Combining Tn-seq with comparative genomics identifies proteins uniquely essential in *Shigella flexneri*

Nikki Freed[1,2,4], Dirk Bumann[1], and Olin Silander[2,3]

[1]Infection Biology, Biozentrum, University of Basel, Basel, Switzerland [2]Institute of Natural and Mathematical Sciences, Massey University, Auckland, New Zealand
[3]Computational and Systems Biology, Biozentrum, University of Basel, Basel, Switzerland [4] presenting author

Introduction

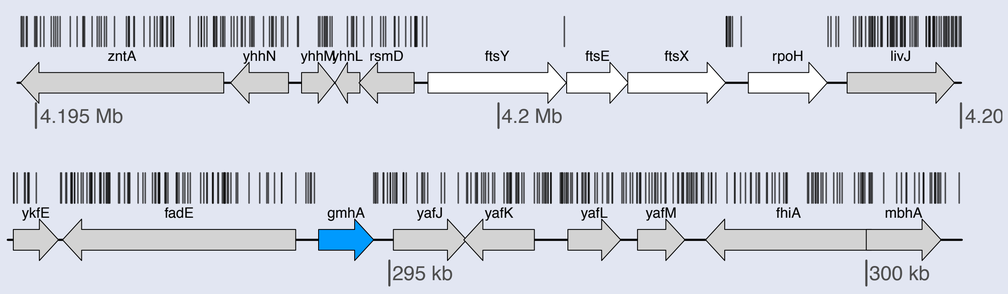
We combined transposon mutagenesis with high-throughput sequencing to **quantify gene essentiality in *Shigella flexneri***, a pathogenic bacteria.

We then compared gene essentiality in *Shigella* with the orthologous genes in *Escherichia coli K12*. These two strains share 99.5% identity in their 16S RNA genes, a common measure of relatedness.

Methods

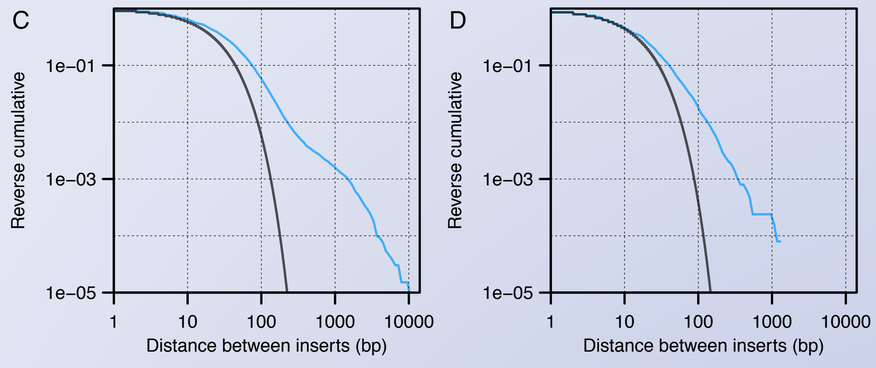
We constructed a high density transposon insertion library in the bacterial pathogen *Shigella flexneri.* We then **sequenced and mapped the locations of all transposon insertion sites** (**188,895 unique insertions**) in the genome and on the ~200Kb virulence plasmid.

Mapping the transposon insertions, we found many regions in which few or no insertions occurred.



Vertical black bars indicate transposon insertion sites across a representative area of the Shigella flexneri *genome. Genes highlighted in white are experimentally shown to be essential in* E. coli*. The lower panel shows a gene highlighted in blue, a gene experimentally shown to be essential in E. coli, yet predicted to be essential computationally.*

Quantitative analyses showed that regions containing no transposon insertions for 100bp or more were considerably enriched.

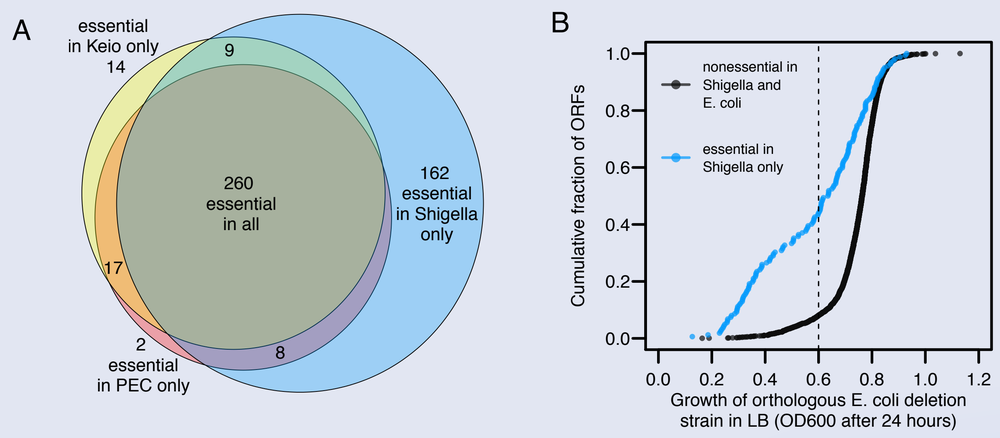


Observed (blue) and expected (negative binomial; black) insertions on the chromosome (left), and the plasmid (right).

It is likely that many of these **regions are critical for cellular growth** in *Shigella*. Indeed, we found that for many of the protein-coding genes in these regions, the orthologous *E. coli* genes are known to be essential. As “gold standards” of essentiality we used data from two studies of the effects of gene deletion on growth in *E. coli K12*: the Keio collection (Baba, Ara et al. 2006) and the PEC study (Kato and Hashimoto 2007).

Results

Our results suggest that 471 protein-coding genes in *Shigella* are critical for cellular growth in rich media. For many of the uninterrupted *Shigella* genes, the orthologous *E. coli* genes are known to be essential, or their deletion results in severe growth defects.



A) A Venn diagram showing the overlap between orthologous ORFs in E. coli *and* Shigella *in their essentiality classifications. B) Cumulative distributions showing the growth phenotypes in* E. coli *K12 for orthologous deletion strains for ORFs that are nonessential in both* E. coli *and* Shigella *(black) or that are nonessential in* E. coli *but essential in* Shigella *(blue). ORFs that we infer to be uniquely essential in* Shigella *consistently have orthologues in* E. coli *with lower growth rates (apparent as a strong left shift in the cumulative). Thus, some may not be strictly essential, but instead be required for robust growth.*

Interestingly, the data provide no evidence that any genes are essential in *E. coli* but not in *Shigella.* We found **a number of genes that are essential in *Shigella* but whose deletion in *E. coli* has no effect on growth.** Several of these genes have been computationally predicted as essential (Joyce et al. 2006); blue arrow in figure at left).

We also found two functionally related sets of genes that stand out as essential in Shigella but whose deletion has no effect on growth in E. coli. First, the genes ***proA, proB, and proC*** are uninterrupted in our *Shigella transposon library*, but their deletion causes little to no growth defect in *E. coli.* One possible explanation for this is that *Shigella* cannot efficiently transport proline, resulting in a requirement for the proline biosynthetic pathway despite growth in rich media. The second set of genes are those in the ***rfb*** operon. These all play a role in sugar nucleotide biosynthesis, suggesting that this pathway is uniquely essential in *Shigella*.

Conclusions

The data here suggest that the essential gene complement of *Shigella* flexneri differs only slightly from the closely related bacterium Escherichia coli K12.

While few, if any, genes are essential in *E. coli* but not *Shigella*, several appear essential in *Shigella* but not *E. coli*. One reason for this change s*Shigella* frequently lives as an intracellular pathogen, and may have lost some of the functional redundancy that is present in *E. coli* K12. This is not necessarily due solely to genetic drift; in some cases, this may have occurred through direct selection for increased virulence, which has resulted in the inactivation of certain genes being selectively advantageous. Finally, we note that the discrepancies in essentiality between these two bacteria may be exploited to develop antibiotics that have strain-specific effects.

.

Combining Tn-seq with comparative genomics identifies proteins uniquely essential in Shigella flexneri

Freed NE

2015-09-03