

## 1 ***SUPPLEMENTARY MATERIALS and METHODS***

2

### 3 **Bacterial Strains and Growth Conditions**

4

5 Serotype M1 GAS strains commonly cause pharyngitis and invasive infection<sup>1</sup>. Strain  
6 MGAS5005 (ATCC # BAA-947), a wild-type (WT) clinical strain (serotype M1), and its isogenic  
7  $\Delta covR$  derivative strain (JRS950) have been described<sup>2-4</sup>. Bacteria were cultured statically on  
8 Trypticase soy agar containing 5% sheep blood agar (Becton Dickinson, Cockeysville, Md.), or in  
9 Todd-Hewitt (TH) broth (Becton Dickinson) containing 0.2% (wt/vol) yeast extract (THY; Difco  
10 Laboratories, Detroit, MI), at 37°C in an atmosphere containing 5% CO<sub>2</sub>.

11

12

### 13 **Mouse Soft-Tissue Infection Model**

14

15 This model of GAS soft-tissue infection has been used extensively to study bacterial-host  
16 interactions<sup>2-5</sup>. Our experimental protocol was approved by the Institutional Animal Care and Use  
17 Committee, National Institute of Allergy and Infectious Diseases (NIAID). Bacterial strains  
18 MGAS5005 and JRS950 were grown to late exponential (LE) phase (OD<sub>600</sub> ~0.75) in THY broth,  
19 washed twice in pyrogen-free Dulbecco's PBS (DPBS), and resuspended as inocula in pyrogen-  
20 free PBS. Cells were harvested at LE phase to limit infectivity differences associated with up-  
21 regulated capsule biosynthesis, which is maximal in the early-to-mid exponential growth phases.  
22 Immediately before inoculation, the animals were weighed and anesthetized with isoflurane  
23 (Aerrane; Ohmeda Caribe, Guayama, P.R.) inhalation. Five-week-old (20- to 25-g) outbred,

1 immunocompetent, hairless male Crl:SKH1-*hrBR* mice (Charles River Breeding Laboratories, Bar  
2 Harbor, Maine) were maintained on standard laboratory food and water *ad libitum*. The animals  
3 were randomly assigned to one of two treatment groups ( $n = 27$  per group), and inoculated  
4 subcutaneously (s.c.) in the dorsal side with either WT or  $\Delta covR$  GAS (about  $3 \times 10^7$  CFU)  
5 contained in 0.1-ml pyrogen-free PBS. The actual number of colony-forming units (CFU) of  
6 viable bacteria inoculated per mouse was verified by growth on blood agar. To blind the  
7 investigator, cage numbers were reassigned after inoculation, and the blind was broken after data  
8 analysis. Length (L) and width (W) values were used to calculate abscess volume [ $V = 4/3\pi(L/2)^2$   
9  $\times (W/2)$ ] and area [ $A = \pi(L/2) \times (W/2)$ ], using equations for a spherical ellipsoid as described<sup>3</sup>.  
10 Mice were euthanized at 53 hr post inoculation and weighed. The infection site was swabbed to  
11 confirm GAS infection, and tissue was obtained from each animal via a biopsy that included  
12 dermis and underlying soft-tissue lesions. Tissues were wrapped in aluminum foil, snap-frozen in  
13 liquid nitrogen, and stored at  $-80^\circ\text{C}$  until total RNA was isolated. Three additional control mice  
14 injected with sterile saline failed to show symptoms of clinical infection, and did not grow GAS  
15 bacterial colonies upon plating.

16

17

## 18 Experimental Design

19

20 A one-factor experimental design with two treatment levels was used for array experiments in this  
21 study. Mice ( $n = 54$ ) were assigned at random to one of the two treatment groups (WT or  $\Delta covR$   
22 mutant GAS strain, each  $n = 27$ ). As data quality is of paramount importance, we took care during  
23 the design of our custom oligonucleotide Affymetrix GeneChip<sup>®</sup> (designated RMLChip herein) to

1 exclude cross-hybridizing sequences and to use randomized blocking in our experimental design.  
2 Samples were randomized before all preparation procedures and we ensured that batches of  
3 sample preparation, hybridizations, and post-hybridization washes were not confounded with  
4 treatment. For example, samples were randomly allocated to positions on a single 96-well plate in  
5 a 6 x 9 rectangle near the center of the plate to protect against positional biases such as thermal  
6 edge effects and other unknown positional effects. In order to protect against the possibility of  
7 random “unlucky” positional allocation, 10,000 random assignments were generated and then each  
8 assignment was tested for right-to-left, top-to-bottom, and edge-to-center imbalances in allocation  
9 of wild type and *covR*-minus. These tests were performed by first computing the distance of each  
10 sample to the top, left, and middle of the plate. For each random assignment, three *t*-tests were  
11 performed to test for a difference between wild type and *covR*-minus in the three distances and the  
12 minimum *P*-value recorded. Finally, the results of the 10,000 random allocations were sorted by  
13 the minimum of the three *P*-values. The randomization with the largest minimum *P*-value was  
14 selected as the actual assignment of samples to positions. To minimize confounding variables,  
15 samples were placed in the same position for each 96-well plate used during sample preparation  
16 (e.g., RNA Extraction II, cDNA synthesis, cDNA cleanup, *in-vitro* transcription, clean-up,  
17 concentration adjustments, and target fragmentation).

18 Batches for RNA extraction I were performed with 12 samples per batch. For all steps  
19 including hybridization and wash batches, samples were run sequentially in such a way that each  
20 batch was balanced to include equal numbers in each batch of both WT- and  $\Delta$ *covR*- infected  
21 samples. An alternating RMLChip scanning order (WT /  $\Delta$ *covR* / WT / etc.) also was used to  
22 balance scanner effects over time evenly across treatment groups. All samples were run on the  
23 same GeneChip<sup>®</sup> lot to avoid unnecessary noise in the data.

24

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21

## RNA isolation

Frozen tissue extracts were divided into three aliquots from which RNA was purified. Tissue extracts were pulverized with a series of sharp blows delivered with a 3-pound drilling hammer (Razor-Back<sup>®</sup>). The extracts (~1 g) were homogenized twice for 20s in Lysing Matrix B tubes (MP Biomedicals) containing 300  $\mu$ l CRSR-Blue (MP Biomedicals) and 300  $\mu$ l acid phenol/chloroform, pH 4.3 (Sigma) in a FastPrep<sup>®</sup> FP 120 (MP Biomedicals) at speed 5.5. Sample mixtures were heated at 65°C for 20 min to complete lysis. Following centrifugation at 16,000 x g for 15 min, glycogen (250  $\mu$ g) was added to the recovered aqueous fractions and the volumes concentrated to 100  $\mu$ l in a vacuum concentrator (Brinkmann). The concentrate was fragmented with a Qiasredder (QIAGEN, Inc.) and the isolated total RNA (containing both bacterial and host RNA) was further purified in 96-well format using a plate centrifugation system (RNeasy 96; QIAGEN), with on-column DNase I treatment and post-treatment with DNAFree (Ambion, Inc.) as described<sup>3,6</sup>. Electrophoretic analysis with an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.) and measurement of the  $A_{260}/A_{280}$  ratios were used to assess RNA integrity. Quantitative-PCR (TaqMan<sup>®</sup>) assays were performed with RNA templates to ensure that contaminating bacterial genomic DNA was absent. Two RNA aliquots were pooled to perform the microarrays; the remaining extract used for real-time RT-PCR validation.

## 1 cDNA Labeling

2  
3 The microarray targets were prepared in 96-well format according to the protocol supplied by the  
4 manufacturer (Affymetrix), with modifications. Each extracted RNA sample was divided into 2-3  
5 aliquots of 10 µg total RNA to which 0.8 µg bacteriophage MS2 carrier RNA (Roche Bioscience)  
6 was added. Control spike transcripts (130 pM) were added to each RNA aliquot, and 5 µg random  
7 primers (Invitrogen) were annealed (10 min at 70°C, 10 min at 25°C). cDNA synthesis reactions  
8 and post-synthesis RNA digestion were performed in a thermocycler otherwise as described<sup>7</sup>. The  
9 resultant cDNA was purified using Qiaquick 96 kit (QIAGEN) according to the manufacturer's  
10 recommendation. For cDNA fragmentation, 3 µg of cDNA and 0.75 U of DNase I (Roche  
11 Bioscience) were used (10 min at 37°C, 10 min at 98°C). The desired cDNA size range of 50-200  
12 bases was verified by separating 200 ng of cDNA on a RNA 6000 Nano LabChip<sup>®</sup> (Agilent) using  
13 the 2100 BioAnalyzer (Agilent) with no added dye in the loading buffer. The fragmented cDNA  
14 was then end-labeled with biotin-ddUTP as per the Enzo BioArray<sup>™</sup> terminal labeling kit (60 min  
15 at 37°C).

## 18 GeneChip<sup>®</sup> Design

19  
20 A custom 18-micron feature size, anti-sense oligonucleotide array representing approximately  
21 249,690 25-mer probe-pairs (16 probe pairs, each consisting of one perfect match probe (PM) and  
22 a mismatch probe (MM), per probe set) manufactured by Affymetrix Inc. was used as described<sup>7-9</sup>.  
23 Each probe-set is used to detect the presence of a single transcript. To facilitate analysis of GAS

1 samples in the presence of host cells, during the design process all probe set sequences were pre-  
2 pruned to exclude cross-hybridizing sequences (those exhibiting sequence similarity) with human,  
3 rat and mouse genome ORFs represented on Affymetrix Inc. arrays, as well as 12 additional  
4 bacterial genome sequences, as described<sup>9</sup>. The RML Affymetrix custom GeneChip<sup>®</sup> array  
5 (designated RMLChip herein) contains 2,636 probe sets (42,351 probe-pairs) for 2,636 predicted  
6 GAS open reading frames (ORFs)<sup>9</sup>. These features represent a composite superset of six GAS  
7 genomic sequences representative of serotypes M1, M3, M5, M12, M18, and M49 (sequenced  
8 strains are designated SF370, MGAS315, Manfredo, MGAS9429, MGAS8232, and CS101,  
9 respectively). Although the RMLChip was not designed based on the genome sequence of strain  
10 MGAS5005 (GenBank Accession No. CP000017)<sup>4</sup>, the composite RMLChip contains 1,893  
11 redundant probe sets (BLAST score match to MGAS5005 ( $E > 0.01$ )) representing more than 90%  
12 coverage of the total number of predicted coding regions (1,869 ORFs) encoded by this M1 GAS  
13 genome<sup>9</sup>. Several GAS genes are represented by more than one probe set. Downstream genome  
14 analysis was accomplished using MicrobesOnline (available at <http://www.microbesonline.org/>)<sup>10</sup>  
15 and in-house bioinformatics analysis.

16

17

## 18 GeneChip<sup>®</sup> Hybridization

19

20 Target hybridizations, washing, staining, and scanning were performed by the NIAID Affymetrix  
21 core facility (SAIC-Frederick, MD) using a GeneChip<sup>®</sup> hybridization oven and the *Pseudomonas*  
22 *aeruginosa* hybridization protocol (Affymetrix, Santa Clara, CA). The hybridization solution  
23 volume used was 200  $\mu$ l because the RMLChip is a standard size array. Each array was scanned  
24 at 570 nm at 3- $\mu$ m resolution with a GeneArray<sup>®</sup> scanner. Scanned DAT-image files were

1 analyzed with Affymetrix<sup>®</sup> Microarray Suite (MAS) 5.0 software (Affymetrix, Santa Clara, CA).  
2 The raw CEL-files have been submitted to Gene Expression Omnibus  
3 ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)).

4

5

## 6 Microarray Data Quality Assessment and Statistical Analysis

7

8 Expression estimates for each gene were obtained using the PM-MM difference model of dCHIP  
9 (available at <http://www.dchip.org/>)<sup>11</sup>. The gene expression estimates were normalized across  
10 samples by quadratic scaling to an artificial array with the median expression for each gene<sup>12</sup>.  
11 Two-dimensional scatterplots of expression estimates were generated for all pairs of samples  
12 within the same treatment group to examine uniformity across samples, and revealed 5 samples  
13 with low correlation to the other within-factor samples (data not shown). Hierarchical clustering  
14 and Principal Components Analysis (PCA) also were performed to identify array outliers and,  
15 except for five cases, the microarray expression data sets clustered according to treatment (GAS  
16 strain) (Supplementary Figure 2). The expression data sets for WT-inoculated mice designated 05,  
17 29, 39, and 43 and the mutant-inoculated mouse designated 38 clustered independently of other  
18 samples within the same treatment group. At least 2 of these 5 outliers were correlated with  
19 poorer quality and yield of RNA. Consequently, these five arrays were removed from the data  
20 matrix and not analyzed further.

21 Randomization and blocking enabled us to assess the respective error variance contributed  
22 by unavoidable (technical) experimental variables. Technical variables included in the analysis  
23 included sample preparation, sample hybridization batches, and post-hybridization wash batches.  
24 For example, during array washing the fluidic station could process only a maximum of 8 slides at

1 a time and wash batch was found to contribute variance (Supplementary Figure 1). Downstream  
2 data analysis was performed on the composite array data with MGAS5005-specific probe sets  
3 using Partek Pro™ (Partek Inc., St. Louis, MO). Image files were converted to text format, and  
4 normalized according to a symmetric square root transformation. To evaluate expression rankings,  
5 the normalized absolute square root expression estimates were integer-ranked such that the most  
6 abundant transcript reported as transcript rank “1”. Integer rank assignments were increased  
7 correspondingly with ever decreasing transcript detection. To investigate expression correlations  
8 between genes, standard Pearson correlation coefficients were determined for select genes *versus*  
9 all other genes.

10 Lastly, to investigate the strain effect, the resultant normalized expression estimates were  
11 analyzed by Analysis of Variance (ANOVA) with treatment (WT versus  $\Delta covR$  strain) as a fixed  
12 effect. Final results were subjected to multiple testing correction using  $Q \leq 0.05$  false discovery  
13 rate (FDR) cutoff values<sup>13</sup>. Using rigorous permutation-based statistics, we also performed  
14 significance analysis of function and expression (termed SAFE)<sup>14</sup> to assess the significance of  
15 multiple gene categories in GAS *in vivo* transcriptional responses across strains. All GAS probe  
16 sets were assigned to one of 17 functional categories (including “Unknown”) as described<sup>7</sup>. For  
17 each gene, the Student’s *t*-statistic was calculated as local statistic measuring the association  
18 between expression and treatment (strain). The Wilcoxon rank sum was then computed as global  
19 statistic to assess how the distribution of local statistics within a functional category differed from  
20 local statistics outside the category. To account for multiple testing of functional categories,  
21 empirical *P*-values for each functional category were obtained by recomputing Wilcoxon statistics  
22 across 10,000 permutations of the array assignments in which treatment assignments were  
23 randomized. Unknown correlation among genes was conserved across permutations using this  
24 approach and allowed computation of permutation-based estimates for the FDR<sup>14</sup> using the



1 Benjamini-Yekutieli method<sup>15</sup>. In the resultant SAFE plots, *P*-values would trace the identity line  
2 (solid diagonal) representing the null hypothesis if no differential expression were present. When  
3 small *P*-values occur with greater frequency than expected by chance, then the cumulative  
4 distribution function for a gene category diverges from the diagonal, and responsible genes are  
5 designated as differentially regulated. Low-ranking, negative *t*-statistics correlate to upregulated  
6 expression in the mutant strain (left side of the SAFE plot); conversely positive *t*-statistics  
7 correlate to downregulated expression in the mutant.

8

9

## 10 Quantitative Real-Time PCR Analysis

11

12 Real-time reverse transcription-PCR (RT-PCR) assays were conducted to validate a subset of the  
13 microarray data<sup>7</sup>. Eight oligonucleotide primer pairs and 6FAM-labeled probe sets (specific for  
14 *cfa*, *dppA*, *emm1*, *sceD*, *sclA*, *sic*, *slo*, and *speA2*) were used to perform target amplification and  
15 detection from cDNA templates in 20 µL multiplex two-step RT-PCR reactions as described<sup>7</sup>.

16 Targets were selected to encompass the full range of expression signal values identified by array

17 transcriptome analysis. Target abundance was normalized to JOE-labeled internal reference

18 transcript *proS*, which is transcribed at constant levels throughout the GAS growth cycle *in vitro*

19 and not affected by *covR* inactivation<sup>3</sup>. Differences in median values were evaluated for statistical

20 significance with the Mann-Whitney Rank Sum Test at the  $P \leq 0.001$  level.

21

22

## 1 Sampling, Histological and Immunohistochemical Assessment

2  
3 Tissue used for histological examination was prepared from mice inoculated s.c. with  $2.4 \times 10^7$   
4 CFU of GAS strains MGAS5005 or  $\Delta covR$  JRS950 ( $n = 16$ , each strain) as described above,  
5 except that 4-week -old (15- to 20-g) female Crl:SKH1-*hrBR* mice (Charles River Breeding  
6 Laboratories) were used. Six animals inoculated with PBS were used as controls. For assessment  
7 of bacterial content, histopathology, and bacterial protein expression, mice were euthanized 48 hrs  
8 post inoculation and the skin and underlying soft-tissue removed from inoculation sites and fixed  
9 in 10% buffered formalin before embedding in paraffin. To assess the presence of bacteria and  
10 pathological changes, formalin-fixed tissues were sectioned and stained with Gram's stain or  
11 hematoxylin and eosin stain (Sigma, St. Louis, MO) according to standard methodologies. An  
12 Olympus model BX51 microscope equipped with a Q-FIRE (Olympus) camera was used for  
13 image capture.

## 16 Immunohistochemical Analysis

17  
18 Rabbit polyclonal anti-GAS antibodies made against purified recombinant GAS proteins were  
19 used for immunostaining<sup>16</sup>. Targets for proteomic confirmation were selected from the full range  
20 of microarray expression values (highest microarray transcript ranking 1<sup>st</sup>/3<sup>rd</sup> for the most  
21 abundant target (SIC); lowest microarray transcript ranking 1780<sup>th</sup>/1602<sup>nd</sup> for IdeS/Mac, for the  
22 WT and mutant strains respectively). As there are only a limited number of GAS antigens for  
23 which highest quality mono- or poly-clonal antisera are available, by necessity targets also were

1 selected based upon available immunological reagents. Paraffin-embedded tissues were cut into  
2 4- $\mu$ m sections and stained with antibodies specific for 16 bacterial antigens [M5005\_Spy ORF  
3 numbers designated in square brackets] (AtmB [0271]; PrtS [0342]; MtsA [0368]; IdeS/Mac  
4 [0668]; [0942]; PstS [0955]; SpeA2 [0996]; MalE [1058]; PrsA [1133]; [1308]; HtsA/SiaA  
5 [1528]; Shp [1529]; DppA [1704]; Lmb [1711]; Fba [1713]; SIC [1718]) using biotinylated  
6 secondary antibodies in combination with HRP-coupled streptavidin (DAKO Corporation,  
7 Carpinteria, CA) and the substrate AEC (BioGenex, San Ramon, CA). To evaluate nonspecific  
8 staining, a polyclonal antibody recognizing a control peptide designated “M3.1/1-24” (Bethyl  
9 Laboratories), representing the *N*-terminal peptide of serotype M3 Emm3.1, was used as negative  
10 control reagent since this peptide is not encoded within the genome of the serotype M1 WT strain  
11 MGAS5005<sup>4</sup>. All immunohistochemically (IHC) stained sections were counterstained with  
12 Mayer’s hematoxylin and mounted using synthetic aqueous-based mounting medium (DAKO  
13 Faramount).

1 **REFERENCES**

2

- 3 1. Musser JM, Krause RM: The revival of Group A streptococcal diseases, with a  
4 commentary on staphylococcal toxic shock syndrome. *Emerging Infections*. Edited by  
5 Krause RM. New York, Academic Press, 1998, pp 185-218
- 6 2. Hoe NP, Nakashima K, Lukomski S, Grigsby D, Liu M, Kordari P, Dou SJ, Pan X,  
7 Vuopio-Varkila J, Salmelinna S, McGeer A, Low DE, Schwartz B, Schuchat A, Naidich S,  
8 De Lorenzo D, Fu YX, Musser JM: Rapid selection of complement-inhibiting protein  
9 variants in group A *Streptococcus* epidemic waves. *Nat Med* 1999, 5:924-929
- 10 3. Graham MR, Smoot LM, Migliaccio CAL, Virtaneva K, Sturdevant DE, Porcella SF,  
11 Federle MJ, Adams GJ, Scott JR, Musser JM: Virulence control in group A *Streptococcus*  
12 by a two-component gene regulatory system: Global expression profiling and *in vivo*  
13 infection modeling. *Proc Natl Acad Sci USA* 2002, 99:13855-13860
- 14 4. Sumby P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM, Sturdevant  
15 DE, Graham MR, Vuopio-Varkila J, Hoe NP, Musser JM: Evolutionary origin and  
16 emergence of a highly successful clone of serotype M1 group A *Streptococcus* involved  
17 multiple horizontal gene transfer events. *J Infect Dis* 2005, 192:771-782
- 18 5. Lukomski S, Hoe NP, Abdi I, Rurangirwa J, Kordari P, Liu M, Dou SJ, Adams GG,  
19 Musser JM: Nonpolar inactivation of the hypervariable streptococcal inhibitor of  
20 complement gene (*sic*) in serotype M1 *Streptococcus pyogenes* significantly decreases  
21 mouse mucosal colonization. *Infect Immun* 2000, 68:535-542
- 22 6. Virtaneva K, Graham MR, Porcella SF, Hoe NP, Su H, Graviss EA, Lemon WJ, Bailey JR,  
23 Parnell MJ, Musser JM: Group A *Streptococcus* Gene Expression in Humans and  
24 *Cynomolgus* Macaques with Acute Pharyngitis. *Infect Immun* 2003, 71:2199-2094

- 1 7. Graham MR, Virtaneva K, Porcella SF, Barry WT, Gowen BB, Johnson CR, Wright FA,  
2 Musser JM: Group A *Streptococcus* transcriptome dynamics during growth in human  
3 blood reveals bacterial adaptive and survival strategies. *Am J Pathol* 2005, 166:455-465
- 4 8. Sumby P, Barbian KD, Gardner DJ, Whitney AR, Welty DM, Long RD, Bailey JR, Parnell  
5 MJ, Hoe NP, Adams GG, DeLeo FR, Musser JM: Extracellular deoxyribonuclease made  
6 by group A *Streptococcus* assists pathogenesis by enhancing evasion of the innate immune  
7 response. *PNAS* 2005:0406641102
- 8 9. Virtaneva K, Porcella SF, Graham MR, Ireland RM, Johnson CA, Ricklefs SM, Babar I,  
9 Parkins LD, Romero RA, G.J. Corn, Gardner DJ, Bailey JR, Parnell MJ, Musser JM:  
10 Longitudinal analysis of Group A *Streptococcus* transcriptome in experimental pharyngitis  
11 in cynomolgus macaques. *Proc Natl Acad Sci USA* 2005, 102:9014-9019
- 12 10. Alm EJ, Huang KH, Price MN, Koche RP, Keller K, Dubchak IL, Arkin AP: The  
13 MicrobesOnline Web site for comparative genomics. *Genome Res* 2005, 15:1015-1022
- 14 11. Li C, Wong WH: Model-based analysis of oligonucleotide arrays: Expression index  
15 computation and outlier detection. *Proc Natl Acad Sci USA* 2001, 98:31-36
- 16 12. Yoon H, Liyanarachchi S, Wright FA, Davuluri R, Lockman JC, de la Chapelle A,  
17 Pellegata NS: Gene expression profiling of isogenic cells with different TP53 gene dosage  
18 reveals numerous genes that are affected by TP53 dosage and identifies CSPG2 as a direct  
19 target of p53. *Proc Natl Acad Sci USA* 2002, 99:15632-15637
- 20 13. Storey JD, Tibshirani R: Statistical significance for genomewide studies. *Proc Natl Acad*  
21 *Sci USA* 2003, 100:9440-9445
- 22 14. Barry WT, Nobel AB, Wright FA: Significance analysis of functional categories in gene  
23 expression studies: a structured permutation approach. *Bioinformatics* 2005, 21:1943-1949

- 1 15. Yekutieli D, Benjamini Y: Resampling based FDR controlling multiple hypotheses testing.  
2 J Stat Plan Infer 1999, 82:171-196
- 3 16. Lei B, Liu M, Chesney GL, Musser JM: Identification of new candidate vaccine antigens  
4 made by *Streptococcus pyogenes*: purification and characterization of 16 putative  
5 extracellular lipoproteins. J Infect Dis 2004, 189:79-89  
6