



Complex genetics of pulmonary diseases: lessons from genome-wide association studies and next-generation sequencing

NIMA POULADI, CHRISTIAN BIME, JOE G. N. GARCIA, and YVES A. LUSSIER

TUCSON, ARIZ AND CHICAGO, ILL

The advent of high-throughput technologies has provided exceptional assistance for lung scientists to discover novel genetic variants underlying the development and progression of complex lung diseases. However, the discovered variants thus far do not explain much of the estimated heritability of complex lung diseases. Here, we review the literature of successfully used genome-wide association studies (GWASs) and identified the polymorphisms that reproducibly underpin the susceptibility to various noncancerous complex lung diseases or affect therapeutic responses. We also discuss the inherent limitations of GWAS approaches and how the use of next-generation sequencing technologies has furthered our understanding about the genetic determinants of these diseases. Next, we describe the contribution of the metagenomics to understand the interactions of the airways microbiome with lung diseases. We then highlight the urgent need for new integrative genomics-phenomics methods to more effectively interrogate and understand multiple downstream “omics” (eg, chromatin modification patterns). Finally, we address the scarcity of genetic studies addressing under-represented populations such as African Americans and Hispanics. (Translational Research 2016;168:22–39)

Abbreviations: DNaseq = DNA sequencing; GWAS = genome-wide association study; NGS = next-generation sequencing

INTRODUCTION

High-throughput technologies have revolutionized our understanding of the molecular mechanisms underlying the development and

progression of complex lung diseases. Importantly, genome-wide association studies (GWASs) have played a critical role as an advanced high-density genotyping approach that allows for characterizing the contribution of single-nucleotide polymorphisms (SNPs) scattered across the genome to the genetic susceptibility of individuals. Specifically, GWAS benefits from the variety of array platform technologies, by them the genotyping of numerous common variants is possible. Until the advent of GWAS in 2007, the genetic research of complex inheritance was labor-intensive and biased, focusing on identifying genetic variants based on prior knowledge. GWAS offered a substantial higher rate of discovery and was not limited by the need for an a priori hypothesis or prior knowledge. However, the limited number of probes in GWAS platforms (eg, 1 million probes) compared with more than 79,000,000 variant sites thus far identified by phase 3 release of 1000 Genomes Project (June 24, 2014) indicates a remaining bias. Importantly, the rapidly reducing cost of next-generation DNA sequencing (DNaseq) paves the way

From the Department of Medicine, University of Arizona, Tucson, Ariz; Center for Biomedical Informatics and Biostatistics, University of Arizona, Tucson, Ariz; BIO5 Institute, University of Arizona, Tucson, Ariz; University of Arizona Health Sciences Center, University of Arizona, Tucson, Ariz; Arizona Respiratory Center, University of Arizona, Tucson, Ariz; Institute for Genomics and Systems Biology, Argonne National Laboratory and University of Chicago, Chicago, Ill.

Submitted for publication February 16, 2015; revision submitted April 27, 2015; accepted for publication April 29, 2015.

Reprint requests: Yves A. Lussier, BIO5 Institute, University of Arizona, 1657 East Helen Street, P.O. Box 210240, Tucson, AZ 85721; e-mail: yves@email.arizona.edu.

1931-5244/\$ - see front matter

© 2016 Elsevier Inc. All rights reserved.

<http://dx.doi.org/10.1016/j.trsl.2015.04.016>

for a paradigm shift in discovery: personal polymorphisms inaccessible by biased methods.

Previous reviews in lung diseases focused primarily on GWAS publications including descriptions of disease-SNP associations observed in only 1 dataset. The present study focused on replicated GWAS discoveries and new discoveries in whole genome sequencing of patients with a variety of lung diseases. We summarize the contribution of GWASs in identifying polymorphisms that confer risk to the development of specific pulmonary disorders and the ones that may influence patient responses to therapy. In addition, we have included a summary of next-generation sequencing (NGS) DNA-seq findings benefiting from various NGS technologies. The advent of NGS has made unbiased sequencing of DNA faster and cheaper than that of Sanger sequencing method. We did not include the lung cancer polymorphisms, as the number of discoveries in that specialty would require an entire different review. Finally, we reflect on the potential considerations and avenues for future studies that would aim to catalog the genetic determinants of complex lung diseases.

METHODS

We used NHGRI GWAS catalog¹ to find the GWAS publications that are related to various primary noncancerous lung diseases; we found 81 curated publications. Among them, we focused on publications with results of discovery cohorts that have been separately and significantly replicated in their replication cohorts. The tables comprise an exhaustive listing of reproducible variants reported in the studied article. The authors might have mapped a certain polymorphism onto more than 1 gene. Therefore, the number of assigned genes might exceed the number of polymorphisms in several studies. We also used query terms of the various pulmonary diseases plus terms related to next-generation, whole exome, or whole genome sequencing in PubMed to find the publications where NGS have been used. We used the original articles and the GeneCards (<http://www.genecards.org/>) to annotate the functions of the genes.

RESULTS

GWAS and complex pulmonary diseases. We divided the studies reviewed into GWAS and NGS categories with GWASs subdivided into 3 sections: obstructive pulmonary diseases (asthma and chronic obstructive pulmonary diseases [COPD]), restrictive pulmonary diseases, and miscellaneous lung diseases. We also provide a systems biology interpretation of discovered variants. The National Human Genome Research Institute (NHGRI) catalog of reproducible published GWAS reports 207

lung disease-associated SNPs from 32 publications. Of these, 134 SNP-disease associations were reproducible in independent cohorts. However, we identified that only 61% of these SNPs were reported by the NHGRI with the correct identifier as originally published in the references. We also identified 13 reproducible lung disease SNPs in these NHGRI-reported publications that were missing in the NHGRI catalog. Here, we review the entirety of these and provide the rectified results in tables.

Obstructive pulmonary diseases. Asthma. Asthma is a common chronic inflammatory disorder of the airways, characterized by episodic and reversible airflow obstruction, airway hyper-responsiveness, and underlying inflammation.² In the United States, children, women, racial and ethnic minorities, residents of inner cities, and economically disadvantaged populations have a disproportionately higher burden of asthma morbidity and mortality when compared with the general population.³⁻⁵ Among children younger than 18 years, asthma is the third major cause of the hospital admissions.⁶ Of note, the pattern of its prevalence is significantly different with respect to various ages, countries with different economic infrastructure, and degree of severity.⁶ Duffy et al estimated that asthma heritability is about 60%,⁷ highlighting the underlying role of genetic determinants in asthma development.

Asthma susceptibility. Asthma genetic variants have been the most studied among pulmonary diseases. Moffatt et al⁸ conducted the first asthma GWAS and uncovered the association of variants of ORMDL3, which downregulates the sphingolipid synthesis, with childhood-onset asthma (Table I). Himes et al subsequently performed a GWAS across asthma patients of various ethnicities²⁴ and discovered novel PDE4D polymorphism in combined populations of Hispanics and European descent but not in African American individuals (Table I). The authors also corroborated the association of the ORMDL3 variants with asthma. DeWan et al in a GWAS found evidence in favor of the association of variants of PDE11A gene with childhood allergic asthma.²⁵ Interestingly, both PDE11A and PDE4D genes belong to the phosphodiesterase superfamily of genes, implying the potential underlying role of this superfamily in the development of asthma.²⁵

In a seminal GWAS comprising about 30,000 cases and controls, Moffatt et al discovered novel predisposing asthma susceptibility variants (Table I).²¹ The list contains interesting genes involved in T-cell responses such as IL2RB and HLA-DQ. Interestingly, the authors found that later-onset asthma is more influenced by major histocompatibility complex (MHC) region (HLA-DQ rs9273349) in contrast to childhood-onset asthma,

Table I. Asthma susceptibility polymorphisms

Study	Sub phenotype		Ancestry				Control + case (replicates included)	Genes	Variants	Reference	
	Susceptibility	Progression	European	AFAM	Hispanic	Asian					Other
GWAS and admixture mapping identify different asthma-associated loci in Latinos: the Genes-environments & Admixture in Latino Americans study	✓				✓			7761	IKZF3, MUC21, MUC22, PBMUCL2	rs907092, 6p21	9
Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype	✓		✓					24,109	ZBTB10, CLEC16A	rs7009110, rs62026376	10
A GWAS identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations	✓	✓	✓				✓	24,913	CDHR3, GSDMB, IL33, RAD50, IL1RL1	rs6967330, rs2305480, rs928413, rs6871536, rs1558641	11
HLA-DQ strikes again: GWAS further confirms HLA-DQ in the diagnosis of asthma among adults	✓		✓		✓			15,054	HLA-DQA1	rs9272346	12
GWASs of asthma indicate opposite immunopathogenesis direction from autoimmune diseases	✓		✓					37,015	IL1RL2, IL1R1, TNIP1	rs3755285, rs12619383, rs1422673	13
GWAS to identify genetic determinants of severe asthma	✓	✓	✓					5855	ORMDL3, IL18R1, IL1R1	rs4794820, rs3771166, rs9807989, rs13035227	14
GWAS identifies PERLD1 as asthma candidate gene	✓						✓	2025	PERLD1	rs2941504	15
Identification of IL6R and chromosome 11q13.5 as risk loci for asthma	✓		✓					57,800	IL6R, LRRC32	rs4129267, rs7130588	16
GWAS identifies 3 new susceptibility loci for adult asthma in the Japanese population	✓						✓	35,083	LOC729675, TSLP, WDR36, NOTCH4, LOC338591, IKZF4	rs7686660, rs1837253, rs404860, rs10508372, rs1701704	17
Meta-analysis of GWASs of asthma in ethnically diverse North American populations	✓		✓	✓	✓		✓	12,561	PYHIN1, IL1RL1, TSLP, IL33, GSDMB	rs1101999, rs3771180, rs1837253, rs2381416, rs11078927	18

GWAS identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations	✓			✓	6420	HLA-DPA1, SLC30A8	rs987870, rs3019885	19
Association between ORMDL3, IL1RL1, and a deletion on chromosome 17q21 with asthma risk in Australia	✓		✓		3436	IL1RL1, ORMDL3	rs10197862, rs6503525	20
PDE11A associations with asthma: results of a genome-wide association scan	✓		✓	✓	607	PDE11A	rs11684634	25
A large-scale, consortium-based GWAS of asthma	✓		✓		30,478	IL18R1, HLA-DQ, IL33, SMAD3, GSDMB, GSDMA, IL2RB, HLA-DRB1, FCER1A, IL13, STAT6, IL4-R/IL21R	rs3771166, rs9273349, rs1342326, rs744910, rs2305480, rs3894194, rs2284033, rs9271300, rs2252226, rs20541, rs167769, rs1859308	21
Variants of DENND1B associated with asthma in children	✓	✓	✓	✓	8956	DENND1B, CRB1	rs2786098, rs1775456	22
A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey	✓		✓		876	SGK493	rs4952590	23
Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene	✓		✓	✓	25,360	PDE4D	rs1588265, rs1544791	24
Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma	✓		✓		4557	ORMDL3	rs7216389	8

Abbreviations: ✓, yes; AFAM, African American; GWAS, genome-wide association study.
The entries of the table have been ordered based on the dates in which the studies were published from the bottom to the top.

which is primarily dominated by 17q region (GSDMA/B) their findings failed to corroborate the role of genetic determinants of lung function in predisposing the individuals to asthma. Afterward, Torgerson et al¹⁸ conducted a large meta-analysis in 2011. Their work included admixed populations of African Americans, African Caribbeans, and Hispanics. They revealed for the first time an association of 4 previously known loci, namely (i) 17q, (ii) Thymic stromal lymphopoietin (TSLP) that enhances the T helper type 2 cell responses, (iii) Interleukin 1 Receptor-Like 1 (IL1RL1), and (iv) Interleukin 33 (IL33) in 3 distinctive ethnicities. Additionally, they discovered the association of novel variants of the PYHIN1 gene in African American asthmatic patients (Table I). The pyrin domain of PYHIN1 is found in proteins that are part of the apoptotic and inflammatory signals and respond to interferon.¹⁸

The contribution of GWAS approaches to uncover the genetic determinants of asthma pathogenesis has not been limited only to cases of European ancestry. Other groups performed GWAS with focus on patients of Asian descent.^{15,17,19} Two groups specifically identified susceptibility polymorphisms of PERLD1 and HLA-DPA1 genes in Chinese and Japanese peoples, respectively (Table I). PERLD1 particularly has some effects on activation and proliferation of T cells.¹⁵ Hirota et al¹⁷ also performed a GWAS in a cohort comprising more than 7000 and 27,000 asthmatic and control adult Japanese individuals, respectively. The authors discovered 5 asthma susceptibility genomic regions, namely 4q31, 5q22, 6p21, 10p14, and 12q13. They mapped these 5 regions onto LOC729675, TSLP-WDR36, NOTCH4, LOC338591, and IKZF4, respectively (Table I). It is of interest that they replicated the correlation of the polymorphism of TSLP with asthma in an additional 499 non-Hispanic individuals of European descent. Furthermore, they corroborated 3 previously known polymorphisms of rs1063355 of HLA-DQ, rs11071559 of RORA, and rs744910 of SMAD3. Recently, Galanter et al⁹ conducted a GWAS that included only Hispanics, who represent an admixed population comprising varying combinations of European, African, and Native American ancestries. Through an admixture-mapping technique, they discovered a novel locus on 6p21 that includes mucin-producing MUC21, MUC22, and PBMUCL2 genes residing in the MHC region. They also corroborated the previously known asthma region 17q21 in which polymorphism of IKZF3 gene, which regulates the development and proliferation of B lymphocyte, was particularly the most significant one (Table I).

Most GWASs have primarily focused on asthma of mild to moderate severity. However, Sleiman et al²² conducted a seminal GWAS in which the target

population comprised children with persistent asthma requiring daily inhaled glucocorticoid therapy. They identified a novel predisposing asthma locus on chromosome 1q31 that is mapped to the gene DENND1B, a modulator of the type 1 helper T-cell cytokine signaling pathway (Table I). They also replicated their findings in a cohort comprising children of African descent. Interestingly, they found that the allele of the discovered variant in the population of African ancestry is opposite to that of children of European descent. Another group also focused on individuals with severe asthma in their GWAS.¹⁴ They found the loci (2q12, 17q12–21) underlying the association with asthma severity overlapped with the ones that had been already discovered in previous GWASs (Table I). Bønnelykke et al hypothesized that acute asthma exacerbation is potentially a distinctive clinical phenotype. Therefore, they identified asthma predisposing polymorphisms of 5 genes, namely CDHR3, IL1RL1, IL33, GSDMB, and RAD50 (Table I).¹¹ The CDHR3 gene that encodes cadherin-related family member 3 was a novel finding, whereas the remainder were previously known to be associated with asthma. Of note, they showed that rs6967330 variant of CDHR3 correlates with higher risk of hospitalization in addition to severe exacerbations. The findings from another GWAS with focus on severe asthmatic patients are also interesting. Li et al found the opposite effects of polymorphisms that correlate with asthma susceptibility and other autoimmune diseases.¹³ For instance, they showed that asthma susceptibility variants of the IL13 and HLA-DRA genes are protective against developing psoriasis and ulcerative colitis, respectively.

Overall, the association of common variants of 7 genes with asthma has been significantly reproduced in at least 2 well-replicated GWASs, namely ORMDL3/GSDMB (17q12),^{8,11,14,18,20,21} TSLP (5q22),^{17,18} IL18R1/IL1R1/IL1RL1 (2q12),^{11,13,14,18,20,21,26} and IL33 (9p24.1) (Tables I and II).^{11,18,21}

Endophenotypes as surrogates for asthma susceptibility.

Specific GWASs assessed the use of the endophenotypes (eg, protein) as surrogates for asthma status. In pioneering work, Ober et al²⁸ conducted a GWAS to identify variants correlating with YKL-40 levels. They had previously shown that YKL-40 (a mediator of airway inflammation) level in blood correlates with the severity of asthma. Their effort uncovered that polymorphisms residing within gene CHI3L1, encoding YKL-40, are significantly associated with quantitative trait YKL-40 level, risk of asthma, and declined lung function (Table II). Gudbjartsson et al²⁶ also considered the circulating eosinophil count as a quantitative trait, because eosinophils are the most observed cell type in the airways of asthmatic patients. They found several

Table II. Asthma susceptibility polymorphisms discovered through endophenotypes

Study	Endophenotype	Ancestry				Control + case (replicates included)	Genes	Variants	Reference
		European	AFAM	Hispanic	Asian				
GWAS of lung function decline in adults with and without asthma	Lung function decline	✓				16,136	DLEU7	rs9316500	27
A large-scale, consortium-based GWAS of asthma	IgE levels	✓				30,478	IL18R1, HLA-DQ, IL33, SMAD3, GSDMB, GSDMA, IL2RB, HLA-DRB1, FCER1A, IL13, STAT6, IL4-R/IL21R	rs3771166, rs9273349, rs1342326, rs744910, rs2305480, rs3894194, rs2284033, rs9271300, rs2252226, rs20541, rs167769, rs1859308	21
Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction	Eosinophil counts	✓			✓	26,722	IL1RL1, IKZF2, GATA2, IL5, SH2B3	rs1420101, rs12619285, rs4143832, rs9494145, rs3184504	26
Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function	YKL-40 levels	✓				1772	CHI3L1	rs4950928	28

Abbreviations: ✓, yes; AFAM, African American; GWAS, genome-wide association study; IgE, immunoglobulin E. The entries of the table have been ordered based on the dates in which the studies were published from the bottom to the top.

novel polymorphisms underlying asthma susceptibility in patients of European descent (Table II). In a novel effort, Imboden et al²⁷ used the rate of the lung function decline as an end point in their GWAS. Although they did not find any significant polymorphism underlying asthma risk, they discovered that a variant of the gene DLEU7 correlates with FEV1 in nonasthmatic individuals (Table II).

Increased circulating levels of immunoglobulin E (IgE) are associated with the development of various inflammatory disorders such as asthma. Therefore, it is of clinical interest to assess whether there is an overlap between the polymorphisms underlying circulating IgE levels and those of asthma susceptibility. Moffatt et al²¹ conducted a large meta-analysis on the individuals of European descent (Table II). The authors identified a novel polymorphism of HLA-DRB1 in addition to confirming previously identified variants influencing IgE level. However, they noticed that the main genetic determinants of asthma susceptibility and those of IgE level were not dependent. Specifically, HLA-DRB1 and HLA-DQB1 that influence the increased IgE level and asthma susceptibility, respectively, are in poor linkage disequilibrium. Importantly, they also showed that controlling for IgE level did not affect the predisposing effect of HLA-DQB1 on asthma. The authors concluded that increment of the IgE level is not probably a causal factor in asthma development but a secondary phenomenon.

Asthma and response to therapeutic interventions. GWAS approaches have been successfully used to uncover variants influencing the response to medications in asthmatic patients. The prevalence of aspirin or nonsteroidal anti-inflammatory drugs sensitivity, especially among patients with glucocorticoid-dependent asthma can be as high as 10%–20%.²⁹ Park et al²⁹ conducted a GWAS and uncovered the association of the polymorphisms of the HLA-DPB1 gene with aspirin intolerance in Korean people (Table III). Managing the signs and symptoms of asthmatic patients is challenging because of the heterogeneity of response to therapy among asthmatic patients.³³ Among various medications, inhaled glucocorticoids are the most commonly prescribed therapy in clinical practice with variable treatment response.³³ Several studies have used GWAS to identify polymorphisms underlying the response to glucocorticoids. Tantisira et al³² showed that rs37972 that maps to the gene GLCCI1 significantly correlates with decreased responses to inhaled glucocorticoids. Another GWAS discovered a novel candidate gene T in asthmatic patients whose SNPs influence the response to inhaled corticosteroids³¹ with a 2- to 3-fold difference in FEV1 response. Wu et al³⁰ showed that inhaled corticosteroids modify the

Table III. Polymorphisms affecting the response to medications in asthmatic patients

Study	Medication response	Ancestry				Control + case (replicates included)	Genes	Variants	Reference
		European	AFAM	Hispanic	Asian				
Inhaled corticosteroid treatment modulates ZNF432 gene variant's effect on bronchodilator response in asthmatics	Bronchodilator response	✓				808	ZNF432	rs12460587, rs3450, rs3752120	³⁰
Genome-wide association study of aspirin-exacerbated respiratory disease in a Korean population	Aspirin exacerbation				✓	1940	HLA-DPB1	rs2281389, rs3117230	²⁹
Genome-wide association identifies the T gene as a novel asthma pharmacogenetic locus	Glucocorticoid response	✓				825	T, AC007922.13-1	rs3127412, rs6456042, rs9955411	³¹
Genome-wide association between GLCCI1 and response to glucocorticoid therapy in asthma	Glucocorticoid response	✓				2104	GLCCI1	rs9955411 rs37972	³²

Abbreviation: ✓, yes; AFAM, African American.

The entries of the table have been ordered based on the dates in which the studies were published from the bottom to the top.

response to β 2-agonist through ZNF432 polymorphisms with patients carrying 2 copies of the mutant allele in the presence of inhaled corticosteroids exhibiting a greater β 2-agonist response.

Chronic obstructive pulmonary diseases. The hallmark of COPD is chronic inflammation of the airways, emphysematous changes in the respiratory bronchioles, and the acini leading to progressive obstruction of airflow, which is largely irreversible. COPD is associated with several comorbidities such as heart failure and diabetes.³⁴ Importantly, the World Health Organization estimates that COPD will be the third principal cause of death by the year 2030.³⁵ Although tobacco smoking is a well-known risk factor for COPD and its related pathogenesis,^{36,37} not all smokers contract COPD, implying the underlying role of genetic factors.³⁸ GWAS has successfully contributed to the discovery of the genetic polymorphisms (Table IV), despite the fact that earlier studies suffer from the lack of replicate cohorts and limited number of sample size. For example, Siedlinski et al identified several loci⁴⁵ but failed to include a replication cohort. Cho et al identified novel loci⁴¹ (FAM13A, RAB4B, CHRNA3/5, IREB2, and HHIP) and recently focused on moderate to severe or severe COPD cases.³⁹ They combined the data from COPDGene, ECLIPSE, NETT/NAS, Norway GenKOLS, and International COPD Genetics Network cohorts in a meta-analysis, which also includes African American patients, to successfully confirm the underpinning roles of previously discovered genes CHRNA3 (a member of the nicotinic acetylcholine receptor family), FAM13A (a GTPase activator), and HHIP (a hedgehog family protein binding), and MMP12 (a member of matrix metalloproteinase family) (Table IV). The authors also found supporting evidence that the genes TGFB2, RIN3, and the locus 19q13 play important roles in COPD susceptibility. For example, they showed that rs4846480 is in strong linkage disequilibrium with rs6684205, which acts in pulmonary tissue as an expression quantitative trait locus for TGFB2.^{39,46} Also, it is of interest that candidate COPD susceptibility genes, such as HHIP and FAM13A, overlap with genes that have a role in the normal function of lung (Tables IV and VI).⁵⁵ It is of interest that Manichaikul et al conducted a GWAS across different ethnic groups in which they considered percentage of emphysema as the outcome. They became able to identify 2 polymorphisms rs7957346 near to SNRPF and rs10947233 within PPT2 genes (Table IV).⁴⁰ However, they replicated the discovered SNPs only in patients of European descent.

Restrictive pulmonary disorders. Interstitial lung diseases. Interstitial lung diseases (ILDs) are a heterogeneous group of pathologic conditions that manifest

clinically and radiologically as diffuse parenchymal lung disease. The pathologic findings of ILDs include various stages and patterns of interstitial inflammation, interstitial fibrosis, and alveolar filling. When a cause is not clearly apparent, the term idiopathic interstitial pneumonia (IIP) is used. Physiologically, ILDs generally manifest as a restrictive ventilatory defect and a gas transfer defect.

Idiopathic pulmonary fibrosis. Idiopathic pulmonary fibrosis (IPF) is the most common and the most severe form of IIP.^{56,57} Despite the dismal prognosis of IPF, a limited number of GWASs have been carried out thus far to uncover the polymorphisms underlying its course of development and progression. Our work and work by others have shown common genetic variants of telomerase reverse transcriptase (TERT) (a regulator of telomerase activity), toll interacting protein (TOLLIP) (an activator of MYD88-dependent nuclear factor of kappa light polypeptide gene enhancer in B-cells (NF- κ B) to adjust the toll-like receptor signaling and membrane trafficking), and MUC5B (an important gel-forming mucin in mucus), to be strongly associated with IPF susceptibility. Mushiroya et al⁴⁹ in a GWAS on a group of Japanese patients with sporadic IPF found an association between the common variants within the gene TERT and IPF (Table V). We previously performed a GWAS to discover novel polymorphisms underlying IPF susceptibility⁴⁷ (Table V) and highlighted the importance of 11p15.5 in the development and progression of IPF. We identified a novel IPF susceptibility variant within TOLLIP and also confirmed the association of previously known MUC5B⁶⁰ with IPF. Of note, we discovered that the minor allele of the novel polymorphism of TOLLIP correlates with the patients' odds of mortality.⁴⁷ Parallel to our work, another group also confirmed the association of IPF with MUC5B at 11p15, TERT at 5p15, and the 3q26 region near telomerase RNA component (TERC).⁴⁸ They also identified other novel polymorphisms both at significant and suggestive level of statistical significance.

Granulomatous disorders. An important subgroup of ILDs includes granulomatous disorders such as sarcoidosis. The pathologic hallmark of sarcoidosis is the presence of noncaseating epithelioid granulomas in the absence of a known etiology. Sarcoidosis is a systemic disorder affecting various organs with a wide range of prognosis from self-limited disease to fatal outcome.⁶¹ The studies focusing on monozygotic twins, families, and various races have provided supporting evidence in favor of putative genetic loci predisposing individuals to sarcoidosis.⁶¹ Previous GWASs have successfully uncovered a limited number of non-HLA common susceptibility loci (Table V). Among various

Table IV. COPD susceptibility polymorphisms

Study	Sub phenotype				Ancestry					Control + case (replicates included)	Genes	Variants	Reference
	Susceptibility	Progression	Endophenotype	Other/ description	European	AFAM	Hispanic	Asian	Other				
Risk loci for COPD: a GWAS and meta-analysis	✓	✓			✓	✓				12,337	CHRNA3, FAM13A, HHIP, MMP12, RAB4B	rs12914385, rs4416442, rs13141641, rs626750, rs7937	39
Genome-wide study of percent emphysema on computed tomography in the general population. The Multi-Ethnic Study of Atherosclerosis Lung/SNP Health Association Resource Study	✓		✓	Percentage of emphysema	✓	✓	✓	✓		7914	SNRPF, PPT2	rs7957346, rs10947233	40
A GWAS of COPD identifies a susceptibility locus on chromosome 19q13	✓				✓					8280	FAM13A, RAB4B, EGLN2, CYP2A6, CHRNA3, CHRNA5, IREB2, HHIP	rs1964516, rs7937, rs11858836, rs13141641	41
Genome-wide association analysis of body mass in COPD	✓				✓					3452	FTO	rs8050136	42
Variants in FAM13A are associated with COPD	✓				✓					9134	FAM13A	rs7671167	43
A GWAS in COPD: identification of 2 major susceptibility loci	✓				✓					5334	CHRNA3, NP_001013641.2, HHIP	rs1051730, rs8034191, rs1828591, rs13118928	44

Abbreviations: ✓, yes; *AFAM*, African American; *COPD*, chronic obstructive pulmonary disease; *GWAS*, genome-wide association study; *SNP*, single-nucleotide polymorphism. The entries of the table have been ordered based on the dates in which the studies were published from the bottom to the top.

Table V. Polymorphisms associated with idiopathic pulmonary fibrosis and sarcoidosis susceptibility

Study	Disease	Sub phenotype		Ancestry					Control + case (replicates included)	Genes	Variants	Reference
		Susceptibility	Progression	European	AFAM	Hispanic	Asian	Other				
Genetic variants associated with IPF susceptibility and mortality: a GWAS	Pulmonary fibrosis	✓	✓	✓					3341	MUC5B, TOLLIP	rs35705950, rs5743890	47
GWAS identifies multiple susceptibility loci for pulmonary fibrosis		✓		✓					8610	DSP, 7q22, 15q14–15, DPP9	rs2076295, rs4727443, rs2034650, rs12610495	48
A GWAS identifies an association of a common variant in TERT with susceptibility to IPF		✓					✓		1711	TERT	rs2736100	49
Genome-wide association analysis reveals 12q13.3-q14.1 as new risk locus for sarcoidosis	Sarcoidosis	✓		✓					9275	OS9	rs1050045	50
A novel sarcoidosis risk locus for Europeans on chromosome 11q13.1		✓		✓					8878	CCDC88B, KCNK4, PRDX5, PCLB3	rs479777	51
A GWAS reveals evidence of association with sarcoidosis at 6p12.1		✓		✓					4138	RAB23	rs10484410	52
GWAS identifies ANXA11 as a new susceptibility locus for sarcoidosis		✓		✓					4470	ANXA11	rs2789679, rs7091565	53
Genome-wide association analysis in sarcoidosis and Crohn's disease unravels a common susceptibility locus on 10p12.2		✓		✓					3582	C10ORF67	rs1398024	54

Abbreviations: ✓, yes; AFAM, African American; IPF, idiopathic pulmonary fibrosis. The entries of the table have been ordered based on the dates in which the studies were published from the bottom to the top.

Table VI. Polymorphisms underlying the normal lung function and susceptibility to miscellaneous pulmonary diseases

Study	Disease/trait	Sub phenotype			
		Susceptibility	Progression	Endophenotype	Other/description
Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension	PVD	✓			
Identification of SPOCK2 as a susceptibility gene for BPD	BPD	✓			
Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function	Normal pulmonary function				Pulmonary function

Abbreviations: ✓, yes; AFAM, African American; BPD, bronchopulmonary dysplasia; pulmonary vascular disease.

genetic determinants that have been discovered thus far, 2 regions are of biological and clinical interest, namely 10q22.3 and 11q13.1. The former contains ANXA11, which is essential for cytokinesis and its product is also detected in the sera of patients with different autoimmune disorders, as a predisposing gene to sarcoidosis that has been replicated in patients of non-European descent, particularly African Americans.⁵³ Other works had shown the association of the latter with other diseases including alopecia areata, Crohn's disease, leprosy, psoriasis, and primary biliary cirrhosis.⁵¹ This region contains CCDC88B, KCNK4, PRDX5, PCLB3 genes with CCDC88B as the most promising putative sarcoidosis susceptibility gene (Table V).⁵¹

Miscellaneous pulmonary diseases. GWAS has successfully interrogated various genomic loci in other pulmonary diseases as well, particularly in bronchopulmonary dysplasia (BPD) and pulmonary arterial hypertension (PAH).

BPD is a chronic form of neonatal respiratory disorders characterized by irreversible airflow obstruction,⁶²

with a significant degree of morbidity.⁶³ About one-third of premature infants in neonatal care units develop BPD.⁶³ Previous twin studies uncovered that the genetic factors strongly underpin the development of BPD.⁶⁴ Despite the clinical importance of BPD, only a limited number of GWASs have identified putative susceptibility genetic loci.^{59,65} Hadchouel et al successfully identified 2 BPD predisposing polymorphisms, rs1245560 and rs1049269 of SPOCK2 gene, in which its expression level is positively correlated with alveolarization of normal rat lung (Table VI).⁵⁹ Importantly, they showed that rs1245560 variant is associated with BPD susceptibility in both populations of European and African descents.

PAH is a rare lung disorder. The pathologic hallmark of PAH is the presence of intimal lesions that consist of eccentric thickening, fibrotic, plexiform, concentric, and dilation lesions in the vessels of the pulmonary arterial circulation. This translates into persistently high pulmonary artery pressures, progressive right heart failure, and a high mortality in untreated patients.^{66,67}

Ancestry					Control + case (replicates included)	Genes	Variants	Reference
European	AFAM	Hispanic	Asian	Other				
✓					2150	CBLN2	rs2217560, rs9916909	58
✓				✓	603	SPOCK2	rs1245560, rs1049269	59
✓					94,612	TNS1, FAM13A, GSTCD- NPNT, HHIP, HTR4, ADAM19, AGER, GPR126, PTCH1, TSHD4, MFAP2, TGFB2, HDAC4, RARB, MECOM, SPATA9, ZKSCAN3, ZNF323, NCR3, ARMC2, CDC123, C10orf11, LRP1, CCDC38, MMP15, CFDP1, KCNE2	rs2571445, rs2869967, rs2045517, rs7671167, rs10516526, rs1032296, rs12504628, rs11100860, rs1980057, rs11168048, rs3995090, rs1985524, rs2277027, rs11134779, rs2070600, rs3817928, rs262129, rs16909859, rs16909898, rs12899618, rs8033889, rs2284746, rs993925, rs12477314, rs1529672, rs1344555, rs153916, rs6903823, rs2857595, rs2798641, rs7068966, rs11001819, rs11172113, rs1036429, rs12447804, rs2865531, rs9978142	55

Previous works had identified the association of several genotypes with PAH, such as *BMPR2* (a member of transmembrane serine or threonine kinases receptor families).⁶⁸ Nevertheless, the finding that mutated *BMPR2* accounts for approximately 80% and 15% of familial PAH and idiopathic PAH patients prompted Germain et al to conduct a GWAS to discover novel susceptibility polymorphisms in these 2 PAH groups.^{58,68} They particularly controlled for *BMPR2* mutation and found 2 novel SNPs, namely rs2217560 and rs9916909, which are in complete linkage disequilibrium. Accordingly, they identified the novel predisposing gene *CBLN2* that resides upstream of the uncovered SNPs and encodes a protein, which binds an adhesion molecule in vascular smooth muscle cells (Table VI).

Systems biology implications of these polymorphisms. Using expression quantitative trait loci (eQTL), we identified 2 interesting findings about the messenger RNAs (mRNAs) in *cis* and *trans* associations

with SNPs reported in Tables I–VI (National Center for Biotechnology Information eQTL Browser; <http://ncbi.nlm.nih.gov/projects/gap/eqtl/index.cgi>). There is a group of 4 asthma risk-associated SNPs in linkage disequilibrium with one another (Fig 1, A) located on the chromosome region 17q12 intragenic to 2 host genes *GSDMB* and *LOC101928947*. These 4 SNPs may be reporting a locus rather than a specific gene, as they adjust the expression values of either *ORMDL3* or *MED24* mRNAs (Fig 1, A). In addition, both asthma risk SNPs (Table I) rs9272346 (2-kb upstream of *HLA-DQA1*) and rs9273349 (2-kb downstream of *HLA-DQB1*) are in eQTL associations with *HLA-DQA1* and *HLA-DQB1* mRNAs. Given the transcription direction of these HLA genes, rs9273349 seems to be within a putative enhancer element, which needs to be experimentally validated. The latter 2 SNPs were also shown in many distinct GWASs that are associated with immune-related diseases (Fig 1, B; National Center for Biotechnology Information eQTL

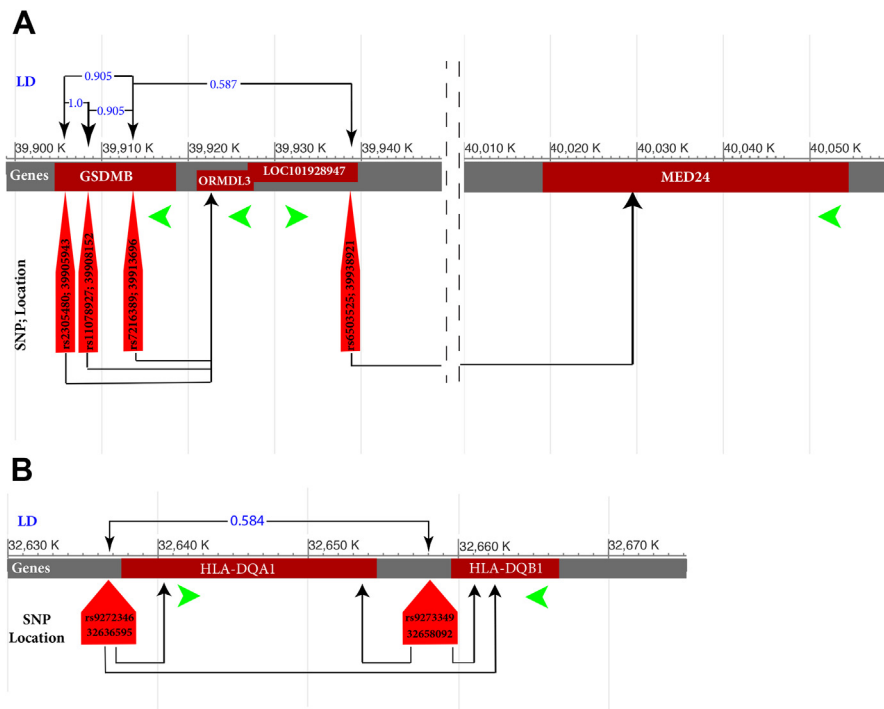


Fig 1. eQTL interaction diagram. The green arrows represent the direction of transcription of the genes. The upper part of each panel corresponds to the degree of linkage disequilibrium values between the SNPs (blue numbers). The arrows at the lower part of each panel show the direction of eQTL effects; the expression values of which genes are adjusted by which SNP (results from eQTL browser). (A) The panel shows that 3 out of 4 well-replicated asthma polymorphisms (Linkage Disequilibrium > 0.9) reside within GSDMB genes when adjusting the expression value of the neighboring ORMDL3 gene. (B) The panel represents how 2 other asthma susceptibility polymorphisms adjust the expression values of HLA-DQA1 and HLA-DQB1. rs9273349 is in an intergenic region, whereas rs9272346 resides 2-kb upstream of the HLA-DQA1. eQTL, expression quantitative trait loci; SNP, single-nucleotide polymorphism. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

browser; celiac diseases, inflammatory bowel diseases, multiple sclerosis, nasal pharyngeal carcinoma, rheumatoid arthritis, schizophrenia, systemic lupus erythematosus, type I diabetes mellitus). We also used the Database for Annotation, Visualization and Integrated Discovery (v6.7; <http://david.abcc.ncifcrf.gov/>), to infer the functional relationship among the gene loci of SNPs reported in Tables I–VI. This analysis shows that the major biological processes to which the candidate lung diseases–associated genes contribute are primarily related to various domains of immune response, such as cytokine receptor activity (Supplement Table 1).

NGS technologies and complex pulmonary diseases. The underlying assumption of GWAS is the “common disease–common variant hypothesis”⁶⁹ in which small number of alleles contribute to the risk of development or progression of a given disease and each allele has small to moderate effect size. Therefore, previous GWAS thus far has aimed to discover the association of genetic variants with minor

allele frequencies (MAF) >5% and in some cases MAF >1% within the trait or disease of interest. However, it is now clear that GWAS discovered variants explain only a small amount of observed heritability of the related traits (missing heritability). In contrast to a common disease–common variant hypothesis, the “common disease, rare variant hypothesis” states that genomic loci that confer risk of disease are cataloged with many rare and independent risk alleles with modest to high effect size that are scattered across the population.⁶⁹ Common disease, rare variant hypothesis implies the presence of an extensive degree of genetic heterogeneity in a putative risk gene or locus and also the lack of linkage disequilibrium among the variants, thereby debunking the inherent assumptions of genotyping-based methods but highlighting the importance of sequencing-based strategies. Several works have used NGS methods to identify different lung disease predisposing genetic variants (Supplement Table 2). Others have used NGS techniques to profile the metagenome of airways and

its interaction with distinct pulmonary ailments (Supplement Table 2).

Predisposing genetic variants to pulmonary diseases using DNaseq. DeWan et al⁷⁰ performed the first whole exome sequencing (WES) with focus on a family enriched with asthma. By resequencing the variants of 3 genes, PDE4DIP, CBLB, and KALRN, they showed that asthmatic members are heterozygote opposite to nonasthmatic ones. They did not find the significant enrichment of selective asthma-related common variants in genes, ORMDL3, PDE11A, PDE4D, and RAD50, within affected members. They were unable to show that the asthmatic members have significantly higher burden of rare variants. Recently, Campbell et al⁷¹ conducted a whole genome sequencing on 16 individuals (8 asthmatic) from a Hutterite population to find novel asthma genetic variants that have not been uncovered in previous GWASs. However, none of their discovered variants met genome-wide significance threshold, specifically mutations within gene NEDD4L, which contributes to the ubiquitination of specific proteins for lysosomal degradation. They were unable to replicate their findings in other populations of Europeans, African Americans, and Hispanics. Wain et al⁷² aimed at finding putative variants protecting smokers from lung function decline. They resequenced the exome of 100 smokers with uncompromised lung function, namely resistant smokers, and 396 individuals from 2 independent control groups. None of their detected signals survived multiple testing corrections. However, they reported rs10859974 as the strongest detected signal in CCDC38 gene, which has been known to be associated with lung function. Although this finding is interesting, it demands to be replicated in additional independent cohorts.

The use of NGS methods has not been limited to only obstructive pulmonary diseases. Austin et al⁷³ used WES to uncover novel hereditary PAH susceptibility variants in a family in which 4 and 8 members were affected and unaffected, respectively. By doing so, they were able to find mutated CAV1 as a new predisposing gene that is a modifier of transforming growth factor beta (TGF- β) signaling at the plasma membrane. Interestingly, they noticed that some unaffected members also are carriers of the same mutation, indicating its incomplete penetrance. Of note, they controlled for the previously identified heritable (HPAH) mutated genes, namely ALK1, ENG, or SMAD9. They replicated their finding in 260 unrelated patients of Caucasian descent with HPAH and idiopathic PAH (IPAH). They also did not detect the mutated CAV1 in the genotype of 1000 ethnically matched healthy Caucasian. Ma et al⁷⁴ also used

WES and studied a family with several affected members. They also ruled out the presence of mutations in previously identified disease predisposing genes. Their effort gave rise to identifying new candidate predisposing gene, KCNK3 that is a potassium channel subfamily K. Importantly, they replicated their result in 92 unrelated patients with familial PAH and also 230 IPH. Although the nuclear families were the target discovery set of the previous 2 studies, Perez et al⁷⁵ interrogated the genomes of 12 unrelated IPAH patients with WES. Consequently, they discovered TopBP1 as a novel IPAH susceptibility gene.

Recently, some works have also used WES techniques to uncover genetic determinants of acute respiratory distress syndrome (ARDS). Shortt et al⁷⁶ studied 96 patients with ARDS (70 Caucasian and 26 African Americans) and identified 76 novel ARDS susceptibility SNPs. They also selectively genotyped 3 of these SNPs, rs3848719 (ZNF335 gene that contributes in proliferation of neural progenitor cells and self-renewal), rs78142040 (ARSD gene that contributes to the development of cartilage and is a member of the sulfatase family), rs9605146 (in XKR3 gene that is a potential membrane transporter), in an additional 117 patients and found the correlation of rs3848719 and rs78142040 with “Acute Physiology and Chronic Health Evaluation II” (APACHE II) score quartile and 60-day mortality.

Metagenome and association of the microbiome with pulmonary diseases. Other works have used NGS techniques to examine the association of microbiome diversity in the respiratory tract with susceptibility to lung disease. In the first effort, Dannemiller et al⁷⁷ assessed the association of the various fungal taxa and their diversity in the house dust with childhood asthma in low-income Hispanic families. Although positive association between various taxa with asthma was not identified, lower fungal diversity, particularly lower diversity within the genus *Cryptococcus*, was positively correlated with asthma susceptibility. Park et al⁷⁸ studied the potential difference of the microbial community of upper respiratory tract of 18 asthmatics and 17 COPD patients compared with 12 healthy controls. Significant difference between the oropharynx microbiome of asthmatics and COPD patients was not discerned, although there was an abundance of *Pseudomonas* spp. from *Proteobacteria* and *Lactobacillus* spp. Garzoni et al⁷⁹ examined the alteration of the composition of the microbiome in upper and lower respiratory tracts of 33 individuals with IIP, sarcoidosis, pneumocystis pneumonia, and healthy controls to determine disruption in respiratory tract microbiota and showed altered upper and lower airways microbiota in 23% patients. Han et al⁸⁰ used the 454 pyrosequencing method and analyzed the degree of the

impact of lung microbiome of 55 IPF patients on the progression of their disease and showed the potential association of a bacterial signature, members within the *Staphylococcus* and *Streptococcus* genera, with the patients' odds of outcome, a correlation that remains significant after adjusting for genetic effects.

DISCUSSION

We have reviewed how GWAS has remarkably contributed to increasing our knowledge about the genetic determinants underlying the development and progression of distinctive pulmonary diseases and their responses to therapeutics. GWAS has helped researchers to discover the polymorphisms that make individuals prone to developing various complex lung diseases such as asthma, COPD, and sarcoidosis. Despite all the successes that GWASs have achieved thus far, the variants discovered with GWASs have yet to explain most of the estimated heritability of various complex diseases and traits,^{81,82} and our understanding about the pathogenesis of pulmonary diseases has remained incomplete.

Although GWAS is a multiplexed interrogation of the genome, the downstream biological effect of many discovered susceptibility variants and the various ways by which they impact on the diseases pathogenesis remain unclear. It is common that researchers map the discovered intergenic variants to the closest neighboring gene. However, the recent discoveries from ENCYClopedia of DNA Elements (ENCODE) Project⁸³ have shown that variants residing in distal regulatory elements might affect the expression of not necessarily neighboring genes through various chromatin modification mechanisms.⁸⁴ Future progress in using new technologies, such as Hi-C, to explain various epigenomic mechanisms in the context of lung diseases will provide pulmonary scientists with a better explanation of the biological impacts of lung disease genetic variants. Integration of other various tiers of data, such as protein interaction⁸⁵ and expression quantitative trait loci data,⁸⁶ is also important to increase the interpretability of GWAS results. Consequently, developing more sophisticated computational techniques and heuristics is necessary.

Much of the GWAS has primarily focused on measuring the significant association of the variants individually with disease susceptibility, progression, or a certain outcome. Nevertheless, it has become exceedingly clear that the discovered lung disease susceptibility genetic variants, similar to other complex diseases, are not capable of explaining all the measured heritability. Part of this missing heritability resides in the interaction between distinctive variants. For example, our work showed that IPF patients might

have a higher rate of mortality despite inheriting the protective allele.⁴⁷ Therefore, devising new innovative experiments to unravel the epistatic interaction is necessary. On the computational front, development of new methods to analyze the GWAS data is also needed, specifically the ones capable of detecting interaction effects effectively in a reasonable computational time.

Another obvious limitation of GWAS, thus far, is the lack of work with a focus on admixed populations, namely African Americans and Hispanics. Most of the GWASs have primarily targeted patients of European descent. For example, we found no GWAS for African American individuals with sarcoidosis in our review. However, some work has reported the role of various ancestries in increment of disease susceptibility and response to therapeutics.⁸⁷ For example, Rumpel et al⁸⁸ found evidence that asthmatic African Americans with the higher African ancestry experience higher rate of severe exacerbation. Another group also reported that Puerto Ricans have a higher asthma prevalence and its related morbidities compared with Mexican Americans,^{89,90} despite the fact that both groups are designated as Hispanics. On the other hand, Choudhry et al⁹¹ found that socioeconomic status (SES) and ancestry significantly interact with each other among asthmatic Hispanic individuals. Particularly, they showed that at the lower spectrum of SES, the European ancestry correlates more with asthma risk, opposite to the other end of the SES in which African ancestry is more associated with the risk of asthma. Therefore, it is crucial to consider patients' SES and ancestry in future GWAS. The under-represented GWAS with focus on admixed populations could also be partly attributed to some limitations in the availability of genotyping technology. Most of the whole-genotyping platforms incorporate polymorphisms specific for subjects of European descent, thereby highlighting the urge for devising more comprehensive genotyping platforms.

Future GWASs should also focus on identifying genetic variants that distinguish various subtypes or clinical manifestations of pulmonary disorders in addition to asthma. For example, although our review identified GWAS results for interstitial pulmonary fibrosis, we found no work related to nonspecific interstitial pneumonia or differentiating between different subtypes of ILD. In addition, our review did not find GWASs identifying lung disease polymorphisms specific to males or females (gender-specific polymorphisms). For instance, although there is a slightly higher prevalence of asthma attack among women (53.5%),⁹² GWASs failed to identify gender-specific asthma polymorphisms when looking for it.^{11,15,21} As another example, the male gender is one of the risk factors for predicting the outcome of IPF,⁹³ yet Fingerlin et al⁴⁸ did not find any significant

association of SNPs with gender. Additionally, although IPH is 1.7 times more common among women,⁶⁷ Germain et al⁵⁸ did not observe any evidence in favor of existing gender-specific polymorphism.

We also summarized the contribution of various NGS methods to uncover the role of rare genetic variants in making individuals prone to various lung diseases. We further explained how these techniques have been helpful for profiling the microbiome of the lung and its interaction with pulmonary diseases. NGS methods are capable of addressing how rare variants confer the disease risk. Of note, the NGS era indeed heralds the path to personalized medicine. The number of published works that have used the novel and massively parallel whole genome sequencing and WES technologies is still limited. However, a gradual drop in the cost during the course of the past few years, which is predicted to continue, will be providing the pulmonary scientist with a great opportunity to search for the rare variants that alter the risk to developing various lung diseases or response to various medications. It is easy to picture that pulmonologists will add the skill of how to interpret the results generated from the NGS methods in the near future to their knowledge repository, subsequently translating them to routine clinical practice.

In view of all these discoveries and the soon to be tsunami of personal polymorphisms, the new challenge is to associate clinically meaningful and actionable meaning with all these discoveries. As precision health requires understanding individuals, this may serve as an additional stimulus to interrogate the genetic makeup of individuals with Hispanic and African American ancestries, as the studies with focus on these populations are lacking. Other strategies include more powerful study designs such as focusing on cases with more extreme and severe phenotypes, more sophisticated GWAS analysis technique, and NGS analyzing methods, as well.

ACKNOWLEDGMENTS

Conflicts of Interest: We appreciate the contribution of Ms Colleen Kenost. Dr Joe G.N. Garcia is the founder, president, and majority shareholder of Aqualung Therapeutics, Corp. Dr Yves A. Lussier is a member of the scientific advisory board of IPSEN corporation (<http://www.ipsen.com/>). All other authors do not have any financial or personal relationship with organizations that could potentially be perceived as influencing the described research. All authors have read the journal's policy on disclosure of potential conflicts of interest.

This manuscript was funded in part by the University of Arizona Health Sciences Center, the University of Arizona Cancer Center (P30CA023074), and the

University of Arizona Center for Biomedical Informatics and Biostatistics.

Supplementary Data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.trsl.2015.04.016>.

REFERENCES

1. Hindorf LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 2009;106:9362–7.
2. National Asthma Education and Prevention Program. Expert Panel Report 3 (EPR-3)—Guidelines for the Diagnosis and Management of Asthma, Summary Report 2007. *J Allergy Clin Immunol* 2007;120(5 suppl 1):S94–138.
3. Moorman JE, Rudd RA, Johnson CA, et al. National surveillance for asthma—United States 1980–2004. Morbidity and Mortality Weekly Report 2007;56:1–54.
4. Moorman JE, Mannino DM. Increasing U.S. asthma mortality rates: who is really dying? *J Asthma* 2001;38:65–71.
5. Moorman JE, Akinbami LJ, Bailey CM, et al. National surveillance of asthma: United States 2001–2010. National Center for Health Statistics. *Vital Health Stat* 2012;3(35):1–67.
6. Eder W, Ege MJ, von Mutius E. The asthma epidemic. *N Engl J Med* 2006;355:2226–35.
7. Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of asthma and hay fever in Australian twins. *Am Rev Respir Dis* 1990;142(6 Pt 1):1351–8.
8. Moffatt MF, Kabesch M, Liang L, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007;448:470–3.
9. Galanter JM, Gignoux CR, Torgerson DG, et al. Genome-wide association study and admixture mapping identify different asthma-associated loci in Latinos: the Genes-environments & Admixture in Latino Americans study. *J Allergy Clin Immunol* 2014;134:295–305.
10. Ferreira MA, Matheson MC, Tang CS, et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *J Allergy Clin Immunol* 2014;133:1564–71.
11. Bønnelykke K, Sleiman P, Nielsen K, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet* 2014;46:51–5.
12. Lasky-Su J, Himes B, Raby B, et al. HLA-DQ strikes again: genome-wide association study further confirms HLA-DQ in the diagnosis of asthma among adults. *Clin Exp Allergy* 2012;42:1724–33.
13. Li X, Ampleford EJ, Howard TD, et al. Genome-wide association studies of asthma indicate opposite immunopathogenesis direction from autoimmune diseases. *J Allergy Clin Immunol* 2012;130:861–8.
14. Wan Y, Shrine N, Artigas MS, et al. Genome-wide association study to identify genetic determinants of severe asthma. *Thorax* 2012;67:762–8.
15. Anantharaman R, Andiappan AK, Nilkanth PP, Suri BK, Wang DY, Chew FT. Genome-wide association study identifies PERLD1 as asthma candidate gene. *BMC Med Genet* 2011;12:170.
16. Ferreira MA, Matheson MC, Duffy DL, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011;378:1006–14.

17. Hirota T, Takahashi A, Kubo M, et al. Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat Genet* 2011;43:893–6.
18. Torgerson DG, Ampleford EJ, Chiu GY, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011;43:887–92.
19. Noguchi E, Sakamoto H, Hirota T, et al. Genome-wide association study identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* 2011;7:e1002170.
20. Ferreira MA, McRae AF, Medland SE, et al. Association between ORMDL3, IL1RL1 and a deletion on chromosome 17q21 with asthma risk in Australia. *Eur J Hum Genet* 2010;19:458–64.
21. Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010;363:1211–21.
22. Sleiman PM, Flory J, Imielinski M, et al. Variants of DENND1B associated with asthma in children. *N Engl J Med* 2010;362:36–44.
23. Castro-Giner F, Bustamante M, González JR, et al. A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey (ECRHS). *BMC Med Genet* 2009;10:128.
24. Himes BE, Hunninghake GM, Baurley JW, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *Am J Hum Genet* 2009;84:581–93.
25. DeWan AT, Triche EW, Xu X, et al. PDE11A associations with asthma: results of a genome-wide association scan. *J Allergy Clin Immunol* 2010;126:871–3.
26. Gudbjartsson DF, Bjornsdottir US, Halapi E, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009;41:342–7.
27. Imboden M, Bouzigon E, Curjuric I, et al. Genome-wide association study of lung function decline in adults with and without asthma. *J Allergy Clin Immunol* 2012;129:1218–28.
28. Ober C, Tan Z, Sun Y, et al. Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 2008;358:1682–91.
29. Park BL, Kim T-H, Kim J-H, et al. Genome-wide association study of aspirin-exacerbated respiratory disease in a Korean population. *Hum Genet* 2013;132:313–21.
30. Wu AC, Himes BE, Lasky-Su J, et al. Inhaled corticosteroid treatment modulates ZNF432 gene variant's effect on bronchodilator response in asthmatics. *J Allergy Clin Immunol* 2014;133:723–8.
31. Tantisira KG, Damask A, Szefer SJ, et al. Genome-wide association identifies the T gene as a novel asthma pharmacogenetic locus. *Am J Respir Crit Care Med* 2012;185:1286–91.
32. Tantisira KG, Lasky-Su J, Harada M, et al. Genomewide association between GLCCII and response to glucocorticoid therapy in asthma. *N Engl J Med* 2011;365:1173–83.
33. Drazen JM, Silverman EK, Lee TH. Heterogeneity of therapeutic responses in asthma. *Br Med Bull* 2000;56:1054–70.
34. Barnes P, Celli B. Systemic manifestations and comorbidities of COPD. *Eur Respir J* 2009;33:1165–85.
35. Berndt A, Leme AS, Shapiro SD. Emerging genetics of COPD. *EMBO Mol Med* 2012;4:1144–55.
36. Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *Lancet* 2009;374:733–43.
37. Shaykhi R, Crystal RG. Early events in the pathogenesis of chronic obstructive pulmonary disease. smoking-induced reprogramming of airway epithelial basal progenitor cells. *Ann Am Thorac Soc* 2014;11(suppl 5):S252–8.
38. Todd JL, Goldstein DB, Ge D, Christie J, Palmer SM. The state of genome-wide association studies in pulmonary disease: a new perspective. *Am J Respir Crit Care Med* 2011;184:873–80.
39. Cho MH, McDonald ML, Zhou X, et al. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med* 2014;2:214–25.
40. Manichaikul A, Hoffman EA, Smolonska J, et al. Genome-wide study of percent emphysema on computed tomography in the general population. The Multi-Ethnic Study of Atherosclerosis Lung/SNP Health Association Resource Study. *Am J Respir Crit Care Med* 2014;189:408–18.
41. Cho MH, Castaldi PJ, Wan ES, et al. A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Hum Mol Genet* 2012;21:947–57.
42. Wan ES, Cho MH, Boutaoui N, et al. Genome-wide association analysis of body mass in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2011;45:304–10.
43. Cho MH, Boutaoui N, Klanderma BJ, et al. Variants in FAM13A are associated with chronic obstructive pulmonary disease. *Nat Genet* 2010;42:200–2.
44. Pillai SG, Ge D, Zhu G, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet* 2009;5:e1000421.
45. Siedlinski M, Cho MH, Bakke P, et al. Genome-wide association study of smoking behaviours in patients with COPD. *Thorax* 2011;66:894–902.
46. Hao K, Bossé Y, Nickle DC, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet* 2012;8:e1003029.
47. Noth I, Zhang Y, Ma S-F, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013;1:309–17.
48. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013;45:613–20.
49. Mushihirota T, Wattanapokayakit S, Takahashi A, et al. A genome-wide association study identifies an association of a common variant in TERT with susceptibility to idiopathic pulmonary fibrosis. *J Med Genet* 2008;45:654–6.
50. Hofmann S, Fischer A, Nothnagel M, et al. Genome-wide association analysis reveals 12q13.3–q14.1 as new risk locus for sarcoidosis. *Eur Respir J* 2013;41:888–900.
51. Fischer A, Schmid B, Ellinghaus D, et al. A novel sarcoidosis risk locus for Europeans on chromosome 11q13.1. *Am J Respir Crit Care Med* 2012;186:877–85.
52. Hofmann S, Fischer A, Till A, et al. A genome-wide association study reveals evidence of association with sarcoidosis at 6p12.1. *Eur Respir J* 2011;38:1127–35.
53. Hofmann S, Franke A, Fischer A, et al. Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. *Nat Genet* 2008;40:1103–6.
54. Franke A, Fischer A, Nothnagel M, et al. Genome-wide association analysis in sarcoidosis and Crohn's disease unravels a common susceptibility locus on 10p12.2. *Gastroenterology* 2008;135:1207–15.
55. Artigas MS, Loth DW, Wain LV, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet* 2011;43:1082–90.
56. Travis WD, Costabel U, Hansell DM, et al. An official American Thoracic Society/European Respiratory Society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2013;188:733–48.
57. Zibrak JD, Price D. Interstitial lung disease: raising the index of suspicion in primary care. *NPJ Prim Care Respir Med* 2014;24:14054.

58. Germain M, Eyries M, Montani D, et al. Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension. *Nat Genet* 2013;45:518–21.
59. Hadchouel A, Durrmeyer X, Bouzigon E, et al. Identification of SPOCK2 as a susceptibility gene for bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2011;184:1164–70.
60. Zhang Y, Noth I, Garcia JG, Kaminski N. A variant in the promoter of MUC5B and idiopathic pulmonary fibrosis. *N Engl J Med* 2011;364:1576–7.
61. Spagnolo P, Grunewald J. Recent advances in the genetics of sarcoidosis. *J Med Genet* 2013;50:290–7.
62. Baraldi E, Filippone M. Chronic lung disease after premature birth. *N Engl J Med* 2007;357:1946–55.
63. Walsh MC, Szefer S, Davis J, et al. Summary proceedings from the bronchopulmonary dysplasia group. *Pediatrics* 2006;117(suppl 1):S52–6.
64. Lavoie PM, Pham C, Jang KL. Heritability of bronchopulmonary dysplasia, defined according to the consensus statement of the national institutes of health. *Pediatrics* 2008;122:479–85.
65. Wang H, Julien KRS, Stevenson DK, et al. A genome-wide association study (GWAS) for bronchopulmonary dysplasia. *Pediatrics* 2013;132:290–7.
66. Liu D, Morrell NW. Genetics and the molecular pathogenesis of pulmonary arterial hypertension. *Curr Hypertens Rep* 2013;15:632–7.
67. McLaughlin VV, McGoon MD. Pulmonary arterial hypertension. *Circulation* 2006;114:1417–31.
68. Simonneau G, Gatzoulis MA, Adatia I, et al. Updated clinical classification of pulmonary hypertension. *J Am Coll Cardiol* 2013;62(25 suppl):D34–41.
69. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev* 2009;19:212–9.
70. DeWan AT, Egan KB, Hellenbrand K, et al. Whole-exome sequencing of a pedigree segregating asthma. *BMC Med Genet* 2012;13:95.
71. Campbell CD, Mohajeri K, Malig M, et al. Whole-genome sequencing of individuals from a founder population identifies candidate genes for asthma. *PLoS One* 2014;9:e104396.
72. Wain LV, Sayers I, Artigas MS, et al. Whole exome re-sequencing implicates CCDC38 and cilia structure and function in resistance to smoking related airflow obstruction. *PLoS Genet* 2014;10:e1004314.
73. Austin ED, Ma L, LeDuc C, et al. Whole exome sequencing to identify a novel gene (caveolin-1) associated with human pulmonary arterial hypertension. *Circ Cardiovasc Genet* 2012;5:336–43.
74. Ma L, Roman-Campos D, Austin ED, et al. A novel channelopathy in pulmonary arterial hypertension. *N Engl J Med* 2013;369:351–61.
75. De Jesus Perez VA, Yuan K, Lyuksytova MA, et al. Whole-exome sequencing reveals TopBP1 as a novel gene in idiopathic pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2014;189:1260–72.
76. Shortt K, Chaudhary S, Grigoryev D, et al. Identification of novel single nucleotide polymorphisms associated with acute respiratory distress syndrome by exome-seq. *PLoS One* 2014;9:e111953.
77. Dannemiller KC, Mendell MJ, Macher JM, et al. Next-generation DNA sequencing reveals that low fungal diversity in house dust is associated with childhood asthma development. *Indoor Air* 2014;24:236–47.
78. Park H, Shin JW, Park S-G, Kim W. Microbial communities in the upper respiratory tract of patients with asthma and chronic obstructive pulmonary disease. *PLoS One* 2014;9:e109710.
79. Garzoni C, Brugger SD, Qi W, et al. Microbial communities in the respiratory tract of patients with interstitial lung disease. *Thorax* 2013;68:1150–6.
80. Han MK, Zhou Y, Murray S, et al. Lung microbiome and disease progression in idiopathic pulmonary fibrosis: an analysis of the COMET study. *Lancet Respir Med* 2014;2:548–56.
81. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–53.
82. Maher B. The case of the missing heritability. *Nature* 2008;456:18–21.
83. ENCODE Project Consortium. The ENCODE (ENCyclopedia of DNA elements) Project. *Science* 2004;306:636–40.
84. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature* 2012;489:109–13.
85. Lee Y, Li H, Li J, et al. Network models of genome-wide association studies uncover the topological centrality of protein interactions in complex diseases. *J Am Med Inform Assoc* 2013;20:619–29.
86. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 2010;6:e1000888.
87. Ortega VE, Meyers DA. Pharmacogenetics: implications of race and ethnicity on defining genetic profiles for personalized medicine. *J Allergy Clin Immunol* 2014;133:16–26.
88. Rumpel JA, Ahmedani BK, Peterson EL, et al. Genetic ancestry and its association with asthma exacerbations among African American subjects with asthma. *J Allergy Clin Immunol* 2012;130:1302–6.
89. Rose D, Mannino DM, Leaderer BP. Asthma prevalence among US adults, 1998–2000: role of Puerto Rican ethnicity and behavioral and geographic factors. *Am J Public Health* 2006;96:880–8.
90. Salari K, Burchard EG. Latino populations: a unique opportunity for epidemiological research of asthma. *Paediatr Perinat Epidemiol* 2007;21(suppl 3):15–22.
91. Choudhry S, Burchard EG, Borrell LN, et al. Ancestry–environment interactions and asthma risk among Puerto Ricans. *Am J Respir Crit Care Med* 2006;174:1088–93.
92. Moorman JE, Person CJ, Zahran HS, Centers for Disease Control and Prevention (CDC). Asthma attacks among persons with current asthma—United States. *MMWR Surveill Summ* 2013;62(suppl 3):93–8.
93. Kolb M, Collard HR. Staging of idiopathic pulmonary fibrosis: past, present and future. *Eur Respir Rev* 2014;23:220–4.