



PATTERN
GENOMICS

Serotype-correlated Markers for *Salmonella enterica* serotype Typhimurium

Daniel Fasulo, PhD.

daniel.fasulo@patterngenomics.com

(203) 533-9362

Abstract

In addition to reporting species, culture-independent detection methods ideally should report the subtype of a targeted pathogen. Serotyping is a commonly-used subtyping scheme. However, the genomic basis for serotype may involve many loci on multiple genes, making a direct PCR assay difficult to develop. While DNA sequencing may be useful to assess serotype from genomic data, it may also be possible to find other genomic markers that correlate strongly with serotypes of interest and yield a simple molecular test.

In this study, Daydreamer™ was utilized for *in silico* design of an identification assay for the *Salmonella enterica* serotype Typhimurium. A total of 265 whole-genome *S. enterica* genomes, including 24 with the Typhimurium serotype, were analyzed. Primers and probe sequences suitable for Life Technologies TaqMan® or similar chemistries were output, and subsequent BLAST analysis suggests that these designs will be sensitive and specific to the desired targets.

Overview of Daydreamer Software

Inputs

Sequences from many isolates

Label for each sample as assay target or non-target

Assay format and design constraints

No user-provided reference genome. No pre-selection of specific target genes.



**Daydreamer
Software**

Outputs

Smallest set of assay elements that meet constraints

Comparative genomic output

Daydreamer™ Project Setup

Inputs

- 12 Typhimurium whole genomes
- 242 other *Salmonella* whole genomes

Label Typhimurium serotype genomes as targets; others as non-targets.

Primer and probe constraints suitable for typical TaqMan® reaction.



**Daydreamer
Software**

Outputs

8 candidate designs, each predicted to provide perfect identification of target class.

Post-Daydreamer™ analysis

BLAST analysis and performance prediction on 12 additional *Salmonella* Typhimurium genomes

Comparative genomic output

Relationships Among Genomes

Daydreamer™ estimates genome similarities using markers that are selected to be informative for finding subtypes, and the serotypes tend to fall into separate clades of *Salmonella* by clustering with these markers. This indicates that distinct serotypes tend to have distinct genomic features as well.



Cluster with Typhimurium serotype

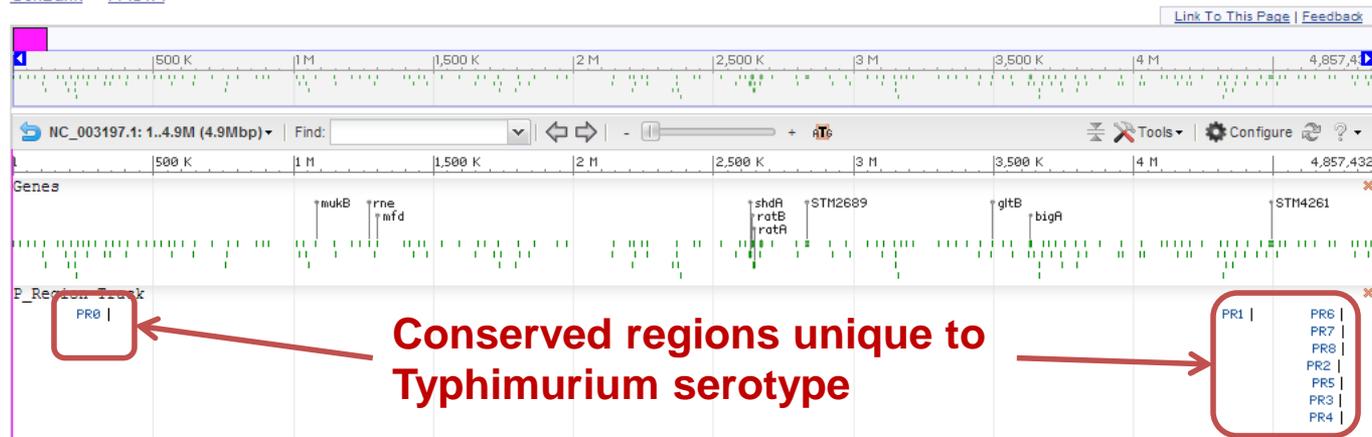
Serotype-specific Regions

In this analysis, Daydreamer™ is configured to identify regions of at least 200bp with high densities of sub-sequences that are unique to Typhimurium. These can be visualized in standard genome browsers such as the NCBI Graphical Sequence Viewer shown below.

Salmonella enterica subsp. enterica serovar Typhimurium str. LT2 chromosome, complete genome

NCBI Reference Sequence: NC_003197.1

[GenBank](#) [FASTA](#)



BLAST Analysis of One Region

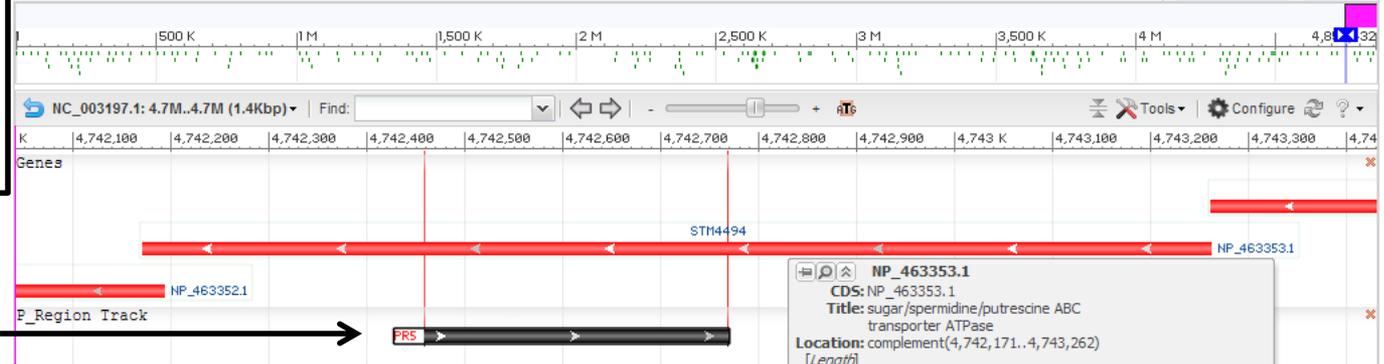
Salmonella enterica subsp. enterica serovar Typhimurium str. LT2 chromosome, complete genome

NCBI Reference Sequence: NC_003197.1

[GenBank](#) [FASTA](#)

[Link To This Page](#) | [Feedback](#)

The black region is a strongly conserved and discriminative sub-region of a gene.



NP_463353.1
 CDS: NP_463353.1
 Title: sugar/spermidine/putrescine ABC transporter ATPase
 Location: complement(4,742,171..4,743,262)
 [Length]
 Span: 1,092
 Product: 363

Download: [NP_463353.1](#)

Links & Tools
 GenBank View: [NC_003197.1 \(4,742,171..4,743,262\)](#), [NP_463353.1](#)
 FASTA View: [NC_003197.1 \(4,742,171..4,743,262\)](#), [NP_463353.1](#)
 BLAST Genomic: [NC_003197.1 \(4,742,171..4,743,262\)](#)
 Graphical View: [NP_463353.1](#)
 BLAST Protein: [NP_463353.1](#)
 BLINK Results: [NP_463353.1](#)

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium strain VNP20009, complete genome	573	573	100%	4e-160	100%	CP007804.2
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. L-3553 DNA, complete genome	573	573	100%	4e-160	100%	AP014565.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium strain 138736, complete genome	573	573	100%	4e-160	100%	CP007581.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. DT2, complete genome	573	573	100%	4e-160	100%	HG326213.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium DT104 main chromosome, complete genome	573	573	100%	4e-160	100%	HF937208.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. 08-1736, complete genome	573	573	100%	4e-160	100%	CP006602.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium var. 5- str. CFSAN001921, complete genome	573	573	100%	4e-160	100%	CP006048.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. U288, complete genome	573	573	100%	4e-160	100%	CP003836.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. 798, complete genome	573	573	100%	4e-160	100%	CP003386.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. UK-1, complete genome	573	573	100%	4e-160	100%	CP002614.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. ST4/74, complete genome	573	573	100%	4e-160	100%	CP002487.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. T000240 DNA, complete genome	573	573	100%	4e-160	100%	AP011957.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium SL1344 complete genome	573	573	100%	4e-160	100%	FQ312003.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S, complete genome	573	573	100%	4e-160	100%	CP001363.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. D23580 complete genome	573	573	100%	4e-160	100%	FN424405.1
<input type="checkbox"/> Salmonella enterica subsp. enterica serovar Typhimurium str. LT2, complete genome	573	573	100%	4e-160	100%	AE006468.1

NCBI BLAST analysis of this region and the surrounding gene confirms that both are well conserved and unique to Salmonella Typhimurium.

Representative qPCR Design

Daydreamer™ outputs an optimized set of primers and probes according to user specifications. One of the designs from the project is depicted below. These primers match 12 additional *Salmonella* Typhimurium sequences that were set aside as a test set, and NCBI BLAST does not show any likely cross-reactivity issues.

Salmonella enterica subsp. enterica serovar Typhimurium str. LT2 chromosome, complete genome

NCBI Reference Sequence: NC_003197.1

[GenBank](#) [FASTA](#)



AGGTAAACCAGATTAGGCAT

5' Primer

GTTGTTCCCTTTGTTCTGCTTACC

Probe

AAAGAGAAATACCCACAGGT

3' Primer

Conclusion

- Daydreamer™ enables rapid, comprehensive comparison of collections of bacterial genomes. The analyses shown on these slides were conducted on a commodity Intel workstation in less than 15 minutes of run time.
- The output can be visualized in typical genome browsers, highlighting primer locations, conserved genomic regions, and discriminative genomic regions.
- Serotype identification is just one of the many applications enabled by Daydreamer™. [Contact us](#) for a free consultation on how Pattern Genomics can help you achieve your assay design goals.



PATTERN
GENOMICS