Canadian Food Inspection Agency | Agence canadienne d'inspection des aliments

Canadian Food Inspection Agency

Our vision:
To excel as a science-based regulator, trusted and respected by Canadians and the international community.

Our mission:
Dedicated to safeguarding food, animals and plants, which enhances the health and well-being of Canada's people, environment and economy.

Strategy for developing a molecular subtyping tool for a foodborne bacterial pathogen using a whole genome analysis approach: the case of *Salmonella* Enteritidis

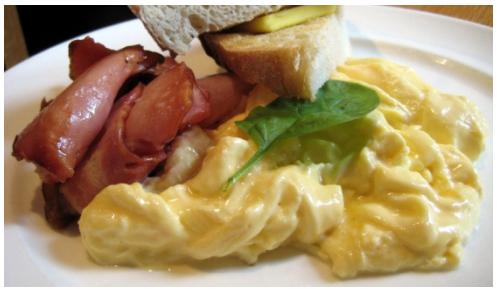
Dele Ogunremi DVM, PhD
Ottawa Laboratory Fallowfield
Dele.Ogunremi@inspection.gc.ca
343-212-0235

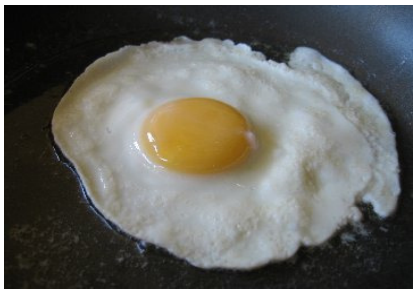
Canada

OUTLINE

- Why a new strategy?
- What is so important about *Salmonella* Enteritidis?
- What is the new strategy?
- Why SNPs?
- Conclusions
- Acknowledgment

Improving food safety





Salmonella is an important foodborne pathogen

- Highest burden of illness ranking among foodborne pathogens
- Wide distribution
- Many serovars of *Salmonella*

  (Total number of serovars = 2,579 serovars)

- *Salmonella enterica* serovar Enteritidis: most common serovar contaminating food, mainly poultry products.

Two key messages

- Whole genome analysis: a mastertool

- Diversity can still be found in a highly clonal organism such as *Salmonella* Enteritidis
  - single nucleotide polymorphism, SNPs
  - prophages

Microbial subtyping

- **Definition**

  means of identifying types of organisms within a species

  = evaluating relatedness

- **An ideal subtyping test**:

  inexpensive

  rapid

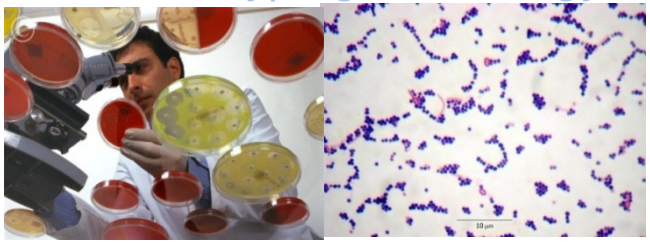
  reproducible

  discriminatory

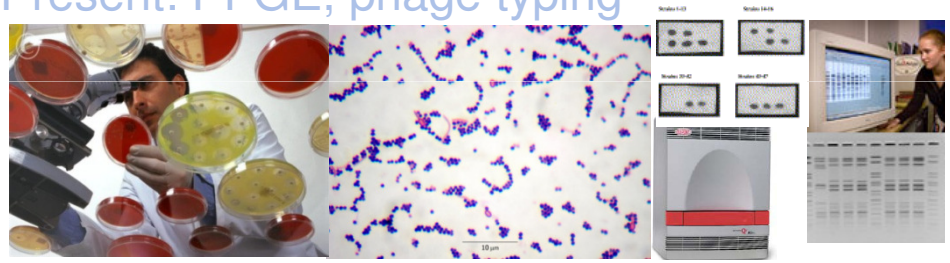
  epidemiologically concordant

The Evolution of Molecular Subtyping of Bacteria

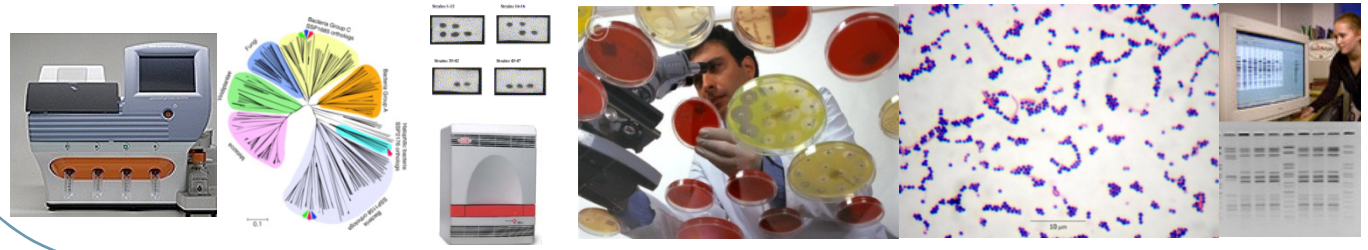
Past: Subtyping technology negligible or unavailable



Present: PFGE, phage typing



New: Whole genomes + PFGE + phage typing

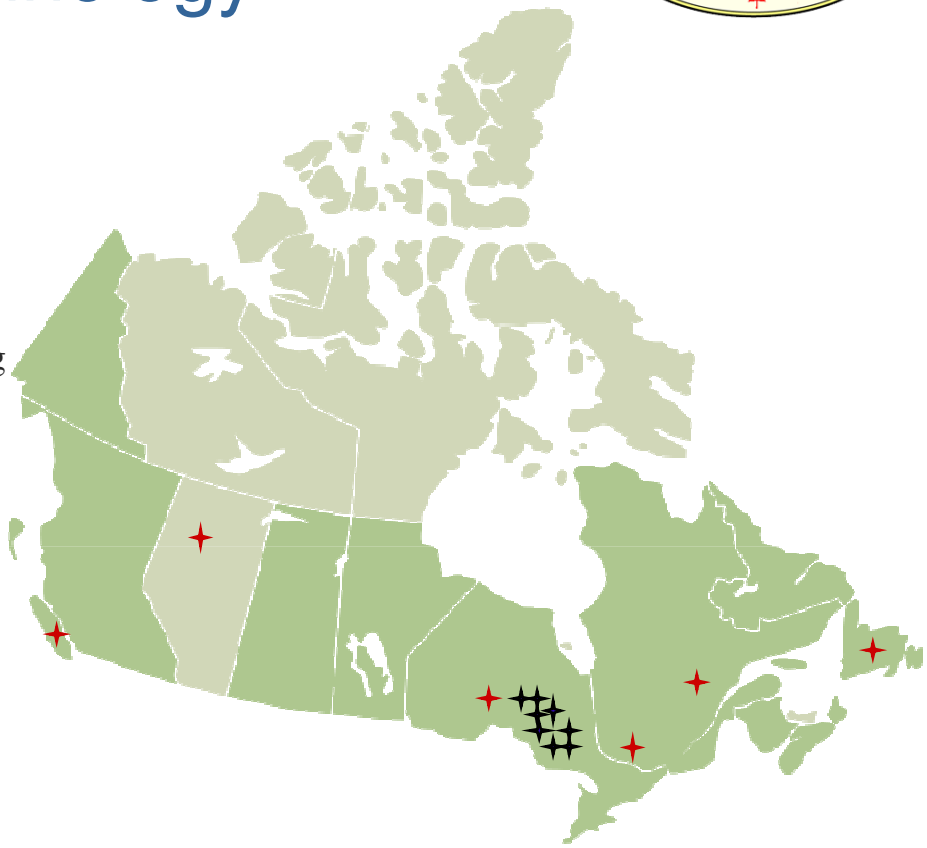


Current Microbial subtyping tools

Subtyping Approach	Assay
Phenotypic	Biotyping
	Serotyping
	Phage typing
Molecular	**Pulse-Field Gel Electrophoresis**
	Variable number tandem repeats (MLVA)
	Multi-locus sequence typing (MLST)
	PCR-RFLP
	Random Amplified Polymorphic DNA typing
	Sequence typing

PulseNet: PFGE technology

- PFGE technology has been used to build an international molecular subtyping network, **PulseNet**, in order to utilize DNA fingerprinting to detect outbreaks of foodborne diseases at the earliest possible stage
- PulseNet Canada serves as a rapid communications platform that has the ability to link laboratories both nationally and internationally when investigating sporadic and outbreak cases as well as identifying and confirming the original source of contamination



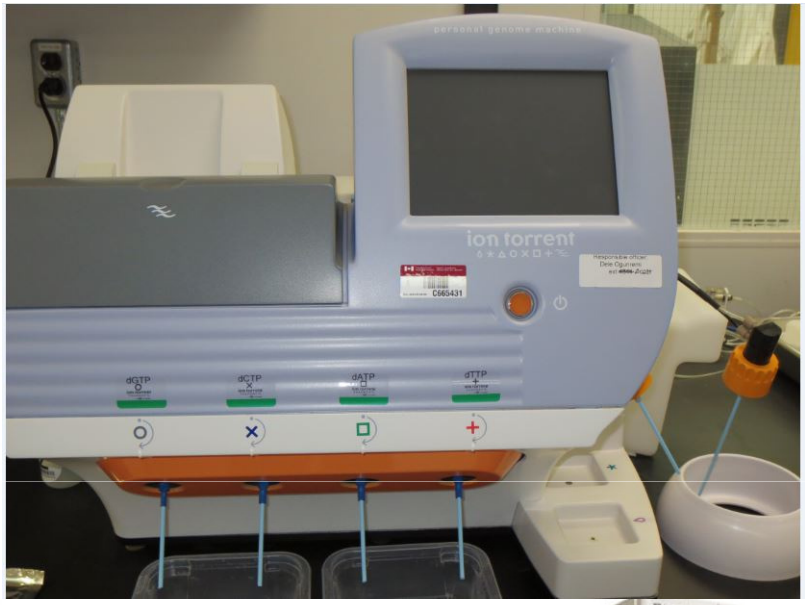
A large outbreak in one place may be obvious

A dispersed outbreak may be difficult to detect without molecular subtyping network-PulseNet

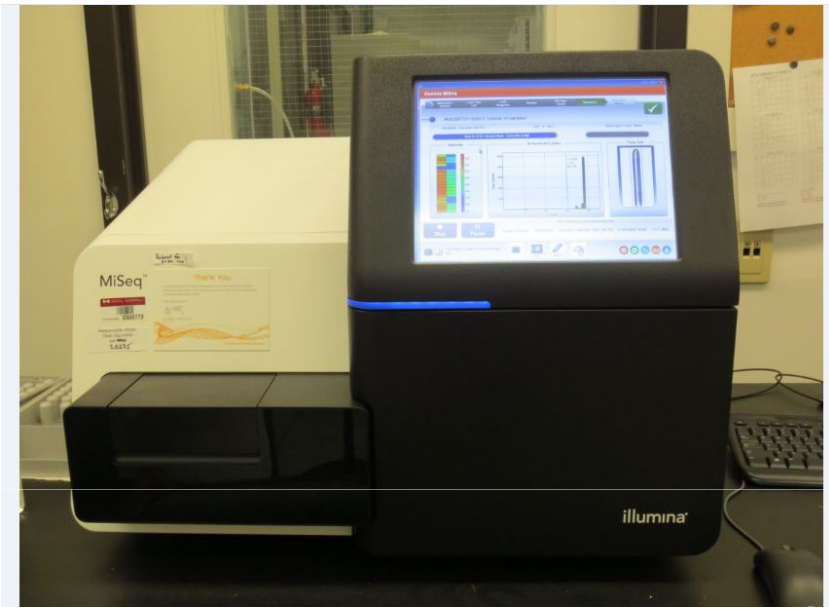
PFGE PATTERNS	Frequency	%	Rank
PRIMARY: SENXAI			
SEN XAI 0001	434	13.1	4
SEN XAI 0002	26	0.8	
SEN XAI 0003	1122	33.8	1
SEN XAI 0004	58	1.7	
SEN XAI 0006	635	19.2	2
SEN XAI 0007	137	4.1	
SEN XAI 0008	264	8.0	
SEN XAI 0009	8	0.2	
SEN XAI 0013	3	0.1	
SEN XAI 0016	3	0.1	
SEN XAI 0019	3	0.1	
SEN XAI 0021	1	0.0	
SEN XAI 0025	1	0.0	
SEN XAI 0035	2	0.1	
SEN XAI 0038	443	13.4	3
SEN XAI 0040	1	0.0	
SEN XAI 0041	57	1.7	
SEN XAI 0053	1	0.0	
SEN XAI 0054	2	0.1	
SEN XAI 0060	8	0.2	
SEN XAI 0062	8	0.2	
SEN XAI 0066	4	0.1	
SEN XAI 0068	44	1.3	
SEN XAI 0069	1	0.0	
SEN XAI 0074	2	0.1	
SEN XAI 0075	8	0.2	
SEN XAI 0076	2	0.1	
SEN XAI 0077	10	0.3	
SEN XAI 0084	3	0.1	
SEN XAI 0088	4	0.1	
SEN XAI 0092	1	0.0	
SEN XAI 0093	5	0.2	
SEN XAI 0094	2	0.1	
SEN XAI 0110	2	0.1	
SEN XAI 0111	1	0.0	
SEN XAI 0114	9	0.3	
	3315	100.0	

Salmonella Enteritidis PFGE analysis
Data source: PulseNet Canada
Analysis: CFIA

Microbial genome sequencers



Ion Torrent PGM



Illumina MiSeq

3730 Genetic analyzer

Strategy for developing a subtyping tool for *Salmonella* Enteritidis: genomics-based

Approaches

- Entire genomes
- Targeting specific parts of the genomes

Advantages

- 1. Accuracy
- 2. Cost-efficiency
- 3. Comprehensive
- 4. Extensible

Strategy for developing a subtyping tool for *Salmonella* Enteritidis

Whole genome analysis: **at least two different platforms**

454, Illumina, Ion Torrent, PacBiosciences

Assemble genome: **Reference assembly**
De novo assembly
Optical mapping

Canadian Food Inspection Agency
Agence canadienne d'inspection des aliments

Canada

Ogunremi et al. BMC Genomics 2014, 15:713
http://www.biomedcentral.com/1471-2164/15/713

RESEARCH ARTICLE Open Access

High resolution assembly and characterization of genomes of Canadian isolates of *Salmonella* Enteritidis

Dele Ogunremi[1*], John Devenish[1], Kingsley Amoako[2], Hilary Kelly[1], Andrée Ann Dupras[1], Sebastien Belanger[1] and Lin Ru Wang[3]

Abstract

Background: There is a need to characterize genomes of the foodborne pathogen, *Salmonella enterica* serovar Enteritidis (SE) and identify genetic information that could be ultimately deployed for differentiating strains of the organism, a need that is yet to be addressed mainly because of the high degree of clonality of the organism. In an effort to achieve the first characterization of the genomes of SE of Canadian origin, we carried out massively parallel sequencing of the nucleotide sequence of 11 SE isolates obtained from poultry production environments (n = 9), a dam and a chicken, assembled finished genomes and investigated diversity of the SE genome.

Results: The median genome size was 4,678,683 bp. A total of 4,833 chromosomal genes defined the pan genome of our field SE isolates consisting of 4,600 genes present in all the genomes, i.e., core genome, and 233 genes absent in at least one genome (accessory genome). Genome diversity was demonstrable by the presence of 1,360 loci showing single nucleotide polymorphism (SNP) in the core genome which was used to portray the genetic distances by means of a phylogenetic tree for the SE isolates. The accessory genome consisted mostly of previously identified SE prophage sequences as well as two, apparently full- sized, novel prophages namely a 28 kb sequence provisionally designated as SE-OLF-10098 (3) prophage and a 43 kb sequence provisionally designated as SE-OLF-10012 prophage.

Conclusions: The number of SNPs identified in the relatively large core genome of SE is a reflection of substantial diversity that could be exploited for strain differentiation as shown by the development of an informative phylogenetic tree. Prophage sequences can also be exploited for SE strain differentiation and lineage tracking. This work has laid the ground work for further studies to develop a readily adoptable laboratory test for the subtyping of SE.

Keywords: *Salmonella* Enteritidis, Genomes, Core, Accessory, Single nucleotide polymorphism, Subtyping, Tracking

Background

Salmonella Enteritidis (SE) has emerged as the most commonly isolated serovar of foodborne *Salmonella* in humans over the last two decades [1-3]. SE belongs to a larger group of pathogens known as non-typhoidal *Salmonella* which ranked the most hazardous when a number of health indices were used to assess the 14 most burdensome foodborne bacteria, viruses and parasites causing diseases in humans [4,5]. In Canada, the proportion of human salmonellosis caused by SE increased from 13% in 2003 to 38% in 2010 [6]. In the US, an outbreak of SE in 2010 resulted in an estimated 1,939 human illnesses (http://www.cdc.gov/salmonella/enteritidis/) and the largest egg recall in the country's history involving over 500 million shell eggs (http://www.fda.gov/Safety/Recalls/MajorProductRecalls/ucm223522.htm [7]).

Early comparative analysis of SE with the serovar Typhimurium, the latter being one of the best studied *Salmonella* serovars because of its enduring importance as a human pathogen and wide host range among

* Correspondence: Dele.Ogunremi@inspection.gc.ca
[1]Ottawa Laboratory Fallowfield, Canadian Food Inspection Agency, 3851 Fallowfield Road, Ottawa, Ontario K2H 8P9, Canada
Full list of author information is available at the end of the article

BMC Genomics 2014, 15: 713

Pan genome of *Salmonella* Enteritidis

- Pan genome: entire set of the genes within the species or serovar
  - Core genome: common to all members of the species or serovar
  - Accessory genome: genes present in at least two strain but absent in at least one strain

Genomes: core vs accessory

	Number of genes (%)	
	Salmonella Enteritidis (Ogunremi et al., 2014 *BMC Genomics* 15: 713)	*Listeria monocytogenes* (Deng et al., 2010 *BMC Genomics* 11:500)
Core genome	4,600 (95%)	2,456 (61%)
Accessory genome	233 (5%)	1,596 (39%)
Total genome	**4,833**	**4,052**

Microbial subtyping tools: old and new

Current	New Generation
Biotyping	SNP-based (whole genomes or PCR)
Serotyping	Prophage
Phage typing	CRISPR
Pulse-Field Gel Electrophoresis	Pseudogenes
Variable number tandem repeats (MLVA)	
Multi-locus sequence typing (MLST)	
PCR-RFLP	
Random Amplified Polymorphic DNA typing	
Sequence typing	

Strategy for developing a subtyping tool for *Salmonella* Enteritidis

Whole genome analysis: at least two different platforms

454, Illumina, Ion Torrent, PacBiosciences

↓

Assemble genome: Reference assembly
De novo assembly
Optical mapping

↓

SNPs

Strategy for developing a subtyping tool for *Salmonella* Enteritidis

Whole genome analysis: at least two different platforms

454, Illumina, Ion Torrent, PacBiosciences

↓

Assemble genome: Reference assembly
De novo assembly
Optical mapping

↓

SNP

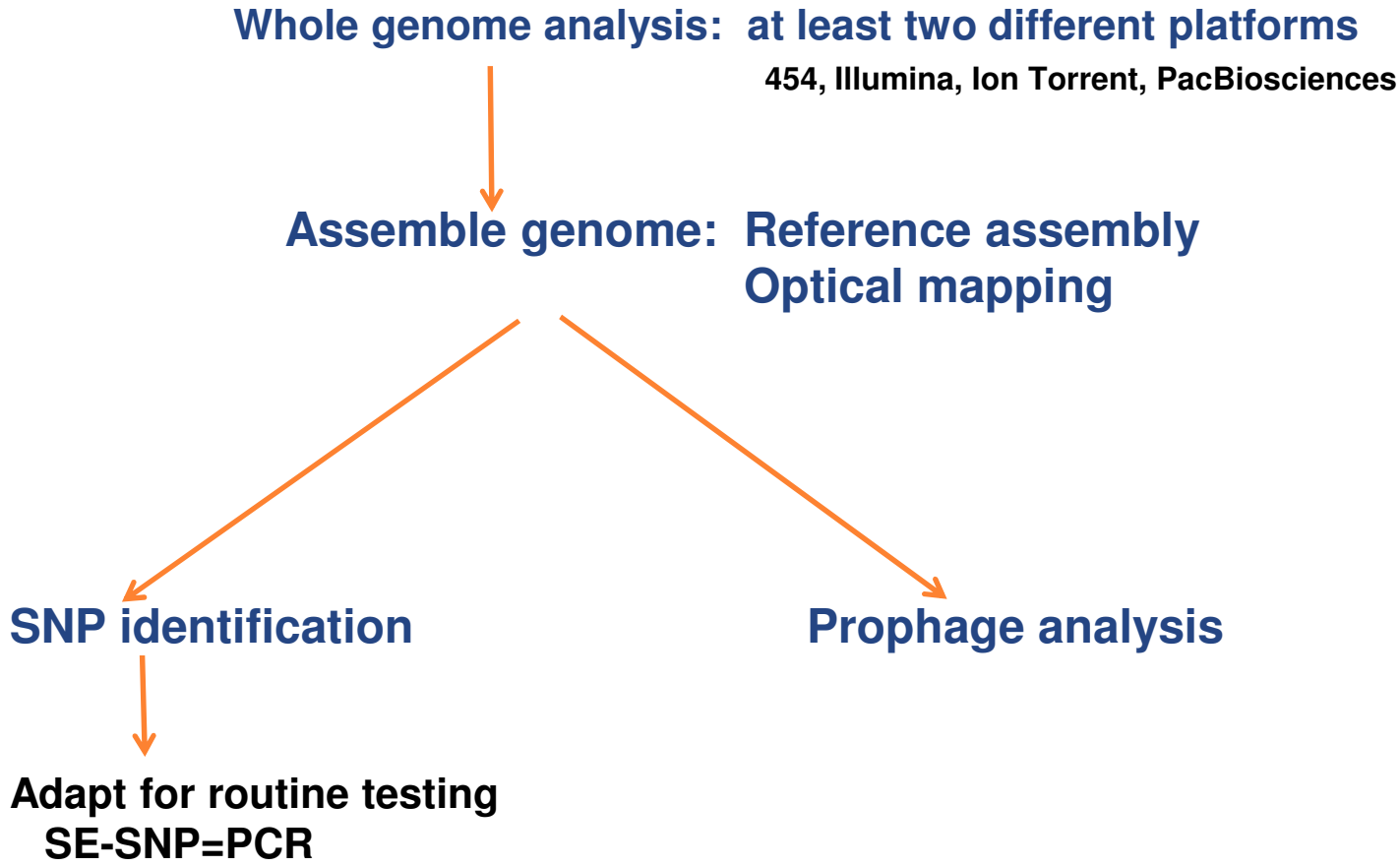
↓

Adapt for routine testing: SE-SNP-PCR

SNP PCR test is highly discriminatory

			SNP CLADES										
			I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
PFGE	Frequency	Rank												
XAI.0001	1		1											
XAI.0003	28	1st		7		1	5	1	5	4	1		1	3
XAI.0006	6	2nd			5				1					
XAI.0007	1										1			
XAI.0009	4		1					2					1	
XAI.0025	2		2											
XAI.0026	1		1											
XAI.0038	5	3rd				5								
XAI.0076	3											3		
XAI.0214	4		2	2										
	55													

Strategy for developing a molecular typing tool for *Salmonella* Enteritidis

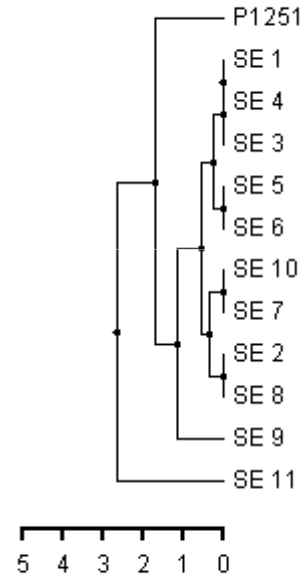


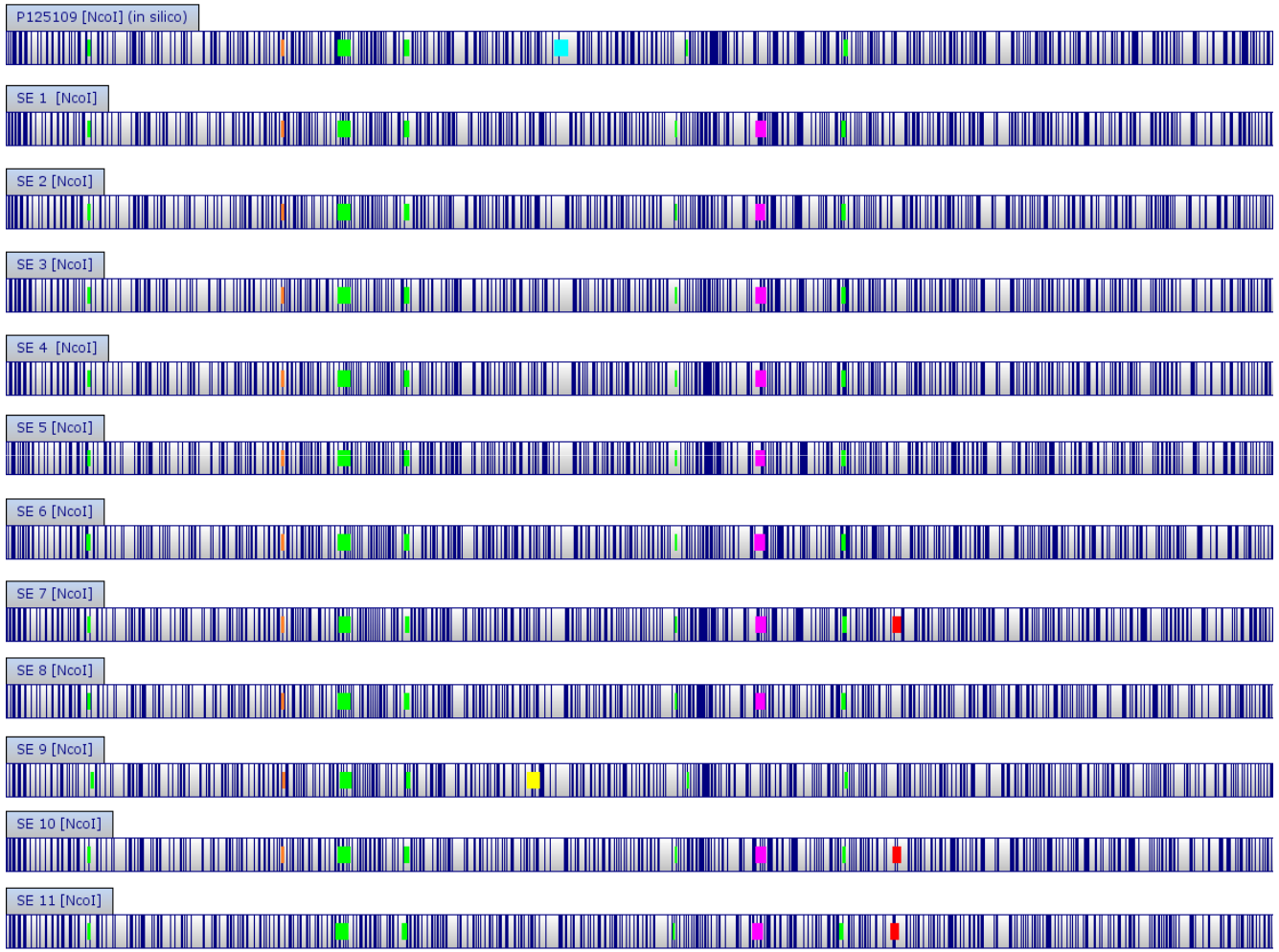
Prophages of *Salmonella* Enteritidis

S. Enteritids P125109 phage type 4 (Windhorst, 2010; Thomson *et al.*, 2008)

	Prophage designation	Size (bp)
	Φ SE 10	8,186
	Φ SE 12/12A	17,753
	Φ SE 14	12,642
	Φ SE 20	40,664
	Total	79, 245

Map Similarity Cluster using Furthest Neighbor

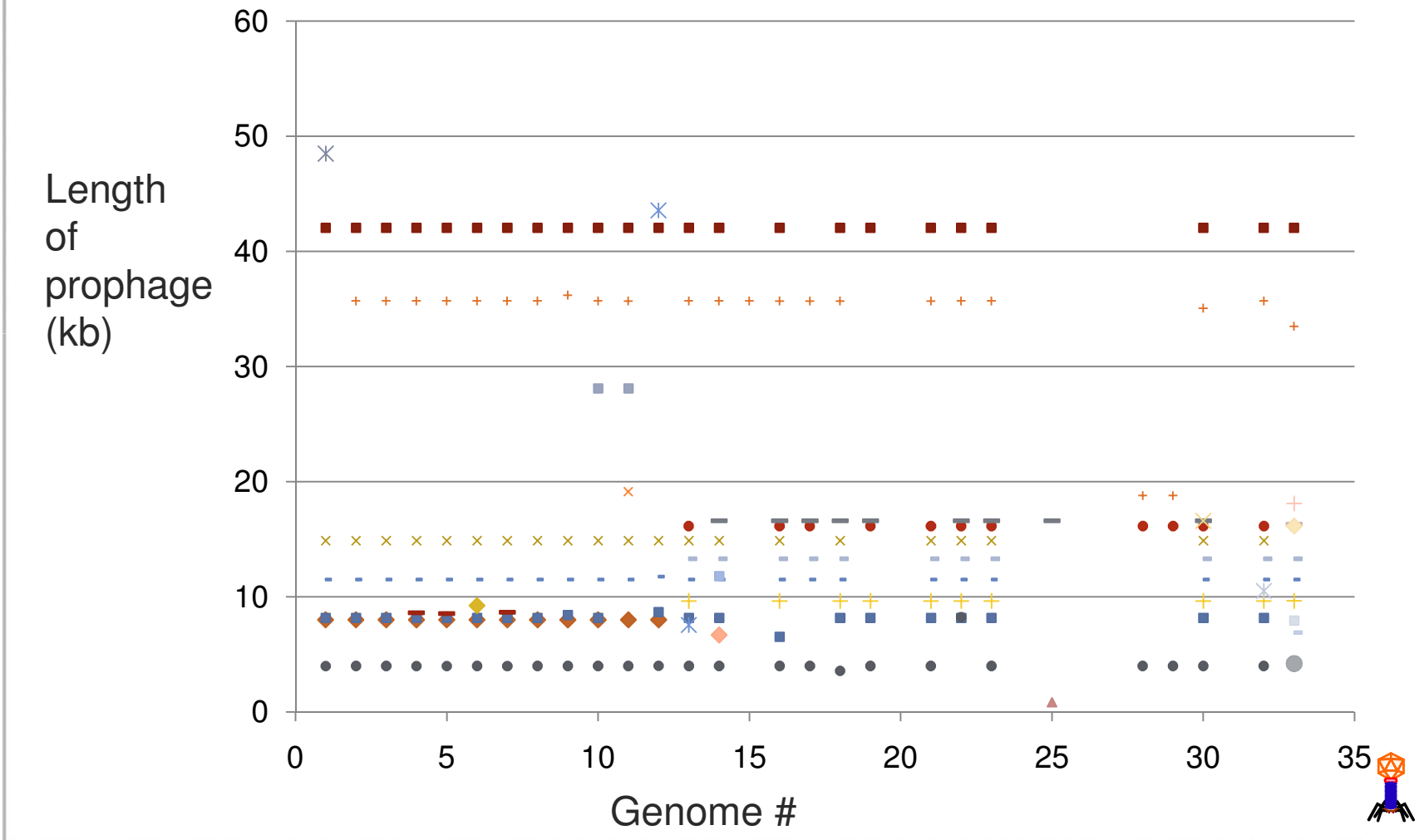




Number and sequence lengths of *Salmonella* Enteritidis prophages

	Salmonella Enteritidis genomes (Reference P125109 and SENT 1-11)											
Prophage	P125109	1	2	3	4	5	6	7	8	9	10	11
1	8,007	8,007	8,007	8,007	8,007	8,007	8,007	8,006	8,007	8,007	8,007	8,007
2	8,168	8,168	8,168	8,168	8,168	8,168	8,168	8,167	8,423	8,168		8,677
3	42,042	42,042	42,042	42,042	42,042	42,041	42,042	42,041	42,041	42,041	42,040	42,042
4	14,877	14,877	14,877	14,877	14,877	14,877	14,877	14,876	14,877	14,877	14,877	14,877
5	48,486											
6	3,987	3,994	3,994	3,987	3,987	3,994	3,987	3,993	3,994	3,994	3,994	4,001
7		35,705	35,705	35,705	35,705	35,705	35,705	35,705	36,209	35,705	35,704	
8	11,502	11,502	11,502	11,501	11,501	11,502	11,501	11,501	11,500	11,503	11,502	11,757
9				8,609	8,551		8,660					
10						9,241						
11										28,108	28,108	
12											19,137	
13												43,544
TOTAL (bp)	137,069	124,295	124,295	132,896	132,838	133,535	132,947	124,289	125,051	152,403	163,369	132,905

Prophages of *Salmonella* Enteritidis

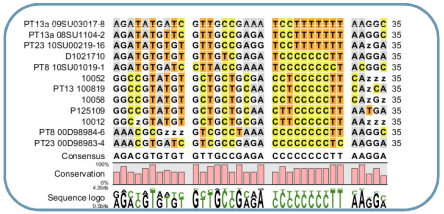


SUMMARY

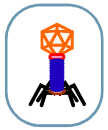
- Why a new strategy?
- What is so important about *Salmonella* Enteritidis?
- What is the new strategy?
- Why SNPs?
- Conclusion

Diversity in the Pan genome of *Salmonella* Enteritidis

- Pan genome: entire set of the genes within the species or serovar
  - Core genome: SNPs



  - Accessory genome: prophages



Acknowledgements

CFIA: Ontario Laboratory Network
Andree Ann Dupras
Hilary Kelly
Sebastien Belanger
Linru Wang
John Devenish
Susan Nadin-Davis
Burton Blais
Olga Andrievskaia

OLF PFGE Unit
Ray Allain
Jennifer Hazelwood
Chris Grenier

OLN Management
Karen Jessett
Ray Theoret
Stephen Norman

CFIA Lethbridge
Kingsley Amoako
Kristen Hahn
Noriko Goji

McGill University
Ken Dewar

Genome Quebec
Alfredo Straffa
Frederick Robidoux
Mathieu Bourgey

Public Health Agency, Lethbridge
Ed Taboada

University of Saskatchewan
Musangu Ngeleka

Public Health Agency, NML
Morag Graham
Gary Van Domselaar
Shaun Taylor
Matt Gilmour (Health Services)

Public Health Agency, Guelph
Roger Johnson
Kim Ziebell

University of Guelph
Shu Chen
Saleema Saleh-Lakha

National Research Council
Luke Masson


Canada

Pairwise SNP counts

	10052	pt8_00d	pt2300d	pt810su	pt13a09	pt13a08	pt2310su	pt13100819	10058	10217	10012
P125109-PT4	589	596	596	604	604	602	614	586	591	597	886
10052		307	304	311	311	304	282	101	70	289	826
pt8_00d			27	207	200	201	247	243	269	246	900
pt2300d				196	193	192	240	238	262	238	900
pt810su					65	69	117	248	270	78	904
pt13a09						12	60	245	270	112	903
pt13a08							59	242	263	105	898
pt2310su								221	241	113	886
pt13100819									60	258	864
10058										252	855
10217											913

PROCEEDINGS

CANADIAN

SALMONELLA ENTERITIDIS CONTROL

SYMPOSIUM AND WORKSHOP

Vancouver, British Columbia

December 1-2, 2010

Prepared for:

The SE Symposium and Workshop Steering Committee

by:

Wylie Thomas and Warren Wilson
Intersol Group

Pan genome of *Salmonella* Enteritidis

Pan genome: entire set of the genes within the species or serovar

- Core genome
- Accessory genome