Accepted Manuscript

Birch pollen allergen immunotherapy reprograms nasal epithelial transcriptome and recovers microbial diversity

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

PII: S0091-6749(19)30204-0

DOI: https://doi.org/10.1016/j.jaci.2019.02.002

Reference: YMAI 13889

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 19 July 2018

Revised Date: 29 January 2019

Accepted Date: 1 February 2019

Please cite this article as: Hanif T, Dhaygude K, Kankainen M, Renkonen J, Mattila P, Ojala T, Joenvaara S, Mäkelä M, Pelkonen A, Kauppi P, Haahtela T, Renkonen R, Toppila-Salmi S, Birch pollen allergen immunotherapy reprograms nasal epithelial transcriptome and recovers microbial diversity, *Journal of Allergy and Clinical Immunology* (2019), doi: https://doi.org/10.1016/j.jaci.2019.02.002.

Birch pollen allergen immunotherapy reprograms nasal epithelial transcriptome and recovers microbial diversity

Tanzeela Hanif, MSci[1*], Kishor Dhaygude, MSci[1*], Matti Kankainen, PhD[2,3], Jutta Renkonen, DDS[1], Pirkko Mattila, PhD[2], Teija Ojala, PhD[4], Sakari Joenvaara, MSci[1], Mika Mäkelä, Prof[5], Anna Pelkonen, MD PhD[5], Paula Kauppi, MD PhD[5], Tari Haahtela, Prof[5], Risto Renkonen, Prof[1,6#], Sanna Toppila-Salmi MD PhD[1, 5#]

[1] Haartman Institute, University of Helsinki, Helsinki, Finland

[2] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

[3] Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[4] Department of Pharmacology, University of Helsinki, Helsinki, Finland

[5] Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[6] HUSLAB, Helsinki University Hospital, Helsinki, Finland

*Shared first author

#Shared last author

Address correspondence and reprint requests to Sanna Toppila-Salmi, Haartman Institute, University of Helsinki, Haartmaninkatu 3, 00014 University of Helsinki, Helsinki, Finland

Phone number: +358 505431421 Fax number +358 9 471 86476

E-mail address: sanna.salmi@helsinki.fi

Conflicts of interest

STS has acted as paid consultant for Mylan Laboratories Ltd., Biomedical systems Ltd. and Roche Products Ltd. All other authors declare no conflicts of interest.

Funding statement

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 (TYH2018103), Paulo Foundation, Sigrid Juselius Foundation, the Tampere Tuberculosis Foundation, the Väinö and Laina Kivi Foundation, the Yrjö Jahnsson Foundation, Academy of Finland (grant no. 292605 and 292635), and Business Finland (Dnro 6113/31/2016).

Key messages / Clinical implications

- Nasal epithelial transcriptome changes in response to season
- Pollen allergen immunotherapy (AIT) alters expression of asthma, chemokine signaling, and toll like receptor signaling related genes
- AIT increases microbial community diversity
- RNA-sequencing enables integrated analysis of microbe and host transcriptomes

Capsule summary

Nasal epithelial transcriptome is altered by the season. Birch pollen allergen immunotherapy recovers microbial community diversity and alters expression of allergy related genes.

Key words

Allergic rhinitis, birch pollen, immunotherapy, nasal epithelium, next generation sequencing, transcriptome

To the Editor,

Airway epithelial cells are known to have an important role in allergic rhinitis (AR) (1-3). They constitute the first line of defense against inhaled aeroallergens and are active mediators of innate and adaptive immune responses (3). Their aberrant functioning is linked with an intake of allergens (2) and their transcriptome is reprogrammed under exposure to pollens (2-3) as well as in AR (3) and atopic asthma (1). Furthermore, epithelial cells interact with and are involved in generating an environmental niche for the respiratory microbiota, whose imbalance has been associated with seasonal AR (4) and childhood rhinitis and asthma (5). 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 (6), decrease in allergen-specific biomarkers, and altered T- and B-cell responses (7).

We collected nasal brushings for RNA-sequencing from five healthy subjects and three birch pollen AR patients with and without AIT at two springs and winters and studied seasonal, AR, and AIT-related alterations in the nasal epithelial and microbial transcriptomes (Fig 1, A, Fig E1, Table E1). Pollen count and AR symptom information was also assessed, revealing the presence of high amounts of birch pollen at spring samplings (Fig 1, B) and a marked improvement of quality of life in AR subjects with AIT compared to controls (p-value < 0.005) and AR subjects without AIT (p-value < 0.03) 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 timepoints within groups and between different groups within timepoints (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 three allergy related pathways that were affected between the spring samplings. An asthma pathway was found to be altered

in AR-noAIT subjects, whereas TLR (Toll like receptor) and chemokine signaling pathways were both affected in AR-noAIT and AR-AIT subjects (Fig 2, C, and Fig E2). 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 eight variants expressed in two or more AR subjects at some time point but in none of the healthy controls (Fig E3).

Further analysis of the gene expression profiles of the three 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 restores epithelial gene expression towards normal and indicate that effectivity of AIT could be screened from nasal epithelium in addition to leukocytes. Specifically, the MHCII components were up-regulated at the second spring in AR-AIT and control 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, C). These findings are in line with the findings that changes in expression of TLR genes are associated with allergic rhinitis and suggest a role for TLR agonists in treatment of AR (3, 7). Notably, 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 towards normal. On average, ~500 CPMs (~16,340 read-pairs) per sample were assigned to microbial taxa, 98.13% of which received a genus-level classification (Fig E4, A). 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 *Alpharetrovirus* (average abundance 4.32%). Similar to previous studies (8), a large sample-to-sample variation was observed (Fig E4, A). Particularly, six samples taken at the second spring varied greatly from the rest (Fig E4, A and E4, B) and were, for instance, the drivers of the greater abundance of viruses at the second spring compared to the other timepoints (Fig 2, H). Interestingly, examination of changes in species abundancies (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 (Fig E5-E7, A-N). This analysis revealed that control subjects primarily had the highest alpha diversity, differing from that seen previously in a study on seasonal allergic rhinitis (4) but similar to that focusing on children with asthma and rhinitis (5). Interestingly, majority of the diversity indices suggested an increase of diversity between the first and second winter in all groups. Most prominent the increase was in the AR-AIT group, while 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 towards 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 allergy season (4) and suggest that AIT may increase microbial diversity and restore it towards 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 towards 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*, MSci, Haartman Institute, University of Helsinki, Helsinki, Finland

Kishor Dhaygude*, MSci, Haartman Institute, University of Helsinki, Helsinki, Finland

Matti Kankainen, PhD, Institute for Molecular Medicine Finland (FIMM), University of Helsinki, and Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Jutta Renkonen, DDS, Haartman Institute, University of Helsinki, Helsinki, Finland

Pirkko Mattila, PhD, Finnish Institute of Molecular Medicine, University of Helsinki, Helsinki, Finland

Teija Ojala, PhD, Department of Pharmacology, University of Helsinki, Helsinki, Finland

Sakari Joenvaara, MSci, Haartman Institute, University of Helsinki, Helsinki, Finland

Mika Mäkelä, Prof, Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Anna Pelkonen, MD PhD, Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Paula Kauppi, MD PhD, Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Tari Haahtela, Prof, Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Risto Renkonen#, Prof, Haartman Institute, University of Helsinki and HUSLAB, Helsinki University Hospital, Helsinki, Finland

Sanna Toppila-Salmi# MD PhD, Haartman Institute, University of Helsinki, and Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

*Shared first author

#Shared last author

References

(1) Poole A, Urbanek C, Eng C, Schageman J, Jacobson S, O'Connor BP, et al. Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. J Allergy Clin Immunol 2014 Mar;133(3):670-8.e12.

(2) Joenvaara S, Mattila P, Renkonen J, Makitie A, Toppila-Salmi S, Lehtonen M, et al. Caveolar transport through nasal epithelium of birch pollen allergen Bet v 1 in allergic patients. J Allergy Clin Immunol 2009 Jul;124(1):135-142.e1-21.

(3) Toppila-Salmi S, van Drunen CM, Fokkens WJ, Golebski K, Mattila P, Joenvaara S, et al. Molecular mechanisms of nasal epithelium in rhinitis and rhinosinusitis. Curr Allergy Asthma Rep 2015 Feb;15(2):495-014-0495-8.

(4) Choi CH, Poroyko V, Watanabe S, Jiang D, Lane J, deTineo M, et al. Seasonal allergic rhinitis affects sinonasal microbiota. Am J Rhinol Allergy 2014 Jul-Aug;28(4):281-286.

(5) Chiu CY, Chan YL, Tsai YS, Chen SA, Wang CJ, Chen KF, et al. Airway Microbial Diversity is Inversely Associated with Mite-Sensitized Rhinitis and Asthma in Early Childhood. Sci Rep 2017 May 12;7(1):1820-017-02067-7.

(6) Bousquet J, Khaltaev N, Cruz AA, Denburg J, Fokkens WJ, Togias A, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). Allergy 2008 Apr;63 Suppl 86:8-160.

(7) Akdis CA, Akdis M. Mechanisms of allergen-specific immunotherapy. J Allergy Clin Immunol 2011 Jan;127(1):18-27; quiz 28-9.

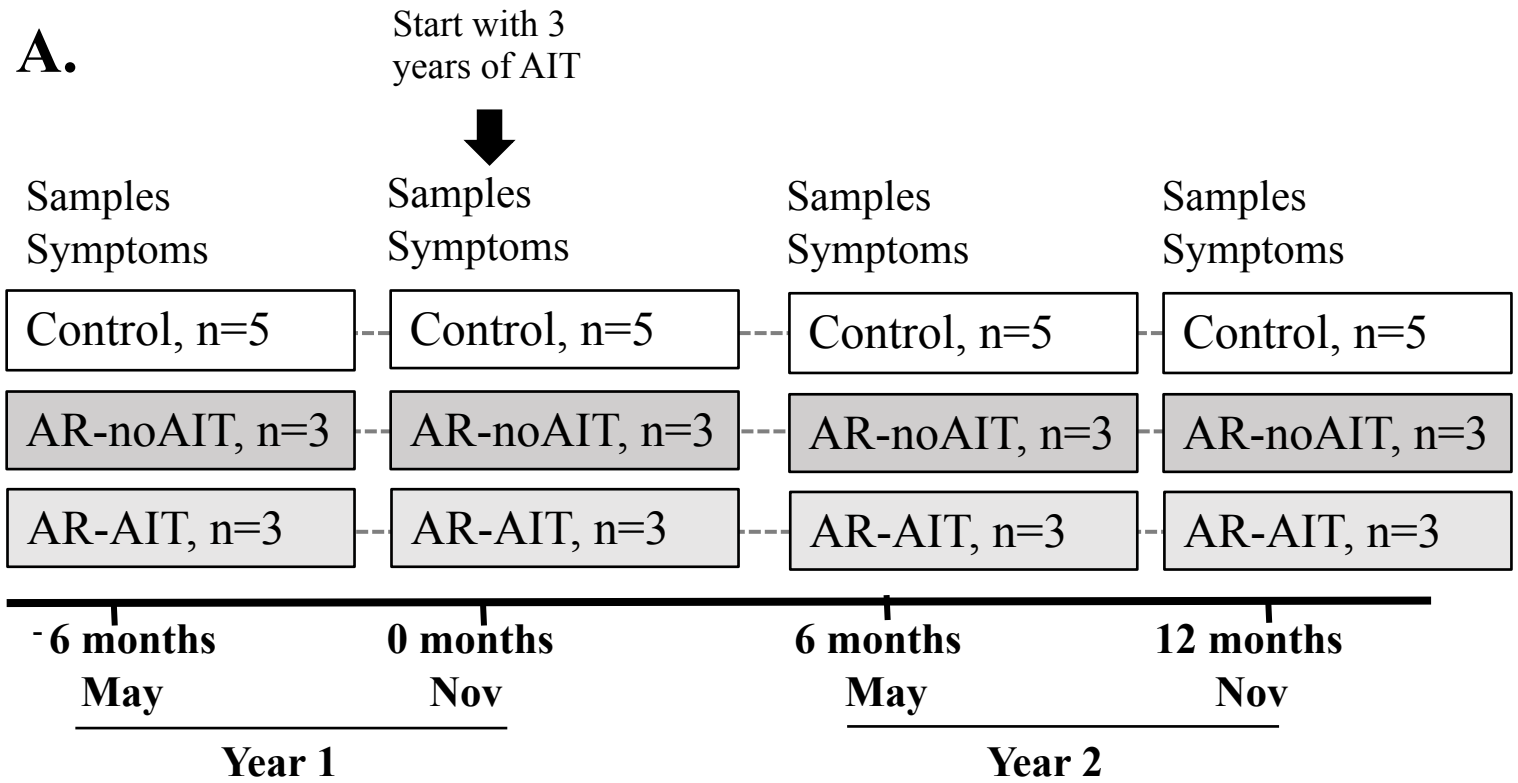
(8) Lal D, Keim P, Delisle J, Barker B, Rank MA, Chia N, et al. Mapping and comparing bacterial microbiota in the sinonasal cavity of healthy, allergic rhinitis, and chronic rhinosinusitis subjects. Int Forum Allergy Rhinol 2017 Jun;7(6):561-569.

Figure legends

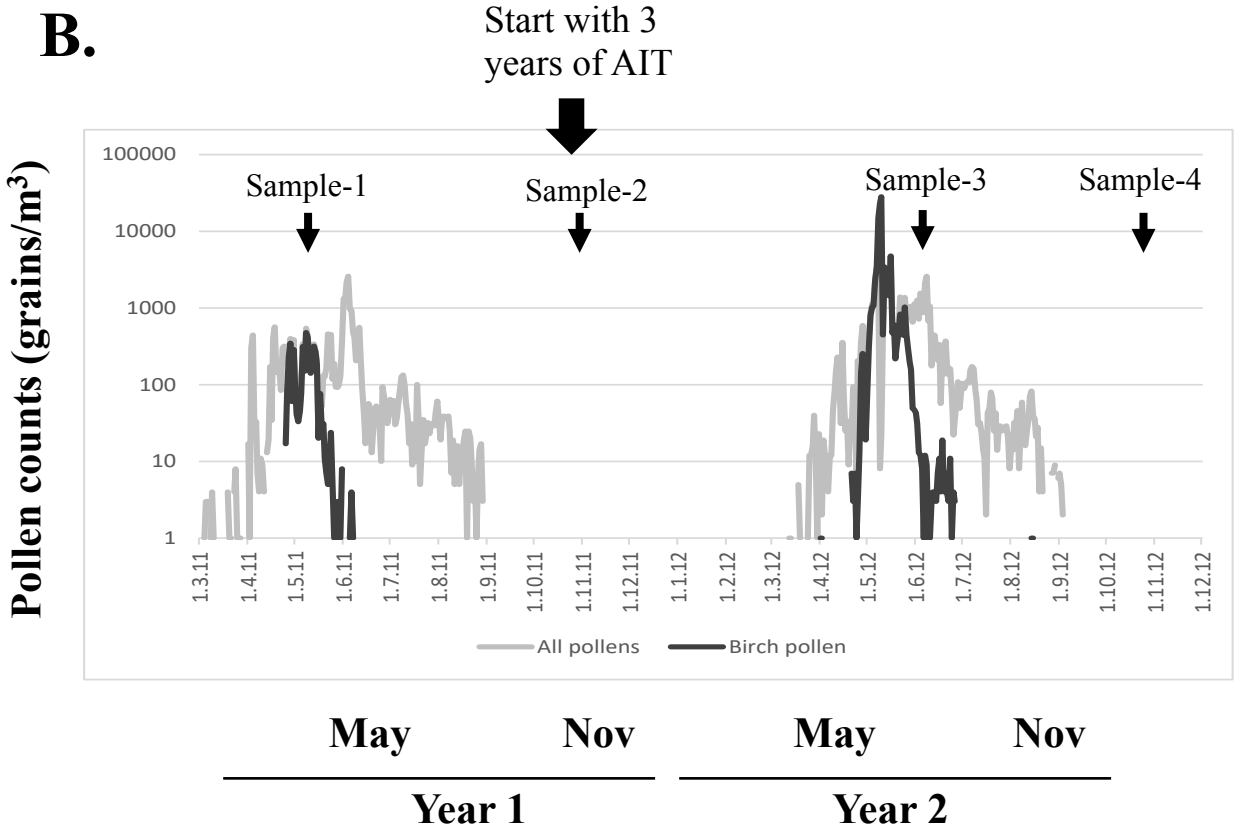
FIG 1. Study overview. A) Study flow chart showing number of subjects, sampling points, and the start of the AIT. Samples were collected at four consecutive sampling points from five healthy control subjects and six AR subjects. Three patients with AR started AIT. All subjects were without medication for at least four weeks before sampling. In springs, symptomatic AR patients were without antihistamines for at least three days prior to sampling. B) Counts of birch and total pollen during the course of the study in grey and black, respectively. Counts of other pollens than birch were under detection level during sampling during the spring samplings. There were no counts of pollen in the air during the winter samplings. C) Total visual analogue scale (VAS) symptom score at the day of sampling. Control and AR-AIT groups (p-value < 0.005) as well as AR-AIT and AR-noAIT groups (p-value < 0.023) differed in interaction by two-way repeated measures analysis of variance (ANOVA). Statistically significant interaction were not observed between control and AR-noAIT groups at the alpha-level of 0.05.

FIG 2. Overview of RNA-sequencing data. A and D) Protein-coding genes statistically differentially expressed (Q-value ≤ 0.1 and absolute $\log_2$ fold-change ≥ 1.5) between spring (A) and winter (D). The heatmap was drawn using $\log_2$ (+1 offset) counts per million expression values, mean centered and scaled by gene averages. Red indicates up-regulation of the gene and blue down-regulation of the gene relative to the average. B) and E) Venn diagrams shows the total number of differentially expressed genes between spring (B) and winter (E) sampling points at each group. C) and F) KEGG pathways enriched among differentially expressed genes per group. Shades of blue and red indicate significance of the enrichment and size of the dot represent gene count. Listed at the bottom in brackets is the total number of differentially expressed genes in each group with an association to some KEGG pathway. G) Number of differentially expressed genes in selected-pairwise comparisons. Comparisons not shown are: AR-noAIT Nov_2 vs. AR-noAIT May_1 (52), AR-AIT Nov_2 vs. AR-AIT May_1 (1), AR-AIT May_1 vs. AR-noAIT May_1 (7), AR-AIT Nov_1 vs. AR-noAIT Nov_1 (0), AR-AIT May_2 vs. AR-noAIT May_2 (6), AR-AIT Nov_2 vs. AR-noAIT Nov_2 (1), Control May_2 vs. Control May_1 (27), Control Nov_2 vs. Control May_1 (7), and Control Nov_2 vs. Control Nov_1 (17). H) Relative abundances of microbial (archaeal, bacterial, and viral) genera and average microbial load per group. Only genera accounting for >5% of the total microbial load in any group are shown. The line denotes the average number of microbe-classified reads within the group.

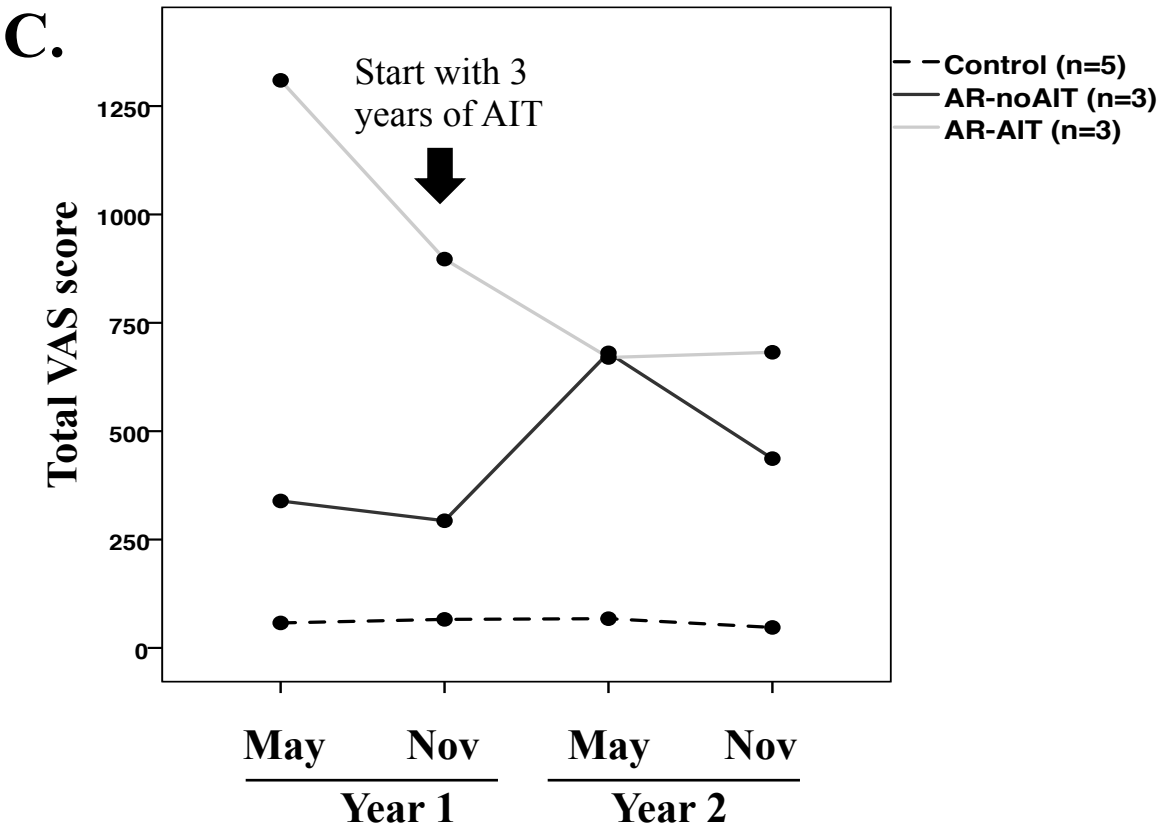
A.


Start with 3
years of AIT
Samples
Symptoms
Samples
Symptoms
Samples
Symptoms
Samples
Symptoms
Control, n=5
Control, n=5
Control, n=5
Control, n=5
AR-noAIT, n=3
AR-noAIT, n=3
AR-noAIT, n=3
AR-noAIT, n=3
AR-AIT, n=3
AR-AIT, n=3
AR-AIT, n=3
AR-AIT, n=3
-6 months
May
0 months
Nov
6 months
May
12 months
Nov
Year 1
Year 2

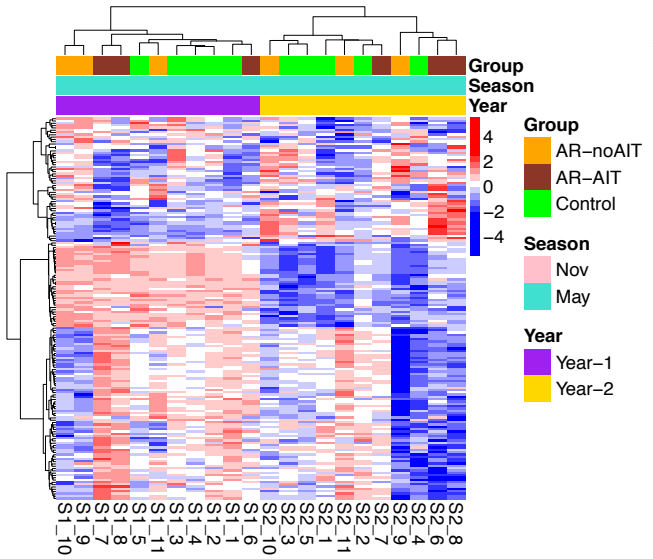
B.


Start with 3
years of AIT
Sample-1
Sample-2
Sample-3
Sample-4
Pollen counts (grains/m3)
100000
10000
1000
100
10
1
1.3.11
1.4.11
1.5.11
1.6.11
1.7.11
1.8.11
1.9.11
1.10.11
1.11.11
1.12.11
1.1.12
1.2.12
1.3.12
1.4.12
1.5.12
1.6.12
1.7.12
1.8.12
1.9.12
1.10.12
1.11.12
1.12.12
All pollens
Birch pollen
May
Nov
May
Nov
Year 1
Year 2

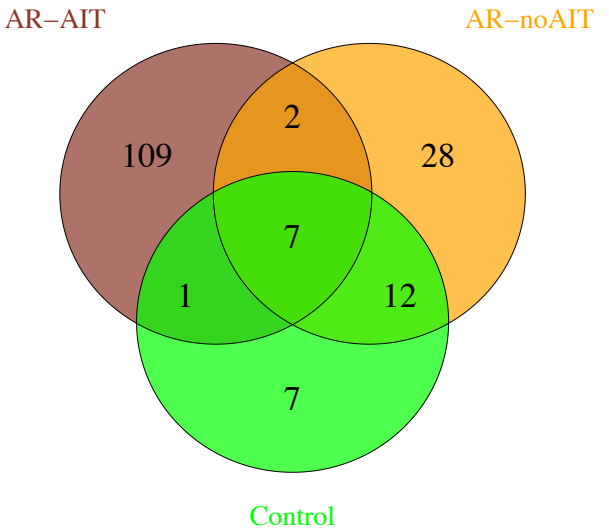
C.


Start with 3
years of AIT
Control (n=5)
AR-noAIT (n=3)
AR-AIT (n=3)
Total VAS score
1250
1000
750
500
250
0
May
Nov
May
Nov
Year 1
Year 2

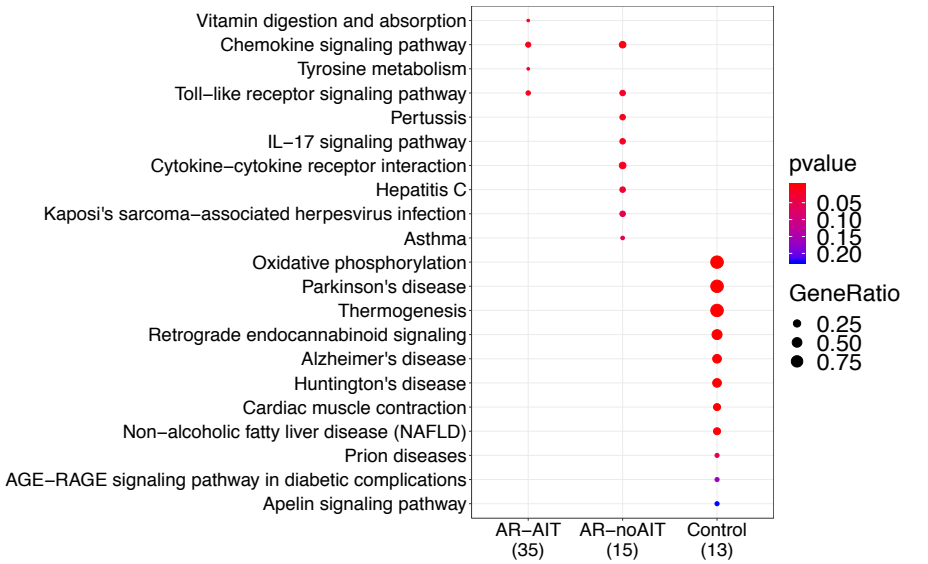
A)


Group
Season
Year
Group
AR−noAIT
AR−AIT
Control
Season
Nov
May
Year
Year−1
Year−2
4
2
0
−2
−4

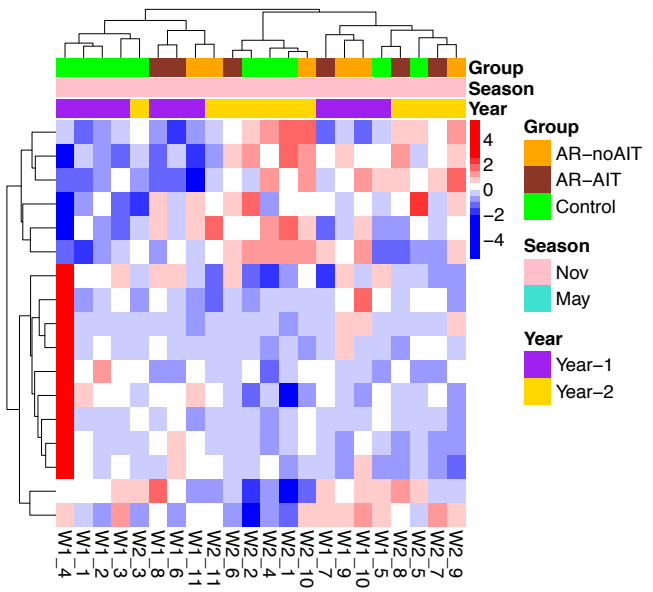
B)


AR−AIT
AR−noAIT
Control
109
2
28
7
1
12
7

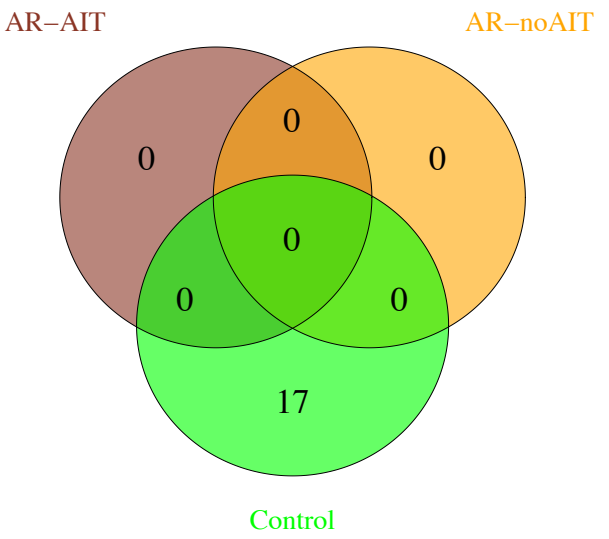
C)


Vitamin digestion and absorption
Chemokine signaling pathway
Tyrosine metabolism
Toll−like receptor signaling pathway
Pertussis
IL−17 signaling pathway
Cytokine−cytokine receptor interaction
Hepatitis C
Kaposi's sarcoma−associated herpesvirus infection
Asthma
Oxidative phosphorylation
Parkinson's disease
Thermogenesis
Retrograde endocannabinoid signaling
Alzheimer's disease
Huntington's disease
Cardiac muscle contraction
Non−alcoholic fatty liver disease (NAFLD)
Prion diseases
AGE−RAGE signaling pathway in diabetic complications
Apelin signaling pathway
AR−AIT (35)
AR−noAIT (15)
Control (13)
pvalue
0.05
0.10
0.15
0.20
GeneRatio
0.25
0.50
0.75

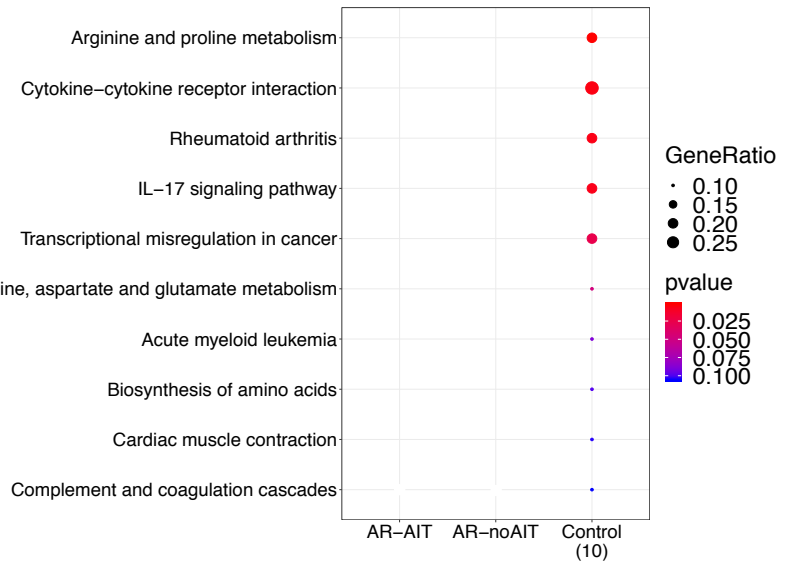
D)


Group
Season
Year
Group
AR−noAIT
AR−AIT
Control
Season
Nov
May
Year
Year−1
Year−2
4
2
0
−2
−4

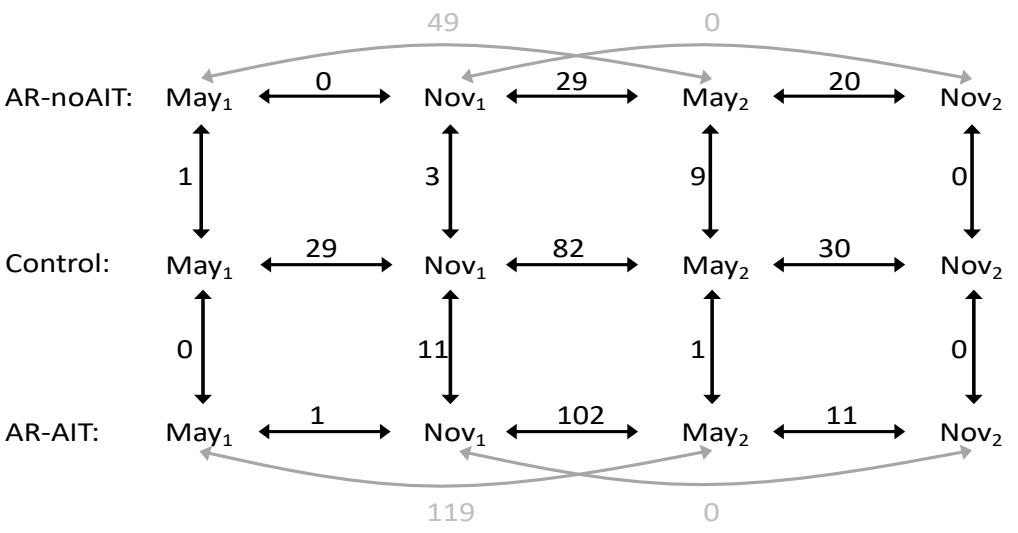
E)


AR−AIT
AR−noAIT
Control
0
0
0
0
0
0
17

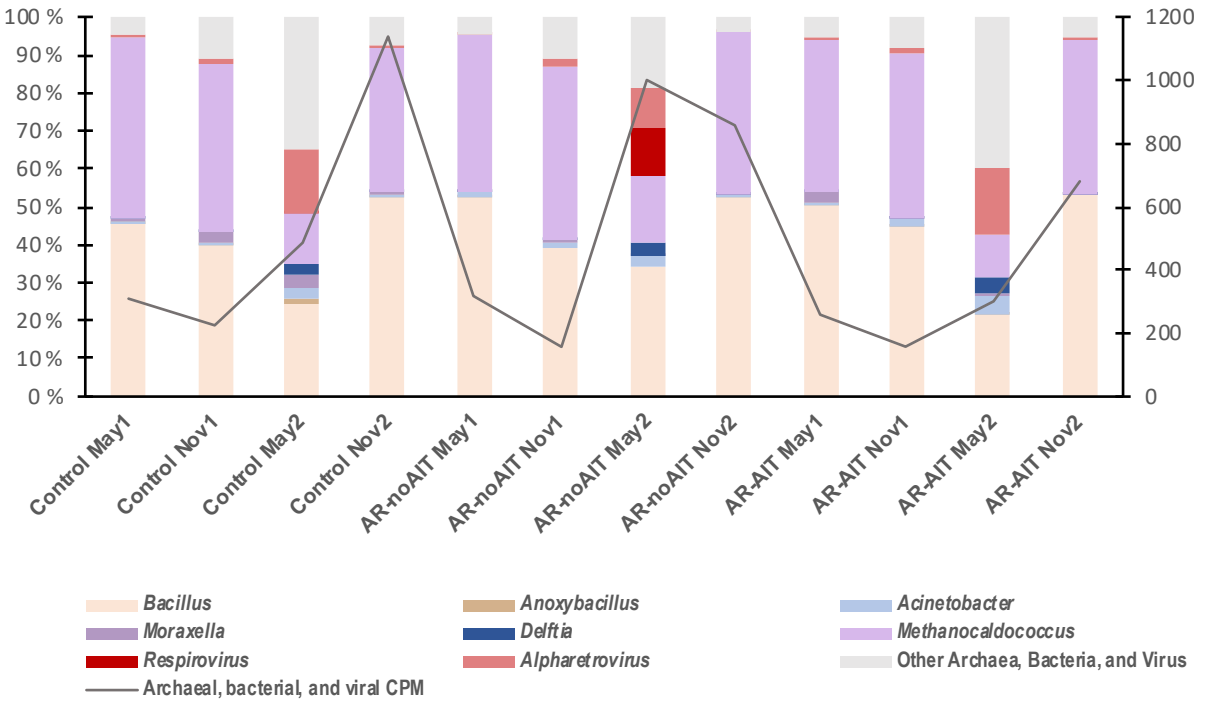
F)


Arginine and proline metabolism
Cytokine−cytokine receptor interaction
Rheumatoid arthritis
IL−17 signaling pathway
Transcriptional misregulation in cancer
Alanine, aspartate and glutamate metabolism
Acute myeloid leukemia
Biosynthesis of amino acids
Cardiac muscle contraction
Complement and coagulation cascades
AR−AIT
AR−noAIT
Control (10)
GeneRatio
0.10
0.15
0.20
0.25
pvalue
0.025
0.050
0.075
0.100

G)


AR-noAIT:
Control:
AR-AIT:
May1
Nov1
May2
Nov2
49
0
0
29
20
1
3
9
0
29
82
30
0
11
1
0
1
102
11
119
0

H)


100 %
90 %
80 %
70 %
60 %
50 %
40 %
30 %
20 %
10 %
0 %
1200
1000
800
600
400
200
0
Control May1
Control Nov1
Control May2
Control Nov2
AR-noAIT May1
AR-noAIT Nov1
AR-noAIT May2
AR-noAIT Nov2
AR-AIT May1
AR-AIT Nov1
AR-AIT May2
AR-AIT Nov2
Bacillus
Anoxybacillus
Acinetobacter
Moraxella
Delftia
Methanocaldococcus
Respirovirus
Alpharetrovirus
Other Archaea, Bacteria, and Virus
Archaeal, bacterial, and viral CPM

Online Repository

Birch pollen allergen immunotherapy reprograms nasal epithelial transcriptome and recovers microbial diversity

Tanzeela Hanif, MSci[1*], Kishor Dhaygude, MSci[1*], Matti Kankainen, PhD[2,3], Jutta Renkonen, DDS[1], Pirkko Mattila, PhD[2], Teija Ojala, PhD[4], Sakari Joenvaara, MSci[1], Mika Mäkelä, Prof[5], Anna Pelkonen, MD PhD[5], Paula Kauppi, MD PhD[5], Tari Haahtela, Prof[5], Risto Renkonen, Prof[1,6#], Sanna Toppila-Salmi MD PhD[1, 5#]

[1] Haartman Institute, University of Helsinki, Helsinki, Finland
[2] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland
[3] Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki,
[4] Department of Pharmacology, University of Helsinki, Helsinki, Finland
[5] Skin and Allergy Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[6] HUSLAB, Helsinki University Hospital, Helsinki, Finland
*Shared first author
#Shared last author

Address correspondence and reprint requests to Sanna Toppila-Salmi, Haartman Institute, University of Helsinki, Haartmaninkatu 3, 00014 University of Helsinki, Helsinki, Finland
Phone number: +358 505431421 Fax number +358 9 471 86476
E-mail address: sanna.salmi@helsinki.fi

Abbreviations

AIT: allergen immunotherapy
AR: Allergic rhinitis
CPM: Count per million
DAVID: Database for Annotation, Visualization, and Integrated Discovery
HC: Hierarchical clustering
KEGG: Kyoto Encyclopedia of Genes and Genomes
PcoA: Principal coordinates analysis
SCIT: subcutaneous immunotherapy
SPT: skin prick test
TMM: Trimmed Mean of M-values
VST: Variance stabilizing transformation
PD15: 15th percentile density
FEV1: Forced expiratory volume
RQLQ: Quality of Life Questionnaire
VAS: Visual analogue scale
ANOVA: analysis of variance
SQ: standardized quality
TLR: Toll like receptor
IQR: Interquartile range
DEG: differentially expressed genes

Materials and methods

Subjects

Study subjects were recruited from 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 number19/13/003/00/11). Written informed consent was received from all subjects and their parents if the age of the participant was under 18-years. The study has been registered in ClinicalTrials.com (nro. NCT01985542). Baseline data of the study subjects is shown in Table E1. The total number of participants entering the study was 23 (Fig E1). AR-AIT group received SCIT in Nov 2011 after the second sampling visit (Fig 1, Fig E1), meaning that two samplings of the AR-AIT group were performed before and two 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 (1). Epithelial cells were collected at four time points, washed once with ice cold nuclease free PBS, and resuspended immediately into RNAlater 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 RNAnano chip (Agilent) was used to evaluate the integrity of RNA and Qubit RNA –kit (Life Technologies) to quantitate RNA in epithelial cell samples. If acceptable in quality (RIN value >7), 1.0 ug of total RNA sample was ribodepleted and prepared to RNA sequencing library by using ScriptSeq v2™ Complete kit (Illumina, Inc., San Diego, CA, USA). RNA sequencing libraries were purified with SPRI beads (Agencourt AMPure XP, Beckman Coulter, Brea, CA, USA). The library QC 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, CA, USA). Planned read amount was 40 million reads per sample.

RNA sequencing data processing

RNA sequencing data were preprocessed as described previously (2). Briefly, Trimmomatics (3) 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 (4) 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 (5), feature counts were converted to expression estimates using Trimmed Mean of M-values (TMM) normalization (6), and lowly expressed genomic features with counts per million (CPM) value ≤1.00 in less than half of controls or birch-pollen patients were removed. Default parameters were used, with 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 (7) 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, while comparisons within subject groups employed also a factor for the subject. The resulting p-values were adjusted Storey's Q-value approach with significance defined as Q-value ≤0.10. A cut-off of absolute $\log_2$ fold-change of ≥ 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 pheatmap R package (8). Hierarchical clusters (HC) 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. Counts per million (CPMs) data were used to generate heatmaps. Venn diagrams were generated using the VennDiagram R package (9). Functional profiles of differentially expressed genes were investigated with clusterProfiler (10) 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

(11). In the process, KEGG gene IDs of the selected pathways were fed along with $\log_2$ 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, *i.e.* values >2 were truncated to 2, and values < -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 (12) and then annotated using Annovar (13) as defined previously (2). Quality control analyses were performed as defined previously (2). Variant calls were further filtered by accepting only those that were present in two or more AR cases, not present in any control case, and predicted to be pathogenic by various pathogen prediction methods part of the Annovar (13) annotation tool. Heatmap was plotted using pheatmap. The functional effects of variants were taken from Annovar (13) outputs. Additionally, we also plotted barplot using CPM expression value of genes in healthy control and AR groups.

Microbial community profiling

Microbial community profiling was performed as previously described (14) 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 (3). Paired-end reads passing the pre-processing were mapped against rRNA sequences from RFAM (15) v12.3 using the Burrows-Wheeler Aligner (BWA) (16) with default settings and reads matching rRNAs were filtered by using samtools (17). Centrifuge (18) 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 (19) database at February 2018. Default parameters were used, with the exception that only one (*i.e.* the lowest common ancestor) assignment was reported for read-pairs with multiple primary assignments. Taxa having <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 DeSeq2 (20) 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 (VST) was used to generate expression estimates for heatmap visualizations. The Storey's Q-value adjustment (21) was used to correct data for multiple hypothesis testing, with significance defined as Q-value ≤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 (22) applied on number of reads assigned directly to the given taxonomic level.

Results

General

AR subjects, and especially AR-AIT cases, had higher total median S-IgE, birch specific S-IgE, and SPT wheel diameter to birch and symptom scores during samplings (Table E1, Fig 1). AR-AIT group reported benefit (Fig 1) and reported no severe side-effects at the end of SCIT three years after start of SCIT (data not shown).

Transcriptome of nasal epithelium

We generated in total of 5164 million raw paired-end transcript reads. Manual inspection of the quality plots generated using FASTQ indicated that sequencing data was 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 were found. To gain insight into cellular processes dysregulated in AR and AIT, differentially expressed genes between sample groups were identified using edgeR (7). In this analysis, we identified altogether 360 genes to be differentially expressed with the Q-value ≤ 0.1 and absolute $\log_2$ fold change ≥ 1.5 between different timepoints within group and between different groups within timepoints.

Alterations in gene expression profiles in the two 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 suggest that greatest transcriptional reprogramming took place in AR-AIT group followed by AR-noAIT group and healthy subjects. Comparison of the two consecutive winters revealed only 17 DEGs among healthy controls and none among AR patients, 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). Out of these, 4 were associated with genes differentially expressed in AR-AIT group, including the chemokine signaling and Toll-like receptor (TLR) signaling pathway (Fig 2, C). Genes differentially expressed in 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 three 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. Firstly, the asthma pathway consisted in total 3 DEGs. The AR-AIT group and healthy controls displayed upregulation of *MHCII* and downregulation of *FcεRI* at the second spring (Fig E2, A). The expression of various other members of the pathway was altered, although unsignificantly, between spring samplings (Fig E2, A). These more borderline findings included *IL-13* that is a T-cell-specific transcription factor and interleukin (23), *IL-4* that is IgE synthesis switch factor (23), and *IL-5* that is an eosinophil growth factor (23). Expression of genes of the asthma pathway between timepoints occurred to opposite directions in AR-AIT and AR-noAIT groups (Fig E2, A). The second pathway of interest was TLR signaling pathway (Fig E2, B) consisting altogether 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, B). The third pathway of relevance to AR was chemokine signaling pathway. This pathway consisted of 49 dysregulated genes (Fig E2, C), including various chemokines and chemokine receptors. Expression of genes between timepoints occurred to opposite directions in AR-AIT and AR-noAIT groups (Fig E2, C).

Transcript variants expressed in AR patients

The GATK best practice for RNA sequencing (12) was employed to identify expressed variants. This approach revealed altogether 3,268,177 (on average 74277 per subject) variants passing GATK filters. Removal of intronic and silent variants and polymorphisms resulted in 8,174 (on average 186 per subject) variants that were further narrowed down to 8 potential candidates expressed in at least two AR subjects 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 to the active nasal microbiota was studied by identifying microbial reads from the RNA sequencing data and performing microbial classification. On average ~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 (1791 CPMs)). Of these microbial reads, on average ~98.13% were classified at 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 ~42.23%, ~35.72%, and ~4.32%, respectively, other genera having average relative proportions of ~1.57% (*Acinetobacter*) or less. The relative proportions of the detected genera varied greatly between the samples, *Bacillus* demonstrating the greatest variation (relative proportions ranging from 0% to ~63.76%). The relative proportions 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 six cases, who were distributed among all groups (Fig E4, A). Examination of the community compositional differences between the samples revealed these six 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 (Fig E5-E7, A-N). Majority of the indices demonstrated an increase in the

diversity for AR-AIT group when the second winter was compared to the first (Fig E6, A-N). Some change of diversity between second and first winters was also seen for the control group, while some indices indicated also an increase for the AR-noAIT group. The increase seen in AR-noAIT group was, however, less than that observed for the AR-AIT group. The second winter diversity indices of the AR-AIT group were also mainly closer to those of the control group than the indices of the AR-noAIT group (Fig E6, A-N). Finally, examination of significant (Q-value ≤ 0.05) changes in species abundancies (Fig E4, C) revealed more changes in all groups between the two 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.

Acknowledgements

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

References

(1) Mattila P, Renkonen J, Toppila-Salmi S, Parviainen V, Joenvaara S, Alff-Tuomala S, et al. Time-series nasal epithelial transcriptomics during natural pollen exposure in healthy subjects and allergic patients. Allergy 2010 Feb;65(2):175-183.

(2) Kumar A, Kankainen M, Parsons A, Kallioniemi O, Mattila P, Heckman CA. The impact of RNA sequence library construction protocols on transcriptomic profiling of leukemia. BMC Genomics 2017 Aug 17;18(1):629-017-4039-1.

(3) Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014 Aug 1;30(15):2114-2120.

(4) Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013 Jan 1;29(1):15-21.

(5) Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res 2013 May 1;41(10):e108.

(6) Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 2010;11(3):R25-2010-11-3-r25. Epub 2010 Mar 2.

(7) Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010 Jan 1;26(1):139-140.

(8) Kolde R. pheatmap: Pretty Heatmaps. R package version 1.0.10. 2018.

(9) Chen H. VennDiagram: Generate High-Resolution Venn and Euler Plots. R package version 1.6.20. 2018.

(10) Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012 May;16(5):284-287.

(11) Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 2013 Jul 15;29(14):1830-1831.

(12) McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010 Sep;20(9):1297-1303.

(13) Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010 Sep;38(16):e164.

(14) Dufva O, Kankainen M, Kelkka T, Sekiguchi N, Awad SA, Eldfors S, et al. Aggressive natural killer-cell leukemia mutational landscape and drug profiling highlight JAK-STAT signaling as therapeutic target. Nat Commun 2018 Apr 19;9(1):1567-018-03987-2.

(15) Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, et al. Rfam 12.0: updates to the RNA families database. Nucleic Acids Res 2015 Jan;43(Database issue):D130-7.

(16) Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009 Jul 15;25(14):1754-1760.

(17) Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009 Aug 15;25(16):2078-2079.

(18) Kim D, Song L, Breitwieser FP, Salzberg SL. Centrifuge: rapid and sensitive classification of metagenomic sequences. Genome Res 2016 Dec;26(12):1721-1729.

(19) O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 2016 Jan 4;44(D1):D733-45.

(20) Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550-014-0550-8.

(21) Storey J.D. The positive false discovery rate: A Bayesian interpretation and the q-value. Annals of Statistics 2003;31(6):2013-2013-2035.

(22) McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013 Apr 22;8(4):e61217.

(23) Pawankar R, Mori S, Ozu C, Kimura S. Overview on the pathomechanisms of allergic rhinitis. Asia Pac Allergy 2011 Oct;1(3):157-167.

(25) Juniper EF, Thompson AK, Ferrie PJ, Roberts JN. Development and validation of the mini Rhinoconjunctivitis Quality of Life Questionnaire. Clin Exp Allergy 2000 Jan;30(1):132-140.

(26) Arvidsson MB, Lowhagen O, Rak S. Effect of 2-year placebo-controlled immunotherapy on airway symptoms and medication in patients with birch pollen allergy. J Allergy Clin Immunol 2002 May;109(5):777-783.

Table E1. Baseline characteristics of the subjects.

	All subjects			AR patients		
	Controls	AR patients	P-value	AR-noAIT	AR-AIT	P-value
Baseline characteristics						
No. of subjects	5	6		3	3	
Age (y), median (IQR)	44 (39-48)	39 (34-46)	.43	34 (32-42)	41 (39-43)	.70
Men/women (n)	1/4	3/3	.54	2/1	1/2	1.00
[1]Spirometry values (% predicted), median (IQR)						
FEV1 baseline	107 (105-109)	95 (94-96)	.23	96 (95-103)	97 (94-95)	.30
FEV1 bronchodilation	110 (105-110)	97 (95-99)	.39	97 (96-104)	97 (95-99)	1.00
FEV1/FVC ratio baseline	104 (101-104)	98 (98-103)	.33	103 (101-107)	95 (92-98)	.30
FEV1/FVC bronchodilation	105 (101-106)	102 (98-104)	.55	104 (101-108)	99.5 (97-102)	.40
PEF baseline	109 (102-125)	109 (101-110)	.57	110 (106-116)	98 (87-109)	.20
PEF bronchodilation	104 (92-125)	112 (92-114)	.74	112 (102-113)	100.5 (86-115)	1.00
Symptoms during sampling in spring$_1$, median (IQR)						
Total RQLQ score	0 (0-6)	48.5 (24-74)	.019	40 (23-49)	74 (49-87)	.40
Total VAS score	15 (10-36)	744 (358-1352)	.009	217 (358-471)	1352 (1128-1512)	.10
Serum values in spring$_1$, median (IQR)						
Total IgE (kU/L)	9 (7-14)	61.5 (39-200)	.015	39 (26-275)	72 (62-136)	.70
Birch-specific IgE (kU/L)	0 (0-0)	16.4 (3.3-24.2)	.004	12 (6-16)	24 (14-28)	.40
Timothy-specific IgE (kU/L)	0 (0-0)	0 (0-0)	.85	0 (0-45)	0 (0-0)	1.00
SPT wheal diameter (mm), median (IQR)						
Negative control	0 (0-0)	0 (0-0)	1.00	0 (0-0)	0 (0-0)	1.00
Histamine (positive control)	5 (5-5)	5 (5-5)	1.00	5 (5-5)	5 (5-5)	1.00
Birch pollen	0 (0-0)	5 (4-6)	.004	4 (3.5-6)	5 (5-5.5)	.60
Timothy grass pollen	0 (0-0)	0 (0-0)	1.00	0 (0-2.5)	0 (0-0)	1.00
Festuca pratensis pollen	0 (0-0)	0 (0-0)	1.00	0 (0-2.5)	0 (0-0)	1.00
Mugwort pollen	0 (0-0)	0 (0-0)	1.00	0 (0-2.5)	0 (0-0)	1.00
Cladosporium herbarum	0 (0-0)	0 (0-0)	1.00	0 (0-0)	0 (0-0)	1.00
Cat dander	0 (0-0)	0 (0-5)	.46	0 (0-2.5)	0 (0-3)	1.00
Dog dander	0 (0-0)	0 (0-4)	.46	0 (0-2)	0 (0-2.5)	1.00
Horse dander	0 (0-0)	0 (0-0)	1.00	0 (0-0)	0 (0-0)	1.00
Dermatophagoides pteronyssinus	0 (0-0)	0 (0-0)	1.00	0 (0-0)	0 (0-0)	1.00

Diagnosis of allergic rhinitis (AR) was based on a typical history, skin prick test (SPT; ALK-Abello, 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 under 12 years, use of tobacco products, nonallergic rhinitis, allergic rhinitis 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. [1)] One subject starting pollen allergen immunotherapy (AIT) was tested for bronchial hyperresponsiveness by

histamine challenge test. He had normal 15th percentile density (PD15) of Forced expiratory volume in 1 second (FEV1) result instead of bronchodilation test. Questionnaire regarding baseline characteristics and symptoms was collected at each sampling visit. We used the 28-item Rhinoconjunctivitis Quality of Life Questionnaire (RQLQ) (25). Subjects filled visual analogue scale (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 4,100. In the analysis, VAS scores $\leq$ 3 mm were regarded as 0. P-values were computed by Kruskal-Wallis and Mann Whitney U-tests (continuous variables) or by Fisher´s exact test (dichotomous variables).

Figures legends

FIG E1. Flowchart of the study. The total number of participants entering the study was 23. Initially, seven allergic rhinitis (AR) patients were assigned with pollen allergen immunotherapy (AIT; AR-AIT group) and nine others to the conventional therapy (AR-noAIT group). The study also included seven control subjects (control group) without allergy. One participant in the AR-AIT group discontinued the study before starting AIT due to a diagnosis of a cardiac disease requiring surgery. The other participants in the AR-AIT group fulfilled the whole AIT scheme of three years with the total dose over 2,000.000 SQ (standardized quality). AIT was performed according to standard 3-year protocol or normal 3-year scheme (26). Subcutaneous injections of birch pollen extract (Betula verrucosa, ALK Abello, Horsholm, Denmark) 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, five out of the 16 subjects, who completed the study, had poor RNA quality in at least one nasal epithelial sample, leading to their exclusion. Thus, a total of 11 subjects (five healthy controls, three AR-AIT, and three AR-noAIT) completed the study. These participants were adults and of white European ancestry except one Chinese male at AR-noAIT group.

FIG E2. Allergy related pathways identified in the pathway analysis. Gene expression profiles of A) asthma, B) toll like receptor (TLR) signaling, and C) chemokine signaling pathways among allergic rhinitis (AR) patients with pollen allergen immunotherapy (AIT), AR patients without AIT, and control subjects. Boxes in the figure represent genes. The gradient colors indicate the $\log_2$ 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, *i.e.* values >2 were truncated to 2, and values < -2 truncated to -2. Asterisk indicates statistically significant difference with a Q-value ≤ 0.10. Green gradient colors indicate up-regulation of the gene at the second spring and red colors indicate up-regulation at the first spring.

FIG E3. Landscape of expressed variants in allergic rhinitis. Short non-synonymous transcript variants identified in at least two AR (allergic rhinitis) 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. Bar-plot at the right indicate the average expression level of the gene in AR and control subjects expressed as counts per million (CPM).

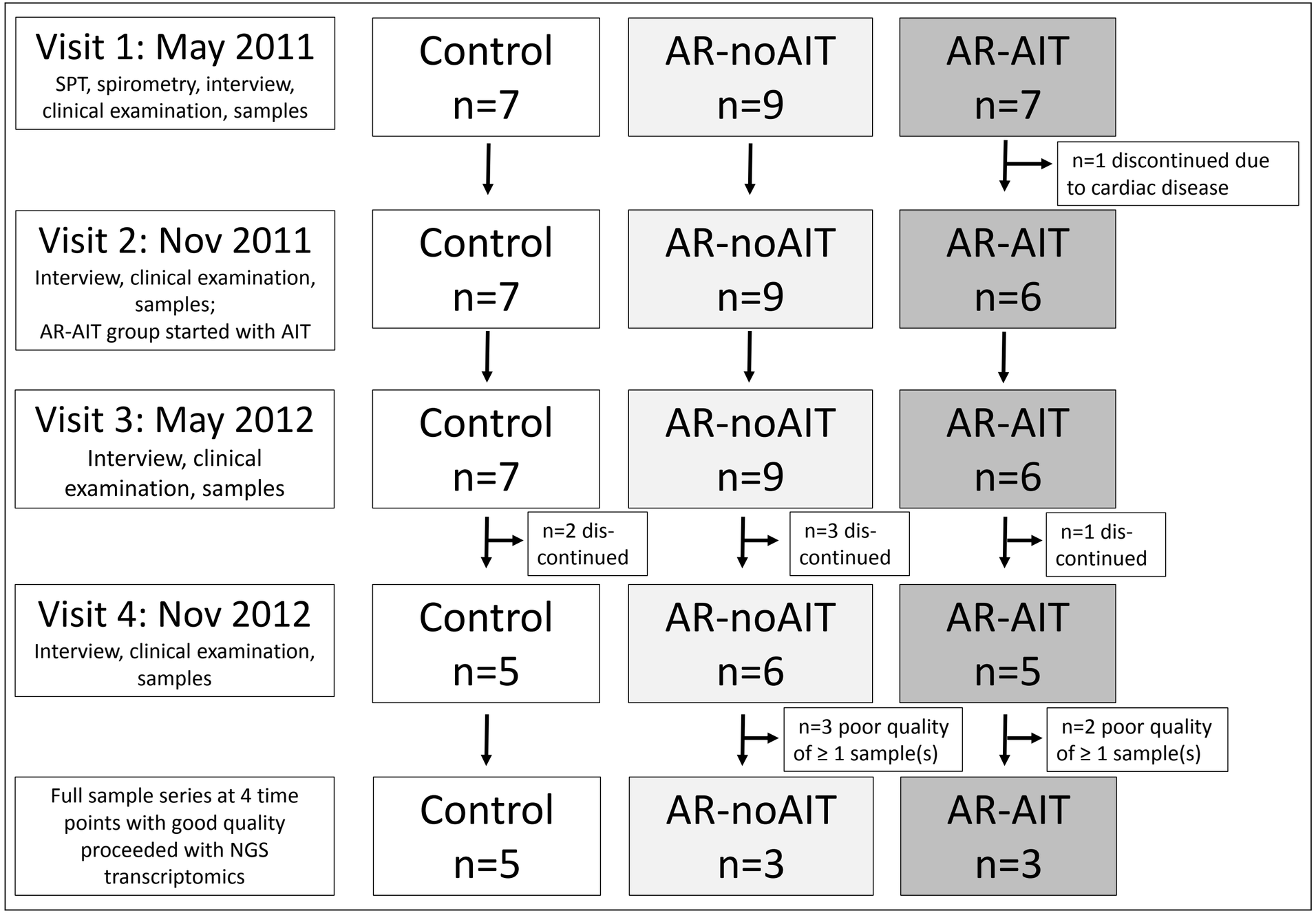
FIG E4. Microbial variation. A) Relative abundances of microbial (archaeal, bacterial, and viral) genera and total microbial load per sample. Only genera accounting for >5% of the total microbial load are shown. The line denotes the total number of microbial reads per sample expressed as counts per million (CPM). B) Principal coordinates analysis plot of microbial community structure based on Bray-Curtis distance. C) A heat map 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 (VST) 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 a variety of alpha diversity metrics. Shown in the figure are alpha diversities computed using Shannon (A, B), ACE (C, D), Chao1 (E, F), observed (G, H), Fisher (I, J), InvSimpson (K, L), and Simpson (M, N) metrics with (B, D, F, H, J, L, N) and without (A, C, E, G, I, K, 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.

FIG E6. Alpha diversity of epithelial microbiotas between groups during winter. Alpha diversity indices of different groups during two consecutive winters using a variety of alpha diversity metrics. Shown in the figure are alpha diversities computed using Shannon (A, B), ACE (C, D), Chao1 (E, F), observed (G, H), Fisher (I, J), InvSimpson (K, L), and Simpson (M, N) metrics with (B, D, F, H, J, L,

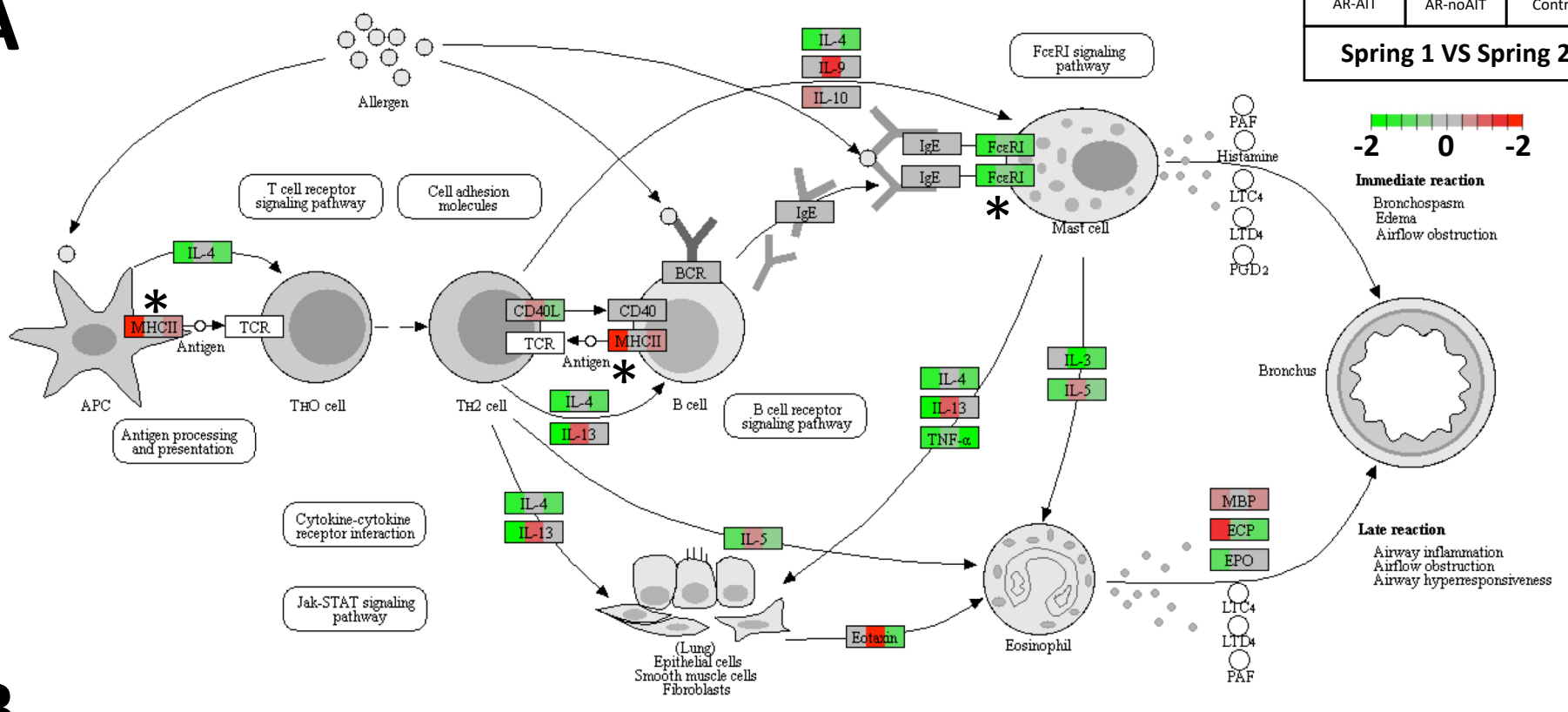
N) and without (A, C, E, G, I, K, M) rarefaction of samples to the minimum sampling depth. In the figure, Nov1 stands for the first winter and Nov2 for the second winter sampling point.

FIG E7. Alpha diversity of epithelial microbiotas between groups during spring. Alpha diversity indices of different groups during two consecutive springs sampling using a variety of alpha diversity metrics. Shown in the figure are alpha diversities computed using Shannon (A, B), ACE (C, D), Chao1 (E, F), observed (G, H), Fisher (I, J), InvSimpson (K, L), and Simpson (M, N) metrics with (B, D, F, H, J, L, N) and without (A, C, E, G, I, K, 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.

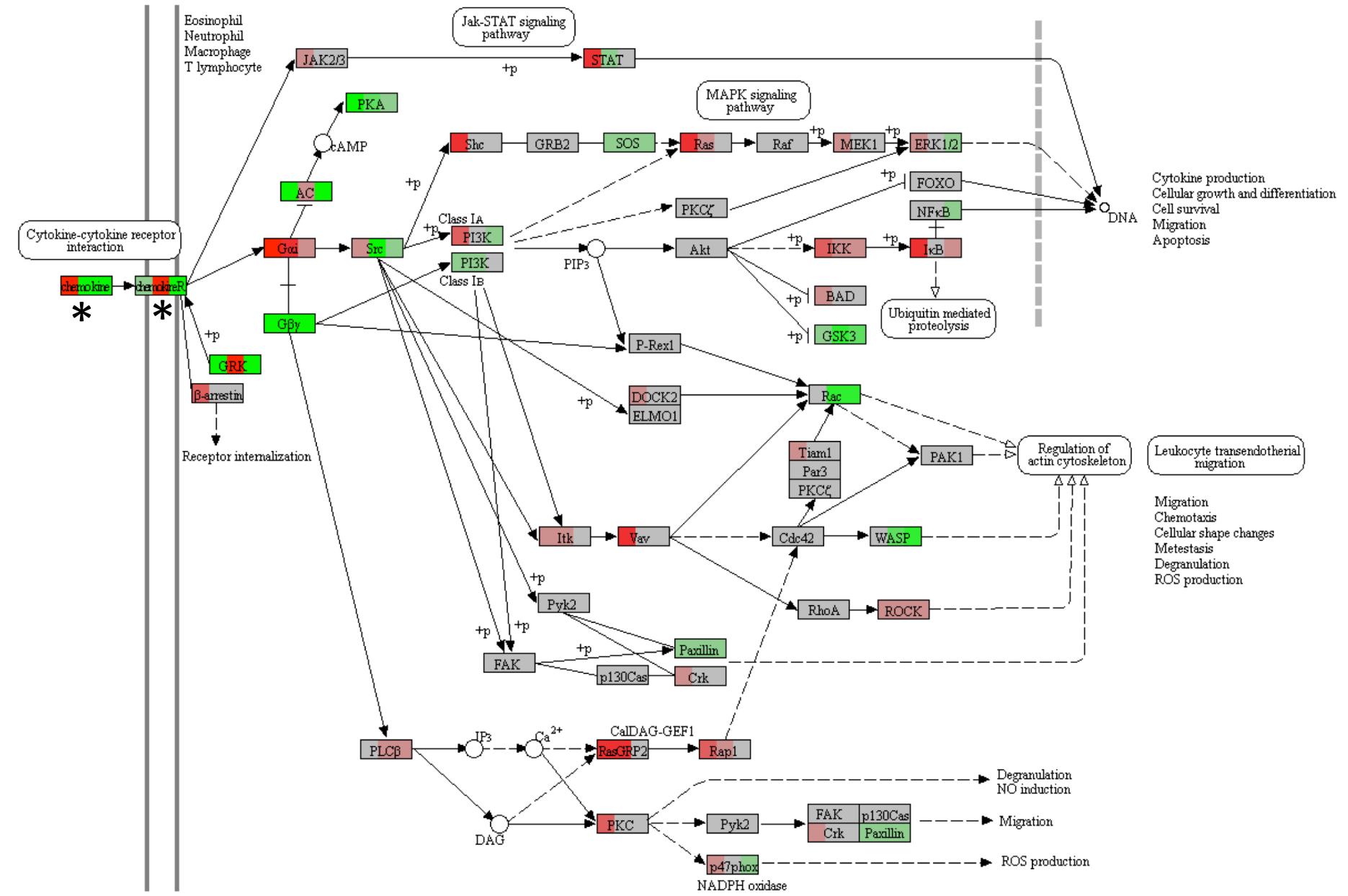

Visit 1: May 2011
SPT, spirometry, interview, clinical examination, samples
Control n=7
AR-noAIT n=9
AR-AIT n=7
n=1 discontinued due to cardiac disease
Visit 2: Nov 2011
Interview, clinical examination, samples;
AR-AIT group started with AIT
Control n=7
AR-noAIT n=9
AR-AIT n=6
Visit 3: May 2012
Interview, clinical examination, samples
Control n=7
AR-noAIT n=9
AR-AIT n=6
n=2 dis-continued
n=3 dis-continued
n=1 dis-continued
Visit 4: Nov 2012
Interview, clinical examination, samples
Control n=5
AR-noAIT n=6
AR-AIT n=5
n=3 poor quality of ≥ 1 sample(s)
n=2 poor quality of ≥ 1 sample(s)
Full sample series at 4 time points with good quality proceeded with NGS transcriptomics
Control n=5
AR-noAIT n=3
AR-AIT n=3

AR-AIT	AR-noAIT	Control
Spring 1 VS Spring 2		

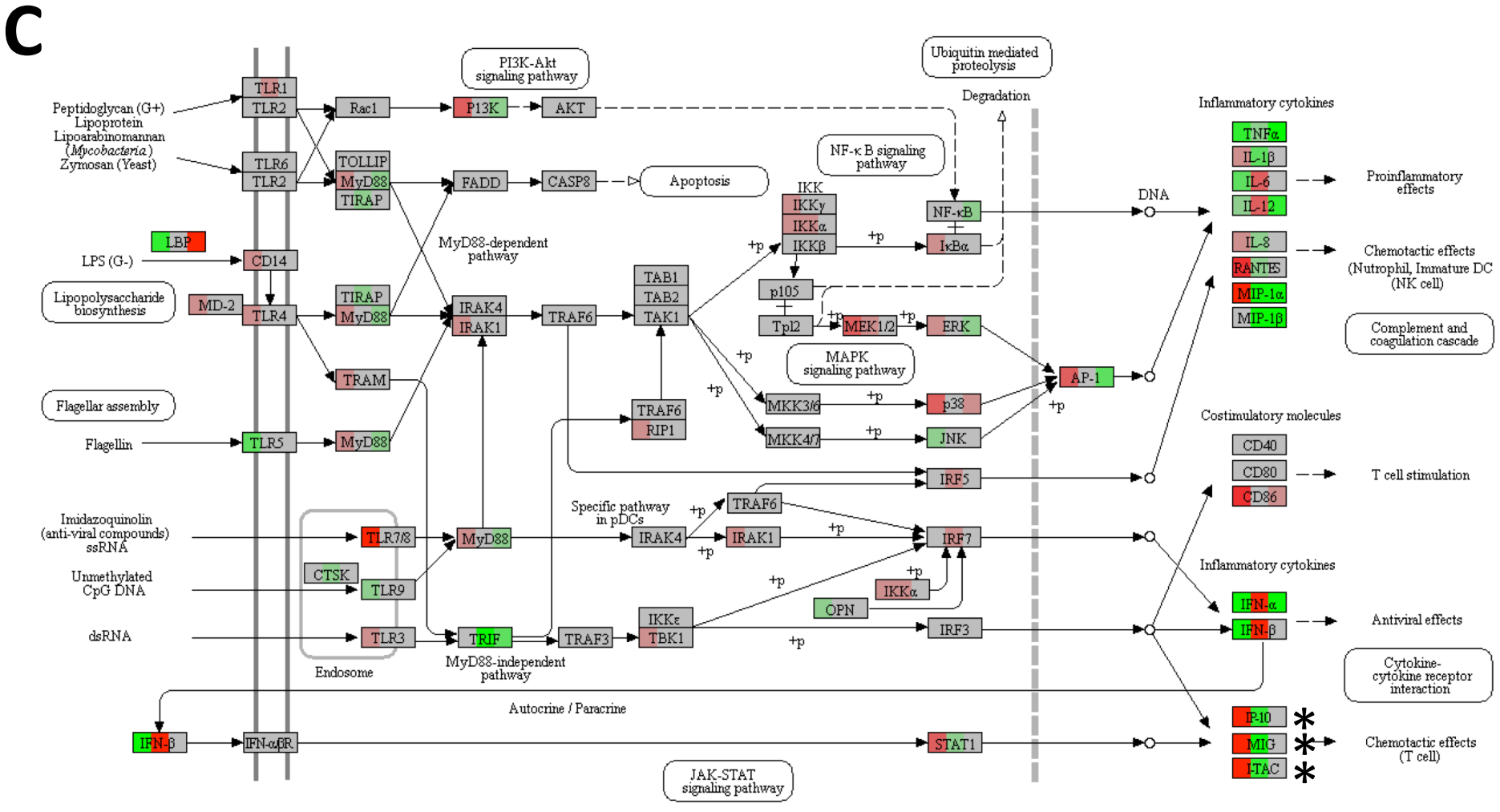
A

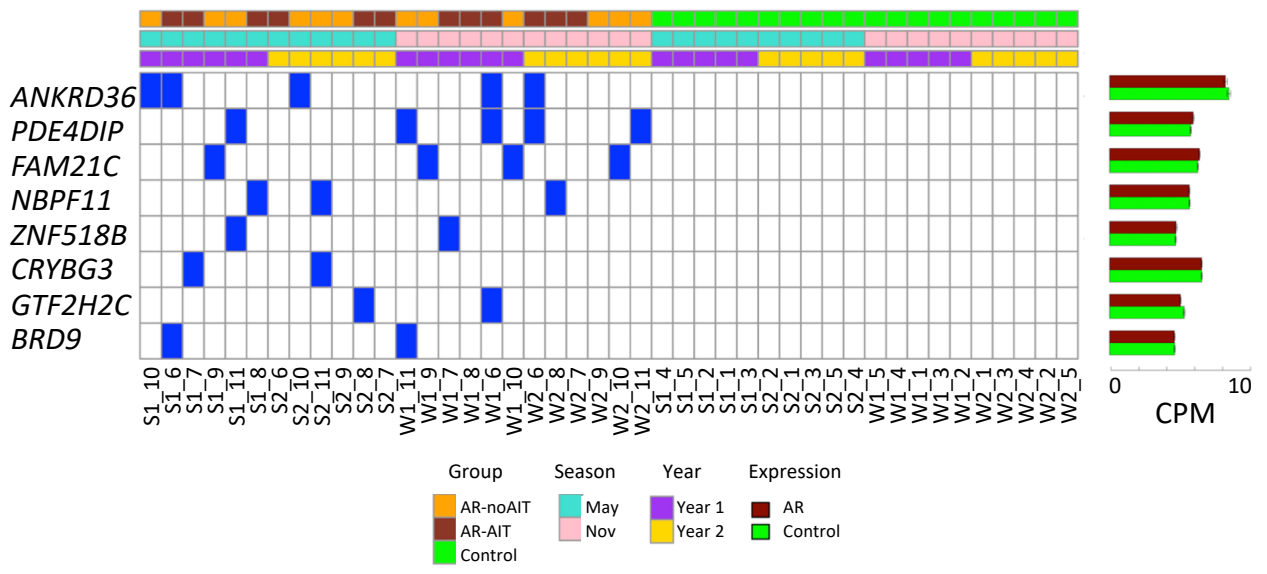


B

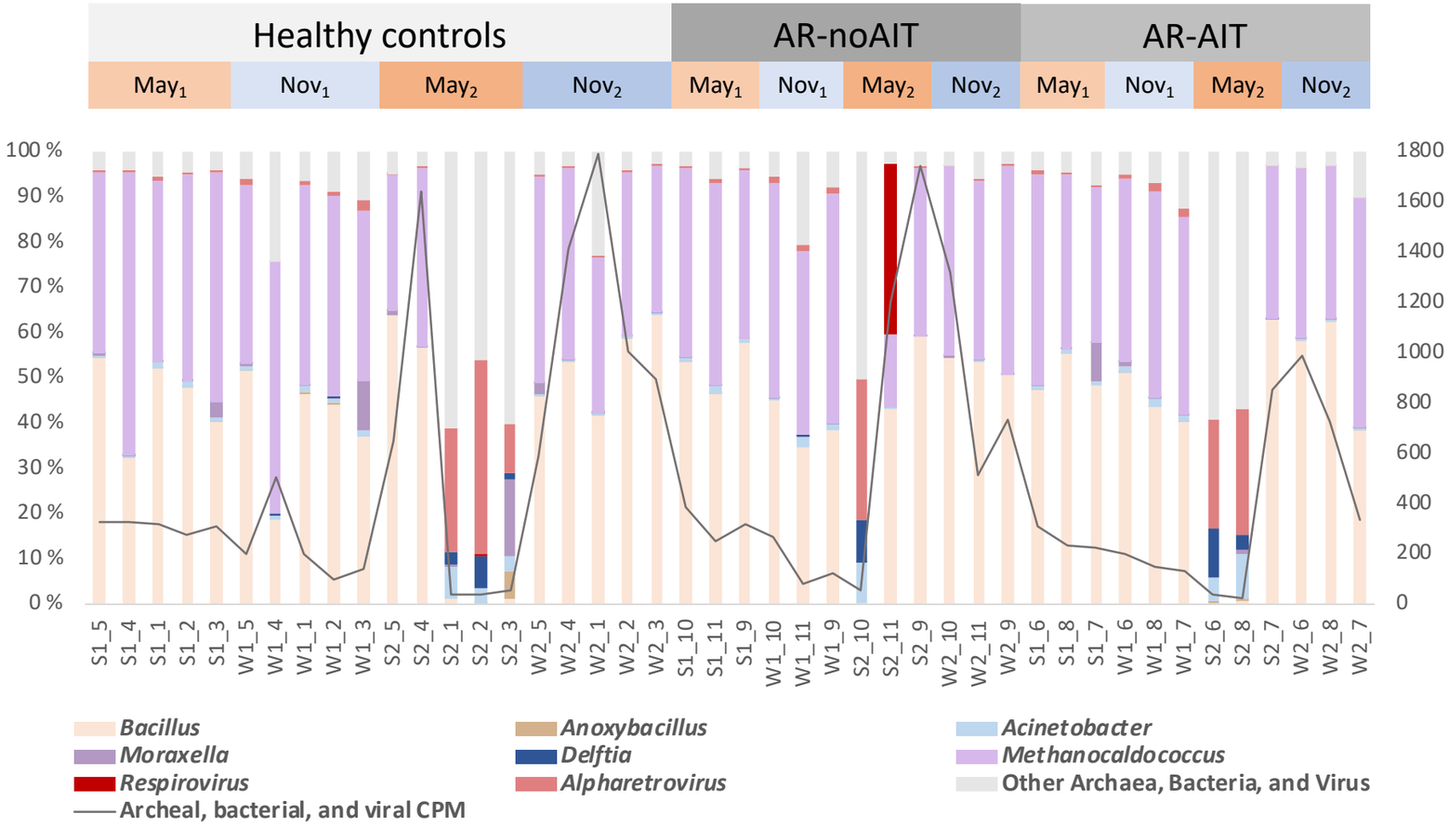


C

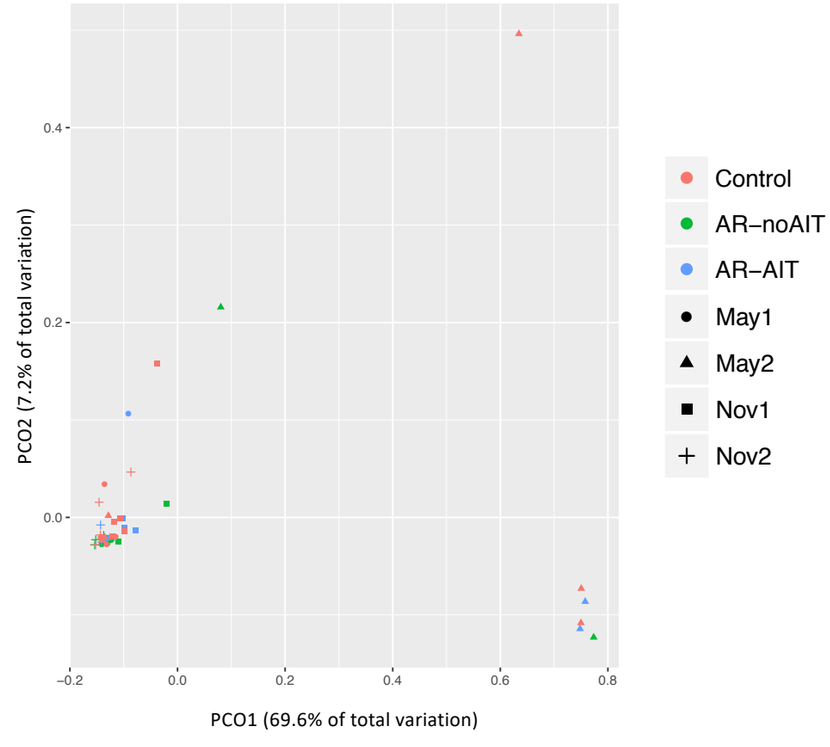



ANKRD36
PDE4DIP
FAM21C
NBPF11
ZNF518B
CRYBG3
GTF2H2C
BRD9
S1_10
S1_6
S1_7
S1_9
S1_11
S1_8
S2_6
S2_10
S2_11
S2_9
S2_8
S2_7
W1_11
W1_9
W1_7
W1_8
W1_6
W1_10
W2_6
W2_8
W2_7
W2_9
W2_10
W2_11
S1_4
S1_5
S1_2
S1_1
S1_3
S2_2
S2_1
S2_3
S2_5
S2_4
W1_5
W1_4
W1_1
W1_3
W1_2
W2_1
W2_3
W2_4
W2_2
W2_5
0
10
CPM
Group
AR-noAIT
AR-AIT
Control
Season
May
Nov
Year
Year 1
Year 2
Expression
AR
Control

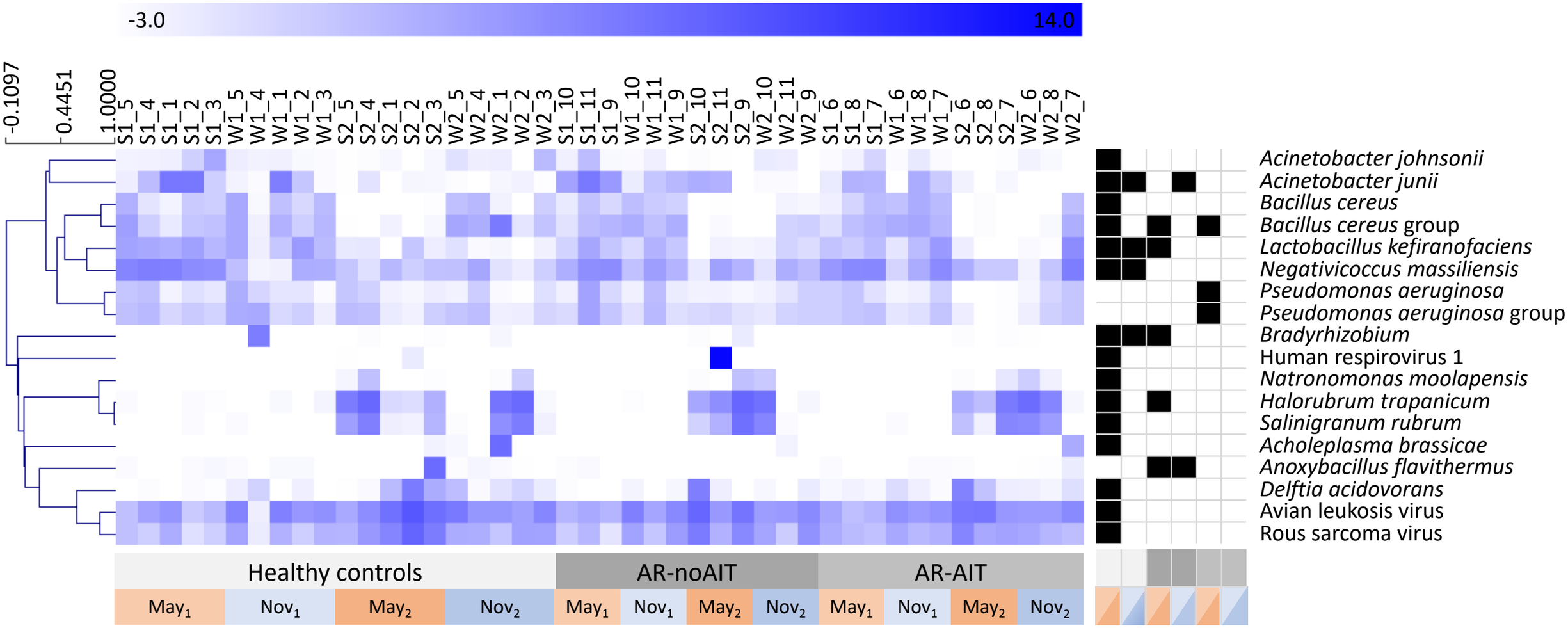
A)

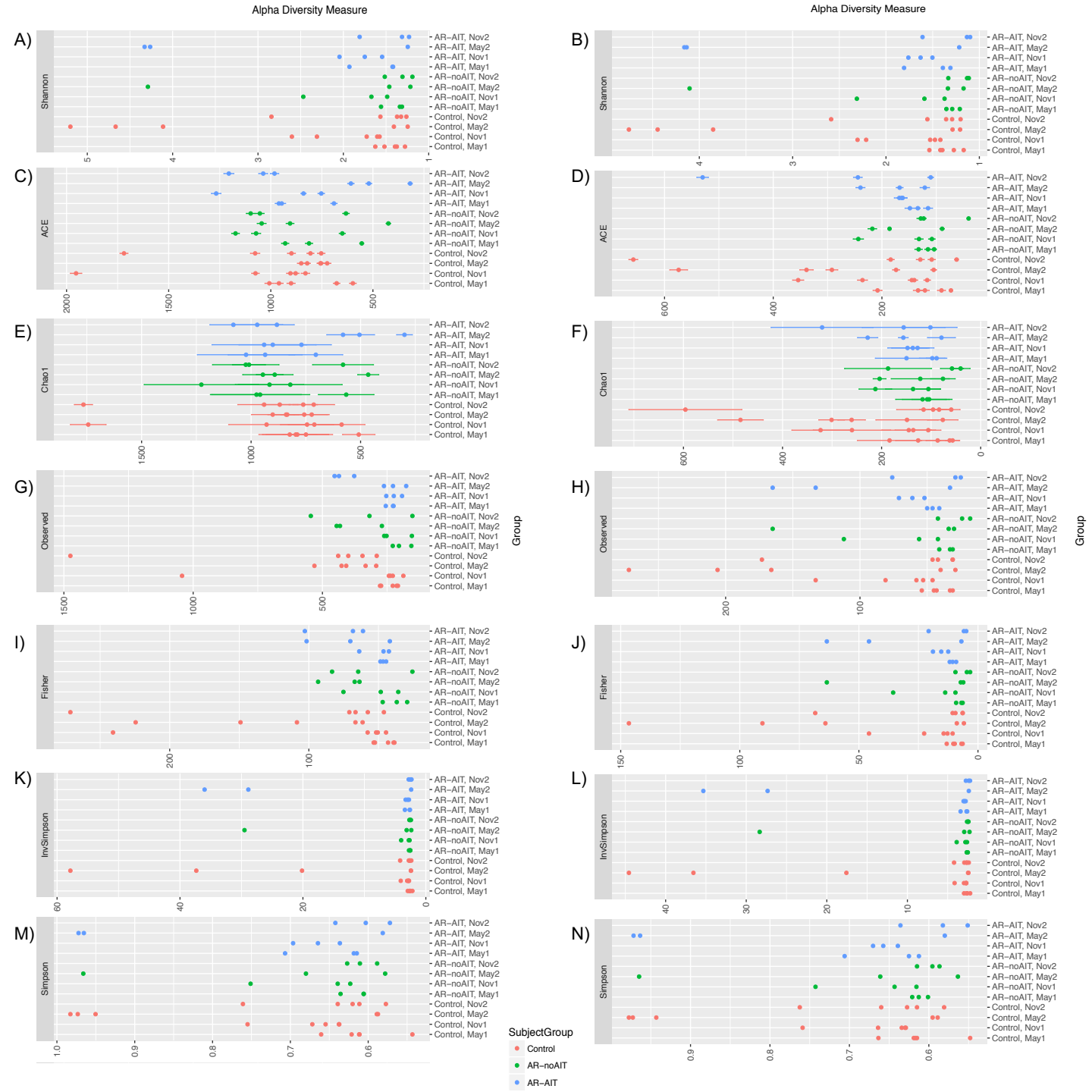

Healthy controls
AR-noAIT
AR-AIT
May1 Nov1 May2 Nov2 May1 Nov1 May2 Nov2 May1 Nov1 May2 Nov2
100 %
90 %
80 %
70 %
60 %
50 %
40 %
30 %
20 %
10 %
0 %
1800
1600
1400
1200
1000
800
600
400
200
0
S1_5 S1_4 S1_1 S1_2 S1_3 W1_5 W1_4 W1_1 W1_2 W1_3 S2_5 S2_4 S2_1 S2_2 S2_3 W2_5 W2_4 W2_1 W2_2 W2_3 S1_10 S1_11 S1_9 W1_10 W1_11 W1_9 S2_10 S2_11 S2_9 W2_10 W2_11 W2_9 S1_6 S1_8 S1_7 W1_6 W1_8 W1_7 S2_6 S2_8 S2_7 W2_6 W2_8 W2_7
Bacillus
Anoxybacillus
Acinetobacter
Moraxella
Delftia
Methanocaldococcus
Respirovirus
Alpharetrovirus
Other Archaea, Bacteria, and Virus
Archeal, bacterial, and viral CPM

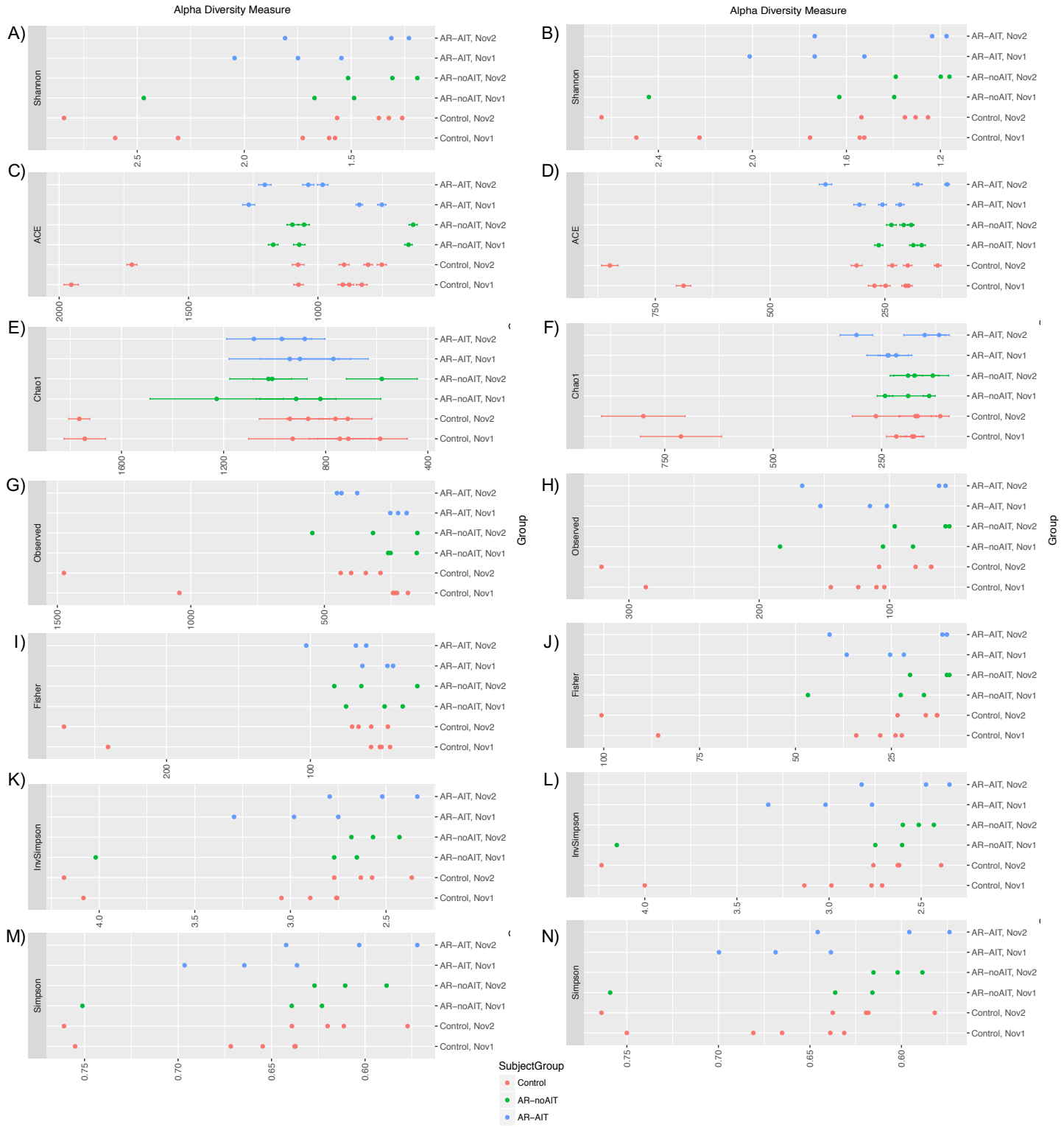
B)

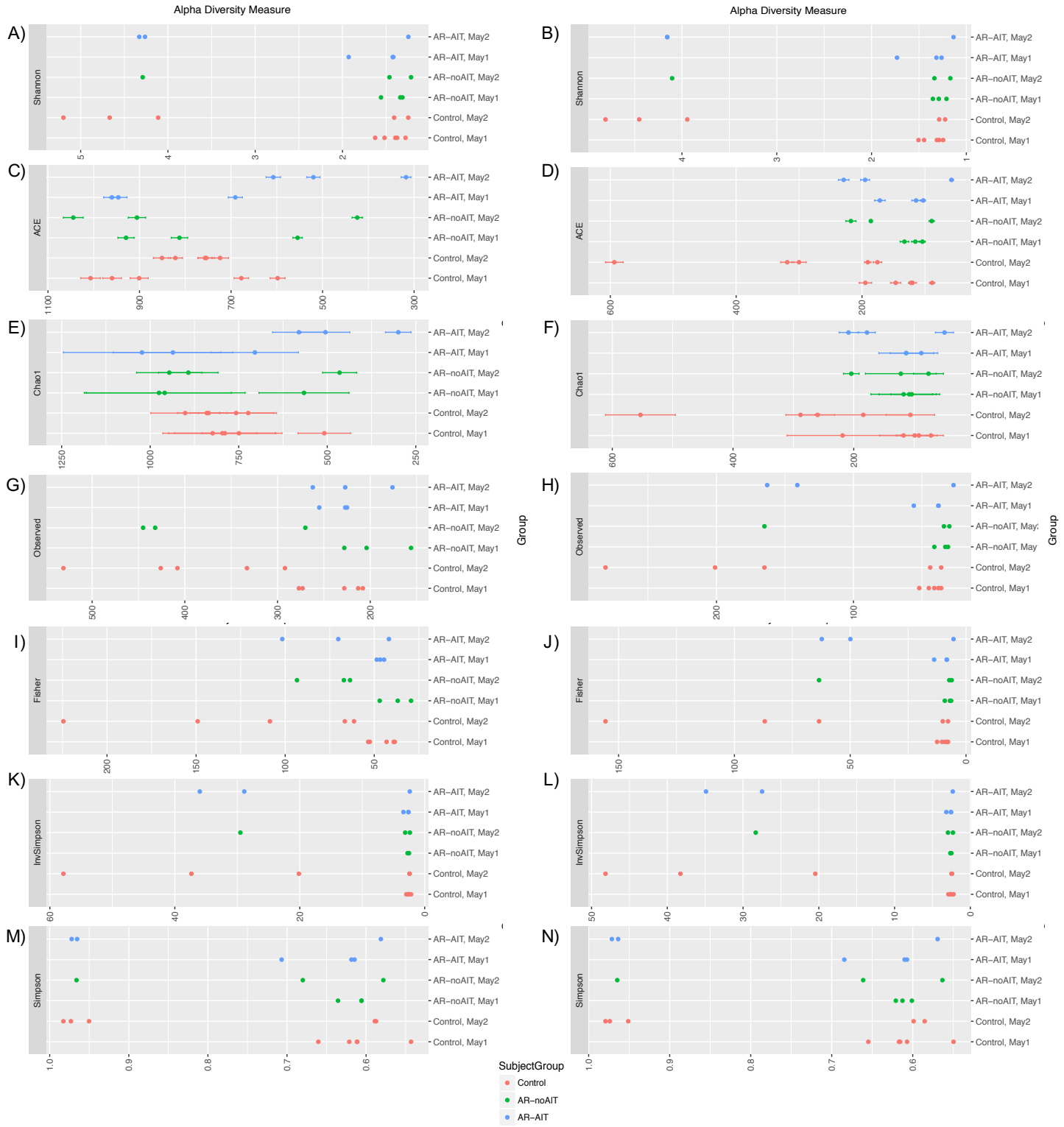

0.4
0.2
0.0
−0.2
0.0
0.2
0.4
0.6
0.8
PCO2 (7.2% of total variation)
PCO1 (69.6% of total variation)
Control
AR–noAIT
AR–AIT
May1
May2
Nov1
Nov2

C)


-3.0
14.0
-0.1097
0.4451
1.0000
S1_5 S1_4 S1_1 S1_2 S1_3 W1_5 W1_4 W1_1 W1_2 W1_3 S2_5 S2_4 S2_1 S2_2 S2_3 W2_5 W2_4 W2_1 W2_2 W2_3 S1_10 S1_11 S1_9 W1_10 W1_11 W1_9 S2_10 S2_11 S2_9 W2_10 W2_11 W2_9 S1_6 S1_8 S1_7 W1_6 W1_8 W1_7 S2_6 S2_8 S2_7 W2_6 W2_8 W2_7
Acinetobacter johnsonii
Acinetobacter junii
Bacillus cereus
Bacillus cereus group
Lactobacillus kefiranofaciens
Negativicoccus massiliensis
Pseudomonas aeruginosa
Pseudomonas aeruginosa group
Bradyrhizobium
Human respirovirus 1
Natronomonas moolapensis
Halorubrum trapanicum
Salinigranum rubrum
Acholeplasma brassicae
Anoxybacillus flavithermus
Delftia acidovorans
Avian leukosis virus
Rous sarcoma virus
Healthy controls
AR-noAIT
AR-AIT
May1 Nov1 May2 Nov2 May1 Nov1 May2 Nov2 May1 Nov1 May2 Nov2


Alpha Diversity Measure
Alpha Diversity Measure
A)
B)
C)
D)
E)
F)
G)
H)
I)
J)
K)
L)
M)
N)
Shannon
ACE
Chao1
Observed
Fisher
InvSimpson
Simpson
AR−AIT, Nov2
AR−AIT, May2
AR−AIT, Nov1
AR−AIT, May1
AR−noAIT, Nov2
AR−noAIT, May2
AR−noAIT, Nov1
AR−noAIT, May1
Control, Nov2
Control, May2
Control, Nov1
Control, May1
Group
SubjectGroup
Control
AR−noAIT
AR−AIT


Alpha Diversity Measure
Alpha Diversity Measure
A)
B)
C)
D)
E)
F)
G)
H)
I)
J)
K)
L)
M)
N)
Shannon
ACE
Chao1
Observed
Fisher
InvSimpson
Simpson
AR−AIT, Nov2
AR−AIT, Nov1
AR−noAIT, Nov2
AR−noAIT, Nov1
Control, Nov2
Control, Nov1
Group
SubjectGroup
Control
AR−noAIT
AR−AIT


Alpha Diversity Measure
A)
B)
C)
D)
E)
F)
G)
H)
I)
J)
K)
L)
M)
N)
Shannon
ACE
Chao1
Observed
Fisher
InvSimpson
Simpson
AR–AIT, May2
AR–AIT, May1
AR–noAIT, May2
AR–noAIT, May1
Control, May2
Control, May1
Group
SubjectGroup
Control
AR–noAIT
AR–AIT