Gene Expression Profiling Identifies MMP-12 and ADAMDEC1 as Potential Pathogenic Mediators of Pulmonary Sarcoidosis.

[*1]Elliott D. Crouser, M.D., *[5]Daniel A. Culver, D.O., [4]Kenneth S. Knox, M.D., [1]Mark W. Julian, M.S., [1]Guohong Shao, M.D., [5]Susamma Abraham, B.S., [2]Sandya Liyanarachchi, M.S., [1]Jennifer E. Macre, B.S., [1]Mark D. Wewers, M.D., [1]Mikhail A. Gavrilin, Ph.D., [3]Patrick Ross, M.D., [3]Abbas Abbas, M.D., [6]Charis Eng, M.D., Ph.D.

*These authors contributed equally to the completion of this manuscript.

On-Line Data Supplement

Methods:

Isolation and Validation of High Quality RNA for Gene Chip Analysis and PCR

Frozen lung tissue or cells derived from bronchoalveolar lavage fluid (BALF) were maintained at -80°C until the day before total RNA isolation at which time the sample was soaked overnight in RNA*later*-ICE (Ambion, Applied Biosystems; Foster City, CA) at -20°C. The samples were then removed from RNA*later*-ICE, and total RNA was isolated using TRIzol reagent (Invitrogen Corp.; Carlsbad, CA) according to the manufacturer's protocol. Briefly, lung tissue samples were intermittently homogenized in 1 ml of TRIzol while on ice using a mini-homogenizer with a sterile RNase-free pestle (Kontes® Glass Company; Vineland, NJ) for up to one minute. Phase separation of the RNA was carried out following the addition of 200 µl of chloroform, vigorous mixing for 15 sec and centrifugation at 12,000 x *g* for 15 min at 4°C. The top RNA-containing phase was carefully transferred to a new tube, mixed with 500 µl of isopropanol, vigorously mixed for 15 sec and centrifuged again at 12,000 x *g* for 10 min at 4°C. Following careful removal of the supernatant, 75% ethanol was added to the RNA pellet which was vortexed and centrifuged at 7500 x *g* for 5 min at 4°C. The ethanol was removed, and the pellet was allowed to air dry. The RNA pellet was then dissolved in RNase-free water by gentle pipetting and incubating at 55-60°C for 5 minutes.

The RNA was purified using the Qiagen RNeasy® Mini Kit (Qiagen Inc.; Valencia, CA). The integrity of these total RNA samples was then assessed qualitatively on an Agilent 2100 Bioanalyzer (Agilent Technologies; Santa Clara, CA). This is a capillary electrophoresis system using fluorescence to characterize size distribution. Finally, total RNA quality was verified by the presence of two discrete electropherogram peaks corresponding to the 28S and 18S rRNA at a ratio approaching 2:1. Sample RNA quality was further evaluated for potential degradation in

each GeneChip (HG-U133 Plus 2, Affymetrix, Inc.; Santa Clara, CA) array used in the Affymetrix analyses. For a given array probeset, the probe intensities will be elevated at the 3' expression end resulting in an increased ratio of 3' expression to 5' expression, indicative of degradation (E1). This ratio corresponds to the slope of the relationship in an RNA degradation plot (Figure E1). Although some degradation is typically seen, it is most important that the arrays demonstrate similar slopes, indicating comparable sample degradation, in order to make valid comparisons within genes across arrays (E1).

Semi-Quantitative RT-PCR and Real-Time PCR

In order to confirm the results obtained from gene array analysis, a two-step semi-quantitative RT (reverse transcription)-PCR method (E2) was used to determine select gene expression in lung tissue samples from an additional set of 11 sarcoidosis and 11 disease-free control patients. Tissue RNA was isolated from tissue samples using TRIzol reagent as described above. One μg of total RNA was reverse-transcribed to cDNA in the presence of 1X RT buffer, 5.5 mM $MgCl_2$, 0.5 mM dNTP, 2.5 μM random hexamers, 1 U RNase inhibitor and 62.5 U MultiScribe™ reverse transcriptase (Applied Biosystems) in a final reaction volume of 50 μl using the TaqMan® reverse transcription kit (Applied Biosystems). The RT reactions were carried out at 25°C for 10 min, 48°C for 30 min, then 95°C for 5 min. To determine the optimal number of cycles needed for PCR, the cDNA for each select gene was amplified by PCR at 20, 25, 28, 30, 32 and 35 cycles then loaded onto an ethidium bromide-stained 1.5% agarose gel for electrophoretic analysis. The appropriate number of cycles for each select gene was chosen based upon the intensity of the PCR product band to ensure that the amplification level was within the exponential range and had not reached saturation or the plateau range (Figure E2).

The primers used for the select human genes (*ADAMDEC1*, *MMP12*, *HLA-DRB1*, *STAT1*, *CXCL9*, *CCL5*, *IL-7*, *IL-15*, *CAP1*, and *GAPDH*), along with their corresponding optimal cycle number and PCR product size, are shown in Table E1. PCR was performed in a 20 µl mixture containing 2 µl of cDNA, 1.5 mM $MgCl_2$, 0.2 mM dNTP, 0.5 µM of each primer and 1 U of native *Taq* DNA polymerase (Invitrogen Corp.). The PCR program profile for each select gene began with 95°C for 3 min, followed by optimally chosen cycles of 95°C for 30 sec, 56°C for 30 sec, 72°C for 30 sec and 72°C for 5 min. 5-10 µl of the PCR product corresponding to each gene for each sample were then analyzed by electrophoresis on 1.5% agarose gels. DNA band intensity for each gene was measured using Quantity One software (Bio-Rad Laboratories; Hercules, CA) and normalized by the corresponding DNA band intensity of the housekeeping genes, *CAP1* and *GAPDH* (Figure E3).

Gene expression for select human genes was further determined in cells derived from BALF using real-time PCR. One µg of total RNA was reverse-transcribed to cDNA by MultiScribe™ reverse transcriptase with random hexamers in a 50 µl reaction volume. The primers used for the human genes *MMP12*, *ADAMDEC1* and *GAPDH* were as follows:

MMP12 (70 bp)	Sense	5'-TGCTGATGACATACGTGGCA-3'
	Anti-Sense	5'-AGGATTTGGCAAGCGTTGG-3'
ADAMDEC1 (86 bp)	Sense	5'-AGTGGCCTTGGTAGGTATGG-3'
	Anti-Sense	5'-GGAAGTTGTCAAACGTGGTG-3'
GAPDH (123 bp)	Sense	5'-TGGTATCGTGGAAGGACTCA-3'
	Anti-Sense	5'- GCAGGGATGATGTTCTGGA-3'

All of the primers spanned an intron, ensuring that gDNA contamination was not present. Primer pairs were validated by real-time PCR and high-resolution gel electrophoresis to have a single band of desired size that was free of primer dimers. PCR was performed in a 20 µl mixture containing 1 µl of cDNA, 5 pmol of each given primer, and 10 µl of SYBR® Green PCR master mix (Applied Biosystems). Amplification and detection were done with The

Applied Biosystems 7900HT Fast Real-Time PCR System with the profile of 50°C for 2 min, 95°C for 10 min and 40 cycles of 95°C for 15 sec and 60°C for 1 min. Relative copy numbers and expression ratios for *MMP12* and *ADAMDEC1* were normalized to the expression of the housekeeping gene, *GAPDH*.

Cytokine Analysis

Cytokine expression was determined in lung tissue homogenates by multiplex quantification. Briefly, frozen lung tissue samples (~0.5 g) were immersed in ice cold homogenization buffer [230 mM mannitol, 70 mM sucrose, 3 mM HEPES (pH 7.4), 1 mM EGTA], minced and repeatedly rinsed. Tissue samples were added to 1.0 mm Zirconia/Silica beads (v/v) and 500 µl of cold homogenization buffer and then lysed using a Mini-Bead Beater cell disrupter (Biospec Products, Inc.; Bartlesville, OK) at 5000 rpm for 20 sec, repeated four times at 5 min intervals. Tissue homogenates were subjected to high-speed centrifugation (10,000 x *g* for 5 min at 4°C), and the supernatant protein concentrations were then determined spectrophotometrically using the Bradford assay. The concentration of each cytokine was analyzed from a 50 µl sample using a custom-made Bio-Plex Human Cytokine panel that included IL-1β, IL-1ra, IL-2, IL-6, IL-7, IL-8, IL-10, IL-12 (p40), IL-12 (p70), IL-13, IL-15, GM-CSF, M-CSF, MIP-1, and TNFα according to the manufacturer's recommendations using the Bio-Plex 200 Analysis System (Bio-Rad Laboratories). Concentrations determined from simultaneously run standards were normalized per mg tissue homogenate protein.

References:

E1. Bolstad BM, Collin F, Brettschneider J, Simpson K, Cope L, Irizarry RA, Speed TP. Quality assessment of Affymetrix GeneChip data. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S, editors. Bioinformatics and computational biology solutions using R and bioconductor (Statistics for biology and health), 1st ed. New York, NY: Springer Verlag; 2005. p. 33-47.

E2. Marone M, Mozzetti S, De Ritis D, Pierelli L, Scambia G. Semiquantitative RT-PCR analysis to assess the expression levels of multiple transcripts from the same sample. *Biol Proced Online* 2001; 3:19-25.

Figure Legends:

Figure E1: RNA degradation plot showing the extent to which the RNA, isolated and purified from lung tissue samples and used for gene array analysis, was degraded. Each line represents a single amplified RNA array wherein the slope corresponds to the ratio of the 3' expression to the 5' expression for the entire probeset in that given array and denotes the degree of degradation. Note that the slopes are quite similar indicating comparable degradation among the samples, thus enabling comparisons within genes across arrays.

Figure E2: Photomicrographs of high resolution gels showing PCR amplification of select human genes wherein the single PCR product band gradually intensified toward maximal with increasing cycle number employed. The optimal cycle number to use for each gene was chosen based upon the corresponding band intensity, which reflected amplification in the exponential (sub-plateau) range, and was then subsequently applied in the final RT-PCR involving all samples.

Figure E3: Photomicrographs of high resolution gels showing the single PCR product band following RT-PCR amplification of select human genes for all (11 sarcoidosis and 11 disease-free control) samples. Subsequent band intensity analysis and normalization to that of the housekeeping genes (*CAP1* and *GAPDH*) demonstrated significantly enhanced select gene expression in the lung tissues from sarcoidosis patients relative to controls.

Table E1: Oligonucleotides Used for Semi-Quantitative RT-PCR.

Gene	Primers (5' → 3')		Optimum Cycle Number	PCR Product Size (bp)
ADAMDEC1	Sense	CTCAGCTTCTCAGCGGGATT	32	150
	Antisense	GGCCCAGCTCATGTGACAT		
MMP12	Sense	GACTACGCAATCCGGAAAGC	28	150
	Antisense	GATTCCACCTTTGCCATCAAA		
HLA-DRB1	Sense	CCACAACCTCCTGGTCTGTTC	25	150
	Antisense	AGCATCACCAGGGTCTGGAA		
STAT1	Sense	CATTCAGAGCTCGTTTGTGG	25	251
	Antisense	TGGACTCCTCCATGTTCATC		
CXCL9	Sense	CAGTAGTGAGAAAGGGTCGCTGTT	25	120
	Antisense	TGATTTCAATTTTCTCGCAGGAA		
CCL5	Sense	CATATTCCTCGGACACCACA	28	121
	Antisense	ACAAAGACGACTGCTGGGTT		
IL-7	Sense	CTGGCCAGGTTAAAGGAAGA	28	180
	Antisense	AGTGTTCTTTAGTGCCCATCAA		
IL-15	Sense	GCTGGCATTCATGTCTTCAT	28	158
	Antisense	GGGTGAACATCACTTTCCGT		
CAP1	Sense	CTCCATATGTGCAGGCATTT	25	126
	Antisense	CCAACTTCAAACCTGTGTGG		
GAPDH	Sense	ATCATCCCTGCCTCTACTGG	25	230
	Antisense	ACCACCTGGTGCTCAGTGTA		

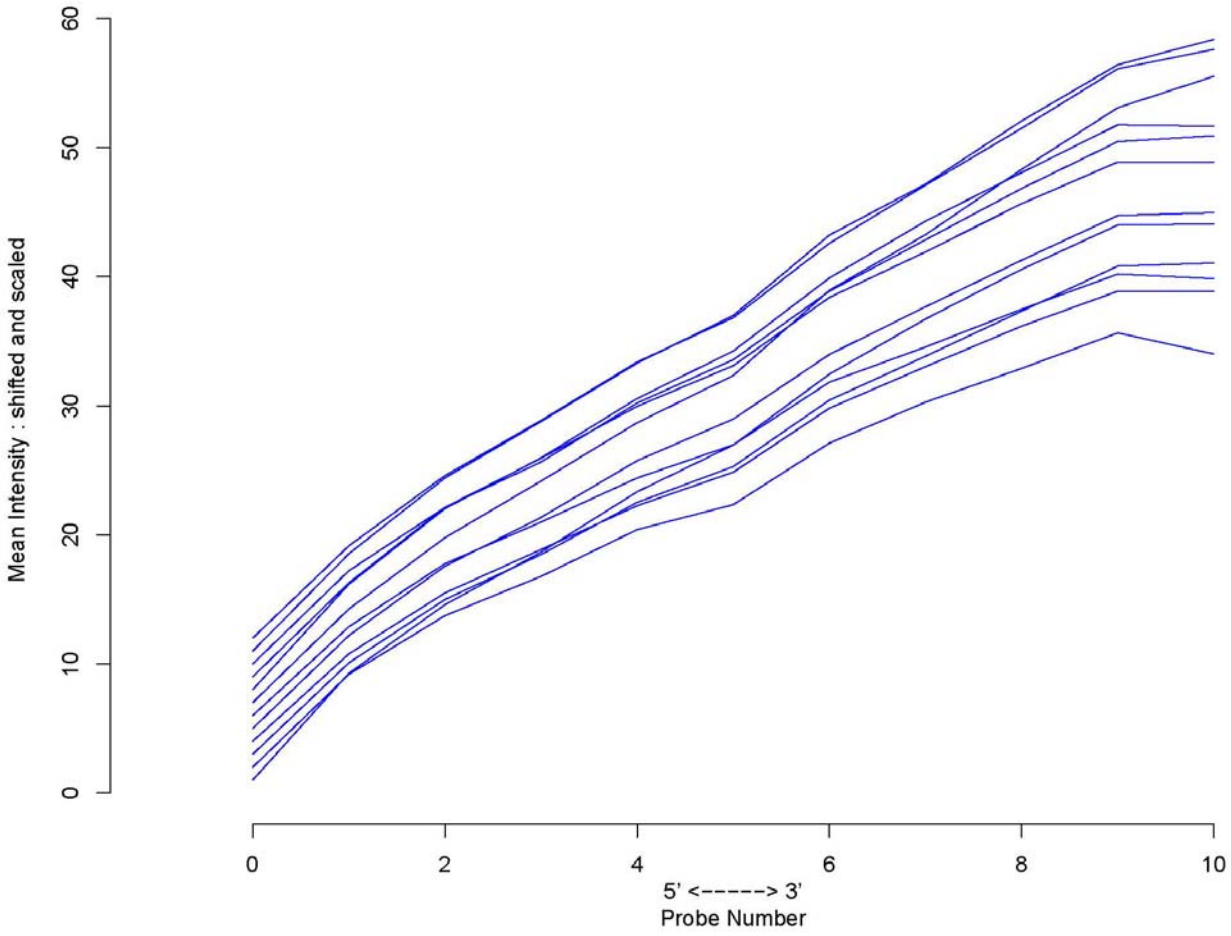
Mean Intensity : shifted and scaled
0
10
20
30
40
50
60
0
2
4
6
8
10
5' <-----> 3'
Probe Number

Figure E1

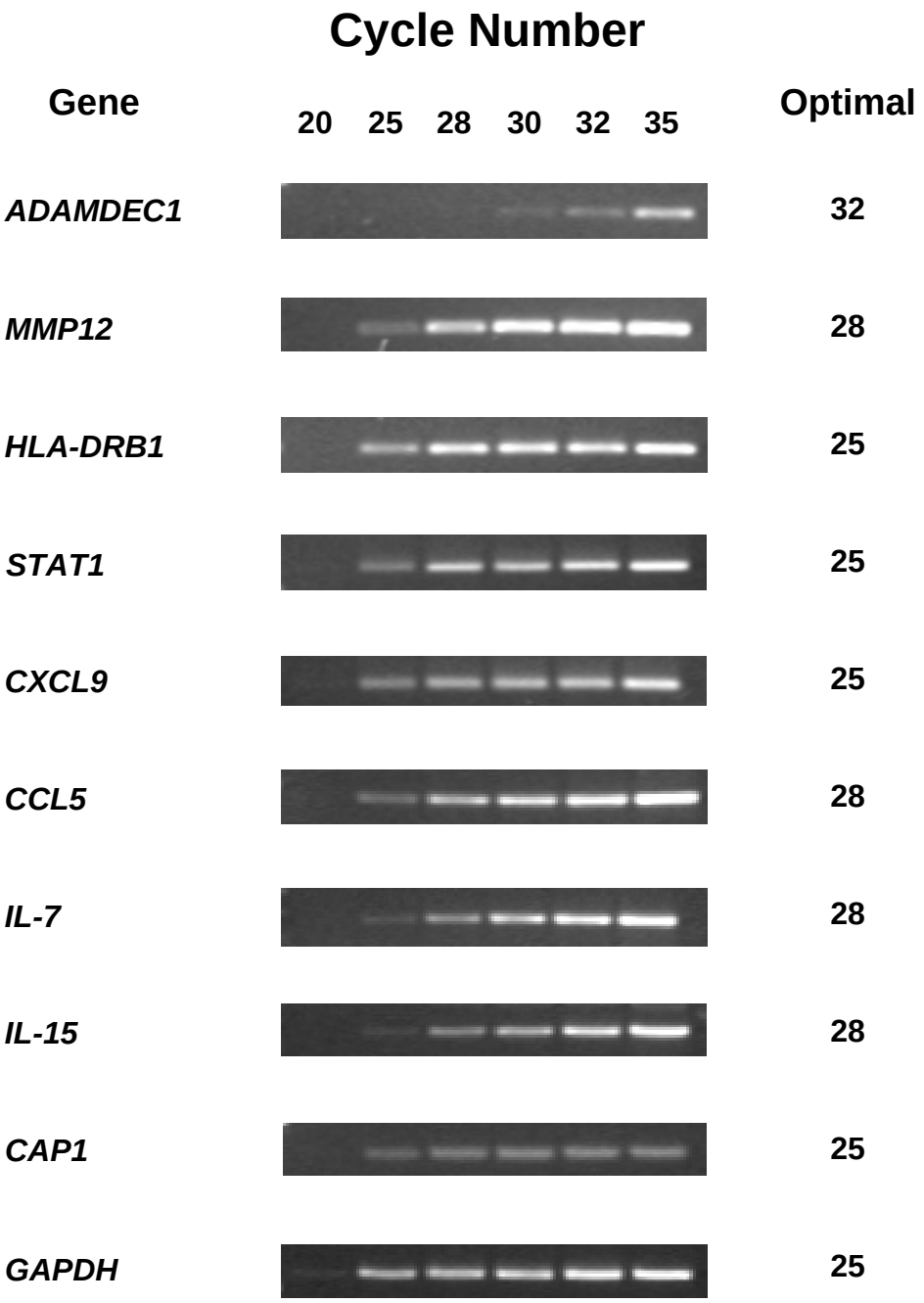
Cycle Number
Gene
20 25 28 30 32 35
Optimal
ADAMDEC1
32
MMP12
28
HLA-DRB1
25
STAT1
25
CXCL9
25
CCL5
28
IL-7
28
IL-15
28
CAP1
25
GAPDH
25

Figure E2

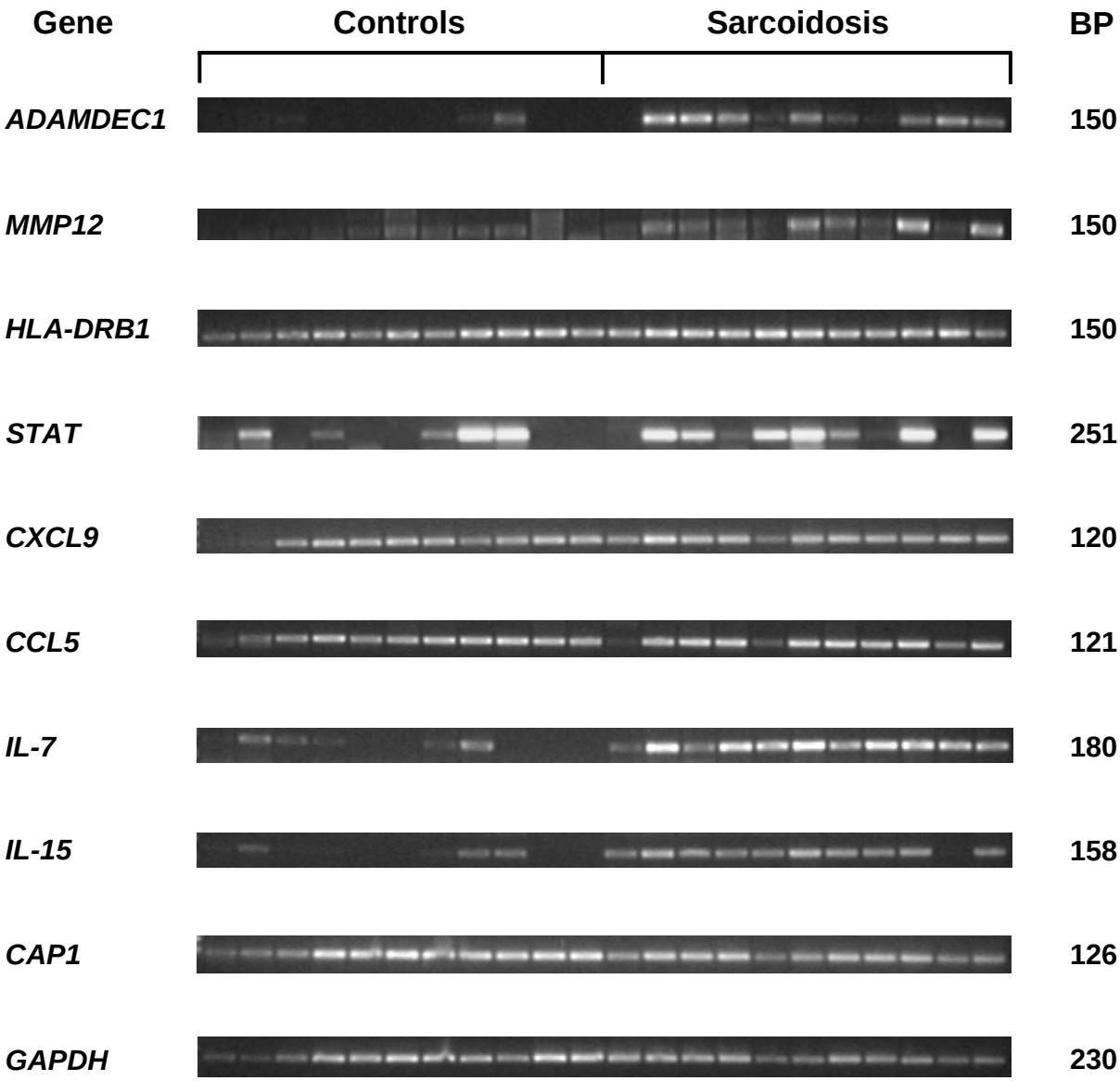
Gene
Controls
Sarcoidosis
BP
ADAMDEC1
150
MMP12
150
HLA-DRB1
150
STAT
251
CXCL9
120
CCL5
121
IL-7
180
IL-15
158
CAP1
126
GAPDH
230

Figure E3