Genomics and Transcriptomics in the Clinical Microbiology Laboratory

Randall J. Olsen, M.D., Ph.D.

Director, Molecular Diagnostics Laboratory
Director, Special Testing Laboratory
Co-Director, Microbiology Laboratory
Department of Pathology and Genomic Medicine
Houston Methodist Hospital and Research Institute
Houston, Texas, USA

Association of Public Health Laboratories

June 6, 2016

HOUSTON Methodist RESEARCH INSTITUTE

The Modern Microbiology Lab

HOUSTON Methodist RESEARCH INSTITUTE

- Rapid identification of pathogenic organisms
- Detection from patient specimens or primary cultures
- Inferred antimicrobial susceptibility and virulence

- PCR/amplification based tests
- MALDI-TOF mass spectrometry
- Next Generation Sequencing

Genomic data may inform treatment decisions to improve patient care

New technologies:
Whole genome sequencing

HOUSTON Methodist RESEARCH INSTITUTE

- First bacteria genome
  - *H. influenzae* strain Rd
    - Published in 1995
    - > 1 year
    - > 1 million US dollars

Flow cell for the Illumina MiSeq sequencer.

- Today,
  - 1 or a few genomes can be generated within 1 day
  - 384 genomes can be generated in 2 weeks for ~$50/each

- The $10 microorganism genome will soon be a reality

Whole genome sequencing in the clinical laboratory

HOUSTON Methodist RESEARCH INSTITUTE

A day in the life of the microbiology laboratory

- 130 samples collected from 116 patient cultures on a single day
- Aerobic bacteria, anaerobic bacteria, acid-fast bacilli and fungi
- Isolated colonies and mixed samples

Long, et al, *J. Clin Microbiol*, 2013

Results: A day in the life of the microbiology laboratory

HOUSTON Methodist RESEARCH INSTITUTE

Whole genome sequencing:

- **Identified most "unknown" organisms**

(88.5% concordance with reference method)

- **Detected *Mycobacterium species* in two samples 10 days before conventional methods**
- **Identified most organisms present in mixed samples**
- **Detected organisms not recovered by our routine culture based methods (e.g. *Gardnerella vaginalis*)**

Long, et al, *J. Clin Microbiol*, 2013

Challenges: A day in the life of the microbiology laboratory

HOUSTON Methodist RESEARCH INSTITUTE

Whole genome sequencing could not identify:

- **10 organisms due to their absence from the reference database**

- **The lack of a comprehensive database of human pathogens was particularly problematic for medically important fungi**

➢ ***Important Research Opportunity***

Long, et al, *J. Clin Microbiol*, 2013

Whole genome sequencing of microorganisms in our clinical lab

HOUSTON Methodist RESEARCH INSTITUTE

First run as a routine clinical test: June 26, 2013

- ~400 genome sequences analyzed in the validation study
- Bacteria, fungi and influenza A virus
- >600 genomes generated as an ordered test

Whole genome sequencing of microorganisms in our clinical lab

HOUSTON Methodist RESEARCH INSTITUTE

Routine cases

- **Organisms that are slow growing, difficult to identify or can not be identified using culture based techniques**
  - ***Mycobacterium species* and fungus**
- **Organisms requiring further classification for optimal patient care or public health efforts**
  - **Serotyping *Salmonella enterica* and *Neisseria meningitidis***

Whole genome sequencing of microorganisms in our clinical lab

HOUSTON Methodist RESEARCH INSTITUTE

Special cases

- **Organisms with unusual antimicrobial resistance**
  - **Investigated the molecular basis of ceftaroline resistance in *S. aureus***

    (Long, et al., *Antimicrob Agents Chemother*, 2014)

High level resistance to ceftaroline

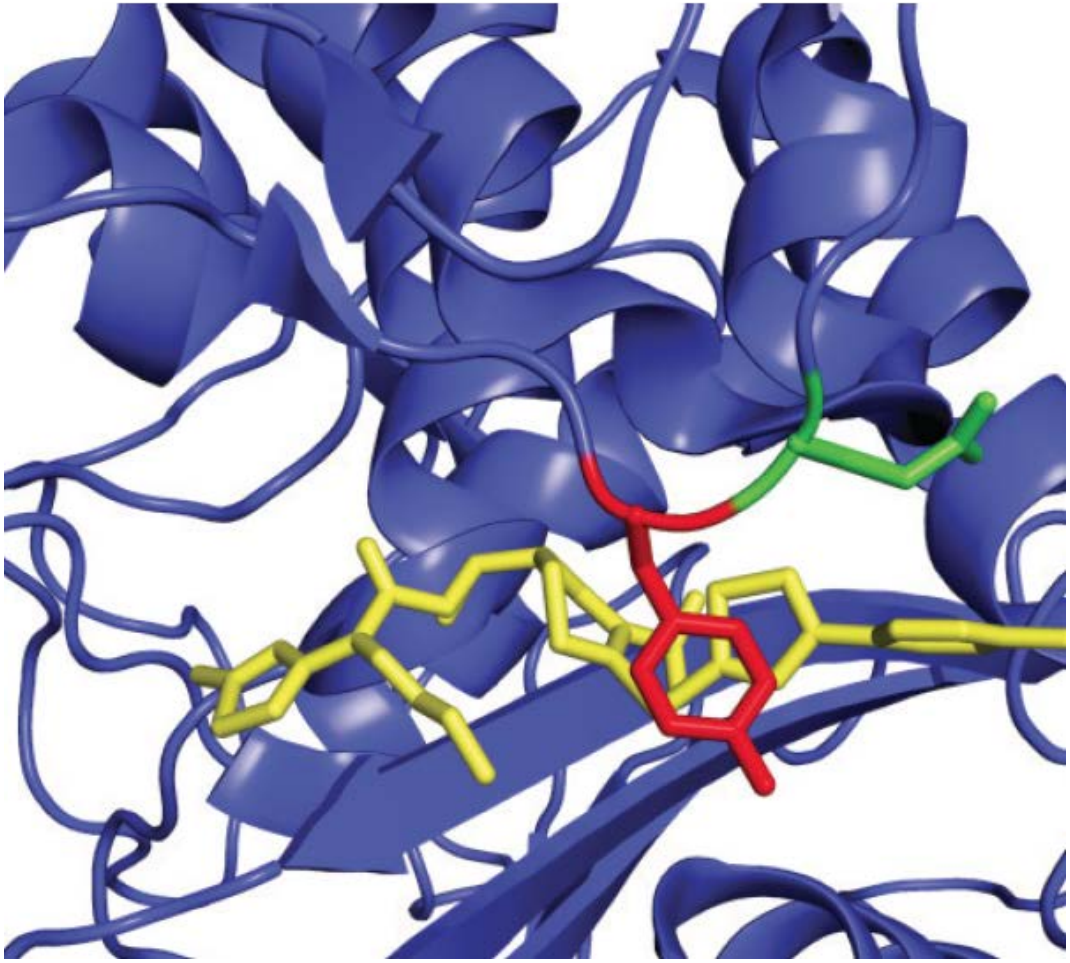
HOUSTON Methodist RESEARCH INSTITUTE

- **We investigated a highly ceftaroline resistant strain of *S. aureus* recovered from a cystic fibrosis patient.**
- **Ceftaroline is a recently approved anti-staphylococcal agent with no reports of highly resistant stains.**
- **Some reports suggest that ceftaroline has a structure making development of resistance unlikely.**

We used whole genome sequencing to discover the molecular basis of ceftaroline resistance

Whole genome sequencing identified the genetic basis of ceftaroline resistance

HOUSTON Methodist RESEARCH INSTITUTE



Genome sequencing identified mutations altering 2 amino acids in the active site of the penicillin binding protein 2a.

These data provide a molecular explanation for an unexpected antibiotic resistance phenotype.

Long, et al, *Antimicrob Agents Chemother*, 2014

Whole genome sequencing of microorganisms in our clinical lab

HOUSTON Methodist RESEARCH INSTITUTE

Special cases

- **Organisms causing unusually severe or unexpected clinical presentations**
  - **Investigated a fatal anthrax-like infection**

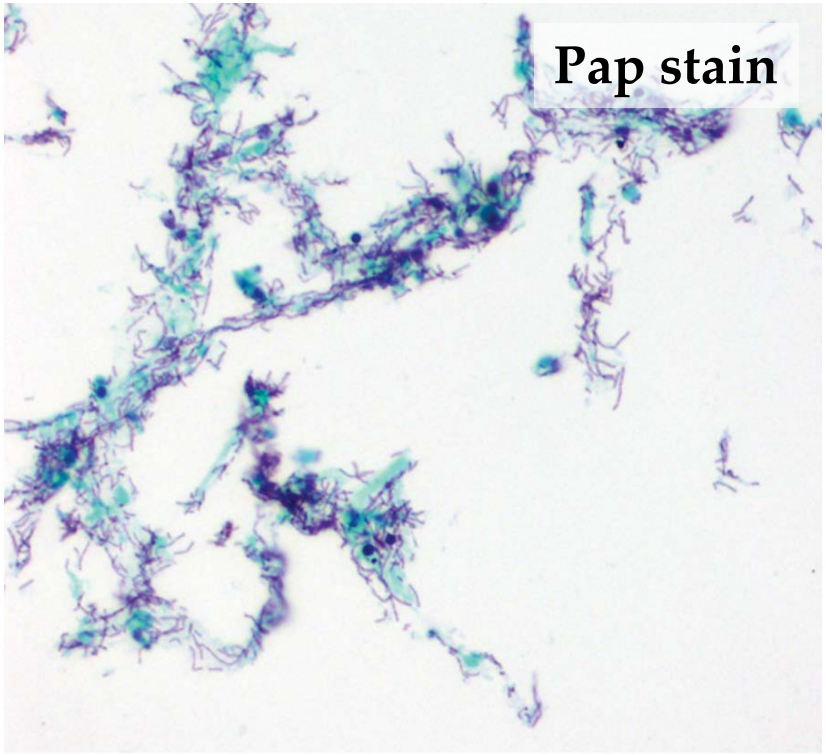
    (Wright, et al., *Arch Pathol Lab Med*, 2011)

Previously healthy male with severe pneumonia and death within 72 h

HOUSTON Methodist RESEARCH INSTITUTE

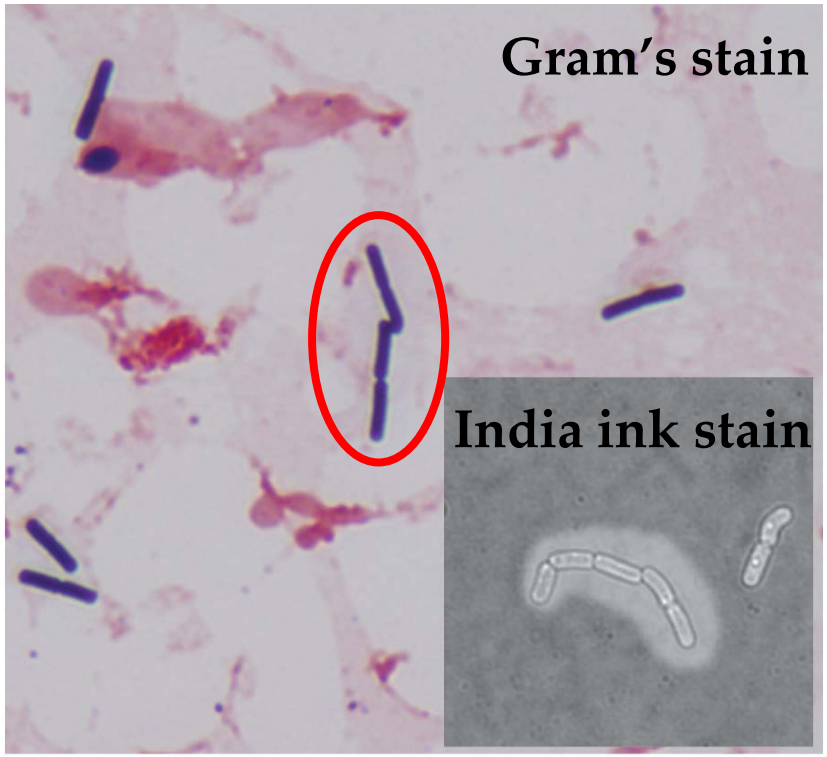
BAL on presentation

with extremely high numbers of rod-shaped bacteria



Blood culture

pure culture of *Bacillus cereus* group organisms



Wright, et al., *Arch Pathol Lab Med*, 2011

The *Bacillus cereus* group troika

HOUSTON Methodist RESEARCH INSTITUTE

Bacillus anthracis

Bacillus cereus

Bacillus thuringiensis

- **Very closely related**
- **Difficult to classify using conventional microbiology techniques**
- **Mobile genetic elements (plasmids) are crucial for virulence**

We performed whole genome sequencing to investigate the molecular basis of this severe, rapidly fatal infection.

Whole genome sequencing identified the causative agent

HOUSTON Methodist RESEARCH INSTITUTE

- **Identified etiologic agent of the fatal Anthrax-like infection**
  - ***B. cereus* with anthrax toxin genes, not *B. anthracis***
- **Ruled out the likelihood of bioterrorism**
  - **No foreign DNA or evidence of other genetic manipulation**
- **Guided infection control response**
  - **No cases of secondary transmission**

Wright, et al., *Arch Pathol Lab Med*, 2011

Whole genome sequencing of outbreak pathogens

HOUSTON Methodist RESEARCH INSTITUTE

- **Outbreaks caused by naturally occurring or nefariously introduced microbial agents can rapidly decimate a population and cause a public health emergency.**

- **Large hospital based laboratories must be highly vigilant since they will likely be the first to recover the pathogen and perhaps recognize the potential for an outbreak.**

- **To this end, our clinical laboratory has validated and implemented whole genome sequencing as a clinical test.**

HOUSTON Methodist RESEARCH INSTITUTE

To assess our clinical laboratory's ability to rapidly respond to a large-scale emergency outbreak scenario, we contrived a "fire drill" using *emm59* Group A Streptococcus (GAS) strains collected by the US CDC.

- **Serotype *emm59* GAS are an uncommon cause of human infection.**
- **A hypervirulent *emm59* clone recently emerged to cause hundreds of severe invasive infections across Canada and some parts of the US.**
- **Epidemiology studies conducted by the US CDC discovered an increased number of *emm59* GAS infections in the state of Minnesota.**

Olsen, et al, *J Clin Microbiol,* 2014

HOUSTON Methodist RESEARCH INSTITUTE

This study was constructed as a best case scenario:

1. **A high quality, closed, well-annotated, *emm59* GAS genome is publically available.**

2. **The research team was familiar with this organism and had sequenced the genome of >500 strains recovered in Canada.**

3. **Compared to other pathogens such as *K. pneumoniae*, GAS has a small genome and few mobile genetic elements.**

4. **The US CDC had already collected the strains, and collaborators at the Minnesota Department of Public Health were ready to act on the genomic findings.**

HOUSTON Methodist RESEARCH INSTITUTE

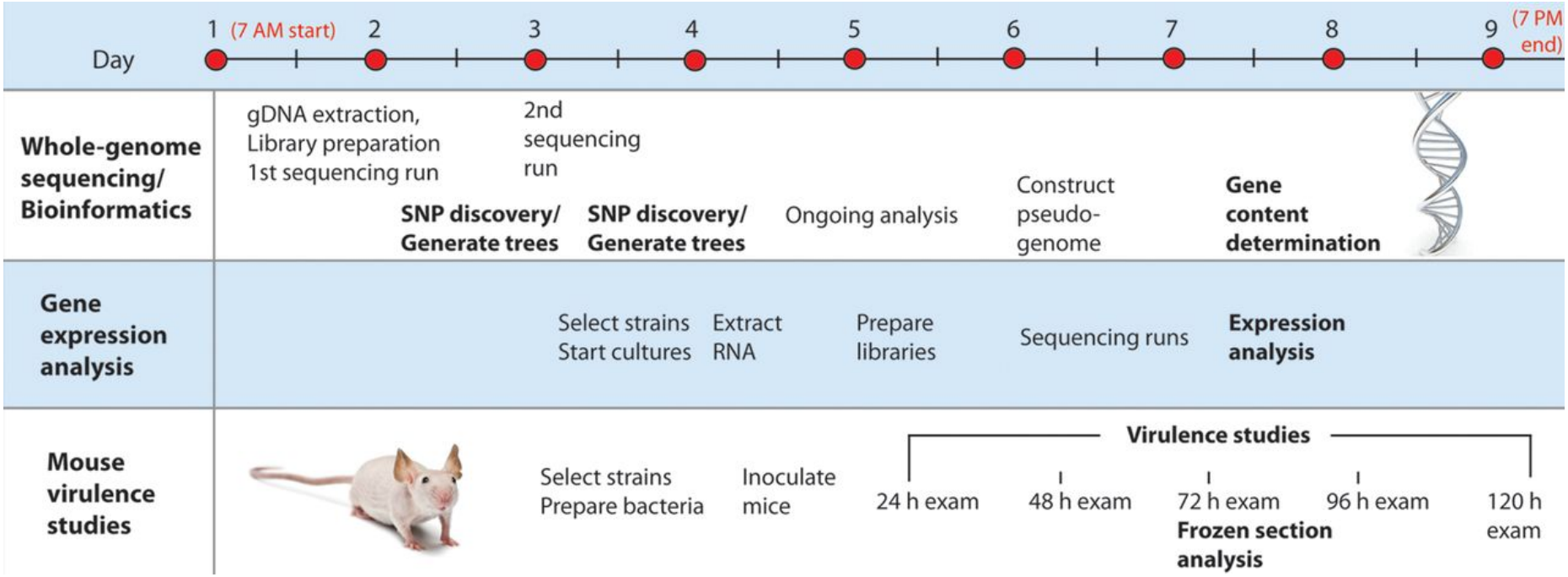
Are these US infections clonally related to the *emm59* GAS outbreak in Canada?

If so, what additional information can be generated in the clinical laboratory to guide patient care?

We performed whole genome sequencing, genome-wide transcript analysis and mouse virulence studies.

Timeline of the rapid-response exercise performed in our clinical laboratory

HOUSTON Methodist RESEARCH INSTITUTE

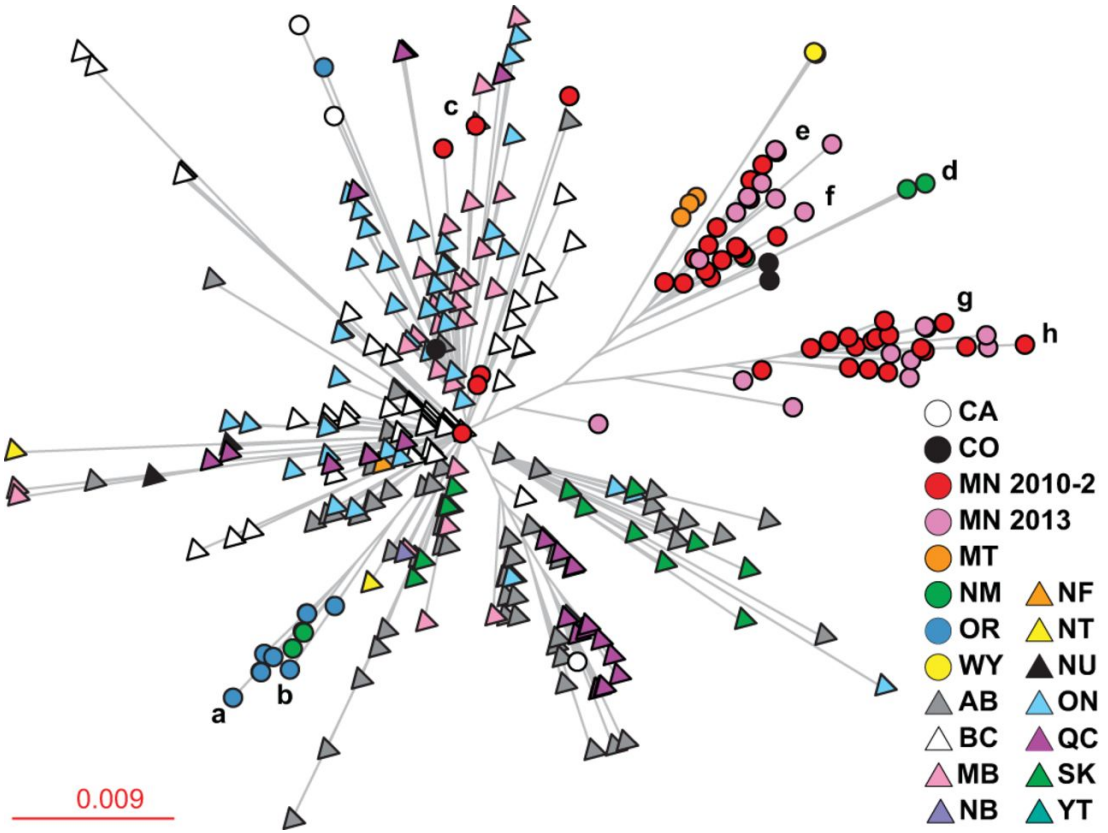


Can we generate actionable information in a clinically relevant timeframe?

Olsen, et al, *J Clin Microbiol,* 2014

Whole genome sequence analysis of *emm59* GAS strains recovered in the United States

HOUSTON Methodist RESEARCH INSTITUTE



Within 3 days, we discovered:

US strains are clonally related to Canadian outbreak strains.

Five sets of genomically indistinguishable strains.

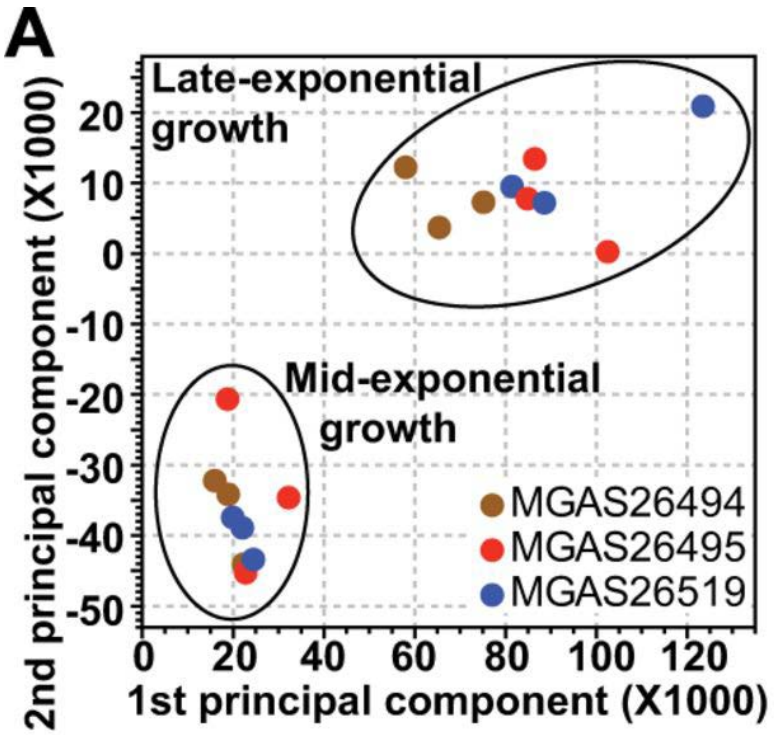
Strains recovered in Minnesota are allied with strains from two different Canadian provinces.

The genomics data informed the epidemiology investigation!

Olsen, et al, *J Clin Microbiol,* 2014

Gene expression analysis of *emm59* GAS strains recovered in the United States

HOUSTON Methodist RESEARCH INSTITUTE



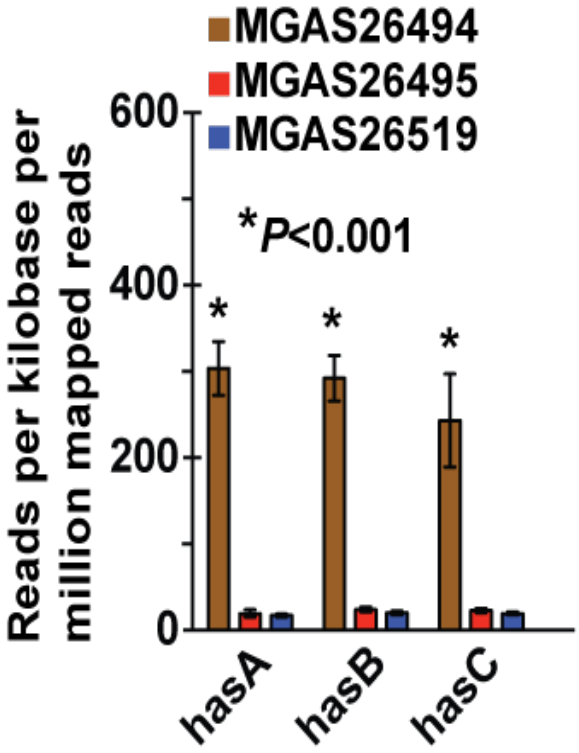
3 strains selected for RNAseq:

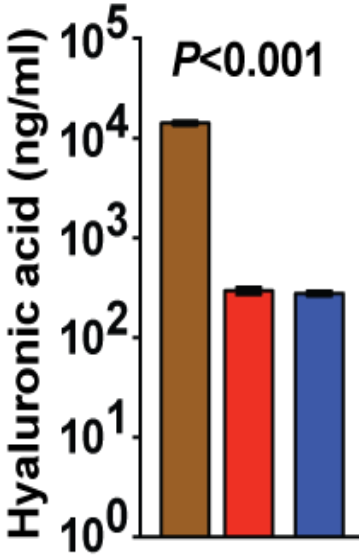
1. Strains recovered from Minnesota to emulate a regional outbreak.
2. Strains were located on different branches of the phylogenetic tree.
3. Strains had wild-type alleles for all major regulators.

Within 8 days, we demonstrated the US strains have a highly similar genome-wide expression profile; however, they were not identical!

Olsen, et al, *J Clin Microbiol,* 2014

Gene expression analysis of *emm59* GAS strains recovered in the United States





RNAseq revealed that:

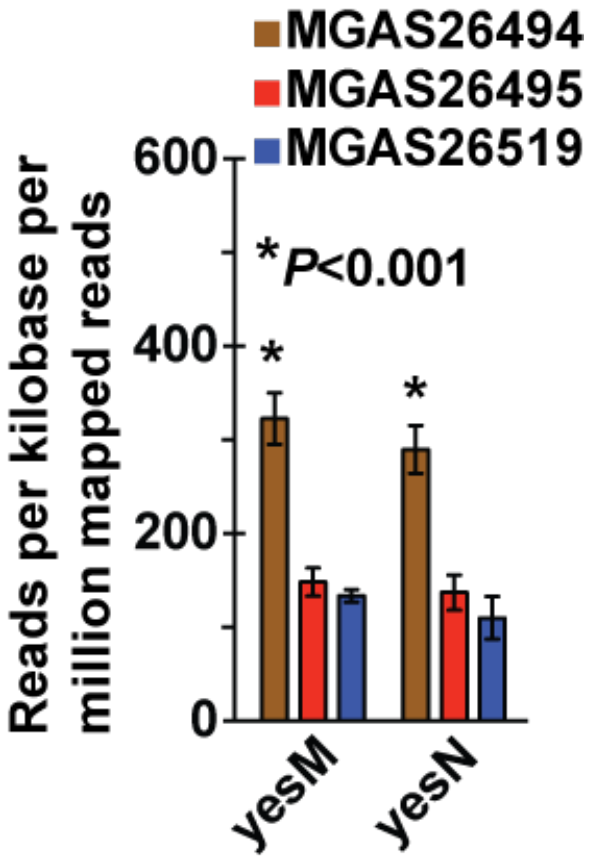
-Strain MGAS26494 unexpectedly overexpresses the genes encoding the hyaluronic acid capsule virulence factor.

-On reanalysis of the genome sequence, we discovered a deletion in the *hasABC* promoter region.

The transcriptomics data informed the genome analysis!

Olsen, et al, *J Clin Microbiol,* 2014

Gene expression analysis of *emm59* GAS strains recovered in the United States



RNAseq also revealed that:

-Strain MGAS26494 overexpresses the genes encoding the *yesM/N* two component system.

-On reanalysis of the genome sequence, we discovered a mixed population, with ~50% of reads having a polymorphism in *yesM/N*.

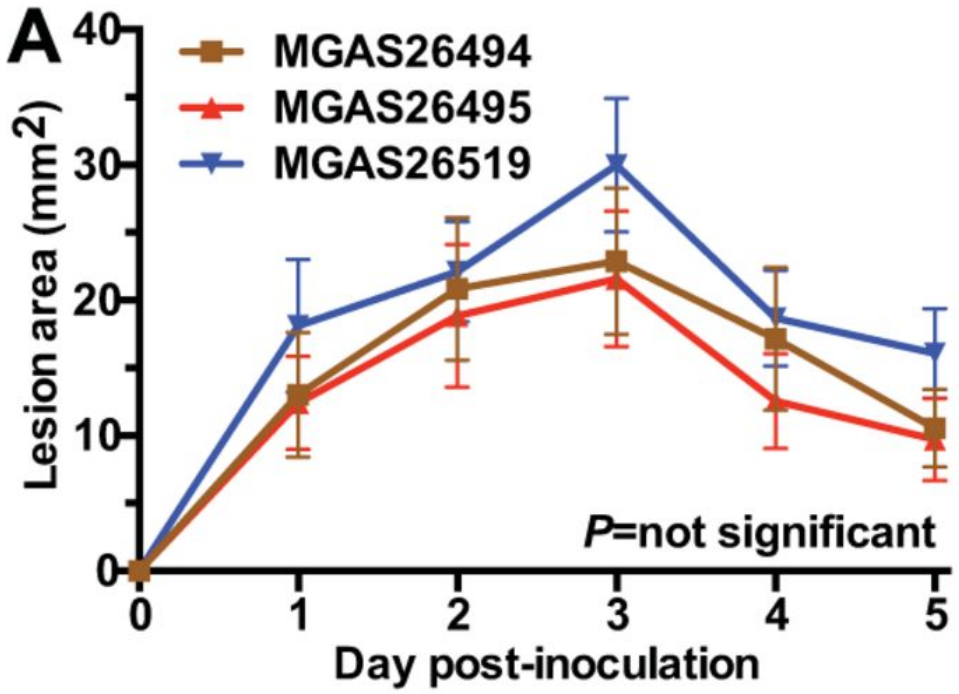
-This mutation also explains the altered expression of 8 genes regulated by YesM/N.

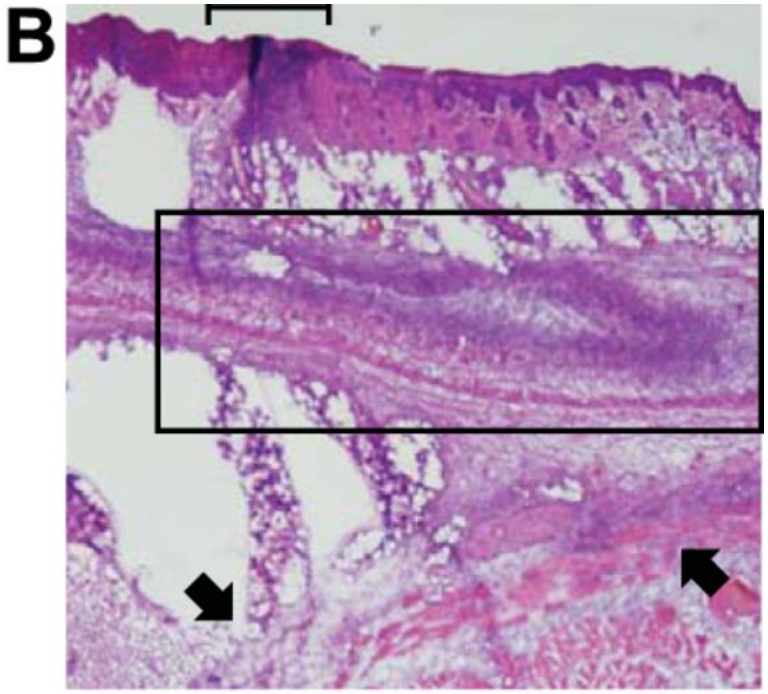
The transcriptomics data informed the genome analysis!

Olsen, et al, *J Clin Microbiol,* 2014

Mouse model of invasive infection using *emm59* GAS strains recovered in the United States

HOUSTON Methodist RESEARCH INSTITUTE





The animal model could be used to rapidly investigate strain virulence or test new vaccines, diagnostics and therapies!

Olsen, et al, *J Clin Microbiol,* 2014

Lessons Learned:

HOUSTON Methodist RESEARCH INSTITUTE

- **Whole genome sequencing can be readily integrated into the routine workflow of a clinical laboratory to support patient care decisions.**

- **We created a standard operating procedure for using genomics, transcriptomics and animal studies as part of our disaster preparedness plan.**

- **We implemented process improvements.**

Acknowledgments

Houston Methodist Hospital

Steve Beres, PhD

Connie Cantu

Priyanka Kachroo, PhD

S. Wesley Long, MD, PhD

James Musser, MD, PhD

Misu Sanson, MD

Chandni Valson

Public Health Ontario

Nahuel Fittipaldi

Minnesota Department of Health

Ruth Lynfield, MD

Kathryn Como-Sabetti, MPH

Centers for Disease Control and Prevention

Chris Van Beneden, MD, MPH