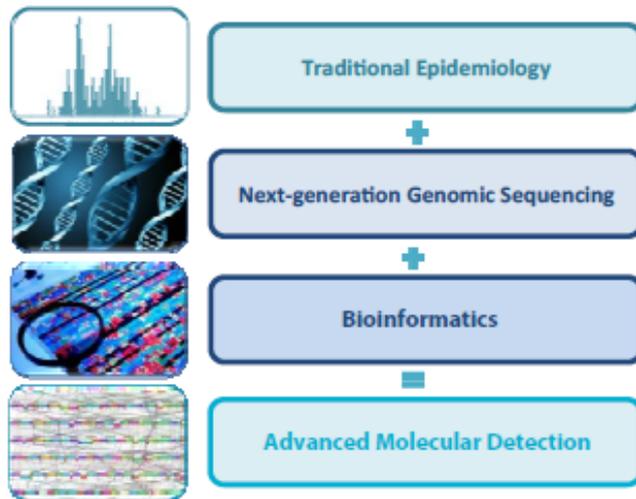


# Whole Genome Sequencing

Delaware Public Health Laboratory

# Background-AMD

Advanced molecular detection (AMD) combines the latest pathogen identification technologies with enhanced capabilities in bioinformatics and advanced epidemiology to be more effective in understanding, preventing, and controlling infectious diseases.



Examples of public health benefits of AMD include more precise and accurate ways of:

- Diagnosing known and emerging infections
- Detecting and responding to outbreaks
- Understanding, characterizing, and controlling antibiotic resistance
- Developing and targeting prevention measures, including vaccines

- Feedback from a [Bioinformatics Blue Ribbon Panel in 2011](#) that deemed that the CDC lacked "the organization, expertise, and computational infrastructure needed for data management and analysis of genomic information it is now capable of acquiring."<sup>1</sup>

1-<https://www.genomeweb.com/sequencing/cdc-scales-use-ngs-technologies-publishes-first-wgs-sequencing->

[data?utm\\_source=SilverpopMailing&utm\\_medium=email&utm\\_campaign=Sequencing%20Bulletin:%20Genentech,%20BioNTech%20Deal%20Pushes%20Personalized%20Cancer%20Vaccines%20Toward%20Market%20-%202010/04/2016%2004:05:00%20PM](https://www.genomeweb.com/sequencing/cdc-scales-use-ngs-technologies-publishes-first-wgs-sequencing-data?utm_source=SilverpopMailing&utm_medium=email&utm_campaign=Sequencing%20Bulletin:%20Genentech,%20BioNTech%20Deal%20Pushes%20Personalized%20Cancer%20Vaccines%20Toward%20Market%20-%202010/04/2016%2004:05:00%20PM)

# Concept

- *Imagine doing a 10,000-piece jigsaw puzzle in the time it takes to finish a 100-piece puzzle. Apply that to infectious disease control, and that's AMD at work. Now imagine putting together that 10,000-piece puzzle when key pieces are missing, disease is spreading, and people are dying. AMD gives CDC scientists the "key pieces" they need to protect people from ever-changing infectious disease threats.*

# DPHL Role

- Replace Foodborne enteric pathogen surveillance testing (Pulse Field Gel Electrophoresis/PFGE) with Whole Genome Sequencing (WGS) utilizing Next Generation Sequencing Technology (NGS)
- CDC using funds to WGS ~100% of all foodborne pathogens.
  - Prioritize certain organisms (Listeria, Non-O157 Shiga-toxin *E. coli* (STEC), representative isolates from outbreaks, rare serotypes)
- Currently, DPHL is submitting Salmonella Newport PFGE #0061 to CDC for WGS.

# Instrument

- Illumina MiSeq- installed April 2016.
- Cost analysis- ~\$150 to WGS bacterial organism<sup>1</sup>
- Must be certified to perform CDC methods (not CLIA regulated)
- Two Microbiologists pending certification at current time
- Funding disseminated to every Public Health Lab to purchase MiSeq before end of 2016
- By 2018, PFGE will be obsolete and no longer performed by any Public Health Lab, only WGS.

# System Parameters

- DPHL will use Kit v2, 2x150 and 2x250 runs.
- Anticipate sequencing up to 80 million bases (Mb) per run (~16 E. coli, Salmonella, or Shigella isolates)

## MiSeq System Performance Parameters

### MiSeq Reagent Kit v2

Read Length	Total Time*	Output
1 × 36 bp	~4 hours	540–610 Mb
2 × 25 bp	~5.5 hours	750–850 Mb
2 × 150 bp	~24 hours	4.5–5.1 Gb
2 × 250 bp	~39 hours	7.5–8.5 Gb

### Reads Passing Filter†

Single Reads	12–15 M
Paired-End Reads	24–30 M

### Quality Scores††

> 90% bases higher than Q30 at 1 × 36 bp  
 > 90% bases higher than Q30 at 2 × 25 bp  
 > 80% bases higher than Q30 at 2 × 150 bp  
 > 75% bases higher than Q30 at 2 × 250 bp

### MiSeq Reagent Kit v3

Read Length	Total Time*	Output
2 × 75 bp	~21 hours	3.3–3.8 Gb
2 × 300 bp	~56 hours	13.2–15 Gb

### Reads Passing Filter†

Single Reads	22–25 M
Paired-End Reads	44–50 M

### Quality Scores††

> 85% bases higher than Q30 at 2 × 75 bp  
 > 70% bases higher than Q30 at 2 × 300 bp

\* Total times include cluster generation, sequencing, and base calling on a MiSeq system enabled with dual surface scanning.

† Install specifications based on Illumina PhiX control library at supported cluster densities between 865–985 k/mm<sup>2</sup> clusters passing filter for v2 chemistry and 1200–1400 k/mm<sup>2</sup> clusters passing filter for v3 chemistry. Actual performance parameters can vary based on sample type, sample quality, and clusters passing filter.

†† The percentage of bases > Q30 is averaged across the entire run.

bp – base pairs, Mb – megabases, Gb – gigabases, M – millions

# Future AMD/CDC/Public Health

- Microbiology Lab Manager attended AMD Day, September 26, 2016.
- Initial WGS: Foodborne Pathogens
- Next step: *M. tuberculosis* complex, Influenza Virus, Hep C, HIV, *N. gonorrhoeae*, etc.
- CDC CLIA certified species level identification for Enteric bacteria using Average Nucleotide Identity (ANI)
- Curating national genome databases for identification, virulence factors, resistance markers, outbreak investigation of all pathogenic organisms (Bacterial, Viral, Parasitic, Fungal)
  - MicrobeNet (ex. Elizabethkingia)
- Direct specimen detection (metagenomics)-NO MORE CULTURE

# Questions?

