Analysis and expansion of the eosinophilic esophagitis transcriptome by RNA sequencing

JD Sherrill, K KC, C Blanchard, EM Stucke, KA Kemme, MH Collins, JP Abonia, PE Putnam, VA Mukkada, A Kaul, SA Kocoshis, JP Kushner, AJ Plassard, RA Korns, PJ Dexheimer, BJ Aronow and ME Rothenberg

Eosinophilic esophagitis (EoE) is an allergic inflammatory disorder of the esophagus that is compounded by genetic predisposition and hypersensitivity to environmental antigens. Using high-density oligonucleotide expression chips, a disease-specific esophageal transcript signature was identified and was shown to be largely reversible with therapy. In an effort to expand the molecular signature of EoE, we performed RNA sequencing on esophageal biopsies from healthy controls and patients with active EoE and identified a total of 1607 significantly dysregulated transcripts (1096 upregulated, 511 downregulated). When clustered by raw expression levels, an abundance of immune cell-specific transcripts are highly induced in EoE but expressed at low (or undetectable) levels in healthy controls. Moreover, 66% of the gene signature identified by RNA sequencing was previously unrecognized in the EoE transcript signature by microarray-based expression profiling and included several long non-coding RNAs (lncRNA), an emerging class of transcriptional regulators. The lncRNA BRAF-activated non-protein coding RNA (BANCR) was upregulated in EoE and induced in interleukin-13 (IL-13)-treated primary esophageal epithelial cells. Repression of BANCR significantly altered the expression of IL-13–induced proinflammatory genes. Together, these data comprise new potential biomarkers of EoE and demonstrate a novel role for lncRNAs in EoE and IL-13–associated responses.

Original Article

INTRODUCTION

Eosinophilic esophagitis (EoE) is a complex food hypersensitivity disorder with an underlying genetic basis that is characterized by intense allergic inflammation within the esophageal mucosa. The symptomology of EoE can progress from feeding difficulties and vomiting to dysphagia and food impactions, underscoring the chronicity of the disease. Moreover, as the prevalence of EoE is increasing on a global scale, with current estimates at 4 in 10,000, further investigation into molecular markers that will facilitate disease diagnosis and treatment is warranted.

EoE presents a unique opportunity to probe markers of allergic inflammation as tissue biopsies acquired during standard-of-care endoscopy can be used for genome-wide analyses. The first detailed molecular analysis of the pathogenic mechanisms in EoE assessed esophageal expression profiles from patient biopsies using probe-based microarrays. An EoE-specific esophageal transcript signature that was highly conserved across patient age, sex and atopic history differentiated EoE patients from control individuals, as well as non-EoE individuals with chronic esophagitis. In particular, 574 transcripts (344 upregulated, 230 downregulated) were differentially regulated in EoE patients compared with healthy controls. Many of these dysregulated transcripts affected epithelial cell function or immune cell activation, with the transcript exhibiting the greatest change (53-fold) being chemokine (C–C motif) ligand 26 (CCL26), which encodes the eosinophil chemoattractant eotaxin 3. Subsequent work demonstrated that the majority of the gene signature in EoE normalized after steroid-induced disease remission. Moreover, a large percentage of the EoE transcript signature (22%) could be induced by interleukin-13 (IL-13) treatment of primary esophageal epithelial cells in vitro, with CCL26 being the most upregulated gene (279-fold).

High-throughput whole-transcriptome (RNA) sequencing offers greater transcriptional resolution compared with traditional probe-based microarrays, as it generates expression profiles that are not reliant upon known transcripts and has greater dynamic range for detection of low-abundance transcripts. In the present study, we utilized RNA sequencing to expand and better define the molecular entities involved in the transcriptional programming of EoE. We observed EoE-specific upregulation of the long non-coding RNA (lncRNA) BRAF-activated non-coding RNA (BANCR), which was recapitulated by IL-13 treatment of primary esophageal epithelial cells. Notably, shRNA silencing of BANCR resulted in the altered expression of other IL-13–regulated proinflammatory genes. These data expand the previously defined EoE transcriptome to a wider transcript set, enriched in genes functionally involved in immunity, atopy and eosinophilia, highlight the ability of RNA sequencing to uncover novel molecular signatures associated with human inflammatory disease and implicate IL-13 as a novel regulator of lncRNA expression.

RESULTS

Comparing disease expression profiles from RNA sequencing and microarray

To obtain an unbiased picture of the transcriptional changes associated with EoE, we used RNA sequencing and analyzed raw
gene expression levels to identify differential transcript signatures in esophageal specimens from patients with active EoE compared with healthy (NL) controls. We identified a total of 1607 transcripts that were dysregulated in EoE ($P < 0.05$, fold change > 2.0; Figures 1a and b). Of these, 1096 genes were upregulated and 511 were downregulated compared with controls. The EoE-dysregulated genes were clustered by their raw expression values in the control samples: upregulated genes that were expressed at high (cluster 1, $n = 392$), medium (cluster 2, $n = 326$) or low (cluster 3, $n = 378$) levels in controls and downregulated genes that were expressed at high (cluster 4, $n = 182$), medium (cluster 5, $n = 155$) and low (cluster 6, $n = 174$) levels in controls. Many of the most highly dysregulated genes (for example, CCL26, ALOX15, CRISP3) were previously identified as part of the EoE transcriptome and fell within cluster 3 (Figures 1a and b). We validated the change in expression of eight representative genes in the same samples by quantitative PCR (qPCR): ALOX15, CCL26, CLC and CPA3 were significantly increased (Figure 1c), whereas SPINK7, SPRR2D, SPRR2B and S100A6 were significantly decreased in EoE (Figure 1d).

### Table: Upregulated and Downregulated Genes in EoE

| Gene Symbol | FC  | P-value |
|-------------|-----|---------|
| CD9         | 3.7 | 1.1E-05 |
| SERPINB3    | 2.5 | 1.4E-02 |
| CAPN14      | 6.0 | 9.8E-04 |
| SERPINB2    | 2.1 | 9.9E-03 |
| AK023178    | 17.9| 1.5E-06 |
| FIGNL1      | 2.1 | 1.4E-04 |
| ARRB2       | 2.6 | 1.3E-04 |
| TMEM63A     | 2.1 | 9.0E-04 |

| Gene Symbol | FC  | P-value |
|-------------|-----|---------|
| ALOX15      | 4.5E-05 | 2.3E-04 |
| POSTN       | 2.3E-05 | 1.4E-04 |
| IGJ         | 3.6E-04 | 5.7E-04 |

| Gene Symbol | FC  | P-value |
|-------------|-----|---------|
| FAM25A      | 2.9E-04 | 2.8E-05 |
| SFTA2       | 2.8E-05 | 1.4E-05 |
| HCG22       | 1.7E-05 | 9.9E-05 |
| MUC21       | 1.3E-05 | 1.3E-04 |

**Figure 1.** Differential gene expression in EoE identified by RNA sequencing. Heatmap showing log2 FPKM values and clustering for 1607 genes identified as dysregulated in EoE (1096 upregulated, 511 downregulated; $P < 0.05$, absolute fold change > 2.0) by RNA sequencing; each column represents an individual patient or control (a). The five most highly expressed genes within each cluster, fold change in EoE, and associated P-value are indicated (right). Volcano plot showing log2 fold-change values (x axis) by −log10 corrected P-values (y axis) for all genes. Genes that are significantly altered ($P < 0.05$, > twofold change) are indicated as red circles. Green lines represent the thresholds used for $P$-value ($P < 0.05$) and fold change (> 2.0) (b). qPCR validation (mean ± interquartile range) of four upregulated (c) and four downregulated (d) candidate genes identified by RNA sequencing. Network displaying enriched biological pathways for gene clusters 1–3 from Figure 1a (e). **$P < 10^{-2}$, ***$P < 5 	imes 10^{-3}$, NL, healthy controls; no., number.
Focusing on the induced genes as potential immunomodulators or immune cell-specific genes within the inflamed esophageal microenvironment, we performed gene enrichment analysis on clusters 1–3 (Figure 1e). Although broad immunological processes were shared across all three clusters, such as immune response (GO:0006955) and immune effector process (GO: 0002252), which were the two most significantly associated biological processes, certain cell-specific functions fell within unique expression clusters. For instance, cluster 1 contained highly expressed genes regulating major histocompatibility complex peptide binding and antigen recognition, whereas cluster 3 contained low-expressed genes involved in immune cell (lymphocytes, mast cells and eosinophils) activation and migration.

We next compared the differential gene signature from RNA sequencing with that identified in a separate cohort of patients by standard microarray. Updated microarray analyses identified a total of 870 dysregulated transcripts in EoE (compared with 574 transcripts as previously reported) with 374 and 496 being upregulated and downregulated, respectively, compared with controls. To compare the differentially expressed genes identified from the RNA sequencing and microarray analyses, we intersected Entrez gene IDs from both datasets and found a substantial overlap (n = 284) in the upregulated genes common to both datasets; notably, this overlap corresponded to 76% and 27% of the total number of upregulated genes identified by microarray and RNA sequencing, respectively (Figure 2a). Comparing the relative fold changes of these 284 upregulated genes between platforms demonstrated a significant correlation (Spearman r = 0.66, P < 10^{-5}; Figure 2b). Similarly, a substantial overlap in downregulated transcripts was observed, with 236 genes common to both the microarray and RNA sequencing profiles. Interestingly, this overlap represented only 48% of both the total downregulated genes from the microarray profile and those from the RNA sequencing profile (Figure 2c). The relative fold changes of the shared downregulated transcripts displayed a significant correlation (Spearman r = 0.48, P < 10^{-5}) between the microarray and RNA sequencing datasets (Figure 2d).

Figure 1. Continued.

Novel EoE transcriptome genes from RNA sequencing
RNA sequencing identified 1028 genes (770 upregulated, 258 downregulated) that were previously unrecognized as part of the EoE transcriptome (Figure 3a). Notably, the expression of the majority of these genes was modestly changed with an absolute fold change of 2.4 (2.2–2.9) (median (interquartile range)), highlighting the ability of RNA sequencing to identify subtle changes (fold change ~2) in the expression of genes that are potentially relevant in disease pathogenesis. However, the expression of a few genes was dramatically altered in EoE, including 42 transcripts with greater than five-fold change (median absolute fold change (interquartile range) = 7.0 (5.7–10.0); Figure 3b). The most highly induced gene unique to the RNA sequencing profile was solute carrier family 9, subfamily A, member 3 (SLC9A3), which was induced 33-fold in EoE. SLC9A3

© 2014 Macmillan Publishers Limited
Genes and Immunity (2014) 361 – 369
encodes the sodium-hydrogen exchanger family member 3 (NH3), and although its function in the esophageal mucosa remains unknown, functional studies in the intestine have shown that NH3 regulates epithelial absorption of NaCl and HCO₃⁻. Expression of chemokine (C–C motif) ligand 24 (CCL24), which encodes the eotaxin family member eotaxin 2, was also identified as being upregulated greater than 15-fold. Notably, qPCR analysis demonstrated that SLC9A3 and CCL24 were significantly upregulated in a separate cohort of EoE patients but not in patients with gastroesophageal reflux disease (GERD) when compared with healthy controls, indicating disease-specific dysregulation (Figure 3c). The most highly downregulated gene unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

Role of IncRNAs in EoE

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

Figure 2. Comparison between EoE transcriptome genes identified by RNA sequencing and microarray. Venn diagrams comparing the number of genes (with Entrez IDs) identified as upregulated (a) and downregulated (c) in EoE across both platforms. Spearman correlation comparing absolute fold-change values for the overlapping upregulated (b) and downregulated (d) genes identified as dysregulated by both platforms. RNA-seq, RNA sequencing.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

Role of IncRNAs in EoE

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.

In an effort to find novel disease biomarkers for EoE, we explored the RNA sequencing profile in EoE for uncharacterized transcripts exhibiting dysregulation. We focused on IncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states. We observed a significant upregulation of the RNA sequencing profile in EoE for uncharacterized transcripts unique to the RNA sequencing profile was family with sequence similarity 25, member A (FAM25A), an uncharacterized gene that was repressed 50-fold in EoE.
DISCUSSION

In the present study, we used RNA sequencing to provide an expanded view of the transcriptional changes occurring within the inflamed esophageal mucosa of patients with EoE. Our data have demonstrated that a large proportion of the RNA sequencing profile (66% or 1028 transcripts) associated with EoE was not identified by traditional microarray analyses. Thus far, expression of non-coding RNAs and their role in EoE has been limited to microRNAs. However, we show that Th2 cytokines (IL-13) are potent stimuli for IncRNA expression, implicating IncRNAs (for example, BANCR) as a novel class of non-coding RNA molecules involved in EoE.

The high degree of overlap of differentially expressed genes in EoE identified between the RNA sequencing and microarray indicates high reproducibility between the two methods; for instance, even when identical samples are analyzed by RNA sequencing and microarray, there is 81% concordance between the two methods. Moreover, the number of overlapping genes also reflects the high conservation in the EoE transcriptome, which is even more impressive given that a different number of samples from different patients were analyzed. Importantly, the microarray analyses presented herein expand upon previously published data from our group investigating the EoE transcriptome by microarray profiling. Here, we identified 870 dysregulated genes.

Figure 3. Novel EoE transcriptome genes identified by RNA sequencing. Heatmap showing log2 FPKM values and clustering for the unique 1028 dysregulated genes in EoE (770 upregulated, 258 downregulated) identified by RNA sequencing only (P<0.05, absolute fold change >2.0) (a). Heatmap showing log2 FPKM values for the 42 unique, most dysregulated genes (fold change >5) (b). qPCR analysis of novel EoE transcriptome genes CCL24, BANCR and SLC9A3 in esophageal biopsies from healthy controls (NL), patients with active EoE and patients with GERD (c). Data in c are represented as the median ± interquartile range. NS, not significant; *P<0.05; **P<5 × 10⁻²⁹; no., number of.
in EoE (374 upregulated, 496 downregulated), which was a 51% increase compared with our previous report that identified 574 dysregulated genes (344 upregulated, 230 downregulated). A likely explanation for the increased total number of genes may be related to the increased number of samples analyzed (n = 28 from the current microarray analysis compared with n = 19 in our previous publication).

The dynamic range of RNA sequencing has expanded the EoE transcriptome through detection of dysregulated genes expressed at levels below detection by microarray analysis. Indeed, the majority of the 1028 EoE transcriptome genes newly identified by RNA sequencing were modestly dysregulated in patients with EoE. For instance, 807 of these genes (78%) showed less than a threefold change with a median (interquartile range) expression level of 6.6 (3.4–16.2) FPKM. These findings reflect previous findings indicating that there is a greater correlation among the more highly expressed genes identified by both RNA sequencing and microarray. However, the minor changes in gene expression of EoE transcriptome genes newly identified by RNA sequencing may have substantial roles in EoE, especially for genes that are critical to or act synergistically within biologically relevant pathways. For instance, an abundance of the newly identified EoE transcriptome genes that are expressed at low or near-undetectable levels in the uninflamed esophagus are enriched for cell-specific immune responses; CCL24 was most notable given its role in the eosinophil chemotaxis. Although CCL24 was not significant by microarray analysis, a significant increase or a generalized upward trend in CCL24 expression in EoE has been reported here and by other groups using alternative methods.

Emerging data indicate a significant role for lncRNAs in diverse cellular processes ranging from growth and differentiation to apoptosis and inflammatory responses, and our RNA sequencing results support involvement of lncRNAs in EoE. Refining the human transcriptome through the GENCODE consortium has uncovered an abundance of lncRNAs (~14880 unique transcripts) that are transcriptionally regulated and processed in similar fashions to coding RNAs. LncRNAs can regulate gene expression through various cis or trans mechanisms, such as by scaffolding multiprotein complexes, mediating DNA–RNA interactions and acting as decoy receptors for microRNAs (reviewed in Mercer et al.).

Figure 4. Dysregulation of the lncRNA BANCR in EoE. Heatmap showing log2 FPKM values for the 36 dysregulated transcripts in EoE identified by RNA sequencing (P < 0.05, absolute fold change > 2) that correlated with BANCR expression (arrow; Spearman r = 0.9) (a). Each column in a represents an individual patient or control. qPCR analysis of BANCR levels in esophageal biopsies from healthy controls (NL) and patients with active EoE (b). Spearman correlation between esophageal levels of BANCR in patients with active EoE (from b) and esophageal eosinophil counts (c) or esophageal IL13 levels in NL controls (squares) and in patients with active EoE (circles) from b (d). Data in b are represented as the median ± interquartile range. **P < 5 × 10⁻³. Eos, eosinophils; HPF, high-power field.
BANCR was originally identified as a BRAF-induced IncRNA associated with melanoma, in which suppressing BANCR induction attenuated CXCL11 expression and reduced melanocyte migration. We demonstrate that BANCR suppression reduced the expression of CASP14 mRNA in IL-13–treated esophageal epithelial cells. CASP14 was recently shown to be downregulated after suppression of another IncRNA termed terminal differentiation–induced non-coding RNA (TINCR). TINCR was demonstrated to promote terminal differentiation of keratinocytes by stabilizing the mRNAs for genes necessary for epidermal differentiation (for example, filaggrin, loricrin) and for the processing of barrier peptides like CASP14. Notably, esophageal epithelial differentiation and barrier function have been reported to be impaired in EoE.

In summary, RNA sequencing has provided greater resolution into the genome-wide transcriptional changes associated with EoE. These RNA sequencing studies, together with our biological data, demonstrate how previously uncharacterized or lowly expressed genes/molecules (for example, BANCR) can be involved in biologically relevant disease processes (for example, IL-13 signaling and inflammatory gene expression). It remains possible that IncRNAs such as BANCR may help in the diagnosis and monitoring of EoE and its disease activity. However, despite these advances in disease-associated transcriptomics, the full power of RNA sequencing in EoE remains untapped as further in-depth analysis of alternative splicing events, novel isoform expression and the effect of genetic variants on gene expression (that is, allele-specific expression) has yet to be explored.

**MATERIALS AND METHODS**

**Cell lines**

Primary esophageal epithelial cells were isolated and cultured as previously described. The immortalized human esophageal epithelial
cell line (EPC2; kindly provided by Dr Anil Rustgi, University of Pennsylvania, Philadelphia, PA, USA) were cultured as previously described.\textsuperscript{25,26}

Patient demographics

EoE patients included for the RNA sequencing and microarray analyses were defined as having a positive EoE diagnosis with \( \geq 15 \) eosinophils per high-power microscopic field (HPF) in a concurrent esophageal biopsy. NL control individuals were defined as having no history of an EoE diagnosis with 0 eosinophils per HPF in a concurrent esophageal biopsy. A total of 10 patients with active EoE and six NL controls were analyzed by RNA sequencing; a negative diagnosis of GERD was made in nine of the 10 patients with EoE following unresponsiveness to proton pump inhibitor (PPI) therapy. A total of 17 patients with active EoE and 11 NL controls were included for the microarray analysis; these samples partially overlapped with those previously reported;\textsuperscript{25} 10 of the 17 EoE patients had a negative diagnosis for GERD following unresponsiveness to PPI therapy. Additional patient cohorts analyzed by qPCR are as follows: Figure 3c, 12 NL control individuals, 12 patients with previous GERD diagnoses and esophageal eosinophil levels = 0–2 while on PPI therapy, and 12 EoE patients with active disease (eosinophils/HPF = 25–160) while on PPI therapy; Figure 4b, 19 NL control individuals and 27 EoE patients with active disease (eosinophils/HPF = 25–394) while on PPI therapy.

RNA sequencing analyses

RNA sequencing analysis was performed by the Cincinnati Children's Hospital Medical Center (CCHMC) Genetic Variation and Gene Discovery Core. In brief, RNA was isolated using the RNeasy kit (QIAGEN Incorporated, Germantown, MD, USA) according to the manufacturer's protocol. Whole transcriptome (RNA) sequencing was performed at the CCHMC Gene Discovery and Genetic Variation Core as previously described.\textsuperscript{11} The paired-end sequencing reads were aligned against the GRCh37 genome model using TopHat 2.0.4 with Bowtie 2.0.\textsuperscript{27,28} The separate alignments were then merged using Cuffmerge\textsuperscript{29} with UCSC gene models as a reference. Raw data were assessed for statistical significance using a Welch t-test with Benjamini–Hochberg false discovery rate correction.

Microarray analyses

Microarray analysis of esophageal biopsies was performed by the CCHMC Gene Expression Microarray Core using the Affymetrix U133 Plus 2.0 array (Affymetrix, Santa Clara, CA, USA) as previously described.\textsuperscript{3} In brief, RNA was isolated using the RNeasy kit (QIAGEN Incorporated) according to the manufacturer's protocol. For microarray analysis of BANC or NSC shRNA–transduced cells, cells were treated with suboptimal doses of IL-13 (10 ng/ml) for 24h in triplicate. RNA quality assessment, library preparation, hybridization to the GeneChip Human Gene 2.0 ST exon array (Affymetrix) and analyses were performed at the CCHMC Gene Expression Microarray Core. Data were normalized to the median of control/untreated samples, and statistical significance was determined using a Welch t-test with Benhami–Hochberg false discovery rate correction (biopsies) or a two-way analysis of variance (BANC shRNA) with a threshold of \( P<0.05 \) and a 2.0-fold (biopsies) or 1.5-fold (BANC shRNA) cut-off filter in GeneSpring GX (Agilent Technologies Incorporated, Clara, CA, USA).

Silencing of BANC by shRNA

Lentivirus production from the pGIPZ lentiviral plasmid (Thermo Scientific, Rockford, IL, USA) containing NSC shRNA or shRNA targeting BANC\textsuperscript{10} (kindly provided by Dr Paul Khavari, Stanford University, Stanford, CA, USA) was performed at the CCHMC Viral Vector Core facility. Stably transduced EPC2 cells were generated by negative selection in puromycin, which was maintained in the culture media throughout all experiments.

qPCR

RNA from esophageal biopsies from the RNA sequencing cohort or separate case–control cohorts and RNA from primary esophageal epithelial cells or shRNA-transduced EPC2 cells were used for cDNA synthesis (iScript, BioRad, Hercules, CA, USA). qPCR analysis using SYBR Green was performed (BioRad). Primer sets used in this study are listed in Supplementary Table 1. All data were normalized to the housekeeping gene GAPDH as previously described.\textsuperscript{30} Normalized data from patient biopsies are presented as the mean ± interquartile range, and statistical analysis was performed using a Welch t-test (for two groups) or Kruskal–Wallis test with Dunn's multiple comparison test (for three groups) with a threshold of \( P<0.05 \). Normalized data derived from cultured cells are presented as mean ± s.e.m., and statistical analysis was performed using a Welch t-test (for two groups) or Kruskal–Wallis test with Dunn's multiple comparison test (for three groups) with a threshold of \( P<0.05 \).

Bioinformatics

Clusters of low, medium and high-expressed Escc transcriptome genes from the RNA sequencing were assessed for GO and Coexpression Atlas enrichment using TopGene.\textsuperscript{31} Enriched terms were considered significant with \( P<0.05 \) after false discovery rate correction.

CONFLICT OF INTEREST

MER is a consultant for Immune Therapeutics and has an equity interest in reslizumab (Teva Pharmaceuticals) and is a consultant for Immune Pharmaceuticals, Pluristem Pharmaceuticals, Novartis and Receptos. The remaining authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported in part by NIH U19 AI070235, NIH R01 DK076893, the PHS Grant P30 DK078392 and the Campaign Urging Research for Eosinophilic Disease (CURED) Foundation, and the Angels for eosinophilic Research Foundation. JDS was supported by the Thrasher Research Fund Early Career Award (NR-0171). We would like to thank Dr Anil Rustgi (University of Pennsylvania, Pennsylvania, PA, USA) for the EPC2-HTERT cell line and Dr Paul Khavari (Stanford University, Stanford, CA, USA) for supplying the BANC shRNA lentiviral plasmids. We would also like to thank Shawn Hottinger for editorial assistance, all of the participating families and clinical research staff of the Cincinnati Center for Eosinophilic Disorders and members of the Division of Allergy and Immunology.

Author contributions: JDS, CB, BJA and MER were involved in study concept and design. JDS, KKC, CB, EMS and KAK were involved in data acquisition and/or sample preparation. MHC, JPA, PEP, VAM, AJP, SAK and JPK collected patient samples or provided histopathological analyses thereof. AJP, PJD and BJA performed computational analysis of RNA sequencing reads. JDS, AJP, PJD, RAK, BJA and MER were involved in data analysis and interpretation and the writing and critical revision of the manuscript. BJA and MER supervised the overall study.

REFERENCES

1. Sherrell JD, Rothenberg ME. Genetic dissection of eosinophilic esophagitis provides insight into disease pathogenesis and treatment strategies. J Allergy Clin Immunol 2011; 128: 23–32.
2. Noel RJ, Putnam PE, Rothenberg ME. Eosinophilic esophagitis. N Engl J Med 2004; 351: 940–941.
3. Noel RJ, Rothenberg ME. Eosinophilic esophagitis. Curr Opin Pediatr 2005; 17: 690–694.
4. Ronkainen J, Talley NJ, Aro P, Storskrubb T, Johansson SE, Lind T et al. Prevalence of oesophageal eosinophils and eosinophilic oesophagitis in adults: the population-based Kalixanda study. Gut 2007; 56: 615–620.
5. Blanchard C, Wang N, Stringer KF, Mishra A, Fulkerson PC, Abonia JP et al. Eotaxin-3 and a uniquely conserved gene-expression profile in eosinophilic esophagitis. J Clin Invest 2006; 116: 536–547.
6. Blanchard C, Mingler MK, Vicario M, Abonia JP, Wu YY, Lu TX et al. IL-13 involvement in eosinophilic esophagitis: transcriptome analysis and reversibility with glucocorticoids. J Allergy Clin Immunol 2007; 120: 1292–1300.
7. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 2008; 18: 1509–1517.
8. Schultheis PJ, Clarke LL, Meneton P, Miller ML, Soleimani M, Gaweins LR et al. Renal and intestinal absorptive defects in mice lacking the NHE3 Na+/H+ exchanger. Nat Genet 1998; 19: 282–285.
9. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Genome Res 2012; 22: 1006–1014.
10. Lu TX, Sherrell JD, Wen T, Plassard AJ, Besse JA, Abonia JP et al. MicroRNA signature in patients with eosinophilic esophagitis, reversibility with glucocorticoids, and assessment as disease biomarkers. J Allergy Clin Immunol 2012; 129: 1064–75 e9.
12 Lu TX, Lim EJ, Wen T, Plassard AJ, Hogan SP, Martin LJ et al. MiR-375 is downregulated in epithelial cells after IL-13 stimulation and regulates an IL-13-induced epithelial transcriptome. *Mucosal Immunol* 2012; 5: 388–396.

13 Lu S, Mukkada VA, Mangray S, Cleveland K, Shillingford N, Schorl C et al. MicroRNA profiling in mucosal biopsies of eosinophilic esophagitis patients pre and post treatment with steroids and relationship with mRNA targets. *PLoS ONE* 2012; 7: e40676.

14 Bhattacharya B, Carlsten J, Sabo E, Kethu S, Meitner P, Tavares R et al. Increased expression of eotaxin-3 distinguishes between eosinophilic esophagitis and gastroesophageal reflux disease. *Hum Pathol* 2007; 38: 1744–1753.

15 Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL et al. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* 2013; 341: 789–792.

16 Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 2012; 22: 1775–1789.

17 Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* 2012; 22: 1760–1774.

18 Khalil AM, Guttman M, Huarte M, Sabo E, Rivea Morales D et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci USA* 2009; 106: 11667–11672.

19 Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 2009; 458: 223–227.

20 Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet* 2009; 10: 155–159.

21 Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell* 2011; 43: 904–914.

22 Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* 2013; 493: 231–235.

23 Blanchard C, Stucke EM, Burwinkel K, Caldwell JM, Collins MH, Ahrens A et al. Coordinate interaction between IL-13 and epithelial differentiation cluster genes in eosinophilic esophagitis. *J Immunol* 2010; 184: 4033–4041.

24 Sherrill JD, Kc K, Wu D, Djukic Z, Caldwell JM, Stucke EM et al. Desmoglein-1 regulates esophageal epithelial barrier function and immune responses in eosinophilic esophagitis. *Mucosal Immunol* 2013; 7: 718–729.

25 Andl CD, Mizushima T, Nakagawa H, Oyama K, Harada H, Chruma K et al. Epidermal growth factor receptor mediates increased cell proliferation, migration, and aggregation in esophageal keratinocytes in vitro and in vivo. *J Biol Chem* 2003; 278: 1824–1830.

26 Harada H, Nakagawa H, Oyama K, Takaoka M, Andl CD, Jacobmeier B et al. Telomerase induces immortalization of human esophageal keratinocytes without p16INK4a inactivation. *Mol Cancer Res* 2003; 1: 729–738.

27 Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 2012; 7: 562–578.

28 Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009; 10: R25.

29 Garber M, Grabherr MG, Guttman M, Trapnell C. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* 2011; 8: 469–477.

30 Sherrill JD, Gao PS, Stucke EM, Blanchard C, Collins MH, Putnam PE et al. Variants of thymic stromal lymphopoietin and its receptor associate with eosinophilic esophagitis. *J Allergy Clin Immunol* 2010; 126: 160–165 e3.

31 Chen J, Bardees EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009; 37(Web Server issue): W305–W311.

Supplementary Information accompanies this paper on Genes and Immunity website (http://www.nature.com/gene)