Supplementary File:
Severe Obstructive Sleep Apnea is Associated with Alterations in the Nasal Microbiome and Increase in Inflammation

Benjamin G. Wu[1], I. Sulaiman[1], J. Wang[1,2], N. Shen[6], J. C. Clemente[6], Y. Li[1], R. J. Laumbach[3,5], S.-E. Lu[4], I. Udasin[4], O. Le-Hoang[3], A. Perez[5], S. Alimokhtari[5], K. Black[5], M. Plietz[3], A. Twumasi[8], H. Sanders[8], P. Melacha[1], B. Kapoor[1], B. D. Scaglione[1], A. Wang[1], C. Blazoski[1], M. D. Weiden[1], D. M. Rapoport[8], D. Harrison[1], N. Chitkara[1], E. Vicente[7], J. M. Marin[7], J. Sunderram[3], I. Ayappa[8], Leopoldo N. Segal[1]*

[1] Division of Pulmonary, Critical Care, & Sleep Medicine, New York University School of Medicine, NY
[2] Beijing Division of Pulmonary and Critical Care Medicine, Beijing Chaoyang Hospital, The Capital University of Medicine, Beijing, China
[3] Division of Pulmonary and Critical Care Medicine, Department of Medicine, Rutgers Robert Wood Johnson Medical School - New Brunswick, NJ/USA
[4] Rutgers School of Public Health - Piscataway, NJ/US
[5] Environment and Occupational Health Sciences Institute, Rutgers Biomedical Health Sciences- Piscataway, NJ/USA
[6] Department of Genetics and Genomic Sciences and Immunology Institute, Icahn School of Medicine at Mount Sinai, NY
[7] Hospital Universitario Miguel Servet, IISAragón & CIBER Enfermedades Respiratorias, Zaragoza, Spain
[8] Division of Pulmonary, Critical Care, & Sleep Medicine, Icahn School of Medicine at Mount Sinai

METHODS

Study design and participants

For this investigation, we utilized two prospective cohorts in whom sleep studies were conducted to diagnose OSA and nasal lavages were obtained. The discovery cohort was recruited from 2 sites of the World Trade Center (WTC) Health programs: NYU and Rutgers, where follow-up occurred after exposure to World Trade Center Dust in 2001 (1). The cohort consists of 472 subjects prospectively enrolled in a sub-study on nasal pathology and OSA (WTCSNORE U01OH010415) who reported no habitual snoring prior to 9/11. This discovery cohort was used to explore the associations between the microbiome, OSA, and inflammation. Exclusion criteria included: 1) gross skeletal alterations affecting the upper airway; 2) unstable chronic medical conditions known to affect OSA (CHF, stroke); 3) pregnancy or intent to become pregnant within the period of the protocol; 4) inability to sign informed consent form; 5) habitual snorer or diagnosis of OSA prior to 9/11/2001. All subjects signed informed consents to participate in this study. The research protocols were independently approved by the New York University and Rutgers University Institutional Review Boards (NYU IRB# i12-02578 and Rutgers IRB# Pro2012002164).

The validation cohort consisted of 93 consecutive subjects with suspected OSA and referred to the Sleep Center at the Hospital Miguel Servet (Zaragoza, Spain) (2). Zaragoza Sleep cohort: Inclusion Criteria: age 18 to 60 years old. Zaragoza Sleep cohort Exclusion Criteria: 1) Current smokers or smoke history of >5 pack/year; 2) Alcohol use of >3 beverages/week; 3) Body mass index ≥ 35 kg/m^2; 4) Chronic metabolic, neurologic, pulmonary, renal, hematologic, gastrointestinal, or genital-urinary disorders; 5) Known

hypertension or blood pressure >140/90 mmHg; 6) Present or past cardiovascular disorders; chronic inflammatory disorders; 7) Active infection or recent infection (<3 months); 8) Malignancy; 9) Dyslipidemia or statins use; 10) Surgery within the previous 3 months; 11) Pregnancy or likely to become pregnant; 12) Atopy, nasal allergy, and nasal polyps; 13) Previous therapy for OSA; 14) Coexistence of other than OSA sleep disorders; 15) Regular use of use of aspirin or other anti-inflammatory agents. Study procedures were approved by the Ethics and Clinical Research Committee of the Aragón Institute of Health Sciences (protocol #10/231) and informed consent was obtained from each participant.

Procedures

Diagnosis of Obstructive Sleep Apnea

All subjects underwent home sleep testing. For the discovery cohort, subjects were issued an ARES^{TM} Unicorder to take home and wear for 2 nights (SleepMed, In, West Palm Beach, FL, USA). The ARES Unicorder is worn on the forehead and measures oxygen saturation and pulse rate from reflectance oximetry, airflow from a nasal cannula/pressure transducer, snoring via acoustic microphone and head movement actigraphy, and head position from accelerometers. The device also provides audible alerts during the study if poor quality airflow or SpO2 is detected so the subject can reposition the device.

The validation cohort used the BITMED NGP 140 (Meditel Ingeniería Médica, Zaragoza, Spain) for one night (3). The montage included airflow measured by nasal cannula, thoracic-abdominal movement by inductance plethysmography, finger pulse oximetry, and body position.

Trained personnel manually scored polygraph data in accordance with American Academy of Sleep Medicine (AASM) guidelines (4). Apneas were defined as a >90% decrease from baseline in airflow lasting for at least 10 seconds. Hypopneas (4%) were defined as a 30-90% decrease from baseline in flow for at least 10 seconds associated with an oxygen desaturation of 4% or higher. AHI4 was calculated as the sum of apneas and hypopneas4% divided by valid recording time. Severity of sleep apnea was categorized by standard AASM criteria: No OSA (AHI4 <5 events/hour), mild OSA (AHI4 5-14 events/hour), moderate OSA (AHI4 15-29 events/hour), and severe OSA (AHI4 ≥30 events/hour) (5).

Nasal Lavage

From both cohorts, nasal lavage samples were collected by trained personnel using different methods. For the discovery cohort, a total of 8 mL of samples was instilled (4 mL/nostril). After brief instruction, the subject received five sprays of 0.9% sterile saline solution (100 µl/spray) into one nostril while occluding the other nostril (6). The subject was asked to inhale gently through the non-occluded nostril with each spray. At the end of five sprays, the subject gently exhaled through the nostril to expel the lavage into a sterile specimen cup (500 µl). Forceful expulsion was discouraged. This process was repeated 8 more times to complete the 4 mL instillation of sterile saline in that nostril. Then the process was repeated with the other nostril. The returned fluid from both nostrils was pooled. The sample was immediately placed on ice and processed within 2 hours. For the validation cohort, nasal lavage was collected by instilling 10 mL of saline from a syringe. This was repeated for a total of three times and collected each time in the same

sterile container (2). In addition, in the validation cohort, nasal lavage was repeated after three months for controls and subjects with OSA.

Longitudinal samples in the Validation Cohort

Longitudinal nasal lavages samples were obtained in all subjects from the validation cohort three months apart from the baseline. These samples were used to evaluate for longitudinal changes in the nasal microbiota of controls and different degree of OSA severity. A subgroup of OSA subjects (1 mild, 4 moderate, and 17 severe OSA) were treated during those 3 months with CPAP. Those subjects had a CPAP titration study with an autoCPAP device (REMstar Auto CPAP, Philips Respironics, Murrysville, Pennsylvania, USA). Adherence during the study was objectively assessed with the time counter on the device. Adherences was defined as CPAP use for an average of 4 hours/night on greater than 70% of the nights (by time counter on device) used and was required for inclusion in longitudinal analysis of subgroup that received this treatment. The average CPAP use was 5.8 hours per night (range 3.2-9.3 hours).

Measurement of inflammatory markers in nasal lavage fluid.

For the discovery cohort, nasal lavage fluid (NLF) was filtered through a 40 µm nylon mesh syringe filter to remove larger particles. An aliquot of whole nasal lavage fluid was separated for microbiome analysis. The remaining sample was then centrifugated at 500g for 10 minutes at 4°C. The cell-free supernatants were aliquoted and stored at −80°C for later analyses of soluble markers. Cells from the pellet were re-suspended in 1 mL of a buffered salt solution. The strainer was rinsed with a sputolysin solution (1:20 sputolysin/DTT) and the filtrate was centrifuged at 500g for 10 minutes at 4°C. The

resulting pellet was added to the re-suspended cells. If the resuspension appeared bloody, the red blood cells were lysed with a RBC lysis buffer. Cell counts were performed on a hemocytometer. The lower limit of detection for the cell concentration was 10,000 cells/mL. For the differential cell counts, cyto-centrifuge slides were prepared, fixed, and stained with a Wright-Geisma stain. On each slide, 200 cells were counted and proportions of epithelial cells, squamous cells, neutrophils, lymphocytes, eosinophils, and basophils were tallied. Cytokines (IL-8 and IL-6) were measured using high-sensitivity ELISA (BD Biosciences, Franklin Lakes, NJ, USA; Cat #BDB550999 and Cat #BDB550799, respectively) and values were expressed in pg/mL. The limits of detection for the assays were 0.8 pg/mL and 2.2pg/mL for IL-8 and IL-6, respectively.

For the validation cohort, nasal lavage fluid was spun down (10 min, 1,700 rpm) to separate cell free fluid from the cell pellet. The cell pellet was re-suspended into sterile saline, and re-centrifuged (5 minutes, 800 rpm). After red blood cells were lysed with 1 mL of lysis buffer, samples were washed with PBS. Lymphocytes were counted by flow cytometry using anti-CD3 antibody (FITC, BD Biosciences, San Jose, CA, USA) and images were acquired in a FacsCanto (BD Biosciences, San Jose, CA, USA). The percentage of lymphocytes were estimated by dividing the number of $CD3^+$ cells over the total viable singletons. Ten thousand lymphocytes were analyzed in a four-color flow cytometer (BD LSR, BD Biosciences, San Jose, CA, USA). In addition, IL-8 and IL-6 concentrations were measured in the cell free nasal lavage fluid using Singleplex Luminex Protein Assays (Affymetrix, Santa Clara, CA, USA; IL-8 Cat #EPX010-10204-901 and IL-6 Cat #EPX010-10213-901). The limits of detection for the assay were 0.3 pg/mL and 0.038 pg/mL, respectively.

Bacterial 16S rRNA gene marker quantitation and sequencing

For the discovery cohort, nasal lavages from all 472 subjects and 16 sterile saline samples from different sterile saline batches used to perform nasal lavages at both NYU and Rutgers were collected. For the validation cohort, nasal lavages (paired baseline and post 3 months) from all 93 subjects and 10 sterile saline sample controls were collected. For both cohorts, samples were sent to NYU where they were processed and sequenced in pool. There was approximately 1-year difference between the processing and sequencing of the samples from the discovery cohort and the processing and sequencing of samples from the validation cohort. For both set of samples, we utilized same DNA isolation approach, library preparation, and sequencing. In each sequencing run, technical controls and mock communities of microbial DNA were run.

For DNA isolation, lysis was ensured by a freeze-thaw cycle, followed by the use of lysozyme, and a heat shock step (56°C) for 60 seconds at the beginning of the DNA isolation process. DNA was extracted with an ion exchange column (Qiagen, Hilden, Germany).

Amplification and detection of 16S rRNA gene by qPCR was performed with the StepOne™ Real-time PCR System (Applied Biosystems, Foster City, CA, USA). The PCR reaction was performed in a total volume of 25 µl using the *Power*SYBR Green PCR Master Mix (Applied Biosystems, REF# 4367659), containing 400 nM of each of the universal forward primer (8F, 5'-AGAGTTTGATYMTGGCTCAG-3') and reverse primer (EUB361R, 5'-CGYCCATTGBGBAADATTCC-3'). The PCR reaction condition for amplification of DNA were initial denaturing at 95°C for 10 minutes, followed by 40 cycles

of denaturation at 95°C for 10 seconds, annealing at 54°C for 30 seconds, and extension at 72°C for 20 seconds.

High-throughput sequencing of bacterial 16S rRNA gene marker amplicons encoding the V4 region (150 base pair read length, paired-end protocol) was performed using a MiSeq Illumina Sequencer (Illumina, San Diego, CA, USA). For each sample, the V4 region of the bacterial 16S rRNA gene marker was amplified in duplicate reactions, using primer set 515F/806R, which nearly universally amplifies bacterial and archaeal 16S rRNA genes (7). Each unique barcoded amplicon was generated in pairs of 25 µl reactions with the following reaction conditions: 11 µl PCR-grade H2O, 10 µl Hot MasterMix (Beckman Coulter, Brea, CA, USA 5' Prime Cat# 2200410), 2 µl of forward and reversed barcoded primer (5 µM) and 2 µl template DNA. Reactions were run on a C1000 Touch Thermal Cycler (Bio-Rad, Hercules, CA, USA) with the following cycling conditions: initial denaturing at 94°C for 3 min followed by 35 cycles of denaturation at 94°C for 45 seconds, annealing at 58°C for 1 minute, and extension at 72°C for 90 seconds, with a final extension of 10 min at 72°C. Amplicons were quantified using Agilent 2200 TapeStation system and pooled. Purification was then performed using Ampure XT (Beckman Coulter, Brea, CA, USA Catalogue #A63882) as per the manufacturer instructions. Sequencing was then performed in MiSeq (Illumina, San Diego, CA, USA) to produce 150 base-paired end reads. For all nasal lavage samples, we obtained >5,000 reads per sample (median[IQR] = 32,680[22,052-45,504]).

The 16S rRNA gene sequences were analyzed using the Quantitative Insights into Microbial Ecology (QIIME version 1.9.1) pipeline for analysis of microbiome data (8). Reads were de-multiplexed and quality filtered with default parameters. Sequences were

then clustered (utilizing closed reference OTU picking) into operational taxonomic units (OTUs) using a 97% similarity threshold with UCLUST (9) and the Greengenes 16S reference dataset and taxonomy (10). After curation and removal of sequences potentially derived from reagent controls, the absolute OTU sequence counts were normalized to obtain the relative abundances of the taxa within each sample. The proportion of reads at the OTU or genus levels was used as a measure of the relative abundance of each type of bacteria. Shannon Diversity Index (SDI) on rarefied data was utilized to evaluate α diversity (within sample diversity). Weighted UniFrac was used to measure β diversity (between sample diversity) of bacterial communities and to perform principal coordinate analysis (PCoA) (11). We used the ade4 package in R to plot a PCoA on weighted UniFrac distances (12). To avoid negative eigenvalues in the analysis, we used the Cailliez method to convert the weighted UniFrac distance matrix into a closest corresponding matrix with Euclidean properties, which was further used for PCoA (13). Discovery cohort data is publicly available in Sequence Read Archive (SRA) under accession number PRJNA419002 and the validation cohort data is publicly available in the SRA accession number PRJNA419003. Codes and metadata utilized for analysis are available at https://github.com/segalmicrobiomelab/OSA_microbiome_repository.

Statistical analysis

Since the distributions of microbiome data are non-normal, and no distribution-specific tests are available, we used non-parametric tests of association for analysis. For association with discrete factors, we used either the Mann-Whitney test (in the case of 2 categories) or the Kruskal-Wallis ANOVA (in case of >2 categories). Paired non-parametric statistics (Wilcoxon rank sum test) was used for comparison between

longitudinal samples (baseline and post 3 months) obtained in the validation cohort. For distribution of frequencies we used Chi-Squared analysis to assess for differences in the presence or absence or population differences in our cohorts. For tests of association with continuous variables, we used non-parametric Spearman (ρ) correlation tests. False discovery rate (FDR) was used to control for multiple testing (14). Multivariate linear regression utilizing clinical covariates (e.g., age, BMI, sex, AHI4, and smoking status) was performed using microbiota signatures identified as associated with OSA as outcome (SPSS version 23.0, IBM, Armonk, NY, USA). Cytokine and inflammatory cell data from both cohorts were obtained using different methods. Therefore, this data was z-transformed and compared using parametric statistics (ANOVA). To evaluate differences in community composition between groups based on 16S data we utilized PERMANOVA (Adonis) with 1000 permutations. Subgroups based on inflammatory biomarkers were defined as above and below the third quartile for their respective measurements. To evaluate for taxonomic differences between groups, we used the non-parametric Kruskal-Wallis group comparisons in order to compare the differences of taxa between groups and to analyze those who were differentially enriched we used linear discriminant analysis (LDA) combined with Effect Size (LEfSe) (15). Features significantly discriminating among groups with LDA score >2.0 were represented as a cladogram produced by LEfSe with default parameters.

For evaluation of correlations between microbiota, clinical, and inflammatory data we utilized co-occurrence network analysis. The most abundant taxa at genus level were selected for this analysis (>2% relative abundance in at least 10% of the samples). Importantly, microbiome data is both sparse (large number of zero values) and

compositional (an increase in the relative abundance of a bacteria results in the decrease in abundance of another bacteria). This leads to an inflation of false positive correlations when using traditional statistics (16). To reduce these effects, we calculated correlations between taxa using SparCC, a tool that significantly reduces artifactual correlations (17). Correlations between taxa were calculated with ten rarefactions on the input table and 500 bootstraps, and removing non-significant correlations ($p<0.05$). Clinical and inflammatory variables were then correlated with the taxa in the co-occurrence network. Significantly correlated clinical and inflammatory variables (FDR<0.20) were kept in the network. The remaining correlations were loaded into Cytoscape v3.6.0 (18), with each node representing a bacterial genus and edges between nodes represents the strength of correlations. The network was displayed using the prefuse forced-directed layout. The co-occurrence network is presenting information on how different bacteria, biomarkers, and AHI4 are correlated: solid lines represent positive correlations (blue: among bacteria; yellow: between bacteria and clinical/inflammatory markers), while dashed lines are negative correlations.

References

1. Sunderram J, Kipen H, Laumbach RJ, Lu S-E, Udasin I, Carson JL, Le-Hoang O, Perez A, Alimokhtari S, Black K, Wu BG, Plietz M, Twumasi A, Scaglione BD, Li Y, Wang A, Gilani J, Rapoport DM, Harrison D, Chitkara N, Gumb T, Lewis C, Ayappa IA. Evaluation of the Nasal Microbiome in World Trade Center Dust Exposed Subjects. A80-B VIEWING OSA PATHOGENESIS IN COLOR INSTEAD OF BLACK AND WHITE. p. A2542-A2542.
2. Vicente E, Marin JM, Carrizo SJ, Osuna CS, Gonzalez R, Marin-Oto M, Forner M, Vicente P, Cubero P, Gil AV, Soler X. Upper airway and systemic inflammation in obstructive sleep apnoea. *Eur Respir J* 2016; 48: 1108-1117.
3. Candela A, Hernandez L, Asensio S, Sanchez-Paya J, Vila J, Benito N, Romero S. [Validation of a respiratory polygraphy system in the diagnosis of sleep apnea syndrome]. *Arch Bronconeumol* 2005; 41: 71-77.
4. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, Harrod CG. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. *J Clin Sleep Med* 2017; 13: 479-504.
5. de la Hoz RE, Aurora RN, Landsbergis P, Bienenfeld LA, Afilaka AA, Herbert R. Snoring and obstructive sleep apnea among former World Trade Center rescue workers and volunteers. *J Occup Environ Med* 2010; 52: 29-32.
6. Laumbach RJ, Fiedler N, Gardner CR, Laskin DL, Fan ZH, Zhang J, Weschler CJ, Lioy PJ, Devlin RB, Ohman-Strickland P, Kelly-McNeil K, Kipen HM. Nasal effects of a mixture of volatile organic compounds and their ozone oxidation products. *J Occup Environ Med* 2005; 47: 1182-1189.
7. Walters WA, Caporaso JG, Lauber CL, Berg-Lyons D, Fierer N, Knight R. PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics* 2011; 27: 1159-1161.
8. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; 7: 335-336.
9. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; 26: 2460-2461.
10. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *Isme J* 2012; 6: 610-618.
11. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 2005; 71: 8228-8235.
12. Dray SaD, A.B. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 2007; 22: 1-20.
13. Cailliez F. The analytical solution of the additive constant problem. *Psychometrika* 1983; 48: 305-310.
14. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 2003; 19: 368-375.

15. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. *Genome biology* 2011; 12: R60.
16. Aitchison J. The Statistical-Analysis of Compositional Data. *J Roy Stat Soc B Met* 1982; 44: 139-177.
17. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol* 2012; 8: e1002687.
18. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 2003; 13: 2498-2504.

Supplementary Table E1. Demographics of Discovery and Validation Cohorts based on OSA categories

Discovery Cohort (n=472)					
	No OSA n=168	Mild OSA n=172	Moderate OSA n=87	Severe OSA n=45	p-values
Demographics					
Age	49.0 [44.0-54.0]	53.0 [48.0-59.0]*	55.0 [47.0-60.5]*	54.0 [49.7-61.2]*	<0.001
Sex (Males)	119 (70.8%)	147 (85.5%)*	80 (92.0%)*	43 (93.0%)*	<0.001
BMI	27.5 [24.6-30.0]	28.9 [26.4-32.7]*	30.1 [27.5-32.4]*	33.3 [30.2-37.3]*	<0.001
Smokers (Smokers)	13 (7.8%)	21 (4.5%)	7 (1.5%)	4 (0.9%)	ns

Validation Cohort (n=93)					
	No OSA n=25	Mild OSA n=19	Moderate OSA n=18	Severe OSA n=31	p-values
Demographics					
Age	41.0 [31.7-54.5]	44.5 [36.2-53.2]	49.0 [37.7-54.7]	46.0 [34.7-58.5]	ns
Sex (Males)	14 (56.0%)	15 (78.9%)	13 (72.2%)	29 (93.5%)*	0.011
BMI	27.0 [25.0-29.5]	29.3 [25.7-31.7]	28.8 [26.0-30.8]	28.8 [24.5-31.0]	ns
Smokers (Smokers)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	ns

All values expressed as Median [IQR] or Total count (percentage of column total)

p values based on Kruskal-Wallis or Chi-Squared

* p values significant for the comparison of OSA group vs. Healthy Subjects.

Supplementary Table E2. Multivariate Linear Regression shows the independent strength of association between AHI4 and Microbiota signatures

Discovery Cohort all Samples										
	Shannon Index		**Streptococcus**		**Veillonella**		**Granulicatella**		**Fusobacterium**	
	Beta	*p-value*	*Beta*	*p-value*	*Beta*	*p-value*	*Beta*	*p-value*	*Beta*	*p-value*
AHI4	**0.095**	*0.040*	**0.116**	*0.012*	0.058	0.209	0.051	0.270	0.085	0.067
Age	0.060	0.170	0.000	0.998	-0.008	0.850	0.019	0.660	0.055	0.208
BMI	0.027	0.550	0.008	0.856	0.029	0.520	0.035	0.434	0.018	0.686
Sex	0.047	0.272	0.043	0.313	0.040	0.349	0.019	0.658	-0.028	0.523
Smoking	-0.014	0.745	**-0.126**	*0.004*	**-0.107**	*0.015*	-0.071	0.105	0.048	0.264

Validation Cohort All Samples										
	Shannon Index		**Streptococcus**		**Prevotella**		**Pseudomonas**		**Haemophilus**	
	Beta	*p-value*	*Beta*	*p-value*	*Beta*	*p-value*	*Beta*	*p-value*	*Beta*	*p-value*
AHI4	0.221	0.057	**0.282**	*0.015*	0.084	0.472	0.016	0.887	**0.241**	*0.036*
Age	-0.079	0.465	-0.035	0.741	0.056	0.613	0.182	0.097	-0.175	0.102
BMI	0.005	0.964	0.021	0.852	-0.067	0.555	0.083	0.459	0.059	0.589
Sex	0.022	0.843	-0.078	0.480	0.064	0.572	0.013	0.912	0.008	0.940

Supplementary Table E3. Inflammatory Markers in the Discovery and the Validation Cohort[¶]

	Discovery cohort (n=472)				
	No OSA	**Mild OSA**	**Moderate OSA**	**Severe OSA**	**p-values[§]**
	n=168	n=172	n=87	n=45	
Inflammatory markers					
Neutrophils	-0.11 ± 1.04	0.14± 1.06	0.08 ± 0.79	0.29 ± 1.15	p=ns
IL-8	-0.18± 1.03	0.16 ± 0.89	0.16 ± 0.87	0.29 ± 1.25	p=0.007
IL-6	-0.14 ± 0.97	0.22 ± 1.00	0.24 ± 1.01	0.36 ± 1.13	p=0.008

	Validation cohort (n=93)				
	No OSA	**Mild OSA**	**Moderate OSA**	**Severe OSA**	**p-values[§]**
	n=25	n=19	n=18	n=31	
Inflammatory markers					
Lymphocytes	-0.98 ± 0.78	-0.32 ± 0.87	0.27 ± 0.36	0.77 ± 0.56	p<0.0001
IL-8	-0.34 ± 0.89	0.12 ± 0.76	0.02 ± 0.87	0.40 ± 0.87	p=0.002
IL-6	-0.04 ± 0.99	0.13 ± 0.85	-0.07 ± 1.06	0.03 ± 1.09	p=ns

[¶] Z transformed data presented

[§] ANOVA calculation

All values expressed as mean ± standard deviation. Lymphocytes not available for the Discovery cohort and Neutrophils not available for the Validation cohort.

Supplementary Figure Legends

Supplementary Figure E1. Bacterial load of samples evaluated by qPCR. (*A***)** Using qPCR, the quantitative bacterial DNA was assessed utilizing universal 16S primers. No differences in bacterial DNA load between nasal samples from no OSA subjects compared to nasal samples from OSA severity in the discovery cohort (Mann-Whitney p=ns). (*B*) Similarly**,** in the validation cohort there were no differences in bacterial load between groups (Mann Whitney p=ns). Sterile saline used for lavages were included (blue dots).

Supplementary Figure E2. Analysis of microbial DNA present in sterile saline used to perform nasal lavage at Rutgers University and New York University. (*A*) β diversity plot based on weighted UniFrac distance showed no significant differences between sterile saline used to perform nasal lavages at NYU and Rutgers (PERMANOVA p=ns). (*B*) Unsupervised hierarchical clustering with heatmap of most abundant genera (relative abundance ≥2%) present in sterile saline.

Supplementary Figure E3. Unsupervised hierarchical clustering with heatmap of most abundant genera in nasal lavage samples from the discovery cohort. Nasal lavage samples were characterized by high relative abundance of *Staphylococcus* and *Corynebacterium*. Subclusters of samples were enriched with *Streptococcus*, *Veillonella*, *Prevotella*, *Peptoniphilus*, *Anaerococcus*, *Moraxella*, and *Neisseriaceae*(u.g.). Nasal lavage samples did not differentially cluster according to location (blue samples were

obtained at Rutgers University and purple samples were obtained at New York University).

Supplementary Figure E4. Unsupervised hierarchical clustering with heatmap of most abundant genera in nasal lavage samples from the validation cohort. Nasal lavage samples from all samples from Zaragoza sleep cohort demonstrate high relative abundance of *Staphylococcus* and *Corynebacterium*. Subclusters were characterized by high relative abundance of *Veillonella*, *Streptococcus*, and *Prevotella*.

Supplementary Figure E5. Taxonomic differences between groups of OSA diagnosis for the discovery and validation cohorts show enrichment in mild and moderate OSA. Linear discriminant analysis (LDA) identified taxa differentially enriched (LDA>2.0). LDA scores were paired with representation of relative abundance in groups of (*A*) mild OSA vs. no OSA subjects, (*B*) moderate OSA vs. no OSA subjects, and (*C*) severe OSA vs. no OSA subjects for the discovery cohort. For the validation cohort, the LDA scores were paired with representation of relative abundance in groups of (*D*) mild OSA vs. no OSA subjects, (*E*) moderate OSA vs. no OSA subjects, and (*F*) severe OSA vs. no OSA subjects.

Supplementary Figure E6. Taxonomic differences between groups of high inflammation and lower inflammation across all samples. Cladograms created using LEfSe demonstrate microbial differences across discovery cohort (*A*) high vs. low neutrophils and (*B*) high vs. low IL-8 levels. In the validation cohort, differences in

taxonomy were found in (*C*) high vs. low lymphocytes and (*D*) high vs. low IL-8 levels. Linear discriminant analysis identified differences among taxa that were differentially enriched (LDA>2.0). LDA scores were paired with representation of relative abundance for (*E*) high vs. low neutrophils in the discovery cohort, (*F*) high vs. low IL-8 in the discovery cohort, (*G*) high vs. low lymphocytes in the validation cohort, and (*H*) high vs. low IL-8 in the validation cohort.

Supplementary Figure E7. High IL-6 vs. low IL-6 levels demonstrate enrichment of microbiota. (*A*) A PCoA plot shows compositional differences using weighted UniFrac between high IL-6 compared to low IL-6 levels of the discovery cohort subjects (PERMANOVA $p<0.05$). Moreover, taxonomic differences demonstrated an enrichment of *Moraxella* (LDA>2.0). (*B*) A PCoA plot shows compositional differences between high IL-6 compared to low IL-6 levels of the validation cohort subjects (PERMANOVA p=ns). *Brochothrix* was enriched in the nasal lavage samples with high IL-6 in the validation cohort (LDA>2.0).

Supplementary Figure E8. Specific taxonomic differences that were found in high vs. low IL-6 samples in the discovery and validation cohorts. (*A*) In the discovery cohort, an increase in relative abundance of *Moraxella* found in samples that had increased IL-6 levels (Mann-Whitney p=0.016) and with increased Akkermansia (Mann-Whitney p=0.01). (*B*) In the validation cohort, increases in *Bronchothrix* (Mann-Whitney p=0.0037), *Lactococcus* (Mann-Whitney p=0.019), and *Enhydrobacter* (Mann-Whitney p=0.019) were observed in high IL-6 samples. Samples with less than $5x10^{-5}$ relative

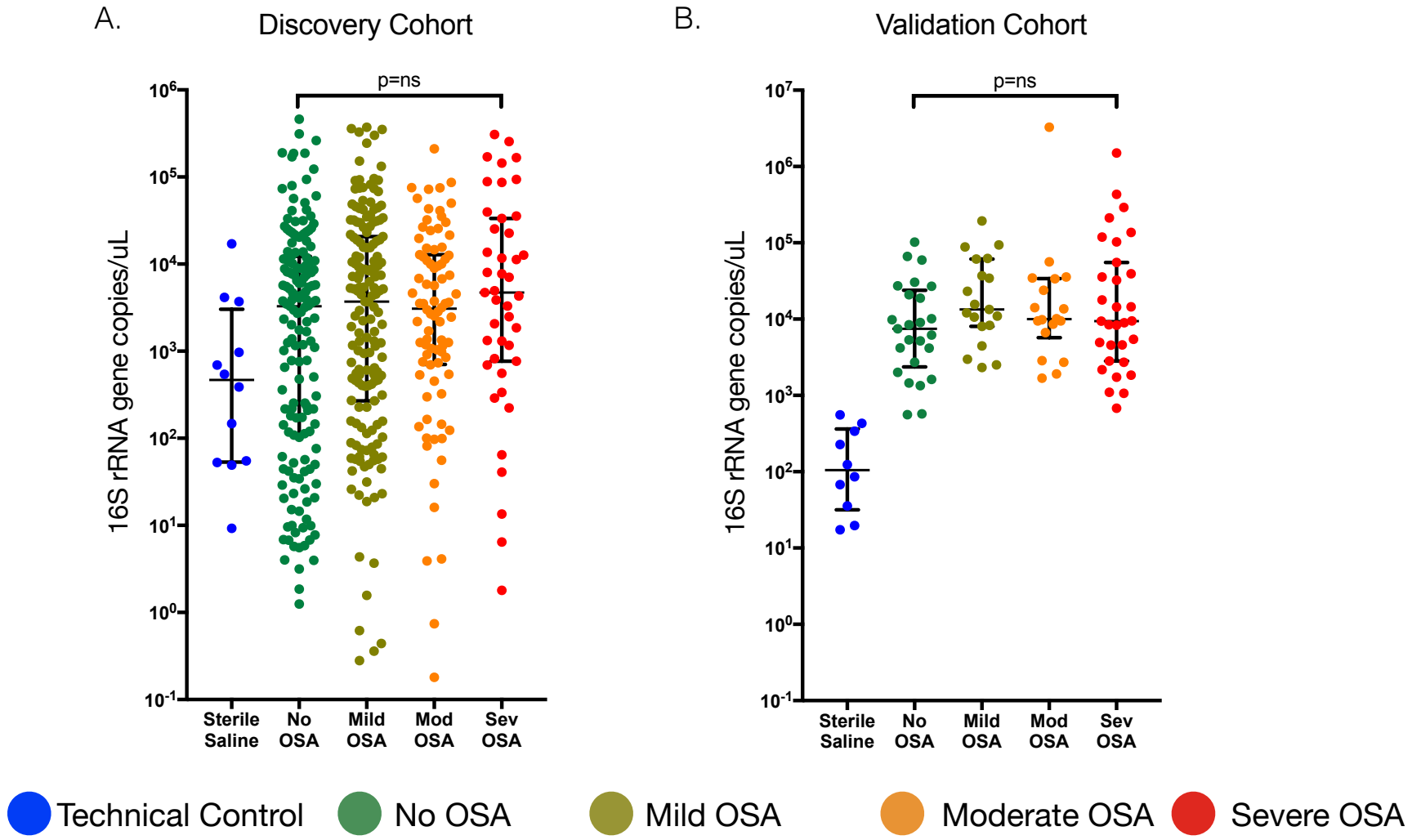
abundance of specific taxa were considered below the lower limit of detection (dotted line).

Supplementary Figure E9. α diversity comparison between high vs. low lymphocytes in the validation cohort. Nasal lavages with high lymphocyte percentage had higher SDI as compared to the low lymphocyte group (Mann-Whitney p=0.005).

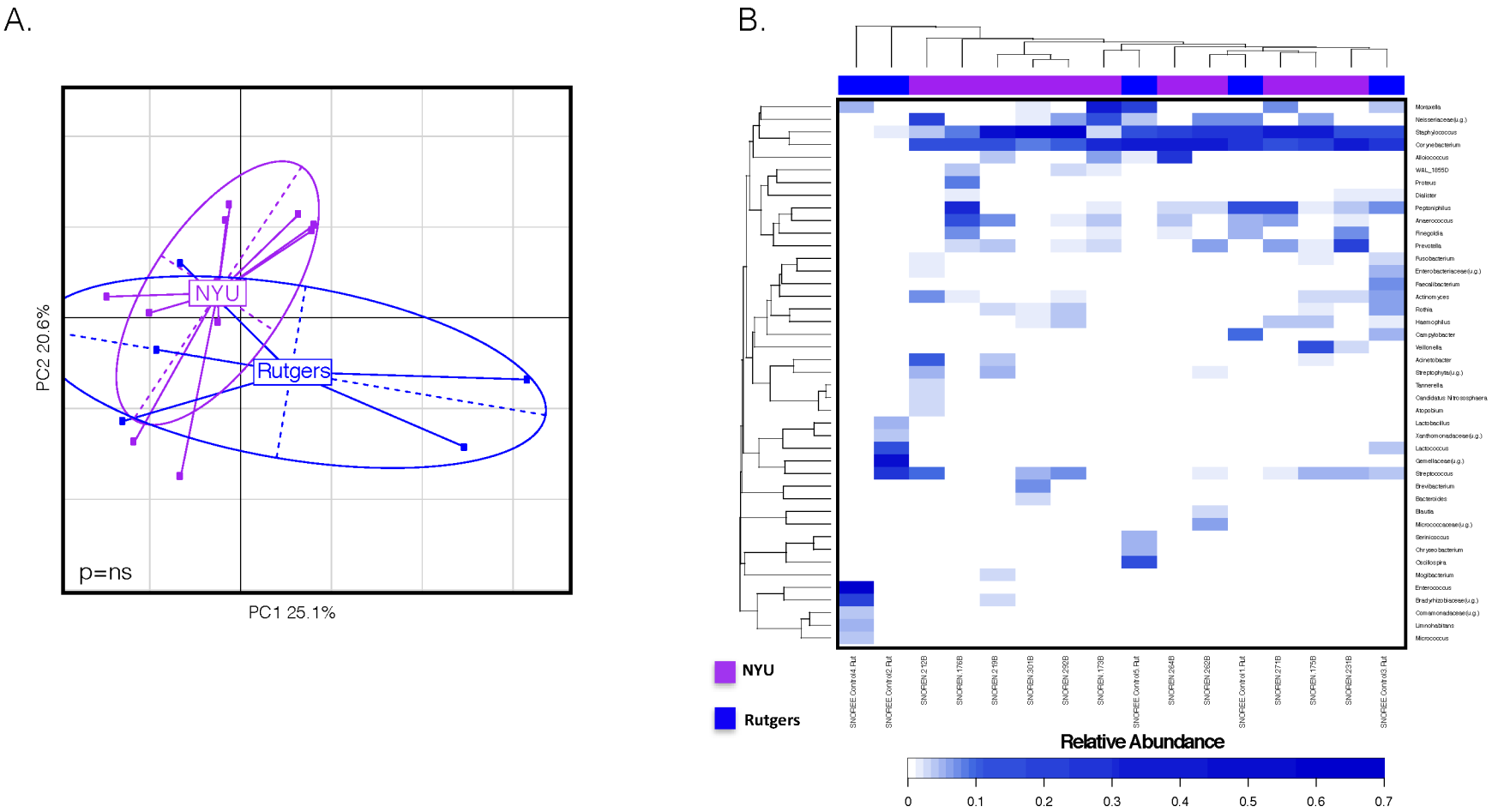
Supplementary Figure E10. **Longitudinal change in α diversity of nasal microbiota in the validation cohort.** Longitudinal change in α diversity (SDI) comparing baseline samples with samples obtained after three months. Paired analysis showed no significant change (Wilcoxon rank sum p=ns for all comparisons).

Supplementary Figure E11. Longitudinal change in β diversity of nasal microbiota in the validation cohort. Longitudinal change in the composition of the nasal microbiota was evaluated for each group of OSA severity and controls using PCoA plots based on weighted UniFrac distances: (*A*) no OSA, (*B*) mild OSA, (*C*) moderate OSA, and (*D*) severe OSA. No significant differences were noted in β diversity (PERMANOVA p=ns for all comparisons).

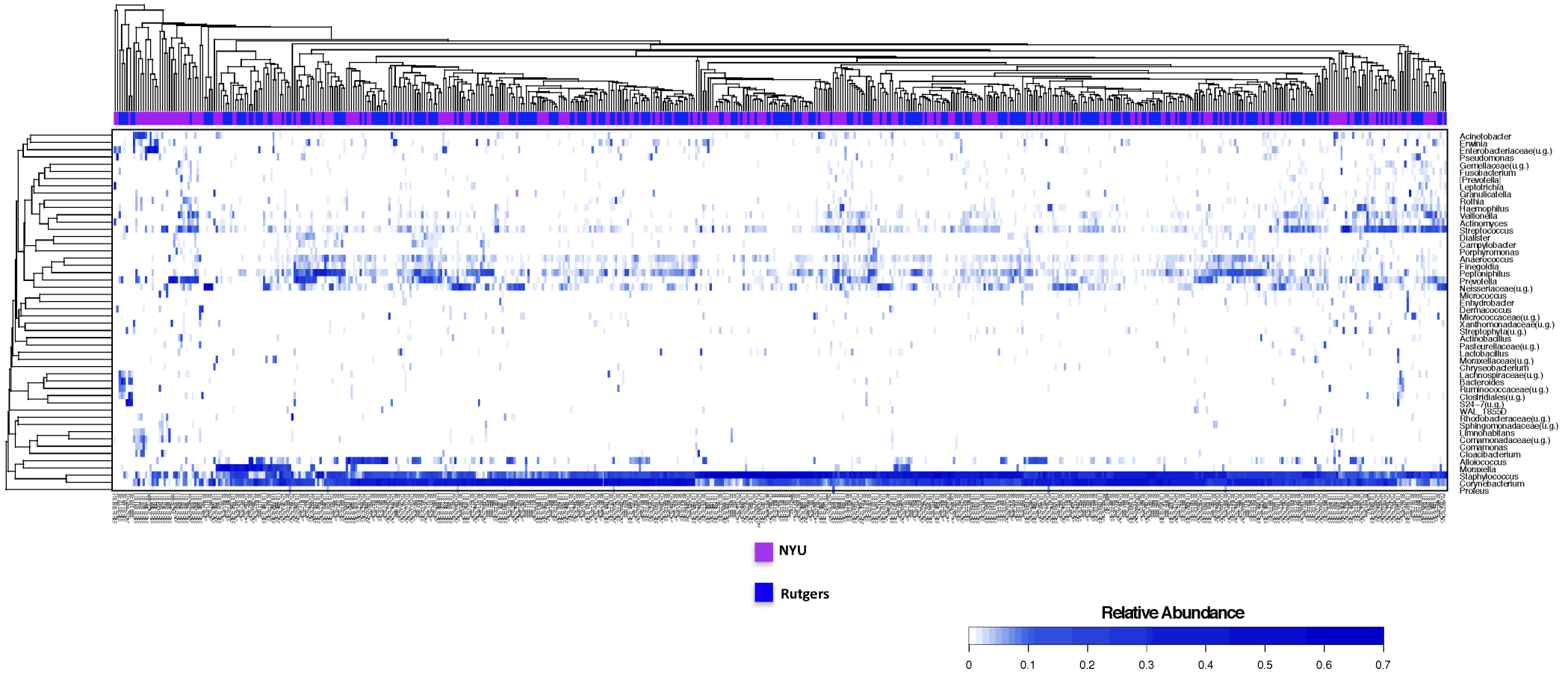
Supplementary Figure E1.



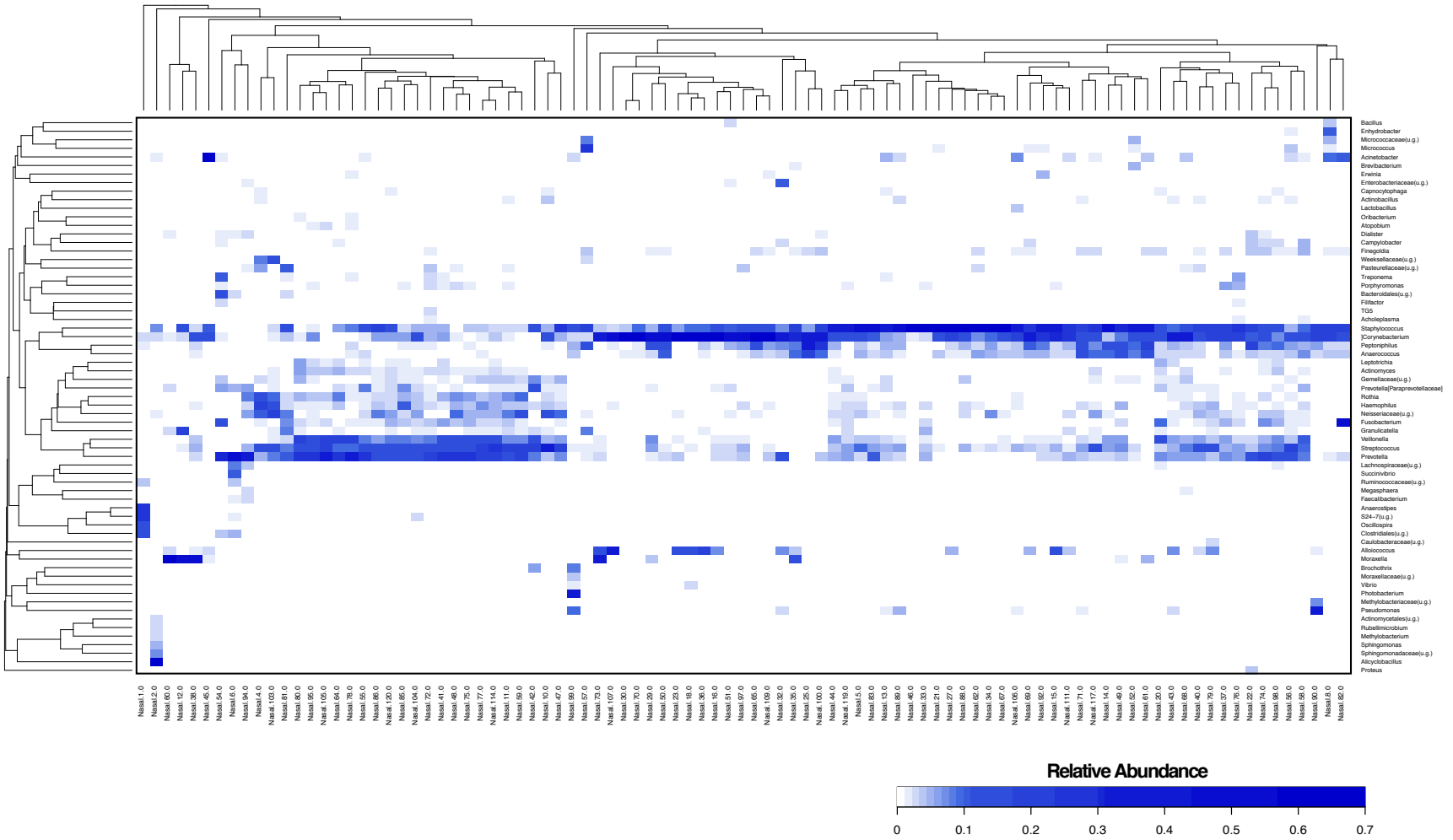
Supplementary Figure E2.



Supplementary Figure E3.



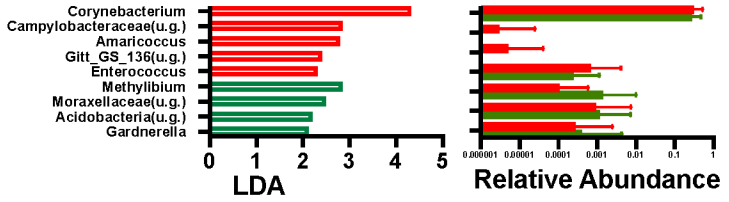
Supplementary Figure E4.



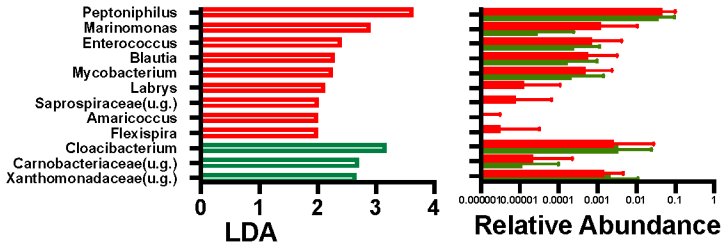
Supplementary Figure E5.

Discovery Cohort

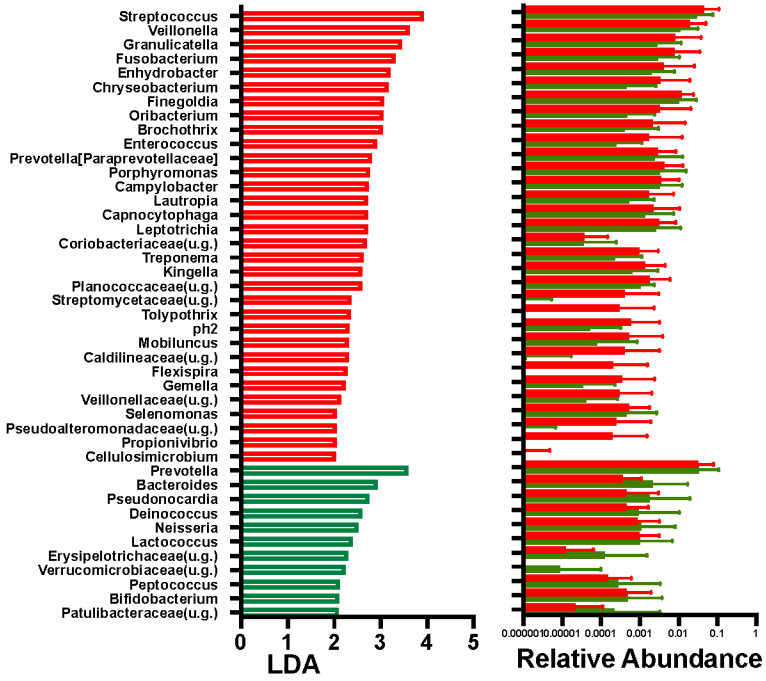
A. Mild OSA vs. No OSA



B. Moderate OSA vs. No OSA

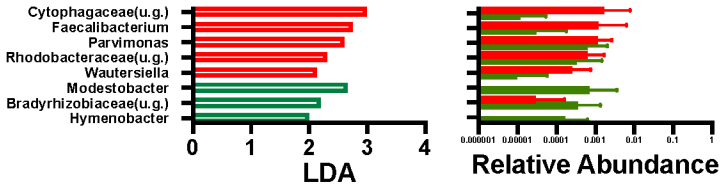


C. Severe OSA vs. No OSA

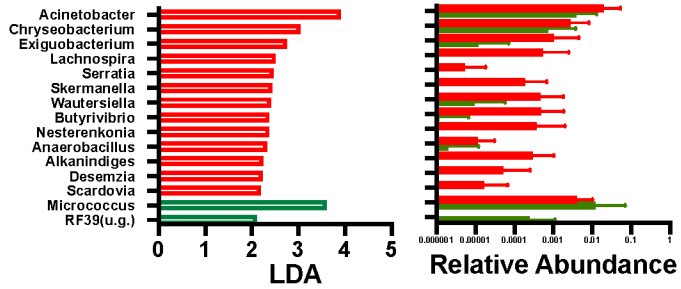


Validation Cohort

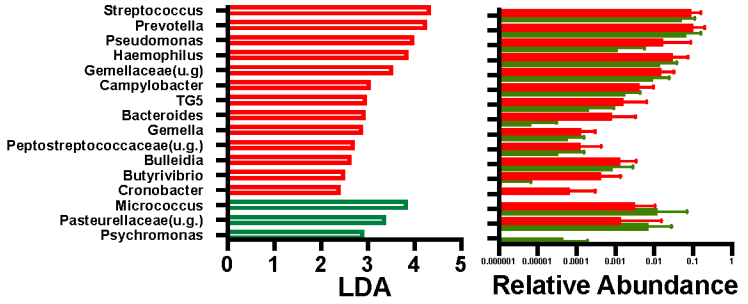
D. Mild OSA vs. No OSA



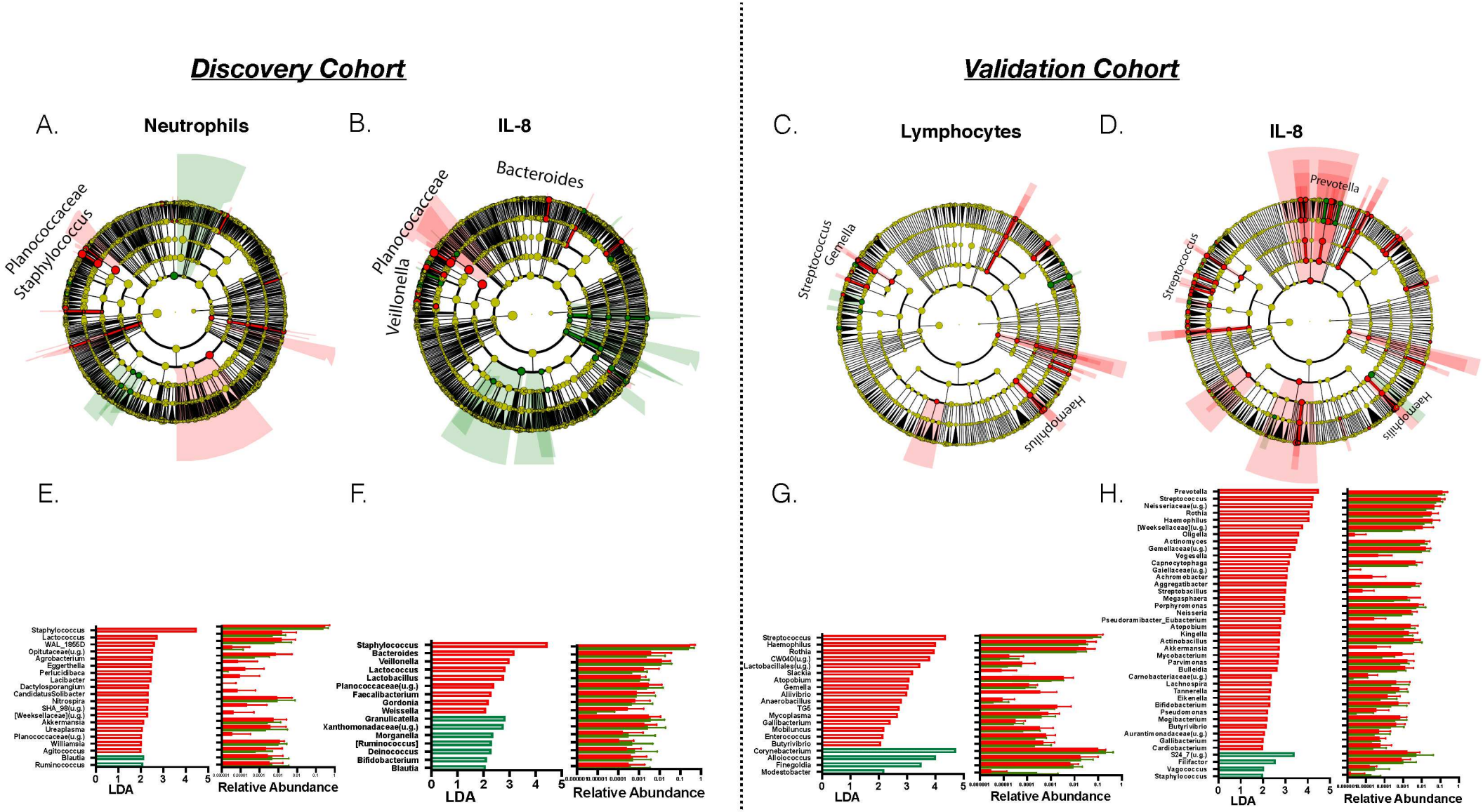
E. Moderate OSA vs. No OSA



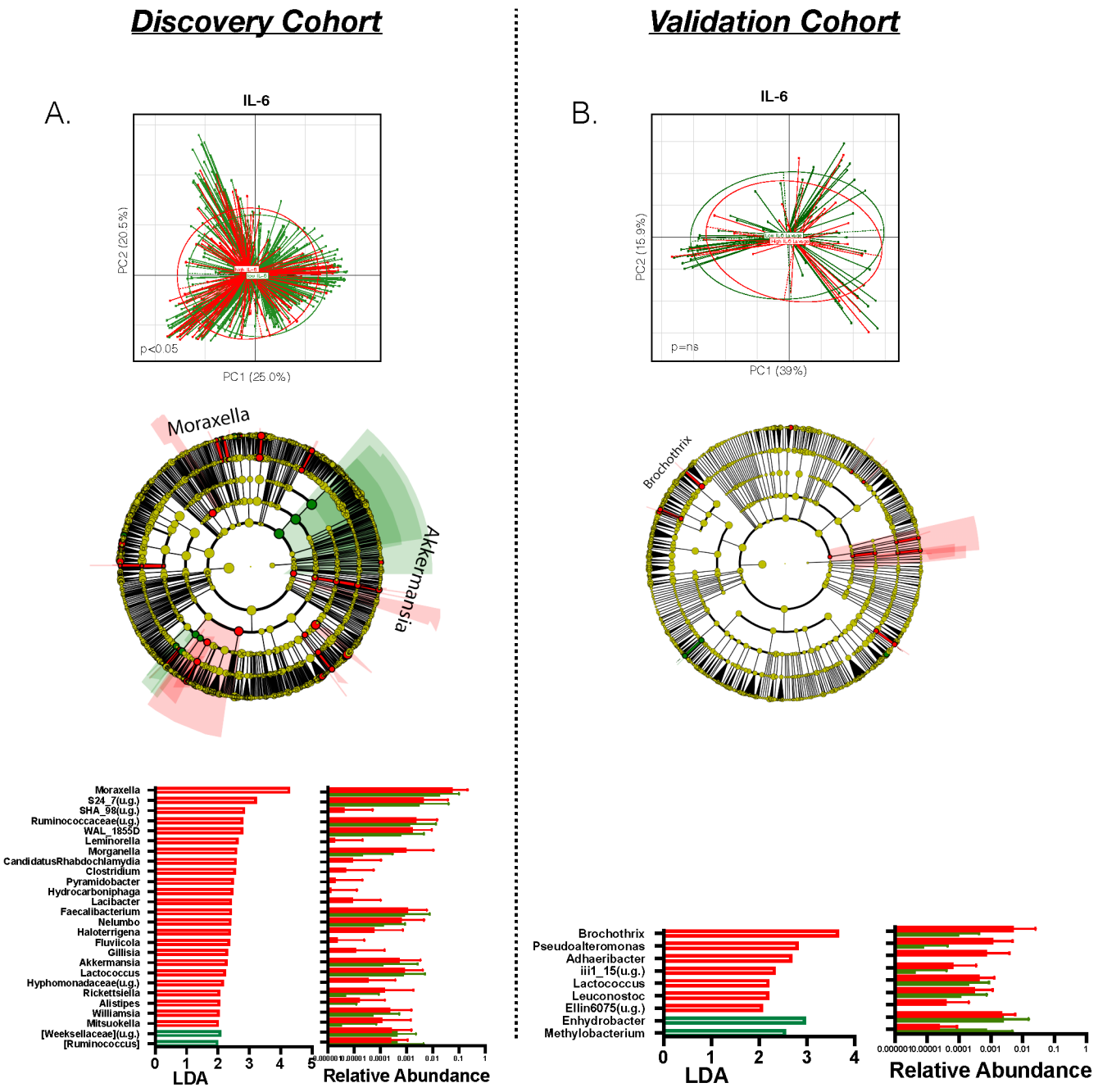
F. Severe OSA vs. No OSA



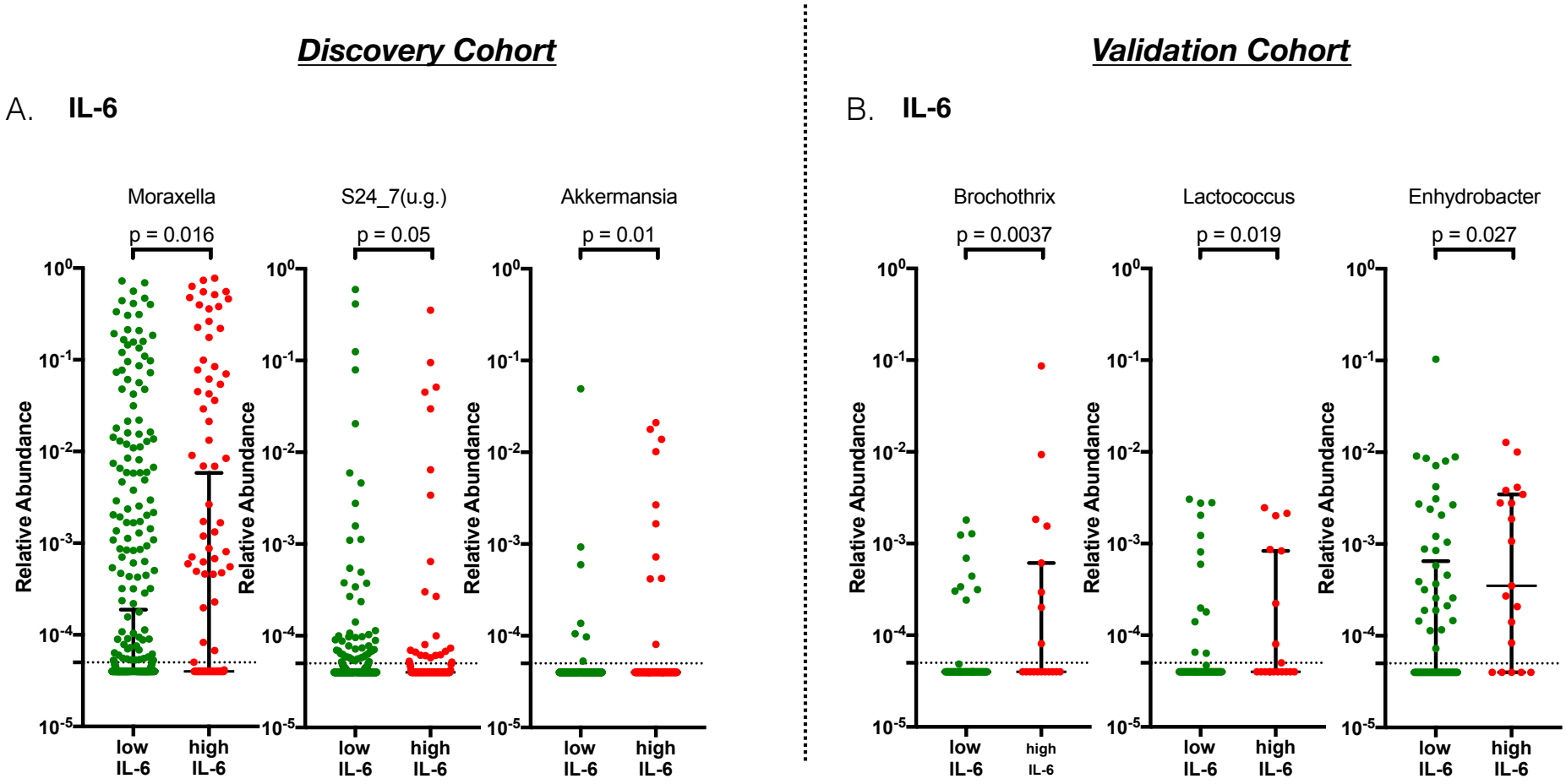
Supplementary Figure E6.



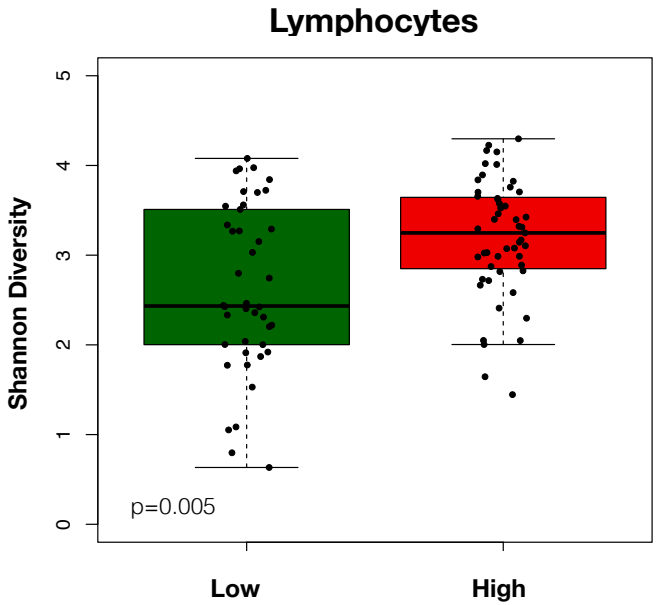
Supplementary Figure E7.



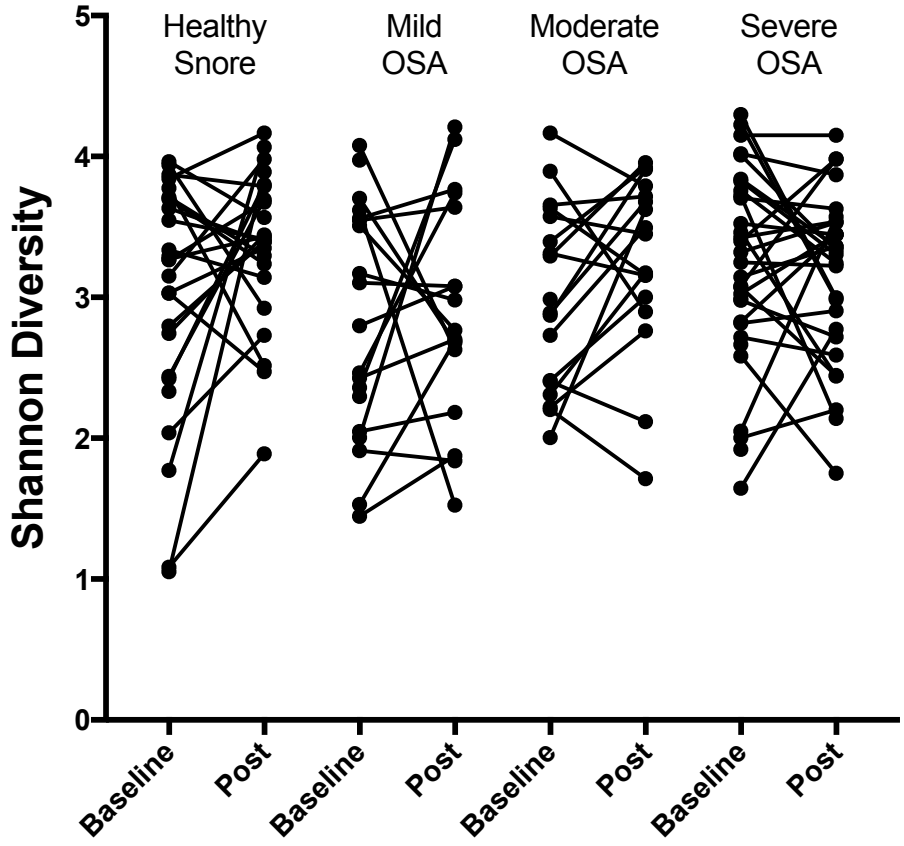
Supplementary Figure E8.



Supplementary Figure E9.



Supplementary Figure E10.



Supplementary Figure E11.