Circulating RNA transcripts identify therapeutic response in cystic fibrosis lung disease

Milene T. Saavedra, Grant J. Hughes, Linda A. Sanders, Michelle Carr, David M. Rodman, Chris D. Coldren, Mark W. Geraci, Scott D. Sagel, Frank J. Accurso, James West, Jerry A. Nick

Online Data Supplement

Material and methods

Patient recruitment: Subjects >18 years of age with CF (based on sweat chloride testing and genotype) were enrolled at the time of admission or home IV antibiotic initiation for a clinically diagnosed pulmonary exacerbation at a large Cystic Fibrosis Foundation accredited adult CF clinic, following obtainment of informed consent from the study subject. All study participants were monitored by the National Jewish Health and University of Colorado Institutional Review Boards. Patients identified for enrollment met CF Foundation Clinical Practice Guidelines of at least 3 of 11 criteria for an acute pulmonary exacerbation(19). All patients were treated with at least 2 antibiotics targeting their specific bacterial pathogens for a minimum of 2 weeks and a maximum of three weeks. The study design utilized within subject comparisons, such that each study subject served as their own control, following treatment with antibiotics. Blood was drawn at the initiation ($\pm$ 2 days) and the completion ($\pm$ 1 week) of intravenous antibiotic therapy. At each timepoint, the following were collected or measured: 1) blood for PBMC isolation 2) sputum for microbiologic analysis 3) simple spirometry for FEV_1 determination according to American Thoracic Society guidelines and 4) plasma isolation from the blood preparation for cytokine measurements. In addition, array data from the PBMCs from 8 healthy volunteers and 6 stable CF patients was utilized to compare transcript abundance to that seen in patients undergoing exacerbation therapy. Following the initial study, a validation study was performed on 14 adult CF patients suffering from an acute pulmonary exacerbation. Patients were identified for enrollment, treated with antibiotics, and subjected to blood, microbiology, and spirometry analysis exactly as described above.

PBMC isolation: Four ml of peripheral blood was collected into sodium citrate tubes (BD Vacutainer® CPT™, BD Biosciences, Franklin Lakes, NJ) utilizing a Ficoll Hypaque density gradient. PBMC's were isolated via density gradient centrifugation (1650 RCF, 30 minutes, 18 °C). Furthermore, cytospins with H&E staining verified that all cells from which RNA was extracted were mononuclear and fewer than 1% neutrophils were identified from the interphase. The RNA was then immediately isolated by Trizol method, then purified using the Rneasy Mini Kit (Qiagen) according to manufacturer's protocol. cDNA was reverse transcribed from total RNA using the Superscript II Reverse Transcriptase™ Kit (Invitrogen, Carlsbad, CA). Complete blood counts were done to quantify cell numbers. Additionally, cell counts were done after each isolation in order to insure that differences in transcript abundance were not due to differences in cell counts, PBMC's were counted at the time of each isolation, and absolute numbers between timepoints evaluated by paired t tests.

Microarray hybridization and data analysis: PBMC RNA isolated from the development cohort, containing 10 patients representing 10 pulmonary exacerbations, was utilized for microarray analysis, in order to identify PBMC transcriptional changes before and after antibiotic therapy. All raw microarray data is available on the NCBI Gene Expression Omnibus Database. Prior to microarray, all RNA was evaluated with an Agilent 2100 Bioanalyzer to confirm high grade RNA quality. PBMC gene expression analysis was performed with Hu133 Plus 2.0 gene chips. Samples were prepared for Affymetrix arrays

using 2.5 μg of total RNA. First and second strand complimentary DNA was synthesized using standard techniques(44). Biotin-labeled antisense complimentary RNA was produced by an in vitro transcription reaction. Target hybridization, washing, staining, and scanning probe arrays were done following manufacturer's protocol as described in the Affymetrix GeneChip Expression Analysis Manual. Affymetrix CEL files were loaded into dChip 2005 array analysis software (Li C, Wong WH 2001), normalized to median brightness, and expression modeled using the perfect match/mismatch (PM/MM) algorithm. Beginning with microarrays containing over 35,000 sequences from 10 patients before and after antibiotic treatment, we focused on sequences assigned to known genes with a minimum expression threshold (Affymetrix "present call"). The study design utilized "within subject comparisons", such that each study subject served as their own control. From these genes, we identified differentially expressed genes in pairwise comparisons between pre- and post- treatment groups, using a non-parametric Wilcoxon signed rank test, with a minimum of 1.4-fold change. Analysis of microarray data using the dChip analysis program yielded 32 candidate genes with both significant expression and significant change ($p<0.05$) between pre- and post- treatment (**Table E1**). Gene ontology analysis was utilized to determine biologic plausibility and narrowed the gene list to 19 genes. The false discovery rate was computed from the gene-wise p values according to Benjamini-Hochberg(45). In addition, high quality PBMC microarray data from 8 normal and 6 stable CF matched controls was utilized to assess baseline transcripts of the gene signature. Relationships between CF pre- and post-antibiotic genes, normal and stable CF PBMC controls were evaluated with ANOVA, with post-hoc

testing determined by Fisher's PLSD (StatView, SAS Institute Inc., Cary, NC). Statistical significance was assigned for p-values ≤ 0.05.

Quantitative RT-PCR validation: Microarray expression data from 19 genes in the development cohort were secondarily confirmed by real-time polymerase chain reaction (RT-PCR), utilizing all sufficient remaining RNA for analysis (6 of 10 total patients had sufficient RNA from both pre- and post- therapy PBMC RNA samples). First strand cDNA was made from 1 μg total RNA using a QuantiTect Reverse Transcription Kit (Qiagen, Valencia, CA). Quantitative real-time PCR was performed using a total reaction volume of 25 μl, containing 5 μl of diluted cDNA, 10 μl H_20, 9.94 μl of the fluorescent indicator Sybrgreen® and 0.03 μl of each nucleotide primer (250 μM). PCR was carried out in a 7300 Real Time PCR System (Applied Biosystems, Foster City, CA), using 40 cycles of 95° C for 15 seconds, followed by 60° C for 1 minute, with a 10 minute 95° C initial soak. Each measurement was made in triplicate and expressed relative to the detection of the housekeeping gene, hypoxanthine guanine phosphoribosyl transferase (HPRT). For quantitative RT-PCR, statistics were performed in StatView (SAS, Cary, NC), utilizing paired *t*-tests, and significance at $P < 0.05$.

Plasma cytokine measurements: Plasma collected at the time of PBMC isolation was processed at The Children's Hospital, University of Colorado GCRC Core laboratory. Multiple cytokine measurements were done using the Fluorokine® MAP Cytokine assay (R&D systems, Minneapolis, MN) on the Luminex 100 system (a dual laser, flow-based sorting and detection platform) to quantify the following cytokines at each blood draw:

IL-1α, IL-1β, IL-1RA, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-17, TNFα, IFNγ, G-CSF, GM-CSF, MIP-1β, MCP-1, VEGF, RANTES, and IL-13. All analyses were performed according to manufacturer's protocols. The mean minimal detectable doses for cytokines are as follows: IL-1α (0.39 pg/ml), IL-1β, (0.27 pg/ml), IL-1RA (2.06 pg/ml), IL-2 (0.89 pg/ml), IL-4, (1.75 pg/ml), IL-5 (0.33 pg/ml), IL-6 (0.36 pg/ml), IL-8 (0.39 pg/ml), IL-10 (0.13 pg/ml), IL-17 (0.39 pg/ml), TNFα (0.47 pg/ml), IFNγ (0.31 pg/ml), G-CSF (0.57 pg/ml), GM-CSF(1.05 pg/ml), MIP-1β (2.12 pg/ml), MCP-1 (0.95 pg/ml), VEGF (0.81 pg/ml), RANTES (1.08 pg/ml), and IL-13 (6.0 pg/ml). Highly sensitive CRP assays were measured utilizing nephelometry of the Dade Behring BNII. Univariate analysis with paired *t*-tests was performed to compare pre- and post- antibiotic values. Correlations between C-reactive protein (CRP), FEV_1, circulating leukocyte counts and gene expression changes pre- and post- antibiotic therapy were assessed by Spearman correlation coefficients.

Endpoints: The primary endpoints of the study were lung function (FEV_1) and gene transcript abundance changes after treatment of an acute pulmonary exacerbation. All patients were evaluated with testing as described above as well as clinical evaluation at the onset and completion of therapy. Patients who responded to therapy by manifesting improvement in FEV_1 % predicted at the completion of antibiotic therapy and not requiring rehospitalization within 1 week with a diagnosis of pulmonary exacerbation were analyzed as described below such that changes in FEV_1% predicted were regressed with transcript changes. Pulmonary function testing was performed in compliance with American Thoracic Society standards [E4].

Statistical methods: Comparison of gene expression values pre- and post- antibiotic therapy were initially performed by paired univariate analyses: *t* test and the Wilcoxon signed rank test. Utilizing Generalized Estimating Equations, multivariate logistic regression models were constructed predicting resolution of acute pulmonary exacerbations as a function of FEV_1% predicted, and the discriminative value of combinations of genes identified by the development cohort (SAS, SAS Institute Inc., Cary, NC). An unstructured correlation structure between time points was utilized, and a stepwise selection procedure chose the most significant combination of FEV_1 improvement and transcript changes. FEV_1% predicted was forced into modeling due to its known clinical reliability, as a standard assessment for response to therapy. To identify the most frugal combination of predictors to predict resolution of inflammation, we required a p value <0.05 for entry into the model, which also allowed for determination of the unique contribution of the genes over and above FEV_1 alone. Finally, receiver operating characteristic (ROC) analyses reflected the overall diagnostic value of the gene markers, in terms of enhanced sensitivity and specificity over FEV_1 alone.

References

E1. CF Foundation, Clinical Practice Guidelines for Cystic Fibrosis 1997.

E2. Golpon, H. A., C. D. Coldren, M. R. Zamora, G. P. Cosgrove, M. D. Moore, R. M. Tuder, M. W. Geraci, and N. F. Voelkel. 2004. Emphysema lung tissue gene expression profiling. *Am J Respir Cell Mol Biol* 31(6):595-600.

E3. Reiner, A., D. Yekutieli, and Y. Benjamini, Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics, 2003. 19(3): p. 368-75.

E4. Pellegrino, R., Viegi G., Brusasco V., Crapo RO., Burgos F., Casaburi R., Coates A., van der Grinten CP., Gustafsson P., Hankinson J., Jensen R., Johnson D.C., Macintyre N., McKay R., Miller M.R., Navajas D., Pedersen OF., Wanger J., Interpretative strategies for lung function tests. Eur Respir J, 2005. 26(5): p. 948-68.

Table E1: Transcripts changed in PBMC's pre- and post- antibiotic therapy. Genes listed in this table meet criteria for differential expression in a paired t-test between log transformed values of pre- and post- antibiotic circulating samples (n=10) (p<0.05).

Gene name	probe set	fold change	p value
ADAM metallopeptidase domain 9 (meltrin gamma)	202381_at	-2.5	0.004
Fc fragment of IgG, high affinity Ia, receptor (CD64)	214511_x_at	-2.4	0.004
Kruppel-like factor 9	203542_s_at	-2.4	0.001
Charcot-Leyden crystal protein	206207_at	-2.3	0.018
histocompatibility (minor) 13	232209_x_at	-2.1	0.014
chondroitin sulfate proteoglycan 2 (versican)	204619_s_at	-1.9	0.001
tweety homolog 3 (Drosophila)	224674_at	-1.8	0.019
hepatocellular carcinoma-associated antigen 112	218345_at	-1.8	0.004
Potassium inwardly-rectifying channel, subfamily J, member 2	231513_at	-1.8	0.014
alanyl (membrane) aminopeptidase (CD13, p150)	202888_s_at	-1.7	0.009
CD36 antigen (thrombospondin receptor)	228766_at	-1.7	0.001
osteoclast-associated receptor	1554503_a_at	-1.7	0.007
C-type lectin domain family 4, member D	1552772_at	-1.7	0.006
plexin D1	212235_at	-1.7	0.005
Trypsin domain containing 1	231422_x_at	-1.7	0.003
CD163 antigen	203645_s_at	-1.6	0.003
ADP-ribosylation factor-like 11	1552691_at	-1.6	0.006
toll-like receptor 2	204924_at	-1.6	0.003
ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4	230836_at	-1.5	0.003
Regulatory factor X-associated ankyrin-containing protein	1560034_a_at	-1.5	0.001
glucosidase, alpha; acid	202812_at	-1.5	0.011
J-type co-chaperone HSC20	223647_x_at	-1.5	0.026
Dmx-like 2	212820_at	-1.5	0.010
CCAAT/enhancer binding protein (C/EBP), alpha	204039_at	-1.4	0.002
BRI3 binding protein	225716_at	-1.4	0.002
heparanase	219403_s_at	-1.4	0.017
Splicing factor, arginine/serine-rich 3	232392_at	1.4	0.057
interleukin 32	203828_s_at	1.4	0.033
Formin binding protein 4	239469_at	1.5	0.000
villin 2 (ezrin)	208621_s_at	1.6	0.005
dehydrogenase/reductase (SDR family) member 3	202481_at	1.9	0.011
chemokine (C-C motif) receptor 5	206991_s_at	2.4	0.007