Lung microbiome and host immune tone in subjects with Idiopathic Pulmonary Fibrosis treated with inhaled interferon-γ

Supplementary materials

Jing Wang[1,2], Melissa Lesko[2], Michelle H. Badri[3], Bianca Kapoor[2], Benjamin G. Wu[2], Yonghua Li[2], Gerald Smaldone[4], Richard Bonneau[3,6,7], Zachary D. Kurtz[5], Rany Condos[2], Leopoldo N. Segal[2]*

[1] Division of Pulmonary and Critical Care Medicine, Beijing Chaoyang Hospital, The Capital University of Medicine, Beijing, China
[2] Division of Pulmonary, Critical Care, & Sleep Medicine, New York University School of Medicine, NY
[3] Department of Biology, Center for Genomics and Systems Biology, New York University, NY
[4] Division of Pulmonary, Critical Care and Sleep Medicine, State University of New York at Stony Brook, NY
[5] Department of Microbiology, New York University School of Medicine, NY
[6] Courant Institute of Mathematical Sciences, New York University, NY
[7] Simons Center for Data Analysis, Simons Foundation, New York, NY

Jing Wang, MD	wangjing@bjcyh.com
Melissa Lesko, DO	Melissa.Lesko@nyumc.org
Michelle H. Badri	mhb383@nyu.edu
Bianca Kapoor	Bianca.Kapoor@med.nyu.edu
Benjamin G. Wu, MD	Benjamin.Wu@nyumc.org
Yonghua Li, MD, PhD	Yonghua.Li@nyumc.org
Gerald Smaldone, PhD	Gerald.Smaldone@stonybrookmedicine.edu
Richard Bonneau, PhD	rb133@nyu.edu
Zachary D. Kurtz, PhD	zk311@nyu.edu
Rany Condos, MD	Rany.Condos@nyumc.org
Leopoldo N. Segal, MD	Leopoldo.Segal@nyumc.org

Corresponding Author/ Address for Reprints:
Leopoldo N. Segal, MD[1] Leopoldo.Segal@nyumc.org

NYU School of Medicine
462 First Ave 7W54
New York, NY 10016
Tel: (212) 562-3752
Fax: (212) 263-7445

Methods

IPF diagnosis, enrollment and exclusion criteria

Diagnosis was made either by open lung biopsy with a histopathologic pattern of Usual Interstitial Pneumonitis (UIP), or based on the HRCT of the Chest utilizing the American Thoracic Society/European Respiratory Society consensus guidelines. Individuals with interstitial lung disease other than IPF or with severe pulmonary hypertension based on trans-thoracic echocardiography were not permitted to participate in the study. Exclusion criteria based on data from pulmonary function tests included a Forced Vital Capacity (FVC) less than 55% predicted, diffusion capacity of carbon monoxide less than 30% predicted, Forced Expiratory Volume in one second (FE1) /FVC of <65%, and a 6 minute walk distance of less than 200 meters. Patients who had a known hypersensitivity to IFN-γ, or medical conditions including severe cardiac disease, peripheral vascular disease, pregnancy, or seizure disorder were also removed from consideration.

Bronchoscopy procedure

Patients received anesthesia with conscious sedation and nebulized lidocaine. The bronchoscope was advanced through the mouth or nose and through the vocal cords. It was wedged into a right middle lobe or lingular segmental airway with the location decided by the physician doing the procedure based on the extent of imaging abnormality on high-resolution CT. Bronchoscopic alveolar lavage (BAL) was done with four instillations of 50 mL of sterile isotonic saline

aliquots for a total of 200 mL and all the collected fluid was pooled. Samples of BAL for microbiome analysis were transported on ice and 5 mL volumes were split into three portions and then centrifuged in 2 mL dolphin-nosed Eppendorf tubes at 13000 revolutions per min in a microcentrifuge at 4°C for 30 min. Supernatants and the pellets were snap frozen in liquid nitrogen and stored at –80°C until required for DNA extraction and measurement the levels of biomarkers.

Bacterial 16S rRNA-encoding genes quantitation and sequencing

DNA was then extracted from samples using an ion exchange column (Qiagen). For each sample, the V4 region of the bacterial 16S rRNA gene was amplified in duplicate reactions, using primer set 515F/806R, which nearly universally amplifies bacterial and archaeal 16S rRNA genes.[1, 2] Each unique barcoded amplicon was generated in pairs of 25µl reactions with the following reaction conditions: 11µl PCR-grade H2O, 10µl Hot MasterMix (5 Prime Cat# 2200410), 2µl of forward and reversed barcoded primer (5µM) and 2µl template DNA. Reactions were run on a C1000 Touch Thermal Cycler (Bio-Rad) 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 Cat# A63882) as per the manufacturer instructions. Sequencing was then performed in MiSeq (Illumina) to produce 150 base-paired end reads. Technical negative

controls included DNA free water (n=2) processed in parallel with BAL samples (which include DNA isolation and library preparation) and elution buffer (from the DNA isolation kit, n=2). The obtained 16S rRNA gene sequences were analyzed using the QIIME package (version 1.9.1) for analysis of community sequence data as previously described.[3, 4] PERMANOVA (Adonis) testing was used to compare β-diversity of groups. The observed OTU counts were normalized to obtain the relative abundances of the microbiota in each sample. These relative abundances at 97% OTU similarity and each of the 5 higher taxonomic levels (phylum, class, order, family, and genus) were tested for univariate associations with clinical variables. To decrease the number of features, we focused only on major taxa and OTUs, defined as those with relative abundance >1% in at least one sample. We used the ade4 package in R to PCoA on weighted UniFrac distances.[5] Principal coordinate analysis (PcoA) was used to visualize potential clustering patterns among samples based on the β-diversity. For comparisons of α-diversity, β-diversity, or taxonomy between groups at baseline, non-parametric (Mann-Whitney) tests were used. To evaluate for changes in α and β diversity post-treatment with IFN-γ, paired non-parametric comparisons were used.

Measurement of biomarkers.

In vivo inflammation was assessed by BAL cell count differential and biomarkers. Forty-seven biomarkers were measured in a Luminex platform (Millipore) in BAL, plasma and BAL cell supernatant (after 24hrs of culture in RPMI). In BAL fluid, measurable levels of biomarkers (defined as above the lowest standard in >50% of the samples) were achieved in 34/47 biomarkers (EGF, eotaxin, FGP-2, Fit-3-

Ligand, Fractalkine, G-CSF, GM-CSF, GRO, IFNα2, IL-13, IL-15, IL-17, IL-1α, IL-1ra, IL-5, IL-6, IL-7, IL-8, IP-10, MCP-1, MCP-3, MDC, MIP-1α, MIP-1β, sCD40L, TGFα, TNFα, TNFβ, VEGF, CXCL9-MIG, ENA-78, PDGF-AA, PDGF-AB-BB, and RANTES). In BAL cell supernatant, measurable levels of biomarkers were achieved in 35/47 (Eotaxin, FGP-2, Fit-3-Ligand, Fractalkine, G-CSF, GM-CSF, GRO, IFNα2, IL-13, IL-15, IL-17, IL-1α, IL-1ra, IL-2, IL-3, IL-5, IL-6, IL-7, IL-8, IP-10, MCP-1, MCP-3, MDC, MIP-1α, MIP-1β, sCD40L, TNFα, TNFβ, VEGF, TGF-β2, CXCL9-MIG, ENA-78, PDGF-AA, PDGF-AB-BB, and RANTES). In plasma, measurable levels of biomarkers were achieved in 30/47 (EGF, Eotaxin, FGP-2, Fit-3-Ligand, Fractalkine, G-CSF, GM-CSF, GRO, IFNα2, IFNγ, IL-13, IL-15, IL-17, IL-1α, IL-1ra, IL-2, IL-5, IL-6, IL-7, IL-8, IP-10, MCP-1, MCP-3, MDC, MIP-1α, MIP-1β, sCD40L, TNFα, TNFβ, and VEGF). Biomarkers were categorized as Th1 (GM-CSF, IFNγ, IL-2, IL-12, and TNFβ), Th2 (IL4, IL5, IL-6, IL-9, IL-10, IL-13, and GM-CSF), Th17 (IL-1β, IL-17, IL-6, and Fractalkine), and IPF (IL-1β, IL-17, IL-13, TGFβ, MCP-1, IFNγ, PDGF-AA, and PDGF-AB-BB).

Network analysis

Using the compPLS framework outlined in Ramanan et al. (2016) we aimed to detect associations between genera in the lung mucosa and host immune phenotypes.[6] To prevent detection of statistically spurious associations we perform a centered log-ratio (clr) transformation on the OTU relative abundance data; apply a variance decomposition to extract within-subject variation (pre- and post-treatment); and finally estimate a sparse linear model via sparse Partial

Least Squares (sPLS) regression. This detects associations between the within subect variation of a sparse set of multi-collinear features (OTUs) and responses (host immune biomarkers).[7] Host immune biomarkers included cytokines, chemokines, and growth factors grouped by BAL fluid, BAL cell supernatant, and plasma from which they were measured. We then used paired pre-post samples, filtered most abundant OTUs at the genus level (>1% relative abundance in at least 30% of the samples). This filtering resulted in a genus level OTU table with 48 taxa. We decomposed the clr-transformed OTUs and host response data using a one-factor variance decomposition to isolate within-patient (fixed) effects. For each sPLS run we set the number of latent components to the number of nonzero singular values in the cross-covariance matrix. To find a sparse set of significant associations between genera and immune profiles in the BAL fluid, BAL cell supernatant, and plasma we applied sPLS and used stability approach to regularization selection (StARS)[8] to select the sparsity level of the linear model and used bootstrap based empirical p-value calculation to assess significance of associations of the StARS-selected support.[9] We then calculated empirical p-values over 5000 bootstraps and set a p-value threshold of 0.05. Feature selection as performed in the compPLS framework reduces false discovery but does not completely eliminate the issue of multiple comparisons. We visualized significant taxa-cytokine associations as a network using the igraph package in R.[10] Of these networks we highlight profiles associated with Th1, Th2, Th17, and IPF immune response.[11]

Supplementary Figures

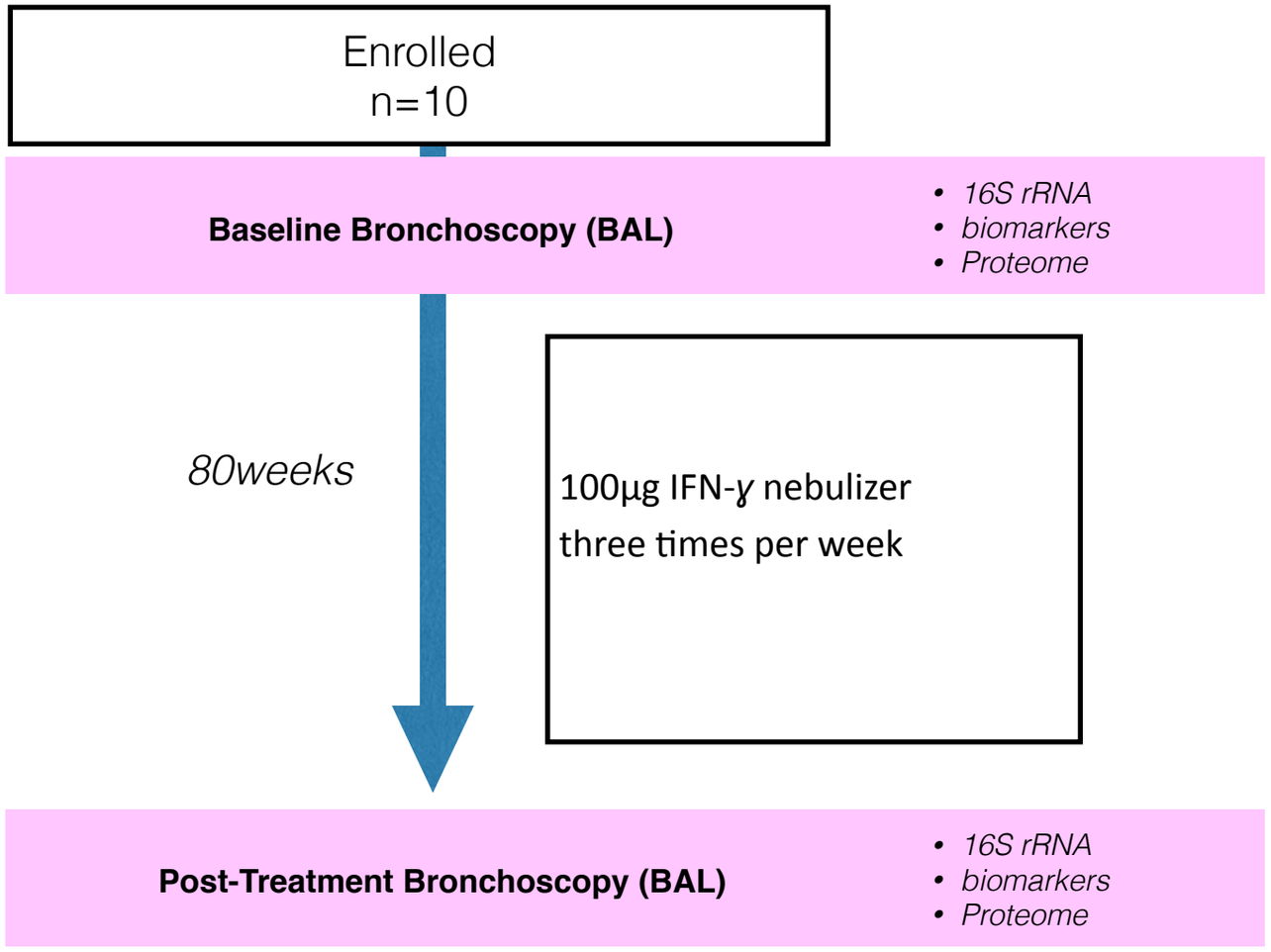
Supplementary Figure 1. Diagram for procedures on subjects. Subjects underwent bronchoscopy with BAL at baseline. After bronchoscopy, subjects were treated with inhaled IFN-γ. Post-treatment bronchoscopy was performed at the end of the treatment period. BAL was utilized to evaluate microbiome (16S) and inflammation.

Supplementary Figure 2. Spearman correlations between taxa and inflammatory markers at baseline. Co-occurrence network for genus-level summarized taxa. Genera (circles) were then correlated with levels of cytokines (triangles) at baseline. 100 permutations with a leave-one-out approach was used to identify and exclude correlations driven by outliers. Cytoscape (version 3.2.181) was used to visualize the network with a prefuse force directed layout and length of edges being 1 – rho for positive correlations and absolute [rho] for negative correlations. Nodes in close proximity are therefore highly positively correlated, while nodes further apart are highly negatively correlated. Taxa that were found to be most prominent in these correlations networks includes microbes characteristics of the upper airways, such as *Prevotella*, *Veillonella*, *Streptococcus*, *Propionibacterium*, and *Porphyromonas*.

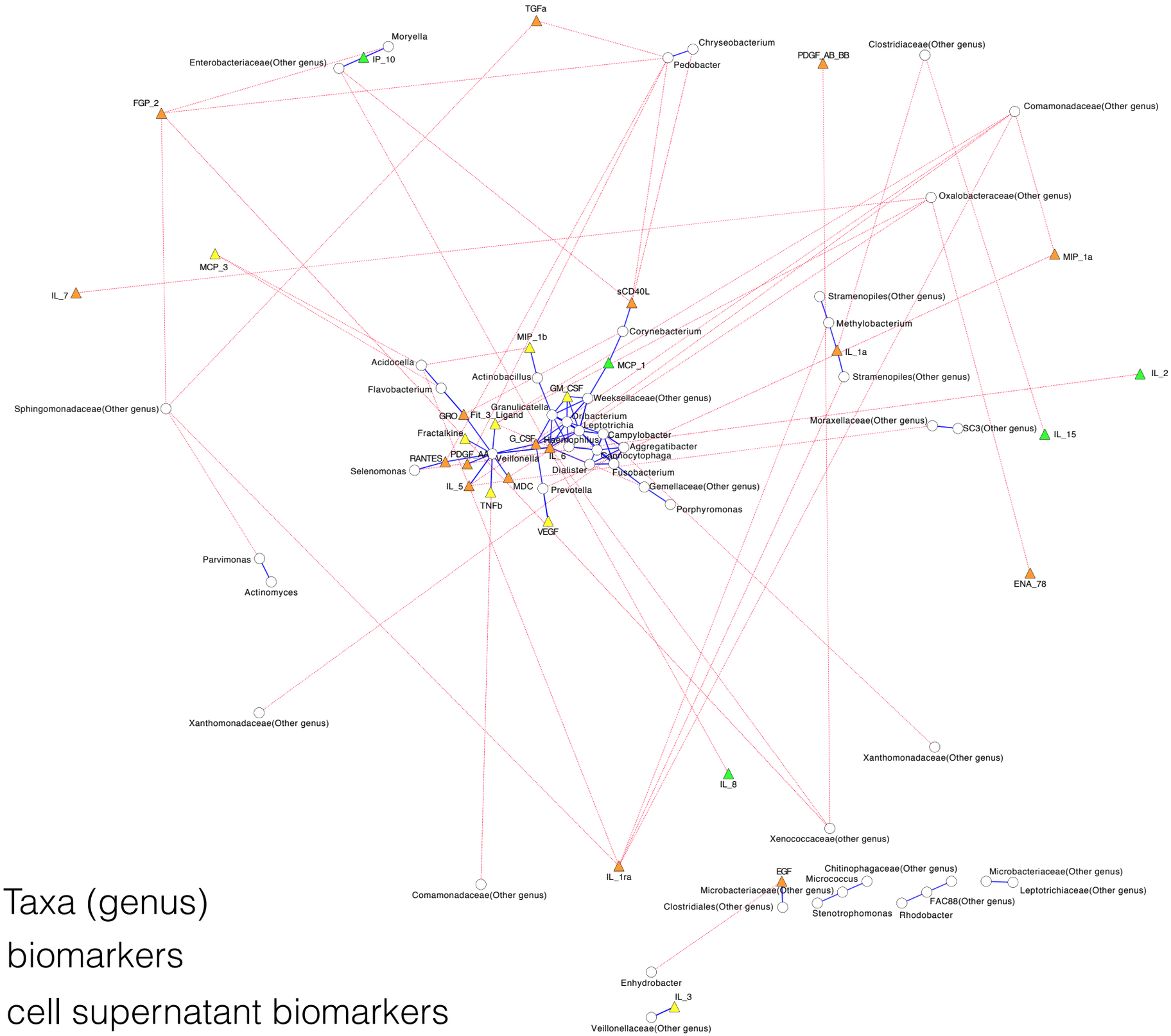
Supplementary Figure 3. Spearman correlations between taxa and inflammatory markers post-IFN-ɣ treatment. Similar to Supplementary figure

2, co-occurrence network for genus-level summarized taxa was constructed and genera (circles) were then correlated with levels of cytokines (triangles).

Supplementary Figure 1



Supplementary Figure 2



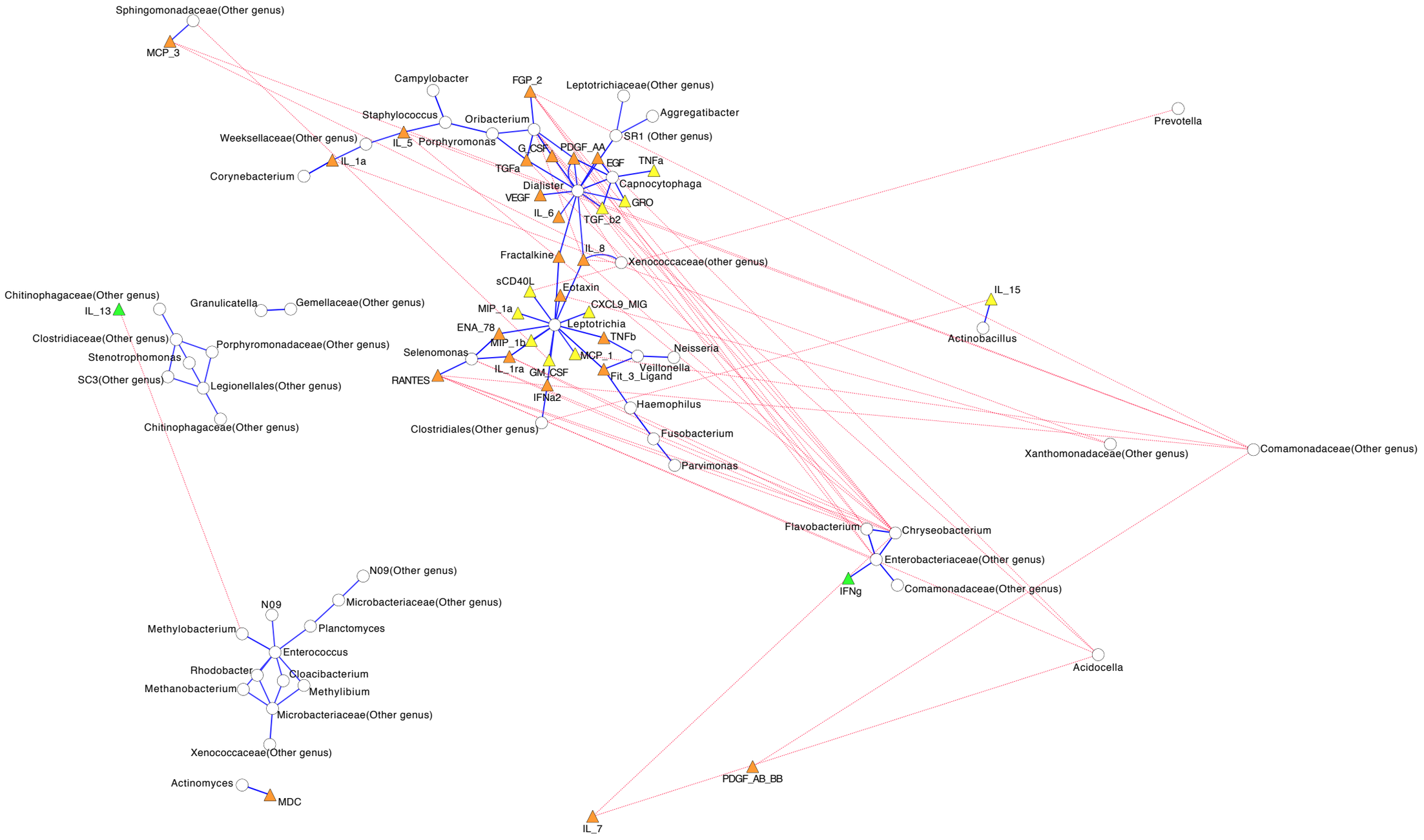
○ BAL Taxa (genus)

▲ BAL biomarkers

▲ BAL cell supernatant biomarkers

▲ Plasma biomarkers

Supplementary Figure 3



○ BAL Taxa (genus)

△ BAL biomarkers

△ BAL cell supernatant biomarkers

△ Plasma biomarkers

References

1. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. Isme J. 2012;6(8):1621-4.
2. 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(8):1159-61.
3. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335-6.
4. Segal LN, Clemente JC, Tsay JC, Koralov SB, Keller BC, Wu BG, et al. Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. Nat Microbiol. 2016;1:16031.
5. Dray SaD, A.B. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software. 2007;22(4):1-20.
6. Ramanan D, Bowcutt R, Lee SC, Tang MS, Kurtz ZD, Ding Y, et al. Helminth infection promotes colonization resistance via type 2 immunity. Science. 2016;352(6285):608-12.
7. Bastien P, Vinzi VE, Tenenhaus M. PLS generalised linear regression. Comput Stat Data An. 2005;48(1):17-46.
8. Liu H, Roeder K, Wasserman L. Stability Approach to Regularization Selection (StARS) for High Dimensional Graphical Models. Advances in neural information processing systems. 2010;24(2):1432-40.
9. Le Cao KA, Martin PG, Robert-Granie C, Besse P. Sparse canonical methods for biological data integration: application to a cross-platform study. BMC bioinformatics. 2009;10:34.
10. Csárdi G, Nepusz T. The igraph software package for complex network research. InterJournal. 2006.
11. Kolahian S, Fernandez IE, Eickelberg O, Hartl D. Immune Mechanisms in Pulmonary Fibrosis. Am J Respir Cell Mol Biol. 2016;55(3):309-22.

Supplementary Table 1: Effects of IFN-γ on local (BAL and BAL cell supernatant) and systemic (plasma) biomarkers

	BAL			BAL Cell Supernatant			Plasma		
	Baseline	**Post IFN**	**P value**	**Baseline**	**Post IFN**	**P value**	**Baseline**	**Post IFN**	**P value**
CXCL9-MIG	31600[11150-68200]	45150[19475-85450]	0.241	3335[1124-8658]	4005[2135-8935]	0.646	n.a	n.a	n.a
EGF	54.20[21.20-81.45]	41.35[25.90-78.45]	0.249	n.a	3.16[0.61-7.79]	0.180	19.00[4.24-28.70]	7.72[7.17-20.00]	0.917
ENA-78	809[549-1080]	891[341-1530]	0.575	2340[824-11125]	793[287-2335]	0.386	n.a	n.a	n.a
Eotaxin	18.55[14.55-58.20]	17.10[8.97-50.00]	0.075	9.92[4.88-15.50]	n.a	n.a	40.80[30.33-56.05]	39.25[29.30-83.57]	0.333
FGP-2	36.45[20.35-53.95]	33.70[19.88-39.35]	0.285	12.80[11.10-18.50]	n.a	n.a	49.60[40.98-54.23]	50.75[44.25-58.15]	0.445
Fit-3-Ligand	8.87[6.81-14.58]	10.30[7.43-13.73]	0.953	5.19[2.79-7.71]	2.32[1.17-6.33]	***0.017***	25.15[20.13-61.40]	23.80[17.63-68.53]	0.610
Fractalkine	42.15[19.43-93.78]	33.65[22.13-67.95]	0.386	13.90[8.70-34.63]	8.94[4.58-16.58]	0.241	58.40[32.05-97.78]	70.15[42.83-103.50]	0.333
G-CSF	566.00[108.73-676.50]	227.50[124.50-556.00]	0.508	10.70[3.66-515.50]	7.71[1.64-19.63]	0.508	4.07[2.59-7.92]	3.92[2.36-6.03]	0.333
GM-CSF	8.60[5.81-15.85]	8.67[7.57-15.45]	0.333	5.57[3.61-88.88]	3.82[1.52-10.08]	0.203	2.44[2.38-2.55]	2.26[1.69-3.10]	0.575
GRO	3585[2823-7303]	4400[3252-6257]	0.646	737[678-2310]	431[159-941]	0.139	284[160-363]	333[265-466]	0.333
IFNα2	8.49[3.93-21.45]	6.83[4.52-9.63]	0.214	3.33[1.63-6.38]	1.31[0.33-2.64]	***0.043***	15.15[2.12-22.93]	13.40[6.15-16.60]	0.721
IFNγ	n.a	294[161-1467]	***n.a***	n.a	28.05[15.68-46.08]	***n.a***	4.10[2.94-4.73]	3.43[2.74-6.00]	0.594
IL-13	7.38[3.77-12.33]	4.85[3.94-12.48]	0.767	1.04[0.43-2.21]	0.56[0.15-2.49]	0.500	1.84[1.20-4.74]	1.64[0.43-2.82]	0.484
IL-15	4.47[2.52-8.19]	5.58[4.03-11.65]	0.074	1.41[0.95-1.65]	1.08[0.97-1.39]	0.575	2.39[1.59-2.72]	2.25[1.72-2.85]	0.799
IL-17	0.99[0.66-1.96]	0.94[0.71-1.28]	0.169	1.63[0.30-2.75]	0.47[0.20-1.88]	0.080	2.91[1.24-4.39]	2.69[1.75-4.28]	0.445
IL-1ra	426.50[269.00-1155.00]	574[266-949]	0.799	13.90[4.64-53.55]	11.30[2.59-18.20]	0.917	5.55[2.58-7.61]	5.24[3.24-6.53]	0.594
IL-1α	56.35[43.90-87.00]	57.30[52.30-86.15]	0.878	39.00[4.77-60.40]	24.70[13.00-40.05]	0.959	17.05[7.27-23.10]	14.55[8.41-22.80]	0.799
IL-2	n.a	0.70[0.35-1.31]	n.a	0.69[0.35-2.03]	n.a	n.a	0.61[0.22-1.04]	0.33[0.29-0.88]	0.889
IL-3	n.a	n.a	n.a	11.20[6.80-15.75]	3.46[2.10-16.10]	0.715	n.a	n.a	n.a
IL-5	1.40[1.14-3.42]	1.36[0.96-2.62]	0.308	0.78[0.71-0.96]	0.73[0.66-0.81]	***0.047***	1.19[1.05-2.03]	1.21[1.07-2.38]	0.359
IL-6	43.30[21.60-163.00]	37.65[16.50-89.23]	0.333	80.65[28.73-825.75]	27.70[5.31-164.50]	0.445	3.88[1.76-7.01]	4.70[2.00-12.32]	0.203
IL-7	10.30[1.85-34.60]	12.80[7.55-21.85]	0.866	11.55[4.93-14.0]	4.44[4.05-7.69]	0.463	4.35[2.69-28.45]	n.a	n.a
IL-8	1630[328-2210]	1305[297-1910]	0.646	2690[2540-2945]	2115[1595-2902]	0.083	5.07[3.98-9.98]	5.98[3.45-9.94]	0.445
IP-10	7860[4608-8315]	8475[7857-8535]	0.093	293[68-3230]	1475[603-3490]	0.241	433[273-728]	485[257-696]	0.799
MCP-1	2540[1513-2998]	2500[1932-2925]	0.508	1880[443-2860]	962[270-2680]	0.386	135[97-156]	124[94-196]	0.114
MCP-3	34.25[17.63-66.55]	34.35[20.83-135.00]	0.169	56.80[20.60-133.50]	27.90[9.36-88.07]	0.575	13.60[5.77-19.90]	11.50[8.18-16.03]	0.721
MDC	253.50[209.75-609.00]	279[232-482]	0.878	919[413-2155]	226[45-1010]	0.169	903[579-1510]	921[576-1242]	0.878
MIP-1α	90.00[26.75-108.50]	48.40[33.20-81.80]	0.374	1775[824-18600]	953[209-9647]	0.515	26.00[14.68-53.70]	32.55[10.14-59.53]	0.484
MIP-1β	74.40[47.25-141.00]	78.60[61.93-107.86]	0.575	490[118-1333]	139[27-477]	0.086	17.65[14.28-23.85]	14.60[13.57-21.00]	0.799
PDGF-AA	415[173-631]	348[199-572]	0.646	18.00[10.75-31.50]	14.00[3.00-17.50]	0.058	n.a	n.a	n.a
PDGF-AB-BB	123[48-569]	223[77-442]	0.878	49.00[31.00-269.50]	156[23-192]	0.866	n.a	n.a	n.a
RANTES	597[155-1168]	972[243-1135]	0.767	160.50[44.75-269.25]	139.00[81.25-345.00]	0.799	n.a	n.a	n.a
sCD40L	176.00[86.13-402.25]	193.5[124.3-413.3]	0.646	25.30[11.25-35.60]	13.10[5.84-27.70]	0.441	2935[1670-4173]	3015[2162-4275]	0.878
TGFα	19.90[8.76-21.90]	14.90[9.27-19.90]	0.345	n.a	n.a	n.a	n.a	n.a	n.a

TGF-β2	n.a	n.a	n.a	9.00[5.75-19.25]	7.50[5.00-13.00]	0.149	n.a	n.a	n.a
TNFα	3.09[2.04-4.38]	3.38[2.95-4.64]	0.721	86.20[45.50-1056.25]	44.40[12.53-121.85]	0.386	3.02[2.45-3.39]	2.80[2.55-3.57]	0.919
TNFβ	1.15[0.71-1.67]	0.94[0.79-1.26]	***0.041***	0.71[0.49-1.62]	0.53[0.36-1.22]	0.069	2.56[1.56-3.00]	2.69[2.10-3.10]	0.203
VEGF	386.00[67.30-469.50]	218.5[65.2-543.0]	0.386	10.20[6.99-28.00]	n.a	n.a	17.55[12.28-74.10]	17.50[7.37-32.92]	0.779
siL-2Ra	n.a	14.60[9.02-19.63]	n.a	n.a	n.a	n.a	n.a	9.37[5.38-16.60]	n.a

Data expressed as median [IQR]. P value based on Wilcoxon Rank Sum Test.
n.a stands for levels not available due to being below the limit of detection.

Supplementary table 2: Taxa associated with various immunological phenotypes

Immunological phenotype	Taxa
Th1	
GM-CSF, IFN-g, IL-2, IL-12, TNF-b	*Prevotella*
	Parvimonas
	Rothia
	Porphyromonas
	Dialister
	Bulleidia
	Butyrivibrio
	Neisseriaceae
	Weeksellaceae
Th2	
IL-4, IL-5, IL-6, IL-9, IL-10, IL-13, GM-CSF	*Prevotella*
	Parvimonas
	Rothia
	Porphyromonas
	Chryseobacterium
	Dialister
	Bulleidia
	Pedobacter
	Butyrivibrio
	Weeksellaceae
	Procabacteriaceae_4458997
Th17	
IL-1b, IL-12, IL-17, IL-6, Fractalkine	*Parvimonas*
	Staphylococcus
	Porphyromonas
	Dialister
	Butyrivibrio
	Weeksellaceae
IPF	
IL-1b, IL-17, IL-13, TGF-b, MCP-1, IFN-g, PDGF_AA, PDGF_AB_BB	*Prevotella*
	Parvimonas
	Parvimonas
	Staphylococcus
	Porphyromonas
	Enterobacteriaceae_782953
	Dialister
	Bulleidia
	Pedobacter
	Butyrivibrio

Aggregatibacter

Weeksellaceae

Procabacteriaceae_4458997

Taxa ranked ordered based on relative abundance (high to low).