

Data-Driven Transcriptional Markers for Classifying *Escherichia coli* and *Staphylococcus aureus*-Induced Sepsis in Adult Patients

Mahnaz Irani-Shemirani^{1, 2*}

¹ Umeå Plant Science Centre (UPSC), Department of Plant Physiology, Umeå University, 901 87 Umeå, Sweden

² Department of Laboratory Medicine, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

***Corresponding author;**

Mahnaz Irani-Shemirani,

Umeå Plant Science Centre (UPSC), Department of Plant Physiology, Umeå University, 901 87 Umeå, Sweden. mahnaz.irani-shemirani@umu.se

Supplementary Table S1. List of genes selected by Lasso regression

Gene	Positive Coefficient	Gene	Negative Coefficient
IFI27	0.002881147	AMFR	-0.003646205
IFIT1	0.038257945	GBP1	-0.009361866
DCK	0.058605274	PI3	-0.000103139
EIF1AY	0.00209795	MARCO	-4.20E-05
DDX3Y	0.026761014	TNFAIP6	-0.054268902
CLC	0.019316271	HIST1H1T	-0.001788501
APOBEC3B	0.018511076	TNNT1	-0.018785541
HBZ	0.008048641	HIST1H2AD	-0.009471913
PF4V1	0.006544974	GUSBP3	-0.026534852
IGHG1	0.039405155	LILRA5	-0.025594504
YLPM1	0.037414014	FCGR1A	-0.056150575
		MS4A4A	-0.04067223
		CD177	-0.06776166
		ERAP2	-0.002085012

Supplementary Table S2. Classification metric and PCA report of selected genes by Lasso regression

Model (samples, n)	Groups	Precision	Recall	f1-score	Support	PC1	PC2	PC3
25 genes (n=94)	E. coli	1	0.8	0.89	5			
	S. aureus	0.88	1	0.93	7			
	Control	1	1	1	7			
	accuracy			0.95	19	35.98	9.84	8.42
22 genes (n=94)	E. coli	1	0.8	0.89	5			
	S. aureus	0.88	1	0.93	7			
	Control	1	1	1	7			
	accuracy			0.95	19	38.4	9.73	6.69
25 genes (n=151)	E. coli	0.9	1	0.95	9			
	S. aureus	1	0.92	0.96	13			
	Control	1	1	1	9			
	accuracy			0.97	31	33.21	14.85	9.69
25 genes (n=228)	E. coli	1	1	1	11			
	S. aureus	1	1	1	19			
	Control	1	1	1	16			
	accuracy			1	46	35.3	14.31	10.01