Whole-Genome Sequencing of Pharmacogenetic Drug Response in Racially Diverse Children with Asthma.

Angel CY Mak
Marquitta J. White
Walter L. Eckalbar
Zachary A. Szpiech
Sam S. Oh

See next page for additional authors

Follow this and additional works at: https://scholarlycommons.henryford.com/internalmedicine_articles

Recommended Citation
Whole-Genome Sequencing of Pharmacogenetic Drug Response in Racially Diverse Children with Asthma

Angel C. Y. Mak1,*, Marquitta J. White4, Walter L. Eckalbar1, Zachary A. Szpiech2, Sam S. Oh1, Maria Pino-Yanes3,4, Donglei Hu1, Pagé Goddard1, Scott Huntsman1, Joshua Galanter1,14, Ann Chen Wu5,6, Blanca E. Himes7, Soren Germer8, Julia M. Vogel8,6, Karen L. Bunting8, Celeste Eng1, Sandra Salazar1, Kevin L. Keys1, Jennifer Liberto1, Thomas J. Nuckton1, Thomas A. Nguyen1, Dana G. Torgerson9, Pui-Yan Kwok10,11, Albert M. Levin12, Juan C. Celedón13, Erick Forno13, Hakon Hakonarson14,15, Patrick M. Sleiman14,15, Amber Dahlin3, Kelan G. Tantisira3, Scott T. Weiss3, Denise Serebrisky16, Emerita Brigno-Buenaventura17, Harold J. Farber18, Kelley Meade19, Michael A. Lenoir20, Pedro C. Avila21,22, Saunak Sen16, Shannon M. Thyne22, William Rodriguez-Cintron23, Cheryl A. Winkler24, Andrés Moreno-Estrada25, Karla Sandoval25, Jose R. Rodriguez-Santiago26, Rajesh Kuma27,28, L. Keoki Williams29,30, Nadav Ahituv2,11, Elad Ziv1, Max A. Seibold31, Robert B. Danell3,32,33, Noah Zaitlen1, Ryan D. Hernandez2,10,34, and Esteban G. Burchard1,2; on behalf of the NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium

1Department of Medicine, 2Department of Bioengineering and Therapeutic Sciences, 3Department of Pediatrics, 4Cardiovascular Research Institute, 5Institute for Human Genetics, and 6Quantitative Biosciences Institute, University of California San Francisco, San Francisco, California; 7Department of Biostatistics, Epidemiology and Informatics and 15Department of Pediatrics, Perelman School of Medicine, Harvard Medical School and Pilgrim Health Care Institute, Boston, Massachusetts; 8New York Genome Center, New York, New York; 12Department of Public Health Sciences, 13Division of Pediatric Pulmonary Medicine, Allergy and Immunology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; 14Center for Applied Genomics, The Children’s Hospital of Philadelphia Research Institute, Philadelphia, Pennsylvania; 15Pediatric Pulmonary Division, Jacobi Medical Center, Bronx, New York; 16Department of Allergy and Immunology, Kaiser Permanente Vallejo Medical Center, Vallejo, California; 17Department of Pediatrics, Baylor College of Medicine and Texas Children’s Hospital, Houston, Texas; 18Children’s Hospital and Research Center, Oakland, California; 19Bay Area Pediatrics, Oakland, California; 20Department of Medicine, Northwestern University, Chicago, Illinois; 21Department of Pediatrics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California; 22Veterans Caribbean Health Care System, San Juan, Puerto Rico; 23Basic Science Laboratory, Center for Cancer Research, National Cancer Institute, Leidos Biomedical Research, Frederick National Laboratory, Frederick, Maryland; 24National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV, Irapuato, Guanajuato, Mexico; 25Centro de Neumología Pediatrica, San Juan, Puerto Rico; 26Division of Allergy and Immunology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois; 27Ann & Robert H. Lurie Children’s Hospital of Chicago, Chicago, Illinois; 28Center for Genes, Environment and Health, Department of Pediatrics, National Jewish Health, Denver, Colorado; and 31Laboratory of Molecular Neuro-Oncology and 32Laboratory of Molecular Neuro-Oncology, Howard Hughes Medical Institute, The Rockefeller University, New York, New York

ORCID IDs: 0000-0002-5372-4198 (A.C.Y.M.); 0000-0001-6372-8224 (Z.A.S.); 0000-0001-6497-9885 (E.F.).

Abstract

Rationale: Albuterol, a bronchodilator medication, is the first-line therapy for asthma worldwide. There are significant racial/ethnic differences in albuterol drug response.

Objectives: To identify genetic variants important for bronchodilator drug response (BDR) in racially diverse children.

Methods: We performed the first whole-genome sequencing pharmacogenetics study from 1,441 children with asthma from the tails of the BDR distribution to identify genetic association with BDR.

Measurements and Main Results: We identified population-specific and shared genetic variants associated with BDR, including genome-wide significant ($P < 3.53 \times 10^{-7}$) and suggestive ($P < 7.06 \times 10^{-6}$) loci near genes previously associated with lung capacity (DNAH5), immunity (NFKB1 and PLCB1), and β-adrenergic signaling (ADAMTS3 and COX18). Functional analyses of the BDR-associated SNP in NFKB1 revealed potential regulatory function in bronchial smooth muscle cells. The SNP is also an expression quantitative trait locus for a neighboring gene, SLC39A8. The lack of other asthma study populations with BDR and whole-genome sequencing data on minority children makes it impossible to perform replication of our rare variant associations. Minority underrepresentation also poses significant challenges to identify age-matched and population-matched cohorts of sufficient sample size for replication of our common variant findings.

Conclusions: The lack of minority data, despite a collaboration of eight universities and 13 individual laboratories, highlights the urgent need for a dedicated national effort to prioritize diversity in research. Our study expands the understanding of pharmacogenetic analyses in racially/ethnically diverse populations and advances the foundation for precision medicine in at-risk and understudied minority populations.

Keywords: albuterol; asthma; minority; NFKB1; Latinos
Asthma is a chronic inflammatory disorder characterized by recurrent respiratory symptoms and reversible airflow obstruction. Asthma is the most common chronic childhood disease (1). In the United States, asthma prevalence is highest among Puerto Ricans (36.5%), intermediate among African Americans (13.0%) and European Americans (12.1%), and lowest among Mexican Americans (7.5%) (2). Asthma mortality is fourfold to fivefold higher in Puerto Ricans and African Americans compared with white persons and Mexican Americans (3).

Inhaled β₂-agonists (e.g., albuterol) are the preferred treatment for acute asthma symptoms. Albuterol produces bronchodilation by causing rapid smooth muscle relaxation in the airways. Albuterol is the most commonly prescribed asthma medication worldwide (4). Among low-income and minority populations, albuterol is often the only medication used for asthma regardless of asthma severity (5, 6).

Spirometry is used to quantify bronchodilator drug response (BDR), which varies significantly among individuals and between populations (7). Specifically, the populations with the highest asthma prevalence and mortality also have the lowest BDR: Puerto Rican and African American children have significantly lower BDR than white and Mexican American children (7, 8). This racial/ethnic variation in BDR may contribute to the observed disparities in asthma morbidity and mortality (9, 10).

BDR is a complex trait, influenced by social, environmental, and genetic factors, with heritability estimates ranging from 47% to 92% (11–13). Genome-wide association studies (GWASs) have identified several common SNPs associated with BDR in populations of European descent (14–16). Only one GWAS of BDR has been conducted among African Americans (17). Although that study identified a novel BDR-associated locus, it did not replicate associations discovered in populations of European descent, suggesting that BDR may be partly determined by population-specific variants. Our genetic investigation of BDR among Latinos identified a
At a Glance Commentary

Scientific Knowledge on the Subject: Asthma is the most common chronic disease among children. Albuterol, a bronchodilator medication, is the first-line therapy for asthma treatment worldwide. In the United States, asthma prevalence is the highest among Puerto Ricans, intermediate among African Americans and white persons, and lowest in Mexicans. Asthma mortality is fourfold to fivefold higher in Puerto Ricans and African Americans compared with Mexicans. Puerto Ricans and African Americans, the populations with the highest asthma prevalence and mortality, also have the lowest albuterol bronchodilator drug response.

What This Study Adds to the Field: We conducted the largest pharmacogenetic study using whole-genome sequencing data from 1,441 minority children with asthma who had extremely high or low albuterol bronchodilator drug response. We identified population-specific and shared pharmacogenetic variants associated with bronchodilator drug response. We prioritized variants in NFκB1 with multiple levels of existing biologic evidence and demonstrated their potential regulatory functions using chromatin immunoprecipitation sequencing, RNA sequencing, and luciferase enhancer assays. Our study reveals the challenges and importance of increasing diversity in research. Our findings help inform the direction of future development of asthma medications and advance the foundation of precision medicine for at-risk, yet understudied, racially/ethnically diverse populations.

Methods

Study Cohorts and Sample Details
This study examined a subset of subjects with asthma from SAGE II (Study of African Americans, Asthma, Genes and Environments) (23) and GALA II (Genes-Environments and Admixture in Latino Americans) studies (18).

A total of 1,441 individuals from three ethnic groups (483 Puerto Ricans, 483 Mexicans, and 475 African Americans) representing the tails of the BDR distribution were selected for WGS (Figure 1). Sequencing quality control metrics are summarized in Table E1. Subject selection and filtering processes are described in the online supplement. Descriptive data of study participants are summarized in Table 1. Detailed descriptions of BDR assessment and analysis are described in the online supplement.

WGS Data Generation, Processing, and Quality Control
Details regarding DNA processing, WGS, variant calling, data quality controls, and variant annotation are described in the online supplement.

Data Availability
TOPMed WGS data are available to download by submitting a data access request through dbGaP. The dbGaP study accession numbers for GALA II and SAGE II are phs000920.v1.p1 and phs000921.v1. p1, respectively.

BDR Association Testing and Replication
Single and multivariate testing for BDR association were performed. Logistic regression analysis (high vs. low responder status) was controlled for age, sex, body mass index category, and the first 10 principal components. We conducted population-specific analyses and also performed a transethnic meta-analysis using METASOFT (24). Multivariate analysis was also performed on individuals combined across all three populations, including local genetic ancestries as additional covariates. The contribution of individual common and rare variants to multivariate association significance was evaluated by excluding common or rare variants in reduced and drop-one analyses. Variation in BDR explained by identified associations was calculated using McFadden pseudo $R^2$ (25). Detailed descriptions of these analyses are described in the online supplement.

We were unable to identify another age- and population-matched asthma cohort of sufficient size to replicate our drug response findings. Replication was nevertheless attempted in smaller cohorts and is described in detail in the online supplement.

Variant Prioritization and Functional Validation
To prioritize BDR-associated variants for further evaluation, an H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq) assay was performed to identify variants overlapping with regulatory regions in primary bronchial smooth muscle cells (BSMCs). The Diverse Convergent Evidence approach (26) was then used to score BDR-associated variants using multiple levels of observational, bioinformatic, and laboratory evidence. Prioritized variants were further validated by luciferase enhancer assays to verify enhancer activity and by RNA sequencing to identify potential expression quantitative trait locus function. Detailed descriptions of these procedures are described in the online supplement.

Results

Descriptive Characteristics of Study Subjects
Descriptive characteristics for all study subjects ($n = 1,441$) are summarized in Table 1. Covariates and demographic variables were assessed for significant
differences between high and low drug responders for each racial/ethnic group. Significant differences were found for age (Mexicans, \( P < 0.001 \)), baseline lung function (pre-FEV\(_1\) % predicted, \( P < 0.001 \)), total IgE (\( P < 0.001 \)), and atopy. Pre-FEV\(_1\) % predicted was defined as the percentage of observed FEV\(_1\), relative to the expected population average FEV\(_1\) based on the Hankinson lung function prediction equations (27). Results and descriptions of genetic ancestry and genetic substructure are shown in Table 1, supplemental text E1, and Figures E2 and E3.

**BDR Association Testing with Common Variants**

Detailed descriptions of variant summary statistics are described in supplemental text E2, Figure E4, and Tables E2 and E3. Study design is described in Figure E5. All subsequent analyses in this study were performed only with biallelic SNPs. Throughout this study, low drug responders were assigned as the reference (i.e., control) group, and high drug responders were classified as cases. We performed genome-wide association testing of common variants with BDR (dichotomized as high/low drug responders) for each population, adjusting by age, sex, body mass index category, and the first 10 principal components. We then performed a transethnic meta-analysis across all three populations.

The commonly used GWAS \( P \) value threshold of \( 5 \times 10^{-8} \) is derived from a Bonferroni correction under the assumption of 1,000,000 independent tests, based primarily on patterns of linkage disequilibrium (LD) from individuals of European descent. This threshold has been shown to be nongeneralizable for genetic studies among populations of non-European descent (28, 29). We therefore empirically calculated the effective number of independent tests for each population and for our transethnic meta-analysis. The resulting genome-wide significance thresholds were \( 1.57 \times 10^{-7} \) for Puerto Ricans, \( 2.42 \times 10^{-7} \) for Mexicans, \( 9.59 \times 10^{-8} \) for African Americans, and \( 3.53 \times 10^{-7} \) for the transethnic meta-analysis.

These numbers are highly concordant with WGS significance thresholds derived from the 1,000 Genomes sequencing data (28). Significance thresholds for discovery analyses in GWASs can often produce false-negative results. To minimize type II error, suggestive associations are often included in replication and functional validation studies. We identified suggestive associations based on the following widely used formula: \( 1/(\text{effective number of tests}) \) (30).

Although no significant associations were identified from the population-specific analyses (see Figure E6), our transethnic meta-analysis identified 10 unique loci (represented by 27 SNPs) significantly \((P < 3.53 \times 10^{-7})\) or suggestively \((P < 7.06 \times 10^{-6})\) associated with BDR status (Figure 2A and Table 2; see Table E4). After LD pruning, these 27 SNPs explain 23\%, 16\%, and 18\% of the variation in BDR status in Puerto Ricans, Mexicans, and African Americans, respectively (see Table E5). To demonstrate that the results of our regression models were robust, a post hoc analysis was performed on the 27 SNPs by including Native American and African local ancestries as covariates. The association results before and after adjusting for local ancestry remained consistent (see Text E3, Table E6). We annotated all 27 SNPs by performing a thorough bioinformatic search in ENCODE, the NHGRI-EBL GWAS Catalog, and PubMed. Their previously reported lung-related phenotype associations and functional annotations are reported in Tables E7 and E8.

Two SNPs located on chromosome 5 (rs17834628 and rs35661809) were significantly associated with BDR \((P = 1.18 \times 10^{-8} \) and \(3.33 \times 10^{-8}\)). The direction of effect for these two variants is concordant across all three populations (Figure 2B; see Table E9). Figure 2C displays a LocusZoom plot of rs17834628 with 400-kb flanking regions. Three of the 27 SNPs were located within genes. Specifically, two SNPs are located in the third and fifth introns of \( NLKFB1 \) (rs28450894 and rs4648006), and a third SNP, rs16995064, mapped to intron 7 of \( PLCB1 \) (Table 2). Based on 1,000 Genomes data, the low BDR-associated T allele of \( NLKFB1 \) (rs28450894) is found predominantly among African populations (minor allele frequency \([\text{MAF}], 8.8\text{–}28.7\%)\), followed by
Table 1. Study Population Description (N = 1,441)

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Puerto Ricans (n = 483)</th>
<th>Mexicans (n = 483)</th>
<th>African Americans (n = 475)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High BDR</td>
<td>Low BDR</td>
<td>P Value</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>239</td>
<td>244</td>
<td>—</td>
</tr>
<tr>
<td>Percent male</td>
<td>53.6</td>
<td>53.3</td>
<td>1.0</td>
</tr>
<tr>
<td>Median (IQR) age, yr</td>
<td>11.6 (9.7–14.8)</td>
<td>12.2 (10.1–15.2)</td>
<td>0.18</td>
</tr>
<tr>
<td>Mean global ancestry proportions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFR</td>
<td>0.24</td>
<td>0.22</td>
<td>0.44</td>
</tr>
<tr>
<td>EUR</td>
<td>0.63</td>
<td>0.64</td>
<td>0.27</td>
</tr>
<tr>
<td>NAM</td>
<td>0.13</td>
<td>0.13</td>
<td>0.93</td>
</tr>
<tr>
<td>BMI category, n</td>
<td>76, 163</td>
<td>67, 177</td>
<td>0.32</td>
</tr>
<tr>
<td>Pre-FEV, % predicted, n&lt;80%, &gt;=80%</td>
<td>149, 90</td>
<td>56, 188</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Median (IQR) ΔFEV1, %</td>
<td>21.2 (18.2–25.7)</td>
<td>5.0 (2.9–6.3)</td>
<td>—</td>
</tr>
<tr>
<td>Median (IQR) IgE, ml</td>
<td>407.5 (126.8–952.8)</td>
<td>191.9 (60.5–542.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Atopy, n</td>
<td>177</td>
<td>118</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Definition of abbreviations: AFR = African ancestry; BDR = bronchodilator drug response; BMI = body mass index; EUR = European ancestry; IQR = interquartile range; NAM = Native American ancestry; pre-FEV, % predicted = percentage of measured FEV1 relative to predicted FEV1 estimated by the Hankinson lung function prediction equations before administration of albuterol; tIgE = measure of total IgE from serum.

Atopy was defined as tIgE measurement ≥100 kU/L. ΔFEV1 is a quantitative measure of BDR, measured as the percent change in baseline FEV1 after administration of albuterol. High and low drug responders were chosen from the extremes of the BDR (ΔFEV1) distribution.
European populations (MAF, 3.7–7.6%) and Puerto Ricans (MAF, 6.2%), and rare in Mexicans (MAF, 1.5%) (see Figure E7).

We were unable to identify age- and population-matched asthma cohorts with albuterol drug response data of sufficient sample size for replication. Nevertheless, we attempted to replicate all 27 SNPs in five independent populations (GALA I, SAGE I, HPR, SAPPHIRE, and CHOP) separately and by meta-analysis (see Tables E10 and E11). Although none of the 27 SNPs were significantly associated with BDR status in our replication analyses (see Table E12), it is important to note that age-specific associations with asthma and asthma-related phenotypes have been reported (31, 32), and that children were not included in our largest replication cohort (SAPPHIRE, median age 28) (see Table E10). All other replication populations were smaller than our discovery populations (ranging from 108 to 414 individuals per study), and some had an imbalance of cases and control subjects (see Table E10), which further diminished power for replication analyses.

Albuterol binds to β2-adrenergic receptors in BSMC, causing rapid onset of airway tissue relaxation and bronchodilation. BSMCs are therefore considered one of the most relevant cell types for molecular studies of BDR (33). We performed H3K27ac ChIP-seq experiments in primary BSMCs to identify potential regulatory regions marked by H3K27ac peaks. We then overlapped the H3K27ac peaks with the 27 BDR-associated SNPs and SNPs that were in high LD with them (R² > 0.8 within 1 Mb in any of our study populations). Variants overlapping with H3K27ac peaks may imply regulatory functions in BSMC (see Table E8).

We applied the Diverse Convergent Evidence (26) approach to prioritize the 27 BDR-associated SNPs for inclusion in further functional analyses (see Table E12). After integrating information from our WGS analysis, publicly available bioinformatics data, and ChIP-seq experiments in BSMCs, the NFKB1 locus had the highest Diverse Convergent Evidence score, indicating that NFKB1 had the strongest evidence of functional relevance to BDR variation (see Figure E8). Therefore, all further functional experiments were focused on variants within this locus.
NFKB1 Functional Assays

Two H3K27ac ChIP-seq regions that overlapped with the BDR-associated NFKB1 locus were tested for enhancer activity using luciferase enhancer assays in BSMCs (see Figure E9A and Table E13). One enhancer, NFKB1 Region 2, showed significantly increased enhancer activity over empty vector (log2 2.24-fold increase, \( P = 1.59 \times 10^{-5}, \) unpaired Wilcoxon test) (see Figure E9B).

Given the relevance of NFKB1 in immune pathways and asthma, we also performed RNA sequencing experiments in African American children with asthma to verify whether the identified intronic NFKB1 SNPs regulate expression of neighboring genes. Among genes within 1 Mb of rs28450894 meeting expression cutoffs, we found that the low-expression neighboring gene, NFKB1 Region 2, showed significantly increased enhancer activity (log2 2.24-fold increase, \( P = 0.327 \)).

Comparison with Previous BDR Association

We observed that two known BDR candidate genes, ADCY9 and CRHR2, which replicated in a previous BDR GWAS performed in the full GALA II population (18), did not replicate in the current study (see Table E14). In that study, imputed GWAS array data were used to evaluate genetic associations with BDR measured as a continuous trait. To determine whether the discrepancy between findings was caused by data type (imputed array-based vs. WGS-based) or study design (continuous trait vs. extreme phenotype), the common variant analysis in the current analysis was repeated among the subset of samples with array-based and WGS data (\( n = 1,414 \) out of 1,441). Based on the top 1,000 BDR-associated SNPs from the common variant analysis, there was perfect correlation between association \( P \) values generated from imputed array-based and WGS-based genotypes (Spearman correlation, 1.0), suggesting that data type is not the cause of the observed discrepancy (see Figure E11A). Nearly all SNPs with high imputation \( R^2 \) exhibited high genotype concordance between array-based and WGS-based genotypes, confirming high imputation quality for most common SNPs (\( \geq 99.7\% \); see Figures E11B and E11C). We also performed linear regression to analyze BDR as a continuous trait (\( \DeltaFEV_1 \)) using imputed array-based data. The most significantly associated SNP identified in the current extreme phenotype analysis displayed the same direction of effect as analyzing BDR as a continuous trait (odds ratio, 1.67 in extreme phenotype analysis; \( \beta = 0.51 \) in continuous analysis). These observations indicate that the discrepancy between findings may be caused by differences in statistical power afforded by the different study designs (continuous trait vs. extreme phenotype).

For common variant analyses, dichotomization of a continuous outcome results in a loss of statistical power (34). The opposite effect is observed in rare variant analyses. The extreme phenotype study design has been shown to increase power and the probability of identifying functional rare variants (35, 36). Also noteworthy is that the previously published results were discovered in one population (Puerto Ricans), whereas the results from our transethnic meta-analysis describe associations that are conserved across three populations (Puerto Ricans, Mexicans, and African Americans).

BDR Association Testing Using Common and Rare Variants

We tested the combined effects of common and rare variants on BDR using SKAT-O (37) to examine variants in 1-kb sliding windows with 500-bp increments. The same covariates used for common variant association testing were applied.

After determining the effective number of samples and adjusting for multiple comparisons on each population separately, we identified three population-specific loci associated with BDR at genome-wide significance levels; two were found in Mexicans on chromosome 1 and chromosome 11, and one in African Americans on chromosome 19 (Figures 3A–3C and Table 3; see Table E15).

We also performed association testing across all three populations in a single analysis to maximize power. To minimize confounding by population substructure, association testing also included local ancestry, defined as the proportions of Native American and African ancestries for the window under testing. Two loci on chromosomes 4 and 8 attained genome-wide significance (\( P < 1.53 \times 10^{-7} \)) (Figure 3D and Table 3). Sixty variants were identified from all SKAT-O regions reported in Table 3. Six of the 60 variants were located within predicted regulatory regions (see Table E16). Three variants on chromosome 11 identified in Mexicans overlap with a CTCF (transcriptional repressor) binding site and comprise a chromatin insulator region. The five regions identified in combined and population-specific SKAT-O analyses independently explained 4–8% of BDR variation in their respective populations (see Table E5).

To investigate whether common and rare variants both contributed to the BDR association \( P \) value, we performed reduced and drop-one SKAT analyses by excluding common or rare variants in the associated region one by one. Because excluding either all common or all rare variants would reduce the significance of the BDR association \( P \) value (see Table E17), the reduced (see Table E18) and drop-one (see Table E19) analyses indicated that both common and rare variants contribute to the significance of the BDR association.

Although we believe the sliding window method is the most appropriate approach for WGS data, we considered alternative grouping strategies for rare variants, including grouping by 1) genes from transcription start sites to transcription end sites, with or without 50-kb flanking regions; 2) transcription start sites with 20-kb flanking regions; and 3) H3K27ac ChIP-seq peaks from airway epithelial cells and airway smooth muscle cells. Association tests...
Table 2. Results from Transethnic BDR Association Tests for Common Variants

<table>
<thead>
<tr>
<th>Chr</th>
<th>Start</th>
<th>rsID</th>
<th>Effect Allele</th>
<th>OR (95% CI)</th>
<th>P Value</th>
<th>Effect Allele Frequency</th>
<th>Nearest Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>12978566</td>
<td>rs17834628</td>
<td>A</td>
<td>1.67 (1.29–2.16)</td>
<td>1.18 × 10⁻⁸⁺</td>
<td>0.32</td>
<td>0.42</td>
</tr>
<tr>
<td>5</td>
<td>12968341</td>
<td>rs35661809</td>
<td>G</td>
<td>1.59 (1.20–2.10)</td>
<td>3.33 × 10⁻⁸⁺</td>
<td>0.34</td>
<td>0.43</td>
</tr>
<tr>
<td>5</td>
<td>12975934</td>
<td>rs17237639</td>
<td>G</td>
<td>1.61 (1.30–2.00)</td>
<td>1.22 × 10⁻⁷</td>
<td>0.31</td>
<td>0.43</td>
</tr>
<tr>
<td>5</td>
<td>12975187</td>
<td>rs1017452</td>
<td>G</td>
<td>1.60 (1.31–1.96)</td>
<td>2.11 × 10⁻⁷</td>
<td>0.31</td>
<td>0.43</td>
</tr>
<tr>
<td>5</td>
<td>12975322</td>
<td>rs1017454</td>
<td>A</td>
<td>1.60 (1.31–1.96)</td>
<td>2.11 × 10⁻⁷</td>
<td>0.31</td>
<td>0.43</td>
</tr>
<tr>
<td>5</td>
<td>12975265</td>
<td>rs1017453</td>
<td>C</td>
<td>1.56 (1.25–1.95)</td>
<td>6.40 × 10⁻⁷</td>
<td>0.31</td>
<td>0.42</td>
</tr>
<tr>
<td>5</td>
<td>12972636</td>
<td>rs17237443</td>
<td>C</td>
<td>1.59 (1.28–1.97)</td>
<td>9.85 × 10⁻⁷</td>
<td>0.29</td>
<td>0.42</td>
</tr>
<tr>
<td>1</td>
<td>209324294</td>
<td>rs10746419</td>
<td>T</td>
<td>1.29 (0.75–2.25)</td>
<td>1.19 × 10⁻⁶</td>
<td>0.49</td>
<td>0.54</td>
</tr>
<tr>
<td>6</td>
<td>12961545</td>
<td>rs17833938</td>
<td>C</td>
<td>1.56 (1.28–1.91)</td>
<td>1.45 × 10⁻⁶</td>
<td>0.30</td>
<td>0.42</td>
</tr>
<tr>
<td>6</td>
<td>104240500</td>
<td>rs13437006</td>
<td>C</td>
<td>1.56 (1.21–2.02)</td>
<td>1.61 × 10⁻⁶</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td>15</td>
<td>101230457</td>
<td>rs1565749</td>
<td>A</td>
<td>1.66 (1.18–2.32)</td>
<td>1.64 × 10⁻⁶</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td>5</td>
<td>12948369</td>
<td>rs34845041</td>
<td>T</td>
<td>1.56 (1.26–1.92)</td>
<td>1.77 × 10⁻⁶</td>
<td>0.30</td>
<td>0.42</td>
</tr>
<tr>
<td>5</td>
<td>12975108</td>
<td>rs1017451</td>
<td>T</td>
<td>1.55 (1.24–1.93)</td>
<td>1.96 × 10⁻⁶</td>
<td>0.30</td>
<td>0.42</td>
</tr>
<tr>
<td>5</td>
<td>12950432</td>
<td>rs62347395</td>
<td>G</td>
<td>1.55 (1.26–1.92)</td>
<td>2.02 × 10⁻⁶</td>
<td>0.30</td>
<td>0.42</td>
</tr>
<tr>
<td>15</td>
<td>101231049</td>
<td>rs57924834</td>
<td>A</td>
<td>1.59 (1.25–2.03)</td>
<td>2.04 × 10⁻⁶</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>5</td>
<td>137382142</td>
<td>rs17048864</td>
<td>A</td>
<td>1.8 (1.18–3.05)</td>
<td>1.22 × 10⁻⁶</td>
<td>0.11</td>
<td>0.14</td>
</tr>
<tr>
<td>5</td>
<td>12959958</td>
<td>rs14362930</td>
<td>G</td>
<td>1.55 (1.24–1.93)</td>
<td>2.73 × 10⁻⁶</td>
<td>0.29</td>
<td>0.42</td>
</tr>
<tr>
<td>20</td>
<td>8635168</td>
<td>rs16995064</td>
<td>G</td>
<td>1.96 (1.12–3.38)</td>
<td>3.30 × 10⁻⁶</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td>12</td>
<td>19821401</td>
<td>rs65644720</td>
<td>T</td>
<td>0.66 (0.55–0.78)</td>
<td>3.66 × 10⁻⁶</td>
<td>0.33</td>
<td>0.37</td>
</tr>
<tr>
<td>6</td>
<td>104235591</td>
<td>rs6926020</td>
<td>C</td>
<td>1.57 (1.25–1.97)</td>
<td>1.68 × 10⁻⁶</td>
<td>0.19</td>
<td>0.23</td>
</tr>
<tr>
<td>4</td>
<td>103453535</td>
<td>rs28450894</td>
<td>T</td>
<td>0.47 (0.34–0.64)</td>
<td>3.75 × 10⁻⁶</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td>10</td>
<td>103461559</td>
<td>rs4648006</td>
<td>T</td>
<td>0.47 (0.34–0.64)</td>
<td>3.75 × 10⁻⁶</td>
<td>0.06</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Definition of abbreviations: AA = African Americans; BDR = bronchodilator drug response; Chr = chromosome; CI = confidence interval; MX = Mexicans; OR = odds ratio; PR = Puerto Ricans.

The top 10 unique loci (represented by 27 SNPs) significantly (P < 5.5 × 10⁻⁵) or suggestively (P < 7.06 × 10⁻⁶) associated with BDR status in our transethnic meta-analysis. The four nearest transcripts from RefSeq were identified and genes with multiple transcripts were reported only once, with the distance to the nearest transcript indicated in parentheses. Negative distances indicate upstream genes. Genes overlapping with BDR-associated SNPs are bold. High drug responders were assigned as cases.

*P values that achieved adjusted genome-wide significance for transethnic meta-analysis (P < 3.53 × 10⁻⁷).
with these grouping strategies identified no further significant associations. Because these analyses were performed post hoc, multiple testing correction for testing alternative grouping strategies was not applied to the results of the sliding window approach.

Discussion

We identified population-specific and shared common and rare variants associated with BDR in three racially and ethnically diverse populations of children with asthma. WGS provides comprehensive detection of common and rare variants in coding and noncoding regions. Combined, the 27 variants (after LD pruning) identified from our common variant analyses (Table 2) explained 23%, 16%,

Figure 3. Manhattan plot of SKAT-O analysis of biallelic common and rare SNPs grouped by 1-kb windows sliding across chromosome 1–22 in (A) Puerto Ricans, (B) Mexicans, (C) African Americans, and (D) all populations combined. Bonferroni-corrected genome-wide and suggestive significance levels are marked by red and blue lines, respectively.
Table 3. Association Testing for Combined Effects of Common and Rare Variants on BDR

<table>
<thead>
<tr>
<th>Chr</th>
<th>Start</th>
<th>Stop</th>
<th>P Value</th>
<th>Population</th>
<th>nCommon</th>
<th>nRare</th>
<th>Nearest Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>114177000</td>
<td>114178000</td>
<td>4.40 × 10^{-9}</td>
<td>MX</td>
<td>2</td>
<td>1</td>
<td>MAGI3 (intron 9), PHTF1 (62 kb), RSNB1 (126 kb) LOC105376671 (~3 kb), LGR4 (~13 kb), LIN7C (8 kb)</td>
</tr>
<tr>
<td>11</td>
<td>27507000</td>
<td>27508000</td>
<td>6.59 × 10^{-9}</td>
<td>MX</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>10424000</td>
<td>10425000</td>
<td>3.12 × 10^{-11}</td>
<td>AA</td>
<td>1</td>
<td>2</td>
<td>ZGLP1 (~4 kb), ICAM5 (~17 kb), FDX1L (intron 3), RAVE1 (2 kb) ADAMTS3 (~43 kb), COX18 (441 kb)</td>
</tr>
<tr>
<td>4</td>
<td>73478000</td>
<td>73479000</td>
<td>6.25 × 10^{-8}</td>
<td>Combined</td>
<td>10</td>
<td>23</td>
<td>SDC2 (~302 kb), CPQ (intron 4), LOC101927066 (37 kb), TSPYL5 (359 kb)</td>
</tr>
<tr>
<td>8</td>
<td>97926000</td>
<td>97927000</td>
<td>1.32 × 10^{-7}</td>
<td>Combined</td>
<td>3</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

Definition of abbreviations: AA = African Americans; BDR = bronchodilator drug response; Chr = chromosome; MX = Mexicans; nCommon = number of common variants; nRare = number of rare variants.

ChIP-seq and functional enhancer assays in BSMCs imply that the regions containing the NFKB1 intronic variants regulate expression of nearby genes, but do not directly suggest that the genetic variants themselves alter the expression of nearby genes. However, the latter was supported by our RNA sequencing data, which showed that individuals with the low BDR-associated T allele displayed reduced expression of the neighboring SLC39A8 gene. Additionally, SLC39A8 has been found to be responsive to cytokine treatment in airway epithelial cells (42) and had reduced expression in mice with allergic airway inflammation (43).

Studies have shown that up-regulation of SLC39A8 is sufficient to protect the lung epithelium against tumor necrosis factor-α–induced cytotoxicity (44). Additionally, the higher frequency of the low BDR-associated allele (T allele of rs28450894 in NFKB1) in African populations suggests that the low BDR-associated allele tracks with African ancestry. This may explain why admixed populations with higher proportions of African ancestry (e.g., African Americans and Puerto Ricans) have lower BDR (8), and by extension may shed light on the higher asthma morbidity and mortality in these populations.

Another intronic variant (chromosome 20, rs16995064, PLCB1 intron 7) suggestively associated with BDR is relevant to childhood asthma. PLCB1 is differentially expressed among children with therapy-resistant asthma versus controlled persistent asthma or age-matched healthy control subjects (45). Silencing PLCB1 inhibited the effect of lipopolysaccharide-induced endothelial cell inflammation by inhibiting expression of proinflammatory cytokines (46). Additional functional studies are necessary to establish the role of NFKB1 and PLCB1 on BDR.

We identified various combined effects of rare variants that were population-specific or shared across populations. Although some nearest genes show no known functional relationship to BDR (MAGI3, LOC105376671, LIN7C, and CPQ), the locus between ADAMTS3 and COX18 may be functionally relevant. The ADAMTS3/COX18 locus was associated with β-adrenergic responses in murine cardiovascular-related traits (47). This locus was significantly associated with cardiac atrial weight in mice treated with the β-blocker atenolol and replicated under β-agonist isoproterenol treatment. These findings suggest that SNPs in this locus may modify β-adrenergic signaling pathways in BDR. In the present study, we also identified BDR association with rare variants within CPQ, which encodes a protein from the carboxypeptidase family. Although no previous BDR association has been identified for CPQ, another member of the carboxypeptidase family, carboxypeptidase A3, is expressed at higher levels in the airway epithelium among subjects with Th2-high asthma (48, 49). Further studies are necessary to determine the role of CPQ in BDR.

GWAS-based BDR-associated common variants in GALA II have previously been reported (18). However, those variants were not significantly associated with extreme BDR in the current study, likely because of differences in study design. The previous BDR GWAS used an array-based genotyping panel to examine children with asthma from the
entire BDR spectrum. In contrast, our current study sequenced the entire genome to investigate only the extremes of the BDR distribution. By repeating our current extreme phenotype analysis using a subset of individuals who had array and WGS data, we confirmed that the major discrepancy between the two studies was caused by study design instead of differences in data type. The contrast in results between GWAS and WGS caused by differences in study design implies that varied study designs are necessary for a comprehensive understanding of variants associated with asthma-related phenotypes and drug response. The extreme phenotype approach is recognized as one of the success factors in the study design of pharmacogenomic GWASs (50).

Furthermore, the power gain from studying extreme phenotypes is much greater in analyses of rare variants compared with common variant studies (51). An extreme phenotype study design provides a cost-effective means for studying common and rare variant associations that may otherwise be missed when sampling across the entire phenotypic spectrum.

We did not identify BDR-associated variants from β2-adrenergic receptors signaling pathways. Instead, most of the BDR-associated genes identified in this study are related to lung function and allergic response, including total IgE levels and cytokine production in mast cells. This suggests that BDR depends in part on the intrinsic state of airway smooth muscle cells. Genetic variation may determine individuals’ intrinsic expression levels of candidate genes, which in turn determine whether their response to albuterol is beneficial.

Including admixed populations in WGS studies has important scientific implications. First, it facilitates discovery of genetic variation of multiple ancestral populations in a single study. Second, studying multiethnic admixed populations with ancestries that are underrepresented in existing genetic repositories can yield novel pharmacogenomic insights. For example, the widely popular PCSK9 inhibitors used to treat hypercholesterolemia were discovered by studying the genetics of African Americans but the drug development eventually benefited patients of all ethnicities (52). Finally, studying admixed populations, such as Mexicans, enhances the understanding of genetic variation in Native American ancestry, a population largely absent in major sequencing efforts.

We and others have documented the implications and challenges posed by the lack of non-European study populations in biomedical research (53–55). We made extensive efforts to test our top BDR-associated variants in other populations, but the unique characteristics of our discovery cohort (minority children with asthma who have BDR and WGS data) posed significant challenges for finding comparable replication cohorts. Testing ancestry-dependent pharmacogenetic variants without age- and ethnicity-matched replication populations poses an impossible statistical bar for publication. We therefore conducted additional analyses involving bioinformatics and experimental assays (56). These challenges highlight the need to include more racially/ethnically diverse populations in all clinical and biomedical research.

In an era of precision medicine, addressing questions about the impact of genetic factors on therapeutic drug response in globally diverse populations is essential for making precision medicine socially and scientifically precise (57). This study advances the understanding of genetic analysis in admixed populations and helps to lay the foundation of precision medicine for understudied and racially and ethnically diverse populations.

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the New York Genome Center investigators and teams for whole-genome sequencing sample preparation, quality control, data generation, data processing, and initial joint genotyping. They gratefully acknowledge the studies and participants who provided biologic samples and data for TOPMed. They also gratefully acknowledge the contributions of the investigators of the NHLBI TOPMed Consortium (https://www.nhlbiwgs.org/topmed-banner-authorship).

References


