Genetic variation in chitinase 3-like 1 (*CHI3L1*) contributes to asthma severity and airway expression of YKL-40

Jose L. Gomez, MD, MS,[a] Gina M. Crisafi, BS,[b] Carole T. Holm, RN,[a] Deborah A. Meyers, PhD,[c] Gregory A. Hawkins, PhD,[c] Eugene R. Bleecker, MD,[c] Nizar Jarjour, MD,[b] for the Severe Asthma Research Program (SARP) Investigators, Lauren Cohn, MD,[a] and Geoffrey L. Chupp, MD[a] *New Haven, Conn, Madison, Wis, and Winston-Salem, NC*

Background: Single nucleotide polymorphisms (SNPs) in the chitinase 3-like 1 (*CHI3L1*) promoter, the gene encoding YKL-40, are associated with circulating YKL-40 levels and asthma prevalence. However, the effects of gene polymorphisms on asthma severity and airway expression of YKL-40 have not been examined.

Objective: We sought to determine the effect of genetic variation in *CHI3L1* on asthma severity and YKL-40 expression in subjects from the Yale Center for Asthma and Airways Disease and the Severe Asthma Research Program.

Methods: SNPs spanning the *CHI3L1* gene were genotyped in 259 Yale Center for Asthma and Airways Disease and 919 Severe Asthma Research Program subjects. Association and haplotype analyses were conducted to identify effects on airflow obstruction, YKL-40 levels, and asthma severity.

Results: Fifteen SNPs in *CHI3L1* were associated with FEV_1, serum YKL-40 levels, or both. rs12141494 (intron 6) was the only SNP in subjects of European ancestry in both cohorts that was associated with serum YKL-40 levels and postbronchodilator FEV_1. Conditional analysis demonstrated that the effect on lung function was independent of the promoter SNP rs4950928, and haplotype analysis demonstrated that G alleles at rs12141494 and rs4950928 are associated with lower YKL-40 expression and higher FEV_1 percent predicted values. In asthmatic subjects the risk allele A at rs12141494 was associated with severe asthma and higher YKL-40 expression in the airway ($P \leq .05$).

Conclusion: In contrast to the promoter SNP rs4950928, the intronic SNP rs12141494 in *CHI3L1* is associated with asthma severity, lung function, and YKL-40 expression in the blood and airway. These data suggest that SNP rs12141494 modulates YKL-40 expression in the airway and contributes to airway remodeling and asthma severity. (J Allergy Clin Immunol 2014;■■■:■■■-■■■.)

Key words: *Asthma, asthma severity, severe asthma, genetic association studies, genetic variation, CHI3L1 protein, human, YKL-40 protein, human, airway remodeling*

YKL-40, a chitinase-like protein, belongs to the chitinase and chitinase-like family of proteins. These evolutionarily conserved molecules interact with chitin, the second most abundant polysaccharide on earth. Human and mechanistic studies have demonstrated that chitinase 3-like 1 *(CHI3L1)*/YKL-40 plays a role in the pathobiology of asthma by modulating innate and adaptive immune and remodeling responses.[1] In the first human asthma studies we demonstrated that increased serum levels of YKL-40 correlate with asthma severity, airway remodeling, and increased thickness of the subepithelial basement membrane in subjects with asthma.[2] This work led to mechanistic studies demonstrating that bronchial epithelial cells exposed to YKL-40 generate higher levels of IL-8 and stimulate smooth muscle proliferation *in vitro*.[3] An additional mechanism of YKL-40–increased bronchial smooth muscle proliferation involves the protease-activated receptor 2.[4] Taken together, these studies established an important role for YKL-40 as a molecule uniquely juxtaposed between environmental exposure, inflammation, and the development of airway remodeling and severe asthma.

Genetic studies have also demonstrated that variation in *CHI3L1* contributes to the pathogenesis of asthma. A genome-wide association study of serum YKL-40 levels in a founder population of European descent by Ober et al[5] identified an association between the rs4950928 promoter single nucleotide polymorphism (SNP) and circulating YKL-40 levels. This SNP was also associated with the risk of asthma and FEV_1 in 3 populations of European ancestry (EA) with mild asthma. Subsequent reports have shown an association with rs4950928 and asthma,[6] although the risk allele was opposite of that reported by Ober et al.[5] Additional publications have been conflicting with respect to the effect of genetic variation in the *CHI3L1* gene and asthma.[7,8] To date, the effect of genetic variation in *CHI3L1* on asthma severity has not been examined.

To determine the effects of genetic variation in the *CHI3L1* gene on YKL-40 expression in the airway, airway remodeling, and asthma severity, we examined 2 cohorts of subjects with

From [a]the Section of Pulmonary, Critical Care and Sleep Medicine and the Department of Internal Medicine, Yale University School of Medicine, New Haven; [b]the Section of Pulmonary and Critical Care Medicine, University of Wisconsin School of Medicine, Madison; and [c]the Center for Human Genomics and Personalized Medicine, Wake Forest School of Medicine, Winston-Salem.

Supported by National Institutes of Health training grants T32HL007778-18 and T15LM007056-26 and R01HL095390-03 (to G.L.C.), R01HL069116, and GCRC RR03186 (to G.M.C. and N.J.) and Flight Attendant Medical Research Institute Young Clinical Scientist Award 113393 (to J.L.G.).

Disclosure of potential conflict of interest: D. A. Meyers, E. R. Bleecker, and G. A. Hawkins' institution has received funding from the National Institutes of Health (NIH; RO1HL69167). J. L. Gomez's institution has received funding from the NIH and the Flight Attendant Medical Research Institute, and he has received compensation for travel and other meeting-related expenses from GlaxoSmithKline. N. Jarjour's institution has received funding from the National Heart, Lung, and Blood Institute. L. Cohn has received compensation for board membership from MannKind Corporation, as well as consultancy fees from Genentech and Novartis, and has received payment for delivering lectures from Genentech. G. L. Chupp has received consultancy fees from MedImmune and MannKind, as well payment for lectures from AstraZeneca, Genentech, and Boston Scientific. C. T. Holm declares no relevant conflicts of interest.

Received for publication June 14, 2014; revised November 11, 2014; accepted for publication November 12, 2014.

Corresponding author: Geoffrey L. Chupp, MD, Yale School of Medicine, Pulmonary and Critical Care Medicine, 300 Cedar St (S441 TAC), New Haven, CT 06520-8057. E-mail: geoffrey.chupp@yale.edu.

0091-6749/$36.00

http://dx.doi.org/10.1016/j.jaci.2014.11.027

Abbreviations used
AA: African Ancestry
ATS: American Thoracic Society
CHI3L1: Chitinase 3-like 1
EA: European Ancestry
LD: Linkage disequilibrium
SARP: Severe Asthma Research Program
SNP: Single nucleotide polymorphism
YCAAD: Yale Center for Asthma and Airways Disease

asthma from the Yale Center for Asthma and Airways Disease (YCAAD) and the Severe Asthma Research Program (SARP). We hypothesized that genetic variation in *CHI3L1* is associated with persistent airflow obstruction, serum YKL-40 levels, asthma severity, and airway expression of YKL-40. To examine this hypothesis, we characterized the effect of SNPs in the *CHI3L1* gene on severe asthma traits, determined the interaction between SNPs using haplotype analysis, and correlated identified SNPs with airway expression of YKL-40. Ultimately, we identified a novel polymorphism in *CHI3L1* that is likely to contribute to airway remodeling and asthma severity through increased production of YKL-40 in the airway.

METHODS

Populations

YCAAD. Study participants in the YCAAD cohort based in New Haven, Connecticut, underwent an extensive phenotypic characterization after institutional review board approval. Inclusion and exclusion criteria and the study protocol have been described previously.[2] Severe asthma was defined as outlined by the SARP clustering algorithm and was modified as follows: *severe asthma* included subjects with a baseline FEV_1 of less than 68% of predicted value (SARP clusters 4 and 5), whereas *nonsevere asthma* was defined by the presence of a baseline FEV_1 equal to or greater than 68% of predicted value (SARP clusters 1-3).[9] For specific details related to definitions and specific study measurements, see the Methods section in this article's Online Repository at www.jacionline.org. A total of 259 subjects from this cohort were analyzed.

SARP. Subjects in the SARP cohort completed study visits by using established standard operating procedures, as previously described.[9] The definition for severe asthma outlined above was also used in this cohort. Institutional review board approval at the SARP institutions was obtained for these studies. The characteristics of these subjects have been reported in previous publications.[9,10] A total of 919 subjects from this cohort were analyzed.

Sputum induction

Subjects in the YCAAD cohort underwent sputum induction with inhaled hypertonic saline. Mucus plugs were removed by using a dissecting microscope and washed to remove squamous cell contamination. The cellular and aqueous compartments were separated, and cell counts, cell differentials, and viability were determined (Diff-Quik and trypan blue exclusion). Aliquots of supernatants were stored until processing. Samples with greater than 20% squamous cells were considered contaminated and were not processed further.

Measurement of serum and sputum YKL-40 levels

YKL-40 levels were measured in duplicate in serum and sputum specimens by using a commercially available ELISA kit for YKL-40 (Quidel, San Diego, Calif). The mean value of the 2 duplicates was used in the statistical analyses. Duplicate samples with coefficients of variation of greater than 20% were reassayed. The mean coefficient of variation for all samples was 10%. The limit of detection is 5.4 ng/mL. Values of less than the limit of detection were defaulted to 0 ng/mL.

SNP genotyping

DNA was extracted from peripheral blood leukocytes by using standard protocols. The Sequenom MassARRAY system (Sequenom, San Diego, Calif) genotyping platform was used. Genotyping was performed according to the manufacturer's iPLEX application guide at the Center for Human Genomics at Wake Forest University Health Sciences and the Yale Keck DNA sequencing facility for SARP and YCAAD samples, respectively.[11] Resequencing of the *CHI3L1* gene was conducted on a subset of 6 subjects of EA in the YCAAD cohort, 3 with the CC genotype and 3 with GG genotype at the promoter SNP rs4950928 and with high and low YKL-40 levels, respectively, to identify rare and novel SNPs that could be associated with YKL-40 expression. Amplicon distribution was as follows: 13 kb upstream-5′, 7.8 kb *CHI3L1* gene, 1.2 kb downstream-3′, as previously described.[12] The identified SNPs were compared with SNPs genotyped in previous publications.[5,6] These SNPs were analyzed with Haploview to identify tagging polymorphisms in the *CHI3L1* gene by using data from the HapMap project (version 3, release 27).[13] A total of 17 SNPs spanning the *CHI3L1* gene identified through sequencing, previous publications,[5,6] and haplotype tagging were genotyped in YCAAD (discovery). Simultaneously, 14 SNPs identified through haplotype tagging by using data from the HapMap project (version 3, release 27)[13] were genotyped in SARP (validation). A total of 9 SNPs present in both data sets tagged all other SNPs reported in this study. Genotyping criteria for SNP quality included the removal of SNPs with low call rates (<80%), SNPs that violated Hardy-Weinberg equilibrium ($P < .05$) and had a minor allele frequency of less than 4%, and samples with low call rates (<98%). Pairwise marker linkage disequilibrium (LD) was estimated by using the Lewontin D′ statistic and r^2 with Haploview.[13]

Statistical analysis

R software (R Foundation for Statistical Computing, Vienna, Austria) was used for data analysis. The R stats (version 3.0.1), SNPassoc (version 1.8-5), haplo.stats (version 1.6.3), and car (version 2.0-21) packages were implemented. Continuous variables were tested by using nonparametric tests, including the Kruskal-Wallis test. Categorical variables were analyzed with χ^2 tests. Results are reported as medians and interquartile ranges, unless otherwise specified. Hardy-Weinberg equilibrium calculation was performed to assess the distribution of the SNPs in each group by using Haploview.[13] The SNPassoc package applies a generalized linear model to generate association *P* values. By using this statistical package, dominant, codominant, recessive, and additive models of association were tested for postbronchodilator FEV_1 percent predicted, which was used as a surrogate marker of airway remodeling, and serum YKL-40 levels on the discovery phase of the study. The additive association model was identified as the best fit based on the Akaike information criterion,[14] and all reported association values were generated by using this model and adjusted for age and sex based on the effect these 2 features have on YKL-40 levels,[15] pulmonary function,[16] and the best fit based on the Akaike information criterion.[14] In the discovery cohort (YCAAD) *P* values were adjusted for multiple testing with the Benjamini-Hochberg procedure (false discovery rate < 0.1)[17] to balance power and the potential for false-positive results. Association values in the validation cohort (SARP) were adjusted for multiple testing with the Bonferroni method ($P < .05$) to provide a more stringent and reliable measure of true association. We conducted a conditional analysis on subjects of EA with the rs4950928 CC genotype to evaluate for independent associations on nonpromoter SNPs; all *P* values on the conditional analyses were adjusted with the Bonferroni method. Haplotypes were inferred by using the expectation maximization algorithm in Haploview.[13] Haplotypes with a frequency of greater than 5% were considered in the analysis. A generalized linear model using the expectation maximization algorithm described by Excoffier and Slatkin[18] was implemented in the haplotype analysis with the haplo.stats R package. Asthmatic and control subjects were considered in the discovery, validation,