

Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes

Katherine J. Baines, PhD,^{a,b} Jodie L. Simpson, PhD,^{a,b} Lisa G. Wood, PhD,^{a,b} Rodney J. Scott, PhD,^c Naomi L. Fibbens, DipPathTech,^{a,b} Heather Powell, MMedSci,^{a,b} Douglas C. Cowan, PhD,^d D. Robin Taylor, DSc,^d Jan O. Cowan,^d and Peter G. Gibson, MBBS^{a,b} *New Lambton Heights, Australia, and Dunedin, New Zealand*

Background: Airway inflammation is associated with asthma exacerbation risk, treatment response, and disease mechanisms. **Objective:** This study aimed to identify and validate a sputum gene expression signature that discriminates asthma inflammatory phenotypes.

Methods: An asthma phenotype biomarker discovery study generated gene expression profiles from induced sputum of 47 asthmatic patients. A clinical validation study ($n = 59$ asthmatic patients) confirmed differential expression of key genes. A 6-gene signature was identified and evaluated for reproducibility ($n = 30$ asthmatic patients and $n = 20$ control subjects) and prediction of inhaled corticosteroid (ICS) response ($n = 71$ asthmatic patients). Receiver operating characteristic curves were calculated, and area under the curve (AUC) values were reported.

Results: From 277 differentially expressed genes between asthma inflammatory phenotypes, we identified 23 genes that showed highly significant differential expression in both the discovery and validation populations. A signature of 6 genes, including Charcot-Leydon crystal protein (*CLC*); carboxypeptidase A3 (*CPA3*); deoxyribonuclease I-like 3 (*DNASE1L3*); IL-1 β (*IL1B*); alkaline phosphatase, tissue-nonspecific isozyme (*ALPL*); and chemokine (C-X-C motif) receptor 2 (*CXCR2*), was reproducible and could significantly

($P < .0001$) discriminate eosinophilic asthma from other phenotypes, including patients with noneosinophilic asthma (AUC, 89.6%), paucigranulocytic asthma (AUC, 92.6%), or neutrophilic asthma (AUC, 91.4%) and healthy control subjects (AUC, 97.6%), as well as discriminating patients with neutrophilic asthma from those with paucigranulocytic asthma (AUC, 85.7%) and healthy control subjects (AUC, 90.8). The 6-gene signature predicted ICS response ($>12\%$ change in FEV₁; AUC, 91.5%). ICS treatment reduced the expression of *CLC*, *CPA3*, and *DNASE1L3* in patients with eosinophilic asthma.

Conclusions: A sputum gene expression signature of 6 biomarkers reproducibly and significantly discriminates inflammatory phenotypes of asthma and predicts ICS treatment response. This signature has the potential to become a useful diagnostic tool to assist in the clinical diagnosis and management of asthma. (J Allergy Clin Immunol 2014;133:997-1007.)

Key words: Asthma, inflammatory phenotypes, gene expression, induced sputum, neutrophils, eosinophils

The airway inflammation that characterizes asthma is heterogeneous and results in the development of different clinical phenotypes of the disease triggered by multiple gene-environment interactions with a varying response to current treatments.¹ Improvements in asthma diagnosis and management require a better understanding of disease heterogeneity and the development of tools for phenotype recognition to guide personalized treatment approaches.² Development of asthma biomarkers are needed and will provide information that enables asthma diagnosis, indication of treatment responsiveness, and monitoring of asthma control. A useful biomarker will represent underlying pathologic mechanisms of disease, marking heterogeneity and thus allowing targeted therapy through personalized medicine.³

Asthma heterogeneity is perhaps best reported and recognized in terms of inflammatory phenotypes of asthma whereby the eosinophilic asthma (EA), neutrophilic asthma (NA), and paucigranulocytic asthma (PGA) phenotypes are identified from the proportions of induced sputum granulocytes.⁴ Molecular phenotyping studies that use unbiased approaches and direct airway samples are needed⁵ because these studies will link gene activity to potential mechanistic pathways and uncover potential biomarkers and have been transformative in the treatment of other diseases, such as cancer. We have recently defined 3 transcriptional asthma phenotypes using unbiased hierarchical clustering of sputum gene expression profiles and showed that these 3 phenotypes are related to eosinophilic, neutrophilic, and paucigranulocytic patterns of airway inflammation.⁶ T_H2-high

From ^athe Priority Research Centre for Asthma and Respiratory Diseases and ^cthe Priority Research Centre of Information Based Medicine, Hunter Medical Research Institute, University of Newcastle, New Lambton Heights; ^bthe Department of Respiratory and Sleep Medicine, John Hunter Hospital, New Lambton Heights; and ^dthe Dunedin School of Medicine, University of Otago, Dunedin.

Supported by the John Hunter Hospital Charitable Trust, University of Newcastle. The Hunter Medical Research Institute provided research grants to support the laboratory testing, and NHMRC and Asthma CRC supported the patient recruitment and sample collection.

Disclosure of potential conflict of interest: K. J. Baines has received research support from the Hunter Medical Research Institute Project Grant, the NHMRC Project Grant, and the John Hunter Hospital Charitable Trust. J. L. Simpson has received research support from the Hunter Medical Research Institute and John Hunter Hospital Charitable Trust and the National Health and Medical Research Council, has received travel support from the University of Newcastle, and is employed by the Australian Respiratory Council and the National Health and Medical Research Council Centre for Respiratory and Sleep Medicine. L. G. Wood has received research support from NSA, LLC and has received payment for lectures from GlaxoSmithKline. D. R. Taylor has received consulting fees and travel support from Aerocrine. P. G. Gibson has received grants from the Australian government NHMRC and has received payment for lectures from AstraZeneca, GlaxoSmithKline, and Boehringer Ingelheim. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication August 23, 2013; revised December 10, 2013; accepted for publication December 17, 2013.

Available online February 28, 2014.

Corresponding author: Katherine J. Baines, PhD, Level 2 West, Hunter Medical Research Institute, Locked Bag 1000, New Lambton, NSW 2305, Australia. E-mail: katherine.baines@newcastle.edu.au.

0091-6749/\$36.00

© 2014 American Academy of Allergy, Asthma & Immunology
<http://dx.doi.org/10.1016/j.jaci.2013.12.1091>

Abbreviations used

ACQ:	Asthma Control Questionnaire
ALPL:	Alkaline phosphatase, tissue-nonspecific isozyme
AUC:	Area under the curve
BIRC3:	Baculoviral IAP repeat containing 3
BPIFB1:	BPI fold containing family B, member 1
CAMP:	Cathelicidin antimicrobial peptide
CCR7:	Chemokine (C-C motif) receptor 7
CEBPE:	CCAAT/enhancer binding protein (C/EBP), epsilon
CLC:	Charcot-Leydon crystal protein
CPA3:	Carboxypeptidase A3
CXCR1:	Chemokine (C-X-C motif) receptor 1
CXCR2:	Chemokine (C-X-C motif) receptor 2
DEFB1:	β -Defensin 1
Δ ct:	Change in cycle threshold
DNASE1L3:	Deoxyribonuclease I-like 3
EA:	Eosinophilic asthma
FENO:	Fraction of exhaled nitric oxide
FPR2:	Formyl peptide receptor 2
GBP1:	Guanylate binding protein 1, interferon inducible
GCHFR:	GTP cyclohydrolase I feedback regulator
HDC:	Histidine decarboxylase
ICS:	Inhaled corticosteroid
IL18RAP:	IL-18 receptor accessory protein
IL1B:	IL-1 β
IL1R2:	IL-1 receptor type II
IL1RN:	IL-1 receptor antagonist
IRAK2:	IL-1 receptor-associated kinase 2
ISG20:	Interferon-stimulated exonuclease gene 20kDa
MGA:	Mixed granulocytic asthma
NA:	Neutrophilic asthma
NAMPT:	Nicotinamide phosphoribosyltransferase
NFKB2:	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)
NFKBIZ:	Nuclear factor of kappa light polypeptide gene enhancer in B-cell inhibitor, zeta
NLRP3:	NLR family, pyrin domain containing 3
ORM1:	Orosomucoid 1
PELI1:	Pellino E3 ubiquitin protein ligase 1
PGA:	Paucigranulocytic asthma
PI3:	Peptidase inhibitor 3, skin derived
PLAU:	Plasminogen activator, urokinase
qPCR:	Quantitative real-time PCR
TFF3:	Trefoil factor 3 (intestinal)
TNFAIP3:	TNF- α -induced protein 3
TNFRSF1B:	TND receptor superfamily, member 1B
TNFSF14:	TND (ligand) superfamily, member 14
VNN2:	Vanin 2

and T_H2 -low phenotypes of asthma have been discovered by using a similar approach.⁷ The T_H2 -high phenotype benefited most from inhaled corticosteroid (ICS) therapies, demonstrating the importance of phenotyping to guide treatment strategies for successful outcomes. Other studies have used defined protein arrays to assay bronchoalveolar lavage fluid⁸ and induced sputum.⁹ This study aimed to develop and validate a robust molecular signature for asthma phenotyping and to understand the clinical utility of this by investigating ICS treatment response.

METHODS**Study design**

We conducted 4 studies: a discovery study ($n = 47$ asthmatic patients), a clinical validation study ($n = 59$ asthmatic patients), a robustness and

reproducibility study ($n = 20$ healthy control subjects, $n = 30$ asthmatic patients having 2 visits spaced a month apart, and $n = 12$ subjects with mixed granulocytic asthma [MGA]), and an ICS response study ($n = 71$ asthmatic patients before and after ICS treatment). The study workflow and design details are summarized in Fig 1. The discovery study assessed differential gene expression in the sputum of 47 patients with stable nonsmoking asthma by using gene expression profiles generated from a previous study.⁶ Selected gene biomarkers underwent technical validation by means of quantitative real-time PCR (qPCR) and were further validated in a separate population of 59 asthmatic patients in a clinical validation study. A biomarker signature of 6 genes was identified and tested for robustness and reproducibility in another study of 20 control subjects and 30 asthmatic patients (2 visits, 1 month apart). In the fourth study ICS treatment response was assessed in 71 asthmatic patients with predominately EA or noneosinophilic asthma (NEA; neutrophilic inflammation was rare in this population) assessed before and after treatment (1000 μ g/d fluticasone for 28 days), as previously published.¹⁰ The studies were approved by the relevant institutional ethics committees: the University of Newcastle Research Ethics Committee and the Lower South Regional Ethics Committee of New Zealand. The ICS response study (ACTRN12606000531516) was conducted to obtain run-in data for 2 larger clinical trials in which establishing the inflammatory phenotype and steroid responsiveness were prerequisites.

Study population

For the discovery and clinical validation and reproducibility studies, adults with stable asthma were recruited from the John Hunter Hospital Ambulatory Care Clinic, New South Wales, Australia. Healthy volunteers were recruited by advertisement. For the ICS response study, adults with stable persistent asthma were enrolled at the Respiratory Research Unit, Dunedin School of Medicine, University of Otago, New Zealand. Asthma was diagnosed according to American Thoracic Society guidelines based on current (past 12 months) episodic symptoms, doctor's diagnosis, and confirmed evidence of airway hyperresponsiveness to hypertonic saline or methacholine or a 12% or greater improvement in FEV₁ after bronchodilator.¹⁰ Exclusion criteria included recent (past month) respiratory tract infection, asthma exacerbation, unstable asthma or change in therapy, and current smoking. Asthma inflammatory phenotype was assigned based on a sputum eosinophil cutoff of greater than 2% and a sputum neutrophil cutoff of greater than 61%.^{4,10}

Sputum induction and analysis

Spirometry (KoKo PD Instrumentation, Louisville, Ky) and sputum induction with hypertonic saline (4.5%) were performed, as previously described.⁴ For inflammatory cell counts, selected sputum (discovery and validation studies) or whole sputum (ICS response study) was dispersed with dithiothreitol, and total cell count viability was determined. Cytospin preparations were prepared and stained (May-Grunwald Giemsa), and a differential cell count was obtained from 400 nonsquamous cells. For gene expression, Buffer RLT (Qiagen, Hilden, Germany) was added to 100 μ L of selected sputum (discovery and validation populations) or the sputum cell pellet (ICS response population) and stored at -80°C until RNA extraction.

Whole-genome gene expression microarrays

RNA was extracted from induced sputum samples, amplified, and hybridized to Illumina HumanRef-8 V2 Expression BeadChips (Illumina, San Diego, Calif), as previously described.⁶ Data were uploaded to GEO Omnibus (reference no. GSE45111).

Ranking of differentially expressed genes as potential biomarkers

Differentially expressed genes caused by asthma inflammatory phenotype were selected based on a number of factors, including level of significance (ANOVA-adjusted P value), degree of fold change, abundance of mRNA expression in sputum, biological plausibility as to the candidate's potential

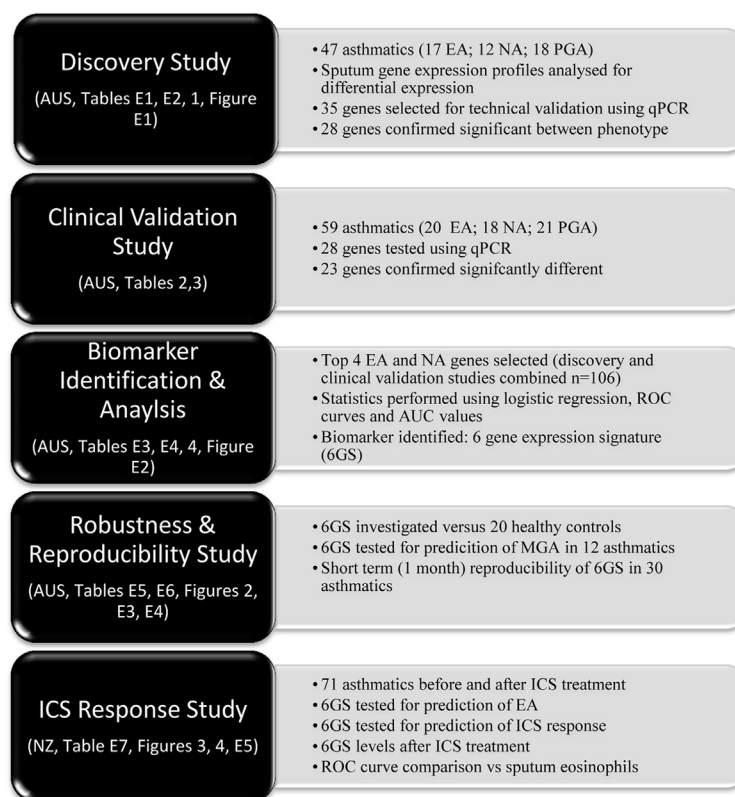


FIG 1. Schematic diagram documenting the biomarker identification and evaluation process undertaken. AUS, Australian population; NZ, New Zealand population; ROC, receiver operating characteristic; 6GS, 6-gene expression signature.

involvement in asthma pathogenesis, potential importance and relevance as a biomarker and potential therapeutic target, and commercial availability of a suitable TaqMan primer/probe set (Applied Biosystems, Foster City, Calif). Priority was given to those genes that could discriminate between EA and NA and were involved in previously identified pathways associated with NA.

qPCR

Sputum RNA (200 ng) was reverse transcribed to cDNA and used to detect gene expression by using standard TaqMan methods, as previously described.⁶ Statistical analysis was performed on the change in cycle threshold (ΔC_t) between the target gene and β -actin, the housekeeping gene. Fold change results were calculated by using $2^{-\Delta\Delta C_t}$ relative to both the housekeeping gene (β -actin) and the mean of the healthy control group.

Statistical analysis

Data were analyzed with STATA 11 (StataCorp, College Station, Tex) and reported as means (SDs) for normally distributed data and medians (quartile 1-quartile 3) for nonparametric data and analyzed by using ANOVA for parametric data and the Kruskal-Wallis test for nonparametric data. A *P* value of less than .05 was considered significant.

Microarray data were exported by using Genome Studio (Illumina) and analyzed with GeneSpring GX11 (Agilent Technologies, Santa Clara, Calif). Data were log transformed, normalized, and baseline converted to the median of all samples. Differential gene expression was determined by using ANOVA, with Tukey *post hoc* testing (*P* < .05 adjusted for multiple comparisons by using the Benjamini-Hochberg method) and a fold change of greater than 2. Technical variation, including batch effects, was assessed by using principal component analysis and corrected for by using regression analysis.

Logistic regression was used (STATA 11) to calculate the predicted value of a subject having the disease phenotype based on their level of expression of a particular gene (simple logistic regression) or combination of genes (multiple logistic regression), as described in detail in the **Methods** section in this article's Online Repository at www.jacionline.org. Receiver operating characteristic curves were generated, and the area under the curve (AUC) was calculated. Bland-Altman plots and intraclass correlation coefficients were used to determine reproducibility.

RESULTS

The study workflow is summarized in **Fig 1**.

Discovery study: Differentially expressed genes between asthma inflammatory phenotypes

Whole-genome gene expression profiles (22,184 entities) were generated from induced sputum samples from 47 asthmatic participants (*n* = 17 patients with EA, *n* = 12 patients with NA, and *n* = 18 patients with PGA), as previously reported.⁶ The clinical details regarding this population have been previously published⁶ and are provided again in **Table E1** in this article's Online Repository at www.jacionline.org. The microarray data analysis is provided in detail in the **Results** section in this article's Online Repository at www.jacionline.org. There were 277 genes differentially expressed between asthma inflammatory phenotypes (*P* < .05, ANOVA, adjusted for multiple comparisons by using Benjamini-Hochberg false discovery rate and fold change >2 in 1 of 3 phenotype comparisons). **Table E2** in this article's Online Repository at www.jacionline.org details the 35

TABLE I. Differential expression of biomarker candidates in both study groups determined by using qPCR¶

Gene symbol	Gene name	Discovery asthma population			
		EA	NA	PGA	P value
No.		17	12	16	
<i>ALPL</i>	Alkaline phosphatase, tissue-nonspecific isozyme	5.82 (1.92)	4.28 (2.32)*	7.64 (2.39)	.001
<i>BIRC3</i>	Baculoviral IAP repeat containing 3	5.42 (2.46)*	5.52 (1.83)‡	7.70 (1.56)	.004
<i>BPIFB1</i>	BPI fold containing family B, member 1	3.25 (1.95)*§	5.32 (1.52)	5.47 (2.12)	.003
<i>CAMP</i>	Cathelicidin antimicrobial peptide	7.93 (1.80)	6.85 (1.63)	7.10 (1.56)	.189
<i>CCR7</i>	Chemokine (C-C motif) receptor 7	5.56 (1.68)*	6.54 (1.79)	8.04 (2.04)	.002
<i>CEBPE</i>	CCAAT/enhancer binding protein (C/EBP), epsilon	8.21 (2.41)*§	10.36 (1.92)*	12.75 (1.44)	<.001
<i>CLC</i>	Charcot-Leydon crystal protein	6.24 (3.01)*§	9.02 (2.10)*	11.97 (2.02)	<.001
<i>CPA3</i>	Carboxypeptidase A3	7.89 (1.96)*	11.12 (2.54)	12.79 (2.84)	<.001
<i>CXCR1</i>	Chemokine (C-X-C motif) receptor 1	9.30 (2.29)	8.12 (2.69)‡	10.38 (2.07)	.050
<i>CXCR2</i>	Chemokine (C-X-C motif) receptor 2	5.52 (1.35)	4.41 (1.68)*	6.45 (1.38)	.003
<i>DEFB1</i>	β-Defensin 1	8.14 (1.62)*	7.76 (2.19)	6.29 (1.29)	.008
<i>DNASE1L3</i>	Deoxyribonuclease I-like 3	10.06 (2.68)*	11.32 (1.61)‡	14.01 (2.27)	<.001
<i>FPR2</i>	Formyl peptide receptor 2	6.38 (2.09)	6.00 (2.68)	7.64 (2.07)	.131
<i>GBP1</i>	Guanylate binding protein 1, interferon inducible	5.08 (1.54)	4.18 (1.74)	5.28 (1.53)	.183
<i>GCHFR</i>	GTP cyclohydrolase I feedback regulator	3.78 (1.36)	3.93 (1.74)	3.10 (0.51)	.173
<i>HDC</i>	Histidine decarboxylase	10.75 (2.58)*	12.58 (2.14)‡	14.82 (1.59)	<.001
<i>IL18RAP</i>	IL-18 receptor accessory protein	8.34 (1.52)	7.54 (1.31)	9.49 (2.85)	.051
<i>IL1B</i>	IL-1β	3.68 (2.38)	1.87 (2.38)*	5.04 (1.67)	.002
<i>IL1R2</i>	IL-1 receptor type II	5.97 (1.74)‡	5.05 (1.71)*	7.82 (1.99)	<.001
<i>IL1RN</i>	IL-1 receptor antagonist	4.80 (2.08)‡	4.05 (2.04)*	6.82 (2.64)	.006
<i>IRAK2</i>	IL-1 receptor–associated kinase 2	4.24 (1.64)	3.42 (3.58)‡	6.04 (1.80)	.015
<i>ISG20</i>	Interferon stimulated exonuclease gene 20kDa	4.55 (1.70)	3.60 (1.51)*	5.84 (1.98)	.006
<i>NAMPT</i>	Nicotinamide phosphoribosyltransferase	3.81 (2.60)	3.28 (1.76)	4.14 (1.76)	.570
<i>NFKB2</i>	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	4.35 (1.63)‡	3.92 (1.89)*	6.08 (1.73)	.004
<i>NFKBIZ</i>	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	4.23 (1.52)‡	3.75 (1.88)*	5.73 (1.57)	.006
<i>NLRP3</i>	NLR family, pyrin domain containing 3	5.60 (1.79)*	5.61 (2.19)*	8.03 (1.70)	<.001
<i>ORM1</i>	Orosomucoid 1	10.08 (3.15)	8.34 (3.29)*	12.23 (2.85)	.007
<i>PELI1</i>	Pellino E3 ubiquitin protein ligase 1	5.17 (1.92)	4.04 (2.03)*	6.55 (1.61)	.004
<i>PI3</i>	Peptidase inhibitor 3	4.95 (3.34)	2.84 (2.91)*	7.17 (2.66)	.002
<i>PLAU</i>	Plasminogen activator, urokinase	4.34 (2.39)	3.85 (1.66)	4.96 (1.73)	.349
<i>TFF3</i>	Trefoil factor 3	4.62 (2.18)*§	7.06 (1.06)	7.41 (2.57)	<.001
<i>TNFAIP3</i>	TNF-α induced protein 3	3.22 (1.72)	2.43 (1.97)*	4.68 (1.66)	.005
<i>TNFRSF1B</i>	TNF receptor superfamily, member 1B	4.04 (1.76)‡	3.48 (2.05)*	5.73 (1.65)	.004
<i>TNFSF14</i>	TNF (ligand) superfamily, member 14	5.58 (1.85)‡	5.46 (2.03)‡	7.59 (1.77)	.004
<i>VNN2</i>	Vanin 2	6.19 (2.20)	4.75 (2.53)*	7.82 (2.60)	.007

**P* < .01, Bonferroni *post hoc* test versus patients with PGA.‡*P* < .05, Bonferroni *post hoc* test versus patients with PGA.§*P* < .05, Bonferroni *post hoc* test versus patients with NA.||*P* < .01, Bonferroni *post hoc* test versus patients with NA.

¶Data are expressed as the ΔCt compared with the housekeeping gene β-actin (mean [SD]). A lower ΔCt corresponds to a stronger expression of the target gene.

differentially expressed genes selected as potential biomarker candidates. Fig E1 in this article's Online Repository at www.jacionline.org shows a Venn diagram containing the differentially expressed genes between phenotypes. Variation explained by batch effects were found to be associated with principal component 2 and therefore were corrected for using regression analysis. All 35 selected genes remained significantly different between phenotypes after correction for principal component 2 (see Table E2).

qPCR was performed to assess the performance of the 35 selected biomarker candidates identified by using microarray analysis. The qPCR results for all genes were significantly correlated with the microarray results (see Table E2). There were 28 genes significantly differentially expressed between asthma inflammatory phenotypes (Table I), with 9 of these genes reaching a significant *P* value of .001 or less. Expression levels of baculoviral IAP repeat containing 3 (*BIRC3*); BPI fold containing

family B, member 1 (*BPIFB1*); chemokine (C-C motif) receptor 7 (*CCR7*); CCAAT/enhancer binding protein (C/EBP), epsilon (*CEBPE*); Charcot-Leydon crystal protein (*CLC*); carboxypeptidase A3 (*CPA3*); deoxyribonuclease I-like 3 (*DNASE1L3*); histidine decarboxylase (*HDC*); and trefoil factor 3 (intestinal) (*TFF3*) were highest in patients with EA. Expression levels of alkaline phosphatase, tissue-nonspecific isozyme (*ALPL*); chemokine (C-X-C motif) receptor 1 (*CXCR1*); chemokine (C-X-C motif) receptor 2 (*CXCR2*); IL-1β (*IL1B*); IL-1 receptor type II (*IL1R2*); IL-1 receptor antagonist (*IL1RN*); IL-1 receptor–associated kinase 2 (*IRAK2*); Interferon-stimulated exonuclease gene 20kDa (*ISG20*); nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100) (*NFKB2*); nuclear factor of kappa light polypeptide gene enhancer in B-cell inhibitor, zeta (*NFKBIZ*); NLR family, pyrin domain containing 3 (*NLRP3*); orosomucoid 1 (*ORM1*); pellino E3 ubiquitin protein ligase 1 (*PELI1*); peptidase inhibitor 3, skin derived (*PI3*);

TABLE II. Clinical characteristics of the participants in the clinical validation population

	Healthy control subjects	Asthma phenotype			P value
		EA	NA	PGA	
No. (% total)	20	20	18	21	
Age (y), mean (SD)	43 (19)	54 (17)	59 (12)*	60 (13)*	.003
Sex, M/F	8/12	8/12	11/7	13/8	.300
Atopy, no. (%)	8 (40)	13 (65)	10 (56)	15 (71)	.198
FEV ₁ (% predicted), mean (SD)	99 (12)	73 (17)*	71 (23)*	87 (22)	<.001
FEV ₁ /FVC ratio (%), mean (SD)	79 (7)	65 (10)*	65 (11)*	69 (7)*	<.001
FENO (ppb), median (Q1-Q3)	16.84 (14.7-236)	49.5 (20.5-71.0)†‡§	18.6 (12.6-26.2)	17.0 (14.1-27.6)	.004
Smoking, exsmoker/never smoker	6/14	13/7	7/11	11/10	.213
Pack years, median (Q1-Q3)	19.5 (29-38.8)	0.6 (0.4-0.7)†§	24.0 (3.6-32.0)‡	10.0 (2.0-25.0)	.004
Asthma control score, median (Q1-Q3)	N/A	1.1 (0.8-1.7)‡	1.1 (0.8-1.5)‡	0.4 (0.3-0.8)	.005
ICS use, no/yes	N/A	3/17	3/15	3/18	.978
ICS dose (μg daily BDP equivalent), median (Q1-Q3)	N/A	1000 (500-2000)	2000 (800-2000)	1000 (1000-2000)	.523
Total cell count × 10 ⁶ /mL, median (Q1-Q3)	3.9 (1.9-5.4)	4.4 (2.7-4.8)	10.8 (6.7-15.5)†‡	2.6 (1.8-4.1)	<.001
Viability, median (Q1-Q3)	75.9 (67.6-85.0)	74.6 (64.6-84.6)	91.4 (87.1-94.6)†‡	67.3 (56.9-75.9)	<.001
Neutrophils (%), median (Q1-Q3)	28.4 (12.1-50.34)	27.4 (18.3-38.8)	76.1 (71.5-87.0)†‡	26.3 (17.0-48.8)	<.001
Eosinophils (%), median (Q1-Q3)	0.3 (0.1-0.8)	11.9 (4.4-20.6)†‡§	0.8 (0.0-1.5)	0.3 (0.0-0.5)	<.001
Macrophages (%), median (Q1-Q3)	68.0 (41.6-74.0)	48.8 (40.3-61.0)	20.8 (11.8-24.3)†‡	67.0 (50.3-76.5)	<.001
Lymphocytes (%), median (Q1-Q3)	1.3 (0.4-3.1)	0.3 (0.3-2.0)	0.5 (0.0-0.8)†	0.3 (0.0-1.0)†	.035
Columnar epithelial cells (%), median (Q1-Q3)	1.9 (0.6-6.4)	1.5 (0.8-2.9)	0.6 (0.0-1.8)	2.0 (0.8-4.0)	.112
Squamous cells (%), median (Q1-Q3)	4.2 (0.8-6.5)	4.1 (2.0-6.1)	1.5 (0.7-2.7)	2.9 (2.0-9.3)	.215

BDP, Beclomethasone dipropionate; FVC, forced vital capacity; N/A, not applicable; Q1-Q3, quartile 1 to quartile 3.

Bonferroni *post hoc* test: **P* < .05 versus healthy control subjects; Kruskal Wallis *post hoc* test; †*P* < .01 versus healthy control subjects; ‡*P* < .01 versus patients with PGA; §*P* < .01 versus patients with NA; and ||*P* < .01 versus patients with EA.

TNF- α -induced protein 3 (*TNFAIP3*); TND receptor superfamily, member 1B (*TNFRSF1B*); and vanin 2 (*VNN2*) were highest in patients with NA. Expression levels of β -defensin 1 (*DEFB1*) was highest in patients with PGA, and expression of TNF (ligand) superfamily, member 14 (*TNFSF14*), were similarly increased in both patients with EA and those with NA compared with those seen in patients with PGA. Expression levels of cathelicidin antimicrobial peptide (*CAMP*); formyl peptide receptor 2 (*FPR2*); guanylate binding protein 1, interferon inducible (*GBP1*); GTP cyclohydrolase I feedback regulator (*GCHFR*); IL-18 receptor accessory protein (*IL18RAP*); nicotinamide phosphoribosyltransferase (*NAMPT*); and plasminogen activator, urokinase (*PLAU*), although correlated with the array results, did not reach significance between inflammatory phenotypes and therefore were not tested in the clinical validation population.

Clinical validation study

We tested the 28 confirmed candidate genes in a clinical validation population of 59 asthmatic patients (Table II). There were 26 genes that remained significantly different between inflammatory phenotypes of asthma (Table III), with 15 of these reaching a significant *P* value of .001 or less. Expression levels of *BPIFB1*, *CLC*, *CPA3*, *DNASE1L3*, and *HDC* were highest in patients with EA, whereas expression levels of *ALPL*, *CXCR1*, *CXCR2*, *IL1B*, *IL1R2*, *IL1RN*, *IRAK2*, *ISG20*, *NFKB2*, *NFKBIZ*, *NLRP3*, *ORM1*, *PELI1*, *PI3*, *TNFAIP3*, *TNFRSF1B*, *TNFSF14*, and *VNN2* were highest in patients with NA. Although expression levels of *BIRC3*, *DEFB1*, and *TFF3* reached significance overall, a *post hoc* analysis of individual phenotypes was not significant. *CCR7* and *CEBPE* did not reach significance in the clinical validation population. This left 23 genes consistently different between inflammatory phenotypes in both the discovery and validation populations.

Biomarker identification and analysis

The top 4 EA and NA genes were selected for testing of their biomarker potential by using logistic regression, receiver operating characteristic curves, and AUC values in the discovery and clinical validation studies combined. *CLC*, *CPA3*, *DNASE1L3*, and *HDC* were further investigated for their potential to predict EA. *CPA3* was the best individual discriminator for EA, followed by *CLC*, *DNASE1L3*, and *HDC* (see Table E3 in this article's Online Repository at www.jacionline.org). *ALPL*, *IL1B*, *CXCR2*, and *PI3* were further investigated for their potential to predict NA. *IL1B* was the best individual discriminator for NA, followed by *ALPL*, *CXCR2*, and *PI3* (see Table E4 in this article's Online Repository at www.jacionline.org).

The combination of 6 genes, including *CLC*, *CPA3*, *DNASE1L3*, *IL1B*, *ALPL*, and *CXCR2*, provided the strongest performance (Table IV). The 6-gene signature was able to discriminate patients with EA from patients with NEA (AUC, 89.6%; *P* < .0001), patients with NA (AUC, 91.4%; *P* < .0001; see Fig E2, A, in this article's Online Repository at www.jacionline.org), and patients with PGA (AUC, 92.6%; *P* < .0001; see Fig E2, B). The 6-gene signature also distinguished patients with NA from patients with PGA (AUC, 85.7%; *P* < .0001; see Fig E2, C). Univariate regression analysis determined that expression levels of the 6 genes were not confounded by age, sex, or smoking history.

Robustness and reproducibility study

This study assessed the expression of the 6-gene signature in healthy control subjects and patients with MGA and the short-term (1-month) reproducibility. The 6-gene expression biomarker signature was able to discriminate asthmatic patients from healthy control subjects (AUC, 84.4%; *P* = .0001;

TABLE III. Clinical validation differential expression of biomarker candidates in both study groups determined by using qPCR¶

Gene symbol	Gene name	Validation asthma population			
		EA	NA	PGA	P value
No.		20	18	21	
<i>ALPL</i>	Alkaline phosphatase, tissue-nonspecific isozyme	5.67 (1.72)	3.42 (1.63)*†	5.73 (1.75)	<.001
<i>BIRC3</i>	Baculoviral IAP repeat containing 3	4.71 (1.85)	4.71 (1.74)	5.86 (1.18)	.034
<i>BPIFB1</i>	BPI fold containing family B, member 1	4.20 (2.11)‡	5.69 (2.43)	5.80 (1.49)	.021
<i>CCR7</i>	Chemokine (C-C motif) receptor 7	6.45 (2.18)	6.04 (2.09)	7.33 (2.39)	.209
<i>CEBPE</i>	CCAAT/enhancer binding protein (C/EBP), epsilon	10.01 (2.40)	10.71 (1.77)	11.56 (1.84)	.052
<i>CLC</i>	Charcot-Leydon crystal protein	6.99 (3.75)*	9.21 (2.75)	10.55 (1.74)	<.001
<i>CPA3</i>	Carboxypeptidase A3	8.55 (2.07)*	10.90 (2.39)	11.81 (2.00)	<.001
<i>CXCR1</i>	Chemokine (C-X-C motif) receptor 1	7.92 (1.49)	6.57 (1.23)*†	8.82 (1.18)	<.001
<i>CXCR2</i>	Chemokine (C-X-C motif) receptor 2	4.08 (1.71)	2.45 (2.01)*†	4.66 (1.39)	<.001
<i>DEFB1</i>	β-defensin 1	6.95 (2.55)	8.65 (2.37)	7.41 (1.25)	.046
<i>DNASE1L3</i>	Deoxyribonuclease I-like 3	9.98 (2.24)*	11.37 (1.76)	12.18 (2.30)	.005
<i>HDC</i>	Histidine decarboxylase	11.89 (1.60)*	12.59 (1.88)‡	14.23 (2.46)	.002
<i>IL1B</i>	IL-1β	2.47 (1.75)	0.78 (1.78)*†	2.82 (1.26)	<.001
<i>IL1R2</i>	IL-1 receptor type II	6.00 (2.19)	3.98 (1.68)*†	6.26 (1.78)	<.001
<i>IL1RN</i>	IL-1 receptor antagonist	4.16 (1.32)§	3.04 (1.83)*	4.57 (1.04)	.004
<i>IRAK2</i>	IL-1 receptor–associated kinase 2	4.55 (1.57)	2.37 (2.36)*†	4.58 (1.72)	<.001
<i>ISG20</i>	Interferon stimulated exonuclease gene 20kDa	5.51 (1.55)	3.55 (1.68)*†	5.66 (1.57)	<.001
<i>NFKB2</i>	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	4.92 (1.00)	3.23 (1.64)*†	4.99 (1.33)	<.001
<i>NFKBIZ</i>	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	4.07 (1.10)	2.62 (1.84)*†	4.68 (1.14)	<.001
<i>NLRP3</i>	NLR family, pyrin domain containing 3	5.78 (1.62)	4.43 (2.13)*	6.43 (1.97)	.007
<i>ORM1</i>	Orosomucoid 1	8.76 (2.88)‡	6.30 (3.16)*	10.16 (2.78)	<.001
<i>PELI1</i>	Pellino E3 ubiquitin protein ligase 1	4.46 (1.92)	3.65 (1.53)‡	5.17 (1.24)	.013
<i>PI3</i>	Peptidase inhibitor 3	4.68 (2.89)	2.02 (2.54)‡†	4.50 (2.36)	.005
<i>TFF3</i>	Trefoil factor 3	5.47 (2.22)	7.16 (2.52)	6.94 (1.80)	.034
<i>TNFAIP3</i>	TNF-α induced protein 3	2.98 (1.47)	1.83 (2.16)‡	3.39 (1.32)	.016
<i>TNFRSF1B</i>	TNF receptor superfamily, member 1B	4.71 (1.10)	3.12 (1.51)*†	4.92 (1.23)	<.001
<i>TNFSF14</i>	TNF (ligand) superfamily, member 14	6.33 (1.60)	3.85 (2.53)*†	6.43 (2.13)	<.001
<i>VNN2</i>	Vanin 2	4.73 (1.70)	3.19 (1.48)*†	5.13 (1.00)	<.001

**P* < .01, Bonferroni *post hoc* test versus patients with PGA.†*P* < .01, Bonferroni *post hoc* test versus patients with EA.‡*P* < .05, Bonferroni *post hoc* test versus patients with PGA.§*P* < .05, Bonferroni *post hoc* test versus patients with NA.||*P* < .01, Bonferroni *post hoc* test versus patients with NA.

¶Data are expressed as ΔCt compared with the housekeeping gene β-actin (mean [SD]). A lower ΔCt corresponds to a stronger expression of the target gene.

see Table E5 in this article's Online Repository at www.jacionline.org), as well as patients with EA (AUC, 97.4%; *P* < .0001; see Fig E3, A, in this article's Online Repository at www.jacionline.org) and patients with NA (AUC, 91.8%; *P* < .0001; see Fig E3, B) from healthy control subjects. Expression levels of *CLC*, *CPA3*, and *DNASE1L3* were significantly increased in patients with EA, and *IL1B*, *ALPL*, and *CXCR2* were significantly increased in patients with NA (Fig 2).

Patients with MGA were initially excluded because of low numbers (*n* = 9 in the discovery study and *n* = 3 in the clinical validation study). However, the ability of the 6-gene signature to identify patients with MGA was investigated by applying the logistic regression equation (detailed in the Methods section in this article's Online Repository) with the coefficients and constant results (Table IV) for patients with EA versus patients with PGA to identify eosinophilic inflammation with a predicted value cut point of 0.369 (sensitivity, 91.9%; specificity, 75.7%) and NA versus PGA to identify neutrophilic inflammation with a predicted value cut point of 0.538 (sensitivity, 73.3%; specificity, 83.8%). Eosinophilic inflammation was identified in 8 (67%) subjects, neutrophilic inflammation was identified in 9 (75%) subjects, and both types of inflammation were demonstrated in 6 (50%) subjects (see Table E6 in this article's Online Repository at www.jacionline.org).

Sputum gene expression of the 6 biomarkers was measured in 30 subjects (*n* = 9 patients with EA, *n* = 7 patients with NA, and *n* = 14 patients with PGA) on 2 occasions an average of 30 days apart. The bias of measurement was small, with equal scatter for all genes (see Fig E4 in this article's Online Repository at www.jacionline.org). Intraclass correlation coefficients were almost perfect for *CLC* (0.84); substantial for *CPA3* (0.71), *DNASE1L3* (0.75), and *IL1B* (0.66); and moderate for *ALPL* (0.55) and *CXCR2* (0.47).

ICS response study

The clinical details for this study have been previously published.¹⁰ The expression of the 6-gene biomarker signature was investigated before ICS treatment (1000 μg of fluticasone for 28 days), as well as their ability to predict inflammatory phenotype and ICS response. To test the ability of the 6-gene biomarker combination to predict patients with EA from patients with NEA by using the logistic regression equation (detailed in the Methods section in this article's Online Repository), we used the coefficients and constant results established for EA versus NEA (Table IV) and a cut point of 0.447 or greater, which had previously shown a sensitivity of 75.7% and specificity of 88.1%. We could successfully predict 92% (44/48) of patients

TABLE IV. Analysis of diagnostic value of *CLC*, *CPA3*, *DNASE1L3*, *IL1B*, *ALPL*, and *CXCR2* in combination for asthma inflammatory phenotype

Comparison	Logistic regression		AUC (%)	Model <i>P</i> value	Minimal false-negative results*			Minimal false-positive results*		
	Coefficient	Constant			Predicted value cut point	Sensitivity	Specificity	Predicted value cut point	Sensitivity	Specificity
EA vs NEA		6.036284	89.6	<.0001	≥0.250	89.2%	73.1%	≥0.519	70.3%	91.0%
<i>CLC</i>	−0.4359487									
<i>CPA3</i>	−0.3900831									
<i>DNASE1L3</i>	−0.1660757									
<i>IL1B</i>	0.149078									
<i>ALPL</i>	0.441533									
<i>CXCR2</i>	0.0254934									
EA vs PGA		10.58183	92.6	<.0001	≥0.361	94.6%	75.7%	≥0.782	70.3%	97.3%
<i>CLC</i>	−0.4280767									
<i>CPA3</i>	−0.7944147									
<i>DNASE1L3</i>	0.1400537									
<i>IL1B</i>	0.0302324									
<i>ALPL</i>	0.23569									
<i>CXCR2</i>	−0.314619									
EA vs NA		3.955467	91.4	<.0001	≥0.372	94.6%	73.3%	≥0.653	78.4%	86.7%
<i>CLC</i>	−0.3120517									
<i>CPA3</i>	−0.4298128									
<i>DNASE1L3</i>	−0.1324042									
<i>IL1B</i>	0.0193606									
<i>ALPL</i>	0.9090019									
<i>CXCR2</i>	0.011632									
NA vs NNA		−0.1190983	84.5	<.0001	≥0.290	86.7%	75.7%	≥0.398	73.3%	82.4%
<i>CLC</i>	0.155038									
<i>CPA3</i>	0.056661									
<i>DNASE1L3</i>	0.0978961									
<i>IL1B</i>	−0.1538187									
<i>ALPL</i>	−0.5583045									
<i>CXCR2</i>	−0.2023545									
NA vs PGA		4.370347	85.7	<.0001	≥0.330	86.7%	62.2%	≥0.543	73.3%	86.5%
<i>CLC</i>	−0.1899702									
<i>CPA3</i>	−0.0530947									
<i>DNASE1L3</i>	0.0924726									
<i>IL1B</i>	−0.2119642									
<i>ALPL</i>	−0.2869583									
<i>CXCR2</i>	−0.272358									
PGA vs NPGA		−7.871375	85.4	<.0001	≥0.244	91.9%	67.2%	≥0.458	70.3%	83.6%
<i>CLC</i>	0.3014908									
<i>CPA3</i>	0.2783498									
<i>DNASE1L3</i>	−0.0220303									
<i>IL1B</i>	0.1057387									
<i>ALPL</i>	0.0231505									
<i>CXCR2</i>	0.2439074									

NNA, Nonneutrophilic asthma; NPGA, nonpaucigranulocytic asthma.

*Minimal false-negative results correspond to the point of the receiver operating characteristic curve with the highest sensitivity (true-positive rate, which is useful for ruling disease out), whereas minimal false-positive results correspond to the point with the highest specificity (false-positive rate, which is useful for ruling disease in).

with EA and 74% (17/23) of patients with NEA. Comparison of the sputum cell counts and the gene signature methods to determine EA and NEA showed an 86% agreement (substantial) with a κ statistic of 0.67 ($P < .0001$).

Further statistical testing demonstrated that the 6-gene biomarker signature could discriminate ICS responsiveness (see Table E7 in this article's Online Repository at www.jacionline.org), particularly a greater than 12% improvement in FEV₁ (AUC, 91.5%; $P < .0001$; Fig 3, A), a 0.5-point decrease in Asthma Control Questionnaire (ACQ) scores (AUC, 85.7%; $P = .0022$; Fig 3, B), or a 40% decrease in fraction of exhaled nitric oxide (FENO) values (AUC, 87.1%; $P < .0001$; Fig 3, C). The 6-gene

signature was better than the sputum eosinophil percentage at predicting ICS responsiveness for all 3 measures (see Fig E5 in this article's Online Repository at www.jacionline.org) but significantly improved ($P = .002$) for detecting a 40% reduction in FENO values (see Fig E5, C).

Measurement of gene expression after ICS treatment showed that *CLC* and *DNASE1L3* expression levels decreased in patients with EA, whereas *CPA3* expression levels decreased in both patients with EA and patients with NEA (Fig 4). *IL1B* gene expression increased after ICS treatment of patients with EA and those with NEA, and *ALPL* expression increased after treatment of patients with EA (Fig 4).

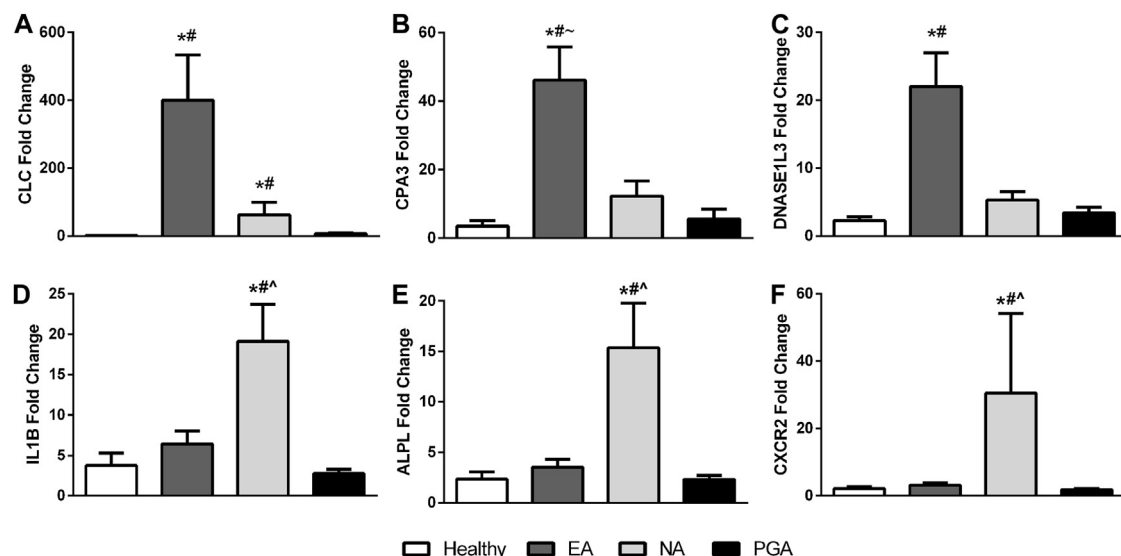


FIG 2. Relative gene expression levels of *CLC* (A), *CPA3* (B), *DNASE1L3* (C), *IL1B* (D), *ALPL* (E), and *CXCR2* (F) in induced sputum samples from patients with inflammatory phenotypes of asthma compared with healthy control subjects. Gene expression is calculated relative to that of β -actin and expressed as a fold change from the healthy control group mean. Bar graphs show means $\pm$ SEMs. * $P < .05$ versus healthy control subjects, # $P < .05$ versus PGA, ^ $P < .05$ versus EA, and ~ $P < .05$ versus NA.

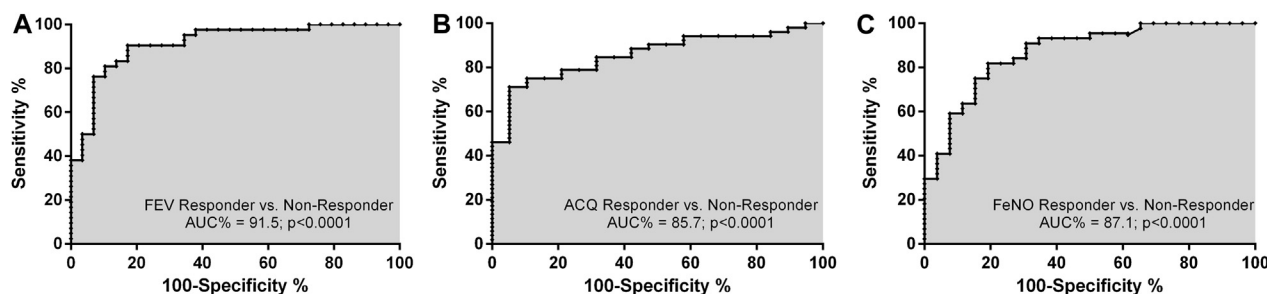


FIG 3. Receiver operating characteristic curves demonstrate that the 6-gene biomarker signature identifies ICS treatment responsiveness determined by a 12% change in FEV₁ (A), a 0.5-point decrease in ACQ score (B), and a 40% reduction in FeNO value (C). AUC, Area under the curve.

DISCUSSION

This study identified a 6-gene expression biomarker signature derived from induced sputum samples that reproducibly discriminates the inflammatory phenotypes of asthma and predicts ICS responsiveness. The top 6 genes (*CLC*, *CPA3*, *DNASE1L3*, *IL1B*, *ALPL*, and *CXCR2*) in combination were shown to accurately identify the asthma inflammatory phenotype, thereby demonstrating their potential use as a clinical biomarker. We show that *CLC*, *CPA3*, and *DNASE1L3* have increased gene expression in patients with EA, whereas *IL1B*, *ALPL*, and *CXCR2* have increased gene expression in patients with NA. This 6-gene expression signature could distinguish patients with EA from those with NEA with substantial agreement and could predict ICS responsiveness defined by a greater than 12% decrease in FEV₁, a 0.5-point decrease in ACQ scores, and a 40% reduction in FeNO values. The 6-gene biomarker signature could predict asthma and the phenotype of asthma from healthy control subjects and demonstrated predominant eosinophilia and neutrophilia in patients with MGA. Expression levels of the 6-gene signature were modified by treatment, with expression of *CLC*, *CPA3*,

and *DNASE1L3* decreased and expression of *IL1B* and *ALPL* increased after ICS treatment of patients with EA.

Several of the biomarkers have an established association with asthma. *CLC* protein, also known as galectin-10, is a major constituent of eosinophils, comprising 5% to 10% of their total protein.¹¹ *CLC* is also expressed in basophils¹² and has more recently been identified on CD4⁺CD25⁺ regulatory T cells, where it plays a role in regulatory T-cell function.¹³ Little is known regarding the function of *CLC*; however, it forms bipyramidal hexagonal crystals, which have long been associated with eosinophilic inflammation.¹⁴ Expression of *CLC* is increased in the peripheral blood of patients with aspirin-induced asthma,¹⁵ in the nasal lavage fluid of patients with aspirin-sensitive respiratory disease,¹⁶ and in the sputum supernatants of asthmatic patients with and without allergic bronchopulmonary aspergillosis.¹⁷ Chua et al¹⁷ also showed relative levels of *CLC* protein detected by means of Western blotting in the sputum supernatant were correlated with sputum eosinophil counts.

Carboxypeptidase A3 (*CPA3*) is a metalloexopeptidase that is specifically expressed by mast cells and contained in the same

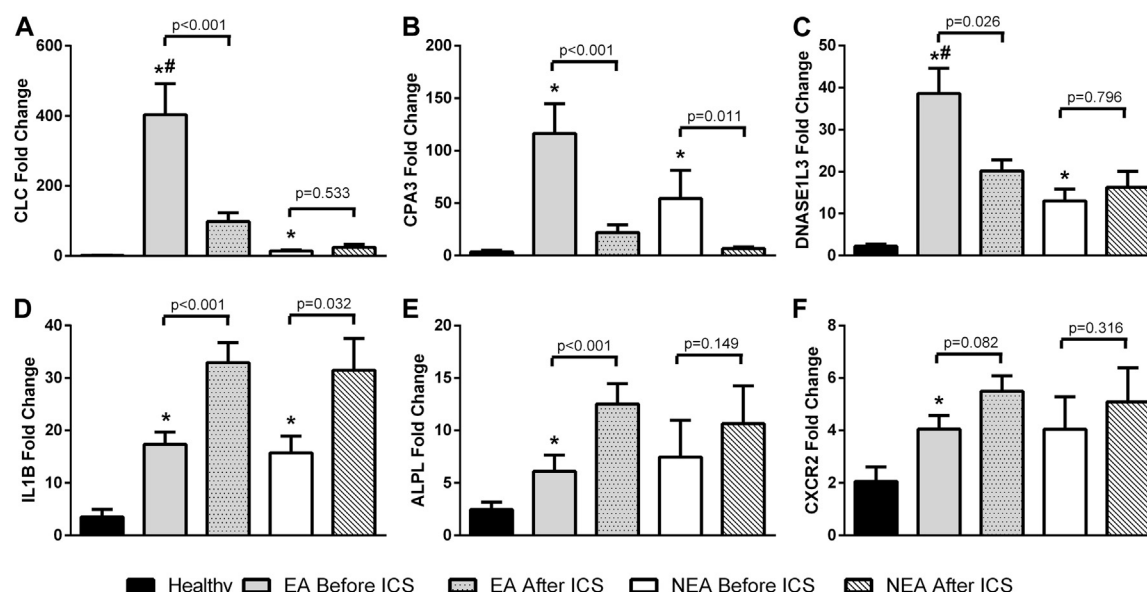


FIG 4. Relative gene expression levels of *CLC* (A), *CPA3* (B), *DNASE1L3* (C), *IL1B* (D), *ALPL* (E), and *CXCR2* (F) in induced sputum samples from subjects with EA or NEA before and after corticosteroid treatment. Gene expression is calculated relative to that of β -actin and expressed as a fold change from the healthy control group mean. * $P < .05$ versus healthy control subjects, # $P < .05$ versus patients with NEA before ICSs.

granules as tryptase and chymase.¹⁸ *CPA3* is expressed in a particular type of mast cell, positive for tryptase and chymase (MC_{TC}), which usually accounts for less than 20% of mast cells in the lung. However, this changes in patients with severe asthma, in whom the MC_{TC} cell type predominates in the submucosa and epithelium.¹⁹ Expression of mast cell proteases in epithelial brushings is increased in asthmatic patients²⁰ and specifically in patients with steroid-responsive T_H2 -high asthma.²¹

Deoxyribonucleases, including DNase1-like 3 (*DNASE1L3*), are endonucleases secreted into body fluids that mediate the breakdown of DNA during cellular apoptosis. Mutations in the gene have been implicated in the development of autoimmune diseases, such as systemic lupus erythematosus.²² Little is known regarding the function or expression of *DNASE1L3* in asthmatic patients. Interestingly, recombinant human DNase has been investigated as a potential treatment for life-threatening asthma, with limited effectiveness.²³ Our study extends the reports in the literature to demonstrate that sputum gene expression of *CLC*, *CPA3*, and *DNASE1L3* is higher in patients with EA and that their expression was decreased after ICS treatment in patients with EA.

IL-1 β (*IL1B*) is a proinflammatory cytokine activated by the NLRP-3 inflammasome, which is a key mediator of inflammation. We^{6,24} and others⁹ have previously reported increased expression of *IL1B* and the IL-1 pathway associated with NA. *IL1B* also directly interacts with many of the differentially expressed genes in this study and is likely to be crucial to the pathogenesis of NA. Further studies are needed to understand the mechanisms of *IL1B* production in the airways, including activation of the inflammasome in patients with NA and whether *IL1B* can be a potential treatment target.

There are 4 known alkaline phosphatases, including intestinal, placental, placenta-like, and tissue nonspecific (*ALPL*). The tissue-nonspecific *ALPL* gene is located on chromosome 1, and

the gene product is a membrane-bound glycosylated enzyme with unknown function. The role of *ALPL* in asthma pathogenesis remains unknown; however, *ALPL* is a member of the TNF- α /nuclear factor κ B pathway and is thought to be involved in responses to LPS, potentially contributing to neutrophilic inflammation in asthmatic patients.^{6,25}

Chemokine (C-X-C motif) receptor 2 is a G protein-coupled receptor that binds IL-8 with high affinity and mediates neutrophil migration to sites of inflammation. Recently, CXCR2 receptor antagonists have been reported to reduce sputum neutrophil counts in patients with severe asthma.²⁶ CXCR2 is thought to be involved in uncontrolled neutrophil influx into the airways in acute lung injury.²⁷ β -Agonists and corticosteroids have been reported to increase expression of CXCR2, along with CXCR1 and CXCL8, without influencing neutrophil chemotaxis to the airways, but the effects of this remain unknown.²⁸ Our study extends the reports in the literature to demonstrate that sputum expression levels of *IL1B*, *ALPL*, and *CXCR2* are higher in patients with NA and that expression levels of *IL1B* and *ALPL* were increased after ICS treatment.

Biomarker development needs to be applicable in predicting response to therapy to improve the treatment of asthma. This study demonstrated that expression of the 6-gene signature was able to predict ICS responsiveness, particularly a greater than 12% improvement in FEV₁, a greater than 0.5-point decrease in ACQ scores, and a 40% reduction in FENO values. The 6-gene signature was better than the sputum eosinophil percentage at predicting ICS response but significantly better at predicting a reduction in FENO values after ICSs. We¹⁰ and others²⁹ have previously reported that the ICS response was greater in patients with EA compared with that seen in those with NEA. Recognition of EA also results in better application of newer therapies³⁰ and improved outcomes compared with guideline-based therapy.³¹ Poor responses might be mediated by an inflammatory pattern

that is insensitive to corticosteroids, which is thought to be the case for airway neutrophilia. Further studies are needed to investigate the prediction of treatment responses in patients with NA. Certainly, corticosteroids can enhance neutrophil survival and thereby increase neutrophil accumulation in the airways.³² Future studies investigating the increase in expression of *IL1B* and *ALPL* caused by ICS treatment of asthma and their relationship to infiltrating neutrophils are warranted.

There are a number of limitations of this study that require further investigation. We established the gene signature in an adult population with stable asthma, and therefore further investigations are required to understand how the expression levels change in exacerbations, whether the signature could predict future risk of loss of asthma control, the effects of covariates and confounders, and how the genes are expressed in children with asthma. Further studies are needed to understand the relationship between gene expression of the 6 markers and their corresponding protein levels, underlying mechanisms, and cell types responsible.

Despite these limitations, there are also a number of strengths to this study, including robust and reproducible validation of the markers in separate populations and at multiple research sites, even though there were slight differences in the methods of sputum processing. Although induced sputum characterizes granulocyte proportions to identify the inflammatory phenotype, this 6-gene expression biomarker might also detect the presence of other cell types, such as mast cells in *CPA3* expression, and likely underlying mechanisms, such as upregulation of *IL1B* in patients with NA. With further development, this PCR-based test is likely to be far superior and wide reaching than sputum cell counts because the standardized methodological techniques are fast, simple, and easily automated.

In summary, we have identified a sputum gene expression signature of 6 markers that identify asthma inflammatory phenotype. Some markers show novel association with asthma and phenotypes of asthma. This signature might assist in the clinical diagnosis and management of asthma, including the prediction of response to ICS treatment. The 6 genes are likely to be involved in underlying mechanisms of asthma and might prove to be potential therapeutic targets for newly developed treatments. Future studies should investigate the diagnostic utility of the gene signature in guiding phenotype-based treatment strategies.

We acknowledge the technical assistance of Bridgette Ridewood, Michelle Gleeson, Kellie Fakes, Joanne Smart, Tom Wright, and Rodney Lea.

Clinical implications: This study has identified a 6-gene expression signature measured in induced sputum samples for use as a diagnostic tool to predict asthma inflammatory phenotype and corticosteroid treatment responsiveness.

REFERENCES

- Ito K, Chung KF, Adcock IM. Update on glucocorticoid action and resistance. *J Allergy Clin Immunol* 2006;117:522-43.
- Szeffer SJ, Martin RJ. Lessons learned from variation in response to therapy in clinical trials. *J Allergy Clin Immunol* 2010;125:285-92.
- Wadsworth S, Sandford A. Personalised medicine and asthma diagnostics/management. *Curr Allergy Asthma Rep* 2013;13:118-29.
- Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006;11:54-61.
- Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012;18:716-25.
- Baines KJ, Simpson JL, Wood LG, Scott RJ, Gibson PG. Transcriptional phenotypes of asthma defined by gene expression profiling of induced sputum samples. *J Allergy Clin Immunol* 2011;127:153-60.
- Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, et al. T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009;180:388-95.
- Fitzpatrick AM, Higgins M, Holguin F, Brown LA, Teague WG, National Institutes of Health/National Heart, Lung, and Blood Institute's Severe Asthma Research Program, et al. The molecular phenotype of severe asthma in children. *J Allergy Clin Immunol* 2010;125:851-7.
- Hastie AT, Moore WC, Meyers DA, Vestal PL, Li H, Peters SP, et al. Analyses of asthma severity phenotypes and inflammatory proteins in subjects stratified by sputum granulocytes. *J Allergy Clin Immunol* 2010;125:1028-36.e13.
- Cowan DC, Cowan JO, Palmay R, Williamson A, Taylor DR. Effects of steroid therapy on inflammatory cell subtypes in asthma. *Thorax* 2010;65:384-90.
- Dvorak AM, Letourneau L, Login GR, Weller PF, Ackerman SJ. Ultrastructural localization of the Charcot-Leyden crystal protein (lysophospholipase) to a distinct crystalloid-free granule population in mature human eosinophils. *Blood* 1988;72:150-8.
- Ackerman SJ, Weil GJ, Gleich GJ. Formation of Charcot-Leyden crystals by human basophils. *J Exp Med* 1982;155:1597-609.
- Kubach J, Lutter P, Bopp T, Stoll S, Becker C, Huter E, et al. Human CD4+CD25+ regulatory T cells: proteome analysis identifies galectin-10 as a novel marker essential for their anergy and suppressive function. *Blood* 2007;110:1550-8.
- Dor PJ, Ackerman SJ, Gleich GJ. Charcot-Leyden crystal protein and eosinophil granule major basic protein in sputum of patients with respiratory diseases. *Am Rev Respir Dis* 1984;130:1072-7.
- Devouassoux G, Pachot A, Laforest L, Diasparra J, Freymond N, Van Ganse E, et al. Galectin-10 mRNA is overexpressed in peripheral blood of aspirin-induced asthma. *Allergy* 2008;63:125-31.
- Nagrete-Garcia MC, Jimenez-Torres CY, Alvarado-Vasquez N, Montes-Vizuet AR, Velazquez-Rodriguez JR, Jimenez-Martinez MC, et al. Galectin-10 is released in the nasal lavage fluid of patients with aspirin-sensitive respiratory disease. *ScientificWorldJournal* 2012;2012:474020.
- Chua JC, Douglass JA, Gillman A, O'Hehir RE, Meeusen EN. Galectin-10, a potential biomarker of eosinophilic airway inflammation. *PLoS One* 2012;7:e42549.
- Trivedi NN, Caughey GH. Mast cell peptidase. Chameleons of innate immunity and host defense. *Am J Respir Cell Mol Biol* 2010;42:257-67.
- Balzar S, Falt ML, Comhair SA, Erzurum SC, Bleeker E, Busse WW, et al. Mast cell phenotype, location and activation. Data from the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2011;183:299-309.
- Woodruff PG, Boushey HA, Dolganov GM, Barker CS, Yang YH, Donnelly S, et al. Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci U S A* 2007;104:15858-63.
- Dougherty RH, Sidhu SS, Raman K, Solon M, Solberg OD, Caughey GH, et al. Accumulation of intraepithelial mast cells with a unique protease phenotype in Th2-high asthma. *J Allergy Clin Immunol* 2010;125:1046-53.
- Al-Mayouf SM, Sunker A, Abdwani R, Arawi SA, Almurshedi N, Al Sonbul A, et al. Loss-of-function variant in DNASE1L3 causes a familial form of systemic lupus erythematosus. *Nat Genet* 2011;43:1186-8.
- Silverman RA, Foley F, Dalipi R, Kline M, Lesser M. The use of rhDNAse in severely ill, non-intubated adult asthmatics refractory to bronchodilators: a pilot study. *Respir Med* 2012;106:1096-102.
- Simpson JL, Grissell TV, Douwes J, Scott RJ, Boyle MJ, Gibson PG. Innate immune activation in neutrophilic asthma and bronchiectasis. *Thorax* 2007;62:211-8.
- Baines KJ, Simpson JL, Bowden NA, Scott RJ, Gibson PG. Differential gene expression and cytokine production from neutrophils in asthma phenotypes. *Eur Respir J* 2010;35:522-31.
- Nair P, Gaga M, Zervas E, Alagha K, Hargreave FE, O'Byrne PM, et al. Safety and efficacy of a CXCR2 antagonist in patients with severe asthma and sputum neutrophils: a randomized, placebo-controlled clinical trial. *Clin Exp Allergy* 2012;42:1097-103.

27. Konrad FM, Reutershan J. CXCR2 in acute lung injury. *Mediators Inflamm* 2012; 2012:740987.
28. Strandberg K, Blidberg K, Sahlander K, Palmberg L, Larsson K. Effect of formoterol and budesonide on chemokine release, chemokine receptor expression and chemotaxis in human neutrophils. *Pulm Pharmacol Ther* 2010;23:316-23.
29. Berry M, Morgan A, Shaw DE, Parker D, Green R, Brightling C, et al. Pathological features and inhaled corticosteroid response of eosinophilic and non-eosinophilic asthma. *Thorax* 2007;62:1043-9.
30. Nair P, Pizzichini MM, Kjarsgaard M, Inman D, Efthimiadis A, Pizzichini E, et al. Mepolizumab for prednisone-dependent asthma with sputum eosinophilia. *N Engl J Med* 2009;360:985-93.
31. Green RH, Brightling CE, McKenna S, Hargadon B, Parker D, Bradding P, et al. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. *Lancet* 2002;30:1715-21.
32. Saffar AS, Ashdown H, Gounni AS. The molecular mechanisms of glucocorticoid-mediated neutrophil survival. *Curr Drug Targets* 2011;12:556-62.

Did you know? The *JACI* has a new website!

You can now personalize the *JACI* website to meet your individual needs. Enjoy these new benefits and more:

- Stay current in your field with Featured Articles of The Week, Articles in Press, and easily view the Most Read and Most Cited articles.
- Sign up for a personalized alerting service with Table of Contents Alerts, Articles in Press Alerts and Saved Search Alerts to notify you of newly published articles.
- Search across 400 top medical and health sciences journals online, including MEDLINE.
- Greater cross-referencing results from your online searches.

Visit www.jacionline.org today to see what else is new online!

ONLINE REPOSITORY

Sputum gene expression of six markers identifies asthma inflammatory phenotype and steroid response

Katherine J Baines, Jodie L Simpson, Lisa G Wood, Rodney J Scott, Naomi L Fibbens, Heather Powell, Douglas C Cowan, D Robin Taylor, Jan O Cowan, Peter G Gibson.

METHODS

Statistical Analysis for quantitative assessment of multiple biomarkers

Logistic regression was used to calculate the predicted value of an individual having the particular asthma inflammatory phenotype based on the expression (cycle threshold (Ct)) of each target gene compared to β -actin (Δ Ct) using simple logistic regression or multiple logistic regression for the combination of 6 genes to generate one set of predicted values. This was generated in the Stata 11 statistical package. The resulting output was a set of predicted value cut points which had ranging sensitivities and specificities plotted on a receiver operating characteristic (ROC) curve, which we could use to calculate the area under the curve (AUC) for the level of accuracy to correctly classify people, as well as a p value of significance for the regression itself.

The regression equation utilised was:

$$\ln(p/(1-p)) = \beta_0 + \beta_i x_i$$

$$p = \exp(\beta_0 + \beta_i x_i) / (1 + \exp(\beta_0 + \beta_i x_i))$$

The predicted value is the probability that the person has the inflammatory phenotype given the level of target gene expression. An individuals' gene expression measurements can be entered into

the logistic regression equation to determine their predicted value, a value above or below the cut point would determine their likelihood of the particular asthma inflammatory phenotype with the defined sensitivity and specificity calculated from the ROC curve.

The significance of the regression models and the AUC were used to assist in determining the best combination of markers. Sensitivity and specificity were used to describe the discriminatory power of the gene expression combinations. We selected cut-points in order to maximise the discriminatory power in any given diagnostic situation.

RESULTS

Discovery Study: Microarray analysis

A gene was considered to be expressed if it was flagged as present in at least half (n=23/47, 49%) of samples tested, and this resulted in 14,573 entities out of 22,184 (65.7%) being expressed. Out of these, 5284 were significantly different due to asthma inflammatory phenotype (ANOVA, $p < 0.05$ adjusted for multiple comparisons using Benjamini Hochberg false discovery rate). There were 277 entities that had a >2 fold change of gene expression in 1 out of 3 phenotype comparisons.

There were 48 genes significantly different between EA and NA, with 22 that had higher expression in EA and 26 that had higher expression in NA. There were 148 genes differentially expressed between EA and PGA with 139 genes higher in EA, and 9 genes higher in PGA. There were 197 genes differentially expressed between NA and PGA with 188 genes higher in NA and 9 genes higher in PGA. A venn diagram detailing the overlap between the lists is shown in Figure E1.

Table E1: Clinical characteristics of the participants in the microarray discovery study

	Asthma			P
	Eosinophilic	Neutrophilic	Paucigranulocytic	
N	17	12	18	
Age years, mean (SD)	59 (18)	59 (11)	56 (11)	0.811
Gender M F	7 10	6 6	7 11	0.825
Atopy, n (%)	16 (94)*	6 (50)	9 (50)	0.004
FEV ₁ % predicted, mean (SD)	67 (15)*	71 (23)	87 (18)	0.005
FEV ₁ /FVC %, mean (SD)	62 (10)*	66 (12)	74 (6)	0.002
FENO (ppb), median (Q1, Q3)	52.4 (36.2, 76.2) ^{†§}	22.8 (14.1, 27.0)	16.3 (11.2, 21.3)	<0.001
Smoking, Ex Never	9 8	5 7	6 12	0.502
Pack years, median (Q1, Q3)	10.0 (0.5, 17.0)	5.0 (5.0, 7.5)	4.4 (0.7, 24.0)	0.897
Asthma Control Score, median (Q1, Q3)	1.3 (0.9, 1.7) [†]	1.0 (0.8, 2.1) [†]	0.5 (0.3, 0.7)	0.005
Inhaled Corticosteroid (ICS) Use, Yes No	11 6	10 2	10 8	0.291
ICS dose (µg daily BDP equivalent) median (Q1, Q3)	800 (400, 1000)	1000 (200, 2000)	750 (400, 2000)	0.738
Total cell count x 10 ⁶ /mL, median (Q1, Q3)	3.2 (2.4, 5.5)	5.0 (2.9, 8.6) [†]	2.5 (1.4, 3.4)	0.044
Viability, median (Q1, Q3)	69.2 (56.6, 88.5)	90.2 (78.2, 95.3) [‡]	81.7 (71.9, 90.8)	0.032
Neutrophils, %, median (Q1, Q3)	23.5 (14.8, 43.3)	78.1 (75.4, 83.3) ^{†‡}	36.2 (21.0, 43.8)	<0.001
Eosinophils, % median (Q1, Q3)	23.8 (13.5, 46.3) ^{†§}	0.6 (0.2, 0.8)	0.0 (0.0, 0.5)	<0.001
Macrophages, %, median (Q1, Q3)	36.8 (26.8, 41.8) [§]	18.9 (15.9, 22.4)	62.3 (55.0, 75.3) ^{§‡}	<0.001
Lymphocytes, %, median (Q1, Q3)	0.3 (0.3, 1.3)	0.1 (0.0, 0.6)	0.8 (0.1, 2.0)	0.060
Columnar epithelial, %, median (Q1, Q3)	1.9 (0.0, 2.5)	0.8 (0.3, 3.6)	1.0 (0.0, 4.6)	0.936
Squamous, %, median (Q1, Q3)	3.4 (1.5, 8.1)	2.6 (0.6, 9.3)	3.8 (2.2, 6.4)	0.965

Bonferroni post hoc test *p<0.017 versus paucigranulocytic asthma; kwallis2 [†]p<0.017 versus paucigranulocytic asthma; [‡]p<0.017 versus eosinophilic asthma; [§]p<0.017 versus neutrophilic asthma; ^{||}Chi squared test p<0.017 versus neutrophilic asthma. BDP=beclomethasone dipropionate.

Table E2: Differentially expression genes between asthma inflammatory phenotypes identified by microarray and validated using qPCR

Gene Symbol	Gene Name	Microarray					qPCR	
		Adjusted P* value	PC2 Adjusted P# value	Fold Change			Taqman probe	Pearson R (ΔCt vs array)
				NA vs. PGA	EA vs. PGA	EA vs. NA		
ALPL	Alkaline phosphatase, tissue-nonspecific isozyme	0.007	0.0006	+2.74	+1.38	-1.98	Hs 01029144_m1	-0.66 ^a
BIRC3	Baculoviral IAP repeat containing 3	0.001	<0.0001	+2.04	+2.71	+1.33	Hs 00154109_m1	-0.70 ^a
BPIFB1	BPI fold containing family B, member 1	0.021	0.0068	+1.06	+3.19	+3.01	Hs 00264197_m1	-0.83 ^a
CAMP	Cathelicidin antimicrobial peptide	0.004	0.0030	-1.87	-2.17	-1.16	Hs 00189038_m1	-0.54 ^b
CCR7	Chemokine (C-C motif) receptor 7	0.003	0.0001	+2 .27	+3.39	+1.49	Hs 01013469_m1	-0.91 ^a
CEBPE	CCAAT/enhancer binding protein (C/EBP), epsilon	0.001	<0.0001	+1.13	+2.83	+2.51	Hs 00357657_m1	-0.82 ^a
CLC	Charcot leydon crystal protein	<0.001	<0.0001	+1.69	+9.38	+5.56	Hs 00171342_m1	-0.89 ^a
CPA3	Carboxypeptidase A3	0.002	0.0001	+1.03	+2.12	+2.07	Hs 00157019_m1	-0.82 ^a
CXCR1	Chemokine (C-X-C motif) receptor 1	0.007	0.0010	+3.45	+1.52	-2.28	Hs 00174146_m1	-0.63 ^a
CXCR2	Chemokine (C-X-C motif) receptor 2	0.006	0.0006	+2.70	+1 .33	-2.01	Hs 00174304_m1	-0.81 ^a
DEFB1	β-defensin 1	0.003	0.0002	-2.07	-2.77	-1.34	Hs 00608345_m1	-0.85 ^a
DNASE1L3	Deoxyribonuclease I-like 3	<0.001	<0.0001	+1 .36	+2.81	+2.07	Hs 00172840_m1	-0.74 ^a
FPR2	Formyl peptide receptor 2	0.002	<0.0001	+2.81	+1 .31	-2.15	Hs 00265954_m1	-0.48 ^b
GBP1	Guanylate binding protein 1, interferon inducible	0.007	0.0009	+1.74	-1.42	-2.47	Hs 00977005_m1	-0.60 ^a
GCHFR	GTP cyclohydrolase I feedback regulator	0.002	0.0001	-2.35	-2.01	+1 .17	Hs 00193360_m1	-0.74 ^a
HDC	Histidine decarboxylase	0.001	<0.0001	+1 .43	+3.83	+2.68	Hs 00157914_m1	-0.88 ^a
IL18RAP	Interleukin-18 receptor accessory protein	0.005	0.0002	+2.93	+1.61	-1.83	Hs 00977695_m1	-0.73 ^a
IL1B	Interleukin-1β	0.003	0.0003	+2.36	+1.27	-1.85	Hs 00174097_m1	-0.68 ^a
IL1R2	Interleukin-1 receptor type II	0.009	0.0014	+3.28	+2.06	-1.59	Hs 01030389_m1	-0.91 ^a
IL1RN	Interleukin-1 receptor antagonist	0.006	0.0009	+2.33	+1.71	-1.36	Hs 00893626_m1	-0.58 ^a
IRAK2	Interleukin-1 receptor-associated kinase 2	0.016	0.0027	+2.77	+1.68	-1.65	Hs 00176394_m1	-0.79 ^a
ISG20	Interferon stimulated exonuclease gene 20kDa	0.005	0.0002	+3.16	+1.70	-1.86	Hs 00158122_m1	-0.90 ^a

NAMPT	Nicotinamide phosphoribosyltransferase	0.004	<0.0001	+2.77	+1.31	-2.12	Hs 00237184_m1	-0.54 ^b
NFKB2	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	0.005	0.0003	+2.03	+1.81	-1.13	Hs 00174517_m1	-0.59 ^a
NFKBIZ	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	0.003	0.0005	+2.35	+1.64	-1.44	Hs 00230071_m1	-0.55 ^b
NLRP3	NLR family, pyrin domain containing 3	0.004	0.0002	+2.72	+1.89	-1.44	Hs 00366465_m1	-0.78 ^a
ORM1	Orosmuroid 1	0.003	0.0008	+4.47	+1.92	-2.33	Hs 01590791_m1	-0.73 ^a
PELI1	Pellino E3 ubiquitin protein ligase 1	0.003	<0.0001	+2.70	+1.60	-1.69	Hs 00221035_m1	-0.65 ^a
PI3	Peptidase inhibitor 3	0.012	0.0030	+3.95	+1.98	-2.00	Hs 00160066_m1	-0.75 ^a
PLAU	Plasminogen activator, urokinase	0.014	0.0028	+2.47	+1.61	-1.53	Hs 01547050_m1	-0.77 ^a
TFF3	Trefoil factor 3	0.017	0.0036	+1.04	+3.08	+2.96	Hs 00173625_m1	-0.87 ^a
TNFAIP3	Tumor necrosis factor α induced protein 3	0.003	0.0010	+2.38	+1.61	-1.47	Hs 00234713_m1	-0.66 ^a
TNFRSF1B	Tumor necrosis factor receptor superfamily, member 1B	0.002	<0.0001	+2.17	+1.77	-1.23	Hs 00961748_m1	-0.63 ^a
TNFSF14	Tumor necrosis factor (ligand) superfamily, member 14	0.003	0.0004	+3.37	+2.59	-1.30	Hs 00542477_m1	-0.89 ^a
VNN2	Vanin 2	0.003	0.0002	+2.80	+1.14	-2.46	Hs 00190581_m1	-0.64 ^a

*ANOVA p value adjusted for multiple comparisons using Benjamini Hochberg FDR correction; NA=neutrophilic asthma; EA=eosinophilic asthma, PGA=paucigranulocytic asthma. ^ap<0.0001, ^bp<0.001

Significance p value for the regression model of each target gene against inflammatory phenotype with the inclusion of the PC2 result to correct for batch effect.

Table E3: Analysis of diagnostic value of 4 individual gene expression biomarkers to diagnose eosinophilic asthma

Comparison	Gene Biomarker	AUC %	P value	Minimal False Negatives*			Minimal False Positives*		
				Predicted Value Cut Point	Sensitivity	Specificity	Predicted Value Cut Point	Sensitivity	Specificity
EA vs. NEA	CLC	80.2	<0.0001	≥ 0.319	78.4%	74.6%	≥ 0.411	73.0%	83.6%
	CPA3	86.1	<0.0001	≥ 0.256	89.2%	67.2%	≥ 0.407	70.3%	88.1%
	DNASE1L3	76.5	<0.0001	≥ 0.351	75.7%	65.7%	≥ 0.422	67.6%	77.6%
	HDC	75.5	<0.0001	≥ 0.346	70.3%	65.2%	≥ 0.419	56.8%	72.7%
EA vs. PGA	CLC	86.6	<0.0001	≥ 0.308	86.5%	67.6%	≥ 0.667	73.0%	91.9%
	CPA3	90.6	<0.0001	≥ 0.294	94.6%	62.2%	≥ 0.634	70.3%	94.6%
	DNASE1L3	80.3	<0.0001	≥ 0.440	86.5%	70.3%	≥ 0.534	75.7%	78.4%
	HDC	84.7	<0.0001	≥ 0.338	91.9%	63.9%	≥ 0.569	70.3%	77.8%
EA vs. NA	CLC	72.3	0.002	≥ 0.496	75.7%	66.7%	≥ 0.544	73.0%	73.3%
	CPA3	80.5	<0.0001	≥ 0.438	86.5%	63.3%	≥ 0.553	73.0%	80.0%
	DNASE1L3	71.9	0.006	≥ 0.569	70.3%	73.3%	≥ 0.572	67.6%	76.7%
	HDC	64.4	0.019	≥ 0.511	67.6%	53.3%	≥ 0.567	51.4%	60.0%

*Minimal false negatives corresponds to the point of the ROC curve with the highest sensitivity (true positive rate, useful for ruling disease out) whereas minimal false positives corresponds to the point with the highest specificity (false positive rate, useful for ruling disease in).

EA=eosinophilic asthma, NEA=non-eosinophilic asthma; PGA=paucigranulocytic asthma; NA=neutrophilic asthma.

Table E4: Analysis of diagnostic value of 4 individual gene expression biomarkers to diagnose neutrophilic asthma

Comparison	Gene Biomarker	AUC %	P value	Minimal False Negatives*			Minimal False Positives*		
				Predicted Value Cut Point	Sensitivity	Specificity	Predicted Value Cut Point	Sensitivity	Specificity
NA vs. NNA	ALPL	71.9	<0.0001	≥ 0.234	86.7%	64.9%	≥ 0.397	63.3%	79.7%
	IL1B	78.9	<0.0001	≥ 0.248	73.3%	64.9%	≥ 0.347	63.3%	77.0%
	CXCR2	77.6	<0.0001	≥ 0.269	76.7%	67.6%	≥ 0.369	66.7%	82.4%
	PI3	75.8	<0.0001	≥ 0.250	73.3%	59.5%	≥ 0.351	56.7%	78.4%
NA vs. PGA	ALPL	81.0	<0.0001	≥ 0.438	83.3%	73.0%	≥ 0.536	70.0%	78.4%
	IL1B	83.1	<0.0001	≥ 0.337	83.3%	64.9%	≥ 0.519	70.0%	78.4%
	CXCR2	81.9	<0.0001	≥ 0.360	83.3%	62.2%	≥ 0.519	70.0%	81.1%
	PI3	78.8	<0.0001	≥ 0.355	83.3%	59.5%	≥ 0.560	56.7%	86.5%
NA vs. EA	ALPL	77.4	<0.0001	≥ 0.495	75.7%	63.3%	≥ 0.648	62.2%	86.7%
	IL1B	74.8	0.0006	≥ 0.529	73.0%	63.3%	≥ 0.632	62.2%	76.7%
	CXCR2	73.2	0.0008	≥ 0.512	81.1%	66.7%	≥ 0.597	62.2%	76.7%
	PI3	72.7	0.0009	≥ 0.539	70.3%	56.7%	≥ 0.612	56.8%	73.3%

*Minimal false negatives corresponds to the point of the ROC curve with the highest sensitivity (true positive rate, useful for ruling disease out) whereas minimal false positives corresponds to the point with the highest specificity (false positive rate, useful for ruling disease in). NA=neutrophilic asthma; NNA=non-neutrophilic asthma; PGA=paucigranulocytic asthma; EA=eosinophilic asthma.

Table E5: Analysis of diagnostic value of *CLC*, *CPA3*, *DNASE1L3*, *IL1B*, *ALPL* and *CXCR2* in combination for asthma from healthy controls

Comparison	Logistic regression		AUC %	Model P value	Minimal False Negatives*			Minimal False Positives*		
	Coefficient	Constant			Predicted Value Cut Point	Sensitivity	Specificity	Predicted Value Cut Point	Sensitivity	Specificity
Asthma vs. Healthy										
CLC	0.5975588									
CPA3	0.1077895									
DNASE1L3	-0.2065524	-6.681488	84.4	0.0001	≥0.177	85.0%	74.0%	≥0.245	65.0%	86.5%
IL1B	0.2038222									
ALPL	-0.2032851									
CXCR2	0.0560672									
EA vs. Healthy										
CLC	0.9637631									
CPA3	1.037619									
DNASE1L3	-0.2943698	-16.01436	97.4	<0.0001	≥0.571	95.0%	94.6%	≥0.862	70.0%	100.0%
IL1B	0.2922089									
ALPL	-0.8238078									
CXCR2	0.3419623									
NA vs. Healthy										
CLC	0.8647747									
CPA3	-0.0772335									
DNASE1L3	-0.1212787	-10.91176	91.8	<0.0001	≥0.180	100.0%	70.0%	≥0.670	65.0%	96.7%
IL1B	0.0470821									
ALPL	0.3482125									
CXCR2	0.3623759									

*Minimal false negatives corresponds to the point of the ROC curve with the highest sensitivity (true positive rate, useful for ruling disease out) whereas minimal false positives corresponds to the point with the highest specificity (false positive rate, useful for ruling disease in). NA=neutrophilic asthma; EA=eosinophilic asthma.

Table E6: Analysis of mixed eosinophilic/neutrophilic asthma using the 6 gene biomarker signature.

	Neutrophil%	Eosinophil%	Eosinophil% C2R	Predicted Values			
				N v P		E v P	
				Predicted value	N or P	Predicted value	E or P
Subject 1	67.3	4.5	6.3	0.542	N	0.554	E
Subject 2	80.0	3.5	4.5	0.415	P	0.969	E
Subject 3	81.5	3.0	2.5	0.508	P	0.767	E
Subject 4	69.3	8.5	8.0	0.788	N	0.938	E
Subject 5	70.0	3.0	4.3	0.575	N	0.300	P
Subject 6	63.0	3.5	3.5	0.178	P	0.368	P
Subject 7	66.5	5.0	6.8	0.745	N	0.255	P
Subject 8	70.8	3.3	4.3	0.750	N	0.434	E
Subject 9	64.0	8.3	NA	0.629	N	0.115	P
Subject 10	76.0	2.5	4.3	0.853	N	0.661	E
Subject 11	76.3	2.0	4.3	0.942	N	0.994	E
Subject 12	72.8	1.8	4.3	0.667	N	0.850	E

N v P predicted value cut off of 0.538 used which had a sensitivity of 73.3% and specificity of 83.8%

E v P predicted value cut off of 0.369 used which had a sensitivity of 91.9% and specificity of 75.7%

P=paucigranulocytic asthma, NP=non-paucigranulocytic asthma, N=neutrophilic asthma, E=eosinophilic asthma, C2R=chromothrope 2R, NA=not available

Table E7: Analysis of diagnostic value of *CLC*, *CPA3*, *DNASE1L3*, *IL1B*, *ALPL* and *CXCR2* in combination for steroid responsiveness

Comparison	Logistic regression		AUC %	Model P value	Minimal False Negatives*			Minimal False Positives*		
	Coefficient	Constant			Predicted Value Cut Point	Sensitivity	Specificity	Predicted Value Cut Point	Sensitivity	Specificity
FEV₁ Responder vs. Non Responder										
CLC	-0.8451629									
CPA3	0.0764054	9.315089	91.5	<0.0001	≥0.353	95.2	65.5	≥0.683	76.2	93.1
DNASE1L3	-0.4695241									
IL1B	0.4592733									
ALPL	0.4098769									
CXCR2	-0.3667733									
ACQ Responder vs. Non Responder										
CLC	-0.2897043									
CPA3	-0.0709981	7.708473	85.7	0.0022	≥0.654	84.6	68.4	≥0.820	69.2	94.7
DNASE1L3	-0.4254671									
IL1B	1.057213									
ALPL	0.0005818									
CXCR2	-0.1353135									
FeNO Responder vs. Non Responder										
CLC	-0.3138916									
CPA3	-0.6297407	7.250561	87.1	<0.0001	≥0.481	90.9	69.2	≥0.786	63.6	88.5
DNASE1L3	0.0848139									
IL1B	1.211543									
ALPL	0.2931872									
CXCR2	-0.5106013									

*Minimal false negatives corresponds to the point of the ROC curve with the highest sensitivity (true positive rate, useful for ruling disease out) whereas minimal false positives corresponds to the point with the highest specificity (false positive rate, useful for ruling disease in).

FEV₁ Responder= subjects that had a >12% improvement FEV₁ after steroid treatment, ACQ Responder= subjects that had a >0.5 point decrease in ACQ score after steroid treatment, FeNO Responder=subjects that had a >40% decrease in F_ENO after steroid treatment.

SUPPLEMENTARY FIGURES

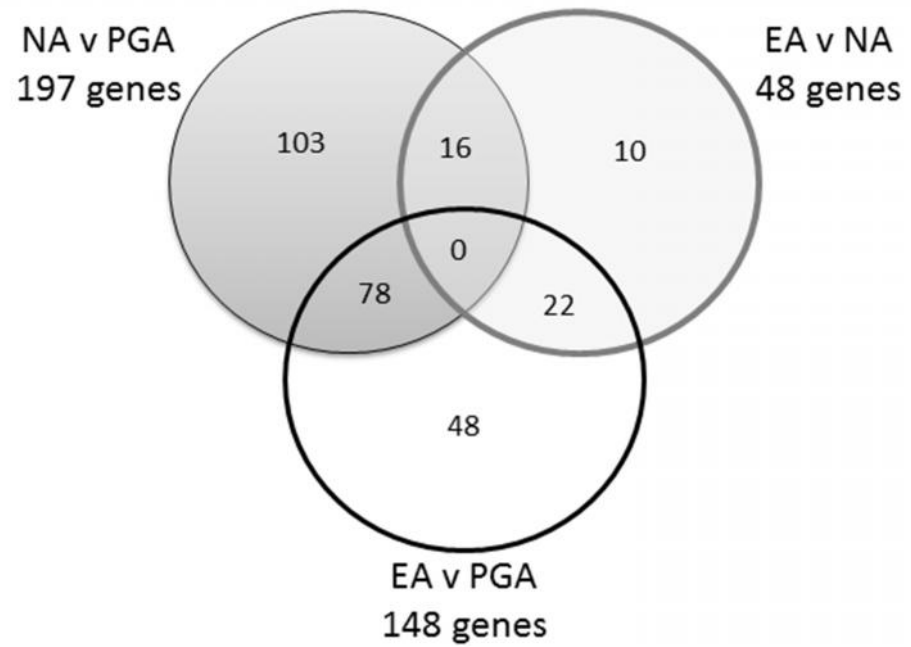


Figure E1: Venn diagram showing differentially expressed genes between asthma inflammatory phenotypes.

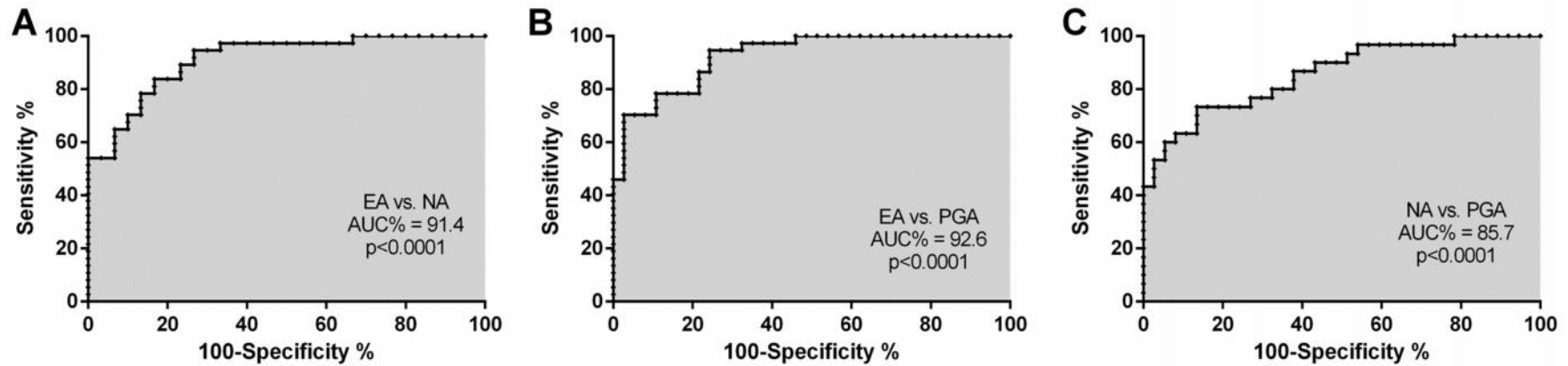


Figure E2: Receiver operating characteristic curves demonstrate that the 6 gene biomarker signature identifies asthma inflammatory phenotype A) eosinophilic (EA) from neutrophilic asthma (NA), B) eosinophilic from paucigranulocytic asthma (PGA) and C) neutrophilic from paucigranulocytic asthma.

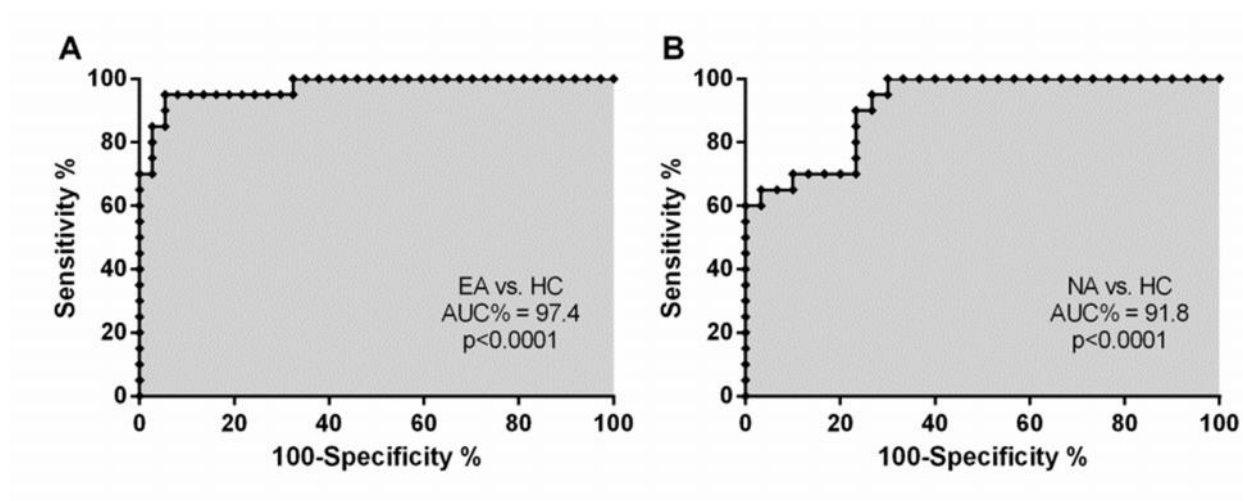


Figure E3: Receiver operating characteristic curves demonstrate that the 6 gene biomarker signature identifies A) eosinophilic asthma (EA) and B) neutrophilic asthma (NA) from healthy controls (HC).

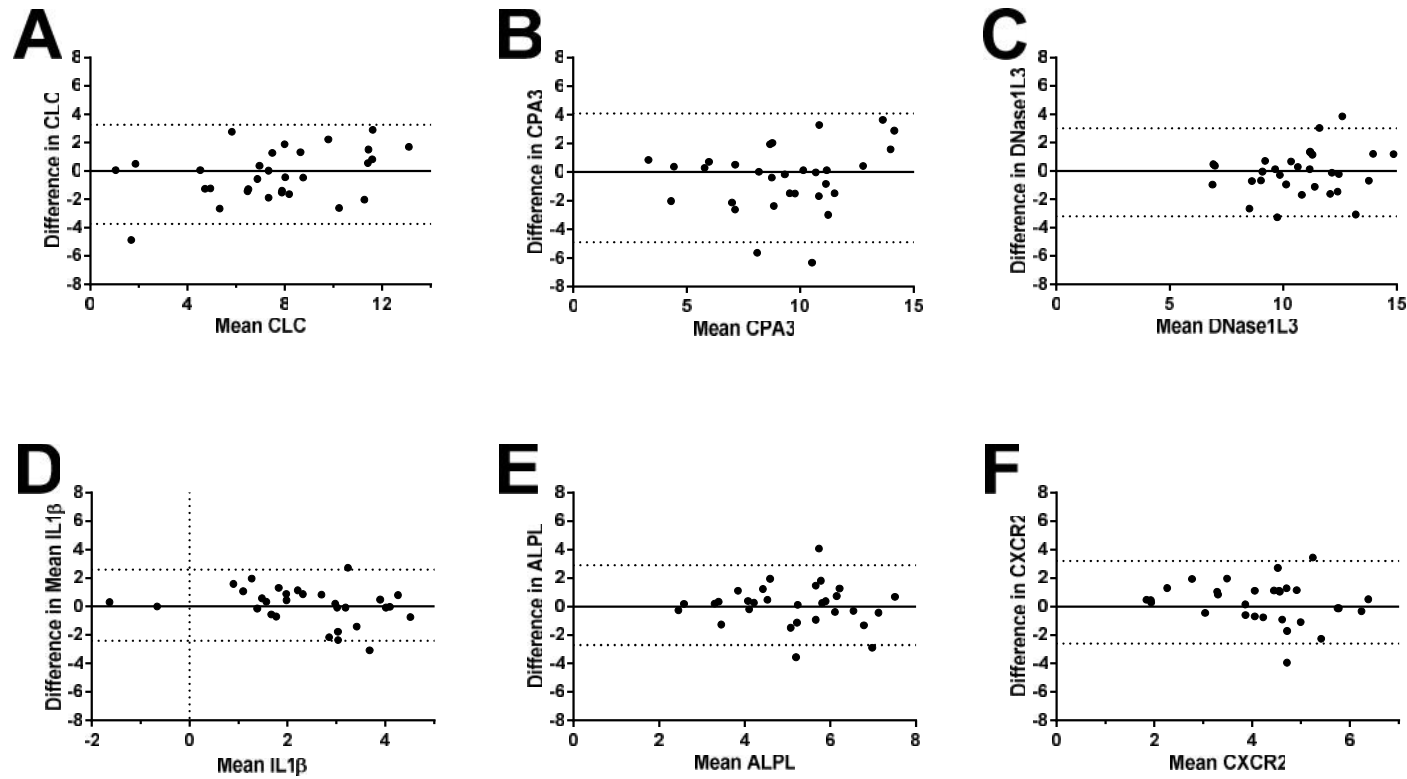


Figure E4: Mean-difference Bland-Altman plots for the EA markers CLC (A), CPA3 (B) and DNASE1L3 (C), and NA markers IL1B (D), ALPL (E) and CXCR2 (F). Bland-Altman plots indicate the mean ΔCt against the absolute difference in ΔCt between visits for each gene. Horizontal dotted lines represent the 95% limits of agreement (mean difference ± 1.96 SD).

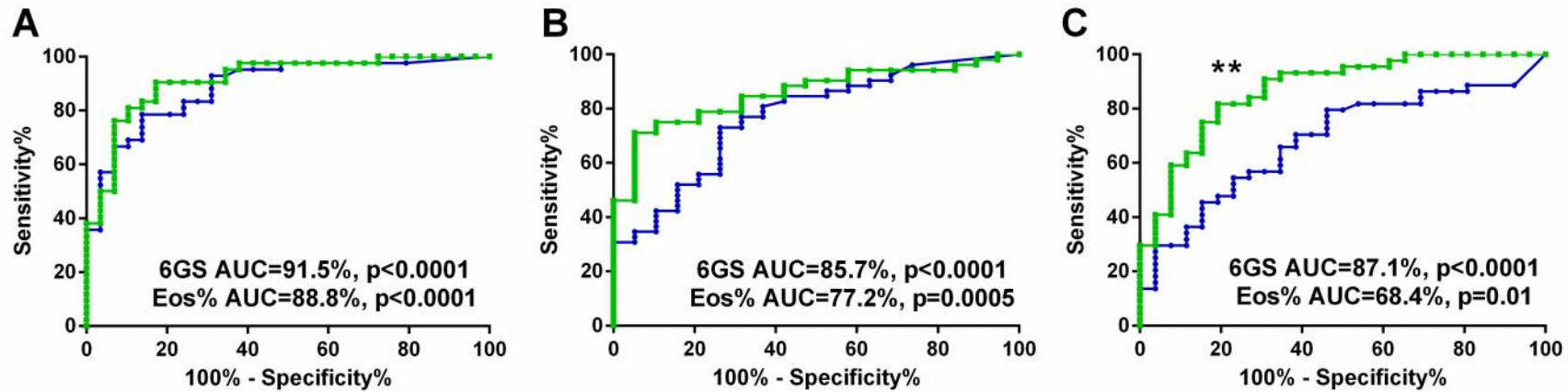


Figure E5: ROC curve comparison for the 6 gene signature (6GS, green line) compared with induced sputum eosinophil % (Eos%, blue line).

The 6GS was better at predicting ICS responsiveness in terms of a change in A) FEV_1 of 12%; B) a 0.5 decrease in ACQ score; and significantly (** $p=0.002$) improved at predicting C) a 40% reduction in F_{ENO} .

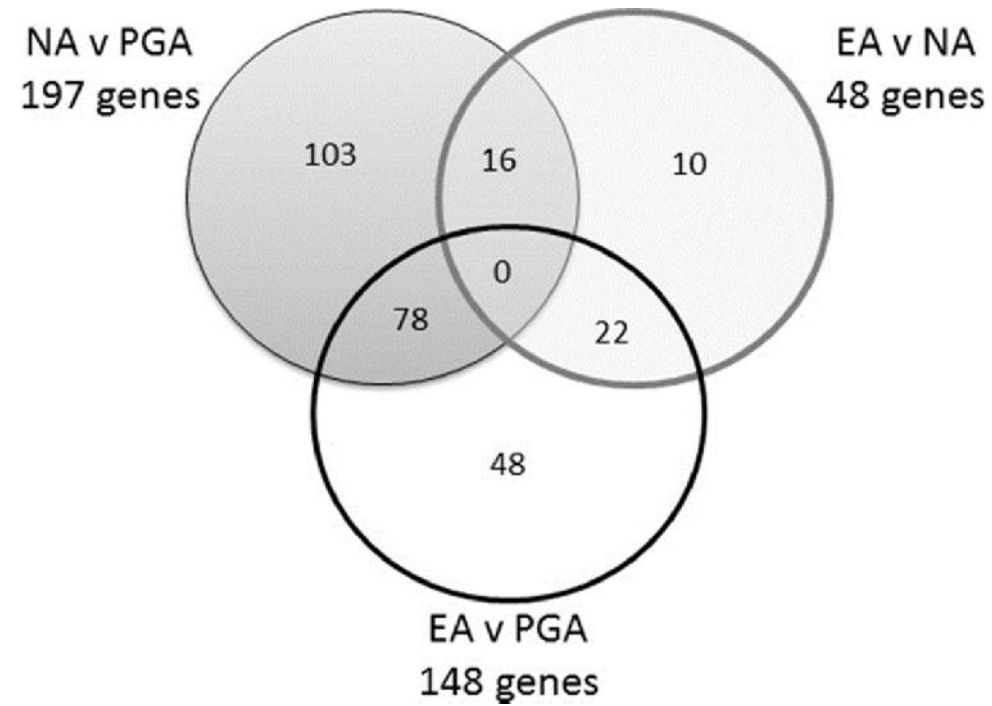


Fig E1

Katherine J. Baines , Jodie L. Simpson , Lisa G. Wood , Rodney J. Scott , Naomi L. Fibbens , Heather Powell , Doug...

Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes

Journal of Allergy and Clinical Immunology, Volume 133, Issue 4, 2014, 997 - 1007

<http://dx.doi.org/10.1016/j.jaci.2013.12.1091>

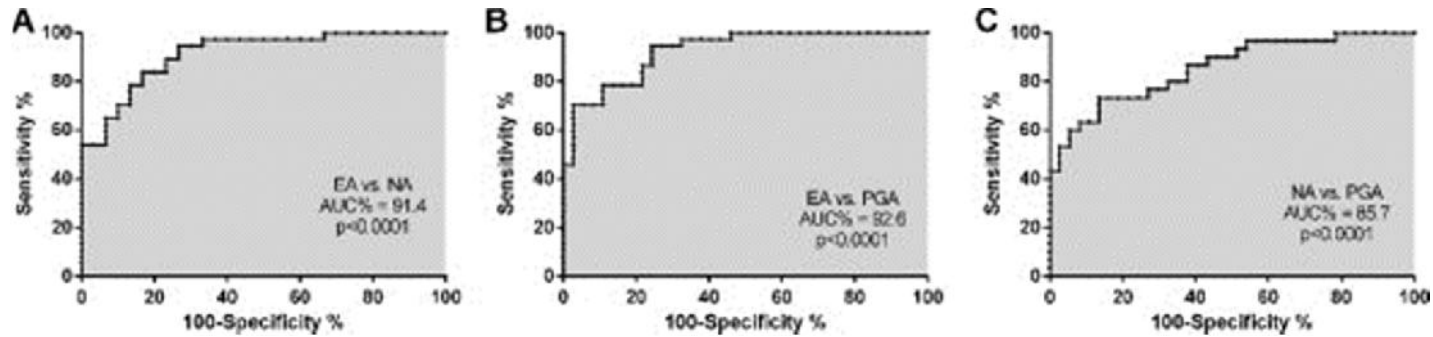


Fig E2

Katherine J. Baines , Jodie L. Simpson , Lisa G. Wood , Rodney J. Scott , Naomi L. Fibbens , Heather Powell , Doug...

Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes

Journal of Allergy and Clinical Immunology, Volume 133, Issue 4, 2014, 997 - 1007

<http://dx.doi.org/10.1016/j.jaci.2013.12.1091>

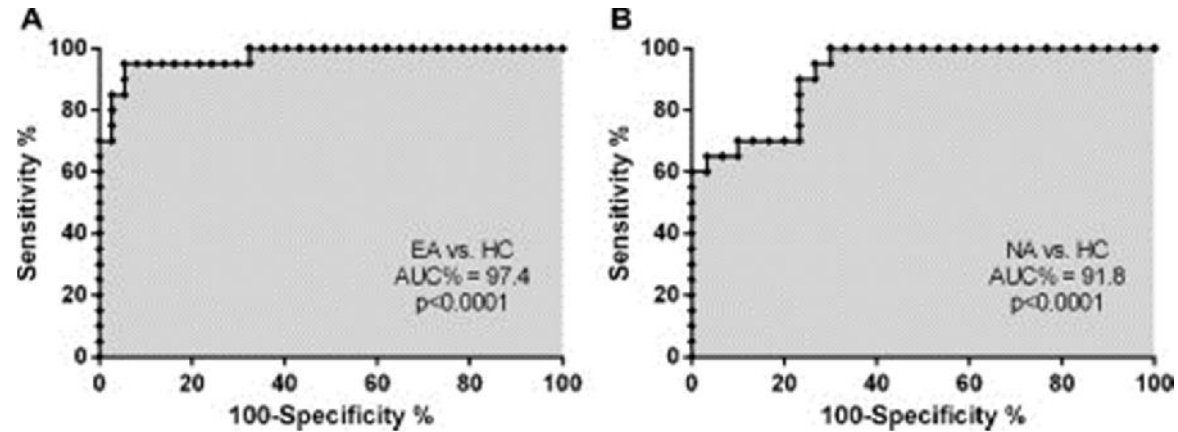


Fig E3

Katherine J. Baines , Jodie L. Simpson , Lisa G. Wood , Rodney J. Scott , Naomi L. Fibbens , Heather Powell , Doug...

Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes

Journal of Allergy and Clinical Immunology, Volume 133, Issue 4, 2014, 997 - 1007

<http://dx.doi.org/10.1016/j.jaci.2013.12.1091>

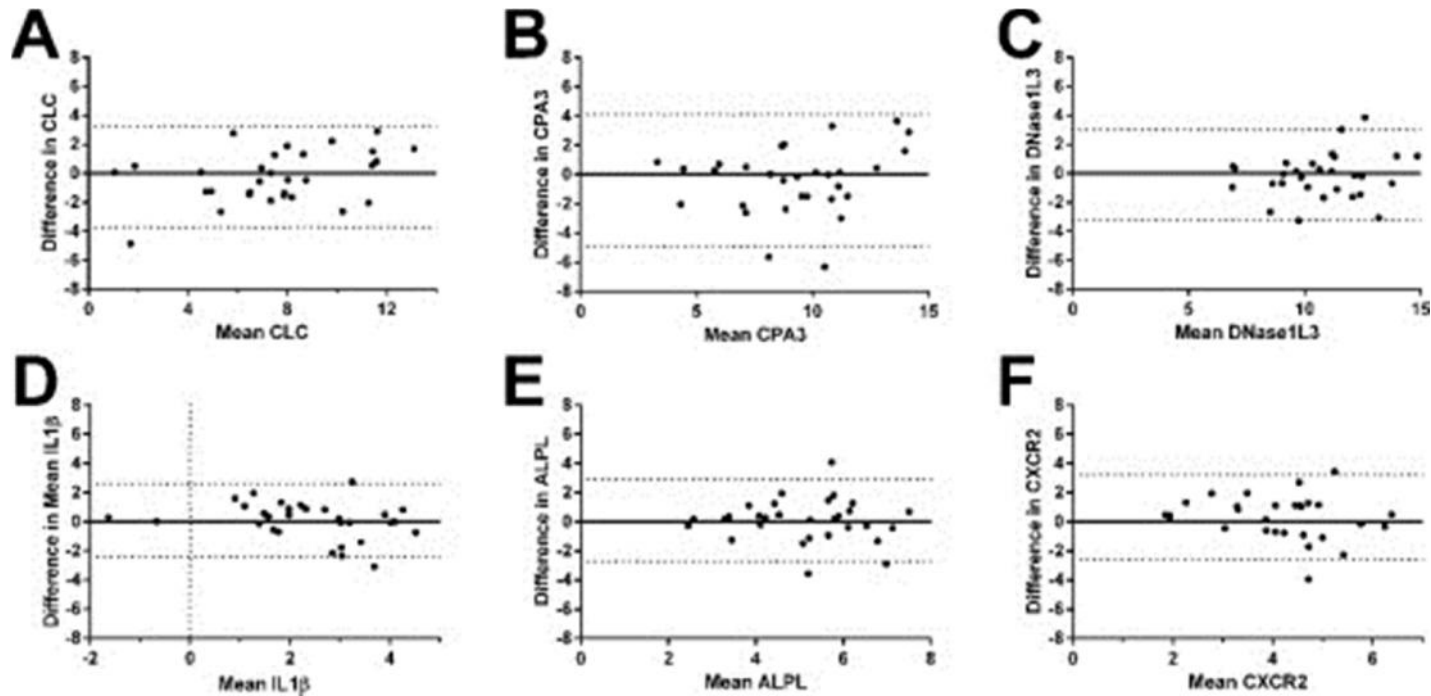


Fig E4

Katherine J. Baines , Jodie L. Simpson , Lisa G. Wood , Rodney J. Scott , Naomi L. Fibbens , Heather Powell , Doug...

Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes

Journal of Allergy and Clinical Immunology, Volume 133, Issue 4, 2014, 997 - 1007

<http://dx.doi.org/10.1016/j.jaci.2013.12.1091>

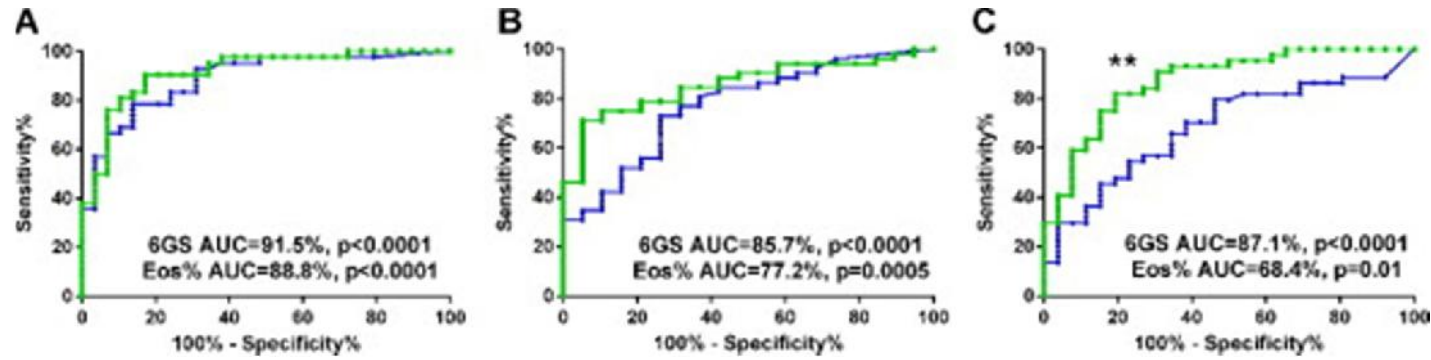


Fig E5

Katherine J. Baines , Jodie L. Simpson , Lisa G. Wood , Rodney J. Scott , Naomi L. Fibbens , Heather Powell , Doug...

Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes

Journal of Allergy and Clinical Immunology, Volume 133, Issue 4, 2014, 997 - 1007

<http://dx.doi.org/10.1016/j.jaci.2013.12.1091>