Molecular, endoscopic, histologic, and circulating biomarker-based diagnosis of eosinophilic gastritis: Multi-site study

Tetsuo Shoda, MD, PhD, Ting Wen, PhD, Julie M. Caldwell, PhD, Margaret H. Collins, MD, John A. Besse, BS, Garrett A. Osswald, BS, J. Pablo Abonia, MD, Nicoleta C. Arva, MD, PhD, Dan Atkins, MD, Kelley E. Capocelli, MD, Evan S. Dellon, MD, MPH, Gary W. Falk, MD, MS, Nirmala Gonsalves, MD, Sandeep K. Gupta, MD, Ikuo Hirano, MD, Vincent A. Mukkada, MD, Philip E. Putnam, MD, Rachel M. Sheridan, MD, Amanda K. Rudman Spergel, MD, Jonathan M. Spergel, MD, Joshua B. Wechsler, MD, PhD, Guang-Yu Yang, MD, PhD, Seema S. Aceves, MD, PhD, Glenn T. Furuta, MD, and Marc E. Rothenberg, MD, PhD, on behalf of the Consortium of Eosinophilic Gastrointestinal Disease Researchers (CEGIR)*

Cincinnati, Ohio, Evanston, Chicago, and Peoria, Ill, Aurora, Colo, Chapel Hill, NC, Philadelphia, Pa, Bethesda, Md, and San Diego, Calif

**GRAPHICAL ABSTRACT**

Molecular, Endoscopic, Histologic and Circulating Biomarker-Based Diagnosis of Eosinophilic Gastritis: Multi-Site Study

**Tissue-Based Molecular Panel**
- 18 significant genes
- Monitor disease activity
- Better diagnosis of ambiguous cases
- Correlations with histology & endoscopy

**Blood-Based Biomarker Panel**
- 4 significant proteins (Eotaxin-3, IL-5, TARC, TSLP)
- Reflect histo-molecular changes in the stomach

From the Division of Allergy and Immunology, the Division of Pathology, and the Division of Gastroenterology and Hepatology, University of Cincinnati College of Medicine and Cincinnati Children’s Hospital Medical Center; the Department of Pathology, Ann & Robert H. Lurie Children’s Hospital of Chicago, Feinberg School of Medicine, and the Department of Pathology and Laboratory Medicine, Northwestern University, Evanston; the Section of Pediatric Allergy and Immunology, Children’s Hospital Colorado, Aurora; the Department of Pathology, Children’s Hospital Colorado, Aurora; the Division of Gastroenterology and Hepatology, University of North Carolina School of Medicine, Chapel Hill; the Division of Gastroenterology, University of Pennsylvania Perelman School of Medicine, Philadelphia; the Division of Gastroenterology & Hepatology, Northwestern University, Feinberg School of Medicine, Chicago; the Division of Pediatric Gastroenterology, Hepatology and Nutrition, University of Illinois College of Medicine/Children’s Hospital of Illinois, Peoria; the Allergy, Asthma and Airway Biology Branch, Division of Allergy, Immunology and Transplantation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda; the Division of Allergy and Immunology, University of Pennsylvania Perelman School of Medicine/Children’s Hospital of Philadelphia; Gastroenterology, Hepatology & Nutrition, Ann & Robert H. Lurie Children’s Hospital of Chicago; the Division of Allergy Immunology, University of California–San Diego, and Rady Children’s Hospital, San Diego; and the Section of Pediatric Gastroenterology, Hepatology and Nutrition, Digestive Health Institute, Children’s Hospital Colorado, Aurora.

*The list of participants is provided in this article’s Online Repository at www.jacionline.org.

N = 185

From the Division of Allergy and Immunology, the Division of Pathology, and the Division of Gastroenterology and Hepatology, University of Cincinnati College of Medicine and Cincinnati Children’s Hospital Medical Center; the Department of Pathology, Ann & Robert H. Lurie Children’s Hospital of Chicago, Feinberg School of Medicine, and the Department of Pathology and Laboratory Medicine, Northwestern University, Evanston; the Section of Pediatric Allergy and Immunology, Children’s Hospital Colorado, Aurora; the Department of Pathology, Children’s Hospital Colorado, Aurora; the Division of Gastroenterology and Hepatology, University of North Carolina School of Medicine, Chapel Hill; the Division of Gastroenterology, University of Pennsylvania Perelman School of Medicine, Philadelphia; the Division of Gastroenterology & Hepatology, Northwestern University, Feinberg School of Medicine, Chicago; the Division of Pediatric Gastroenterology, Hepatology and Nutrition, University of Illinois College of Medicine/Children’s Hospital of Illinois, Peoria; the Allergy, Asthma and Airway Biology Branch, Division of Allergy, Immunology and Transplantation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda; the Division of Allergy and Immunology, University of Pennsylvania Perelman School of Medicine/Children’s Hospital of Philadelphia; Gastroenterology, Hepatology & Nutrition, Ann & Robert H. Lurie Children’s Hospital of Chicago; the Division of Allergy Immunology, University of California–San Diego, and Rady Children’s Hospital, San Diego; and the Section of Pediatric Gastroenterology, Hepatology and Nutrition, Digestive Health Institute, Children’s Hospital Colorado, Aurora.

*The list of participants is provided in this article’s Online Repository at www.jacionline.org.
Background: Eosinophilic gastritis (EG) is a clinicopathologic disorder with marked gastric eosinophilia and clinical symptoms. There is an unmet need among patients with EG for more precise diagnostic tools.

Objective: We aimed to develop tissue- and blood-based diagnostic platforms for EG.

Methods: Patients with EG and control subjects without EG were enrolled across 9 Consortium of Eosinophilic Gastrointestinal Disease Researchers–associated sites. An EG Diagnostic Panel (EGDP; gastric transcript subset) and EG blood biomarker panel (protein multiplex array) were analyzed. EGDP18 scores were derived from the expression of 18 highly dysregulated genes, and blood EG scores were derived from dysregulated cytokine/chemokine levels.

Results: Gastric biopsy specimens and blood samples from 185 subjects (patients with EG, n = 74; control subjects without EG, n = 111) were analyzed. The EGDP (1) identified patients with active EG (P < .0001, area under the curve ≥ 0.95), (2) effectively monitored disease activity in longitudinal samples (P = .0078), (3) highly correlated in same-patient samples (antrum vs body, r = 0.85, P < .0001), and (4) inversely correlated with gastric peak eosinophil levels (r = −0.83, P < .0001), periglandular circumferential collars (r = −0.73, P < .0001), and endoscopic nodularity (r = −0.45, P < .0001). For blood-based platforms, eotaxin-3, thymus and activation-regulated chemokine, IL-5, and thymic stromal lymphopoietin levels were significantly increased. Blood EG scores (1) distinguished patients with EG from control subjects without EG (P < .0001, area under the curve ≥ 0.91), (2) correlated with gastric eosinophil levels (plasma: r = 0.72, P = .0002; serum: r = 0.54, P = .0015), and (3) inversely correlated with EGDP18 scores (plasma: r = −0.64, P = .0015; serum: r = −0.46, P = .0084). Plasma eotaxin-3 levels strongly associated with gastric CCL26 expression (r = 0.81, P < .0001).

Conclusion: We developed tissue- and blood-based platforms for assessment of EG and uncovered robust associations between specific gastric molecular profiles and histologic and endoscopic features, providing insight and clinical readiness tools for this emerging rare disease. (J Allergy Clin Immunol 2020;145:255-69.)

Key words: Biomarker, diagnostic panel, eosinophil, eosinophilic gastritis, transcriptome

Eosinophilic gastritis (EG) is one of the eosinophilic gastrointestinal disorders (EGIDs) characterized by marked eosinophil accumulation into the gastrointestinal tract, with an estimated prevalence of about 6.3 patients per 100,000 subjects, which is likely increasing. Although studies of eosinophilic esophagitis (EoE) have elucidated specific molecular, cellular, and immune mechanisms involved in its pathogenesis, EG is more rare than EoE and thus less well understood, with few publications even addressing diagnostic criteria, genetics, biomarkers, or disease pathogenesis. Unlike in the esophagus, eosinophils normally reside in the gastric mucosa during homeostasis, adding complexity to disease diagnosis and monitoring.

Substantial progress has been made using whole-genome transcript expression profiling (transcriptome) of tissue biopsy specimens from patients with EGIDs, particularly EoE. We previously developed a molecular EoE diagnostic panel, a set of 96 informative transcripts that can distinguish, monitor, and endotype EoE, however, such molecular diagnostic profiles are lacking for the other EGIDs.

To date, EG studies have identified a prominent and conserved gastric transcriptome that is largely distinct from the EoE transcriptome. In this context, we hypothesized that a tissue-based diagnostic platform based on the EG transcriptome would provide more clarity to EG diagnosis than would isolated eosinophil counts, would correlate with gastric endoscopic and pathologic parameters, and would align with specific clinical findings.

EGID diagnosis and management require procuring tissue biopsy specimens during endoscopy, an invasive procedure that adds costs and risk. Two unmet needs are to more precisely and objectively diagnose EG and develop monitoring tools for disease.
management and clinical trial readiness. Circulating biomarkers have the potential to facilitate noninvasive disease diagnosis and monitoring and represent a pressing need in the field. Previous studies suggest that EG has potential systemic markers, such as concurrent peripheral blood eosinophilia. However, the peripheral expression of EG-related inflammatory markers and their potential to function as surrogate disease markers in patients with EG has not been examined.

Here we aimed to develop tissue- and blood-based diagnostic platforms for EG; validate their utility for diagnosis, monitoring, and management; assess their clinical significance by means of concurrent analysis of histologic and endoscopic findings; and provide insight into disease pathogenesis. To approach this aim, we examined patients with EG across multiple sites associated with the Consortium of Eosinophilic Gastrointestinal Disease Researchers (CEGIR) and an independent replication cohort from a single center.

METHODS
Study design and participants
This study was conducted within the context of CEGIR, a national collaborative network of 16 academic centers caring for adults and children with EGIDs. The CEGIR clinical study Outcomes Measures in Eosinophilic Gastrointestinal Disorders Across the Ages is a longitudinal cohort study aimed at understanding the natural history of EoE, EG, and eosinophilic colitis during routine clinical care. Demographic, clinical, endoscopic, and histologic data, as well as gastrointestinal and blood samples, were prospectively collected starting from 2015; all samples from any CEGIR site that contributed patients with EG were used (n = 9 sample-providing institutions). The clinical features of subjects were determined during a standard-of-care evaluation by using standardized intake forms (see the supplementary data provided in this article’s Online Repository at www.jacionline.org for details). All subjects’ clinical data were stored at the Data Management and Coordinating Center at the University of South Florida in Tampa, Florida. Data were systematically extracted from the databases. Pediatric subjects were defined as having an age of less than 18 years. Atopy was defined on the basis of self-report of allergic rhinitis, dermatitis, asthma, or food allergy. For the validation cohort of the tissue-based platform and the plasma cohort of the blood-based platform, children and adults with EG presenting for standard care between 2007 and 2016 were enrolled in an independent cohort at Cincinnati Children’s Hospital Medical Center (CCHMC) using the same disease definitions. This study was approved by the institutional review boards of the participating institutions through a central institutional review board at CCHMC.

Patients were defined as having EG if they had 30 or more eosinophils/high-power field (hpf) in 5 or more hpfs. There were no other known causes of gastric mucosal eosinophilia as defined by standard of care, such as negative test results for other potential causes, including stool culture for pathogenic bacteria or parasites, viral antibody titers and/or PCR, celiac and inflammatory bowel disease serology, and staining for Helicobacter pylori infection. Active EG was defined as gastric biopsy specimens that showed 30 or more eosinophils/hpf in 5 or more hpfs, intermediate EG was defined as gastric biopsy specimens that showed 30 or more eosinophils in 1 to 4 of 5 hpfs, and inactive EG was defined as less than 30 eosinophils/hpf in all hpfs in patients with a previous history of EG. Patients with EG were allowed to be included if they had gastrointestinal eosinophilia outside of the stomach.

Control subjects without EG from the Cincinnati Center for Eosinophilic Disorders EGID database between 2015 and 2018 included children and adults who had undergone endoscopy, had no history of EG, had no pathologic evidence of EG surveyed during the index endoscopy, were not taking systemic glucocorticoid treatments in the period immediately before the index endoscopy, and had gastric biopsy specimens, blood samples, or both collected for research purposes during the index endoscopy. For the tissue-based and blood-based platforms, control subjects included those with atopic comorbidities, chronic gastritis, and concurrent active EoE because these control subjects with different T_{h2} baseline levels would aid in identifying transcriptional changes that are specific to EG. Control subjects were selected with an effort of closely matching sex and age. Gastric biopsy specimens in the discovery and validation cohorts did not overlap.

RNA sequencing analysis
Fresh biopsy specimens collected from patients with EG and control subjects without EG were stored in RNAater until they were subjected to RNA isolation by using the miRNeasy kit (Qiagen, Valencia, Calif), according to the manufacturer’s instructions. Samples for RNA sequencing were selected from the total cohort on the basis of RNA quality and quantity. RNA sequencing acquired 20 million mappable, 75-bp reads from paired-end libraries and was performed at the DNA Sequencing and Genotyping Core Facility at CCHMC. Data were aligned to the GRCh37 build of the human genome by using the Ensembl annotations as a guide for TopHat. Expression analysis was performed with DESeq2 in CLC Genomics Workbench software (CLC bio, Waltham, Mass). Reads per kilobase of exon per million reads mapped were assessed for statistical significance by using a Welch t test with Benjamini-Hochberg false discovery rate (FDR) and a threshold P value of less than .05 and a 2-fold–change cutoff filter. Gene ontology enrichment analysis was performed with the ToppGene suite (https://toppgene.cchmc.org/).

Tissue-based diagnostic platform
RNA from fresh gastric biopsy specimens (from the antrum, body, or both) was isolated from patients with EG and control subjects without EG, as described above. RNA was reverse transcribed with the iScript cDNA Synthesis Kit (170-8891; Bio-Rad Laboratories, Hercules, Calif), according to the manufacturer’s protocol. The transcriptomic signature of gastric biopsy specimens was obtained by using an EG Diagnostic Panel (EGDP) comprising a set of 48 gastric transcripts (including 2 housekeeping transcripts, see Table E1 in this article’s Online Repository at www.jacionline.org). In addition to the significantly dysregulated gene transcripts highly reproducible by using microarrays and RNA sequencing, we considered the magnitude of dysregulation, the direction of dysregulation, and cell-type tissue origins to optimize the diagnostic algorithm. Moreover, we aimed to reveal the presence and function of different biological processes and cell types that are known to be involved in gastrointestinal T_{h2} allergic disorders, so that we would be able to diagnose EG using a personalized medicine approach.

TaqMan reagents for amplification of major EG signature genes were obtained from Applied Biosystems (Foster City, Calif), and TaqMan real-time PCR amplification was performed on the Quant Studio 7 (Life Technologies, Grand Island, NY). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an expression control for all analyzed genes. Samples with a GAPDH value of less than 30 cycle threshold (CT) value were considered.

Abbreviations used
AUC: Area under the curve
CCHMC: Cincinnati Children’s Hospital Medical Center
CEGIR: Consortium of Eosinophilic Gastrointestinal Disease Researchers
CT: Cycle threshold
EG: Eosinophilic gastritis
EGDP: EG Diagnostic Panel
EGID: Eosinophilic gastrointestinal disorder
EoE: Eosinophilic esophagitis
FDR: False discovery rate
GAPDH: Glyceraldehyde-3-phosphate dehydrogenase
hpf: High-power field
IL-4R: IL-4 receptor
ROC: Receiver operating characteristic
TARC: Thymus and activation–regulated chemokine
TSLP: Thymic stromal lymphopoietin
FIG 1. Schematic illustration of the study and genome-wide screening of gastric tissue for identifying representative biomarkers. A, Flow chart of the study design and strategy. B, Gastric transcriptomic data on
acceptable for analysis. The expression CT value of the housekeeping gene GAPDH was subtracted from each EG gene-of-interest CT value to determine the ΔCT. The EGDp score was calculated by summing ΔCT values of the most highly dysregulated genes (ΔΔCT), as described previously and expanded upon later.

Blood-based diagnostic platform

Peripheral blood samples were collected before endoscopy and separated into serum, plasma, or both, and aliquots were frozen and stored at −80°C. The major difference between plasma and serum is the depletion of coagulation components present in the blood, a process that might alter detection of cytokine profiles in the blood. Levels of blood cytokines/chemokines, eotaxin-1 (CCL11), eotaxin-2 (CCL24), eotaxin-3 (CCL26), IL-1α, IL-4, IL-5, IL-13, IL-33, thymus and activation-regulated chemokine/chemokine ligand 17 (TARC/CCL17), and thymic stromal lymphopoietin (TSLP) were assayed by using customized immunoassays quantified on a Sector Imager 6000 (Meso Scale Discovery, Gaithersburg, Md), according to the manufacturer’s instructions. Scoring systems for plasma and serum were established separately based on their respective dysregulated biomarker levels. One point was added to a score for each instance of upregulation of specific cytokines (upregulation is defined as having a value greater than the cutoff value based on the comparison between active EG and non-EG [cutoff values—plasma: eotaxin-3, >168 pg/mL; IL-5, >1.4 pg/mL; and TARC, >87 pg/mL; serum: eotaxin-3, >32 pg/mL; IL-5, >1.5 pg/mL; and TSLP, >6.7 pg/mL]). A plasma or serum EG score is the sum of the assigned scores for each biomarker assessed, ranging from 0 to 3.

Endoscopic and histologic data

Endoscopic features were prospectively recorded in real time by using a classification and grading system specifically developed for EG by CEGIR, including erosion/ulceration, granularity, raised lesion nodularity, erythema, friability/blending, and fold thickness. These findings were scored (erosion/ulceration, 0-6; granularity, 0-2; raised lesion nodularity, 0-2; erythema, 0-2; friability/blending, 0-2; and fold thickness, 0-1) for the fundus, body, and antrum. The total score for each feature was calculated as the sum of the scores for the 3 anatomic sites. The overall global assessment of endoscopic severity for the stomach, ranging from 0 to 10, was evaluated for each patient.

Gastric biopsy specimens were assessed by peak eosinophil counts and the histologic features of EG. These features included lamina propria eosinophil sheets, periglandular circumferential collars, eosinophils in surface epithelium, eosinophil glandulitis, eosinophil gland abscesses, eosinophils in muscularis mucosa, lamina propria fibroplasia, lamina propria smooth muscle hyperplasia, reactive epithelial changes, acute inflammatory cells, and surface erosion/ulceration (see Table E2 in this article’s Online Repository at www.jacionline.org). Each feature was scored by using a 3-point scale (0, absent; 1, mild/moderate; and 2, marked). The maximum score for each biopsy specimen is 22 because it is the summation of the 11 feature scores if each feature were to be scored at 2. The final score of histologic severity, which ranges from 0 to 1, is the ratio of the sum of the assigned scores for each evaluated feature (0-22) divided by the maximum possible score for that biopsy specimen (22).

For example, if all 11 features have maximum scores of 2, the final score is 22/22 = 1. If a feature were not evaluated, the maximum possible score was reduced by 2 per feature not evaluated; most score reductions occurred because the gastric muscularis mucosa was not present in the biopsy specimen.

Statistical analysis

Statistical analyses were performed with JMP (version 13.1; SAS Institute, Cary, NC), CLC Genomics Workbench (CLC bio), GeneSpring GX 12.6 (Agilent Technologies, Santa Clara, Calif), and GraphPad Prism 8 (GraphPad Software, San Diego, Calif) software. Data are presented as numbers (percentages) or medians (interquartile ranges), unless otherwise stated. Missing data were excluded from all formal statistical analyses.

The sample size was estimated to provide 90% power to conclude that 0.8 of the area under the curve (AUC) for an individual marker was significantly greater than 0.5 while controlling the type I error rate at 1% for multiplicity. Statistical significance comparing 2 different groups was determined by using the Mann-Whitney U test (nonparametric test, 2 groups), the Kruskal-Wallis test followed by a Dunn multiple-comparison test (nonparametric test, 3 groups), or a paired t test (for quantification of longitudinal data and different site of biopsy specimens in the same patients). Nonparametric correlations were calculated by using Spearman correlations. Receiver operating characteristic (ROC) curves were constructed, and AUCs were calculated to determine the utility of the developed platforms for distinguishing patients with EG from control subjects without EG. The optimal cutoff points were determined by using the Youden index. We measured sensitivity, specificity, positive predictive value, and negative predictive value to assess the diagnostic accuracy of the scoring systems. A significant P value was defined as less than .05.

RESULTS

Patients’ characteristics

A total of 185 subjects (patients with EG, n = 74; control subjects without EG, n = 111) provided 201 gastric biopsy specimens (RNA sequencing, n = 21; EGDp discovery, n = 104; EGDp validation, n = 76) and 155 blood samples (plasma, n = 81; serum, n = 74) for analyses. A flow chart of the analysis is shown in Fig 1A. Demographic and clinical characteristics of the study cohort and subsets, including patients with EG and clinically relevant control subjects without EG, are detailed in Table 1.

Among all of the study subjects, age ranged from 1 to 67 years, with 124 (67%) pediatric and 61 (33%) adult subjects. There was a similar proportion of both sexes, with 90 male (48.6%) and 95 female (51.4%) subjects; the majority of subjects were white (91.4%). Peak gastric eosinophil counts ranged from 0 to 352 eosinophils/hpf (patients with active EG, 36-352 eosinophils/hpf; patients with inactive EG, 2-29 eosinophils/hpf; control subjects without EG, 0-28 eosinophils/hpf). Control subjects without EG (n = 111) included patients with atopic comorbidities (n = 47 [42.3%]), chronic gastritis (n = 44 [39.6%]), and active EoE without EG (n = 20 [18.0%]).

There were no significantly different baseline demographic features among the cohorts for the tissue-based (n = 124) and blood-based (n = 108) platforms. Focusing on patients with EG (n = 74), 46 (62%) had concurrent eosinophilia in the esophagus, 2 (3%) had concurrent eosinophilia in the colon, and 3 (4%) had concurrent eosinophilia in both the esophagus and colon. In the tissue- and blood-based platforms active EG
did not reveal any significant differences from control subjects without EG in age, sex, race, atopic status, or proton pump inhibitor therapy at the time of biopsy, whereas patients with active EG had significantly greater levels in the disease parameters (peak/average gastric eosinophil counts, endoscopic severity, and histologic severity; \( P < .01 \), respectively) and a greater rate of treatment (ongoing diet therapy, topical steroid therapy, and systemic steroid therapy).

**RNA sequencing of gastric tissue for identifying representative biomarkers**

To obtain the molecular foundation for developing the EGDP, we aimed to identify a gene set that was conserved across multiple experimental platforms. Accordingly, we generated an RNA sequencing data set of gastric tissue from patients with active EG (\( n = 9 \)) and control subjects without EG (\( n = 12 \)). Unsupervised principal component analysis demonstrated robust
FIG 2. Development of a tissue-based platform (EGDP) and EGDP_18 score on the basis of differentially expressed genes. A, Heat map (yellow, upregulated; blue, downregulated) based on the 18 core genes (FDR $P < .01$ and fold change $\geq 10$-fold change) in the discovery cohort. B, Three-dimensional presentation using principal component analysis between samples based on the 18 core genes (blue, control subjects without EG; red, patients with EG). C, Comparison of the EGDP_18 score between patients with EG and control subjects without EG in the discovery and validation cohort. D, ROC curve analysis showing the utility of the EGDP_18 score for the diagnosis of EG. E, Correlation between peak gastric eosinophil counts and EGDP_18 scores. F, Longitudinal changes of peak gastric eosinophil counts and EGDP_18 scores in patients with EG at active and inactive states. G, Correlation of EGDP_18 scores between the gastric antrum and body mucosa from the same subjects. H, EGDP_18 score as a function of different patient groups, including patients with EG with involvement of 1 to 5 hpf. NPV, Negative predictive value; PCA, principal component analysis; PPV, positive predictive value.
FIG 3. Gastric transcript associations with histologic and endoscopic features. A, Associations between individual genes of EGDP and diagnostic parameters. Negative log_{10} FDR P value of Spearman correlation.
Development of tissue-based platform (EGDP) and EGDP$_{18}$ scores based on differentially expressed genes

We manually curated the EG transcriptome with the aim of selecting 48 informative genes that could be embedded into a multiplex PCR-based panel for serially diagnosing and probing clinical samples. Accordingly, an EG Diagnostic Panel (EGDP) was generated based on the following considerations: dysregulation between patients with EG and control subjects without EG defined by $P$ values and fold changes, bidirectional changes of gene expression, and inclusion of genes in pathways that were likely to be involved, such as type 2 immunity. The major functional categories represented included those associated with antimicrobial defense, cell adhesion, cytokines and chemokines, epithelial-to-mesenchymal transition, hypoxia, eosinophils, epithelium, fibrosis, inflammatory response, ion transportation, mast cells, neurosensory, neutrophils, and stomach-related processes (see Table E1).

Using this set of 48 informative genes, we aimed to determine the minimal number that would successfully distinguish patients with active EG ($n = 21$) from control subjects ($n = 23$) in a discovery cohort (see Table E3 in this article’s Online Repository at www.jacionline.org). Using relatively stringent criteria ($\geq 10$-fold change, $FDR P < 0.05$), 18 differentially expressed genes completely separated the 2 groups (Fig 2, A and B). Among the 18 genes, 8 were upregulated genes related to cytokines/chemokines ($CCL26$, $CCL18$, $ILI3RA2$, and $IL5$), eosinophilia ($CLC$), cell adhesion ($CDH26$), antimicrobial defense ($KLK7$), and the epithelium ($MUC4$), and 10 were downregulated genes related to antimicrobial defense ($DEFB1$), fibrosis ($BMP3$ and $COL2A1$), ion transportation ($SLC26A7$), neurosensory activity ($GABRA1$, $GLDN$, $NPY$, and $TAC1$), and stomach-related processes ($ATP4A$ and $SST$).

With the goal of developing a quantitative diagnostic cutoff, the EGDP$_{18}$ score was developed to distinguish patients with EG versus control subjects without EG and to quantify the severity of EG. On the basis of the 18 significant and reproducible differential genes, we made CT sums of the upregulated genes and downregulated genes separately and then combined the 2 sums considering their different direction of dysregulation. The EGDP$_{18}$ score was significantly decreased in patients with active EG compared with that in control subjects without EG in the discovery cohort ($P < 0.0001$) and similarly decreased in the validation cohort ($P < 0.0001$; Fig 2, C). ROC analysis demonstrated an excellent diagnostic merit ($P < 0.0001$, AUC $\geq 0.95$) in both cohorts (Fig 2, D). After investigation by setting optimal cutoff points, a score of less than 0 resulted in a positive predictive value of 100% and a negative predictive value of greater than 94% (Fig 2, D). Of note, the EGDP$_{18}$ score is inversely correlated with disease severity, as defined by eosinophil counts when analyzed cross-sectionally ($r = -0.83$, $P < 0.0001$; Fig 2, E) and longitudinally ($P = 0.0078$; Fig 2, F). The EGDP$_{18}$ score showed comparable levels and high correlation between the gastric antrum and body ($n = 8$, $r = 0.85$, $P < 0.0001$; Fig 2, G). Among patients with active EG, the EGDP$_{18}$ score showed consistency across geographically diverse sites (see Fig E2 in this article’s Online Repository at www.jacionline.org) and comparable levels across ages (pediatric vs adult patients), atopic status (atopy vs no atopy), coexistence with EoE (EG only vs EG with EoE), and treatment status at biopsy (ongoing therapy including diet and steroids vs no therapy, see Fig E3 in this article’s Online Repository at www.jacionline.org).

Interestingly, the EGDP$_{18}$ score was able to classify patients with intermediate tissue eosinophil counts (ie, the number of hpf with $\geq 30$ eosinophils, $n = 1$–4 hpf). When these patients were analyzed by using the EGDP$_{18}$ score ($n = 8$, all of them were clinically symptomatic), 5 (63%) patients were molecularly equivalent to having active EG (Fig 2, H, and see Fig E4 in this article’s Online Repository at www.jacionline.org).

Gastric transcript associations with histologic and endoscopic features

Significant correlations were noted between specific genes within the EGDP and the peak gastric eosinophil level, histologic severity, and overall global assessment of endoscopic severity (Fig 3, A, and see Table E4 in this article’s Online Repository at www.jacionline.org). The top 10 genes that tracked with tissue eosinophilia were $CCL26$, $CLC$, $ILI3RA2$, $BMP3$, $IL5$, $CDH26$, $CCL18$, $NPY$, $HPGDS$, and $SST$; those that tracked with histologic severity were $CCL26$, $ILI3RA2$, $CLC$, $SST$, $BMP3$, $IL5$, $CDH26$, $GLDN$, $ANXA1$, and $DEFB1$; and those that tracked with endoscopic severity were $CCL26$, $GLDN$, $ILI3RA2$, $SST$, $DEFB1$, $GABRA1$, $IL5$, $TAC1$, $CLC$, and $ILI3$. Notably, these gene groups...
FIG 4. Development of blood-based platforms by using a multiplex protein array. A and B, Among the 10 biomarkers embedded in the platform, a statistical screening was performed between the control subjects...
included genes that overlapped between tissue eosinophilia, histologic severity, and endoscopic severity (ie, CCL26, CLC, IL13RA2, IL5, and SST).

Individual components of the histologic and endoscopic features associated with the EGDP18 score. Associations were noted between the EGDP18 score and several histologic features (Spearman \( r \) range = −0.05 to −0.73; Fig 3, B, left panel), with periglandular circumferential collars showing the greatest magnitude of correlation with the EGDP18 score (\( r = −0.73, \) FDR \( P < .0001 \)). Associations were also observed between the EGDP18 score and several endoscopic features (Spearman \( r \) range = −0.22 to −0.45; Fig 3, B, right panel). The EGDP18 score inversely correlated the most with granularity (\( r = −0.45, \) FDR \( P < .0001 \)) and nodularity (\( r = −0.45, \) FDR \( P < .0001 \)).

Three histologic features (lamina propria eosinophil sheets, periglandular circumferential collars, and eosinophil glandulitis) showed higher correlations based on hierarchical clustering of Spearman correlations (Fig 3, C, left panel), suggesting that they might be more effective than other features at capturing biological processes underlying the EGDP. At the gene level, CCL26 showed the strongest correlation with histologic features, most notably periglandular circumferential collars (\( r = 0.74, \) \( P = 7.0E-21 \)) and eosinophil glandulitis (\( r = 0.68, \) \( P = 2.0E-16 \)); in terms of correlation strength, CCL26 was followed by IL13RA2, which correlated most notably with periglandular circumferential collars (\( r = 0.67, \) \( P = 7.0E-16 \)) and lamina propria eosinophil sheets (\( r = 0.67, \) \( P = 5.0E-16 \)). Interestingly, although mucosal muscosa eosinophilia showed relatively weak associations compared with epithelial and lamina propria changes, DUOX2 and DUOX2A showed unique association with mucosal muscosa eosinophilia (\( r = 0.32, \) \( P = 3.8E-3; \) see Fig E5 and Table E5 in this article’s Online Repository at www.jacionline.org). Some features, such as acute inflammation and erosion/ulcer, were uncommon and, possibly for that reason, did not show gene correlations.

In contrast to histology, for which gene transcripts showed association with only a limited set of histologic features, all recorded endoscopic features correlated with specific gastric transcripts (Fig 3, C, right panel). CCL26 showed the strongest correlation with any endoscopic features, most notably nodularity (\( r = 0.55, \) \( P = 2.8E-5 \)) and granularity (\( r = 0.53, \) \( P = 4.9E-5 \)), followed by IL33, which inversely correlated most notably with granularity (\( r = −0.46, \) \( P = 6.1E-4 \)) and friability and bleeding (\( r = −0.39, \) \( P = 3.9E-5 \)). Interestingly, clustering separated endoscopic features into 2 general groups, one was associated with endoscopic changes, including friability/bleeding and erythema, and the other was associated with endoscopic changes including nodularity and granularity. Endoscopic changes, including friability/bleeding and erythema, were associated with downregulation of molecular signatures (ATP4A, IL33, and SLC26A7), whereas endoscopic changes, including nodularity and granularity, were associated with upregulation of type 2 immunity and eosinophil-associated pathways (CCL26, IL13RA2, and IL5; see Fig E6 and Table E6 in this article’s Online Repository at www.jacionline.org). A heat map showing \(-\log_{10} \) FDR \( P \) values determined by differential expression between patients with EG with or without a specific endoscopic feature by using the Mann-Whitney \( U \) test also supported this finding (see Fig E7 in this article’s Online Repository at www.jacionline.org).

Development of blood-based platforms and blood EG scores based on significantly increased biomarker levels

We explored the possibility that systemic levels of cytokines/chemokines might be increased in patients with EG. Focusing on plasma and serum samples from patient cohorts with and without EG, we designed a multiplex immunoassay containing 10 EG-relevant cytokines/chemokines, particularly those based on type 2 immunity, as reflected in the functional predictions found in the EG transcripts (Fig 1, E). Notably, patients with active EG showed significantly greater levels of 3 cytokines in the plasma and 3 cytokines in the serum (plasma eotaxin-3/CCL26, IL-5, and TARC/CCL17 and serum TSLP, eotaxin-3/CCL26, and IL-5, respectively; Fig 4, A and B, and see Table E7 in this article’s Online Repository at www.jacionline.org), suggesting that the activity of the disease consistently affects these cytokines systemically.

On the basis of the levels of these dysregulated cytokines and chemokines, we developed a circulation-based EG biomarker scoring system for plasma and serum (see Fig E8 in this article’s Online Repository at www.jacionline.org). The blood-based EG score differentiated patients with active EG from control subjects without EG in both the plasma and serum cohorts (\( P < .0001; \) Fig 4, C and D). Notably, patients with active EG had significantly higher scores than did patients with inactive EG (plasma EG score: \( P < .0001; \) serum EG score: \( P = .0012; \) Fig 4, E and F).

To determine their diagnostic performances, ROC analyses were constructed to investigate the use of blood EG scores and cytokine/chemokine levels alone (Fig 4, G and H). The plasma EG score yielded an AUC of 0.93, whereas levels of eotaxin-3 alone yielded an AUC of 0.89, levels of TARC alone yielded an AUC of 0.82, and levels of IL-5 alone yielded an AUC of 0.80. The serum EG score yielded an AUC of 0.91, whereas levels of TSLP alone yielded an AUC of 0.86, levels of eotaxin-3 alone yielded an AUC of 0.80, and levels of IL-5 alone yielded an AUC of 0.77.

Associations among local and systemic molecular expressions

We were interested in exploring the association between tissue local gene expressions and circulating systemic molecular expressions. Notably, plasma eotaxin-3 exhibited a greater magnitude of correlation with the EGDP than any other protein (\( P < .01; \) Fig 5, A, upper panel). By using Spearman \( r \) for the correlation between the EGDP gene expressions and plasma and...
FIG 5. Associations among local and systemic features. A, Associations between EGDP and blood cytokine/chemokine levels. Using Spearman r for the correlation between the EGDP gene expressions and plasma cytokine levels. B, Correlation between EGDP and peak gastric eosinophil counts. Spearman r = 0.72, P = .0002. C, Correlation between EGDP and EGDP score. Spearman r = -0.64, P = .0015.
serum protein biomarkers, a Spearman $r$-based heat diagram for the correlation at the gene level was generated (Fig 5, A, lower panel). Focusing on plasma eotaxin-3, we observed that plasma eotaxin-3 levels correlated with genes related to gastric cytokine/chemokines (CCL26, IL13RA2, IL1RL1, IL4, and IL5), eosinophilia (CLC and CCR3), cell adhesion (CDH26), mast cells (CPA3 and HPGD), inflammatory response (ANXA1), “other” (ITLN1), neurosensory features (GLDN), fibrosis (BMP3), and antimicrobial defense (DEFB1).

The blood EG score (the circulation-based biomarker) showed significant correlations with gastric eosinophil counts (plasma: $r = 0.72, P = .0002$; serum: $r = 0.54, P = .0015$; Fig 5, B) and EGDP$_{18}$ scores (plasma: $r = -0.64, P = .0015$; serum: $r = -0.46, P = .0084$; Fig 5, C), suggesting that systemic circulating biomarker levels reflect the local gastric inflammatory process defined by eosinophil levels and transcript expression profiles.

**DISCUSSION**

In this study, we have established molecular diagnostic criteria for EG using gastric mRNA transcript and circulating protein levels. We (1) developed a set of EG transcripts composed of 48 genes that robustly distinguishes patients with EG from control subjects without EG whose expression is equivalent across independent and geographically diverse sites and across patient ages (including children and adults) and medical therapy; (2) determined that a diagnostic score limited to changes in a subset of 18 genes, referred to as the EGDP$_{18}$ score, is sufficient to allow EG diagnosis relative to control subjects (sensitivity of 88% to 95% in the discovery and validation cohort and specificity of 100%), including control subjects without EG and patients with EGID limited to the esophagus; (3) determined that the EGDP$_{18}$ score can robustly separate patients with active EG from those with inactive EG, strongly correlates with gastric eosinophil levels ($r = -0.83, P < .0001$), and potentially aids in diagnostic classification of patients with intermediate eosinophil levels; (4) determined that expression of specific gene tracks with tissue eosinophilia, namely CCL26, CLC, IL13RA2, BMAP3, IL5, CDH26, CCL18, NPY, HPGD, and SST; (5) linked the magnitude of molecular changes to endoscopic changes, most notably associating nodularity and granularity with a subset of type 2 inflammatory genes, including CCL26 and IL13RA2, respectively; (6) linked the magnitude of molecular changes to histologic changes, with CCL26 levels most notably strongly associated with periglandular circumferential collars ($r = 0.74, P = 7.0E-21$) and eosinophil glandulitis ($r = 0.68, P = 2.0E-16$), whereas IL13RA2 correlated most notably with periglandular circumferential collars ($r = 0.67, P = 7.0E-16$) and lamina propria eosinophil sheets ($r = 0.67, P = 5.0E-16$); (7) identified circulating biomarkers that reflect local changes in the stomach, most notably the gastric eosinophilia; and (8) demonstrated that combined levels of plasma eotaxin-3, TARC, and IL-5 have the capacity to diagnose EG disease and monitor disease activity with high sensitivity and specificity (100% and 72%, respectively).

Here we analyzed more than 200 gastric tissue samples and assessed the overlap among molecular profiles. Although it is conceivable that EG and EoE share a common Th2 molecular pathogenesis, as we published earlier, the EG and EoE transcriptomes (as assessed by using microarray analysis) only overlap by 7% despite a common IL-13–induced signature. In this study, using independent methods of RNA sequencing and quantitative PCR arrays, we confirmed our previous observation that the overall gene expression profiles of EoE and EG are distinct at a transcription level. Previous histopathologic studies indicate that the eosinophilic infiltration in patients with EG can be patchy and that the minimum threshold number of gastric eosinophils required for the diagnosis of EG varies, ranging from 20 eosinophils in 1 hpf to 70 eosinophils per hpf in at least 3 hpfs. Our results were obtained by using only 1 RNA sample per patient, suggesting that molecular diagnosis is a relatively promising and sensitive method for disease diagnosis and monitoring. The EGDP$_{18}$ score algorithm indicated that 63% of the histologically intermediate patients were molecularly equivalent to patients with active EG, providing evidence that 30 eosinophils/hpf in less than 5 hpfs is still associated with robust molecular inflammatory processes. These data suggest that analysis of less than 5 hpfs might be sufficient for diagnosis.

Beyond diagnostic merits, to understand disease pathogenesis, we also assessed correlations between molecular profiles and histologic and endoscopic features. For EG histologic features, regardless of the distribution, eosinophilic features (periglandular circumferential collars, eosinophil glandulitis, lamina propria eosinophil sheets, and eosinophils in surface epithelium) were highly associated with the EG transcriptome (especially the EGDP core 18 genes), with the strongest association occurring in periglandular circumferential collars. Not all histologic features showed strong associations with the EGDP; possibly because of the insufficient depth of biopsy specimens, resulting in many of them not including muscularis mucosa and/or the low occurrence of some histologic features (lamina propria fibroplasia, surface erosion/ulceration, eosinophil gland abscess, and acute inflammatory cells). Moreover, certain features of endoscopic changes, such as nodularity and granularity, were notable as features uniquely related to transcript changes, particularly those enriched in inflammatory responses involving upregulation of type 2 immunity and eosinophil–related pathways (IL13RA2, CCL26, and IL5). Both CCL26 and IL13RA2 are IL-13–inducible genes; controversy exists as to whether the latter is an activating or possibly inhibitory signaling molecule (including a potential inhibitor role of soluble IL-13 receptor [IL-13R] α2). The prominent role of type 2 immunity–related responses provides the scientific basis for therapeutic intervention with dupilumab (anti–IL-4 receptor (left) and serum cytokine/chemokine levels (right); magnitudes of correlation with the EGDP are shown (upper). A Spearman $r$-based heat diagram for correlation at the gene level is shown (lower). Genes shown on the $y$-axis are organized within functional groupings. Darker red shades indicate stronger positive correlations, whereas darker blue shades indicate stronger negative correlations. B, Correlation between blood EG scores (left; plasma; right; serum) and peak gastric eosinophil counts, with Spearman r and P values shown. C, Correlation between blood EG scores (left; plasma; right; serum) and EGDP$_{18}$ scores, with Spearman r and P values shown. *P < .01 versus all other proteins.
α, which inhibits IL-4 receptor α/IL-13Rα1), anti–IL-13Rα1, and/or anti–IL-13 (eg, RPC4046 inhibiting IL-13 interactions with both IL-13Rα1 and IL-13Rα2). Increased expression levels of type 2 immune/eosinophil-associated pathways can be seen in patients with other atopic disorders associated with nodularity, such as chronic rhinosinusitis with nasal polyps, suggesting that they were not to be EG specific but might function to generate these histologic features in certain tissue/conditions.

Interestingly, endoscopic changes, such as friability and erythema, were associated with downregulation of IL13 (epithelium-derived, proinflammatory alarmin), SLC26A7 (anion exchange transporter), and ATP4A (proton pump; gastric H, K-ATPase alpha subunit). Decreased expression levels of these genes might suggest injured mucosa because of tissue inflammation. Our current findings showed only minimal overlap (eg, AREG, CXCL8, SST, and TGFBR1) compared with a prior report limited to 8 patients with EG from a single site, probably because of the differences in sample size, molecular platform, and definition of endoscopic features. However, in this large cross-sectional cohort of patients with EG, we could identify specific findings with differences in potential pathways.

Biopsy specimen procurement is currently required to establish a definitive diagnosis of EGID. The field urgently calls for developing noninvasive biomarkers. Prior findings suggested that EG is more systemic than EoE based on the co-occurrence of EG with circulating eosinophilia; therefore, we hypothesized that circulating biomarkers might be present in patients with EG. Indeed, eotaxin-3 and IL-5 were significantly upregulated in both serum and plasma of patients with EG compared with control subjects without EG, and circulating levels of eotaxin-3 were particularly coregulated with tissue expression of CCL26. Of note, plasma eotaxin-3 levels in patients with EG were not reflective of an atopic state in general because they were not increased in patients with EoE. Furthermore, average circulating eotaxin-3 levels in patients with EG appear to be substantially greater than levels reported in patients with other atopic diseases, such as chronic rhinosinusitis with high-eosinophil-count mucosal infiltration (plasma: 122.6 vs 481.2 pg/mL seen in our study) but in the same range for serum as seen for another rare eosinophilic disease, eosinophilic granulomatosis with polyangiitis. We speculate that the stomach (related disease EG) might contribute to higher circulating eotaxin-3 levels than does the esophagus (related disease EoE) because of differences in the underlying tissue architecture, with the gastric mucosa having relatively increased proximity to the vasculature, a relatively large surface area, and resident eosinophil populations during homeostasis. The reasons for the selective increase in circulating eotaxin-3 levels in patients with EG compared with other EGIDs deserves further attention.

To our knowledge, this is the first EGID study simultaneously addressing tissue signatures and circulating cytokine profiles in the same disorder with autologous samples across different collecting centers. This study was not intended to replace the histologic method but rather to provide at least 2 alternative platforms to more precisely and sensitively diagnose EG. It is conceivable that the circulating markers could serve as an early noninvasive test during EGID/EG screening, whereas the tissue signature profiling (EGDP) could be used for definitive diagnostic confirmation. The combination of both would provide molecular tools to diagnose, monitor, and potentially further subtype (eg, endotype) knowledge of EG. Future studies should examine the utility of the blood-based platform to identify disease remission with treatment, which would prevent the need for repeat endoscopy.

Our study has several strengths. First, we analyzed samples from multiple sites across the United States, which increases the generalizability of the results. Second, participants were assessed with several diagnostic assessments, allowing us to examine associations between gene expression and endoscopic and histologic parameters. Third, we assessed not only gene expression, but also circulating blood protein levels. Fourth, we validated gene expression differences between EG diagnoses in an independent cohort.

Our study also has limitations. First, our findings include patients with active EG with mixed treatment status or who have disease that is refractory to treatment, which might influence the results. However, patients still exhibited signs of disease clinically, histologically, and molecularly.

Second, most of the analyses for gene and biomarker expression were restricted to 48 genes included in the EGDP and 10 blood biomarkers. Unbiased, genome-wide transcriptome and proteome approaches would likely reveal additional genes of interest, biomarkers, and optimal combinations.

Finally, data are limited by the cross-sectional approach, highlighting the importance of additional replication, particularly in prospective and longitudinal studies.

In conclusion, we have developed and validated diagnostic panels that can diagnose EG by using biopsy and blood samples. CCL26/eotaxin-3 emerged as the strongest single tissue and circulating disease biomarker. We have uncovered robust associations among the EG molecular profile, periglandular circumferential collars, and endoscopic granularity/nodularity, providing insight into the better understanding of the pathogenesis for EG. Further work is required to apply these platforms to a prospective trial in a different clinical setting, explore the feasibility and further validation, and optimize platforms for disease stratification.

We thank all the patients who participated in the study. We are also grateful to their colleagues and clinical support staff for procuring biopsy specimens, blood samples, and clinical data. Shawna Hottinger provided editorial assistance as a medical writer funded by CCHMC.

Clinical implications: We have developed tissue- and blood-based platforms for diagnosing and monitoring EG and uncovered likely molecular pathogenesis that accounts for the distinct endoscopic and histologic features of the disease.

REFERENCES


