Lung-enriched organisms and aberrant bacterial and fungal respiratory microbiota following lung transplant

Emily S. Charlson, Joshua M. Diamond, Kyle Bittinger, Ayannah S. Fitzgerald, Anjana Yadav, Andrew R. Haas, Frederic D. Bushman* and Ronald G. Collman*

* joint corresponding authors

Online Data and Methods Supplement

Supplemental Methodological Information:

Fungal ITS Sequence Analysis

ITS reads were assigned taxonomy with the BROCC classifier (*Dollive* et al, submitted). BROCC is a consensus-based assigner designed for taxonomic classification of ITS tag sequencing reads, which present a particular challenge to existing classifiers. The program takes as input a file of Blast hits against the nr database from NCBI. Each OTU was represented by the most abundant sequence for purposes of classification. BROCC forms a consensus of assigned Blast hits that exceed a specified threshold of percent identity, after filtering hits for uneven coverage and generic classifications in the database.

Additional Statistical Methods.

For the Procrustes analysis (1, 2), the goodness of fit (M^2 value) was measured by summing over the residuals, and statistical significance assessed by the Monte-Carlo label permutation method, which showed an optimal fit when samples from each individual were aligned together ($p<0.0001$, $M^2=0.126/0.475$ unweighted/weighted UniFrac, $M^2=0.034/0.0.366$ binary/abundance Jaccard).

To detect OTUs enriched in BAL samples relative to OW, we use a Dirichlet-multinomial distribution. A maximum-likelihood estimate of the parameters is determined by numerical optimization. For each OTU, the Dirichlet-multinomial

parameter estimates are used to construct a marginal beta-binomial distribution (3). The marginal form represents the distribution of OTU counts under the null hypothesis that a particular OTU is not enriched in the BAL sample. The p-values reported for BAL enrichment are generated from a one-sided test against the null distribution. This method is conservative, because genuine BAL-enriched OTUs will artificially increase the estimated level of over-dispersion, and thus raise the bar for detection.

References

1. Gower JC. Generalized procrustes analysis. *Psychometrika* 1975;40.
2. Hurley JR, and Cattell RB. The procrustes program: Producing direct rotation to test a hypothesized factor structure. *Behav Sci* 1962;7.
3. Weir BS, Hill WG. Estimating f-statistics. *Annu Rev Genet* 2002;36:721-750.

Supplementary Tables:

Table E1. Full clinical and microbiological data.

BAL samples were obtained at the indicated time following transplant for the reasons indicated. BAL bacterial culture results are indicated, along with the dominant taxa on 16S profiling. If taxa matching culture results were not the dominant species, they are indicated in parentheses. The presence or absence of "normal respiratory tract flora" in culture is also indicated. BAL fungal culture and sequencing results are similarly shown, along with quantification of BAL ITS and OW ITS amplification products. Pre-transplant lung diseases are indicated as: IPF, idiopathic pulmonary fibrosis; ILD, interstitial lung disease associated with collagen vascular disease; CF, cystic fibrosis; A1AT, alpha-1 antitrypsin deficiency; PCD, primary cilliary dyskinesia; CHD, congenital heart disease; PHTN, pulmonary hypertension. Immunosuppression regimens are indicated as: MMF, Mmycophenolate mofetil; Tac, tacrolimus; Aza, azathioprine; Pred, prednisone; CSA, cyclosporin A. Antimicrobial regimens are indicated as: TMP/S, trimethorpim/sulfamethoxazole; Azithro, azithromycin; Valgan, valgancyclovir; Valcyc, valcyclovir; Vori, voriconazole; Nystatin, nystatin oral swish.

Table E2. 16S OTU table

This table is available in Excel format and is accessible from this issue's table of contents online at www.atsjournals.org.

Table E3. ITS OTU table

Supplementary Figures:

Figure E1. 16S heatmap with extraction controls and prewash samples

Figure E2. ITS heatmap with extraction controls and prewash samples

Figure E3. Within-subject BAL-OW distance by transplant group. Weighted UniFrac distances between BAL and OW were calculated for each sample pair. Mean group distances were then calculated for subjects transplanted for suppurative or non-suppurative indications (A), as well as suppurative, COPD/emphysema (n=7) or ILD/IPF indications (B). Differences between groups do not reach statistical significance (Wilcoxon rank sum).

Figure E4. Complete set of Per-Patient Reports.