon expression may be involved in asthma exacerbations commonly caused by rhinovirus infections in asthmatic patients [49].

Numerous genetic and molecular studies have been carried out in the field of asthma in both the children and in adults previously [50–60]. Among genetic studies, exome/NG sequencing studies has been documented throughout the global populations [61, 62]. The confirmed observations along with the genetic conclusions confirm the attractive genetic susceptibility factors in asthma patients. It is possible that genetic and non-genetic novel and documented variants might play a major role in the Saudi asthma patients. However, the limited patients number and missing the screening of novel variants present a limitation of this study.

Conclusions

In conclusion, this is the initial exome sequenced study implemented in the Saudi children diagnosed with asthma. Based on early studies and the results we found in this study, we assume that genetic variants might play a role in the increase susceptibility for the development of asthma. Other variants present in this study cannot be avoided considering the high number of loci and its specific genetic role involved within the disease in the global population. Future studies recommend to screen more patients for novel variants within the Saudi population to rule out its role in the asthma disease.

Abbreviations

DNA

Deoxyribonucleic acid

EDTA

Ethylene Diamine tetra acetic acid

SNP

Single Nucleotide Polymorphism

QC

Quality control

KSA

Kingdom of Saudi Arabia

Declarations

Ethics statement

This study was approved by both the Medical Ethical Committee of the Faculty of Medicine, Umm Al Qura University (Ethical approval certificate No 18 dated 8/2/1435H) and the Committee of Ethics and Medical Research in the Maternity & Children Hospital Makkah (Certificate of Approval No 47/25/107862 dated 11/12/2013).

Consent for publication

All the team of the study pleased to submit an original research article entitled "Next Generation Exome Sequencing of paediatric Asthma Disease Identifies Rare and Novel Variants in Candidate Genes for consideration for publication in BMC Genetics journal.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its additional information files:

1. Additional file: Table 1: Mean depth, coverage at 10x depth across target region and Verify Bam ID free mix values for each sample.
2. Additional file: Table 2: Gene coverage for 107 genes from the asthma gene panel.
3. Additional file: Table 3 list of selected genes associated with asthma

Competing interests

All investigators have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Funding

This Project was funded by the National Plan for Science, Technology, and Innovation

(MAARIFAH)-King Abdul-Aziz City for Science and Technology-the kingdom of Saudi Arabia,

award number (13-BIO890-10).

Neda M Bogari^{1*}, Amr A Amin^{2,3*}, Husni H Rayes⁴, Ahmed Abdelmotelb⁵, Mohiuddin M Taher^{1,6}, Faisal A. Al-Allaf¹, Abdellatif Bouazzaoui^{1,6}, Luke O'Gorman⁷, John W Holloway^{7*}.

Authors' Contributions

Conceptualization: Neda M Bogari, Ahmed Abdelmotelb

Data curation: Hosni El-Rayes

Funding acquisition: Neda M Bogari

Investigation: Hosni El-Rayes, Amr Amin

Methodology: Mohiuddin M, Amr A. Amin, Hosni El-Rayes

Resources: Neda M Bogari

Supervision: John W Holloway, Neda Bogari

Data analysis: Luke O’Gorman, Neda M Bogari

Writing ± original draft: Amr A. Amin, Husni H Rayes

Review & editing: Amr A. Amin, Abdellatif Bouazzaoui, Faisal A. Al-Allaf

Validation: Neda M Bogari, Amr A. Amin, John W Holloway

Acknowledgements

We are thankful to Dr. Ahmed Shawky and the staff of the Science and Technology Unit (STU) at Umm-Al-Qura university, Makkah for their continuous support. The authors also acknowledge the support of Mr. Abdulmoniem Gowda, and Mr. Mohammad Adil of Integrated Gulf Biosystems as well as Dr Muhmmmed Noor and Mr. Soud Abdulraof A Khogeer from the Department of Biochemistry, Umm Al-Qura University, Makkah, KSA.

References

1. Braman SS. The global burden of asthma. *Chest* 2006; 130, 4 s-12 s, doi:10.1378/chest.130.1_suppl.4S.
2. Huang SK, Zhang Q, Qiu Z, Chung KF. Mechanistic impact of outdoor air pollution on asthma and allergic diseases. *J Thorac Dis.* 2015;7:23–33. doi:10.3978/j.issn.2072-1439.2014.12.13.
3. Melen E, Nyberg F, Lindgren CM, Berglund N, Zucchelli M, Nordling E, Hallberg J, Svartengren M, Morgenstern R, Kere J, et al. Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease. *Environ Health Perspect.* 2008;116:1077–84. doi:10.1289/ehp.11117.
4. Haagerup A, Bjerke T, Schiøtz PO, Binderup HG, Dahl R, Kruse TA. Asthma and atopy - a total genome scan for susceptibility genes. *Allergy.* 2002;57:680–6.
5. Salam MT, Lin PC, Avol EL, Gauderman WJ, Gilliland FD. Microsomal epoxide hydrolase, glutathione S-transferase P1, traffic and childhood asthma. *Thorax.* 2007;62:1050–7. doi:10.1136/thx.2007.080127.
6. Liu AH, Leung SJ DYM *Nelson textbook of pediatrics*; Elsevier: New Delhi, 2004.
7. Malerba G, Pignatti PF. A review of asthma genetics: gene expression studies and recent candidates. *J Appl Genet.* 2005;46:93–104.
8. Vercelli D. Discovering susceptibility genes for asthma and allergy. *Nat Rev Immunol.* 2008;8:169–82. doi:10.1038/nri2257.
9. Cusanovich DA, Billstrand C, Zhou X, Chavarría C, De Leon S, Michelini K, Pai AA, Ober C, Gilad Y. The combination of a genome-wide association study of lymphocyte count and analysis of gene expression data reveals novel asthma candidate genes. *Human molecular genetics.* 2012;21:2111–23. doi:10.1093/hmg/dds021.
10. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, Martins RA, Kennedy BA, Hassell RG, Visser ME, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet.* 2010;42:684–7. doi:10.1038/ng.628.
11. Holloway JW, Dunbar PR, Riley GA, Sawyer GM, Fitzharris PF, Pearce N, Le Gros GS, Beasley R. Association of beta2-adrenergic receptor polymorphisms with severe asthma. *Clin Exp Allergy.* 2000;30:1097–103.
12. Soler Artigas M, Loth DW, Wain LV, Gharib SA, Obeidat M, Tang W, Zhai G, Zhao JH, Smith AV, Huffman JE, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet.* 2011;43:1082–90. doi:10.1038/ng.941.
13. Wan YI, Strachan DP, Evans DM, Henderson J, McKeever T, Holloway JW, Hall IP, Sayers I. A genome-wide association study to identify genetic determinants of atopy in subjects from the United Kingdom. *J Allergy Clin Immunol.* 2011;127(231.e221–223):223–31. doi:10.1016/j.jaci.2010.10.006.
14. Yao TC, Kuo ML, See LC, Chen LC, Yan DC, Ou LS, Shaw CK, Huang JL. The RANTES promoter polymorphism: a genetic risk factor for near-fatal asthma in Chinese children. *J Allergy Clin Immunol.* 2003;111:1285–92.
15. Kabesch M. Next generation genetics in allergy. *Curr Opin Allergy Clin Immunol* 2010, 10, 407, doi:10.1097/ACI.0b013e32833dc779.
16. Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C, Cruz-Gordillo P, Knoechel B, Asmann YW, Slager SL, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci USA.* 2012;109:3879–84. doi:10.1073/pnas.1121343109.
17. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson W. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med.* 2010;363:1211–21. doi:10.1056/NEJMoa0906312.
18. Schroer KT, Biagini Myers JM, Ryan PH, LeMasters GK, Bernstein DI, Villareal M, Lockey JE, Reponen T, Grinshpun S, Khurana Hershey GK. Associations between multiple environmental exposures and Glutathione S-Transferase P1 on persistent wheezing in a birth cohort. *J Pediatr.* 2009;154(408.e401):401–8. doi:10.1016/j.jpeds.2008.08.040.
19. Akhbari L, Sandford AJ. Genome-wide association studies for discovery of genes involved in asthma. *Respirology.* 2011;16:396–406. doi:10.1111/j.1440-1843.2011.01939.x.

20. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, von Berg A, Bufe A, Rietschel E, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature*. 2007;448:470–3. doi:10.1038/nature06014.
21. Verlaan DJ, Berlivet S, Hunninghake GM, Madore AM, Lariviere M, Moussette S, Grundberg E, Kwan T, Ouimet M, Ge B, et al. Allele-specific chromatin remodeling in the ZBP2/GSDMB/ORMDL3 locus associated with the risk of asthma and autoimmune disease. *Am J Hum Genet*. 2009;85:377–93. doi:10.1016/j.ajhg.2009.08.007.
22. Lluis A, Schedel M, Liu J, Illi S, Depner M, von Mutius E, Kabesch M, Schaub B. Asthma-associated polymorphisms in 17q21 influence cord blood ORMDL3 and GSDMA gene expression and IL-17 secretion. *J Allergy Clin Immunol*. 2011;127:1587–94.e1586. doi:10.1016/j.jaci.2011.03.015.
23. Lu LQ, Liao W. Screening and functional pathway analysis of genes associated with pediatric allergic asthma using a DNA microarray. *Mol Med Rep*. 2015;11:4197–203. doi:10.3892/mmr.2015.3277.
24. Hebbiring SJ. The challenges, advantages and future of phenome-wide association studies. *Immunology*. 2014;141:157–65. doi:10.1111/imm.12195.
25. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–303. doi:10.1101/gr.107524.110.
26. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164. doi:10.1093/nar/gkq603.
27. Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2005;33:D501–4. doi:10.1093/nar/gki025.
28. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–91. doi:10.1038/nature19057.
29. Ng PC, Henikoff SSIFT. Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31:3812–4.
30. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15:1034–50. doi:10.1101/gr.3715005.
31. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. 2010;6:e1001025. doi:10.1371/journal.pcbi.1001025.
32. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46:310–5. doi:10.1038/ng.2892.
33. Jun G, Flickinger M, Hetrick KN, Romm JM, Doheny KF, Abecasis GR, Boehnke M, Kang HM. Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am J Hum Genet*. 2012;91:839–48. doi:10.1016/j.ajhg.2012.09.004.
34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–9. doi:10.1093/bioinformatics/btp352.
35. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–2. doi:10.1093/bioinformatics/btq033.
36. Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, Liu X. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Human molecular genetics*. 2015;24:2125–37. doi:10.1093/hmg/ddu733.
37. Randhawa AK, Shey MS, Keyser A, Peixoto B, Wells RD, de Kock M, Lerumo L, Hughes J, Hussey G, Hawkrigde A, et al. Association of human TLR1 and TLR6 deficiency with altered immune responses to BCG vaccination in South African infants. *PLoS pathogens*. 2011;7:e1002174. doi:10.1371/journal.ppat.1002174.
38. Flannick J, Thorleifsson G, Beer NL, Jacobs SB, Grarup N, Burt NP, Mahajan A, Fuchsberger C, Atzmon G, Benediktsson R, et al. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet*. 2014;46:357–63. doi:10.1038/ng.2915.
39. Barnes KC, Grant AV, Hansel NN, Gao P, Dunston GM African Americans with asthma: genetic insights. *Proceedings of the American Thoracic Society* 2007, 4, 58–68.
40. Ammit AJ, Hastie AT, Edsall LC, Hoffman RK, Amrani Y, Krymskaya VP, Kane SA, Peters SP, Penn RB, Spiegel S. Sphingosine 1-phosphate modulates human airway smooth muscle cell functions that promote inflammation and airway remodeling in asthma. *The FASEB journal*. 2001;15:1212–4.
41. Castaldi PJ, Cho MH, Litonjua AA, Bakke P, Gulsvik A, Lomas DA, Anderson W, Beaty TH, Hokanson JE, Crapo JD. The association of genome-wide significant spirometric loci with chronic obstructive pulmonary disease susceptibility. *Am J Respir Cell Mol Biol*. 2011;45:1147–53.
42. Törmänen S, Korppi M, Teräsjarvi J, Vuononvirta J, Koponen P, Helminen M, He Q, Nuolivirta K. Polymorphism in the gene encoding toll-like receptor 10 may be associated with asthma after bronchiolitis. *Scientific reports*. 2017;7:2956.
43. Bunyavanich S, Schadt EE. Systems biology of asthma and allergic diseases: a multiscale approach. *Journal of Allergy Clinical Immunology*. 2015;135:31–42.
44. Manni ML, Mandalapu S, Salmeron A, Lora JM, Kolls JK, Alcorn JF. Bromodomain and Extra-Terminal protein inhibition attenuates neutrophil-dominant allergic airway disease. *Scientific reports*. 2017;7:43139.
45. Cheng L, Zhang D, Zhou L, Zhao J, Chen B. Association between SLC30A8 rs13266634 polymorphism and type 2 diabetes risk: a meta-analysis. *Medical science monitor: international medical journal of experimental clinical research*. 2015;21:2178.
46. Kim J-H, Cheong HS, Park JS, Jang A-S, Uh S-T, Kim Y-H, Kim M-K, Choi IS, Cho SH, Choi BW. A genome-wide association study of total serum and mite-specific IgEs in asthma patients. *PLoS One*. 2013;8:e71958.
47. Ghosh S, Erzurum SC. Modulation of asthma pathogenesis by nitric oxide pathways and therapeutic opportunities. *Drug Discovery Today: Disease Mechanisms*. 2012;9:e89–94.

48. White MJ, Risse-Adams O, Goddard P, Contreras MG, Adams J, Hu D, Eng C, Oh SS, Davis A, Meade K. Novel genetic risk factors for asthma in African American children: Precision Medicine and the SAGE II Study. *Immunogenetics*. 2016;68:391–400.
49. Akbarshahi H, Menzel M, Ramu S, Mahmutovic Persson I, Bjermer L, Uller L. House dust mite impairs antiviral response in asthma exacerbation models through its effects on TLR 3. *Allergy*. 2018;73:1053–63.
50. Al-Herz W. A Systemic Review on the Prevalence of Atopic Diseases in Children in the Arabian Peninsula. *Medical principles practice: international journal of the Kuwait University Health Science Centre* 2018, 10.1159/000493267, doi:10.1159/000493267.
51. Cao J, Tian L, Li Z, Zhang C, Ji Q, Zhang C, Qian T. Interleukin-7 gene polymorphism rs766736182 associates with the risk of asthma in children. *Journal of clinical laboratory analysis* 2018, 10.1002/jcla.22675, e22675, doi:10.1002/jcla.22675.
52. Charrad R, Kaabachi W, Rafrafi A, Berraies A, Hamzaoui K, Hamzaoui A. IL-8 Gene Variants and Expression in Childhood Asthma. *Lung*. 2017;195:749–57. doi:10.1007/s00408-017-0058-6.
53. Fu G, Fu L, Cai Y, Zhao H, Fu W Association between polymorphisms of glucocorticoid receptor genes and asthma: A meta-analysis. *Cellular and molecular biology (Noisy-le-Grand, France)* 2018, 64, 13–23.
54. Kim JH, Lee SY, Kang MJ, Yoon J, Jung S, Cho HJ, Kim HB, Hong SJ. Association of Genetic Polymorphisms with Atopic Dermatitis, Clinical Severity and Total IgE: A Replication and Extended Study. *Allergy Asthma Immunol Res*. 2018;10:397–405. doi:10.4168/aaair.2018.10.4.397.
55. Lee YL, Chen JH, Wang CM, Chen ML, Hwang BF Association of Air Pollution Exposure and Interleukin-13 Haplotype with the Risk of Aggregate Bronchitic Symptoms in Children. *EBioMedicine* 2018, 29, 70–77, doi:10.1016/j.ebiom.2018.02.008.
56. Liu Z, Li J, Wang K, Tan Q, Tan W, Guo G. Association Between TGF-beta1 Polymorphisms and Asthma Susceptibility Among the Chinese: A Meta-Analysis. *Genet Test Mol Biomarkers*. 2018;22:433–42. doi:10.1089/gtmb.2017.0238.
57. Mei Q, Qu J, Interleukin-13 + 2044. G/A and + 1923C/T polymorphisms are associated with asthma susceptibility in Asians: A meta-analysis. *Medicine*. 2017;96:e9203. doi:10.1097/md.00000000000009203.
58. Sun FJ, Zou LY, Tong DM, Lu XY, Li J, Deng CB. Association between ADAM metalloproteinase domain 33 gene polymorphism and risk of childhood asthma: a meta-analysis. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas* 2017, 50, e6148, doi:10.1590/1414-431x20176148.
59. Zheng Y, Wang H, Luo L, Liao L, You L, Wang J, Li Q. A meta-analysis of the association between CTLA-4 genetic polymorphism and susceptibility of asthma. *Medicine*. 2018;97:e11380. doi:10.1097/md.00000000000011380.
60. Zhu Y, Yan X, Zhai C, Yang L, Li M. Association between risk of asthma and gene polymorphisms in CHI3L1 and CHIA: a systematic meta-analysis. *BMC pulmonary medicine*. 2017;17:193. doi:10.1186/s12890-017-0515-2.
61. Dannemiller KC, Mendell MJ, Macher JM, Kumagai K, Bradman A, Holland N, Harley K, Eskenazi B, Peccia J. Next-generation DNA sequencing reveals that low fungal diversity in house dust is associated with childhood asthma development. *Indoor air*. 2014;24:236–47.
62. DeWan AT, Egan KB, Hellenbrand K, Sorrentino K, Pizzoferrato N, Walsh KM, Bracken MB. Whole-exome sequencing of a pedigree segregating asthma. *BMC medical genetics*. 2012;13:95. doi:10.1186/1471-2350-13-95.

Supplementary Files

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

- [AdditionalFile1.docx](#)
- [AdditionalFile2.docx](#)
- [AdditionalFile3.docx](#)