Published in final edited form as:

*Genes Immun*. 2014 September ; 15(6): 361–369. doi:10.1038/gene.2014.27.

# Analysis and expansion of the eosinophilic esophagitis transcriptome by RNA sequencing

**Joseph D. Sherrill**[1], **KC Kiran**[1], **Carine Blanchard**[6], **Emily M. Stucke**[1], **Katherine A. Kemme**[1], **Margaret H. Collins**[2], **J. Pablo Abonia**[1], **Philip E. Putnam**[3], **Vincent A. Mukkada**[3], **Ajay Kaul**[3], **Samuel A. Kocoshis**[3], **Jonathan P. Kushner**[5], **Andrew J. Plassard**[4], **Rebekah A. Karns**[4], **Phillip J. Dexheimer**[4], **Bruce J. Aronow**[4], and **Marc E. Rothenberg**[1,†]

[1]Division of Allergy and Immunology, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA [2]Division of Pathology and Laboratory Medicine, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA [3]Division of Gastroenterology, Hepatology, and Nutrition, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA [4]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA [5]Department of Internal Medicine, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA [6]Nutrition and Health / Allergy Group, Nestlé Research Center, PO Box 44, CH-1000 Lausanne 26, Switzerland

## Abstract

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 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 1 607 significantly dysregulated transcripts (1 096 upregulated, 511 downregulated). When clustered by raw expression levels, an abundance of immune-cell specific transcripts that are highly induced in EoE are 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 *BANCR* was upregulated in EoE and induced in



†**Corresponding Author** Marc E. Rothenberg, MD, PhD, Division of Allergy and Immunology, Cincinnati Children's Hospital Medical Center, MLC 7028, 3333 Burnet Avenue, Cincinnati, OH 45229.

Author Contributions

J.D.S., C.B., B.J.A., and M.E.R. were involved in study concept and design. J.D.S., K.K., C.B., E.M.S., and K.A.K. were involved in data acquisition and/or sample preparation. M.H.C., J.P.A., P.E.P., V.A.M., A.J., S.A.K., and J.P.K. collected patient samples or provided histopathological analyses thereof. A.J.P., P.J.D., and B.J.A. performed computational analysis of RNA sequencing reads. J.D.S., A.J.P., P.J.D., R.A.K., B.J.A., and M.E.R. were involved in data analysis and interpretation and the writing and critical revision of the manuscript. B.J.A. and M.E.R. supervised the overall study.

Conflicts of Interest

M.E.R. 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. All remaining authors have no disclosures.

IL-13–treated primary esophageal epithelial cells. Repression of *BANCR* significantly altered the expression of IL-13–induced pro-inflammatory genes. Together, these data comprise new potential biomarkers of EoE and demonstrate a novel role for lncRNAs in EoE and IL-13–associated responses.

**Keywords**

Eosinophilic esophagitis; RNA sequencing; lncRNA; IL-13

## 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.[1] The symptomology of EoE can progress from feeding difficulties and vomiting to dysphagia and food impactions, underscoring the chronicity of the disease.[2, 3] Moreover, as the prevalence of EoE is increasing on a global scale, with current estimates at 4 in 10 000[4], 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 individuals with chronic, non-eosinophilic esophagitis.[5] In particular, 574 transcripts (344 upregulated, 230 downregulated) were differentially regulated in EoE patients compared to healthy controls.[5] 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.[5] Subsequent work demonstrated that the majority of the gene signature in EoE normalized after steroid-induced disease remission.[6] 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).[6]

Recent technical advancements for elucidating transcript profiles, such as high-throughput whole-transcriptome (RNA) sequencing, have been made. RNA sequencing offers greater transcriptional resolution compared to traditional probe-based microarrays, as it generates transcript profiles that are not reliant upon known transcripts and has greater dynamic range for detection of low-abundance transcripts.[7] 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 pro-inflammatory 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 to from healthy (NL) controls. We identified a total of 1 607 transcripts that were dysregulated in EoE ($P < .05$, fold change > 2.0) (Fig. 1A and B). Of these, 1 085 genes were upregulated and 511 were downregulated compared to controls. We also clustered the EoE dysregulated genes 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 (e.g., *CCL26, ALOX15, CRISP3*) were previously identified as part of the EoE transcriptome and fell within cluster 3 (Fig. 1A and B).[5] We validated the change in expression of eight representative genes in the same samples by quantitative PCR: *ALOX15*, *CCL26*, *CLC*, and *CPA3* were significantly increased (Fig. 1C), whereas *SPINK7*, *SPRR2D*, *SPRR2B*, and *S100A6* were significantly decreased in EoE (Fig. 1D).

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 (Fig. 1E). While 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 MHC peptide binding and antigen recognition whereas cluster 3 contained low expressed genes involved in immune cell (lymphocytes, mast cells, and eosinophils) activation and migration.

In a separate cohort of patients, we compared the differential gene signature from RNA sequencing to that identified by expression profiling by standard microarray. Updated microarray analyses identified a total of 870 dysregulated transcripts in EoE (compared to 574 transcripts as previously reported[5]), with 374 and 496 being upregulated and downregulated, respectively, compared to controls. To compare the differentially expressed gene signatures 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 data sets; notably, this overlap corresponded to 76% and 27% of the total number of upregulated genes identified by microarray and RNA sequencing, respectively (Fig. 2A). Comparing the relative fold changes of these 284 upregulated genes between platforms demonstrated a significant correlation (Spearman r = 0.66, $P < 10^{-4}$)

(Fig. 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 (Fig. 2C). The relative fold changes of the shared downregulated transcripts displayed a significant correlation (Spearman r = 0.48, $P < 10^{-4}$) between the microarray and RNA sequencing data sets (Fig. 2D).

## Novel EoE transcriptome genes from RNA sequencing

RNA sequencing identified 1 028 genes (770 upregulated, 258 downregulated) that were previously unrecognized as part of the EoE transcriptome (Fig. 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 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 a greater than 5-fold change (median absolute fold change [interquartile range] = 7.0 [5.7 – 10.0]) (Fig. 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* encodes the sodium-hydrogen exchanger family member 3 (NHE3), and although its function in the esophageal mucosa remains unknown, functional studies in the intestine have shown that NHE3 regulates epithelial absorption of NaCl and $HCO_3^-$.[8] 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 *SLC9A3* and *CCL24* were significantly upregulated in a separate cohort of EoE but not in patients with gastroesophageal reflux disease (GERD) when compared healthy controls, indicating disease-specific dysregulation (Fig. 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 lncRNAs 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 lncRNAs, an emerging class of non-coding RNA species that have been shown to significantly alter gene expression under both homeostatic and disease states.[9] We observed a significant upregulation of the lncRNA termed BRAF-activated non-protein coding RNA (*BANCR*)[10], which was induced 3.7 fold in EoE compared to in NL controls by RNA sequencing (Fig. 4A) and specifically upregulated in EoE but not GERD (Fig. 3C). To further characterize *BANCR* expression in EoE, we probed the entire RNA sequencing profile in EoE for transcripts exhibiting similar expression profiles to that of *BANCR*. Although the overall expression of *BANCR* was low in EoE (FPKM median [interquartile range] = 3.4 [0.1 – 26]), the expression of 36 transcripts was highly correlated with *BANCR* levels (Spearman r > 0.90), with the genes encoding *ALOX15* and periostin (*POSTN*) being the most correlated with *BANCR* expression in patients with EoE (Fig. 4A, lower panel). We subsequently validated EoE-specific *BANCR* overexpression by quantitative PCR in a large cohort of NL controls (n = 23) and patients with active EoE (n = 34) (Fig. 4B). In 4 of the 23

NL controls, *BANCR* failed to amplify, demonstrating that the low abundance observed in the RNA sequencing data was also observed in the NL control group. In samples with detectable expression, *BANCR* was significantly induced in EoE ($P < 10^{-4}$, fold change = 43.5). Interestingly, *BANCR* expression showed a significant correlation with esophageal eosinophil levels (Spearman r = 0.58, $P = 3 \times 10^{-4}$) (Fig. 4C) in EoE patients and was highly correlated with expression of *IL13* in both NL controls and EoE patients (Spearman r = 0.77, $P < 10^{-4}$) (Fig. 4D).

Given the inducible nature of *BANCR* and its correlation with both relevant transcriptional (e.g., *ALOX15, POSTN*) and biological (eosinophil levels) markers of EoE, we hypothesized that IL-13, a critical cytokine capable of inducing a large proportion of the EoE transcriptome[6], may be responsible for triggering *BANCR* expression in esophageal epithelial cells. Indeed, primary esophageal epithelial cells stimulated with IL-13 showed 288-fold and 66-fold increases in *BANCR* expression at 24 and 48 hours post treatment, respectively, compared to unstimulated cells (0 hours) (Fig. 5A). In order to address the potential biological significance of *BANCR* in EoE pathogenesis, we used shRNA to block the IL-13– mediated induction of *BANCR* in an esophageal epithelial cell line. *BANCR* induction was blocked by approximately 76% in *BANCR* shRNA–treated cells compared to cells treated with a non-silencing control (NSC) shRNA ($P < 0.05$) (Fig. 5B). Microarray analyses identified 40 transcripts that showed differential expression between *BANCR* and NSC shRNA–treated cells after IL-13 stimulation ($P < 0.05$, fold change > 1.5) (Fig. 5C). Here, interferon-induced protein 44-like (*IFI44L*) was induced 2.4 fold in IL-13–treated *BANCR*-deficient cells; in contrast, caspase 14 (*CASP14*), which was unchanged after IL-13 stimulation of NSC shRNA–treated cells, was downregulated 2.6 fold in IL-13–treated cells depleted for *BANCR*. Quantitative PCR analysis further supported that loss of *BANCR* attenuates *CASP14* expression following IL-13 treatment by 5.6 fold (Fig. 5D).

## 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 1 028 transcripts) associated with EoE was not identified by traditional microarray expression profiles. Thus far, expression of non-coding RNAs and their role in EoE has been limited to microRNAs.[11-13] However, we show that Th2 cytokines (IL-13) are potent stimuli for lncRNA expression, implicating lncRNAs (e.g., *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.[7] 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 analysis.[5] We

identified 870 dysregulated genes in EoE (374 upregulated, 496 downregulated), which was a 51% increase compared to a previous analysis[5] 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 to n = 19 in our previous publication[5] ).

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 1 028 EoE transcriptome genes newly identified by RNA sequencing were modestly dysregulated in patients with EoE. For instance, 807 of these genes (78%) showed < 3-fold 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.[7] 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.[5, 14]

Emerging data indicate a significant role for lncRNAs in diverse cellular processes ranging from growth and differentiation to apoptosis and inflammatory responses,[15-18] 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 (~14 880 unique transcripts) that are transcriptionally regulated and processed in similar fashions to coding RNAs.[16, 17, 19] LncRNAs can regulate gene expression through various *cis* or *trans* mechanisms, such as by scaffolding multi-protein complexes, mediating DNA-RNA interactions, and acting as decoy receptors for microRNAs (reviewed in [20] and [21]). *BANCR* was originally identified as a BRAF-induced lncRNA associated with melanoma, in which suppressing *BANCR* induction attenuated *CXCL11* expression and reduced melanocyte migration.[10] We demonstrate that *BANCR* suppression reduced the expression of caspase 14 mRNA in IL-13 treated esophageal epithelial cells. *CASP14* was recently shown to be downregulated after suppression of another lncRNA termed terminal differentiation–induced non-coding RNA (*TINCR*).[22] *TINCR* was demonstrated to promote terminal differentiation of keratinocytes by stabilizing the mRNAs for genes necessary for epidermal differentiation (e.g., filaggrin, loricrin) and for the processing of barrier peptides like *CASP14*.[22] Notably, esophageal epithelial differentiation[23] and barrier function[24] 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 (e.g., *BANCR*) can be involved in biologically relevant disease processes (e.g.,

IL-13 signaling and inflammatory gene expression). It remains possible that lncRNAs 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 (i.e., allele-specific expression) has yet to be explored.

## Materials and Methods

### Cell lines

Primary esophageal epithelial cells were isolated and cultured as previously described.[23] 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.[25, 26]

### Patient demographics

EoE patients included for the RNA sequencing and microarray analyses were defined as having a positive EoE diagnosis with ≥ 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 6 NL controls were analyzed by RNA sequencing; a negative diagnosis of gastroesophageal reflux disease (GERD) was made in 9 of the 10 patients with EoE following unresponsiveness to proton pump inhibitor 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[5]; 10 of the 17 EoE patients had a negative diagnosis for GERD following unresponsiveness to proton pump inhibitor therapy. Additional patient cohorts analyzed by quantitative PCR 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) according to the manufacturer's protocol. Whole transcriptome (RNA) sequencing was performed at the CCHMC Gene Discovery and Genetic Variation Core as previously described.[11] The paired-end sequencing reads were aligned against the GRCh37 genome model using TopHat 2.04 with Bowtie 2.03.[27, 28] The separate alignments were then merged using Cuffmerge[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 and a threshold of $P < 0.05$ and a 2.0-fold cut-off filter in GeneSpring® GX (Agilent Technologies Incorporated, Clara, CA).

### 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) as previously described.[5] In brief, RNA was isolated using the RNeasy kit (QIAGEN Incorporated, Germantown, MD) according to the manufacturer's protocol. For microarray analysis of *BANCR* or NSC shRNA–transduced cells, cells were treated with suboptimal doses of IL-13 (10 ng/mL) for 24 hours in triplicate. RNA quality assessment, library preparation, hybridization to the GeneChip® Human Gene 2.0 ST exon array (Affymetrix, Santa Clara, CA), 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 Benjamini-Hochberg false discovery rate (biopsies) or a 2-way ANOVA (*BANCR* shRNA) with a threshold of $P < 0.05$ and a 2.0-(biopsies) or 1.5-fold (*BANCR* shRNA) cut-off filter in GeneSpring® GX (Agilent Technologies Incorporated, Clara, CA).

### Silencing of BANCR by shRNA

Lentivirus production from the pGIPZ lentiviral plasmid (Thermo Scientific, Rockford, IL, USA) containing NSC shRNA or shRNA targeting *BANCR*[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.

### Quantitative PCR

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). Quantitative PCR analysis using SYBR Green was performed (BioRad, Hercules, CA). Primer sets used in this study are listed in Supplementary Table 1. All data were normalized to the housekeeping gene *GAPDH* as previously described.[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 ± standard error of the mean (SEM), 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 EoE transcriptome genes from the RNA sequencing were assessed for GO and Coexpression Atlas enrichment using ToppGene.[31] Enriched terms were considered significant with $P < 0.05$ after false discovery rate correction.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported in part by NIH U19 AI070235, NIH R01 DK076893, the PHS Grant P30 DK0789392, and the Campaign Urging Research for Eosinophilic Disease (CURED) Foundation, and the Angels for Eosinophilic Research Foundation. J.D.S. 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 *BANCR* shRNA lentiviral plasmids. We would also like to thank Shawna 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 for improving the health of children with allergic and immune conditions through innovative research, outstanding clinical care, and education of the current and next generation of leaders in healthcare and research.

## References

1. Sherrill JD, Rothenberg ME. Genetic dissection of eosinophilic esophagitis provides insight into disease pathogenesis and treatment strategies. The Journal of allergy and clinical immunology. 2011; 128(1):23–32. quiz 33-4. [PubMed: 21570716]
2. Noel RJ, Putnam PE, Rothenberg ME. Eosinophilic esophagitis. N Engl J Med. 2004; 351(9):940–1. [PubMed: 15329438]
3. Noel RJ, Rothenberg ME. Eosinophilic esophagitis. Current opinion in pediatrics. 2005; 17(6):690–4. [PubMed: 16282772]
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(5):615–20. [PubMed: 17135307]
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. The Journal of clinical investigation. 2006; 116(2):536–47. [PubMed: 16453027]
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. The Journal of allergy and clinical immunology. 2007; 120(6):1292–300. [PubMed: 18073124]
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 research. 2008; 18(9):1509–17. [PubMed: 18550803]
8. Schultheis PJ, Clarke LL, Meneton P, Miller ML, Soleimani M, Gawenis LR, et al. Renal and intestinal absorptive defects in mice lacking the NHE3 Na+/H+ exchanger. Nat Genet. 1998; 19(3): 282–5. [PubMed: 9662405]
9. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Nature. 2012; 482(7385):339–46. [PubMed: 22337053]
10. Flockhart RJ, Webster DE, Qu K, Mascarenhas N, Kovalski J, Kretz M, et al. BRAFV600E remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. Genome research. 2012; 22(6):1006–14. [PubMed: 22581800]
11. Lu TX, Sherrill 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. The Journal of allergy and clinical immunology. 2012; 129(4):1064–75. e9. [PubMed: 22391115]
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 immunology. 2012; 5(4):388–96. [PubMed: 22453679]
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(7):e40676. [PubMed: 22815788]
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. Human pathology. 2007; 38(12):1744–53. [PubMed: 17900656]

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(6147): 789–92. [PubMed: 23907535]
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 research. 2012; 22(9):1775–89. [PubMed: 22955988]
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 research. 2012; 22(9): 1760–74. [PubMed: 22955987]
18. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009; 106(28):11667–72. [PubMed: 19571010]
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(7235): 223–7. [PubMed: 19182780]
20. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nature reviews. Genetics. 2009; 10(3):155–9. [PubMed: 19188922]
21. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Molecular cell. 2011; 43(6):904–14. [PubMed: 21925379]
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(7431):231–5. [PubMed: 23201690]
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. Journal of immunology. 2010; 184(7):4033–41.
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 immunology. 2013
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(3):1824–30. [PubMed: 12435727]
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(10):729–38. [PubMed: 12939398]
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. Nature protocols. 2012; 7(3):562–78. [PubMed: 22383036]
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(3):R25. [PubMed: 19261174]
29. Garber M, Grabherr MG, Guttman M, Trapnell C. Computational methods for transcriptome annotation and quantification using RNA-seq. Nature methods. 2011; 8(6):469–77. [PubMed: 21623353]
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. The Journal of allergy and clinical immunology. 2010; 126(1):160–5. e3. [PubMed: 20620568]
31. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009; 37(Web Server issue):W305–11. [PubMed: 19465376]

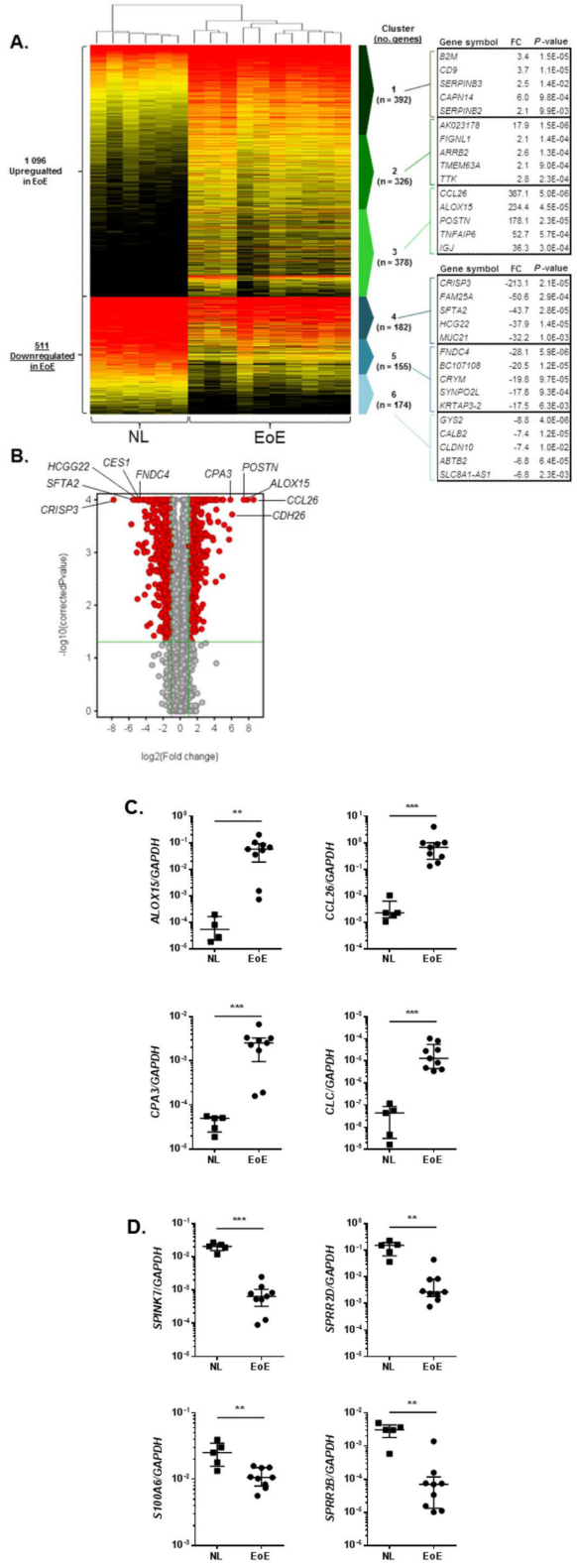


**Figure 1. Differential gene expression in EoE identified by RNA sequencing**

Heatmap showing log2 FPKM values and clustering for 1 607 genes identified as dysregulated in EoE (1 096 upregulated, 511 downregulated; $P < .05$, absolute fold change > 2) 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 ($P < 0.05$, > 2 fold change) are indicated as red squares. Green lines represent the thresholds used for *P*-value ($P < 0.05$) and fold change (> 2) (B). Quantitative PCR validation (mean ± interquartile range) of 4 upregulated (C) and 4 downregulated (D) candidate genes identified by RNA sequencing. Network displaying enriched biological pathways for gene clusters 1-3 from Fig. 1A (E). **, $P < 10^{-2}$. ***, $P < 5 \times 10^{-3}$. EoE, eosinophilic esophagitis; NL, healthy controls; no., number.

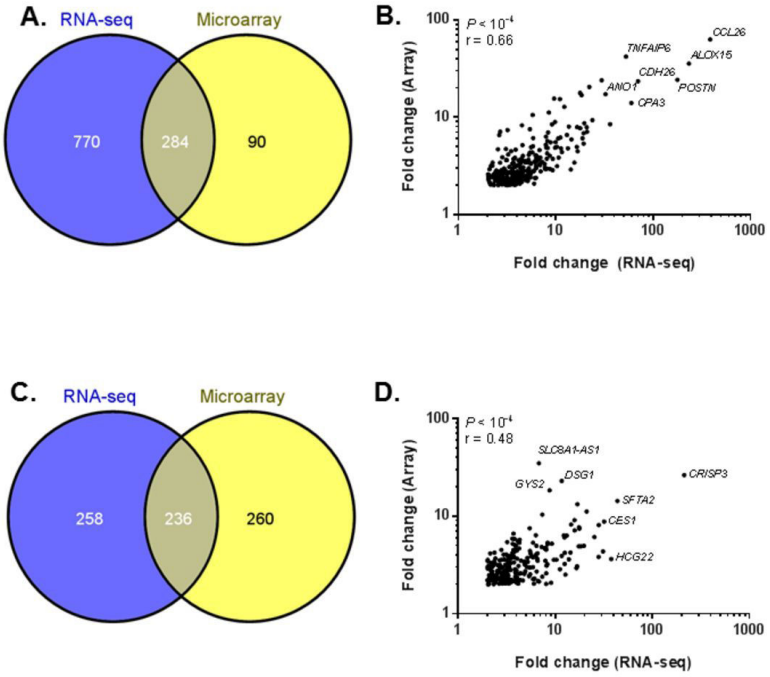


**Figure 2. Comparison between EoE transcriptome genes identified by RNA sequencing and microarray**

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.

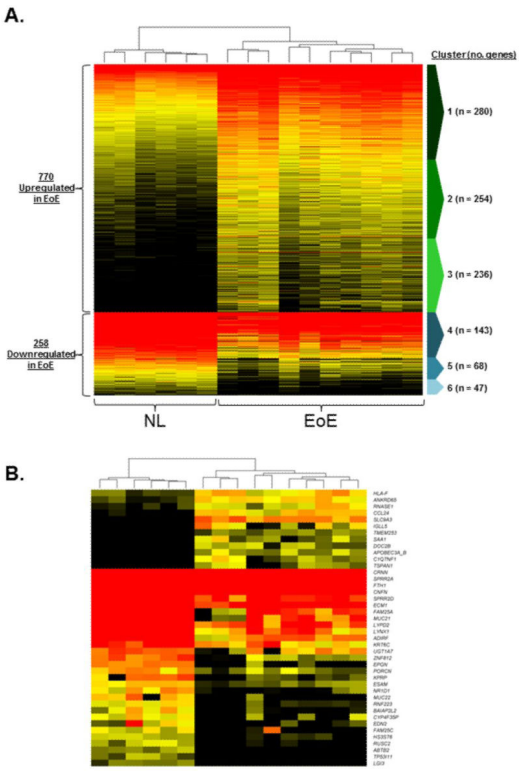


**Figure 3. Novel EoE transcriptome genes identified by RNA sequencing**

Heatmap showing log2 FPKM values and clustering for the unique 1 028 dysregulated genes in EoE (770 upregulated, 258 downregulated) identified by RNA sequencing only ($P < 0.05$, absolute fold change > 2) (A). Heatmap showing log2 FPKM values for the 42 unique, most dysregulated genes (fold change > 5) (B). Quantitative PCR 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; *, $P < 0.05$; **, $P < 5 \times 10^{-3}$; EoE, eosinophilic esophagitis; NL, healthy controls; GERD, gastroesophageal reflux disease; no., number.

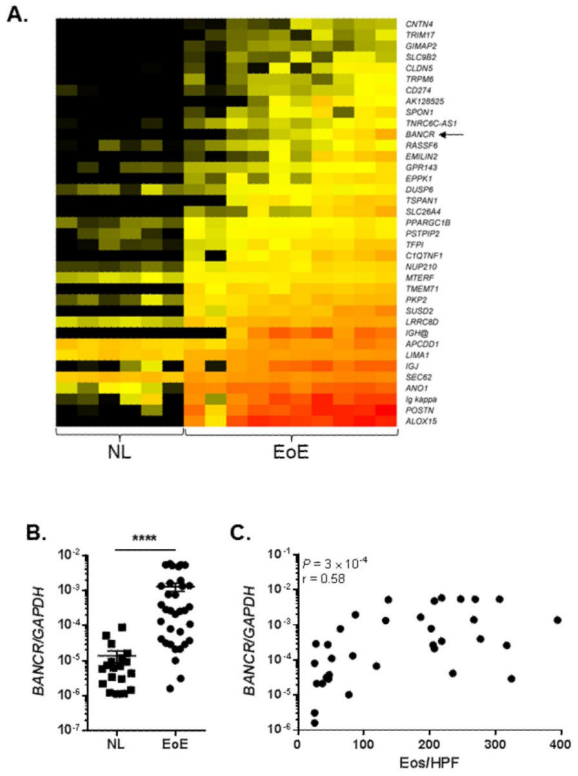


**Figure 4. Dysregulation of the lncRNA *BANCR* in EoE**

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. Quantitative PCR 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. **, $P < 5 \times 10^{-3}$. EoE, eosinophilic esophagitis; Eos, eosinophils; HPF, high-power field; NL, healthy controls.

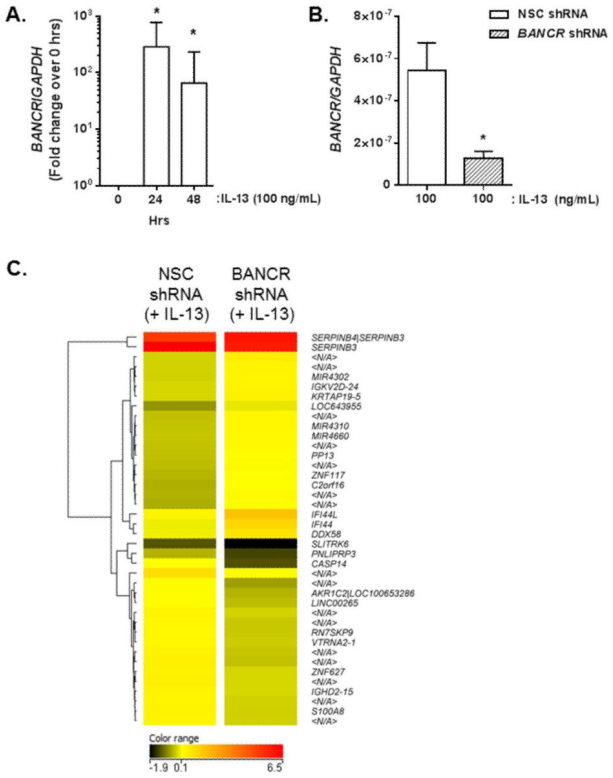


**Figure 5. *BANCR* induction by IL-13 and regulation of innate immune genes**

Quantitative PCR analysis of *BANCR* induction in IL-13–treated primary esophageal epithelial cells (A). Quantitative PCR analysis of suppression of IL-13–induced *BANCR* expression by targeted shRNA (hatched bar) compared to by non-silencing control (NSC) shRNA (open bar) in IL-13–treated esophageal epithelial cells (B). Heatmap showing log2 expression values and clustering of the 40 dysregulated transcripts in IL-13–treated esophageal epithelial cells transduced with *BANCR* shRNA compared to with NSC shRNA ($P < 0.05$, absolute fold change > 1.5) (C) Entities in (C) are labelled by gene symbol where appropriate; N/A, not available. Quantitative PCR analysis of *CASP14* expression in IL-13–treated esophageal epithelial cells transduced with *BANCR* shRNA or NSC shRNA (fold change compared to untreated cells) (D). Data in (A), (B), and (D) were from three independent experiments performed in duplicate or triplicate and expressed as the mean + SEM. Ns, not significant; *, $P < 0.05$; **, $P < 5 \times 10^{-3}$.