Analysis of the Messenger RNAs in Spores of *Bacillus subtilis*

George Korza[1], Emily Camilleri[1], Joshua Green[1], Janelle Robinson[1], Katja Nagler[2,3], Ralf Moeller[4], Melissa J. Caimano[1,5] and Peter Setlow[1]#

Supplemental Material

References

1. Nagler K, Krawczyk AO, De Jong A, Madela K, Hoffmann T, Laue M, Kuipers OP, Bremer E, Moeller R. 2016. Identification of differently expressed genes during *Bacillus subtilis* spore outgrowth in high-salinity environments using RNA sequencing. Front Microbiol 7:1564.

2. Swarge BN, Roseboom W., Zheng L. Abhyankar WR, Brul S, de Koster CG, de Koning LJ. 2018. "One-Pot" sample processing method for proteome-wide analysis of microbial cells and spores. Proteomics Clin Appl 12:e1700169.

3. Zheng L. Abhyankar W, Ouwerling N, Dekker HL, van Veen H, Rosenboom W, de Koning, L, Brul S, de Koster CG. 2016. *Bacillus subtilis* spore inner membrane proteome. J Proteome Res 15:585-594.

Table S1

Characteristics of RNA-seq data in the current work

Sample	Description*	Total reads	# aligned	# unique	# not aligned	% Unique aligned	Avg depth
1-1DPL	Plates-1d-dc	14766388	14525805	14472994	293394	99.64	34
2-7DPL	Plates-7d-hp-dc	14752008	14563140	14509380	242628	99.63	23
4-7DL	Liquid-7d-hp-dc	12315554	12167080	12129116	186438	99.69	19
Setlow_2018_08_27_S1	Plates-7d-hp	14721666	11028926	10217288	4504378	92.64	307
Setlow_2018_08_27_S2	Plates-7d-hp	14055316	11363461	10652056	3403260	93.74	1161
Setlow_2018_08_27_S3	Plates-7d-hp	12932458	12538604	12393640	538818	98.84	238
Setlow_2018_10_31_S1	Liquid-7d-lp	17290054	15800103	15735356	1554698	99.59	401
Setlow_2018_10_31_S2	Liquid-7d-lp	19099550	18494815	18452576	646974	99.77	109
Setlow_2018_10_31_S3	Liquid-7d-hp	14277172	13927158	13891238	385934	99.74	96

* Abbreviations used are: d, day; dc, chemically decoated; hp, highly purified; and lp, less well purified, with the regimens for spore production and purification described in Methods. The number of days indicates how long spores were held at 4°C while they were being purified.

Table S2*

Most abundant *B. subtilis* spore mRNAs identified in published work, their σ-factor dependence and whether encoded proteins have been detected in spores

1		**2**		**3**	
Gene	**σ^G dependence?**	**Gene**	**σ^G dependence?**	**Gene**	**σ^G dependence?**
sspF	yes	***yhcQ***	yes	***yozQ***	yes
yfhD	yes	***yjzC***	yes	ybxF	no
yrzQ	yes	***ypzI***	yes	ypzE	no
ytzL	yes	rpmA	no	***yrrD***	yes
ypzG	yes	***yhcM***	yes	***sspH***	yes
ykzP	yes	***yxeD***	yes	mtlF	no
sspE	yes	***yuzA***	yes	iseA	no
yhcV	yes	***yhcN***	yes	licH	no
ypzF	yes	**ytzK**	yes	mtlD	no
sspM	yes	gmuB	no	tatAY	no
ykzE	yes	*sspD*	yes	ylqC	no
ytzC	yes	gmuA	no	*rpsR*	no
sspA	yes	ysxB/prp	no	veg	no
yqfX	yes	*rpsU*	no	rpmH	no
sspB	yes	*rplU*	no	*yfhS*	yes
sspK	yes	sboA	no	*rpsO*	no
sspN	yes	*licA*	no	gmuC	no
sspJ	yes	**sspL**	yes	**ykzD**	yes

62	***sspO***	yes	*rpsT*	no	ykzV	no
63	***tlp***	yes	***yoyE***	yes	yhfH	no
64	***yizC***	yes	*licB*	no	***yusG***	yes
65	*coxA*	yes	yyzE	no	rbfA	no
66	***ymfJ***	yes	abrB	no	mleA	no
67	***yrzR***	yes	*yxiE*	no	yrhC	no
68	***yhdB***	yes	*rpsP*	no	*bglA*	no
69	***sspP***	yes	*rpsF*	no	***sspI***	yes

70 *Data are from reference 1 in order of mRNA abundance in columns 1-3; σ^G-dependence was
71 determined as described in Methods; note that some σ^G-dependent genes are also transcribed under
72 σ^F control. mRNAs from genes in bold were detected in the most abundant 78 mRNAs determined
73 in the current work by RNA-seq on RNA from spores prepared on plates and highly purified. The
74 proteins from italicized genes have been identified in *B. subtilis* spores (2,3).

Table S3*

Average molecules/spore for different abundance mRNAs in less well purified *B. subtilis* spores

	A	B	C	D	E	F	G	H
Avg RPKM/ Group	Genes #	Length nt[1]	RPKM	Reads/ gene[2]	RelmRNAs/ gene[3]	mRNA nt/ spo[4]	mRNA nt/ spo ($\Sigma=10^6$)[5]	mRNA/spo[6]
>1300	37	327	$1.2x10^5$	$3.8x10^4$	7170	$8.7x10^7$	$9.9x10^5$	86
290-1299	14	304	455	138	26	$1.1x10^5$	1300	0.5
29-289	67	652	69	45	8.5	$3.7x10^5$	4200	0.13
3-28	474	880	6	5.3	1	$4.2x10^5$	4800	0.013
Total = 592								

*RPKM values are from RNA-seq analysis on two RNA preparation isolated from dormant *B. subtilis* 168 spores (spo) prepared in liquid and less well purified as described in Methods. Values in Columns B-E and H are averages for all mRNAs in a group, while values in Columns A, F and G are totals for all genes or mRNAs in a group.

[1]mRNAs lengths were adjusted for the presence of 5'- and 3'-non-coding sequences by adding 50 nt as described in the legend to Table 2.

[2]Reads/gene are average RPKM values corrected for the average length of mRNAs in this group as (C x B/1000). These values are the relative number of mRNAs for genes in each group.

[4]RelmRNAs/gene are relative levels of individual mRNAs in groups calculated as D/average Reads/gene value for the mRNAs in the lowest abundance group which was set as 1.

[5]Values for mRNA nt/spore were calculated as (E x A x B) for each group.

[6]Calculated as (F x 0.01138) to get a total of 10^6 nt for all 4 groups.

[7]These values were calculated as (G/A x B), and are the average numbers of individual mRNA/spore in each group.

Supplemental Figure Legends.

Fig. S1. Graphical representation of mapped read data for all genes. The *y*-axis indicates RPKM values for individual genes (*x*-axis). Genes are sorted based on RPKM values in descending order. Graphs in panels A and C are based on the average (avg) RPKM values from at least 2 biologically-independent dormant spore preparations, while data for panels B and D-F represent RPKM values with individual spore samples. Abbreviations are: d, day; dc, chemically decoated; hp, highly purified; and lp, less well purified. The regimens for spore production and purification are described in Methods. The number of d indicates how long spores were held at 4°C while they were being purified.

Fig. S2. Representative cartoons depicting read coverage from RNA-seq on RNA from sample Setlow_2018_08_27_S1 (Plates-7d-hp), generated using MacVector Assembler software. Green arrows are used to indicate genes of interest and orientation. Genome coordinates are indicated in red. Cumulative mapped read data for a given gene are shaded light blue. The y-axis for each plot indicates the average running read depth. Regions with no mapped reads are shaded pale green.

Fig. S3. High Sensitivity RNA ScreenTape analysis of total RNA from mid log phase *B. subtilis* vegetative cells grown in LB medium and dormant spores prepared on plates and highly purified over 7 d at 4°C. RNA samples were run on an Agilent TapeStation 4200 instrument.

Fig. S1.

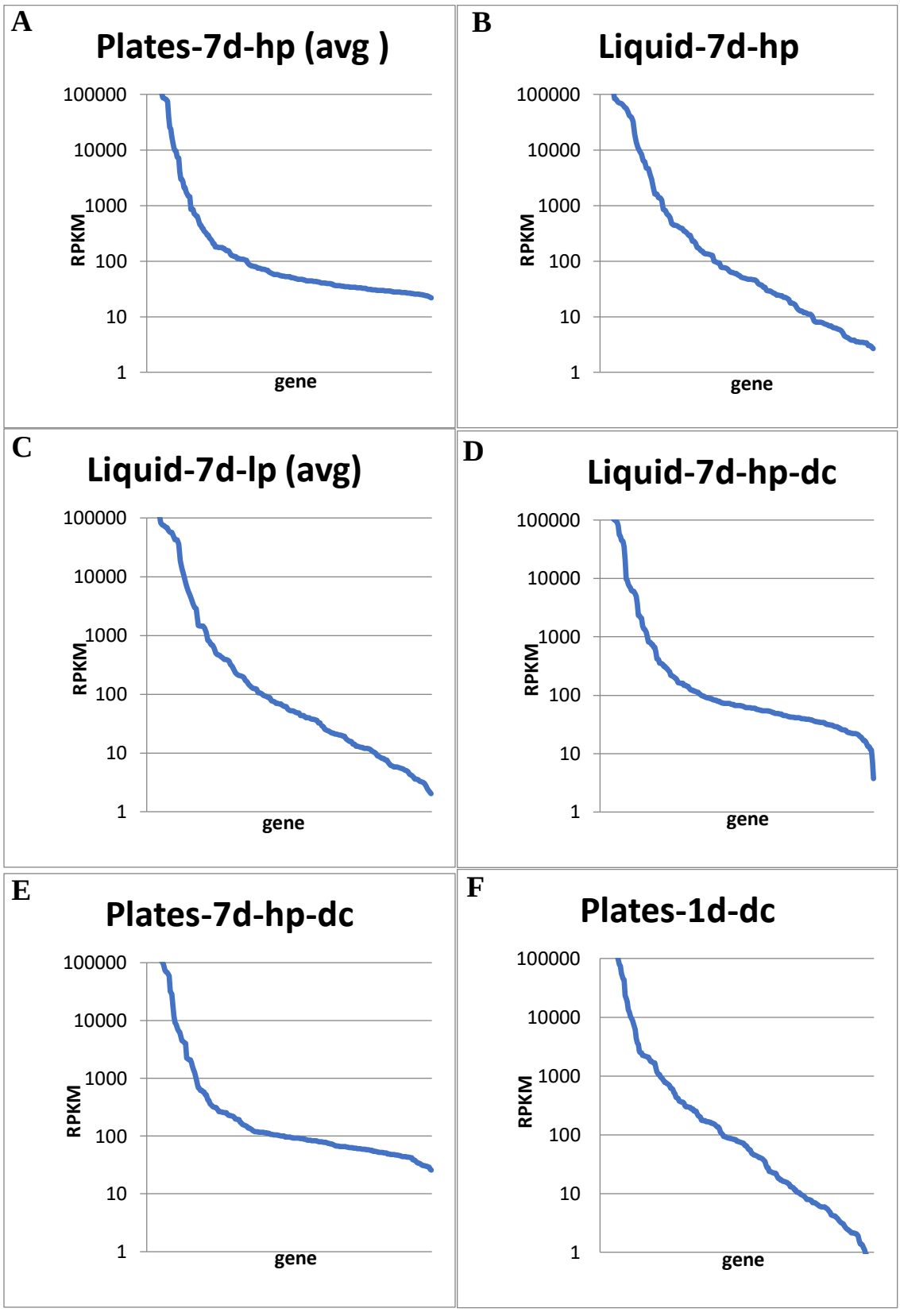
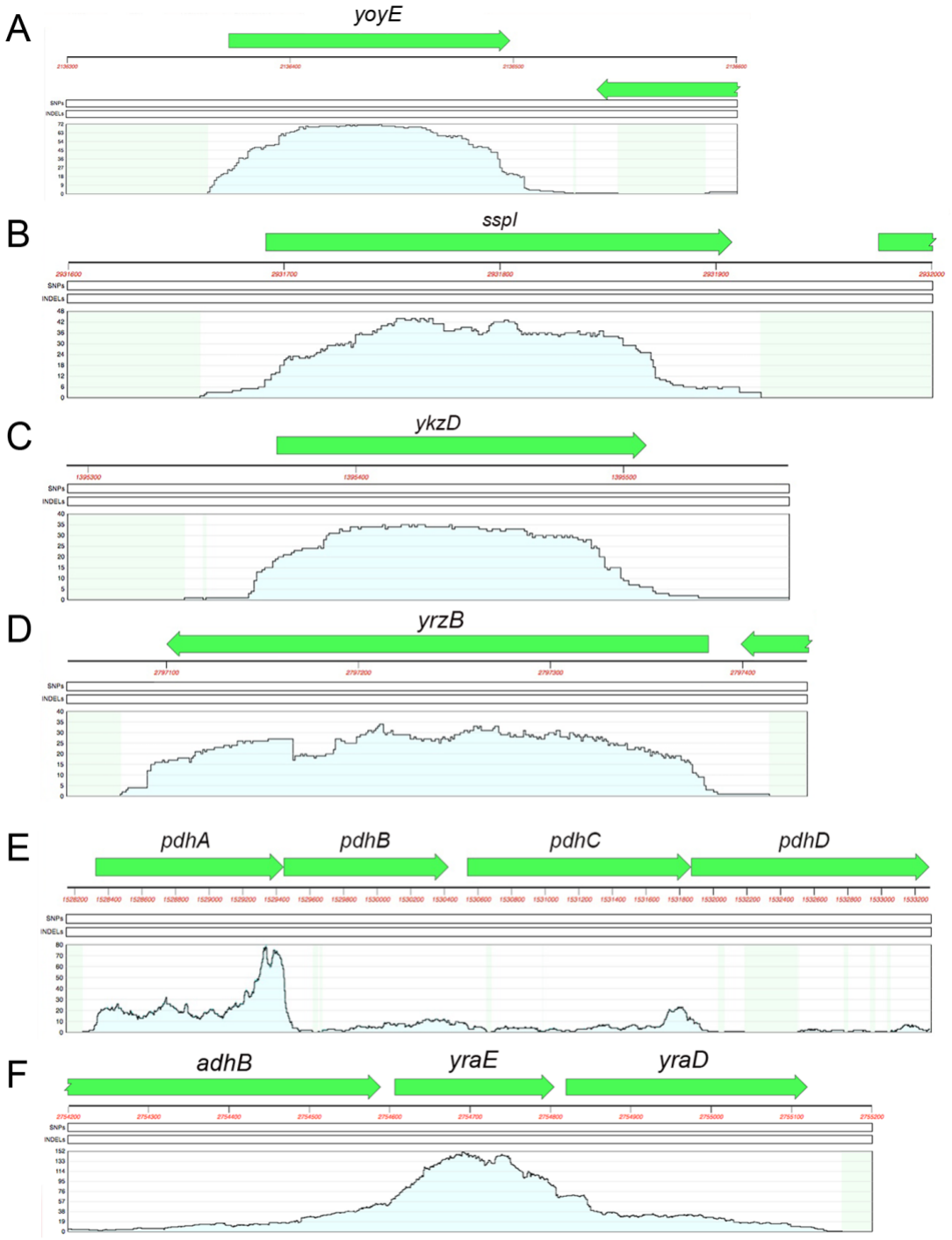
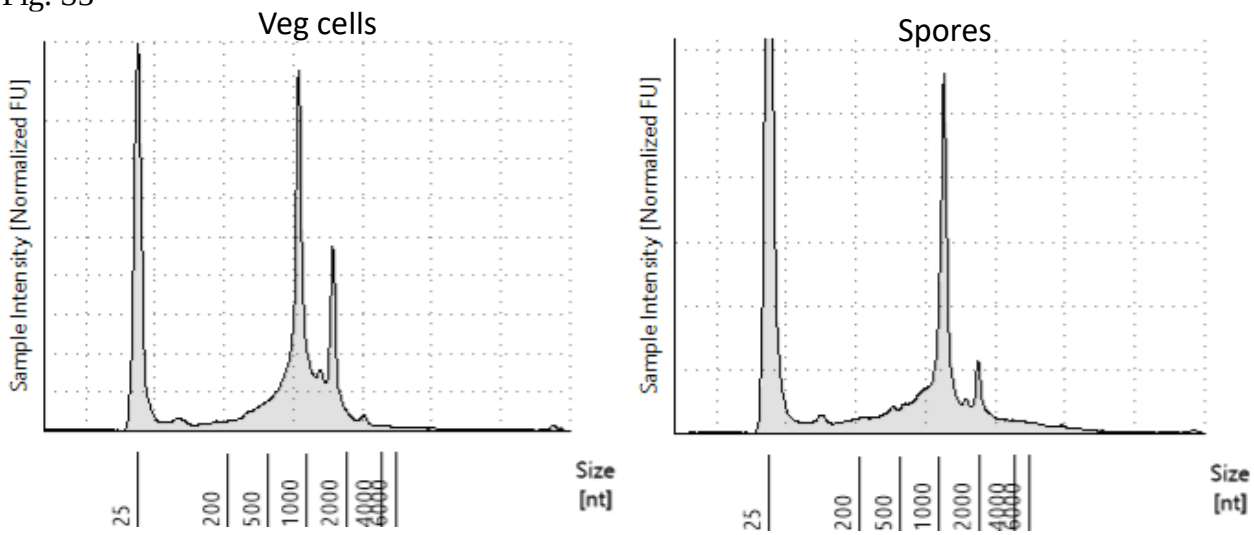

A
Plates-7d-hp (avg)
B
Liquid-7d-hp
C
Liquid-7d-lp (avg)
D
Liquid-7d-hp-dc
E
Plates-7d-hp-dc
F
Plates-1d-dc
RPKM
100000
10000
1000
100
10
1
gene

Fig. S2.


A
yoyE
B
sspI
C
ykzD
D
yrzB
E
pdhA
pdhB
pdhC
pdhD
F
adhB
yraE
yraD

126 Fig. S3



127
128