To the Editor:

Airway epithelial cells are known to have an important role in allergic rhinitis (AR). They constitute the first line of defense against inhaled aeroallergens and are active mediators of innate and adaptive immune responses. Their aberrant functioning is linked with an intake of allergens, and their transcriptome is associated with seasonal AR and childhood rhinitis and asthma. However, the precise functions of epithelial host cells and respiratory microbes in AR are still largely elusive, especially during pollen allergen immunotherapy (AIT) that is associated with symptom reduction, decrease in allergen-specific biomarkers, and altered T- and B-cell responses.

We collected nasal brushings for RNA sequencing from 5 healthy subjects and 3 birch pollen AR patients with and without AIT at 2 springs and winters and studied seasonal, AR, and AIT-related alterations in the nasal epithelial and microbial transcriptomes (Fig 1, A; see Fig E1 and Table E1 in this article’s Online Repository at www.jacionline.org). Pollen count and AR symptom information were also assessed, revealing the presence of high amounts of birch pollen at spring samplings (Fig 1, B) and a marked improvement in quality of life in subjects with AR with AIT compared with controls (P < .005) and subjects with AR without AIT (P < .03) but not between other groups (Fig 1, C).

RNA sequencing resulted in 90 million mappable reads per sample on average. Of all the annotated human protein-coding genes, 17,347 were deduced expressed and 360 differentially expressed between different time points within groups and between different groups within time points (Fig 2, G). Identified were also 166 (Fig 2, A and B) and 17 (Fig 2, D and E) protein-coding genes with an altered expression between the consecutive springs and winters, respectively. Notably, we identified the greatest transcriptional reprogramming between springs in the AR-AIT group, indicating that AIT alters epithelial expression in the presence of allergens. Analyses also revealed 3 allergy-related pathways that were affected between the spring samplings. An asthma pathway was found to be altered in AR-noAIT subjects, whereas Toll-like receptor (TLR) and chemokine signaling pathways were both affected in AR-noAIT and AR-AIT subjects (Fig 2, C; see Fig E2 in this article’s Online Repository at www.jacionline.org). Pathway enrichment analysis of winter data revealed pathways with coordinated expression change only in healthy controls (Fig 2, F). Analysis of expressed variants pinpointed in turn 8 variants expressed in 2 or more subjects with AR at some time point but in none of the healthy controls (see Fig E3 in this article’s Online Repository at www.jacionline.org).

Further analysis of the gene expression profiles of the 3 allergy pathways between the spring samplings highlighted marked similarities in the AR-AIT and control groups that were not seen in the AR-noAIT group (Fig E2). These results imply that AIT may restore epithelial gene expression toward normal and indicate that effectiveness of AIT could be screened from nasal epithelium in addition to leukocytes. Specifically, the MHCII components were upregulated at the second spring in AR-AIT and control groups but not in the AR-noAIT group (Fig E2, A), indicating that AIT restores the compromised antigen-presenting capacity of epithelial cells in AR. We also found that genes that are downstream effectors of the chemokine signaling or pattern recognition and provide proinflammatory, antiviral, chemotactic, and T-cell stimulatory effects behaved alike between the AR-AIT and control groups (Fig E2, B and C). These findings are in line with the findings that changes in expression of TLR genes are associated with AR and suggest a role for TLR agonists in treatment of AR. Notably, the expression of several asthma-related genes was found to be in opposite between the AR-AIT and AR-noAIT subjects (Fig E2).

Microbial classification of sequencing data was performed to explore whether AR alters nasal microbiota (archaeal, bacterial, and viral) and whether AIT could restore microbial imbalances toward normal. On average, approximately 500 counts per million (~16,340 read-pairs) per sample were assigned to microbial taxa, 98.13% of which received a genus-level classification (see Fig E4, A, in this article’s Online Repository at www.jacionline.org). The classification showed that bacteria, archaea, and viruses were part of the active nasal microbiota, the most common genera being Bacillus (average abundance, 42.23%), Methanocaldococcus (average abundance, 35.72%), and Alphaproteobacterium (average abundance, 4.32%). Similar to previous studies, a large sample-to-sample variation was observed (Fig E4, A). Particularly, 6 samples taken at the second spring varied greatly from the rest (Fig E4, A and B) and were, for instance, the drivers of the greater abundance of viruses at the second spring compared with the other time points (Fig 2, H). Interestingly, examination of changes in species abundances (Fig E4, C) pinpointed Pseudomonas aeruginosa to be more abundant in the first spring in comparison to the second spring in the AR-AIT group.
We also computed alpha diversities to evaluate the effect of AR and AIT on the microbial diversity of nasal epithelia (see Figs E5-E7, A-N, in this article’s Online Repository at www.jacionline.org). This analysis revealed that control subjects primarily had the highest alpha diversity, differing from that seen previously in a study on seasonal AR but similar to that focusing on children with asthma and rhinitis. Interestingly, most diversity indices suggested an increase in diversity between the first and second winters in all groups. Most prominent was the increase in the AR-AIT group, whereas some increase was also detectable in the control and AR-noAIT groups (Fig E6, A-N). The diversity at the second winter in the AR-AIT group also changed more toward that of the control group than what was the corresponding change in the AR-noAIT group (Fig E6, A-N). These findings are largely in line with the previous studies noting that the bacterial diversity varies during the allergy season and suggest that AIT may increase microbial diversity and restore it toward normal.

Limitations of this study include the small subject number, lack of placebo group, differences in baseline allergic symptoms between the groups, differences in pollen seasons, differences in air quality, and technical differences in sampling, which may in part have compromised results. Yet, the study provided interesting insights into the epithelial transcriptome during AIT and revealed that AIT causes subtle but significant alterations in asthma, TLR signaling, and chemokine signaling-related genes and may as well recover microbiological diversity toward normal.

Seasonal heterogeneity represented the largest source of variation in transcriptomes, indicating a need for novel biomarkers in AIT treatment monitoring that accommodate inherent heterogeneity and seasonal variation.

Tanzeela Hanif, MSc
Kishor Dhaygude, MSc
Matti Kankainen, PhD
Jutta Renkonen, DDS
Pirkko Mattila, PhD
Teija Ojala, PhD
Sakari Joenväärä, MSc
Mika Mäkelä, MD
Anna Pelkonen, MD, PhD
Paula Kauppi, MD, PhD
Tari Haahtela, MD
Risto Renkonen, MD
Sanna Toppila-Salmi, MD, PhD

From the Haartman Institute, University of Helsinki, the Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, the Department of Pharmacology, University of Helsinki, the Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, and HUSLAB, Helsinki University Hospital, Helsinki, Finland. E-mail: sanna.salmi@helsinki.fi.

*Shared first author.
†Shared last author.

The study was supported in part by research grants from the Finnish Association of Otorhinolaryngology and Head and Neck Surgery, Finnish Medical Foundation, the Finnish Medical Society Duodecim, the Finnish Society of Allergology and Immunology, the Jane and Aatos Erkko Foundation, the Finnish Cultural Foundation, State
funding for university-level health research (grant no. TYH2018103), the Paulo Foundation, the Sigrid Juselius Foundation, the Tampere Tuberculosis Foundation, the Väinö and Laina Kivi Foundation, the Yrjö Jahnsson Foundation, the Academy of Finland (grant nos. 292605 and 292635), and Business Finland (grant no. Dnr no. 6113/31/2016).

Disclosure of potential conflict of interest: S. Toppila-Salmi has acted as a paid consultant for Mylan Laboratories Ltd, Biomedical systems Ltd, and Roche Products Ltd. The rest of the authors declare that they have no relevant conflicts of interest.

REFERENCES

To the Editor:

Recently, a novel syndrome of combined immunodeficiency, allergy, and “auto”inflammation caused by mutations in the ARPC1B gene has been reported. Analysis of patient-derived hematopoietic cells has shown a defect in actin polymerization, which resulted in a wide range of clinical manifestations and immunologic-hematologic features. We report on the immunologic, cellular, and molecular phenotypes in 14 patients with biallelic ARPC1B mutations and variable clinical presentations (Fig 1, A and B; see Fig E1, Table E1 in this article’s Online Repository at www.jacionline.org; for case descriptions, see this article’s Online Repository at www.jacionline.org), helping to delineate the broad spectrum of this novel disease and presenting unreported insights into cell-intrinsic defects involving regulatory T (Treg) cells and natural killer (NK) cells, potential players in the immune dysregulation and susceptibility to viral infections observed in these patients. The disease-causing variants are diverse and scattered throughout the gene (Fig E1, B; Table E1). Patient (P) 4, P12, and P14 have Nepalese ancestry and share the same variant, suggesting a founder mutation. In all patient samples tested, ARPC1B protein was undetectable by Western blotting and we identified an increased—although variable—expression of the ARPC1A isoform (see Fig E2 in this article’s Online Repository at www.jacionline.org).

The disease is characterized by a very early clinical onset (mean, 2 months; range, 1-6 months) (see Table E2 in this article’s Online Repository at www.jacionline.org). Presenting symptoms included skin rash, infections, and gastrointestinal bleeding. Most patients (79%) (Fig 1, A) suffered from recurrent or severe bleeding episodes, most frequently represented by enterorrhagia. Platelet counts were reduced (see Table E3 in this article’s Online Repository at www.jacionline.org), with normal volume in most cases. An increased rate and/or abnormal severity of respiratory...
METHODS

Subjects

Study subjects were recruited from the Skin and Allergy Hospital of Helsinki University Hospital. The study plan was approved by the ethical committee of Hospital District of Helsinki and Uusimaa, Finland (permission no. 19/13/003/00/11). Written informed consent was received from all subjects and their parents if the age of the participant was less than 18 years. The study has been registered in ClinicalTrials.com (no. NCT01985542). Baseline data of the study subjects are presented in Table E1. The total number of participants entering the study was 23 (Fig E1). The AR-AIT group received subcutaneous immunotherapy in November 2011 after the second sampling visit (Fig 1 and Fig E1), meaning that 2 samplings of the AR-AIT group were performed before and 2 during AIT.

Nasal brushings and RNA extraction

Nasal epithelial brushing was performed to middle meatus of both sides of nasal cavity after slight blowing of nose without local anesthesia as described. Epithelial cells were collected at 4 time points, washed once with ice-cold nuclease-free PBS, and resuspended immediately into RNA Later RNA stabilization reagent (Qiagen, Hilden, Germany) to preserve RNA profiles. The epithelial RNA isolation was done next day using Qiagen RNeasy Mini Kit with the optional DNase treatment included.

Library preparation and RNA sequencing

Agilent Bioanalyzer RNA Nano chip (Agilent, Santa Clara, Calif) was used to evaluate the integrity of RNA and Qubit RNA kit (Life Technologies, Carlsbad, Calif) to quantitate RNA in epithelial cell samples. If acceptable in quality (RIN value >7), 1.0 μg of total RNA sample was ribodepleted and prepared to RNA sequencing library by using ScriptSeq v2 Complete kit (Illumina, Inc, San Diego, Calif), RNA sequencing libraries were purifed with SPIR beads (Agencourt AMPure XP, Beckman Coulter, Brae, Calif). The library quality control was evaluated on high-sensitivity chips by Agilent Bioanalyzer (Agilent). Paired-end sequencing of sequencing libraries with 100-bp read length was performed using Illumina HiSeq technology (HiSeq 2000, Illumina Inc, San Diego, Calif). Planned read amount was 40 million reads per sample.

RNA sequencing data processing

RNA sequencing data were preprocessed as described previously. Briefly, Trimmomatic was used to correct read data for low quality, Illumina adapters, and short read length. Filtered paired-end reads were aligned to the human genome (GRCh38) using the STAR with the guidance of Ensembl v82 gene models. Default 2-pass per-sample parameters were used, except that the overhang on each side of the splice junctions was set to 99. The alignments were then sorted and PCR duplicates were marked using Picard, feature counts were computed using SubRead. Feature counts were converted to expression estimates using Trimmed Mean of M-values normalization, and lowly expressed genomic features with counts per million (CPM) value less than or equal to 1.00 in less than half the controls or birch pollen patients were removed. Default parameters were used, with the exception that reads were allowed to be assigned to overlapping genome features in the feature counting.

Host gene expression analysis

Differential expression testing was performed using the edgeR software and included testing of differential expression between and within groups at different sampling points. In the statistical testing, comparisons between subject groups used a combined factor of subject group and sampling point, whereas comparisons within subject groups also used a factor for the subject. The resulting P values were adjusted Storey’s Q-value approach, with significance defined as a Q value of less than or equal to 0.10. A cutoff value of absolute log2 fold-change of greater than or equal to 1.5 and Ensembl v82 biotype annotations were used as additional filters to select differentially expressed genes (DEGs) with protein-coding annotation for the downstream analysis. Heatmaps of differentially expressed protein-coding genes were produced with heatmap R package. Hierarchical clusters were generated using the spearman correlation and ward.D2 as the linkage method, with the exception of using ward.D2 and Euclidean distance for genes that were differentially expressed between different sampling years at springs and using complete linkage and spearman correlation for genes that were differentially expressed between different sampling years at winters. CPM data were used to generate heatmaps. Venn diagrams were generated using the VennDiagram R package. Functional profiles of DEGs were investigated with clusterProfiler using functions enrichGO and enrichKEGG. Outputs of enrichment analyses were visualized using dotplot function in clusterProfiler. Biologically relevant pathways found by clusterProfiler were visualized using pathview R package. In the process, Kyoto Encyclopedia of Genes and Genomes (KEGG) gene IDs of the selected pathways were fed along with log, fold-change values from relevant comparisons. Color codes on the pathway map were used to illustrate genes that were differentially expressed and the direction of their expression changes. Fold-change values beyond that range were truncated to the closest extreme; that is, values greater than 2 were truncated to 2, and values less than −2 truncated to −2. Downstream analyses were performed using R 3.3.1 with Bioconductor 3.0.

Variant analysis

Transcript variants were called from STAR alignments using the GATK best practices workflows for transcriptome data and then annotated using Annotate command defined previously. Quality control analyses were performed as defined previously. Variant calls were further filtered by accepting only those that were present in 2 or more AR cases, not present in any control case, and predicted to be pathogenic by various pathogen prediction methods part of the Annotator annotation tool. Heatmap was plotted using heatmap. The functional effects of variants were taken from AnnoVAR outputs. In addition, we plotted bpplot using CPM expression value of genes in healthy control and AR groups.

Microbial community profiling

Microbial community profiling was performed as previously described with some modifications. Specifically, RNA sequencing data were preprocessed for adapter trimming, low-quality bases filtering, and removal of reads less than 36 bp in length by using Trimmomatic. Paired-end reads passing the preprocessing were mapped against rRNA sequences from RPAM v12.3 using the Burrows-Wheeler Aligner with default settings, and reads-matching rRNAs were filtered by using samtools. Centrifuge was then used to classify paired-end reads to microbial taxa. Alignment data were converted to kraken-style output. In the classification, reads were aligned against 27,127 known complete bacterial, archaeal, and viral genome assemblies, the human genome, and 10,615 technical artifact sequences that were available in the RefSeq database at February 2018. Default parameters were used, with the exception that only 1 (ie, the lowest common ancestor) assignment was reported for read-pairs with multiple primary assignments. Taxa having less than 100 read-pairs assigned to them in any sample were removed. Pairwise comparisons between and within groups at different sampling points were performed by applying DeSeq on the number of reads covered by the clade rooted at the given taxon level. In the analyses, size factors were estimated by using the poscounts method, comparisons between subject groups were done using a combined factor of subject group and sampling point, and comparisons within subject groups with a model where individuals were nested within subject groups. Each taxonomic level was analyzed separately and variance stabilizing transformation was used to generate expression estimates for heatmap visualizations. The Storey’s Q-value adjustment was used to correct data for multiple hypothesis testing, with significance defined as a Q value of less than or equal to 0.05. Finally, alpha diversity (Observed, Chao1, ACE, Shannon, Simpson, InvSimpson, and Fisher), beta diversity (Bray-Curtis dissimilarity), and rarefaction analyses were done using the Phyloseq software applied on the number of reads assigned directly to the given taxonomic level.
RESULTS

General

Subjects with AR, and especially AR-AIT cases, had higher median of total serum IgE, median of birch-specific serum IgE and skin prick test wheel diameter to birch and symptom scores during samplings (Table E1 and Fig 1). The AR-AIT group reported benefit (Fig 1) and reported no severe side effects at the end of subcutaneous immunotherapy 3 years after start of subcutaneous immunotherapy (data not shown).

Transcriptome of nasal epithelium

We generated in total 5164 million raw paired-end transcript reads. Manual inspection of the quality plots generated using FASTQ indicated that sequencing data were of excellent quality. On average, 90 million reads were mapped to the reference per subject. Mapped reads were then used to generate CPM expression estimates, revealing expression of 13,873 protein-coding genes among healthy controls and 17,347 across all the 44 samples. Altogether, expression of 34,896 genes was found. To gain insight into cellular processes dysregulated in AR and AIT, DEGs between sample groups were identified using edgeR.17 In this analysis, we identified altogether 360 genes to be differentially expressed with the \( \log_2 \) fold-change greater than or equal to 1.5 between different time points within group and between different groups within time points.

Alterations in gene expression profiles in the 2 following springs and winters

Comparison of the transcriptome profiles between the springs revealed 119 DEGs between the samplings in the AR-AIT group, 49 between the samplings in the AR-noAIT group, and 27 between the samplings in healthy controls (Fig 2, B), which suggests that the greatest transcriptional reprogramming took place in the AR-AIT group followed by the AR-noAIT group and healthy subjects. Comparison of the 2 consecutive winters revealed only 17 DEGs among healthy controls and none among patients with AR, suggesting that AIT alters epithelial expression only in the presence of allergens (Fig 2, E).

Differential expression of immune response and signaling pathways

We performed KEGG pathway enrichment analysis to discover functional themes shared by DEGs. This analysis revealed altogether 21 KEGG pathways with coordinated expression change between the spring samplings (Fig 2, C). Of these, 4 were associated with genes differentially expressed in the AR-AIT group, including the chemokine signaling and TLR signaling pathway (Fig 2, C). Genes differentially expressed in the AR-noAIT group were in turn associated with 8 pathways, including IL-17 signaling and asthma pathway that were discovered only in this comparison (Fig 2, C). The healthy group genes were enriched in 11 KEGG pathways (Fig 2, C). Altogether we detected 3 allergy-related pathways, of which asthma was discovered only in the AR-noAIT comparison and TLR signaling and chemokine signaling pathways were discovered in AR-noAIT and AR-AIT comparisons (Fig 2, C). Pathway enrichment analysis of winter comparison data revealed pathways with coordinated expression change only in healthy controls (Fig 2, F).

In-depth analysis of allergy-related pathways

Allergy-related pathways found in the pathway analysis (Fig E2) were analyzed more in-depth to study these mechanisms. First, the asthma pathway consisted in total 3 DEGs. The AR-AIT group and healthy controls displayed upregulation of MHCII and downregulation of FcerI at the second spring (Fig E2, A). The expression of various other members of the pathway was altered, although un-significantly, between spring samplings (Fig E2, A). These more border-line findings included IL-13, which is a T-cell–specific transcription factor, and interleukin IL-4, which is IgE synthesis switch factor,18 and IL-5, which is an eosinophil growth factor.19 Expression of genes of the asthma pathway between time points occurred to opposite directions in AR-AIT and AR-noAIT groups (Fig E2, A). The second pathway of interest was chemokine signaling pathway. This pathway consisted of 49 dysregulated genes (Fig E2, B), including various chemokines and chemokine receptors. Expression of genes between time points occurred to opposite directions in AR-AIT and AR-noAIT groups (Fig E2, B). The third pathway of relevance to AR was the the TLR signaling pathway (Fig E2, C) consisting altogether of 67 dysregulated genes, such as p38, TNF-α, IL-12, and INF-α genes. Expression of genes between springs happened to opposite directions in AR-AIT and AR-noAIT groups (Fig E2, C).

Transcript variants expressed in patients with AR

The GATK best practice for RNA sequencing12 was used to identify expressed variants. This approach revealed altogether 3,268,177 (on average 74,277 per subject) variants passing GATK filters. Removal of intronic and silent variants and polymorphisms resulted in 8174 (on average 186 per subject) variants that were further narrowed down to 8 potential candidates expressed in at least 2 subjects with AR at any time point but in none of the healthy control samples (Fig E3).

Functional characterization of microbiome showed no AR-related changes

The effect of AR and AIT on the active nasal microbiota was studied by identifying microbial reads from the RNA sequencing data and performing microbial classification. On average approximately 16,340 read-pairs (500 CPMs) per sample were assigned to bacterial, archaeal, or viral taxa with a high sample-to-sample variation (minimum of 791 reads [24 CPMs] and maximum of 69,428 reads [1,791 CPMs]). Of these microbial reads, on average approximately 90.13% were classified at the genus level and 67.08% at the species level. The genus-level classifications (Fig 2, H, and Fig E4) showed that overall the most abundant genera were Bacillus, Methanocaldococcus, and Alpharetrovirus, with average relative proportions of approximately 42.23%, approximately 35.72%, and approximately 4.32%, respectively, and other genera having average relative proportions of approximately 1.57% (Acinetobacter) or less. The relative proportions of the detected genera varied greatly between the samples, with Bacillus demonstrating the greatest variation (relative proportions ranging from 0% to ~63.76%). The relative proportion of viruses, particularly Alpharetrovirus, was greater in the second spring sampling point than in the other sampling points for all the groups, whereas the relative proportion of Bacillus was reduced (Fig 2, H, and Fig E4, A). This variation, however, was mainly driven by the second spring samples of 6 cases, who were distributed among all groups.
(Fig E4, A). Examination of the community compositional differences between the samples revealed these 6 samples to be the most dissimilar from the rest, which in turn had rather similar community compositions (Fig E4, B). To further investigate microbial richness, we estimated the alpha diversity measures for all the samples (Figs E5-E7, A-N). Most indices demonstrated an increase in the diversity for the AR-AIT group when the second winter was compared with the first (Fig E6, A-N). Some change in diversity between second and first winters was also seen in the control group, whereas some indices also indicated an increase for the AR-noAIT group. The increase seen in the AR-noAIT group was, however, less than that observed for the AR-AIT control group, whereas some indices also indicated an increase for the AR-AIT group when the second winter was compared with the first (Fig E6, A-N). Finally, examination of significant (Q value ≤ 0.05) changes in species abundances (Fig E4, C) revealed more changes in all groups between the 2 spring sampling points than between the winter sampling points. The winter comparisons revealed only increased species for the control and AR-AIT groups, whereas for the AR-noAIT group only decreased species were reported. The AR-AIT group had significantly less *Pseudomonas aeruginosa* in the second spring sampling point than the first. For the other groups, the comparisons revealed no significant changes in the abundance of this bacterium.

We acknowledge Anne-Maria Konkola, Docent Sanna Korkonen, Docent Pekka Malnberg, Leena Petman, Mirja-Liisa Sipola, BDent Emma Terna, Tanja Urttiainen, and Docent Jan Weckström, personnel at the Sequencing and Bioinformatics Units at FIMM Technology Centre supported by the Helsinki Institute of Life Science and Biocenter Finland, and the volunteer subjects and their family members for making this study possible.

REFERENCES


FIG E1. Flowchart of the study. The total number of participants entering the study was 23. Initially, 7 patients with AR were assigned to pollen AIT (AR-AIT group) and 9 others to conventional therapy (AR-noAIT group). The study also included 7 control subjects (control group) without allergy. One participant in the AR-AIT group discontinued the study before starting AIT because of a diagnosis of a cardiac disease requiring surgery. The other participants in the AR-AIT group fulfilled the whole AIT scheme of 3 years, with the total dose being more than 2000.000 standardized quality (SQ). AIT was performed according to standard 3-year protocol or normal 3-year scheme. Subcutaneous injections of birch pollen extract (Betula verrucosa, ALK-Abello) were administered in the clinic and included an induction phase with increasing dosing starting with a dose of 20 SQ. The maintenance-phase dose was 100,000 SQ. Six subjects discontinued the study before the last follow-up. Moreover, 5 of the 16 subjects who completed the study had poor RNA quality in at least 1 nasal epithelial sample, leading to their exclusion. Thus, a total of 11 subjects (5 healthy controls, 3 AR-AIT, and 3 AR-noAIT) completed the study. These participants were adults and of white European ancestry except 1 Chinese male in the AR-noAIT group. NGS, Next Generation Sequencing; SPT, skin prick test.
FIG E2. Allergy-related pathways identified in the pathway analysis. Gene expression profiles of (A) asthma, (B) chemokine signaling, and (C) TLR signaling pathways among patients with AR with pollen AIT, patients with AR without AIT, and control subjects. Boxes in the figure represent genes. The gradient colors indicate the log2 fold-change of the gene between the spring samplings in AR-AIT (I/left), AR-noAIT (II/middle), and control (III/right) groups. Fold-change values beyond the range (from -2 to 2) were truncated to the closest extreme; that is, values greater than 2 were truncated to 2, and values less than -2 truncated to -2. Asterisk indicates statistically
A significant difference with a $P$ value of less than or equal to 0.10. Green gradient colors indicate upregulation of the gene at the second spring, and red colors indicate upregulation at the first spring. AC, Adenyl cyclase; AKT, AKT Serine/Threonine Kinase; AP-1, Activator protein 1; APC, Antigen presenting cell; BAD, Bcl-2-antagonist of cell death; BCR, BCR, RhoGEF And GTPase Activating Protein; Ca$^{2+}$, Calcium ion; CASP8, Caspase 8; CD14, Monocyte differentiation antigen CD14; CD40, Tumor necrosis factor receptor superfamily member 5; CD40L, Tumor necrosis factor ligand superfamily member 5; CD80, CD80 antigen; CD86, CD86 antigen; CD642, Cell division control protein 42; Crk, Crk Proto-Oncogene, Adaptor Protein; DAG, Diacylglycerol; DC, Dendritic cell; DOCK2, Cdc42 and Rac guanine exchange protein; ECP, Eosinophil cationic protein; ELMO1, Engagement and cell motility protein 1; EPO, Eosinophil peroxidase; ERK, Mitogen-activated protein kinase; FADD, FAS-associated death domain protein; FAK, Focal adhesion kinase; FC, High affinity immunoglobulin epsilon receptor; FOXO, Forkhead box protein O; G, Alpha-beta-gamma subunit; G, Beta-gamma complex; GRB2, Growth factor receptor-binding protein 2; GRK, Rhodopsin kinase; GSK3, Glycogen synthase kinase 3; IFN, Interferon alpha; IFNβR, Interferon receptor; IFNβ, Interferon beta; IgE, Immunoglobulin E; I kappaB, NFκB Inhibitor Alpha; IKK, Inhibitor of nuclear factor kappa-B; IKKα, Inhibitor of nuclear factor kappa-B kinase subunit alpha; IKKβ, Inhibitor of nuclear factor kappa-B kinase subunit beta; IKKε, Inhibitor of nuclear factor kappa-B kinase subunit epsilon; IKKγ, Inhibitor of nuclear factor kappa-B kinase subunit gamma; IP10, Interferon Gamma-Induced Protein 10; IP3, Inositol trisphosphate; IRAK1, Interleukin-1 receptor-associated kinase 1; IRS, Interferon regulatory factor; i-TAC, Interleukin-6-like Inducible T-cell alpha chemotactic; ITK, I22-inducible T-cell kinase; Jak, Janus kinase; JNK, Mitogen-activated protein kinase B9/10 (c-Jun N-terminal kinase); LBP, Lipopolysaccharide-binding protein; LTD4, Leukotriene C4; LTA, Leukotriene D4; MBP, Bone marrow proteoglycan; MD2, Lymphocyte antigen 96; MEK1, Mitogen-activated protein kinase 1; MEK2, Mitogen-activated protein kinase; MHCII, Major histocompatibility complex, class II; MG, Monokine induced by gamma interferon; MIP1, MCK2-interacting protein; MKK, Mitogen-activated protein kinase; MYD88, Myeloid differentiation primary response protein; NADPH, 47-kilodalton cytosolic subunit of the multi-protein complex known as NADPH oxidase; NFκB, Nuclear factor NF-kappa-B; NFκB, Nuclear factor NF-kappa-B; NO, Nitric oxide; OPN, Secreted phosphoprotein 1; p105, NF-kappa-B p105 subunit; p130Cas, Cas Family Scaffold Protein; P13K, Phosphoinositide-3-kinase; P38, p38 MAP kinase; p47phox, NADPH oxidase; PAF, Platelet activating factor; PKA1, p21-activated kinase 1; PAR3, Partitioning defective protein 3; PGD2, Prostaglandin D2; PI3K, Phosphoinositide 3-kinase; PI3P, Phosphatidylinositol 3,4,5-trisphosphate; PKA, Protein kinase A; PKK, Protein kinase C; PLC, Phospholipase C; PI3K, Phosphoinositide-3-kinase; pIκB, Interferon beta; TAK1, Mitogen-activated protein kinase 2; TANK, TAK1-binding protein; TAK1, Mitogen-activated protein kinase kinase kinase 7; TANK, TANK-binding kinase 1; TH1, Th1 cytokine (Th1); TH2, Th2 cytokine (Th2); TLR, Toll-like receptor; TNFα, Tumor necrosis factor alpha; TOLLIP, Toll-interacting protein; Tsp2, Tumor progression locus 2; TRAF3, TNF receptor-associated factor 3; TRAF6, TNF receptor-associated factor 6; TRAM, Translocating chain-associated membrane protein; TRIF, Toll-like receptor adapter molecule 1; WASP, Wiskott-Aldrich syndrome protein; VAV, Guanine nucleotide exchange factor VAV.

**Figure 3.** Landscape of expressed variants in AR. Short nonsynonymous transcript variants identified in at least 2 subjects with AR but in none of the control cases. Variants are grouped by gene. Column annotations from top to bottom: subject group, sampling season, and sampling year. Barplot at the right indicate the average expression level of the gene in subjects with AR and control subjects expressed as CPM.
FIG E4. Microbial variation. A, Relative abundances of microbial (archaeal, bacterial, and viral) genera and total microbial load per sample. Only genera accounting for more than 5% of the total microbial load are shown. The line denotes the total number of microbial reads per sample expressed as CPM. B, Principal co-ordinates analysis plot of microbial community structure based on Bray-Curtis distance. C, A heatmap of differentially abundant microbial species ($Q$ value < 0.05) between any season- and group-matched year comparison. The different shades of blue illustrate the variance stabilizing transformation values. Darker shades of blue indicate higher values. The species were hierarchically clustered using average linkage and distance metric Pearson correlation. Acronym S1 in the sample name stands for first spring (May), S2 for the second spring, W1 for the first winter (November), and W2 for the second winter. Comparisons that reached the statistical significance are indicated by black boxes next to the heatmap.
FIG E5. Alpha diversity of epithelial microbiotas across study groups during winter and spring. Alpha diversity indices of different sample groups using various alpha diversity metrics. Shown in the figure are Fisher (A and B), ACE (C and D), Chao1 (E and F), observed (G and H), Fisher (I and J), InvSimpson (K and L), and Simpson (M and N) metrics with (Fig E5, B, D, F, H, J, L, and N) and without (Fig E5, A, C, E, G, I, K, and M) rarefaction of samples to the minimum sampling depth. In the figure, Nov1 stands for the first winter, Nov2 for the second winter, May1 for the first spring, and May2 for the second spring sampling point. ACE, Abundance-based coverage estimators.
FIG E6. Alpha diversity of epithelial microbiotas between groups during winter. Alpha diversity indices of different groups during 2 consecutive winters using various alpha diversity metrics. Shown in the figure are alpha diversities computed using Shannon (A and B), ACE (C and D), Chao1 (E and F), observed (G and H), Fisher (I and J), InvSimpson (K and L), and Simpson (M and N) metrics with (Fig E6, B, D, F, H, J, L, and N) and without (Fig E6, A, C, E, G, I, K, and M) rarefaction of samples to the minimum sampling depth. In the figure, Nov1 stands for the first winter and Nov2 for the second winter sampling point. ACE, Abundance-based coverage estimators.
FIG E7. Alpha diversity of epithelial microbiotas between groups during spring. Alpha diversity indices of different groups during 2 consecutive springs sampling using various alpha diversity metrics. Shown in the figure are alpha diversities computed using Shannon (A and B), ACE (C and D), Chao1 (E and F), observed (G and H), Fisher (I and J), InvSimpson (K and L), and Simpson (M and N) metrics with (Fig E7, B, D, F, H, J, L, and N) and without (Fig E7, A, C, E, G, I, K, and M) rarefaction of samples to the minimum sampling depth. In the figure, May1 stands for the first spring and May2 for the second spring sampling point. ACE, Abundance-based coverage estimators.
# TABLE E1. Baseline characteristics of the subjects

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All subjects</th>
<th>Patients with AR</th>
<th>$P$ value</th>
<th>AR-noAIT</th>
<th>AR-AIT</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of subjects</td>
<td>5</td>
<td>6</td>
<td>.43</td>
<td>3</td>
<td>3</td>
<td>.70</td>
</tr>
<tr>
<td>Age (y), median (IQR)</td>
<td>44 (39-48)</td>
<td>39 (34-46)</td>
<td></td>
<td>34 (32-42)</td>
<td>41 (39-43)</td>
<td>.70</td>
</tr>
<tr>
<td>Men/women (n)</td>
<td>1/4</td>
<td>3/5</td>
<td>.54</td>
<td>2/1</td>
<td>1/2</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Spirometry values (% predicted), median (IQR)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV$_1$ baseline</td>
<td>107 (105-109)</td>
<td>95 (94-96)</td>
<td>.23</td>
<td>96 (95-103)</td>
<td>97 (94-95)</td>
<td>.30</td>
</tr>
<tr>
<td>FEV$_1$/FVC ratio baseline</td>
<td>110 (105-110)</td>
<td>97 (95-99)</td>
<td>.39</td>
<td>97 (96-104)</td>
<td>97 (95-99)</td>
<td>1.00</td>
</tr>
<tr>
<td>FEV$_1$/FVC bronchodilation</td>
<td>104 (101-104)</td>
<td>98 (98-103)</td>
<td>.33</td>
<td>103 (101-107)</td>
<td>95 (92-98)</td>
<td>.30</td>
</tr>
<tr>
<td>PEF baseline</td>
<td>109 (102-125)</td>
<td>109 (101-110)</td>
<td>.57</td>
<td>110 (106-116)</td>
<td>98 (87-109)</td>
<td>.20</td>
</tr>
<tr>
<td>PEF bronchodilation</td>
<td>104 (92-125)</td>
<td>112 (92-114)</td>
<td>.74</td>
<td>112 (102-113)</td>
<td>100.5 (86-115)</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Symptoms during sampling in spring$_1$, median (IQR)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total RQLQ score</td>
<td>0 (0-6)</td>
<td>48.5 (24-74)</td>
<td>.019</td>
<td>40 (23-49)</td>
<td>74 (49-87)</td>
<td>.40</td>
</tr>
<tr>
<td>Total VAS score</td>
<td>15 (10-36)</td>
<td>744 (358-1352)</td>
<td>.009</td>
<td>217 (358-471)</td>
<td>1352 (1128-1512)</td>
<td>.10</td>
</tr>
<tr>
<td><strong>Serum values in spring$_1$, median (IQR)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total IgE (kU/L)</td>
<td>9 (7-14)</td>
<td>61.5 (39-200)</td>
<td>.015</td>
<td>39 (26-275)</td>
<td>72 (62-136)</td>
<td>.70</td>
</tr>
<tr>
<td>Birch specific IgE (kU/L)</td>
<td>0 (0-0)</td>
<td>16.4 (3.3-24.2)</td>
<td>.004</td>
<td>12 (6-16)</td>
<td>24 (14-28)</td>
<td>.40</td>
</tr>
<tr>
<td>Timothy specific IgE (kU/L)</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>.85</td>
<td>0 (0-45)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
<tr>
<td>SPT wheal diameter (mm), median (IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative control</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Histamine (positive control)</td>
<td>5 (5-5)</td>
<td>5 (5-5)</td>
<td>1.00</td>
<td>5 (5-5)</td>
<td>5 (5-5)</td>
<td>1.00</td>
</tr>
<tr>
<td>Birch pollen</td>
<td>0 (0-0)</td>
<td>5 (4-6)</td>
<td>.004</td>
<td>4 (3.5-6)</td>
<td>5 (5-5.5)</td>
<td>.60</td>
</tr>
<tr>
<td>Timothy grass pollen</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
<td>0 (0-2.5)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Festuca pratensis pollen</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
<td>0 (0-2.5)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Mugwort pollen</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
<td>0 (0-2.5)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Cladosporium herbarum</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Cat dander</td>
<td>0 (0-0)</td>
<td>0 (0-5)</td>
<td>.46</td>
<td>0 (0-2.5)</td>
<td>0 (0-3)</td>
<td>1.00</td>
</tr>
<tr>
<td>Dog dander</td>
<td>0 (0-0)</td>
<td>0 (0-4)</td>
<td>.46</td>
<td>0 (0-2)</td>
<td>0 (0-2.5)</td>
<td>1.00</td>
</tr>
<tr>
<td>Horse dander</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Dermatophagoides pteronyssinus</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>1.00</td>
</tr>
</tbody>
</table>

FVC, Forced vital capacity; IQR, interquartile range; PEF, peak expiratory force; RQLQ, Rhinoconjunctivitis Quality of Life Questionnaire; SPT, skin prick test; VAS, visual analogue scale.

Diagnosis of AR was based on a typical history, SPT (ALK-Abelló, Hørsholm, Denmark), total serum IgE, and serum birch and timothy allergen specific IgE antibodies. Healthy volunteers did not have symptoms and were negative for SPT of common aeroallergens and serum birch and timothy allergen specific IgE antibodies. Exclusion criteria were age less than 12 y, use of tobacco products, non-AR, AR symptoms caused by other than seasonal allergens, asthma, and general disease requiring regular medication. Asthma was excluded by absence of typical symptoms and by normal values in spirometry with bronchodilation test. One subject starting pollen AIT was tested for bronchial hyperresponsiveness by histamine challenge test. He had normal 15th percentile density of FEV$_1$ result instead of bronchodilation test. Questionnaire regarding baseline characteristics and symptoms was collected at each sampling visit. We used the 28-item RQLQ. Subjects filled the VAS so that they were without medication. The questionnaire included 18 questions concerning airway symptoms and 23 questions concerning general health. Value 0 (mm) indicated no symptoms, and value 100 (mm) indicated the worst case. The total maximum score of the 41 questions was 4100. In the analysis, VAS scores less than or equal to 3 mm were regarded as 0. $P$ values were computed by Kruskal-Wallis and Mann Whitney $U$ tests (continuous variables) or by Fisher exact test (dichotomous variables).