Published in final edited form as:
Clin Exp Allergy. 2012 February ; 42(2): 265–274. doi:10.1111/j.1365-2222.2011.03873.x.

African ancestry, early life exposures, and respiratory morbidity in early childhood

R. Kumar[1], **H.-J. Tsai**[2,3], **X. Hong**[2], **C. Gignoux**[4], **C. Pearson**[5], **K. Ortiz**[5], **M. Fu**[5], **J. A. Pongracic**[1], **E. G. Burchard**[4], **H. Bauchner**[5,*], and **X. Wang**[2,*]

[1]Division of Allergy and Immunology, Children's Memorial Hospital, Chicago, IL, USA

[2]The Mary Ann and J. Milburn Smith Child Health Research Program, Children's Memorial Hospital and Children's Memorial Research Center, Chicago, IL, USA

[3]Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan

[4]Departments of Bioengineering & Therapeutic Sciences and Medicine, UCSF, San Francisco, CA, USA

[5]Department of Pediatrics, Boston University School of Medicine, and Boston Medical Center, Boston, MA, USA

Summary

Background—Racial disparities persist in early childhood wheezing and cannot be completely explained by known risk factors.

Objective—To evaluate the associations of genetic ancestry and self-identified race with early childhood recurrent wheezing, accounting for socio-economic status (SES) and early life exposures.

Methods—We studied 1034 children in an urban, multi-racial, prospective birth cohort. Multivariate logistic regression was used to evaluate the association of genetic ancestry as opposed to self-identified race with recurrent wheezing (>3 episodes). Sequential models accounted for demographic, socio-economic factors and early life risk factors. Genetic ancestry, estimated using 150 ancestry informative markers, was expressed in deciles.

Results—Approximately 6.1% of subjects (mean age 3.1 years) experienced recurrent wheezing. After accounting for SES and demographic factors, African ancestry (OR: 1.16, 95% CI: 1.02–1.31) was significantly associated with recurrent wheezing. By self-reported race, hispanic subjects had a borderline decrease in risk of wheeze compared with African Americans (OR: 0.44, 95% CI: 0.19–1.00), whereas white subjects (OR: 0.46, 95% CI: 0.14–1.57) did not have. After further adjustment for known confounders and early life exposures, both African (OR: 1.19, 95% CI: 1.05–1.34) and European ancestry (OR: 0.84, 95% CI: 0.74–0.94) retained a significant association with recurrent wheezing, as compared with self-identified race (OR_{whites}: 0.31, 95% CI: 0.09–1.14; $OR_{hispanic}$: 0.47, 95% CI: 0.20–1.08). There were no significant interactions between ancestry and early life factors on recurrent wheezing.

Correspondence: R. Kumar, Division of Allergy and, Clinical Immunology, Children's Memorial Hospital, 2300 Children's Plaza, Box No. 60, Chicago, IL 60614, USA. rkumar@childrensmemorial.org.
*These authors contributed equally to this manuscript.

Competing Interests: None of the authors has any conflict of interest related to the work.

Conclusions and Clinical Relevance—In contrast to self-identified race, African ancestry remained a significant, independent predictor of early childhood wheezing after accounting for early life and other known risk factors associated with lung function changes and asthma. Genetic ancestry may be a powerful way to evaluate wheezing disparities and a proxy for differentially distributed genetic and early life risk factors associated with childhood recurrent wheezing.

Keywords

early childhood; genetic ancestry; race; wheezing

Introduction

There are disparities in childhood asthma with higher levels of prevalence, morbidity and mortality in African American populations [1–4]. However, recurrent wheezing in early childhood is a heterogeneous phenotype with a significant public health burden with racial disparities present throughout childhood [5]. While many children develop asthma wheeze in the first few years of life, there are many of these children who do not go on to have asthma despite having recurrent wheezing [6]. This early onset of symptoms raises the question whether differentially distributed pre-natal or early life environmental factors associated with lung growth and development contribute to observed racial disparities either alone or in combination with differentially distributed genetic factors. Genetic ancestry may serve as a proxy for both differentially distributed genetic factors and early life risk factors, which may differ by race. Indeed, genetic ancestry has been associated with a number of asthma phenotypes including asthma risk [7, 8], and serum IgE [7]. However, the relative contribution of early life risk factors in conjunction with ancestry-related factors on early life wheezing is unstudied.

There are a number of perinatal factors, such as chorioamnionitis and preterm birth [9–11], which are more prevalent in African American populations or which, like tobacco smoke exposure [12] may have differential effect in African Americans. Like *in utero* smoke exposure [13, 14], some of the effects of prematurity and chorioamnionitis on early life respiratory morbidity have been hypothesized to be due to changes in lung function [15, 16]. Just as early life factors may influence respiratory morbidity via altered lung function [17], genetically determined estimates of African ancestry are also associated with plateau or maximally achieved lung function as shown in a large national cohort [18]. As such, genetic ancestry may affect respiratory morbidity alone or in combination with other early life factors, prominent in inner-city populations, which affect lung function.

We sought to determine whether African ancestry would be more closely associated with recurrent wheezing than self-identified race, and whether these associations persist after accounting for socio-economic status (SES), as well as pre-natal and early life risk factors, some of which are common in the African American population. We secondarily sought to determine if the early life factors of preterm birth, chorioamnionitis, or maternal patterns of smoking interact with African ancestry to increase the risk of recurrent wheezing in the first years of life in a multi-ethnic, inner-city birth cohort.

Methods

Patient population

The Boston Birth Cohort (BBC) is an ongoing general population, multi-ethnic cohort, which includes subjects from a range of socio-economic strata that include urban poor and a smaller number of middle class subjects. The BBC was originally initiated in 1998 at the Boston Medical Center to study adverse birth outcomes, particularly preterm birth [19, 20].

Any woman admitted to the Labor and Delivery floor at the Boston Medical Center who delivered a singleton live infant was eligible. Mother–infant pairs were recruited 24–72 h post-delivery. After obtaining signed written informed consent, we interviewed subjects using a structured questionnaire. We also reviewed maternal and infant medical records using a standardized abstraction form to obtain clinical data and birth outcomes. Individuals recruited by BBC and who identified themselves as planning to receive primary care at the Boston Medical Center were invited to participate in the post-natal follow- up study, the Children's Health Study (CHS), starting in 2004 [20]. After written informed consent was obtained from the biological mother, visits were scheduled at 6–12 months, 2, 4 and 6 years consistent with the child's paediatric primary care visit schedule. Mothers were interviewed using structured questionnaires at each visit.

The CHS enrolment includes 1794 subjects (enrolment of 91% of all those approached in the BBC). Of those enrolled in the CHS, genetic data and key covariate data were available in 1196 subjects (66.7%, 1196/1794). Aside from being slightly older (mean ± SD of 3.9 ± 2.8 years), these individuals did not differ from those not included in terms of key socio-demographic variables including gender, self-identified race and SES measures including maximal level of maternal education achieved and annual household income (data not shown). For the purposes of this analysis, the subjects were limited to the three main ethnic groups (white, African American, and Hispanic) as exclusion of 160 subjects in the 'other' ethnicity group did not alter our findings and the 'other' group included heterogeneous ethnicities with different ancestral proportions such as Asians and Caribbean Islanders. These groups included too few subjects to analyse separately. One extreme outlier in each of the African American and white groups was removed as potentially misclassified. (This included one individual with an African ancestry of 3.4% in the African American group and one individual with 93% African Ancestry in the white group). Before exclusion, a sensitivity analysis was carried out to ensure that this did not change estimates. This exclusion left 1034 subjects, the sample for this analysis.

The parent study protocol was approved by the Children's Memorial Hospital (CMH) Institutional Review Board (IRB) and the Boston Medical Center IRB. The Boston Medical Center IRB and CMH IRB also approved the study protocol and consent process whereby all infants enrolled in the BBC are eligible for the postnatal follow-up study to determine post-natal health outcomes as detailed below.

Determination of gestational age and preterm birth status

Gestational age was assessed based on both the first day of the last menstrual period (LMP) as recorded in the maternal medical record and early (<20 weeks) prenatal ultrasound. This approach has been used in large hospital-based preterm studies and in our ongoing funded preterm studies [19, 21, 22]. Subjects were categorized as having preterm delivery if they were <37 weeks gestation at delivery.

Determination of presence of chorioamnionitis

For this study, chorioamnionitis was considered present if the pregnancy was associated with either the presence of intrapartum fever (>38°C) [23] based on review of medical record or by the presence of placental histological changes associated with chorioamnionitis. A detailed description of placental pathology criteria has been published by our group [24].

Determination of self-identified race

The race/ethnicity variable was based on maternal self-report of race in keeping with the policy of the National Center for Health Statistics [25]. For this analysis, race was categorized as non-Hispanic black, non-Hispanic white, Hispanic, or 'other'. Because of

ethno-cultural differences, if Hispanic/Latino ethnicity was reported, individuals were classified as Hispanic/Latino.

Determination of genetic ancestry

Estimates of global genetic ancestry were determined using a panel of 150 ancestry informative markers (AIMs), which were highly informative among three ancestral HapMap populations (African, European and Asian including Japanese and Han Chinese) with averaged δ (allele frequency difference between two ancestral populations) >0.5. This panel is described in more detail in a previous publication from our group [26]. This number of AIMs was chosen based on previous studies by other groups who found that a set of markers as small as 64 AIMs was sufficient to both control for population substructure and correctly identify West African ancestry with adequate precision [27]. With the STRUCTURE program (v. 2.3.1, http://pritch.bsd.uchicago.edu/structure.html), we calculated individual ancestral proportion based on 144 AIMs having a call rate >98.0%. The variables for individual African and European ancestry were transformed to percentage (fraction of ancestry multiplied by 100), and then expressed in deciles to make the results more interpretable in regression models. For all evaluations of interactions, ancestry variables and interaction terms used mean centred ancestry variables to decrease any potential collinearity between the base variable and the interaction term.

Assessment of sensitization to aeroallergens

Specific IgE (sIgE) concentration in plasma for each of five aeroallergens (cockroach, mouse, alternaria, cat, dog, and the two common dust mites Dermatophygoides pteronyssinus, and Dermatophygoides farinae) was measured using ImmunoCAP® at Quest Diagnostics (Madison, NJ, USA) as per manufacturer protocol (range 0.35 to ≥100 kU_A/L). Sensitization was defined as sIgE ≥ 0.35 kU_A/L. An individual was considered sensitized to aeroallergen if he/she was positive to any one of these allergens.

Other variables

Maternal atopy was considered present if the mother had reported a diagnosis of one of the following conditions: atopic dermatitis, allergic rhinitis, or asthma. Maternal smoking during pregnancy was assessed based on responses to the maternal questionnaire at delivery. Mothers who answered yes to smoking during 6 months before pregnancy or smoking during the first trimester and those who answered yes to smoking during the second or third trimester in the maternal questionnaire were classified as having ongoing *in utero* smoke exposure through the pregnancy. In the follow-up CHS study, those children whose mothers answered yes to smoking at the time of follow-up at any visit were classified as having post-natal exposure from the mother. Socio-economic status was approximated by highest reported level of maternal education (categorized as less than high school, high school (GED), and some college) and annual household income (categorized as above or below $30 000 or unknown.) This level was chosen based on the federal poverty threshold. For families who refused to report income, these individuals were classified as unknown. However, these families were headed by more mothers who had less than high school education compared with families with less than $30 000/year (40% vs. 31%, $P < 0.001$), suggesting that these families were in many cases lower SES families.

Outcome measures

Given the mean age of our subjects and the natural history of wheezing illness, a physician diagnosis of asthma is uncertain in this age range [28]. Indeed, as lung growth and remodelling may be affected by both chorioamnionitis and genetic background, wheezing irrespective of later asthma development was felt to be a more important clinical outcome to

evaluate the effect of these pre-natal factors and genetic ancestry. Recurrent wheezing (more than three episodes of medically attended wheezing illness in the subject's lifetime) was the primary outcome. If an individual had a second episode or had multiple types of encounters (office, emergency room, or hospitalization) within the space of 2 weeks, these were counted as a single episode. More than three episodes of medically attended wheezing illness by age 3 years is a morbidity criterion in asthma predictive indices [29]. Individuals classified as recurrent wheezers in our cohort would be similar in level of morbidity to other studies that have used asthma predictive indices in this age range [28, 29]. Episodes were determined prospectively using a unified electronic medical record (with data extracted on an ongoing basis) to evaluate all instances of contact with the Boston Medical Center system for each subject, including well and ill/ unscheduled visits (office visits, urgent care, ED and inpatient encounters). This is the most accurate and complete method to assess wheezing episodes requiring medical attention of CHS subjects who receive urgent and emergent medical care at Boston Medical Center.

Statistical analyses

A total of 1034 children from the BBC self-identified as white, Hispanic, or African American, completed at least one post-natal follow-up visit and had data on key covariates as well as genetic ancestry data available. For comparison of baseline characteristics of the population by self-identified race, ANOVA and chi-square test were used to characterize all continuous variables and all categorical variables respectively. In parallel multivariate logistic regression models, we examined the association of recurrent wheezing with African ancestry and self-identified race. The basic models included the key determinant variables, demographic measures (child age and child gender) and measures of SES (maternal age, highest maternal level of education achieved and annual household income). The second models evaluated other potential confounding variables based on the literature (caesarean section delivery, first born child, breastfeeding and maternal atopy). In the third set of models, potential early life risk factors were added in addition to the variables in model 2, including exposure to tobacco smoke *in utero*, and ongoing passive smoke exposure in the home, prematurity and chorioamnionitis. Finally, in a subgroup analysis, we evaluated a subset of 636 children; we also had aeroallergen data to determine whether the addition of aeroallergen sensitization in the final ancestry models modified the associations observed between ancestry and wheezing.

All analyses were performed using statistical software STATA for Windows 9.20 (STATA, College Station, TX, USA).

Results

Demographic characteristics of the cohort

The BBC is a predominantly low income, multi-ethnic population. In this sample, most (67.3%, $n = 696$) mothers were African American by self-report and 25.3% ($n = 262$) were Latino (Table 1). Approximately, one-third (29.7%) had not completed high school, but 34.3% had some post-secondary education. The mean age of children at follow up was 3.08 (SD 2.29) years, with roughly equal numbers of males and females. More than three episodes of wheezing were present in 68 (6.1%) of the children. The highest frequency of recurrent wheezing was present in African American children at 7.6% ($P = 0.01$). There were a number of differences by race. The African American children were slightly older. There were more white mothers with some post-secondary education, more white households had annual incomes greater than $30 000 and more white infants who were firstborn children. There were differences in all measures of smoke exposure (antenatal and postnatal maternal smoking, and other smokers in the home) by race with the highest rates

of exposure being in whites. Finally, while there are significant differences between the self-identified racial groups in terms of genetic estimates of individual ancestry, there was also a wide range of African ancestry across each of the groups except for self-identified whites. The mean and inter-quartile ranges for percent African ancestry are 86.1 (79.3–93.4) for African Americans, 1.1 (0.6–7.0) for whites, and 21.2 (6.5–38.7) for Latino Americans.

Crude associations

In Table 2, we show the bivariate associations between infant and maternal variables and recurrent wheezing, and crude associations of African ancestry with recurrent wheezing and key variables.

Compared with African American race, self-identified Hispanic race/ethnicity was associated with a lower risk of recurrent wheezing (OR: 0.33, 95% CI: 0.15–0.74), while white race was not significantly different despite a similar magnitude of effect (OR: 0.50, 95% CI: 0.15– 1.64). African ancestry (expressed in units of 10% African ancestry) was associated with an increase in risk of recurrent wheezing (OR: 1.16, 95% CI: 1.05–1.28). African ancestry was not associated with prematurity, chorioamnionitis or gestational age.

With respect to the early life factors, preterm birth, chorioamnionitis and maternal smoking during pregnancy were all individually associated with recurrent wheezing. Preterm birth and chorioamnionitis were also associated with recurrent wheezing independently of one another in a crude model. There was no evidence of a multiplicative interaction between prematurity and chorioamnionitis on the outcome of recurrent wheezing.

In terms of other potential confounders, maternal atopy was associated with recurrent wheezing on univariate analyses (OR: 2.02, 95% CI: 1.21–3.37) as was a less than high school maternal education level (OR: 0.41, 95% CI: 0.20–0.86). Other factors such as c-section delivery, first born child status, exclusive breastfeeding and annual household income were not significantly associated with recurrent wheezing.

Role of ancestry vs. self-identified race

Parallel multivariate models accounting for SES and demographic measures were used alternately to evaluate the association of self-identified race or African ancestry with recurrent wheezing in this multi-ethnic cohort. These are presented in the basic model depicted in Table 3. In the basic model (Table 3a), self-reported white race (OR: 0.46, 95% CI: 0.14–1.57) was not associated with a lower risk of recurrent wheezing compared with African Americans, but there was a borderline lower risk for Latino Americans (OR: 0.44, 95% CI: 0.19–1.00). In contrast, African ancestry (expressed in deciles) was associated (OR: 1.16, 95% CI: 1.02–1.31) with an increased risk of recurrent wheezing. Similarly, European ancestry (expressed in increments of 10%) was associated (OR: 0.85, 95% CI: 0.76–0.96) with a decreased risk of recurrent wheezing. Table 3b includes known risk factors from the literature (caesarean section, firstborn status, exclusive breastfeeding and maternal atopy). After inclusion of these risk factors, the associations for self-identified white and Hispanic race remained non-significant, and the associations for African ancestry (OR: 1.17, 95% CI: 1.03–1.32) and European ancestry (OR: 0.85, 95% CI: 0.75–0.96) did not change significantly. Furthermore, we also evaluated models with family history of asthma (both maternal and paternal) with and without the inclusion of maternal atopy and the results were identical. As such, only maternal atopy and not family history of asthma was included in final models in Table 3c.

Self-identified race, African ancestry and differentially distributed early life exposures on recurrent wheezing

In Table 3c, we ran final models with known risk factors from the literature (caesarean section, firstborn status, exclusive breastfeeding and maternal atopy) as well as early life exposures which are either documented in the literature to be differentially distributed by race (chorioamnionitis and preterm birth) or with a potential differential effect (cigarette smoke exposure) by race. Self-identified white and Hispanic race/ethnicity was again not significantly different from African American race. On the other hand, both African (OR: 1.19, 95% CI: 1.05–1.34) and European (OR: 0.84, 95% CI: 0.74– 0.94) ancestry remained significant even after inclusion of these factors with minimal change in the magnitude of the association, suggesting that these variables are not intermediate factors or confounders in this association.

In separate models assessing interaction terms, the interaction terms for preterm birth, chorioamnionitis, antenatal smoke exposure and post-natal maternal smoking with African ancestry respectively were not significant.

Sensitivity analysis

We performed a revised analysis within only blacks and whites to determine whether our estimates of ancestry effect were affected by the inclusion of Hispanics in the model. For this new analysis, we re-estimated ancestry for self-identified whites and blacks only with an assumed $k = 2$ ancestral groups in STRUCTURE using the same AIMs. The correlation of the re-estimated ancestry variables with the original variables was 0.99 for both European and African ancestry. Accordingly, despite some decrease in significance due to reduced sample size ($n = 771$), the estimates for the effects of African ancestry (OR: 1.20, 95% CI: 1.004–1.43) and European ancestry (OR: 0.83, 95% CI: 0.70–0.996) remained significant and were similar in magnitude and direction to the primary analysis.

Subgroup analyses

To evaluate whether the association of ancestry and recurrent wheezing was confounded by an association between ancestry and sensitization, we repeated the final ancestry models from Table 3c, restricting the analysis to those individuals ($n = 566$) who had aeroallergen data available. This resulted in a beta of 1.19 (95% CI: 1.02–1.39), which is similar in direction and magnitude to the main model. We then added aeroallergen sensitization in the model and found that this made no difference to the association seen in the restricted sample (β: 1.19, 95% CI: 1.02–1.38).

We also performed a subgroup analysis within African American subjects only ($n = 693$) and observed a similar effect for ancestry on recurrent wheezing (OR: 1.22, 95% CI: 0.87–1.69). However, this effect is no longer significant, potentially due to the reduction of numbers of subjects included in this analysis by approximately half.

Discussion

In this study, we sought to evaluate whether self-identified race or genetic ancestry were more robustly associated with early childhood wheezing morbidity, and whether these associations would persist after accounting for socio-economic and early life risk factors associated with airway calibre. Contrary to other studies [1–4], self-identified race was not independently associated with respiratory morbidity, although our cohort had only 7.4% white children, which affected the power for this analysis. However, African ancestry within the population as a whole was highly significantly (OR for 10% increment: 1.19 CI 1.05–1.34) associated with recurrent wheeze independently from SES, known confounders and

early life exposures. European ancestry was, on the other hand, associated with a decreased risk of recurrent wheezing (OR for 10% increment: 0.84 CI 0.74–0.94). There was no interaction between genetic ancestry and early life factors felt to affect lung function including preterm birth, chorioamnionitis, antenatal maternal smoking and post-natal maternal smoking.

There are a number of points which bear discussion. First, there was a highly significant association between African ancestry and risk of recurrent wheezing. There was also a decreased risk of wheezing for European ancestry. The strength of the association with African ancestry may in part be due to a wide range of African ancestry in multiple racial groups aside from self-identified African Americans. As such, ancestry across a multi-ethnic sample may be a better proxy than self-identified race for differentially distributed environmental factors (such as discrimination, perceived stress, or early life nutrition) or differentially distributed genetic factors, which affect wheezing morbidity. Furthermore, self-identified race is an imprecise categorical variable that will have less statistical power to evaluate these types of associations compared with ancestry [18]. This contention is supported by the fact that while self-identified race showed an effect in the same direction, it never attained significance.

The lack of a significant association for self-identified race and recurrent wheezing contrasts with other studies in the literature regarding racial disparity and childhood asthma,[1–4], partly because of the small number of white children, which lowered the power to find a significant effect. In our analyses, white children had a 54–69% reduction in the risk for wheezing, although these did not reach statistical significance. In addition, this may be because wheezing in early childhood is a heterogeneous disease. Many children have transient wheezing associated with initially smaller airway calibre that improves over time [6]. As African ancestry allows for more precise predictions of plateau lung function compared with self-identified race [18], it is also likely that ancestry is associated with lung function and hence airway calibre in the first few years of life, which in turn would have an impact on a significant proportion of early childhood wheezing. However, in the absence of infant lung function, this remains speculative. As the children become older, wheezing episodes will be more representative of asthma morbidity as opposed to airway calibre. It would not be surprising if the associations for both race and ancestry become significant in older children given clear racial disparities in asthma morbidity [1, 5]. Further studies in larger multi-ethnic birth cohorts over the range of childhood years are needed to resolve this question.

As stated above, the degree of significance of the association compared with that for self-identified race may suggest that ancestry may be a more precise proxy for the early life environmental exposures and genetic factors for which race is often used as a proxy. This issue will become progressively more important with the relative increase in the US census of self-identified non-white racial/ethnic groups [30, 31]. The changing pattern in the US census is important given the admixed ancestral origins of these populations. While African Americans have approximately 80% African ancestry, this varies widely within African American populations [32–34]. Similarly, various Latino groups have wide variation in African, European, and Native American ancestry both across and within populations [35, 36]. Also, while the current census allows reporting of multiple races, since 1989, the national center for health statistics [25] and hence large national cohorts [37, 38] have used mother’s race as a proxy for that of the infant. Even with the change in census forms, self-identified race will become increasingly less precise with more inter-racial marriages and more parents themselves have mixed racial heritage. Even though ancestry will vary within admixed populations, self-identified race will correlate with other ethno-cultural risk factors which may not vary within admixed populations. As we have found for other diseases, both

variables may provide useful information serving as proxies for different factors relevant to disease susceptibility within subgroups of a population [39].

As recently discussed by Rotimi and Jorde, population categories are inadequate to describe the range of genetic variation among individuals [40]. Furthermore, risk alleles for 26 diseases in 11 populations showed significant variability in prevalence across populations [41]. Thus, ancestry may have more relevance to genetic bases of disease than self-identified race. A caution needs to be made about interpreting ancestry as differentially distributed genetic factors. Any ancestry association will need to be further investigated to determine if specific differentially distributed risk alleles can be determined [42].

There are a number of limitations in our study. African ancestry may co-vary with other pre-natal exposures, which are also associated with asthma and respiratory morbidity resulting in residual confounding such as lower levels of vitamin D levels [43–47]. Similarly, discrimination is correlated with skin pigmentation [48] even within African Americans and this in turn may impact SES [49]. While we account for SES with annual household income and level of maternal education achieved, it remains possible that there are some more subtle differences in SES that are not accounted for. Secondly, the ancestry variable is derived from a panel of 150 AIMs, which is sufficient to control for population stratification for African ancestry as noted by other investigators [27]. While GWAS-based estimates may be more precise, reasonable accuracy may be obtained using estimates from a dramatically smaller numbers of AIMs [50]. We have evaluated the differences in such estimates compared to those generated from AIMs panels in national cohorts and found these estimates of ancestry to be highly correlated (data not shown). However, a larger panel of AIMs allowing estimation of Native American as well as African and European ancestry would improve precision in estimation of effect. Also, the use of ancestry across a population with multiple self-identified ethnicities is less common. However, this approach has been used in the past in multi-ethnic populations [22, 51], and has been shown to result in estimates, which are highly correlated with the results of a principal component- based analysis [51].

There are a number of limitations to this study, which relate to the population of interest. There are a limited number of white subjects in inner-city samples, which may decrease power to detect associations with self-identified race. However, wheezing-related racial disparities are a particular problem for inner-city populations. Our study shows that using genetic ancestry to help delineate associations with wheezing disparities has utility even within a population such as ours where a majority of subjects self-identify as African American. Despite the large size of our cohort, it was a general population cohort unselected by parental risk of asthma, which limits our sample size of children with recurrent wheezing. The general population approach increases relevance to the population and phenotype of interest. We were interested in the factors associated with early childhood recurrent wheezing in the inner city as opposed to simply evaluating risk factors for morbidity in children with a high risk of early asthma. Another limitation is that there may have been some subjects who may have acute care, which occurred outside of our healthcare system. This is likely to be a small number. Subjects were only included in the cohort if they intended to receive primary care at Boston Medical Center and its affiliated clinics. There was a proportion of subjects without all key variables who were not included in the analysis, but these individuals did not differ in key socio-demographic variables. Finally, our sample is derived from an inner-city multiethnic population. It remains to be seen whether similar findings would be seen in a more affluent population.

In conclusion, we found that African ancestry was associated with recurrent wheezing within a multi-ethnic sample independently of SES. This raises the question whether

ancestry will serve as a powerful and novel method to examine wheezing disparities, and serve as a useful proxy for differentially distributed genetic or early life exposures, which affect disease. Also, there were no interactions between ancestry and early life risk factors associated with airway calibre including preterm birth, chorioamnionitis and patterns of maternal pre-natal and post-natal smoking on the outcome of respiratory morbidity. However, other more precise phenotypes like early life lung function may be better able to investigate these associations. Further investigation is needed in this area.

Acknowledgments

Funding: The parent study is in part supported by the March of Dimes PERI grants [PI: Wang, 20-FY02-56]; and NIEHS [PI: Wang, R21 ES011666]; and NICHD [PI: Wang, R01 HD041702]. The follow-up study is in part supported by the Food Allergy Initiative and the Sunshine Charitable Foundation, the NIAID [PI Wang, R21AI079872; U01AI090727] and the Department of Defense [PI Wang, W81XWH-10-1-0123]. Dr Kumar is also supported by the NHLBI [PI: Kumar, K23HL093023]. Dr Burchard and Gignoux are supported by HL088133.

References

1. Akinbami LJ, LaFleur BJ, Schoendorf KC. Racial and income disparities in childhood asthma in the United States. Ambul Pediatr. 2002; 2:382–7. [PubMed: 12241134]
2. Claudio L, Stingone JA, Godbold J. Prevalence of childhood asthma in urban communities: the impact of ethnicity and income. Ann Epidemiol. 2006; 16:332–40. [PubMed: 16242960]
3. Haselkorn T, Lee JH, Mink DR, Weiss ST, Group TS. Racial disparities in asthma-related health outcomes in severe or difficult-to-treat asthma. Ann Allergy Asthma Immunol. 2008; 101:256–63. [PubMed: 18814448]
4. Litonjua AA, Carey VJ, Weiss ST, Gold DR. Race, socioeconomic factors, and area of residence are associated with asthma prevalence. Pediatr Pulmonol. 1999; 28:394–401. [PubMed: 10587412]
5. Akinbami LJ, Moorman JE, Garbe PL, Sondik EJ. Status of childhood asthma in the United States, 1980–2007. Pediatrics. 2009; 123(Suppl 3):S131–45. [PubMed: 19221156]
6. Morgan WJ, Stern DA, Sherrill DL, et al. Outcome of asthma and wheezing in the first 6 years of life: follow-up through adolescence. Am J Respir Crit Care Med. 2005; 172:1253–8. [PubMed: 16109980]
7. Vergara C, Caraballo L, Mercado D, et al. African ancestry is associated with risk of asthma and high total serum IgE in a population from the Caribbean Coast of Colombia. Hum Genet. 2009; 125:565–79. [PubMed: 19290544]
8. Choudhry S, González Burchard E, Borrell LN, et al. Ancestry-environment interactions and asthma risk among Puerto Ricans. Am J Resp Crit Care Med. 2006; 174:1088–93. [PubMed: 16973984]
9. Fiscella K. Race, perinatal outcome, and amniotic infection. Obstet Gynecol Surv. 1996; 51:60–6. [PubMed: 8657398]
10. Alexander GR, Slay M. Prematurity at birth: trends, racial disparities, and epidemiology. Ment Retard Dev Disabil Res Rev. 2002; 8:215–20. [PubMed: 12454897]
11. Fiscella K. Racial disparities in preterm births. The role of urogenital infections. Public Health Rep. 1996; 111:104–13. [PubMed: 8606905]
12. Berg JZ, Mason J, Boettcher AJ, Hatsukami DK, Murphy SE. Nicotine metabolism in African Americans and European Americans: variation in glucuronidation by ethnicity and UGT2B10 haplotype. J Pharmacol Exp Ther. 2010; 332:202–9. [PubMed: 19786624]
13. Dezateux C, Stocks J, Wade AM, Dundas I, Fletcher ME. Airway function at one year: association with premorbid airway function, wheezing, and maternal smoking. Thorax. 2001; 56:680–6. [PubMed: 11514687]
14. Gilliland FD, Berhane K, McConnell R, et al. Maternal smoking during pregnancy, environmental tobacco smoke exposure and childhood lung function. Thorax. 2000; 55:271–6. [PubMed: 10722765]
15. Getahun D, Strickland D, Zeiger RS, et al. Effect of chorioamnionitis on early childhood asthma. Arch Pediatr Adolesc Med. 2010; 164:187–92. [PubMed: 20124149]

16. Rona RJ, Gulliford MC, Chinn S. Effects of prematurity and intrauterine growth on respiratory health and lung function in childhood. BMJ. 1993; 306:817–20. [PubMed: 8490372]
17. Kumar R. Prenatal factors and the development of asthma. Curr Opin Pediatr. 2008; 20:682–7. [PubMed: 19005336]
18. Kumar R, Seibold MA, Aldrich MC, et al. Genetic ancestry in lung-function predictions. N Engl J Med. 2010; 363:321–30. [PubMed: 20647190]
19. Wang X, Zuckerman B, Pearson C, et al. Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight. JAMA. 2002; 287:195– 202. [PubMed: 11779261]
20. Kumar R, Yu Y, Story RE, et al. Prematurity, chorioamnionitis, and the development of recurrent wheezing: a prospective birth cohort study. J Allergy Clin Immunol. 2008; 121:878–84 e6. [PubMed: 18313129]
21. Tsai HJ, Liu X, Mestan K, et al. Maternal cigarette smoking, metabolic gene polymorphisms, and preterm delivery: new insights on GxE interactions and pathogenic pathways. Hum Genet. 2008; 123:359–69. [PubMed: 18320229]
22. Tsai HJ, Yu Y, Zhang S, et al. Association of genetic ancestry with preterm delivery and related traits among African American mothers. Am J Obstet Gynecol. 2009; 201:94 e1–10. [PubMed: 19446788]
23. Gibbs RS, Blanco JD, St Clair PJ, Castaneda YS. Quantitative bacteriology of amniotic fluid from women with clinical intraamniotic infection at term. J Infect Dis. 1982; 145:1–8. [PubMed: 7033397]
24. Gupta M, Mestan KK, Martin CR, et al. Impact of clinical and histologic correlates of maternal and fetal inflammatory response on gestational age in preterm births. J Matern Fetal Neonatal Med. 2007; 20:39–46. [PubMed: 17437198]
25. Office of Management and Budget US. Revisions to the standards for the classification of federal data on race and ethnicity. Office of Management and Budget US; Washington, DC: 1997.
26. Hong X, Tsai HJ, Liu X, et al. Does genetic regulation of IgE begin in utero? Evidence from T(H)1/T(H)2 gene polymorphisms and cord blood total IgE. J Allergy Clin Immunol. 2010; 126:1059–67. 67 e1. [PubMed: 21050946]
27. Kosoy R, Nassir R, Tian C, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat. 2009; 30:69–78. [PubMed: 18683858]
28. Castro-Rodriguez JA, Holberg CJ, Wright AL, Martinez FD. A clinical index to define risk of asthma in young children with recurrent wheezing. Am J Respir Crit Care Med. 2000; 162(4 Pt 1): 1403–6. [PubMed: 11029352]
29. Guilbert TW, Morgan WJ, Zeiger RS, et al. Atopic characteristics of children with recurrent wheezing at high risk for the development of childhood asthma. J Allergy Clin Immunol. 2004; 114:1282–7. [PubMed: 15577824]
30. [accessed September 2011] census.gov. http://www.census.gov/. [updated 2009; cited 2009]; Available from: http://www.census.gov/popest/national/asrh/NC-EST2009-srh.html
31. U.S. Census Bureau. National Population Estimates, Characteristics, National Demographic Components of Change by Hispanic or Latino Origin. Washington, DC: U.S. Census Bureau; 2003.
32. Halder I, Yang BZ, Kranzler HR, Stein MB, Shriver MD, Gelernter J. Measurement of admixture proportions and description of admixture structure in different U.S. populations. Hum Mutat. 2009; 30:1299–309. [PubMed: 19572378]
33. Parra EJ, Kittles RA, Argyropoulos G, et al. Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. Am J Phys Anthropol. 2001; 114:18–29. [PubMed: 11150049]
34. Yaeger R, Avila-Bront A, Abdul K, et al. Comparing genetic ancestry and self-described race in african americans born in the United States and in Africa. Cancer Epidemiol Biomarkers Prev. 2008; 17:1329–38. [PubMed: 18559547]
35. Choudhry S, Seibold MA, Borrell LN, et al. Dissecting complex diseases in complex populations: asthma in latino americans. Proc Am Thorac Soc. 2007; 4:226–33. [PubMed: 17607004]

36. Lai CQ, Tucker KL, Choudhry S, et al. Population admixture associated with disease prevalence in the Boston Puerto Rican health study. Hum Genet. 2009; 125:199–209. [PubMed: 19107526]
37. Davis RR, Hofferth SL. The association between inadequate gestational weight gain and infant mortality among U.S. infants born in 2002. Matern Child Health J. 201010.1007/s10995-010-0713-5
38. Wingate MS, Barfield WD. Racial and ethnic variations in temporal changes in fetal deaths and first day infant deaths. Matern Child Health J. 201010.1007/s10995-010-0665-9
39. Kumar R, Tsai HJ, Hong X, et al. Race, Ancestry and Development of Food Allergen Sensitization in Early Childhood. Pediatrics. 201110.1542/peds.2011-0691
40. Rotimi CN, Jorde LB. Ancestry and disease in the age of genomic medicine. N Engl J Med. 2010; 363:1551–8. [PubMed: 20942671]
41. Adeyemo A, Rotimi C. Genetic variants associated with complex human diseases show wide variation across multiple populations. Public Health Genomics. 2010; 13:72–9. [PubMed: 19439916]
42. Scanlon PD, Shriver MD. “Race correction” in pulmonary-function testing. N Engl J Med. 2010; 363:385–6. [PubMed: 20647189]
43. Brehm JM, Schuemann B, Fuhlbrigge AL, et al. Serum vitamin D levels and severe asthma exacerbations in the Childhood Asthma Management Program study. J Allergy Clin Immunol. 2010; 126:52–8 e5. [PubMed: 20538327]
44. Freishtat RJ, Iqbal SF, Pillai DK, et al. High prevalence of vitamin D deficiency among inner-city African American youth with asthma in Washington, DC. J Pediatr. 2010; 156:948– 52. [PubMed: 20236657]
45. Miyake Y, Sasaki S, Tanaka K, Hirota Y. Dairy food, calcium and vitamin D intake in pregnancy, and wheeze and eczema in infants. Eur Respir J. 2010; 35:1228–34. [PubMed: 19840962]
46. Searing DA, Zhang Y, Murphy JR, Hauk PJ, Goleva E, Leung DY. Decreased serum vitamin D levels in children with asthma are associated with increased corticosteroid use. J Allergy Clin Immunol. 2010; 125:995– 1000. [PubMed: 20381849]
47. Signorello LB, Williams SM, Zheng W, et al. Blood vitamin d levels in relation to genetic estimation of African ancestry. Cancer Epidemiol Biomarkers Prev. 2010; 19:2325–31. [PubMed: 20647395]
48. Klimentidis YC, Miller GF, Shriver MD. Genetic admixture, self-reported ethnicity, self-estimated admixture, and skin pigmentation among Hispanics and Native Americans. Am J Phys Anthropol. 2009; 138:375–83. [PubMed: 18951390]
49. Krieger N, Sidney S, Coakley E. Racial discrimination and skin color in the CARDIA study: implications for public health research. Coronary Artery Risk Development in Young Adults. Am J Public Health. 1998; 88:1308–13. [PubMed: 9736868]
50. Tian C, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. Hum Mol Genet. 2008; 17:R143–50. [PubMed: 18852203]
51. Wang H, Haiman CA, Kolonel LN, et al. Self-reported ethnicity, genetic structure and the impact of population stratification in a multiethnic study. Hum Genet. 2010; 128:165–77. [PubMed: 20499252]

Table 1

Baseline characteristics of cohort by self-identified race/ethnicity

	Total sample (n = 1034)	**Black (n = 696)**	**White (n = 76)**	**Hispanic (n = 262)**	
	***N* (%)**	***N* (%)**	***N* (%)**	***N* (%)**	***P*-Value**
Outcome variable					
Recurrent wheezing	63 (6.1)	53 (7.6)	3 (4.0)	7 (2.7)	**0.01**
Key exposure variables					
Preterm birth	280 (27.1)	184 (26.4)	22 (29.0)	74 (28.2)	0.80
Inflammation/chorioamnionitis	148 (14.3)	106 (15.2)	12 (15.8)	30 (11.5)	0.30
Maternal post-natal smoking	161 (15.6)	102 (14.7)	28 (36.8)	31 (11.8)	**<0.0001**
Antenatal smoke exposure	103 (10.0)	62 (8.9)	26 (34.2)	15 (5.7)	**<0.0001**
Other variables					
Male gender	530 (51.2)	356 (51.2)	45 (59.2)	129 (49.2)	0.31
Age of child, years (mean, SD)	3.08, 2.29	3.21, 2.31	2.78, 2.45	2.82, 2.18	**0.03**
Maternal atopy	346 (33.5)	232 (33.3)	33 (43.4)	81 (30.9)	0.13
c-section delivery	341 (33.1)	232 (33.5)	22 (29.0)	87 (33.3)	0.72
Firstborn child	428 (41.4)	279 (40.1)	46 (60.5)	103 (39.3)	**0.002**
Exclusive breastfeeding	55 (5.3)	30 (4.3)	7 (9.1)	18 (6.9)	0.08
Highest level of Maternal Education Achieved					
<HS	307 (29.7)	178 (25.6)	13 (17.1)	116 (44.3)	
HS	372 (36.0)	265 (38.1)	25 (32.9)	82 (31.3)	
some college	355 (34.3)	253 (36.4)	38 (50.0)	64 (24.4)	**<0.0001**
Annual Household income					
<$30000	512 (49.5)	360 (51.7)	30 (39.5)	122 (46.6)	
≥$30000	302 (29.2)	196 (28.2)	34 (44.7)	72 (27.5)	
unknown	220 (21.3)	140 (20.1)	12 (15.8)	68 (26.0)	**0.008**
Ancestry, median% (IQR)					
European	12.6 (3.8–46)	6.7 (2.2–13.7)	95.1 (87.1–97.8)	51.3 (40–62.6)	**<0.0001**
African	79.7 (37.5–89.6)	86.1 (79.3–93.4)	1.1 (0.6–7.0)	21.2 (6.5–42.9)	**<0.0001**
Asian	5.1 (2.5–11.3)	4.4 (2.2–8.2)	2.8 (1.3–6.6)	15.9 (6.5–38.7)	**<0.0001**

All *P*-values represent differences by race. *P*-values for categorical variables are determined by chi-square statistic. All *P*-values for continuous variables are determined by ANOVA. *P*-values less than 0.05 are in bold.

Table 2

Crude associations between key variables

	OR	CI	*P*-value
Univariate associations of non-ancestry variables			
Associations with recurrent wheezing			
Preterm birth	2.62	1.57–4.39	**<0.0003**
Inflammation/chorioamnionitis	2.59	1.45–4.61	**0.002**
Antenatal smoke exposure	2.02	1.02–4.01	**0.007**
Maternal post-natal smoking	1.60	0.86–2.98	0.15
Maternal atopy	2.02	1.21–3.37	**0.007**
Caesarean delivery	1.17	0.69–1.99	0.55
Firstborn child	1.14	0.68–1.91	0.26
Exclusive breastfeeding	0.57	0.14–2.39	0.44
Maximal level of maternal education			
Some college	1.00	–	–
HS	0.91	0.52–1.60	0.75
<HS	0.41	0.20–0.86	**0.02**
Annual household income			
≥30K	1.00	–	–
<30K	1.45	0.79–2.69	0.24
Unknown	1.10	0.51–2.41	0.80
Self-identified race			
Black	1.00	–	–
White	0.50	0.15–1.64	0.25
Hispanic/Latino	0.33	0.15–0.74	**0.007**
Chorioamnionitis with prematurity			
Inflammation/chorioamnionitis	2.36	1.64–3.39	**>0.0001**
Crude (non-univariate) models for non-ancestry variables			
Association of both preterm birth and chorioamnionitis with recurrent wheezing			
Preterm birth	2.36	1.39–3.99	0.001
Inflammation/chorioamnionitis	2.20	1.21–3.96	0.009
Association of interaction term for preterm birth and chorioamnionitis with recurrent wheezing			
Preterm birth	2.16	1.17–4.01	0.01
Inflammation/chorioamnionitis	1.83	0.73–4.58	0.20
Preterm*chorioamnionitis	1.38	0.41–4.65	0.53
Univariate associations of African ancestry (per 10% African ancestry)			
With recurrent wheezing	1.16	1.05–1.28	0.003
With chorioamnionitis	1.01	0.96–1.07	0.55
With preterm birth	0.98	0.94–1.02	0.34
	β	CI	*P*-value
With gestational age in weeks	−0.02	−0.08,0.05	0.62

All ancestry analyses included estimates of individual global ancestry with variables for African and Asian ancestry, assuming three ancestral populations.

Table 3

(a) Ancestry, race and recurrent wheezing accounting for socio-economic status and demographic factors. (b) Model with inclusion of known confounders including maternal atopy, caesarean section, firstborn status, exclusive breastfeeding. (c) Model with inclusion of known confounders listed above and also early life risk factors including preterm birth, chorioamnionitis, antenatal smoke exposure and post-natal smoke exposure

	Models with self-reported race			Models with genetic ancestry*		
Variables	**OR**	**CI**	***P*-value**	**OR**	**CI**	***P*-value**
(a) Base models including SES and Demographic factors						
White	**0.46**	**0.14–1.57**	**0.22**			
Hispanic	**0.44**	**0.19–1.00**	**0.05**			
African ancestry				**1.15**	**1.04–1.28**	**0.007**
European ancestry				**0.85**	**0.76–0.96**	**0.01**
(b) Addition of known confounders						
White	**0.40**	**0.11–1.39**	**0.15**			
Hispanic	**0.46**	**0.20–1.05**	**0.06**			
African ancestry				**1.16**	**1.05–1.29**	**0.005**
European ancestry				**0.85**	**0.75–0.96**	**0.03**
(c) Addition of known confounders and Early life exposures						
White	**0.31**	**0.09–1.14**	**0.08**			
Hispanic	**0.47**	**0.20–1.08**	**0.08**			
African ancestry				**1.17**	**1.05–1.31**	**0.005**
European ancestry				**0.84**	**0.74–0.94**	**0.004**

*Ancestry is mean centred for interaction models and expressed in 10% increments. Models for African and European ancestry carried out separately.