Translating Sequencing to the Public Health Laboratory

Michael Valentine
Research Associate II
ORAU/FDA Research Fellow
Translational Genomics Research Institute-North

tgen® NORTH

DIAGNOSTIC ▪ ANALYTIC ▪ FORENSIC ▪ EPIDEMIOLOGIC
TRANSLATIONAL PATHOGEN GENOMIC RESEARCH

-A non-profit research institute dedicated to making and translating discoveries into advances in human health.

Profile:

- Est. April 13, 2007

- Staff size: 36
- PhD, MS, MBA, MD, MPH, BS

- Funded Projects: 17

Translating Sequencing to the Public Health Laboratory

Topics

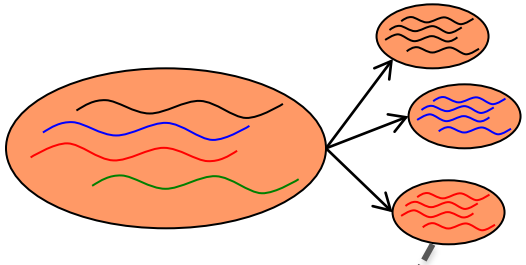
- Quick overview of Next Gen sequencing (NGS) process
- Whole genome sequencing (WGS) applications
- WGS analysis for foodborne outbreak investigations

tgen NORTH

PH & Clinical Microbiology – Pre-Next Gen

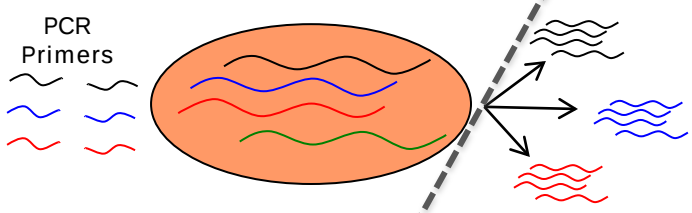
DIRECT SAMPLE ANALYSIS

Direct culture of mixed population



- Isolated organisms speciated by morph
- biochem, molecular, mass spec

PCR Detection in mixed population



- Highly sensitive
- specific

ISOLATE ANALYSIS

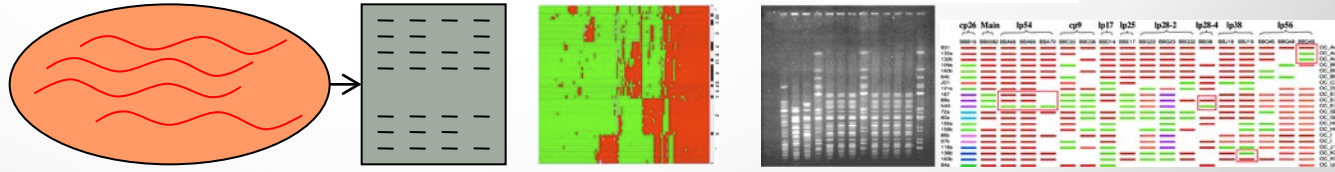
Phenotype Analysis

Antibiotic Resistance

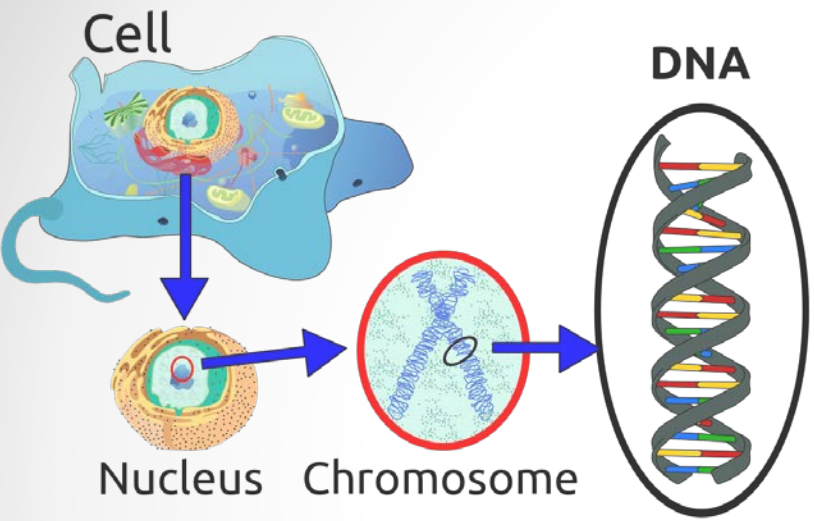
Genotype Analysis

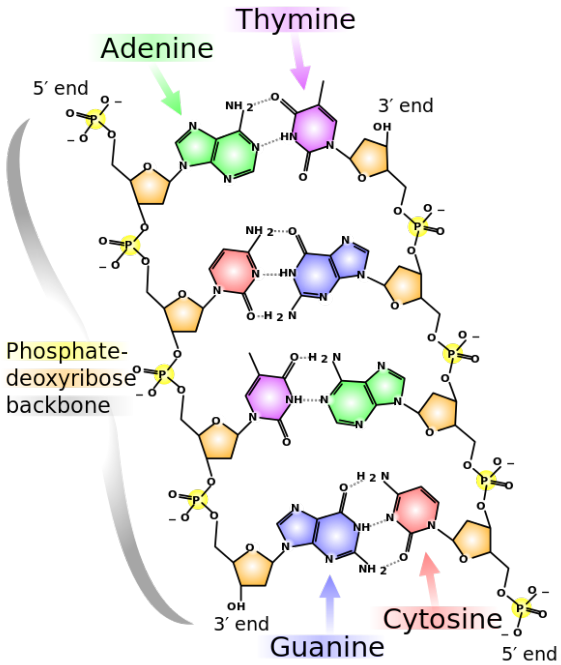
RFLP
AFLP
Primers
RE

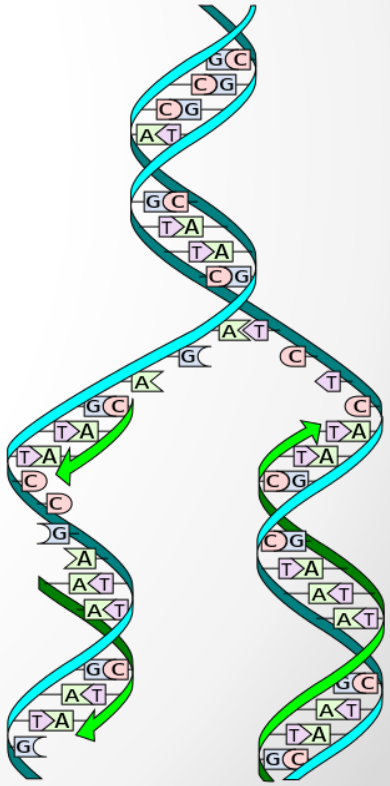
MLST
Primers



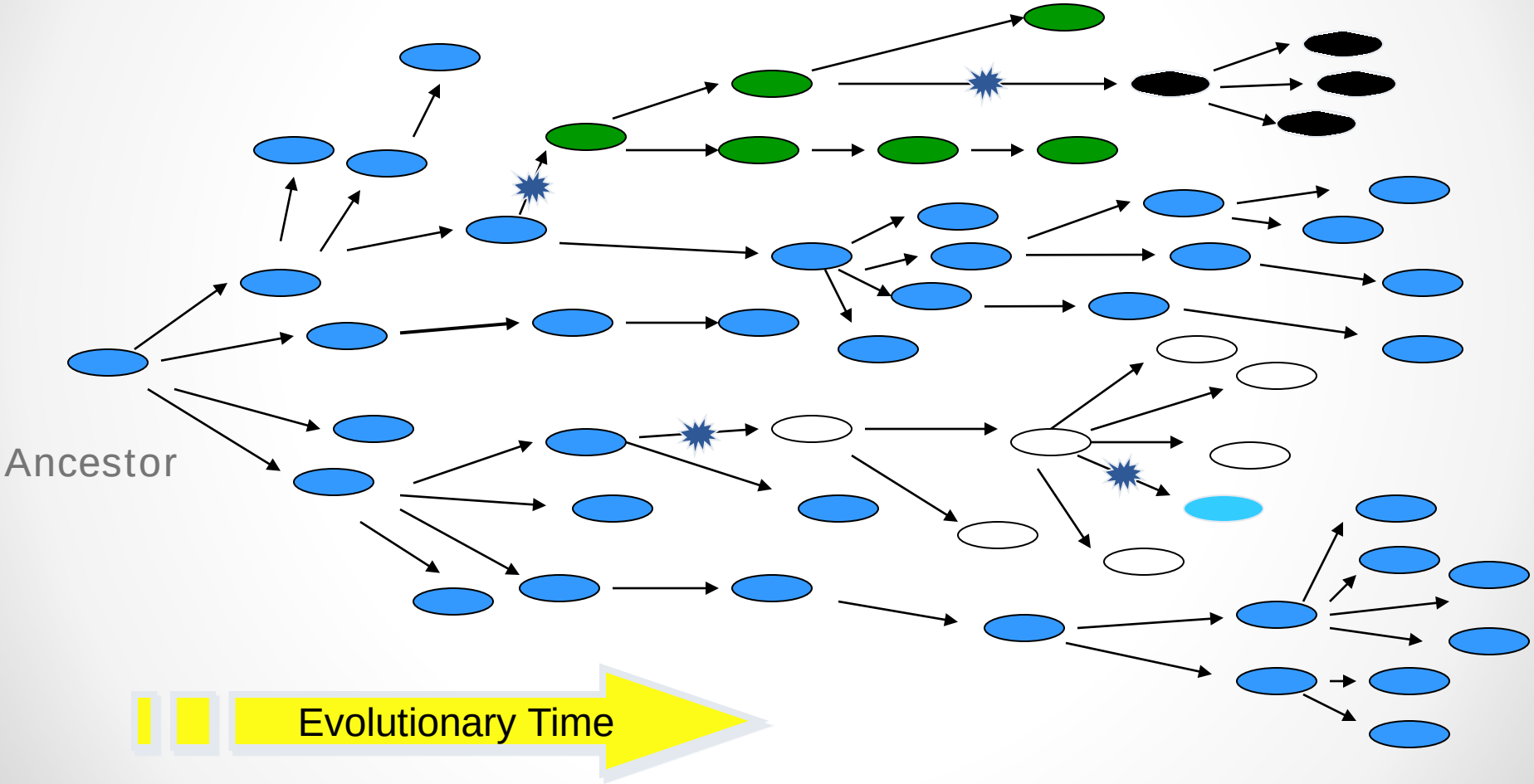
DNA Basics


Cell
DNA
Nucleus
Chromosome


Thymine
Adenine
5′ end
3′ end
Phosphate-
deoxyribose
backbone
3′ end
Guanine
Cytosine
5′ end



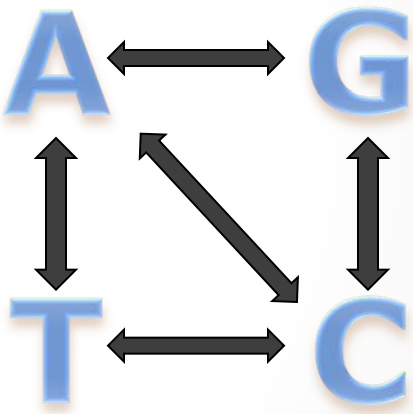
Pathogen and Genomic Diversity



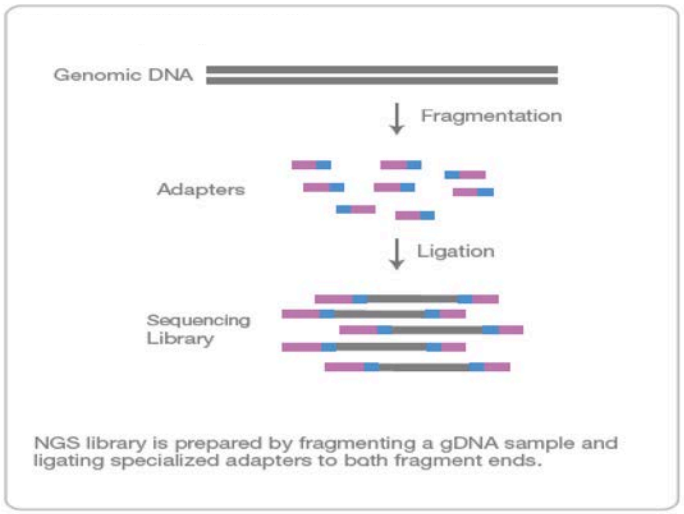
= Mutational Event

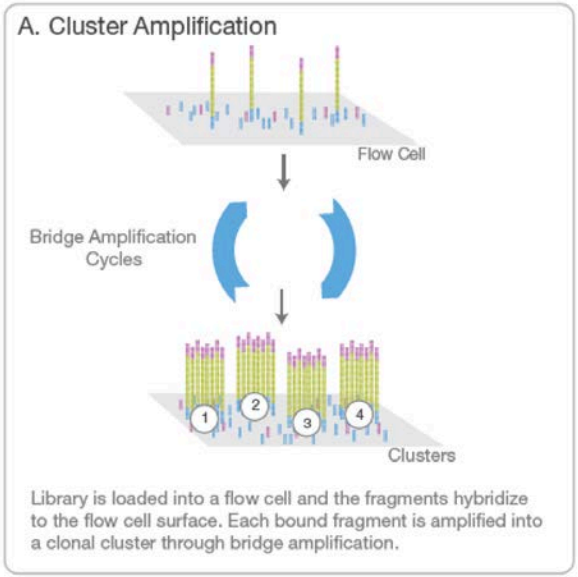
SNP
Single Nucleotide Polymorphism

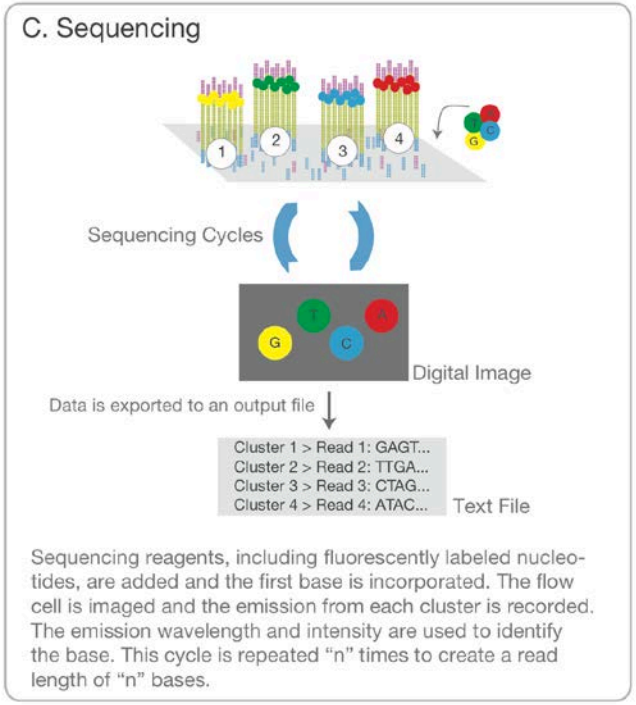
- The mutation of one DNA nucleotide into another
- Some SNPs cause change in protein structure; some cause no changes at all
- All are useful for determining relationships between organisms



Sequence Process



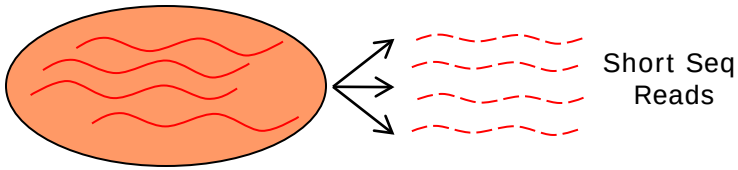




PH & Clinical Microbiology – Next Gen Era

ISOLATE ANALYSIS

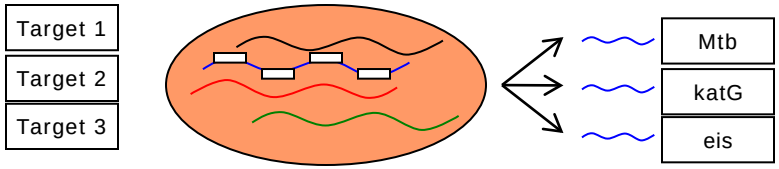
Whole Genome Sequencing



- Genotyping/Phylogenetics
- Ab res, Virulence
- Expression

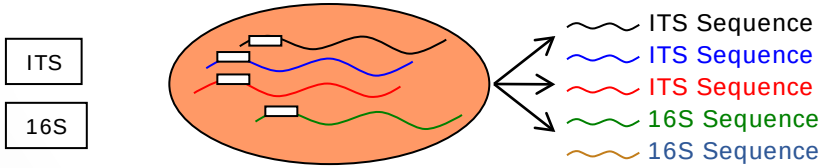
DIRECT SAMPLE ANALYSIS

Amplicon Sequencing



- Highly Multiplexed
- Targeted sequencing

Mycobiome/ Microbiome Profiling



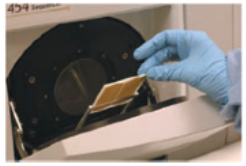
- NextGen sequencing of amplified ITS/16S

tgen NORTH

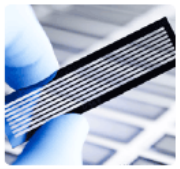
Next Gen Sequencing

Capillary electrophoresis (Sanger)
Between 96 and 384 samples
(76 - 308 Kb/run)

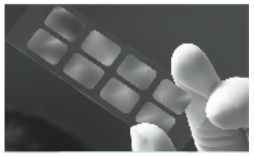
454
400,000 samples
(120Mb / run)



Solexa
32M samples
(2Gb / run)



Solid
40M samples
(2-6 Gb / run)



Up to 15Gb – 24-60hrs

1GB – 4.5hrs

?

WGS Bioinformatic Toolkit

SNP Callers
SolSNP
Varscan

Assemblers
Velvet
oases
phrap
SSPACE
edena v3
Amos
MIRA
SPADES

Next-Gen Aligners
bwa
STAMPY
Novoalign
NGen

Whole Genome Aligners
Mauve
MUMmer
Mugsy

Phylogenetic Software
circos
arb
mothur
FigTree
MetaTree

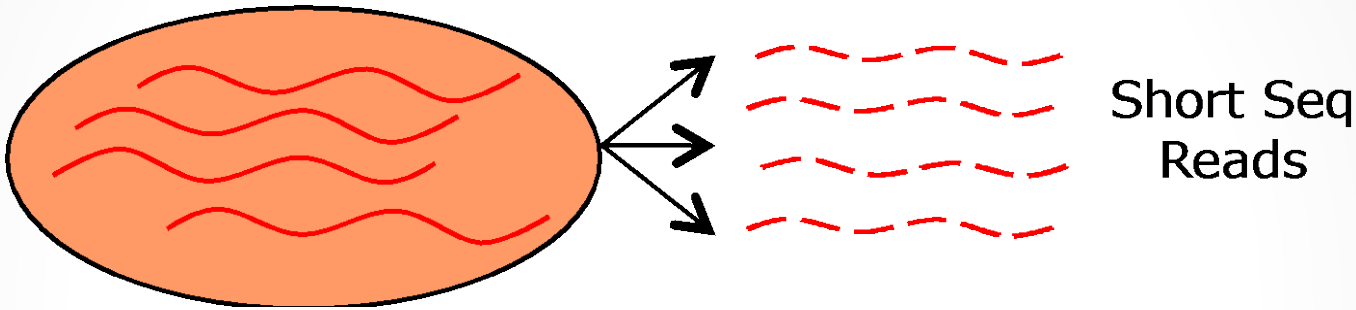
FastTree
RaxML
PAUP 4.0b10
MEGA

Next-Gen Misc Tools
sickle
htseq-count
DESeq
Qiime
GATK
Bedtools
FASTX Toolkit
Picard
IGV
SAMTools

CLC Genomic Workbench

Whole-Genome Misc Tools
glimmer
transeq
Jalview
MPI-BLAST

WGS analyses
Disease Outbreaks



- Prison *Salmonella enteriditis*
- Cucumbers and *Salmonella poona*

AZ Prison Outbreaks

- AZDH identified cases of *Salmonella enteritidis* with PFGE pattern JEGX01.0004 from three different prisons in AZ
  - Apr '15- Jan '16'

Are the outbreaks related between facilities?

- Prison #1→ 8 lab confirmed cases spread over a period of weeks
  - Chicken tested positive for Salmonella
  - Serotype differed from clinical strains

Are the outbreaks within each facility related ?

- Prison #2→ outbreak point-source, hundreds of ill inmates over the course of a few days
  - Salad (+) for Salmonella, individual ingredients (-)
  - Probably the source?

- Prison #3→ only one case identified (not an outbreak)

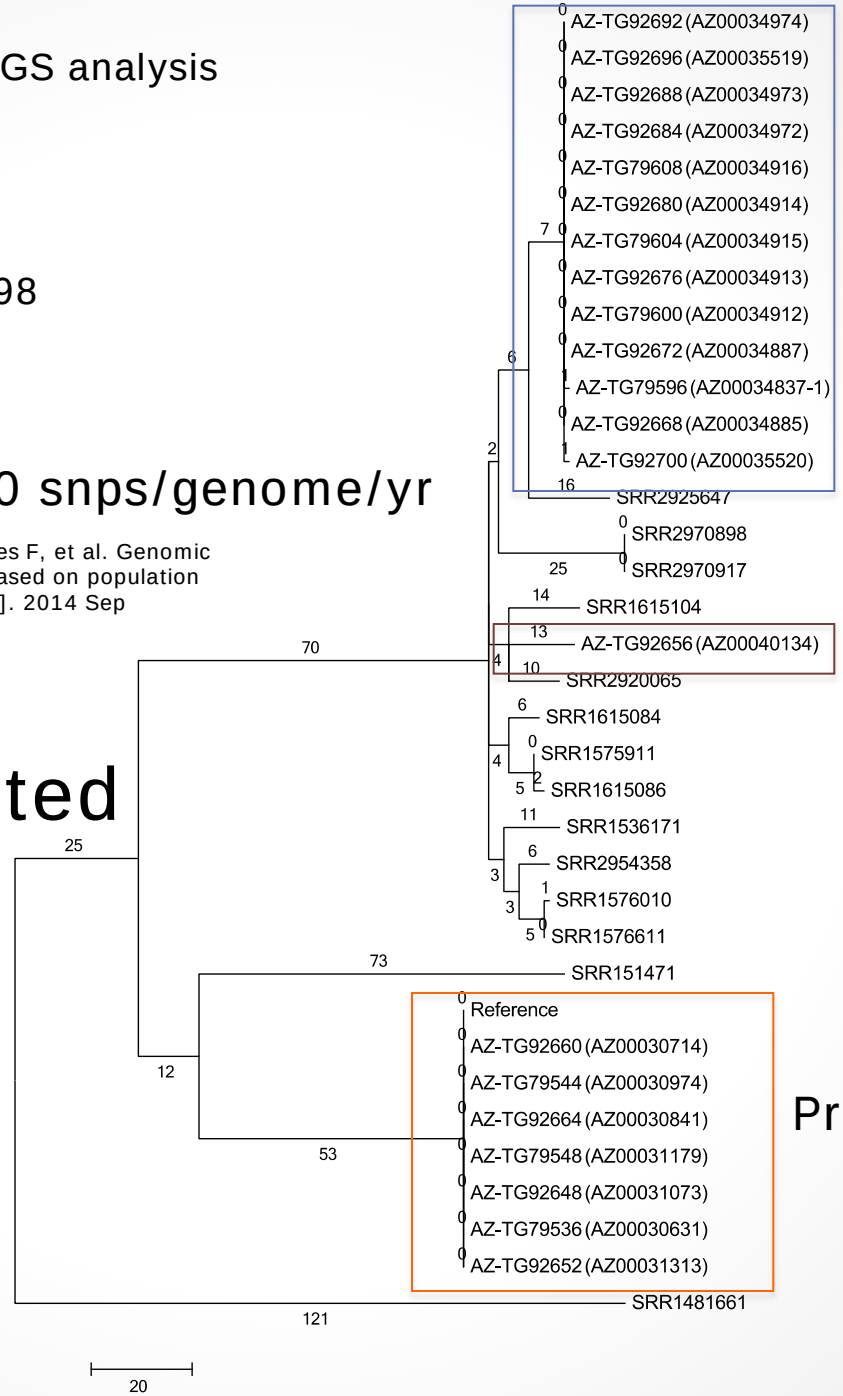
S. enteritidis prison outbreak WGS analysis
PFGE pattern: JEGX01.0004

Reference=AZ-TG79536
CI=0.995
Parsimony Informative SNPs=198
Total SNPs=495
Quality Breadth=97.20%

S. enteritidis SNP rate=1.0 snps/genome/yr

Deng X, Desai PT, den Bakker HC, Mikoleit M, Tolar B, Trees F, et al. Genomic epidemiology of Salmonella enterica serotype Enteritidis based on population structure of prevalent lineages. Emerg Infect Dis [Internet]. 2014 Sep

Outbreaks unrelated



Salmonella poona Outbreak Linked to Cucumbers

- A total of 907 people ill, reported from 40 states.
- 204 ill people were hospitalized-6 deaths
- Epi, lab and traceback investigations identified possible source-cucumbers from Mexico
  - distributed by Andrew & Williamson Fresh Produce
  - Two recalls issued for cucumbers
- Illnesses reported after the peak did not identify an additional food linked to illness.
- The exact source of contamination for the cucumbers distributed by Andrew & Williamson has not been identified.

Source: CDC.gov/salmonella/poona-09-15

tgen NORTH

Salmonella poona Outbreak-AZ

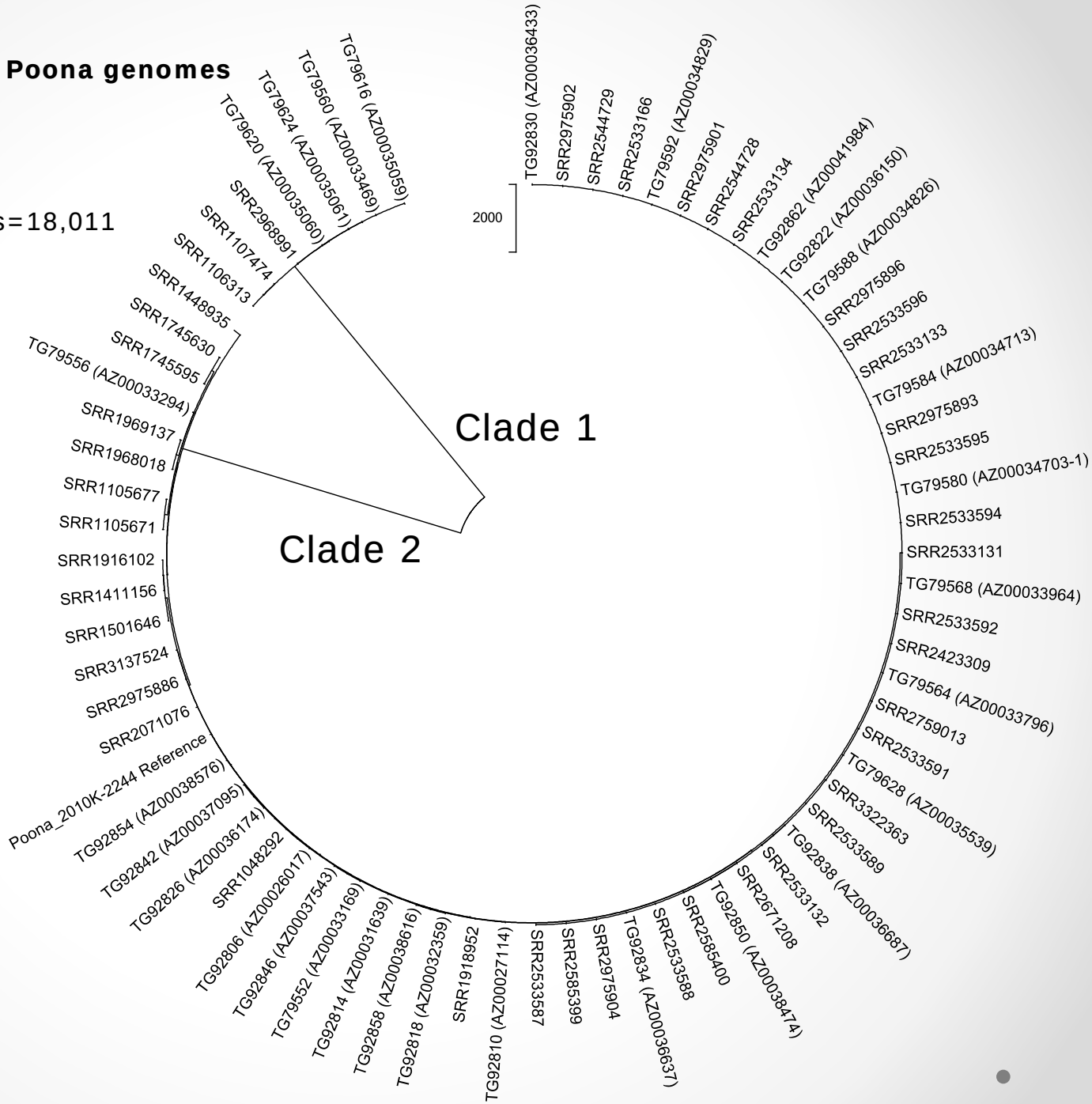
How were AZ strains related to cucumber strains?

Is the outbreak ongoing in AZ?

- Salmonella Poona with PFGE pattern primarily JL6X01.0018
- Others: JL6X01.0375, JL6X01.0133, JL6X01.0778, JL6X01.0129, JL6X01.784
- CDC called this cluster 1508MLJL6-1, 2nd cluster called 1508MLJL6-2
- Estimated and reported illness onset dates ranged from 7/3/2015 to 2/15/2016
- 140 lab confirmed cases in Arizona.

AZ Poona and select SRA Poona genomes

NASP: Novoalign, GATK
Maximum Parsimony
Ref=Poona_2010K_2244
Quality Breadth=90.03%
Parsimony Informative SNPs=18,011
Total SNPs=18,833
CI=0.999



AZ Poona and select SRA Poona genomes
Clade 1 Zoom

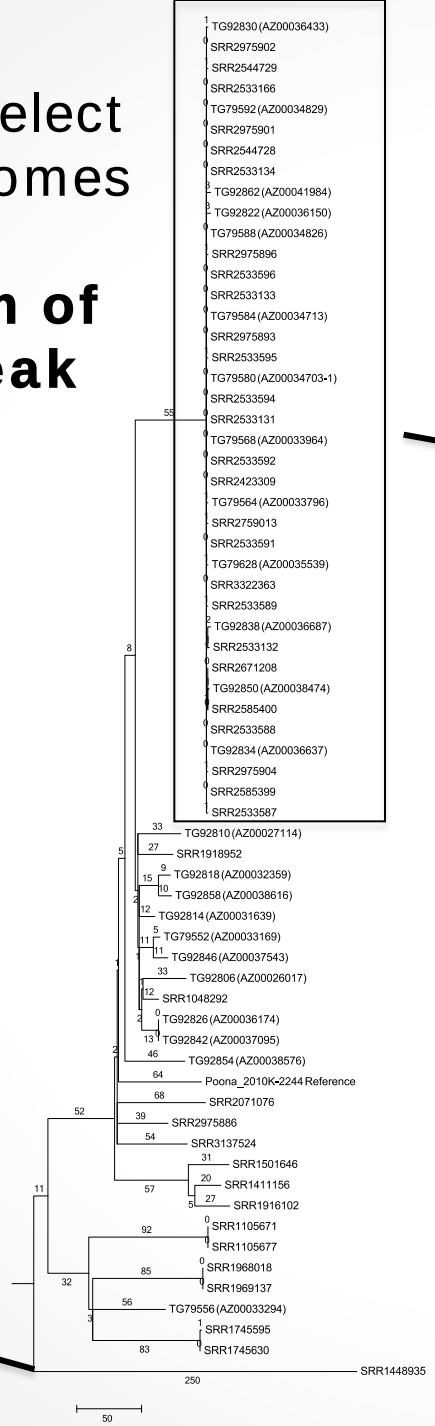
CDC outbreak pattern
1508MLJL6-2
PFGE: JL6X01.0104

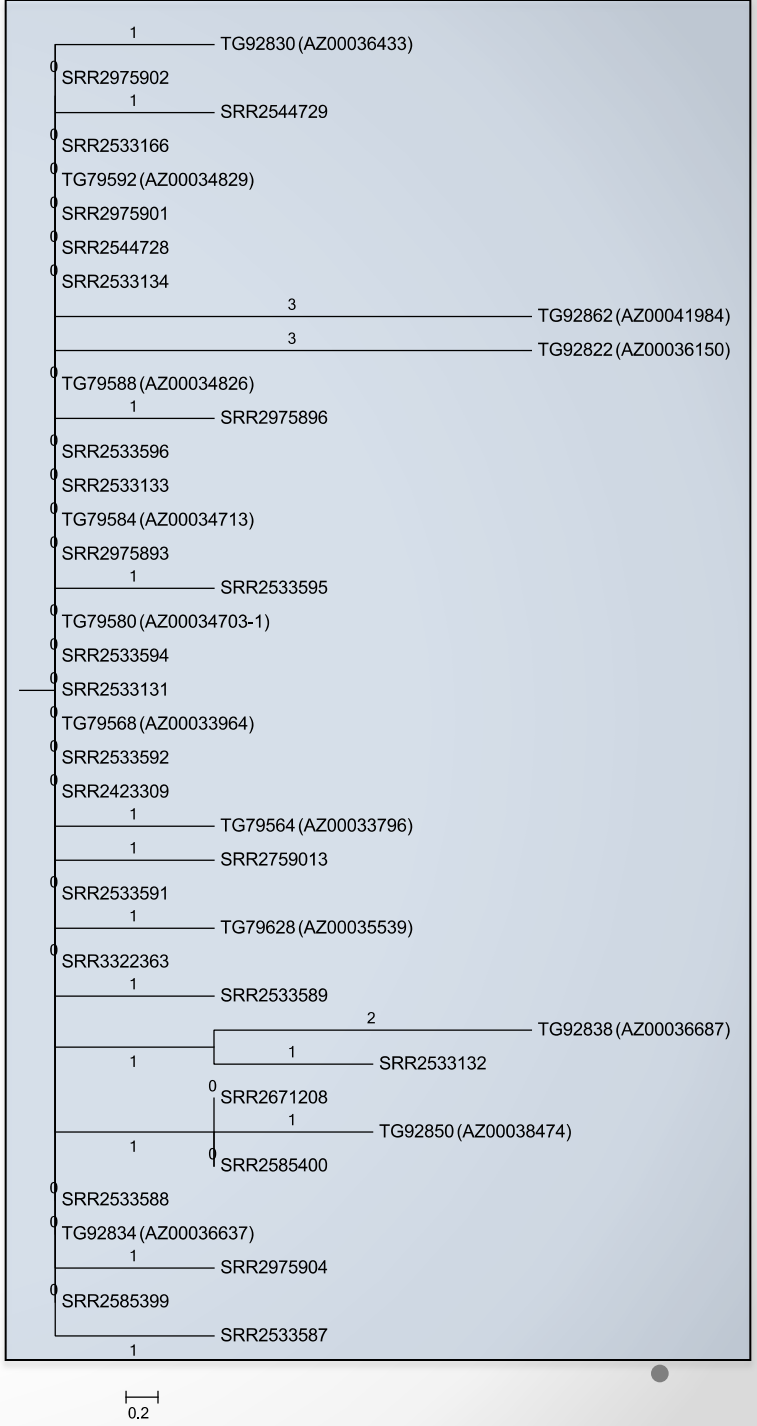
- TG79560-3yr old 7/20/15
- TG79616-animal 9/10/15
- TG79620-aquarium
- TG79624-aquarium water
- Same household?

- SRR2968991 submitted by CDC/Pulsenet
- Collected Aug 2015. HHS region 9 (AZ, CA, HI, NV)

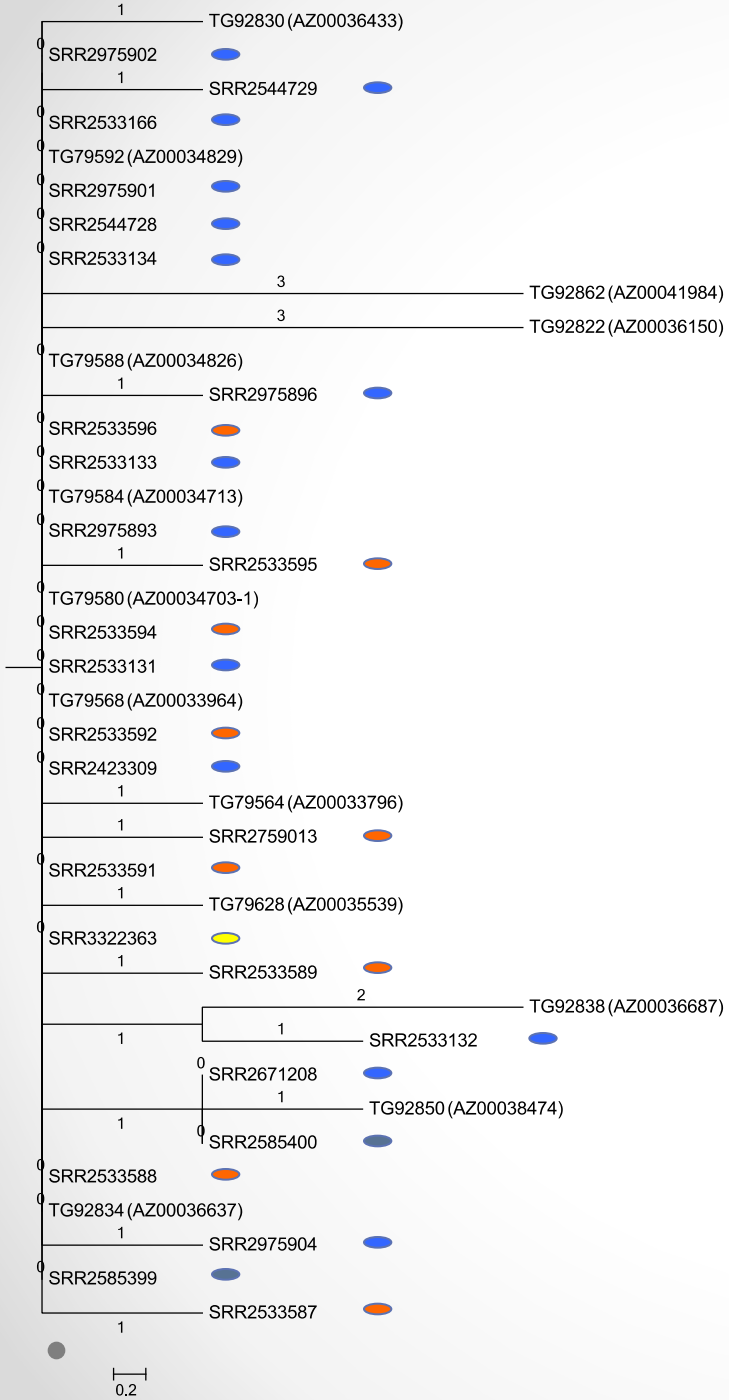
AZ Poona and select SRA Poona genomes

Clade 2: zoom of largest outbreak group





FL-University of FL
Collected Dec 2013
PFGE: JL6X01.0678



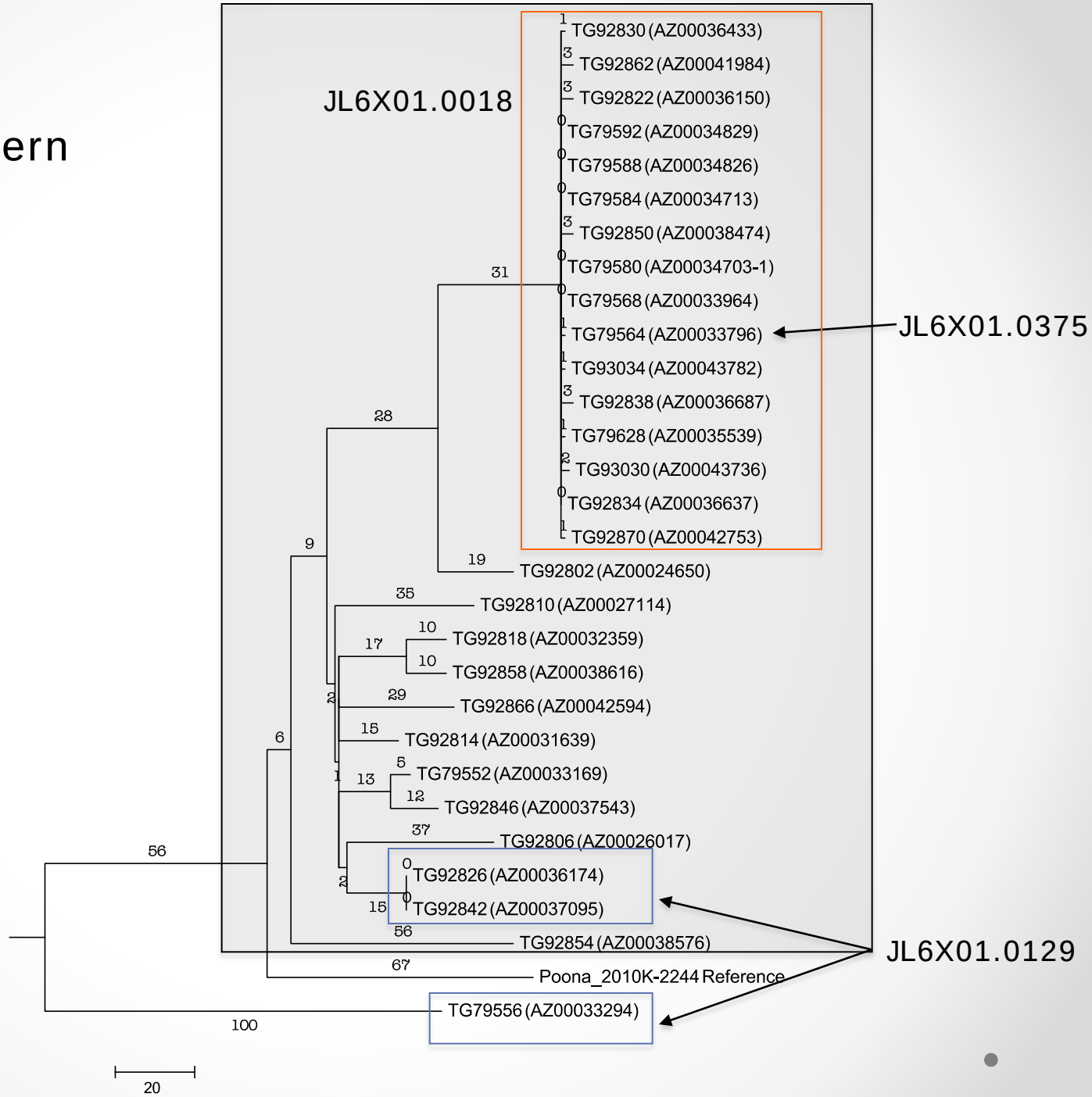
- FDA/Pulsenet-Collected Aug-Sept 2015 Clinical Samples HHS region 5
- FDA-Collected 9/7/15, Utah-Cucumbers
  PFGE: JL6X01.0018
  Secondary: JL6A26.0007
- CA Public Health FDLAB-Collected 9/10/15 Cucumbers
  PFGE: JL6X01.0018
  Secondary: JL6A26.0007
- U.K. Public Health-Collected Sept 2015 Clinical Sample

AZ Poona

CDC outbreak pattern

1508MLJL6-1

- TG79552
- TG79556
- TG79564
- TG79568
- TG79580
- TG79584
- TG79588
- TG79592
- TG79628



Translating Sequencing to the Public Health Laboratory Summary:

- WGS/NGS much higher resolution
- Able to better understand relatedness
- Tool to focus epi investigations
- Rapidly becoming gold standard
- Tgen-north is a resource

The Promise of Next Gen for PH and Clinical Laboratory

- Detection and Genotyping – Now
- Antimicrobial Resistance Profiling - Now
- Virulence Phenotyping - Now
- Complex Infections - Now
- Host Factors – Soon
- 3-8 Hour Sequencing – Now
- 15 min Sequencing – Soon
- Kit-based + Automation - Now

tgen NORTH

Collaborators

- Dave Engelthaler
- Jolene Bowers
- Veronica Harrison
- Chandler Roe
- Darrin Lemmer
- Jason Travis
- Elizabeth Driebe
- Paul Keim

- Victor Waddell
- Marilee Kellis
- Stacy White
- Sherry Gower
- Ken Komatsu

Thank you!