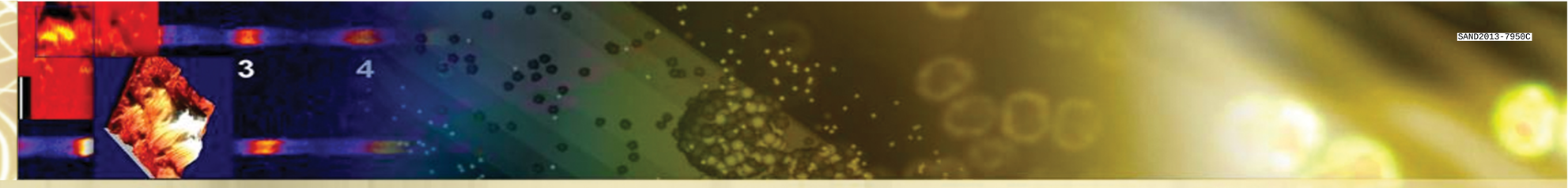
SANDIA BIOSCIENCE



An RNA-Seq analysis of the *Yersinia enterocolitica* biovar 1B transcriptome at early time points during the infection of murine macrophage cells

Zachary W. Bent[1], David M. Brazel[3], Anupama Sinha[1], Deanna J. Curtis[1], Glenn M. Young[4] and Steven S. Branda[2]

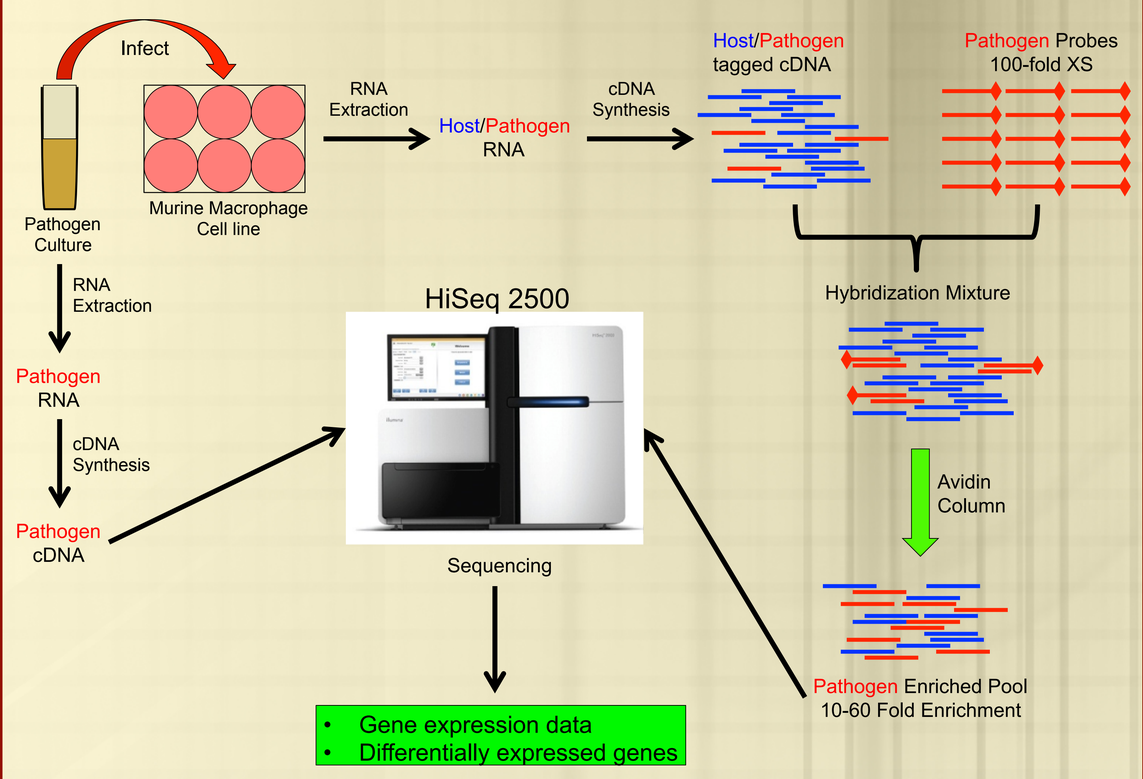
Departments of [1]Systems Biology and [2]Biotechnology and Bioengineering, Sandia National Laboratories, Livermore, CA, [3]IQ Biology Graduate Group, University of Colorado, Boulder, and [4]Department of Food Science and Technology, University of California, Davis

Abstract

Yersinia entercolitica is a zoonotic extracellular pathogen that typically causes a self-limiting gastroenteritis in humans. Among the six biovars of *Y. enterocolitica*, biovar 1B is of particular interest because it causes the most severe illness in humans and, distinct from the other biovars, is lethal to mice. Like all successful bacterial pathogens, *Y. enterocolitica* is able to rapidly respond to changing conditions as it encounters different microenvironments within the host and infects different host cell types. This ability to appropriately respond to the challenges of infection requires rapid and global shifts in gene expression patterns. In this study we analyze these shifts using a new pathogen transcript enrichment strategy followed by whole transcriptome sequencing (RNA-Seq) of *Y. enterocolitica* infecting murine macrophages. By focusing on four early time points during infection, we were able to observe the transcriptional shifts that occur as the bacteria move from log-phase growth in a nutrient rich medium at 26°C to an active infection of mammalian cells. We used growth in filter-sterilized spent tissue culture media as a control to focus our analysis on genes that are expressed upon contact with host cells. We also compared the transcriptomes of bacteria located on the surface of host cells to that of cells that have been internalized by the macrophage cells revealing numerous genes involved in intracellular survival. Finally, we analyzed the expression of biovar 1B-specific genes such as those of the plasticity zone to determine how they contribute to its exceptional virulence. This study provides insight into the spatially and temporally dynamic expression patterns of known virulence factors as well as revealing several genes encoding unknown products that likely play roles during infection. This is also the first study to analyze the complete transcriptome of *Y. enterocolitica* during an infection using next-generation sequencing technologies.

Methods

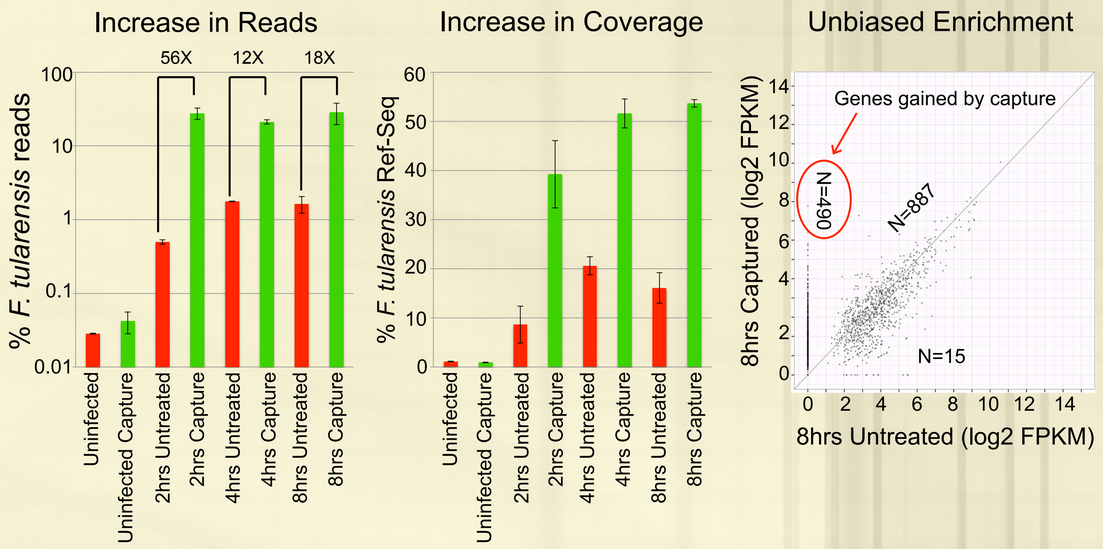
(Pathogen Capture and RNA-Seq)



Bacteria are grown to mid-log phase in a culture medium and then used to infect murine macrophage cells at an MOI of 10. RNA is extracted from the infected sample as well as the inoculum and converted to cDNA with short non-identical tags on the 5' and 3' ends of each molecule. This mixed host/pathogen cDNA pool is then combined with biotinylated probes generated against the entire genome of the pathogen of interest, insuring that all possible transcripts can be captured. The hybridization mixture is denatured and allowed to re-anneal overnight. The mixture is then passed through a monomeric avidin column to remove unbound (host) transcripts followed by elution of the probe bound (pathogen) transcripts from the column. We refer to this technique as Pathogen Capture. Libraries are sequenced using an Illumina HiSeq 2500. Gene expression data is generated for both the inoculum and the bacteria involved in the infection and compared using DESeq2 to determine genes that are differentially expressed.

Proof of Concept

(*F. tularensis* infections)



These figures demonstrate the ability of our capture technique to enrich for bacterial transcripts. Murine macrophage cells were infected with *F. tularensis* and following RNA extraction and cDNA synthesis, subjected to capture and compared to untreated controls. **A)** Percentage of reads from infected murine macrophage cells that map to the *F. tularensis* genome plotted on a log scale. Capture can increase the percentage of bacterial reads by over 50 fold. **B)** The increase in the number of bacterial reads also leads to a large increase in the percentage of the *F. tularensis* genome that is covered. C) The enrichment of bacterial transcripts is unbiased with the expression of very few genes significantly altered by the capture process. In this case, capture is also able to increase the number of genes with expression data by 490, a 55% increase in genes with only an 18 fold increase in reads.

Results

(*Y. enterocolitica* infections)

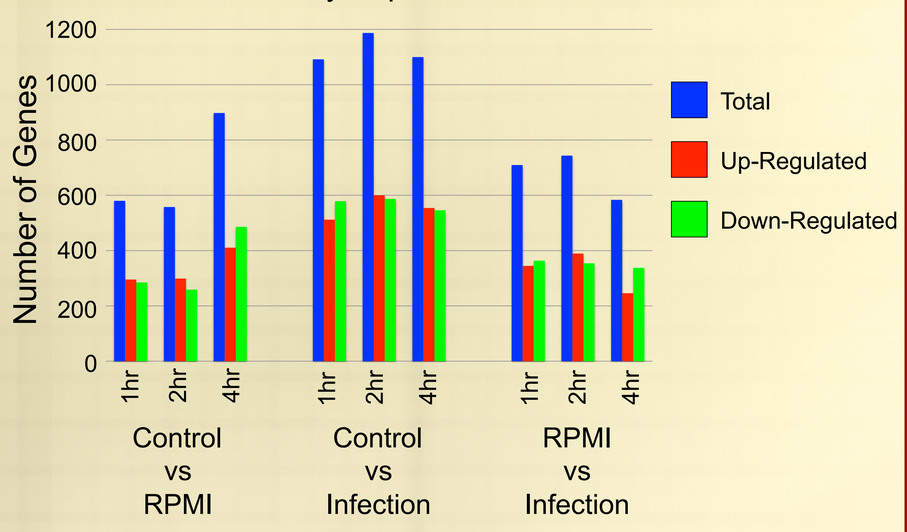
Y. enterocolitica Growth Conditions

Control = 26°C in LB w/o salt

RPMI = 37°C in filter-sterilized spent RPMI

Infection = 37°C on P388 cells in RPMI

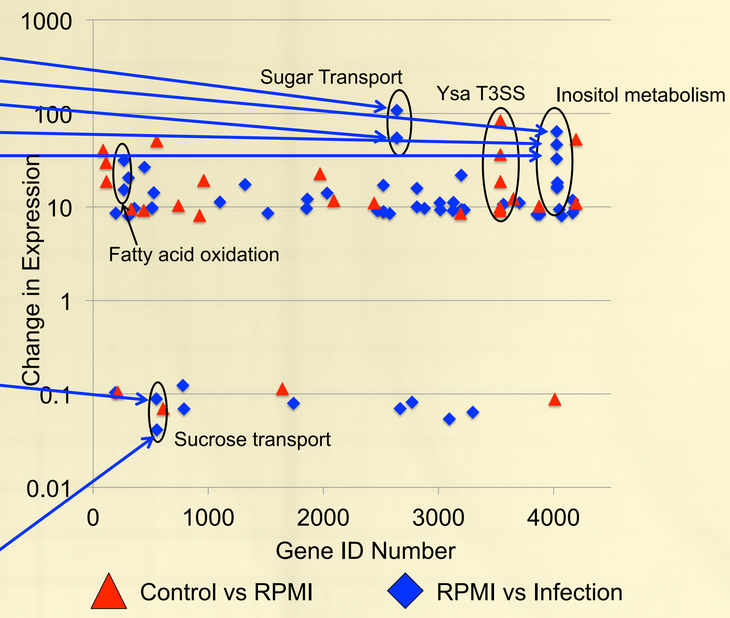
Wild type *Y. enterocolitica* strains were grown to mid-log phase and inoculated into 6-well plates of either fresh salt-free LB medium (control), filter-sterilized spent RPMI (RPMI), or centrifuged onto P388 cells (Infection) and grown at the temperatures indicated above. After RNA-Seq analysis comparing the different conditions at each time point, a significant number of differentially expressed genes were discovered. Interestingly, there was a near equal number of up and down regulated genes in each condition analyzed. Note that ~25% of all *Y. enterocolitica* genes are differentially expressed when the control transcriptome is compared to the infection.



Most Differentially Expressed Genes 1hr RPMI vs Infection

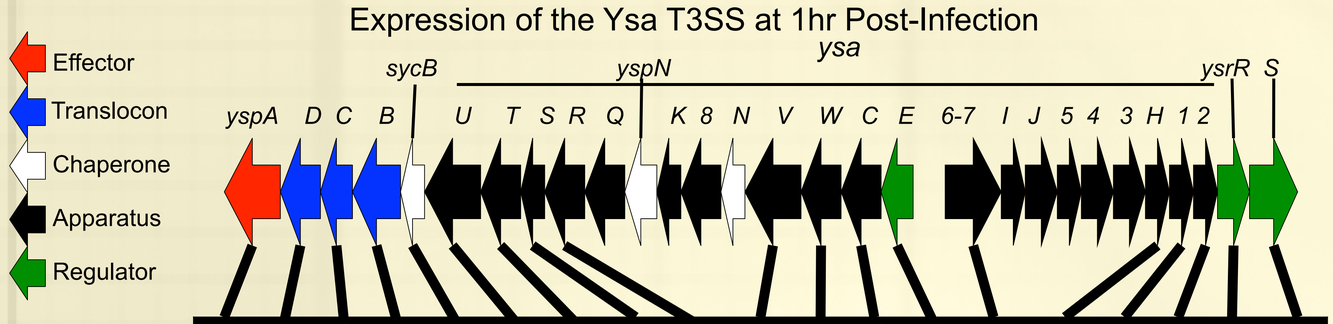
Gene ID	Name	Function	Fold Change
YE2639	-	Sugar Transporter	107.65
YE4027	-	Epi-inositol hydrolase	63.44
YE2638	-	Sugar Transport	54.79
YE4025	-	Inosose dehydratase	46.21
YE4026	*idh*	Myo-inositol 2-dehydrogenase	32.73
YE0267	*fadA*	Fatty acid oxidation	31.52
YE0447	-	Unknown membrane transporter	26.63
YE3198	*aglB*	6-phospho-α-glucosidase	21.75
YE0309	*acs*	acetyl-CoA synthetase	20.38
YE4031	-	aldehyde dehydrogenase	18.21
YE0781	*mrfC*	Fimbrial protein	-8.13
YE0195	*yigB*	Flavin mononucleotide phosphatase	-9.69
YE0552	*scrA*	Sucrose transporter	-11.30
YE2772	*hisI*	ATP pyrophosphatase	-12.19
YE1740	-	Unknown	-12.59
YE2667	-	Unknown	-14.32
YE0789	-	thiol:disulfide interchange	-14.42
YE3297	*xni*	exonuclease	-15.63
YE3092	-	Unknown	-18.47
YE0553	*scrY*	sucrose porin	-24.15

Clustering of Differentially Expressed Genes



To determine the genes that are most affected by contact with host cells, the bacteria grown in the filter-sterilized spent RPMI were compared to the infection at the 1hr time-point. The 10 genes with the largest up and down-regulation are shown in the table. Strikingly, several of the most up-regulated genes were involved in sugar transport and inositol metabolism. Inositol is commonly found within eukaryotic cellular membranes and has been implicated in the activation of the Ysc T3SS secreted effector YopJ. The graph depicts the clustering of all genes with ≥8-fold change in expression across the *Y. enterocolitica* chromosome. Although the Ysa T3SS has previously been shown to be expressed only at 26°C, the clustering analysis reveals that several genes in the system are significantly up-regulated in the filter-sterilized spent RPMI at 37°C

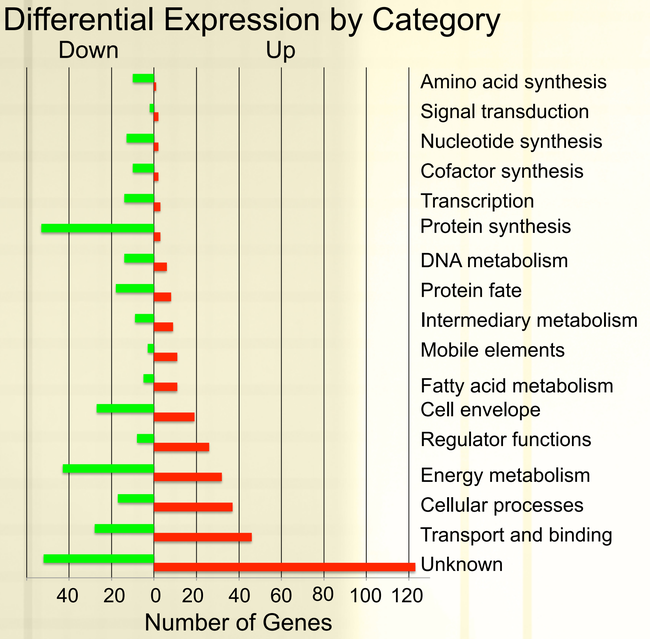
Expression of the Ysa T3SS at 1hr Post-Infection



	yspA	D	C	B	scyB	U	T	S	R	V	W	C	E	6-7	H	1	2	ysrR	S
Control FPKM	5.2	13.6	3.1	1.5	0.8	1.3	1.6	3.1	.5	0.5	0.5	0.2	1.0	2.5	1.7	14.9	15.1	19.4	6.7
RPMI FPKM	41.1	78.4	62.3	54.9	147.2	0.9	0.5	0	0.7	1.0	0.9	0.1	0	1.3	1.2	16.4	12.4	10.8	2.9
Infection FPKM	97.1	162.2	125.1	119.0	335.6	21.7	7.2	0	0	4.3	0	0.2	2.5	16.3	21.7	220.2	144.5	52.1	15.2

To further analyze the expression of the Ysa T3SS during the infection of murine macrophage cells, FPKM (fragments per kilobase of transcript per million mapped reads) values were determined for several representative genes within the *ysa* locus. While most of the genes making up the apparatus do not change much in RPMI, they tend to go up dramatically during infection indicating that the Ysa T3SS system is contact dependent. Genes encoding proteins that make up the translocon are significantly up-regulated in RPMI alone with additional increases in expression during the infection.

Differential Expression by Category



All differentially expressed genes comparing the RPMI to Infection at 1hr post-infection are grouped according to their cellular function. As would be expected, there is significant down regulation of genes involved in protein synthesis and energy metabolism during the infection. This is likely due to decreased growth rate caused by the innate immune response of the murine macrophage cells. Interestingly, there are over 120 genes with unknown function that are up-regulated in this comparison. However, it should be noted that many known virulence factors are categorized as unknown in this analysis

Conclusions

- The pathogen capture technique leads to significant enrichment of bacterial transcripts
- Using RNA-Seq it is possible to identify hundreds of differentially expressed genes during an infection
- During infection *Y. enterocolitica* up-regulates many genes involved in inositol metabolism
- The Ysa T3SS system is expressed at 37°C in a contact-dependent manner during the course of an infection
- While infecting cells Y. enterocolitica down-regulates genes invovled in bacterial replication and up-regulated virulence factors and transport/binding proteins

zbent@sandia.gov

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories