

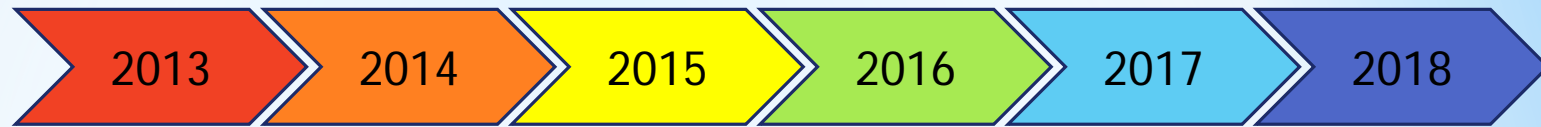
# **Integration of WGS into an Established Public Health Laboratory Testing Process – Virginia's Perspective**



**Leigh-Emma Lion**

**Molecular Detection & Characterization Principal Scientist**

# DCLS WGS CAPACITY

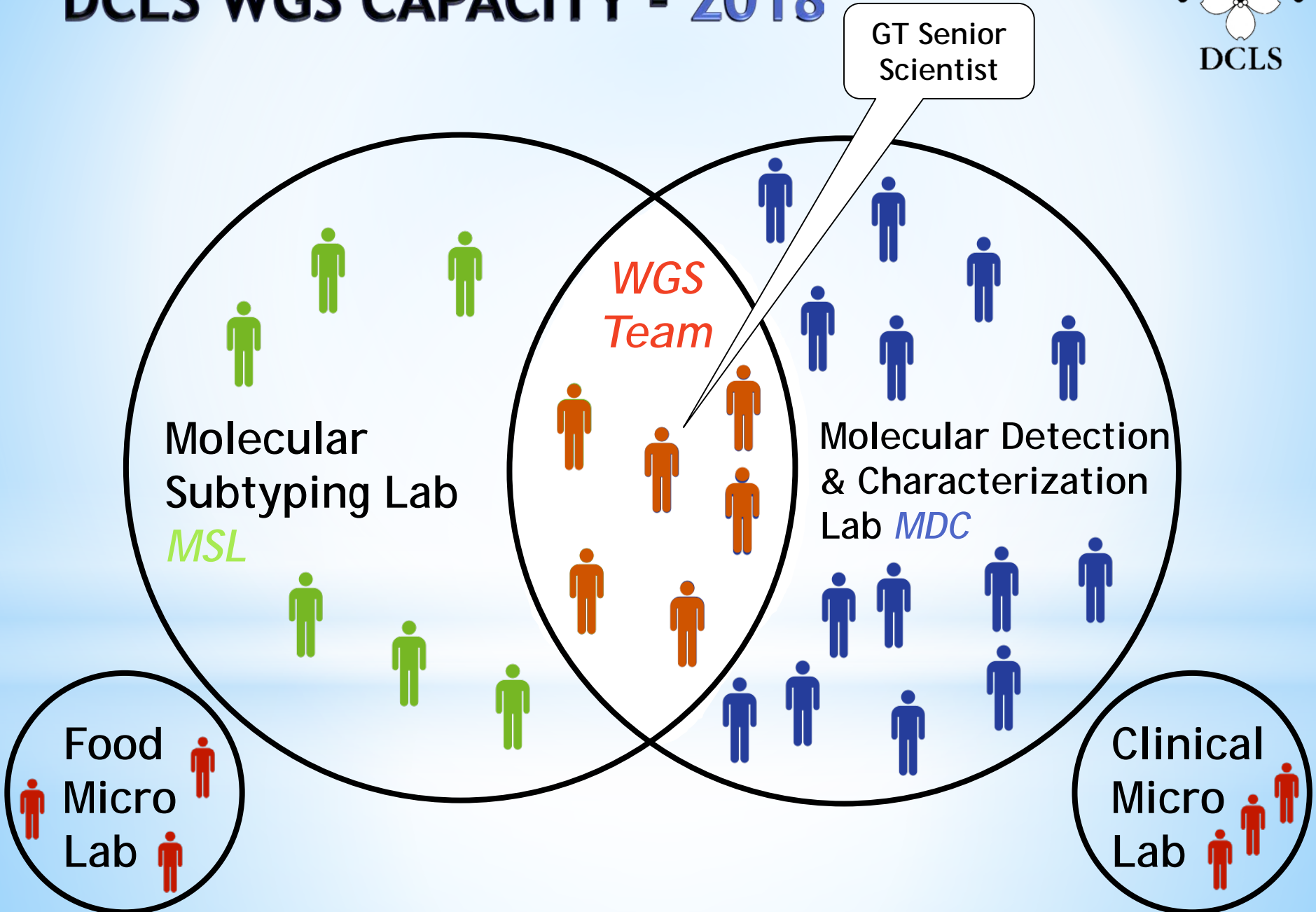


	2013	2014	2015	2016	2017	2018
Staff trained to sequence						
Staff trained to extract						
Total samples sequenced	12	498	738	1063	2365	1677 so far
GT samples sequenced	12	461	292	220	224	94 so far
Samples pulsed	1294	1596	1661	1636	2054	1221 so far
Instruments						
Bioinformatics staff						
GenomeTrakr						
PulseNet						
Tuberculosis						

Actively participating employee

Trained employee working on other projects

# DCLS WGS CAPACITY - 2018



# DCLS PFGE WORKFLOW OVERVIEW



Isolate identified  
from clinical,  
food or  
environmental  
sample

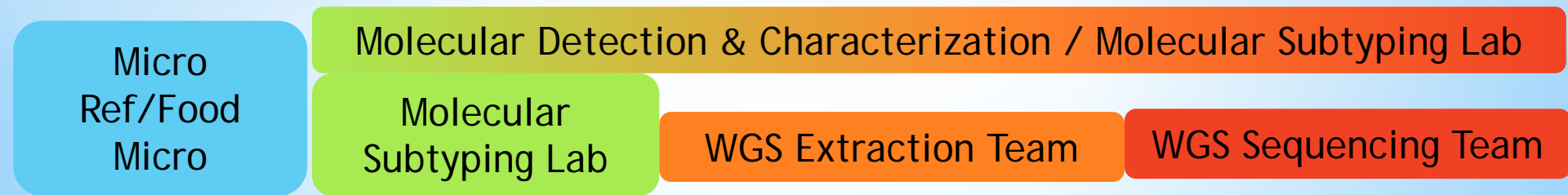
PFGE

Data review  
and data  
sharing

Clinical Micro/  
Food Micro

Molecular Subtyping  
Lab (MSL)

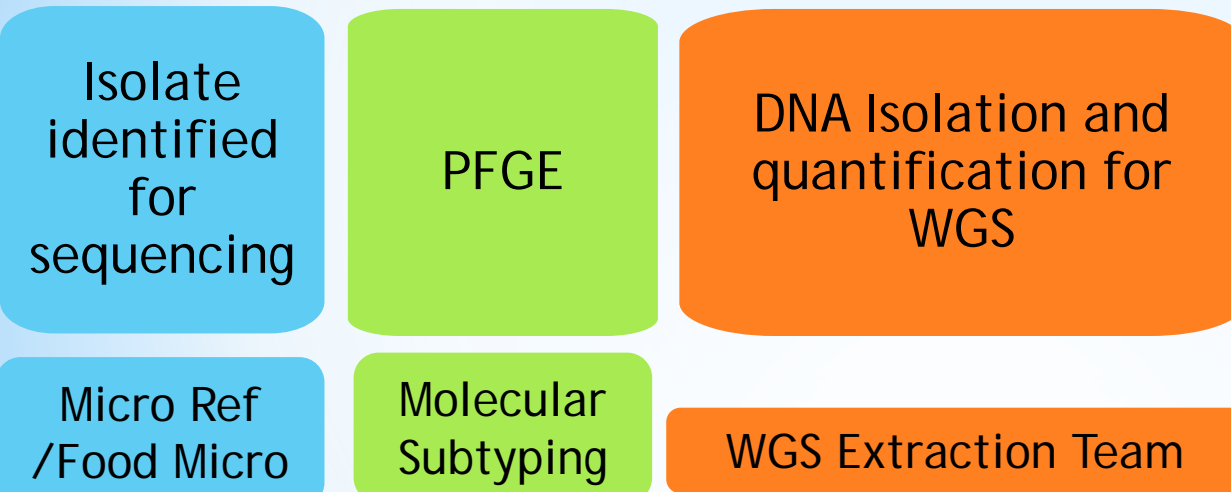
# DCLS WGS WORKFLOW OVERVIEW



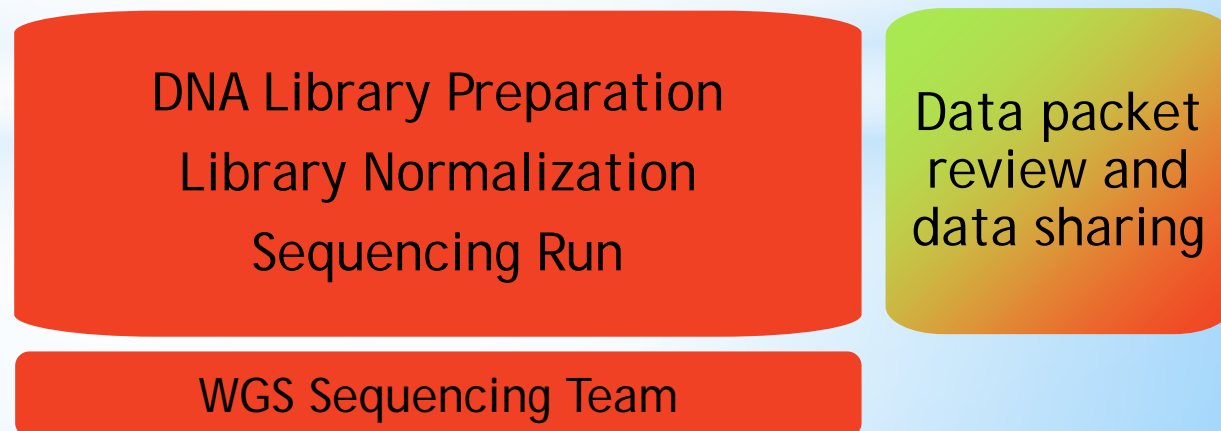
# DCLS WGS WORKFLOW PROCESS DIVISION



## Identification and Extraction Process



## Sequencing and Data Analysis Process

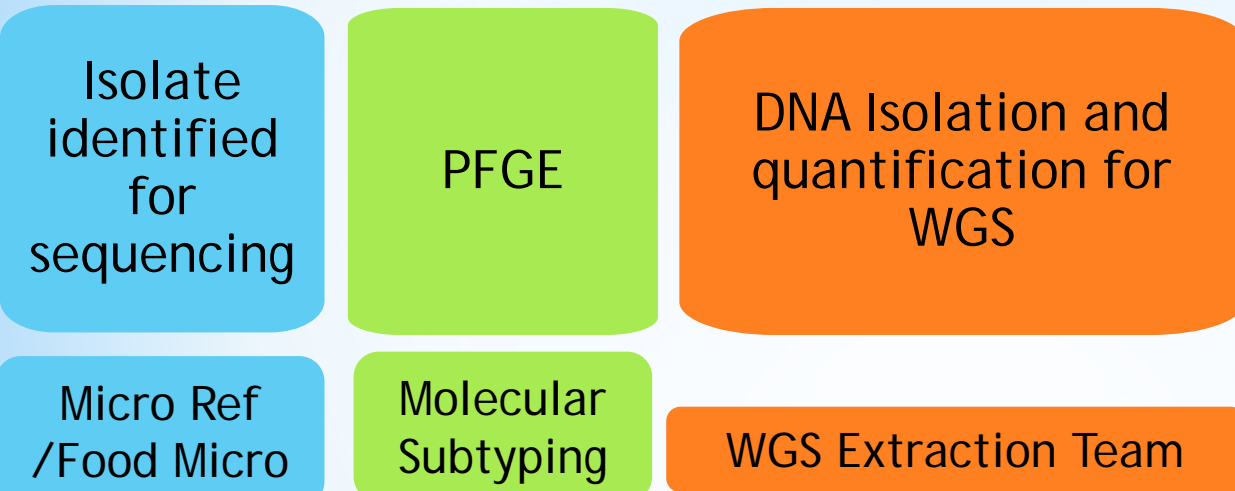




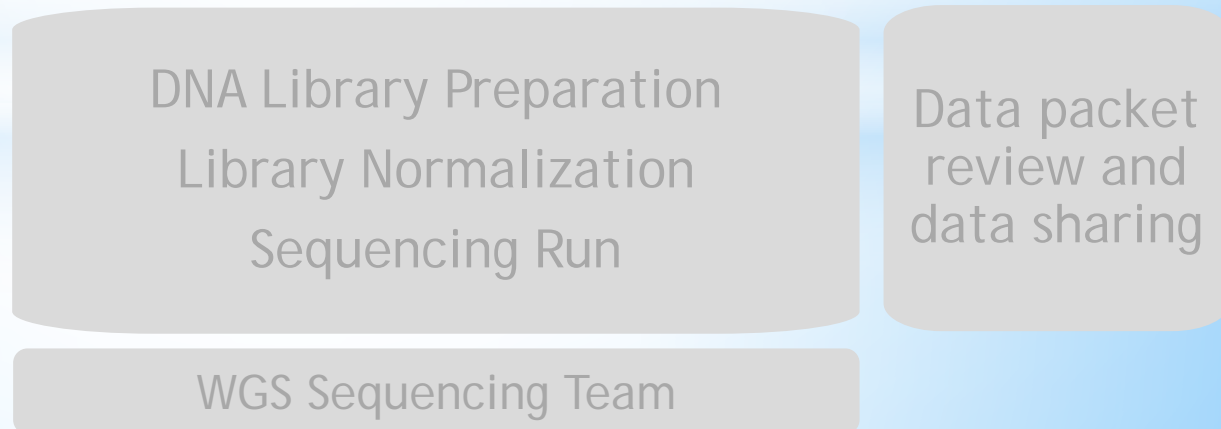
# DCLS WGS WORKFLOW PROCESS DIVISION



## Identification and Extraction Process



## Sequencing and Data Analysis Process



# DCLS WGS SAMPLE MANAGEMENT

## Extraction



Organism	Serotype	Outbreak Code	PFGE/WGS	Source type	WGS_ID	Other ID	request receive date	DNA extraction date	Freezer location	Problems	DNA quant. value
Salmonella	Undetermined		PFGE/WGS	Human	PNUSAS053407			8/31/2018			29.1
Salmonella	Undetermined		PFGE/WGS	Human	PNUSAS053408			8/31/2018			17.2
Salmonella	Typhimurium		PFGE/WGS	Human	PNUSAS053409			8/31/2018			8.32
Salmonella	Typhimurium		PFGE/WGS	Human	PNUSAS053410			8/31/2018			5.47
Salmonella	Panama		PFGE/WGS	Human	PNUSAS053404			8/31/2018			30.4
Salmonella	Panama		PFGE/WGS	environmenta	VA-WGS-18032	E18M024327-01		8/31/2018	46-D3		16.6
Salmonella	Mississippi		PFGE/WGS	environmenta	VA-WGS-18058	E18M023878-01		8/31/2018	46-D4		18
Salmonella	Enteritidis	VA cluster	PFGE/WGS	Human	PNUSAS053413			8/31/2018	46-D5		5.53
Salmonella	Undetermined	VA cluster	PFGE/WGS	Human	PNUSAS053414			8/31/2018	46-D6		11.7
Salmonella	Enteritidis		PFGE/WGS	Human	PNUSAS053415			8/31/2018			9.83
Salmonella	London		PFGE/WGS	Human	PNUSAS053416			8/31/2018			11.9
Salmonella	Poona		PFGE/WGS	Human	PNUSAS053417			8/31/2018			12.7
Salmonella	Undetermined		PFGE/WGS	Human	PNUSAS053418			8/31/2018			10.4
Salmonella	Undetermined		PFGE/WGS	Human	PNUSAS053421			8/31/2018			9.04
Salmonella	Typhimurium	VA cluster	PFGE/WGS	Human	PNUSAS053422			8/31/2018	46-D7		7.4
Salmonella	Typhimurium		PFGE/WGS	Human	PNUSAS053424			8/31/2018			12.6
Salmonella	Enteritidis		PFGE/WGS	Human	PNUSAS053411			8/31/2018			4.84
Salmonella	Javiana		PFGE/WGS	Human	PNUSAS053429			8/31/2018			
Salmonella	Undetermined		PFGE/WGS	Human	PNUSAS053433			8/31/2018			
Salmonella	Hartford		PFGE/WGS	Human	PNUSAS053434			8/31/2018			
Salmonella	Hartford		PFGE/WGS	Human	PNUSAS053435			8/31/2018			
Salmonella	Bareilly		PFGE/WGS	Human	PNUSAS053436			8/31/2018			
E.coli	non-O157		PFGE/WGS	Human	PNUSAE018097			8/31/2018			
E.coli	non-O157		PFGE/WGS	Human	PNUSAE018099			8/31/2018			
E.coli	non-O157		PFGE/WGS	Human	PNUSAE018100			8/31/2018			
E.coli	non-O157		PFGE/WGS	Human	PNUSAE018101			8/31/2018			
E.coli	non-O157		PFGE/WGS	Human	PNUSAE018102			8/31/2018			

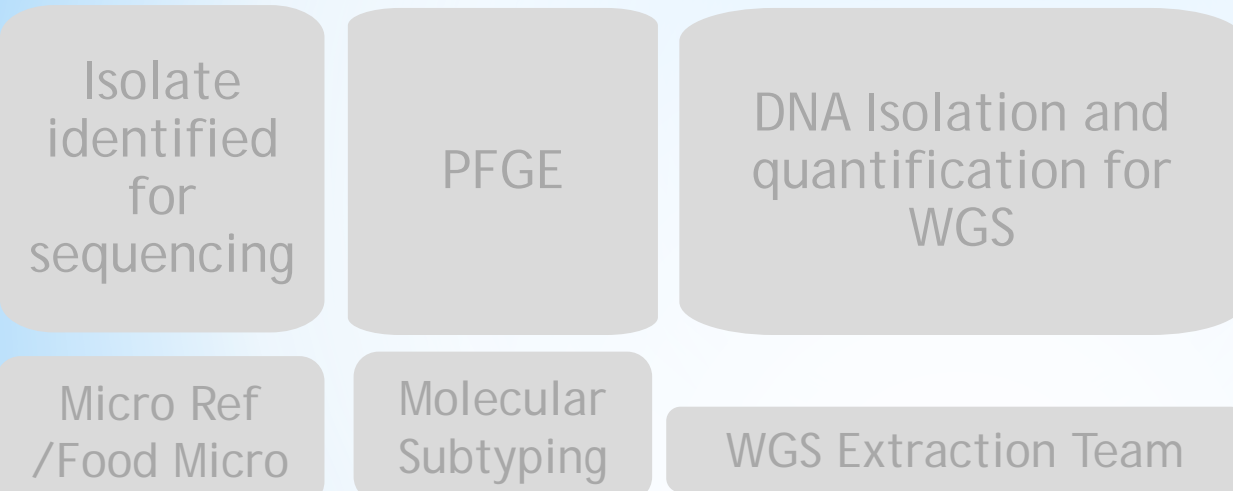




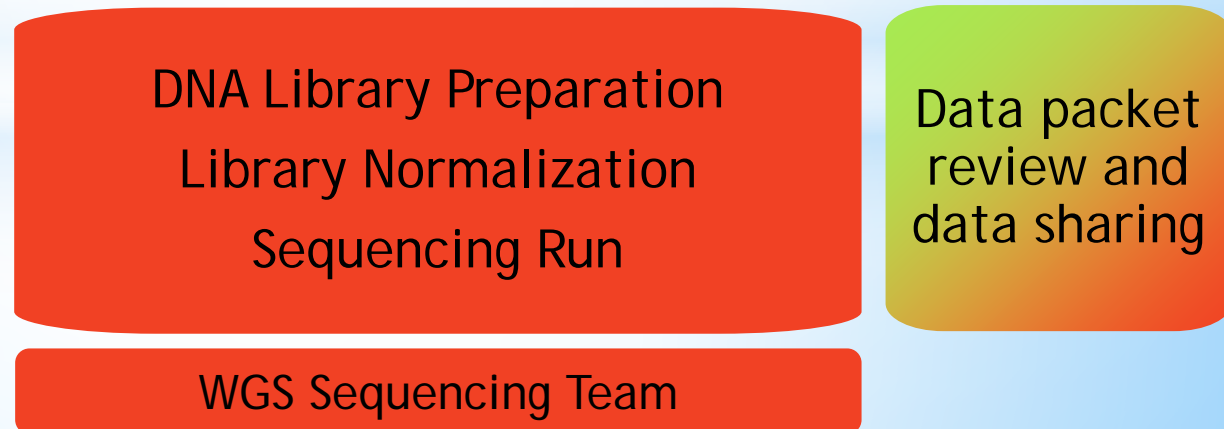
# DCLS WGS WORKFLOW PROCESS DIVISION



## Identification and Extraction Process



## Sequencing and Data Analysis Process



# DCLS WGS SAMPLE MANAGEMENT

## Sequencing



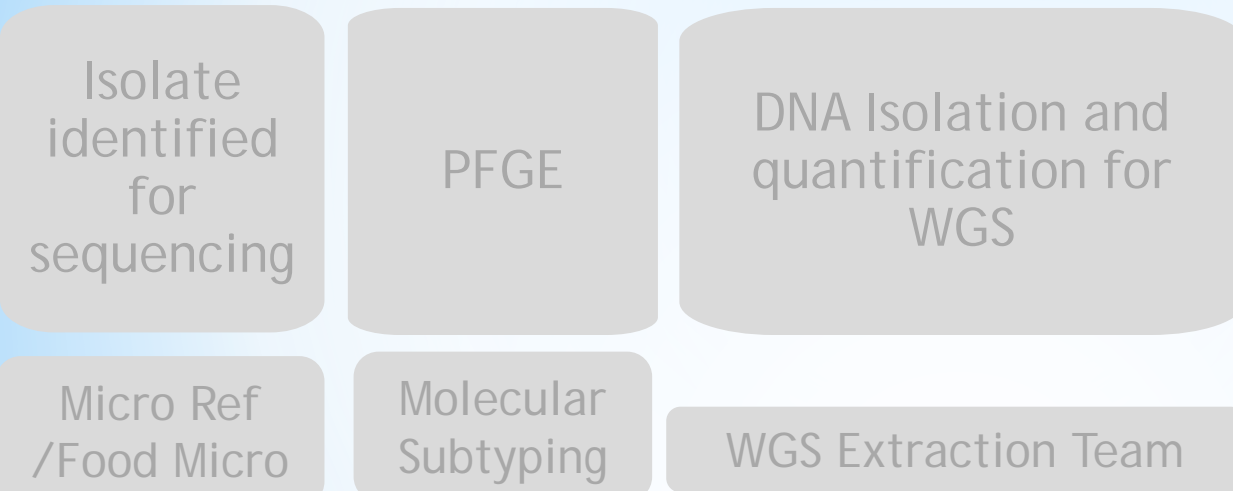
Organism	WGS_ID	Other ID	request receive date	DNA extraction date	Freezer location	Problems	DNA quant. value	Index 1	Index 2	Library Date	MiSeq run date	MiSeq Run #	NCBI Accession	SRR #
Salmonella	PNUSAS053407			8/31/2018			29.1							
Salmonella	PNUSAS053408			8/31/2018			17.2							
Salmonella	PNUSAS053409			8/31/2018			8.32							
Salmonella	PNUSAS053410			8/31/2018			5.47							
Salmonella	PNUSAS053404			8/31/2018			30.4							
Salmonella	VA-WGS-18032	← E18M024327-01		8/31/2018	46-D3		16.6	N712	S502	9/5/2018	9/5/2018	VA-M04590-180905		
Salmonella	VA-WGS-18058	← E18M023878-01		8/31/2018	46-D4		18	N709	S503	9/5/2018	9/5/2018	VA-M04590-180905		
Salmonella	PNUSAS053413	←		8/31/2018	46-D5		5.53	N710	S503	9/5/2018	9/5/2018	VA-M04590-180905		
Salmonella	PNUSAS053414	←		8/31/2018	46-D6		11.7	N711	S503	9/5/2018	9/5/2018	VA-M04590-180905		
Salmonella	PNUSAS053415			8/31/2018			9.83							
Salmonella	PNUSAS053416			8/31/2018			11.9							
Salmonella	PNUSAS053417			8/31/2018			12.7							
Salmonella	PNUSAS053418			8/31/2018			10.4							
Salmonella	PNUSAS053421			8/31/2018			9.04							
Salmonella	PNUSAS053422	←		8/31/2018			7.4							
Salmonella	PNUSAS053424			8/31/2018			12.6							
Salmonella	PNUSAS053411			8/31/2018			4.84							
Salmonella	PNUSAS053429			8/31/2018										
Salmonella	PNUSAS053433			8/31/2018										
Salmonella	PNUSAS053434			8/31/2018										
Salmonella	PNUSAS053435	←		8/31/2018										
Salmonella	PNUSAS053436	←		8/31/2018										
E.coli	PNUSAE018097	←		8/31/2018										
E.coli	PNUSAE018099			8/31/2018										
E.coli	PNUSAE018100			8/31/2018										
E.coli	PNUSAE018101			8/31/2018										
E.coli	PNUSAE018102			8/31/2018										

Highlighted samples = priorities

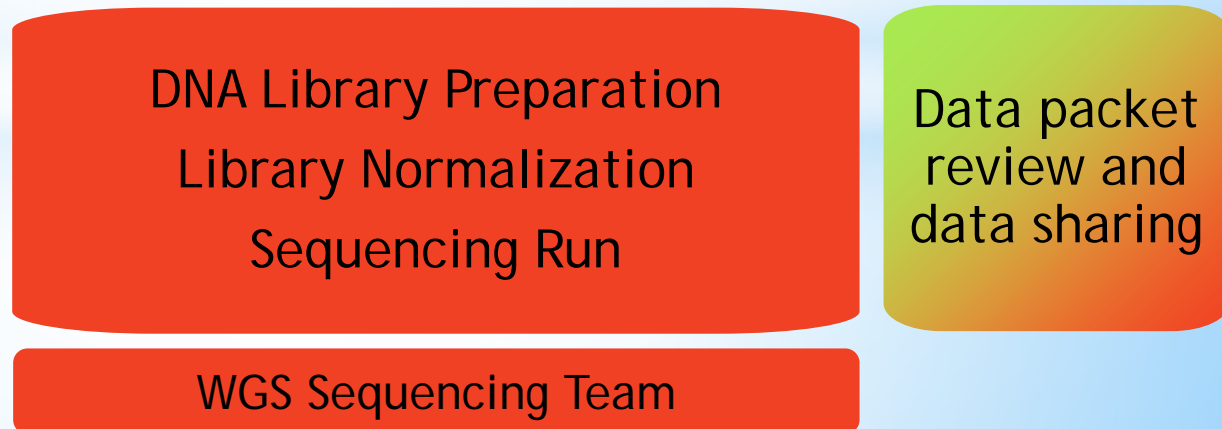
# DCLS WGS WORKFLOW PROCESS DIVISION



## Identification and Extraction Process



## Sequencing and Data Analysis Process



# DCLS WGS GT DATA APPROVAL



## WGS Run Review Checklist

APPROVED FOR SHARING \_\_\_\_\_

Run # \_\_\_\_\_  
Reviewer / Date \_\_\_\_\_

- Circle type(s) of samples included on the run: PN GT MTB Other
- Confirm that the log is accurate
  - PN isolates - Confirm WGSID emails match log printout (WGS ID and R# should be checked)
  - GT isolates - Confirm that ID's have been submitted and match log printout
  - MTB Confirm that TB log and LIMS match log printout
  - Confirm indices on log match BaseSpace sample sheet
  - Check run date and run number for accuracy in log
- Confirm top sheet of Qualtrax ID # 12711 is accurate
  - Confirm that R# and WGS ID match the log
  - Confirm organism matches the log
  - Confirm correct BioProject ID was used (N/A for MTB)
  - Confirm that the indices match the log
  - Verify the # of reads on BaseSpace or FastQ printouts (MTB)
  - Verify coverage calculations
  - Verify coverage minimum is correct per organism
  - Verify that the calculated coverage exceeds the minimum
  - Check FastQC printout for any samples that are less than 20% above the minimum
  - Confirm that the run metrics on the top sheet are correct using the BaseSpace Run and Lane Metrics printout or Run Summary (MTB)
- Confirm pages 2-5 of Qualtrax ID # 12711 are accurate
  - Confirm lots, calculations, temps and expirations are correct on additional pages of Qualtrax ID # 12711
  - Confirm Qubit control chart numbers have been entered.
- Confirm Qualtrax ID # 12658 is accurate
  - Confirm that the samples match the samples on the top sheet of Qualtrax ID # 12711.
  - Confirm organism matches the top sheet of Qualtrax ID # 12711.
  - Confirm that the genome sizes are correct.
  - Confirm that the input stock concentrations match the log.
  - Confirm that all normalization calculations are correct.
  - Confirm that the output DNA concentration matches the top sheet of Qualtrax ID # 12711.
  - Confirm that all pooling calculations are correct.
- Confirm QC, notify team that the run is ready to correct, and share
  - Confirm that any Listeria samples on the run have passed QC check in BioNumerics
  - Email sent to request corrections
- GT isolates only:
  - Copy of run made for MDC binder
  - SAMN #'s confirmed from RedMine for QC check
  - Data shared with submitter

### ISSUES TO CORRECT PRIOR TO SHARING

(date and initial when complete)

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_

# DCLS / GT PARTNER ENGAGEMENT



An open conversation is held with the partner to discuss the

- Partnership and associated metadata requirements
- Special requirements and turn around time needs
- Shipment requirements and specifics



Documents are shared

- MOU and TOA
- DCLS Standardized Sample Submission Guidance document
- FDA Metadata spreadsheet



An internal meeting is held to discuss the requirements and orchestrate a plan to allow us to meet the partner's terms.



# DCLS WGS GT WORKFLOW OVERVIEW

*DCLS' policy is to confirm all isolates to at least the genus level before proceeding with WGS.*

## Identification and Extraction Process - EXTERNAL SUBMITTERS

- Isolate received
- Accessioned
- Plated



DNA Isolation and quantification for WGS

GenomeTrakr Team

Micro Lab

WGS Extraction Team

---

## Sequencing and Data Analysis Process

- DNA Library Preparation
- Library Normalization
- Sequencing Run

Data packet review and data sharing

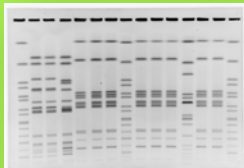
WGS Sequencing Team

# DCLS WGS GT WORKFLOW OVERVIEW

*DCLS performs sequencing for isolates that are representative of pathogens in food samples by PFGE*

## Identification and Extraction Process - INTERNAL SUBMITTER

Isolate identified from food or environment



PFGE

DNA Isolation and quantification for WGS

Micro Lab/Food Micro

MSL

WGS Extraction Team

---

## Sequencing and Data Analysis Process

- DNA Library Preparation
- Library Normalization
- Sequencing Run

Data packet review and data sharing

WGS Sequencing Team

# GT DATA ANALYSIS AND SHARING



## Data review

- Data packets are prepared by the Scientist
- Data packets are reviewed by a Senior Scientist (that did not perform the run)
- QC metrics are assessed



## Data sharing

- Data is shared with the FDA by the packet reviewer
- Data is shared with the partner based on pre-determined preference - NCBI identifiers, BaseSpace, GalaxyTrackr



## Follow-Up

- RedMine is checked weekly to determine which samples did not pass QC metrics at the FDA and require repeating
- Follow up conversation with partners if their samples are implicated

# NEXT STEPS



## LIMS integration

- Move from Excel to LIMS for data management in 2018



## QC improvements

- Validate bioinformatics tools to expand QC checks available and expedite the repeating of samples



## Process improvements

- V3 chemistry validation
- Tapestation analysis integration

# ACKNOWLEDGMENTS



FDA GenomeTrakr Grant - 5U18FD006188-02  
CDC ELC Grant - NU50CK000387-04-01



James Madison University  
North Carolina State University  
Pennsylvania State University  
Texas A & M University  
Virginia Tech



Dr. Denise Toney - *Director*  
Dr. Angela Fritzingler - *Deputy Director*  
Dr. Lauren Turner - *Lead Scientist*  
Arian Nassiri - *GenomeTrakr Senior Scientist*  
Patricia Croscutt  
Robert Barnes  
Shannon Levesque  
Kevin Libuit  
Gretchen Wilson





**THANK  
YOU**

**QUESTIONS?**





# DCLS WGS GT WORKFLOW JOB AID



## GenomeTrakr Sample Processing Job Aid

### Sample Requesting:

1. When requesting GenomeTrakr (GT) isolates, guidance can be found in the "Sample Submission Guidelines for WGS External Isolates" (Qualtrax ID # 13035).
2. Freezer isolates are the preferred sample submission type and the FDA Metadata Spreadsheet should be completed by the submitter (consult with Lead Scientist if you need assistance finding the spreadsheet).
3. The "DCLS Molecular Subtyping Submission Form" (Qualtrax ID # 8891) may also be completed by the submitter and included with the shipment if the sample batch being submitted is small (perhaps <5).

### Sample Receiving and Accessioning:

1. Once shipment arrives, GT Scientist will retrieve the package and open in a BSC to assess the package contents.
2. All samples in the package should be reconciled with the shipping manifest.
3. GT Scientist will accession the samples in LIMS. See LIMS accessioning details below.
4. GT Scientist will enter the pertinent sample information (LIMS ID, WGS ID, organism, source, received date, collection date, collected by, and submitter ID) in the GenomeTrakr Log (T:\Whole Genome Sequencing\WGS Log\GT logs).
5. GT Scientist will add the samples to the WGS Log (T:\Whole Genome Sequencing\WGS Log), making sure to use the "\_GT" designation following the sample ID. A note should be added that these samples are not yet ready for extraction and Enteric should be contacted for an expected pick up date.

### Sample Processing:

1. GT Scientist will deliver the contents of the package, and a printout of the samples from the log attached to the "DCLS Molecular Subtyping Submission Form" (Qualtrax ID # 8891), to the Enteric Group who will prepare cultures for confirmatory ID (ex. Maldi), WGS extraction and PFGE if they require pulsing.
2. If samples require PFGE analysis, MSL staff will send an email indicating such to the DCLS\_WGS-users email address and it will be noted in the WGS Log that PFGE is required. Consult with the Lead or Principal Scientist to determine if sequencing should proceed in parallel. In addition, MSL will send an email to the DCLS\_WGS-users email address when PFGE analysis is complete.
3. Enteric staff will notify MDC when the samples are ready to pick up for extraction and will submit the printout given to them in Sample Processing Step 1, with the samples.
4. MDC WGS extraction competent staff will prepare freezer stocks of all GT samples at the time of extraction and maintain them for long term storage at -80°C. WGS Freezer Stock Creation Form (Qualtrax ID # 22868) should be used to document this process.
5. Prioritization of samples in the sequencing queue will vary with submission type. Check with Lead Scientist, Principal Scientist or GT Senior for priority and TAT requirements.
6. Once GT isolates are selected for library preparation they must be accessioned with FDA CFSAN per the DCLS WGS SOP Qualtrax ID # 17639.
7. Following sequencing and data upload, the FDA's RedMine site should be checked to ensure that the QC has passed and the SAMN numbers should be recorded in the WGS Log.

Select Panel

STARS LIMS

Sample Type: None

- Adult Sickle Cell
- Arbovirus Human
- Bacteriology Reference Culture
- Blood Parasites
- Enteric Clinical