Phenotype - genotype correlation of *Salmonella enterica* to identify markers associated with colonization/invasion in chickens, and antibiotic resistances in the poultry production environment

Jean Guard, DVM, PhD
Veterinary Medical Officer
USDA, ARS, National Poultry Research Center
Athens, GA
Jean.guard@ars.usda.gov

FROM USDA

- Craig Parker
- Cesar Morales
- Roxana Sanchez-Ingunza
- Tod Stewart
- Richard Gast
- Kim Ingram
- Joyce Jacks
- Eric Line
- Nathan Likens

FROM University of Georgia

- Russ Carlson
- Mahbubur Rahman
- Kokila Asokan

FROM FDA CFSAN

- Guijie Cao
- Marc Allard
- Yan Luo
- Guodong Zhang
- Jie Zheng
- Eric Brown

University of Pennsylvania and/or Pennsylvania Department of Agriculture

- David Henzler
- Sherrill Davison

FROM Washington State University, Pullman

- Devendra Shah
- Kim Chiok
- Jake Elder

FROM Mississippi State University

- Martha-Pulido Landinez

FROM University of California, Irvine

- Michael McClelland

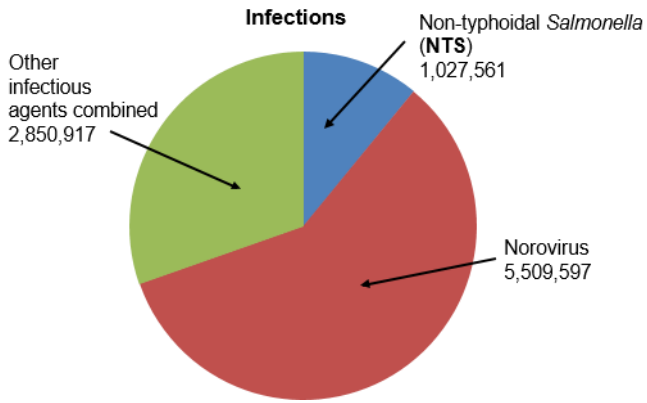
Outline

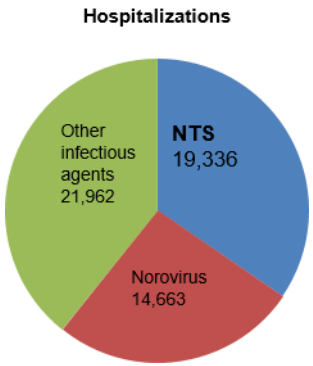
Defining virulence for foodborne *Salmonella enterica* subspecies I vs host-restricted serotypes

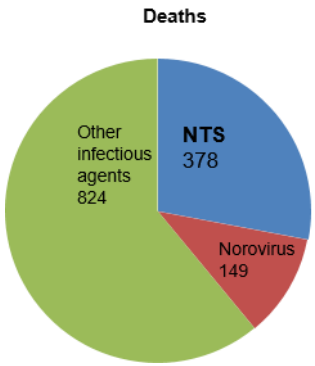
Collaborative work with Devendra Shah and colleagues on SPI-13 & 14, *ksgA* of serotype Enteritidis (SE), and acid resistance (collaboration with R Gast)

Genes observed pre- and post-NGS linked to egg contamination by SE (*sefD* & *mocR*-like HTH regulator), collaborators FDA and USDA student Nathan Likens

Infectious agents responsible for human food-borne illnesses in the US (2010)







Slide courtesy of Jake Elder, Washington State University, Pullman, WA

ARS PROJECT:

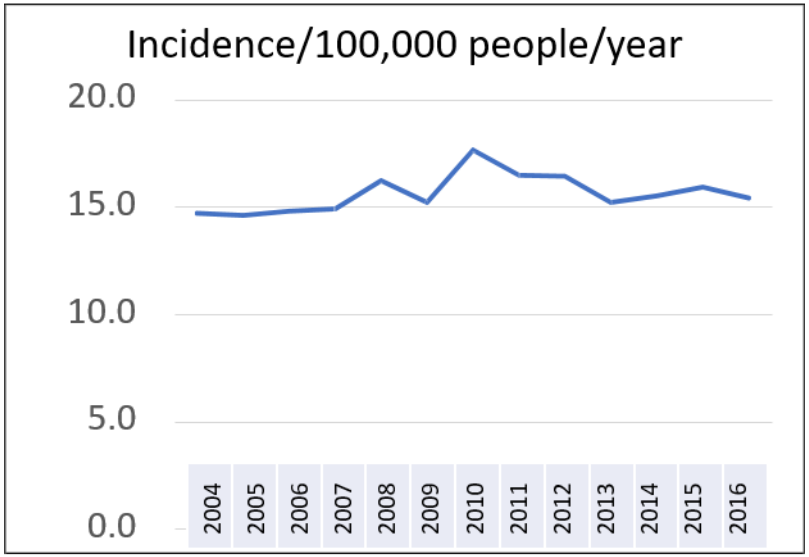
Reduction of Invasive *Salmonella enterica* in Poultry through Genomics, Phenomics and Field Investigations of Small Multi-Species Farm Environments

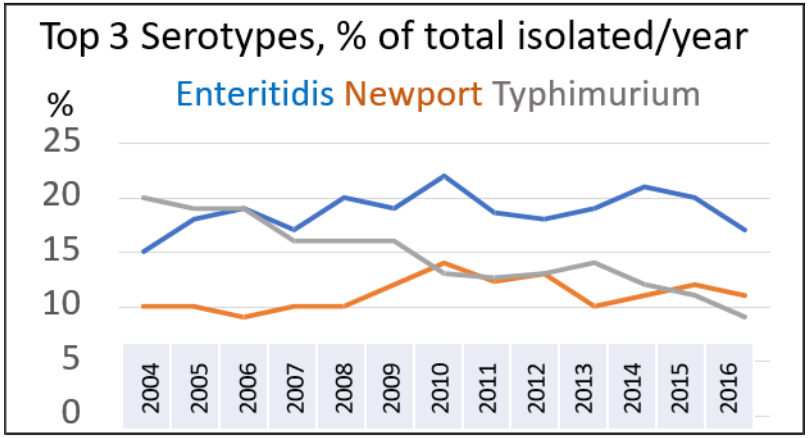
Jean Guard & Michael Rothrock

PER YEAR *Salmonella enterica* causes:

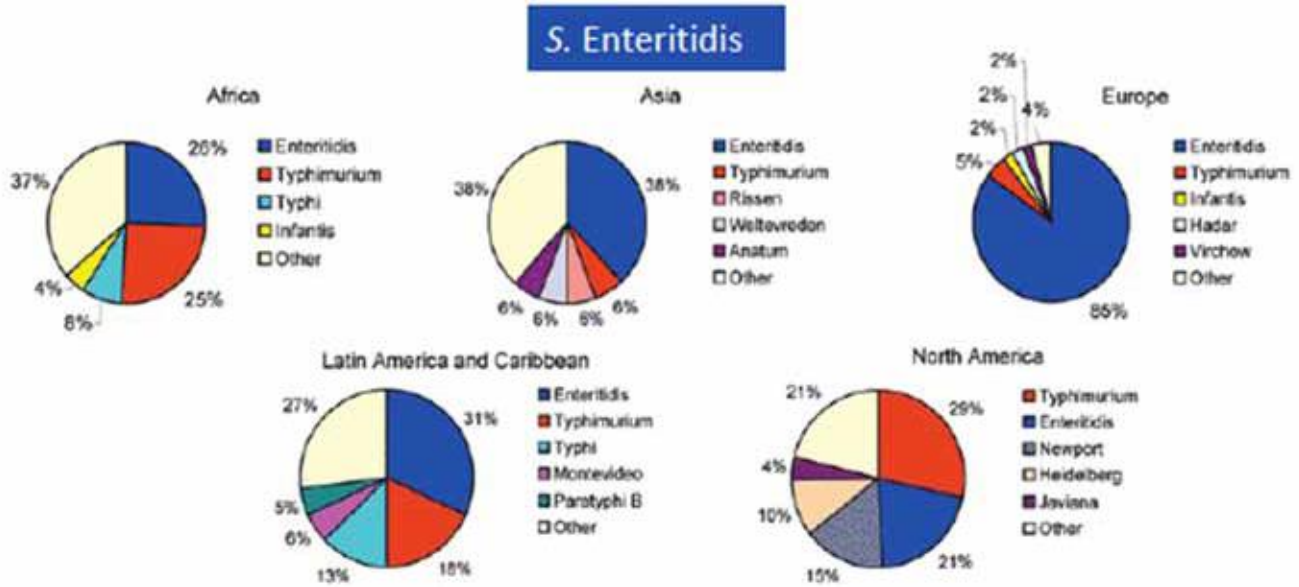
- 1.2 million illnesses per year in the US
- 23,000 hospitalizations
- 450 deaths
- Reduction in illnesses has plateaued the past several years in the US and Europe (per 100,000 laboratory confirmed cases).

Source of data for graphs:
https://www.cdc.gov/foodnet/index.html





Principle *Salmonella* serotypes causing human disease worldwide



Galanis et al., 2006

Slide courtesy of Jake Elder, Washington State University, Pullman, WA

Defining virulence for foodborne *Salmonella enterica* subspecies I as a species

- Can contaminate multiple foodstuffs, though some preferences exist among serotypes,
- Has a broad host range, though host preferences do occur among serotypes,
- Shows high growth density in nutrient broth relative to host restricted serotypes (e.g. 10exp9 vs 10exp6 CFU/ml)
- Is complement resistant, but O-antigen structure and mass varies between serotypes (Enteritidis produces HMM O-antigen),
- Is generally motile, but motility varies, ranging from non-motile to phase variation between 2 flagellin genes,
- Metabolizes a broad range of Carbon & Nitrogen sources, which contributes to successful competition in the gut,
- Able to attach to and invade gut epithelium, causing cytoskeletal rearrangements (Kentucky does not do this)
- Causes inflammatory diarrhea, (Kentucky does not do this),
- Might invade beyond GI tract, but not a requirement
  - Typhimurium, 6%; Enteritidis, 6%; Heidelberg, 11%

- Serotypes ENTERITIDIS, TYPHIMURIUM, HEIDELBERG, NEWPORT, INFANTIS are defined by these parameters
- Serotypes GALLINARUM, TYPHI and KENTUCKY are not

Host restricted pathology is beyond inflammatory gastroenteritis

- Replication & evasion of immune response,
- Uptake by macrophages, with/without replication,
- Migration & dissemination of phagocytic cells (systemic invasion),
- Wide distribution in the lymphoid and nonlymphoid tissues,
- Severe systemic illness, sepsis, organ failure, death can result,
- Similar processes are involved in egg contamination by Enteritidis although hens do not usually become ill following dissemination into organs and lymphoid system, and in some ways, respond favorably to low level infection.
- There is some evidence that the unique properties of capsular like O-antigen contribute to the overall ability of Enteritidis to contaminate and survive in the internal contents of eggs without compromising the avian host.

SUMMARY

Across over 1500 serotypes of *Salmonella enterica* subspecies I, there is a spectrum of outcomes, ranging from increased performance in one host to death in another.

The most frequent outcome <u>we observe</u> is inflammatory gastroenteritis.

HO: There is a continuum of evolution impacting characteristics of the bacterium and how it interacts with the characteristics of a particular host and the environment of that host that determines where any one serotype most frequently falls on the spectrum.

Salmonella pathogenicity islands (SPIs)

Characteristics of SPIs

- Gene clusters
- Acquired horizontally
- Lower GC content
- Located near tRNA genes

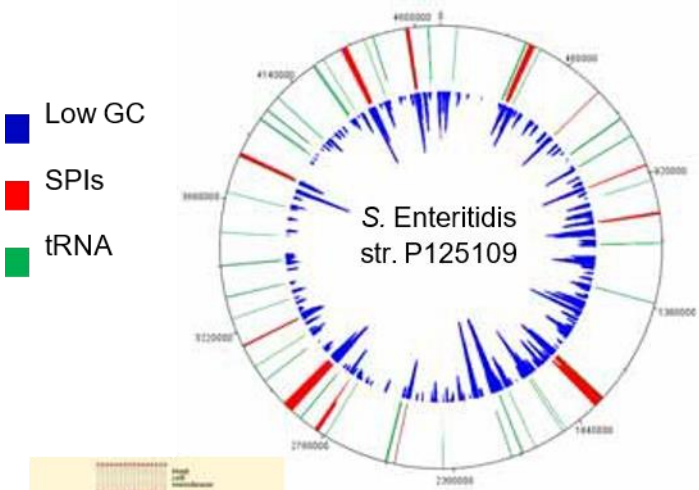
SPI1-SPI23 have been annotated

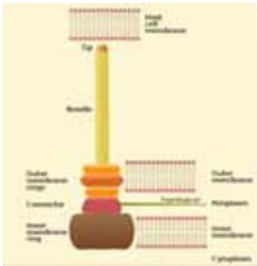
SPI1 and SPI2

- Encode T3SS and effectors
- Best characterized

SE genome contains 14 SPIs

- 1-6, 9, 10, 12, 13, 14, 16, 17, 19





The laboratory of Devendra Shah, Washington State University, Pullman, WA characterizes SPI13 & SPI14

Distribution of SPI13 and SPI14 among *Salmonella enterica* subspecies I

Group	Serotype	Hosts	SPI13	SPI14
NTS				
	Enteritidis	Multiple	Complete	Complete
	Typhimurium	Multiple	Complete	Complete
	Other NTS	Multiple	Complete	Complete
Typhoidal				
	Typhi	Humans	Partial	Absent
	Paratyphi A	Humans	Partial	Absent
	Gallinarum	Avian	Complete	Complete

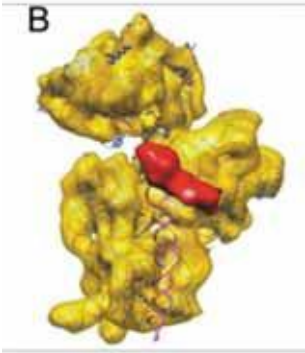
RESULTS FROM MUTATIONAL ANALYSIS

- Mutant and wildtype did not differ in pathogenicity in day-old chicks (Elder et al, 2016, Gut Pathog)
- Mutant did cause more organ invasion and pathology in streptomycin-treated mice (animal model for human inflammatory gastroenteritis),
- No major differences in metabolism were found between mutant and wildtype.
- SPI-13 may play a role in host adaptation and/or restriction of *Salmonella* serovars
- SPI-14 is implicated in serotypes Enteritidis and host-restricted Gallinarum in systemic infection using a Caco-2 cell infection model (Shah et al, 2012 I&I)

ksgA (ribosomal maturation, DNA mis-match repair, KaSuGAmycin resistance)

- Oral challenge of day old chicks with wildtype and mutant, 10exp8, samples collected at 24h, 48h, 4 days, 8 days and 12 days post-infection (some experimental variation in collection dates between 2 trials) (Shah laboratory)

- In comparison to wildtype, ksgA::Tn5 mutant had:
  - Reduced intestinal colonization and organ invasiveness,
  - Reduced respiratory activity with respect to several C, N, P, and S sources as determined by phenotype microarray (PM),
  - Severe growth defect in high-osmolarity (6.5% NaCl) at 37 and 42 C (PM),
  - Resistance to kasugamycin (aminoglycoside targeting ribosomal function) but increased resistance to chloramphenicol (inhibits protein synthesis)

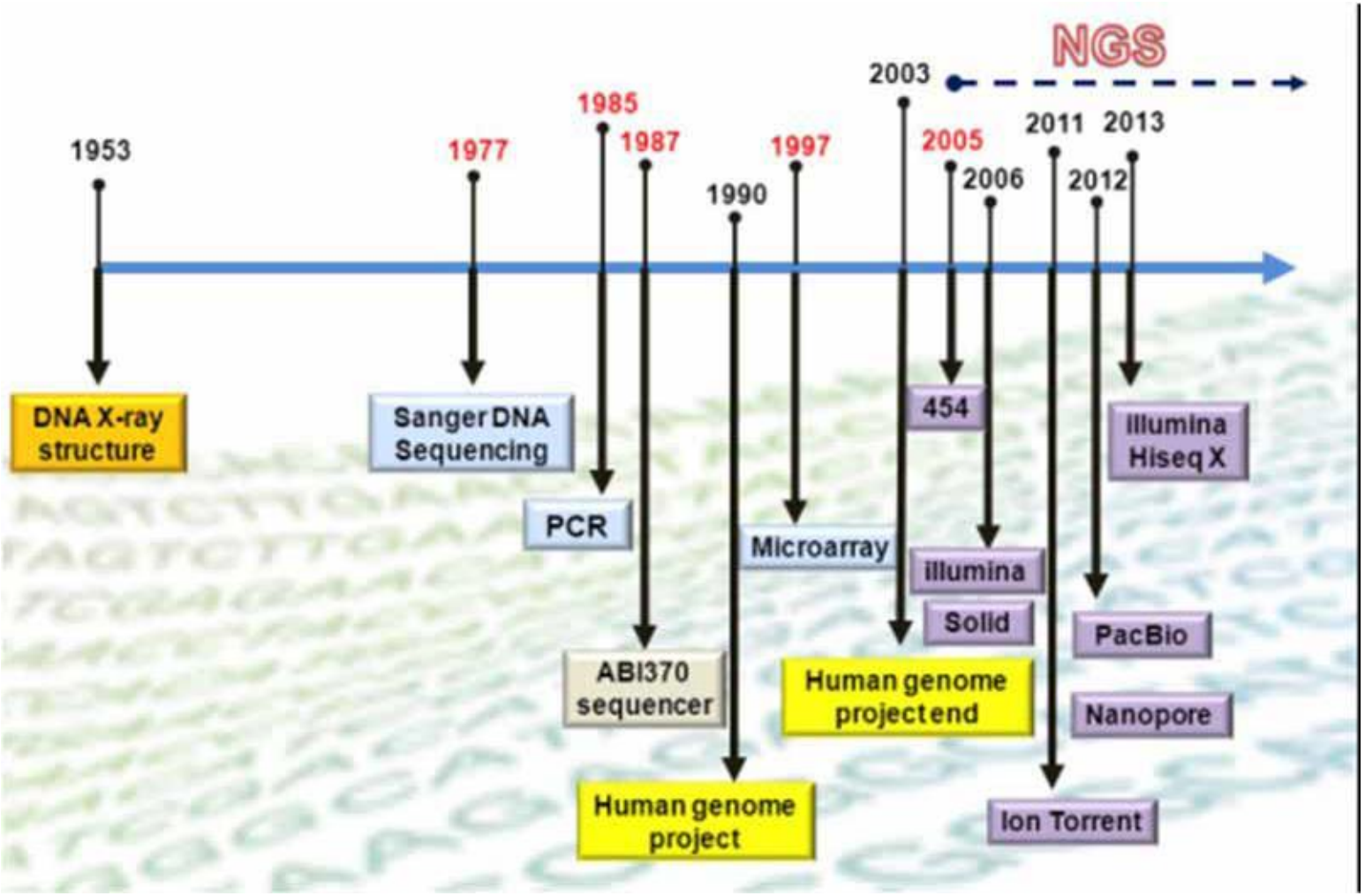


Acid resistance was not linked to virulence

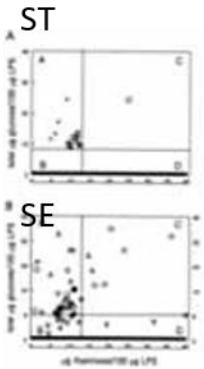
- 51 strains analyzed by phenotype microarray (PM), pH gradient 3.5 – 9.0,
  - Enteritidis, Typhimurium, Infantis, Heidelberg, Senftenberg
  - 1st four are poultry associated, Senftenberg associated with turkeys
  - Internal control strains consisted of characterized strains from my lab and the Shah laboratory, some had complete genome sequences available,
- 3 layers of statistical analysis were used to classify resistance,
  - Ttest ($p<0.03$), paired cluster analysis, effect size
  - PM runs were repeated to determine range of results for control strains
- Acid resistance was linked to increasing resistance to osmotic stress
  - Sodium lactate (strong association)
  - Sodium chloride (possible association)

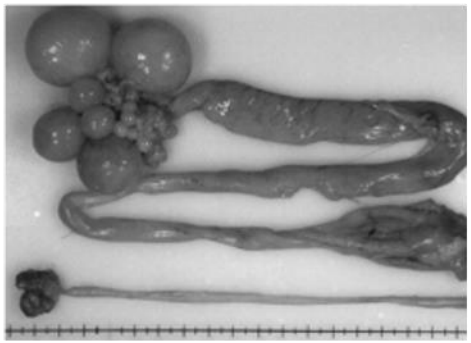
SUMMARY: Acid resistance may be contributing to issues of contamination within the processing of poultry products, rather than being an on-farm effect.

Timeline for whole genome sequencing



Egg-contaminating Salmonella Enteritidis has reproductive tract tropism in birds





SE produces a specialized capsular-like lipopolysaccharide (LPS) O-antigen, which...

- mitigates signs of disease in the hen
- contributes to long term survival of cells in the egg
- may contribute to persistent colonization of the gall bladder
- is linked to high cell density growth and organ invasion
- is linked to robust metabolic properties (e.g. nitrogen utilization).

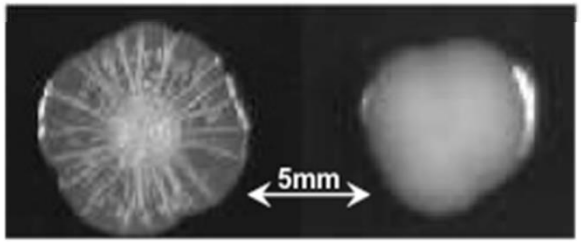
Craig Parker, USDA & Russ Carlson + colleagues, UGA

- The change in the O-antigen is mostly a change in mass, not antigenicity!
- Presence of the capsule, and other linked traits, is strain dependent.

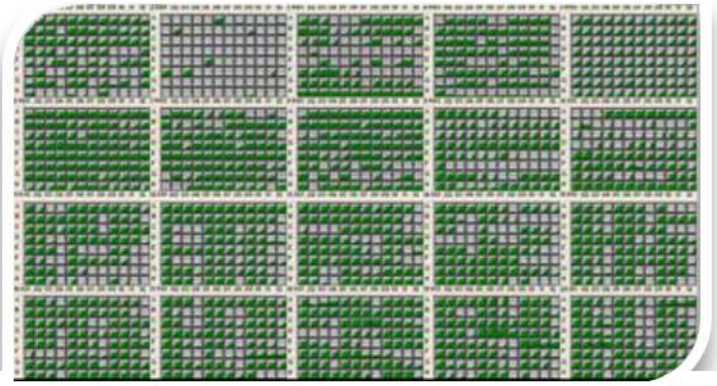
Subpopulation biology of *S.* Enteritidis

- Subpopulations share >99.99% of genome content
  - Not distinguishable by DNA-DNA microarray hybridization.
  - Not always distinguishable using fingerprinting methodology such as PFGE, ribotyping etc.
  - Occur naturally, frequently and can colonize spleens
  - Vary in their ability to grow to high cell density, to produce biofilm and to contaminate eggs.

- Subpopulations of S. Enteritidis have minimal genetic distance but maximal phenotypic distance.

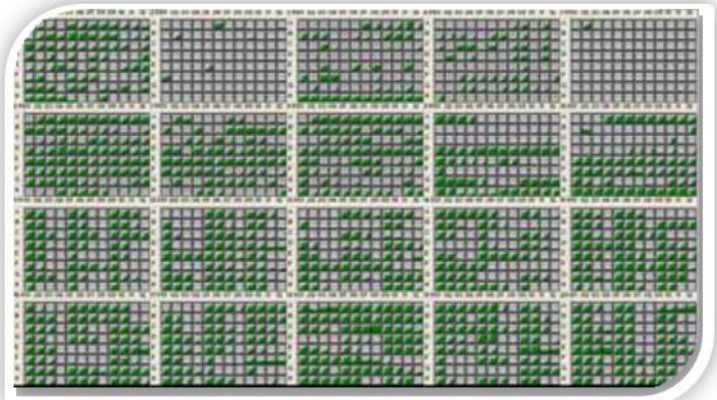


*Phenotype Microarray

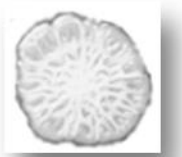


21046

The two subpopulations were found to differ in over 20% of test conditions performed in Phenotype Microarr.



21027



Morales et al. 2005

Currently conducting SNP analysis of SE obtained pre-NGS to post-NGS*

PRE-NGS, $15,000 per isolate POST-NGS, Collaborators take care of cost

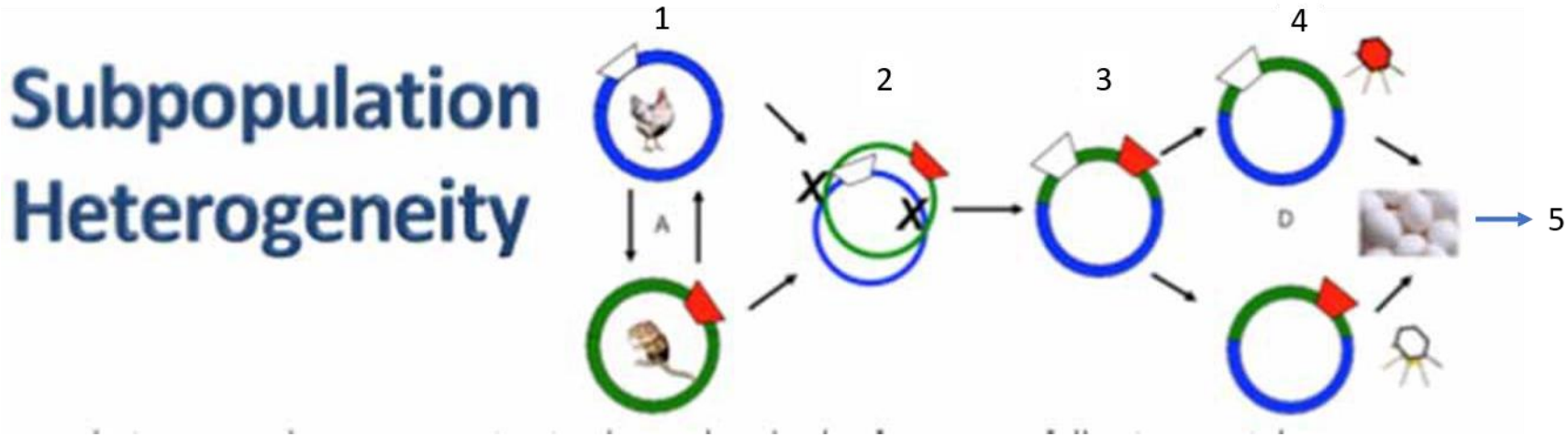
- 2005 – NGS just beginning: Sanger Institute allowed access to their draft genome sequence of SE subject to their publishing the genome first, access allowed me to do SNP analysis of genomes sequenced using primer walking (Nimblegen),
- 2008 – visited the Sanger Institute to discuss how to contribute the most to the rapidly evolving field of high throughput sequencing, decided my role was correlating genotype to phenotype, had completed SNP analysis of two strains of Enteritidis that varied in egg contamination and biofilm formation.
- 2008 – Sanger published PT4 Enteritidis genome in comparison to Gallinarum
- 2011 – SNPs linking phenotype to genotype in SE PT13a/8 strains 21027 & 21046 was published (Cesar Morales, PhD thesis)
- 2016 -2018 -Collaborated with FDA CFSAN (Guojie Cao & Marc Allard) through the Genome Trakr pipeline to sequence a collection of 91 SE obtained from mice caught live on farms during the height of the egg contamination problem
- 2018 – at NCBI:
  - Approximately 639 complete Salmonella genomes are available for BLAST,
  - 1.1 million contigs from about 7,900 genomes are available for BLAST,
  - Approximately 105,000 isolates submitted to Pathogen Detection will be linked for BLAST over next year or so,
- Can now revisit the prevalence of the previously detected ORF-disrupting SNPs in a much larger database
  - Of 107 complete genomes, have 74 high quality complete genomes of SE available for BLAST now (confirmed by cyaA alignments),
  - Approximately 234,000 contigs suggest >1500 draft genomes of SE available for BLAST (to be confirmed by concordant detection of cyaA and dkgB-linked intergenic ribosomal sequences that identify SE)
  - Linking the Pathogen Detection database to BLAST will increase stringency of analysis

*This slide was edited after the meeting following discussions with Bill Klimke at NCBI

Table 1: Genes of *Salmonella enterica* with coding sequences that vary between serotypes as detected by comparative whole genome analysis of *S.* Enteritidis subpopulations

Type of mutation as compared to reference genome *S. enteritidis* PT4	Location in S. Enteritidis PT4 genome	*Salmonella enterica* serovar Enteritidis gene with mutation: Gene with altered open reading frame	PT13a21046(D1)	PT13a21027(D1)	PT422079(D1)	Other *Salmonella enterica* subsp. I serotypes (O-serotype)[2,3]: Dublin(D1)	Gallinarum(D1)	Typhi(D1)	Javiana(D1)	Paratyphi(A)	Paratyphi(B)	TyphimuriumLT2(B)	TyphimuriumDT104(B)	Typhimurium4,[5];12;i:-(B)	Saintpaul(B)	Schwarzengrund(B)	Heidelberg(B)	Agona(B)	Choleraesuis(C1)	Infantis(C1)	Hadar(C2)	Newport(C2)	Kentucky(C3)	Outlier genomes: *S.enterica* arizonae	*Salmonellabongori*	E. coli 0157:H7 Sakai
A) Mutation that distinguishes biofilm negative PT13a 21046 from biofilm positive PT13a 21027 and S. enteritidis *PT4*																										
substitution: ORF termination	226922	fhuA (SeAg_B0227) (SeKA_4496)	0	+	+	78	78	0	78	0	99	78	78	78	78	78	99	100	78	78	78	78	99	97	99	93
deletion: 92 bp	4189388-4189479	mocR (SG3312) (SeD_A4510)	0	+	+	99	100	98	0	98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
deletion: 80bp of 5' end + 135 bp upstream	4642252-4642331	DGC (STM4551)	0	+	+	99	99	99	99	99	99	99	99					99	99	99	99	99	99	99	99	91	0	0
B) Mutation that distinguishes non-egg-contaminating PT13a 21027 from egg-contaminating PT13a 21046 and S. enteritidis *PT4*																										
deletion: 1bp	393529	foxA (STM0364)	+	0	+	99	0	99	99	99	99	99	99	99	0	99	99	99	99	99	99	99	99	0	0	0
substitution: ORF termination	1095448	putP (STM1128)	+	0	+	100	99	0	100	100	99	99	99	99	99	99	100	0	0	100	99	100	0	99	0	0
deletion: 1b p	1099079	kdgM (STM1131)	+	0	+	99	99	97	98	99	99	99	99	99	99	98	99	37	99	99	99	99	0	96	0	26
deletion: 6 bp	1554930-1554935	ydcZ (STM1591)	+	0	+	99	100	99	99	99	99	99	99	99	99	100	98	99	99	98	99	99	99	97	0	0
substitution: ORF termination	1681327	unknown (STM1472)	+	0	+	99	0	99	98	98	98	99	99	na	98	98	99	99	98	99	99	99	99	0	96	0
substitution: ORF termination	1828544	yojN (STM1320)	+	0	+	99	99	99	99	99	100	100	100	100	100	99	100	99	99	100	100	99	100	99	99	96
deletion: 12 bp in-frame	2939833-2939844	lysR (STM2912)	+	0	+	97	99	98	99	97	97	98	98	98	98	99	98	97	0	97	98	98	97	93	86	0
substitution: ORF termination	2957802	cysN (STM2934)	+	0	+	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	95	92	87
deletion: 10 bp	3880104-3880113	dsdA (STM3802)	+	0	+	99	100	99	99	99	99	99	99	99	99	99	100	99	99	99	99	100	99	97	0	0
deletion: 12 bp in-frame	4368126-4368137	lrgB (STM4272)	+	0	+	99	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	98	97	0
deletion: fusion of yjfK and yfjL	4472579-4472589	yjfK (STM4373)	+	0	+	99	99	99	99	99	100	99	100	100	100	99	100	99	99	100	100	100	99	0	0	0
deletion: fusion of yjfK and yfjL	4472579-4472589	yjfL (STM4374)	+	0	+	100	0	99	99	100	100	100	100	99	100	99	100	99	99	100	100	100	99	0	99	0
deletion: single bp	4574262	sefD (SeD_A4882)	+	0	+	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

[1] Genes of S. Enteritidis with coding sequences that vary between and within phage type due to a deletion or a change in termination; +, ORF predicted to result in a complete transcript; 0, ORF is disrupted; na, possible error in contig assembly or database information or sequence not available.



1. **Two strains of serotype Enteritidis (SE) were co-evolving during middle ages,**
    1. Each strain had a preferred host (HO mouse vs chicken) but association was not exclusive,
    2. Each strain had its own lysogenic bacteriophage (red vs white),
2. **Sometime from the 1950s to 1970s changing agricultural practices may have placed selection pressure on SE to undergo a rare homologous recombination event resulting in a 40%/60% swap in the genome between strains,**
3. **For a brief time, a strain existed that had two bacteriophages in one genome, which is an inherently unstable,**
4. **The strain split into two evolutionary lineages with similar chromosomes but that varied in phage type,**
    1. PT4 SE established itself on the European continent, PT13a/8 established itself in the United States
5. **Eventually Phage type varied around the world and included other phage types as evolution continued since the first widespread emergence documented during the 1980s; phenotype and SNP content varied.**

***Salmonella enterica* subspecies I is a dynamically evolving pathogen capable of changing its evolutionary trajectory and fine-tuning its genomic content to a myriad of hosts, food sources, and environments.**

Post NGS Era

- Have released 64 genomes from SE isolated from mice caught live on poultry farms from the NE of the US during the 1990s,
- Will eventually have 91 isolates in the collection (at NCBI pending release date)
- Opportunity to revisit the occurrence of 16 different mutations that differentiated two phenotypes of SE (will look at two today)
  - sefD, SEF14 fimbrial structural gene
  - mocR, HTH transcriptional regulator upstream of amino acid synthesis genes
- Opportunity to search for antibiotic resistance genes
- Have several other serotypes to similarly characterize.

Genome Announc. 2017 Sep 7;5(36). pii: e00953-17. doi: 10.1128/genomeA.00953-17.
Draft Genome Sequences of 64 Salmonella enterica Serotype Enteritidis Isolates Obtained from Wild Mice.
Guard J[1], Cao G[2], Kastanis GJ[2], Davison S[3], McClelland M[4], Sanchez Leon M[2], Zheng J[2], Brown E[2], Allard MW[2].

91 mouse isolate genomes (FDA CFSAN collaboration)

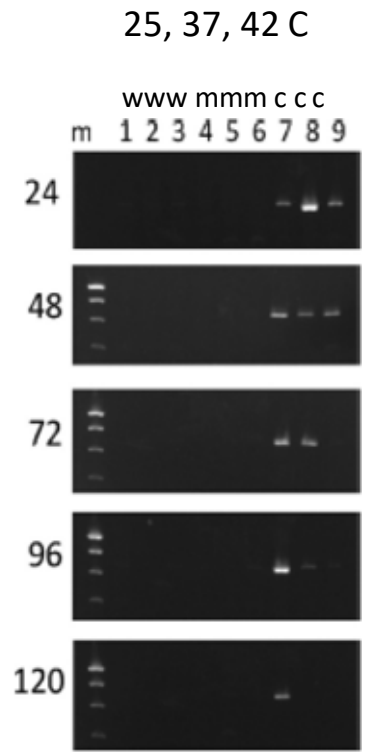
The number of SNP differences (standard deviation) between groups using MEGA7 (bootstrap 1000)

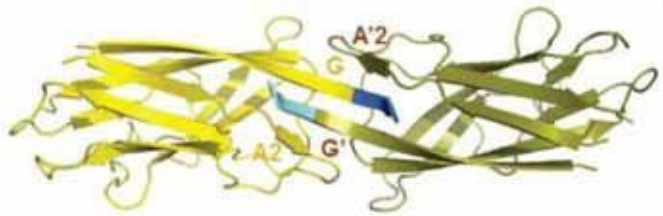
	A1	A2	A3	A4	A5	A6	A7	A8	B1	B2	B3
A2	44(5)										
A3	50(6)	45(6)									
A4	40(5)	36(5)	39(6)								
A5	44(5)	40(5)	43(6)	32(5)							
A6	54(6)	50(6)	55(7)	46(6)	50(6)						
A7	51(6)	47(6)	52(6)	43(6)	47(6)	57(7)					
A8	50(6)	45(6)	51(6)	41(6)	45(6)	55(7)	24(5)				
B1	92(8)	90(8)	95(8)	86(8)	90(8)	99(8)	97(8)	95(8)			
B2	84(8)	82(8)	87(8)	78(8)	82(8)	91(8)	89(8)	87(8)	33(5)		
B3	137(9)	135(9)	140(10)	131(9)	135(9)	144(10)	142(10)	140(10)	86(8)	78(8)	
B4	133(9)	131(9)	137(9)	127(9)	132(9)	141(10)	138(10)	137(9)	83(8)	75(7)	46(6)

- Two major lineages: A and B
- A divided into 8 sublineages
- B divided into 4
- Previous work focused on isolates in B1 & B2 which have a near minimal genetic distance amongst the collection

SefD, tip adhesion of Fimbria SEF14

- SefD is essential for uptake or survival of SE in macrophages
- It is a stable dimer, with the monomers packing against a hydrophobic face stabilized through domain-swapping via the A2 and G beta-strands
- Only in Group D salmonella (Enteritidis, Pullorum, Gallinarum, Typhi)
- BF 21027, but not EC 21046, had one nucleotide deleted, which disrupted the ORF
- However, EC 21046 was a functional mutant, negligibly expressing the gene under conditions analyzed
- Complementation with the gene on a constitutive expressing plasmid was required to recover SefD
- Best expression was at ambient temperature



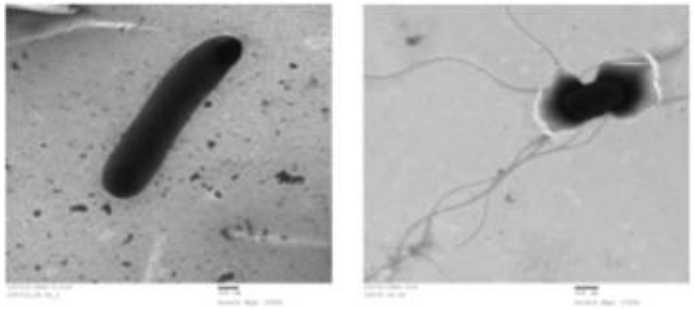


Liu et al, 2012. Biochem Biophys Res Commun 421:2018-213
Promoting Crystallization of the Salmonella enteritidis fimbriae 14 pilin SefD using Deuterium Oxide

Making a better vaccine is tough

SefD

TRANSMISSION ELECTRON MICROSCOPY



SefD + SefD-

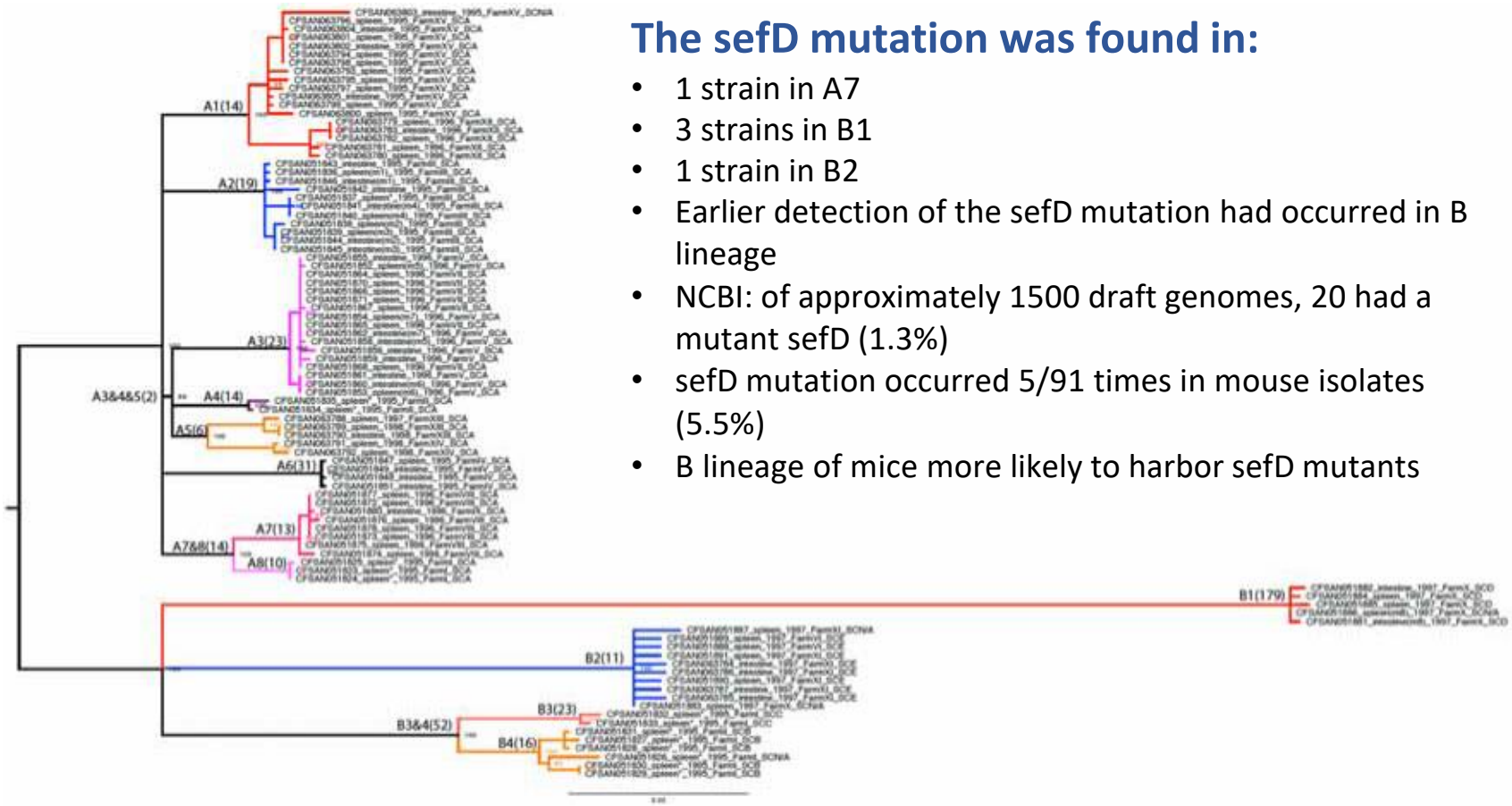
FIGURE 1 LEGEND: Morphological differences between *S. enterica* serovar Enteritidis that varies in the presence of SefD. Constitutive expression of the protein is associated with cell elongation, lack of flagellation, and absence of capsular-like O-antigen.

Knocking down HMM LPS& flagella, and increasing piliation made a better vaccine!

Multiplication in Egg Yolk and Survival in Egg Albumen of Genetically and Phenotypically Characterized *Salmonella* Enteritidis Strains. Gast RK, J Guard, R Guraya and A Locatelli. Journal of food protection. Acceptd 1/18/18.

- Collaboration with Richard Gast
- Only two phenotypes influenced growth in yolks or survival in albumen ($p<0.01$)
  - Mutation of fimbrial gene sefD decreased growth
    - Suggests there is positive selection pressure to maintain sefD, and by inference, SEF14
  - Antibiotic resistance genes bla_{TEM-1}, tet(A) were associated with less survival in albumen IF resistance was being expressed
    - A strain was found that had the genes but was not expressing resistance
    - Suggests there is negative selection pressure for antibiotic resistances in Enteritidis
    - Antibiotic resistance genes were found only in the B2 lineage within the mouse collection

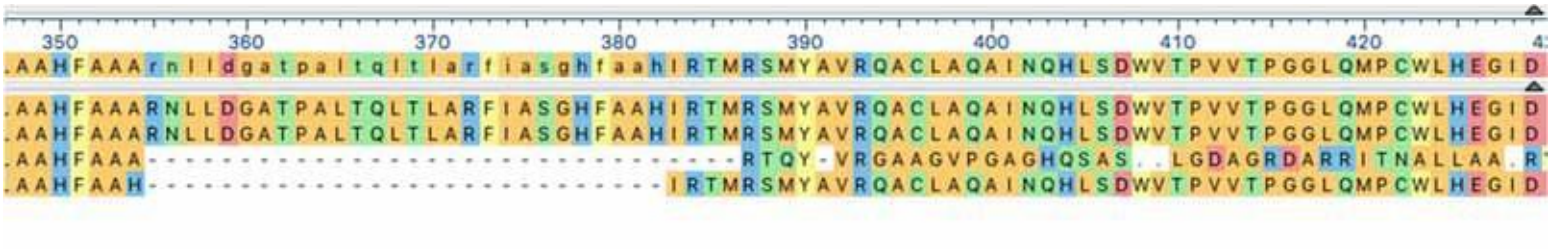
sefD, structural protein of SEF14, gain in virulence mutant, single bp deletion 1st discovered in lineage B1 strain 21027

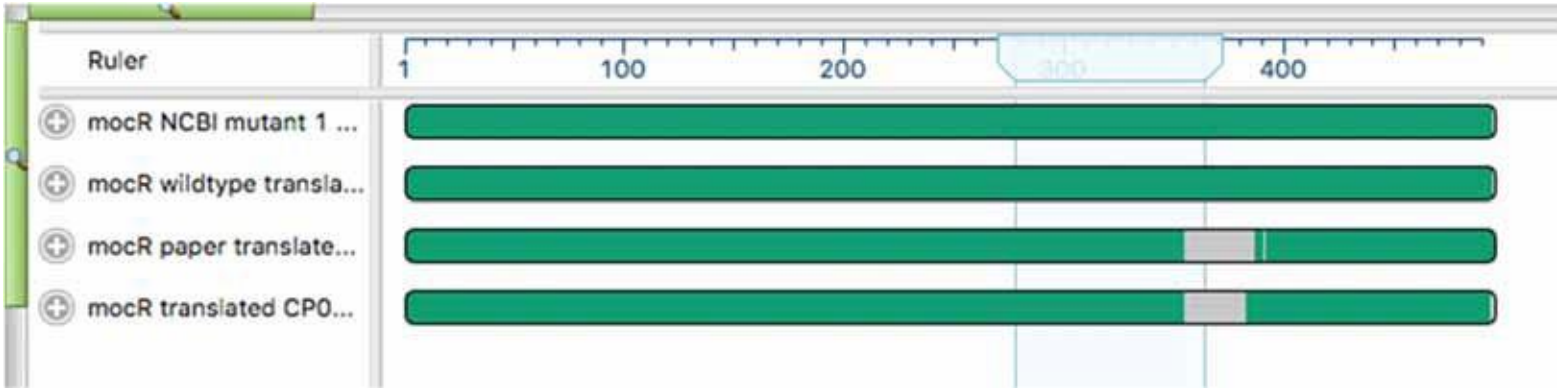


The sefD mutation was found in:

- 1 strain in A7
- 3 strains in B1
- 1 strain in B2
- Earlier detection of the sefD mutation had occurred in B lineage
- NCBI: of approximately 1500 draft genomes, 20 had a mutant sefD (1.3%)
- sefD mutation occurred 5/91 times in mouse isolates (5.5%)
- B lineage of mice more likely to harbor sefD mutants

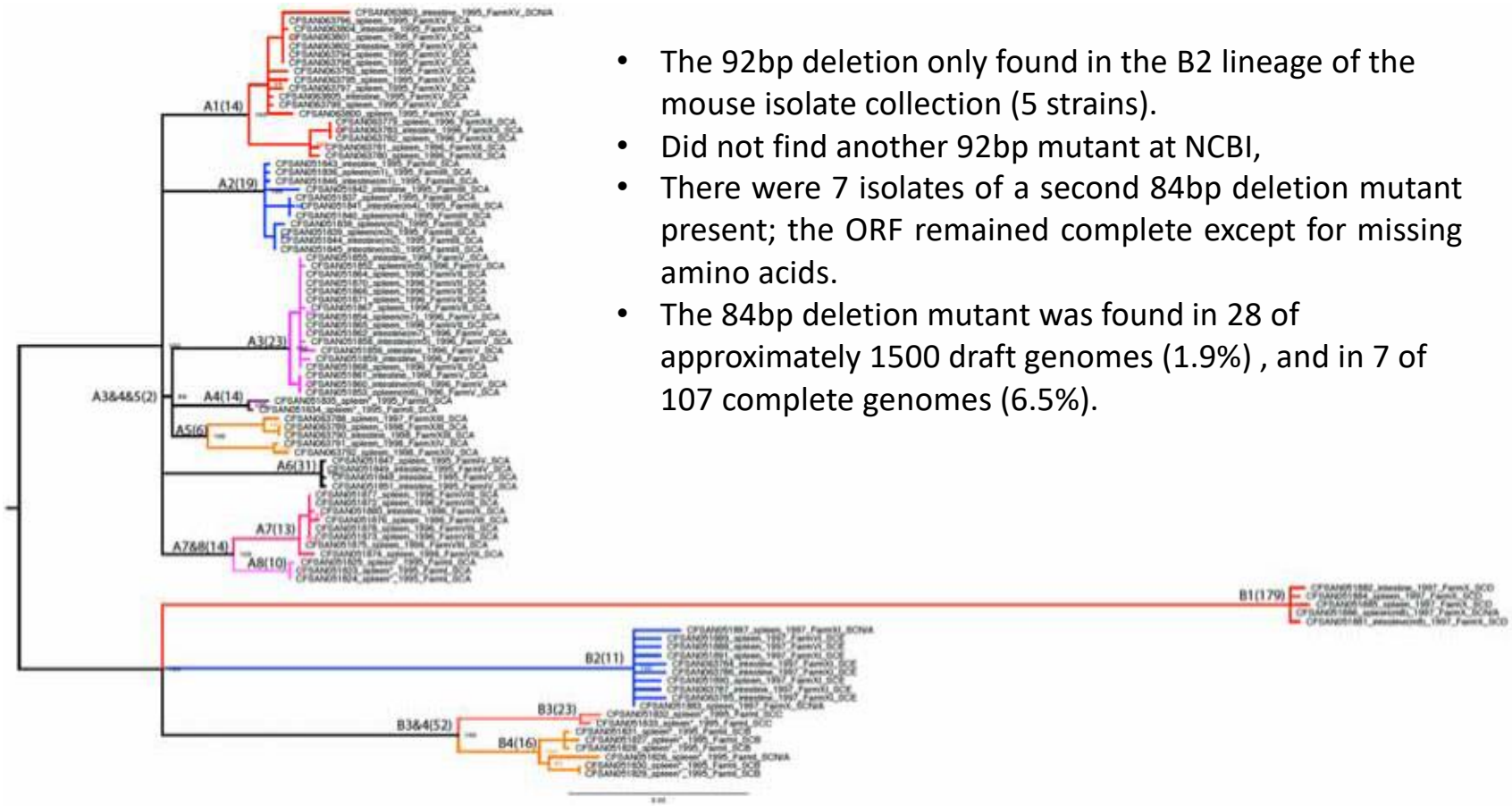
mocR, helix-turn-helix transcriptional regulator of amino acid synthesis (predicted)





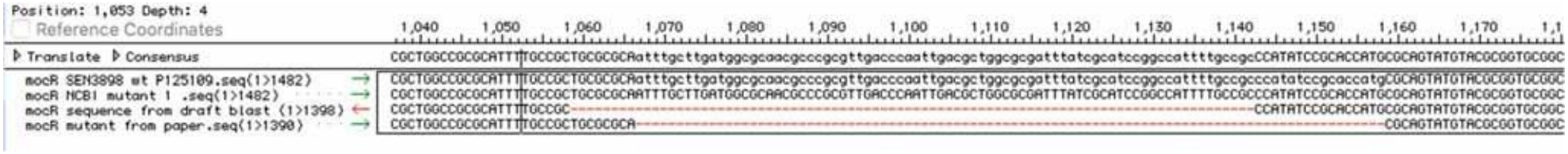
mocR, transcriptional regulator, amino acid synthesis

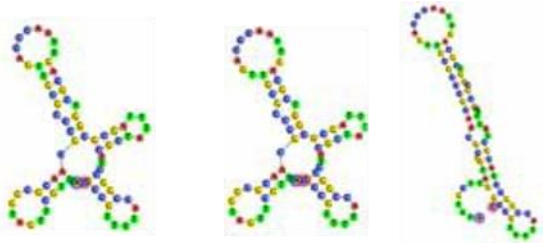
92bp deletion 1st discovered in B2 lineage (strain 21046)



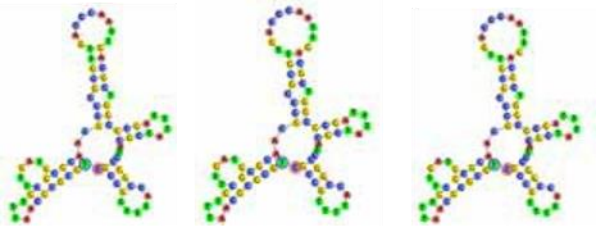
- The 92bp deletion only found in the B2 lineage of the mouse isolate collection (5 strains).
- Did not find another 92bp mutant at NCBI,
- There were 7 isolates of a second 84bp deletion mutant present; the ORF remained complete except for missing amino acids.
- The 84bp deletion mutant was found in 28 of approximately 1500 draft genomes (1.9%) , and in 7 of 107 complete genomes (6.5%).

Melting profiles of *mocR, a helix-turn-helix transcriptional regulator,* vary by temperature and the type of deletion present





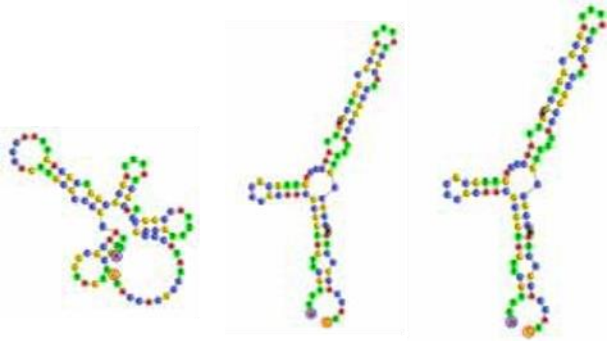
76 bp deletion



84 bp deletion

28 C 37 C 42 C

100 bp deletion



28 C 37 C 42 C

92 bp deletion

Software: DNA Matthews 2004

Summary (2)

- Current analysis of mutations of SE that were previously linked to phenotype confirm that mouse isolates had some unusual attributes,
- Although some of these mutations were not prevalent in mouse isolates, they were not rare,
- Incidence of these mutations in the current database of NCBI genome entries indicates that these mutations are less common in the overall universe of SE genomes than they were in mouse isolates,
- Results suggest that strains were accumulating mutations as did avian-restricted Gallinarum,
- The path to host restriction may have some intermediates with phenotypes impacting association with egg contamination, biofilm formation and organ invasion in hens.