Transcriptome analysis of controlled and therapy-resistant childhood asthma reveals distinct gene expression profiles

Helena Persson, PhD,[a]‡ Andrew T. Kwon, PhD,[b,c] Jordan A. Ramilowski, PhD,[b,c] Gilad Silberberg, PhD,[d] Cilla Söderhäll, PhD,[a,e] Christina Orsmark-Pietras, PhD,[a]* Björn Nordlund, RN, PhD,[e,f,g] Jon R. Konradsen, MD, PhD,[e,f,g] Michiel J. L. de Hoon, PhD,[b,c] Erik Melén, MD, PhD,[e,h,i] Yoshihide Hayashizaki, MD, PhD,[b,j] Gunilla Hedlin, MD, PhD,[e,f,g] Juha Kere, MD, PhD,[a,e,k,l] and Carsten O. Daub, PhD[a,b,c]

Stockholm, Sweden, Yokohama and Wako, Japan, and Helsinki, Finland

Background: Children with problematic severe asthma have poor disease control despite high doses of inhaled corticosteroids and additional therapy, leading to personal suffering, early deterioration of lung function, and significant consumption of health care resources. If no exacerbating factors, such as smoking or allergies, are found after extensive investigation, these children are given a diagnosis of therapy-resistant (or therapy-refractory) asthma (SA).

Objective: We sought to deepen our understanding of childhood SA by analyzing gene expression and modeling the underlying regulatory transcription factor networks in peripheral blood leukocytes.

Methods: Gene expression was analyzed by using Cap Analysis of Gene Expression in children with SA (n = 13), children with controlled persistent asthma (n = 15), and age-matched healthy control subjects (n = 9). Cap Analysis of Gene Expression sequencing detects the transcription start sites of known and novel mRNAs and noncoding RNAs.

Results: Sample groups could be separated by hierarchical clustering on 1305 differentially expressed transcription start sites, including 816 known genes and several novel transcripts. Ten of 13 tested novel transcripts were validated by means of RT-PCR and Sanger sequencing. Expression of RAR-related orphan receptor A *(RORA)*, which has been linked to asthma in genome-wide association studies, was significantly upregulated in patients with SA. Gene network modeling revealed decreased glucocorticoid receptor signaling and increased activity of the mitogen-activated protein kinase and Jun kinase cascades in patients with SA.

Conclusion: Circulating leukocytes from children with controlled asthma and those with SA have distinct gene expression profiles, demonstrating the possible development of specific molecular biomarkers and supporting the need for novel therapeutic approaches. (J Allergy Clin Immunol 2015;■■■:■■■-■■■.)

Key words: *Therapy-resistant asthma, childhood asthma, peripheral blood leukocytes, transcriptome, long noncoding RNA*

Asthma is the most common chronic disease in children[1] and can be defined as mild, moderate, or severe depending on the extent of medication needed to control symptoms.[2] Problematic severe asthma is characterized by poor disease control, even with high doses of inhaled corticosteroids and additional therapies. It affects approximately 5% of all asthmatic children[3] and can cause extensive personal suffering, early deterioration of lung function, and significant consumption of health care resources.[4] Problematic severe asthma in children can often be explained by exposure to exacerbating factors in the environment, such as smoking or allergens, but therapy-resistant (or therapy-refractory) asthma (SA) is believed to exist in a subgroup of children in whom no such factors are found, despite extensive investigation.[5]

Global gene expression in human asthma has been studied in isolated leukocyte populations,[6-9] bronchial and epithelial biopsy specimens,[10-14] and nasal lavage samples.[15] These studies have provided important mechanistic insights for mild-to-moderate and atopic asthma, but few studies have addressed severe

From [a]the Department of Biosciences and Nutrition and Center for Innovative Medicine (CIMED), [d]the Unit of Computational Medicine, Department of Medicine, [e]the Centre for Allergy Research, [g]the Department of Women's and Children's Health, and [h]the Institute of Environmental Medicine, Karolinska Institutet, Stockholm; [b]Omics Science Center,§ RIKEN Yokohama Institute, Yokohama; [c]the Division of Genomic Technologies, RIKEN Center for Life Science Technologies, Yokohama; and [f]Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm; [i]Sachs' Children's Hospital, Stockholm; [j]the Preventive Medicine and Diagnosis Innovation Program, RIKEN Research Cluster for Innovation, Wako; [k]the Folkhälsan Institute of Genetics, Helsinki; and [l]the Research Programs Unit, University of Helsinki.

*Dr Orsmark-Pietras is currently affiliated with the Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, Lund, Sweden.

‡Dr Persson is currently affiliated with the Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Lund, Sweden.

§The RIKEN Omics Science Center ceased to exist as of April 1, 2013, because of RIKEN reorganization.

Supported by a research grant for the RIKEN Omics Science Center from MEXT (to Y.H.) and grants from the Swedish Foundation for Strategic Research (RBc08-0027) and the Swedish Research Council (2009-5091; to J.K.). Patient sample collection was supported by grants from the Freemason Child House Foundation in Stockholm, the Konsul Th. C. Bergh's Foundation, the Swedish Asthma and Allergy Association's Research Foundation, the Centre for Allergy Research at Karolinska Institutet, the Swedish Heart-Lung Foundation, Karolinska Institutet, and the Bernard Osher Initiative for Research on Severe Asthma.

Disclosure of potential conflict of interest: J. R. Konradsen has received research support from Novartis; has received lecture fees from Novartis, Thermo Fisher Scientific, and Meda; and has received travel support from Thermo Fisher Scientific. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication December 19, 2013; revised November 23, 2014; accepted for publication February 3, 2015.

Corresponding author: Juha Kere, MD, PhD, Karolinska Institutet, Department of Biosciences and Nutrition and Center for Innovative Medicine (CIMED), Hälsovägen 7-9, SE-141 83 Huddinge, Sweden. E-mail: juha.kere@ki.se.

0091-6749

http://dx.doi.org/10.1016/j.jaci.2015.02.026

Abbreviations used
CA: Controlled persistent asthma
CAGE: Cap Analysis of Gene Expression
CTRL: Healthy control subject
GLM: Generalized linear model
GO: Gene Ontology
KEGG: Kyoto Encyclopedia of Genes and Genomes
LD: Linkage disequilibrium
lincRNA: Long intergenic noncoding RNA
MAPK: Mitogen-activated protein kinase
MARA: Motif Activity Response Analysis
NK: Natural killer
RORA: RAR-related orphan receptor A
SA: Therapy-resistant asthma
SNP: Single nucleotide polymorphism
TC: Tag cluster
TF: Transcription factor
TSS: Transcription start site

asthma[16] or SA,[17] for which few clinically well-characterized cohorts are available.[18]

Our aim in the present study was to identify the gene pathways that are deregulated in children with SA. Therefore we analyzed gene expression in peripheral blood leukocytes from children with SA (n = 13) and controlled persistent asthma (CA; n = 15), as well as age-matched healthy control subjects (CTRLs; n = 9). We used Cap Analysis of Gene Expression (CAGE), a technology based on next-generation sequencing that quantitatively measures gene expression for individual transcription start sites (TSSs), including mRNAs and long noncoding RNAs.[19] The asthmatic children were selected from Swedish Search, a national multicenter, cross-sectional study that aims to identify children with problematic severe asthma and characterize those who are difficult to treat or resistant to therapy.[20] Although these samples had previously been analyzed with expression microarrays,[21] we reasoned that the promoter-level resolution of CAGE sequencing data and the unbiased detection of novel genes and alternative promoters could provide new and interesting information.

METHODS

Children with CA (n = 15) and SA (n = 13) were selected from the Swedish Search study; details on inclusion criteria and clinical examination procedures are described elsewhere.[20,21] CTRLs (n = 9) were recruited at Astrid Lindgren Children's Hospital, Stockholm, Sweden, from children admitted for elective surgical procedures unrelated to asthma. Buffy coat isolation from blood samples and RNA extraction were described previously.[21] For statistical analysis of clinical parameters, normally distributed data are summarized by means and SDs and compared by using the Student *t* test. Nonnormally distributed data are presented as medians and interquartile ranges and compared with the Mann-Whitney *U* test. A comprehensive description of the methods used in this study is provided in the Methods section in this article's Online Repository at www.jacionline.org. Primers for validation of novel TSSs are listed in Table E1 in this article's Online Repository at www.jacionline.org. Informed consent was obtained from all participating children and their parents, and the study was approved by the local ethics committee.

RESULTS

We analyzed the transcriptome of peripheral blood leukocytes from children with CA (n = 15) and those with SA (n = 13), as well as age-matched CTRLs (n = 9) using CAGE sequencing. The asthmatic children were selected from the Swedish Search study,[20] and all cases of SA had been classified as SA in the absence of any identified aggravating factors. Most clinical variables did not differ between the CA and SA groups, but children with SA had significantly stronger responses to methacholine provocation and increased blood neutrophil counts compared with those with CA (Table I), as previously reported in children with severe asthma.[22] Leukocyte cell counts are provided for individual samples in Table E2 in this article's Online Repository at www.jacionline.org. A flow chart illustrating the different analyses performed in this article is provided in Fig E1 in this article's Online Repository at www.jacionline.org.

CAGE measures gene expression at the TSS of known and novel transcripts

CAGE sequencing detects the 5′ end of transcripts with a 5′ 7-methylguanosine cap, such as mRNAs and many long noncoding RNAs. Use of the HeliScope single-molecule sequencer (Helicos BioSciences, Cambridge, Mass) bypasses the need for PCR amplification of sample libraries to produce highly quantitative expression data.[19] Expression for a particular TSS is estimated by grouping nearby sequencing reads (tags) after mapping to the human genome to tag clusters (TCs) and counting the number of reads within each cluster. Sequencing library sizes are shown in Table E2. We limited expression analysis to a set of 45635 TCs that were detected by at least 3 reads in at least 9 samples (ie, the size of the smallest sample group). When compared with GENCODE[23] transcript annotation, 49% of the TCs mapped within 500 bp of a known TSS, and another 24% mapped in exons of known genes. A further 27% of the TCs were not located within exons of known coding or noncoding genes; these might represent novel genes or transcript isoforms. TC annotation is summarized in Fig E2 in this article's Online Repository at www.jacionline.org.

Children with CA and those with SA have distinct gene expression profiles and cluster separately

We performed differential expression analysis for all pairwise comparisons: children with CA versus CTRLs, children with SA versus CTRLs, and children with SA versus those with CA. A total of 1305 TCs were differentially expressed with a Benjamini-Hochberg–adjusted *P* value of less than .1 in any comparison, and 1029 of these mapped to 816 unique gene symbols (ie, some genes had >1 cluster). The remaining TCs were intergenic, intronic, or antisense to known transcripts. Hierarchical clustering of the samples based on differentially expressed TCs clearly separated the sample groups, with the exception of a single CTRL sample that clustered with the SA group (Fig 1). Statistically significant TCs with GENCODE transcript annotation are provided in Tables E3-E5 in this article's Online Repository at www.jacionline.org.

These results imply that differences in the underlying disease biology of CA and SA are reflected as distinct molecular fingerprints. However, because we measured gene expression for a complex population of leukocytes, the profiles also reflect differences in cell composition between sample groups, such as the significantly increased neutrophil counts in children with SA