

1 [Supplementary Figure Legends]

2

3 Supplementary Figure 1. GAS microarray outliers revealed by hierarchical clustering.
4 Depicted is the clustering dendrogram generated by hierarchical clustering of the *in vivo* profiling
5 data set. Expression array data sets were clearly discriminated by strain, except for WT-inoculated
6 animals designated 05, 29, 39, and 43 and the mutant-inoculated mouse designated 38, which
7 clustered independent of other samples within their same treatment groups. Consequently, these
8 five arrays were removed from the expression data matrix and not further analyzed.

9

10

11 Supplementary Figure 2. Analysis of Variance (ANOVA) reveals a significant batch effect in
12 the microarray expression data set. An ANOVA model was applied to an absolute square root
13 transform of the dCHIP PM-MM expression estimates with the fixed effect treatment (Strain: WT
14 or $\Delta covR$) using Partek Pro™ v6.01 (Partek Inc.). Sample preparation, sample hybridization, and
15 post-hybridization wash batches were investigated in the ANOVA model for their contributions to
16 the variance. Strain is the biological factor of interest, whereas wash batch is considered an
17 undesirable batch effect. As wash batch was observed to contribute, it was included in the
18 ANOVA model as a block. A linear sliding scaling was performed to remove the wash batch effect
19 without impact upon strain or model error terms. Differentially expressed genes were identified
20 from the resultant wash batch-adjusted expression data set. Depicted in (A) is an example of the
21 effect of batch washing (in subgroups over two days) upon detected *dppA* transcript levels. Bar
22 charts (depicted in B and C) summarize the main factors contributing variance in the microarray
23 expression data set. Note variance contributed by wash batch is vastly improved by linear scaling
24 (C) compared to the raw data (B).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Supplementary Figure 3. Cluster image map relating the effect of strain (treatment) upon GAS *in vivo* gene expression. Depicted is the cluster image map for the *in vivo* profiling data set as generated by agglomerative hierarchical clustering of the absolute square root-transformed expression estimates for the significant 76 differential GAS transcripts. Although some subject (mouse-to-mouse) variation in transcript expression is evident, the expression array data sets are clearly discriminated by strain, with WT and Mutant clustering independently during mouse soft-tissue infection.

Supplementary Figure 4. Principal component analysis (PCA) clustering of GAS differentially expressed transcripts *in vivo* during soft-tissue infection. Depicted is the two-dimensional projection onto the plane spanned by the first and third principal components for the *in vivo* profiling data set, as generated with 76 differential transcripts (by strain). Colored ellipsoids represent two standard deviations from the mean of each treatment group. Overall the expression array data sets are clearly discriminated by strain using these markers, with WT and Mutant clustering independently. A limited number of deviating outliers are apparent, as indicated by their exclusion from ellipsoids delineating each treatment.

Supplementary Figure 5. Statistical analysis of function and expression for GAS Virulence during model soft-tissue infection. The SAFE plot shows the cumulative distribution function (CDF) for ranked *t*-statistics of genes within the selected functional category Virulence (dashed

1 line) relative to all expressed GAS genes (solid line), when testing for the effect of strain. If the
2 degree of differential expression of genes within a category is the same as that of all other GAS
3 genes, the distribution will trace the identity line (solid line). If small P -values occur with greater
4 frequency than by chance among genes in a category, then the distribution curve diverges from the
5 diagonal identity line (solid line), with the most differentially regulated genes shifted to the right.
6 Tick marks along the top of the plot show the ranks of individual transcripts in the category.
7 Shaded regions correspond to genes that pass a nominal level of significance (empirical P -values \leq
8 0.05). Genes exhibiting significant differential regulation are indicated according to their assigned
9 gene names or M5005 ORF numbers. Transcripts differentially expressed with treatment (strain)
10 that exhibit a leftward shift (negative t -statistic) indicate upregulation, whereas those showing a
11 rightward shift indicate down-regulation. The set of 69 genes involved in Virulence primarily
12 exhibit upregulation in the $\Delta covR$ mutant versus the WT strain during soft-tissue infection
13 Virulence (Empirical $P = 0.009$; FDR = 0.055), confirming previous *in vitro* findings.

14

15

16 Supplementary Figure 6. Statistical analysis of function and expression for GAS Stress
17 Adaptation during model soft-tissue infection. The SAFE plot shows the cumulative distribution
18 function (CDF) for ranked t -statistics of genes within the selected functional category Virulence
19 (dashed line) relative to all expressed genes (solid line), when testing for the effect of strain. If the
20 degree of differential expression of genes within a category is the same as that of all other GAS
21 genes, the distribution will trace the identity line (solid line). If small P -values occur with greater
22 frequency than by chance among genes in a category, then the distribution curve diverges from the
23 diagonal identity line (solid line), with the most differentially regulated genes shifted to the right.
24 Tick marks along the top of the plot show the ranks of individual transcripts in the category.

1 Shaded regions correspond to genes that pass a nominal level of significance (empirical P -values \leq
2 0.05; $|t| \geq 2.42$). Genes exhibiting significant differential regulation are indicated according to
3 their assigned gene names or M5005 ORF numbers. Transcripts differentially expressed with
4 treatment (strain) that exhibit a leftward shift (negative t-statistic) indicate upregulation, whereas
5 those showing a rightward shift indicate down-regulation. In contrast to previous *in vitro* findings,
6 the set of 35 genes involved in Stress Adaptation show a significant downregulation on average in
7 the $\Delta covR$ mutant versus the wild-type strain during soft-tissue infection (empirical $P = 0.005$;
8 FDR = 0.044).