



Department of Health
Wadsworth Center

Threat Agnostic Sentinel Surveillance (TASS) Pilot in New York: Evaluating Metagenomic Sequencing

Michael Perry, MSED, DrPH
Director, Biodefense Laboratory

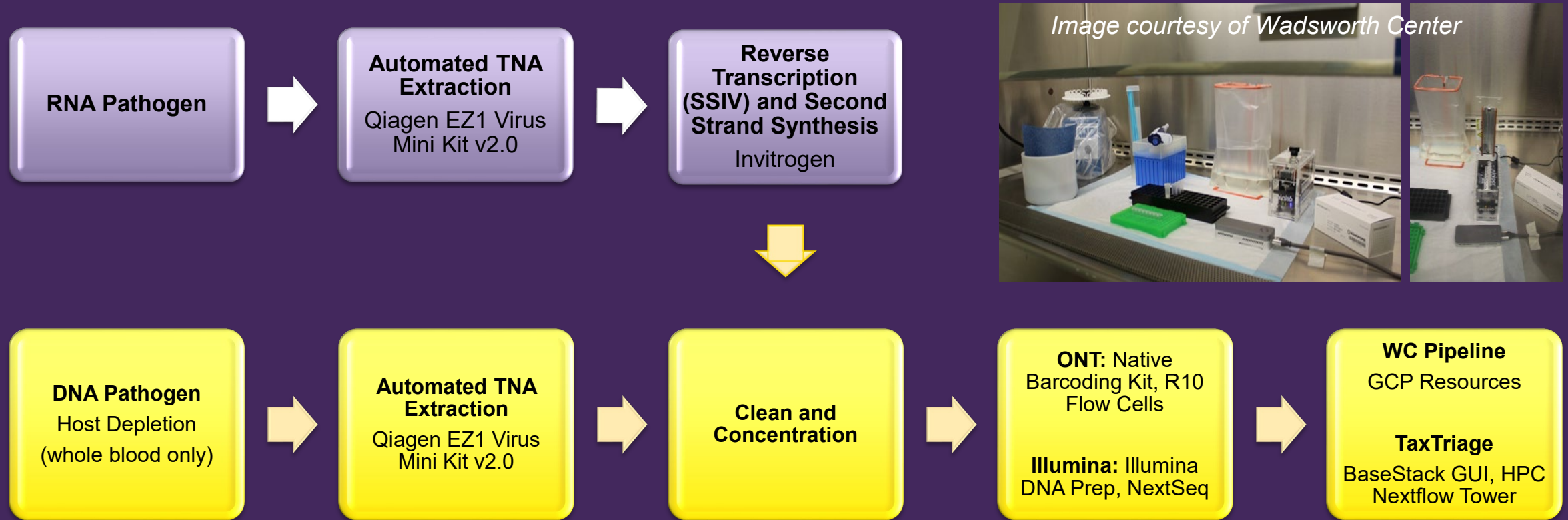
MARCH 26, 2025 | APHL ID LAB CON 2025 | PASADENA, CA

Overview

- **Threat Agnostic Sequencing Project Overview**
 - Matrices
 - Optimization
 - Pilot Challenge Panel
 - Analysis
- **Pilot Results**
- **Sequencing Moving Forward**



GENERAL WET LAB WORKFLOW



Department of Health
Wadsworth Center

RNA – Ribonucleic acid
TNA – Total nucleic acid
DNA – Deoxyribonucleic acid
ONT – Oxford Nanopore

WC – Wadsworth Center
GCP – Google Cloud Pipeline
GUI – Graphical user interface
SS – SuperScript
HPC – High-performance computing

WORKFLOW DEVELOPMENT AND OPTIMIZATION: EXTRACTION



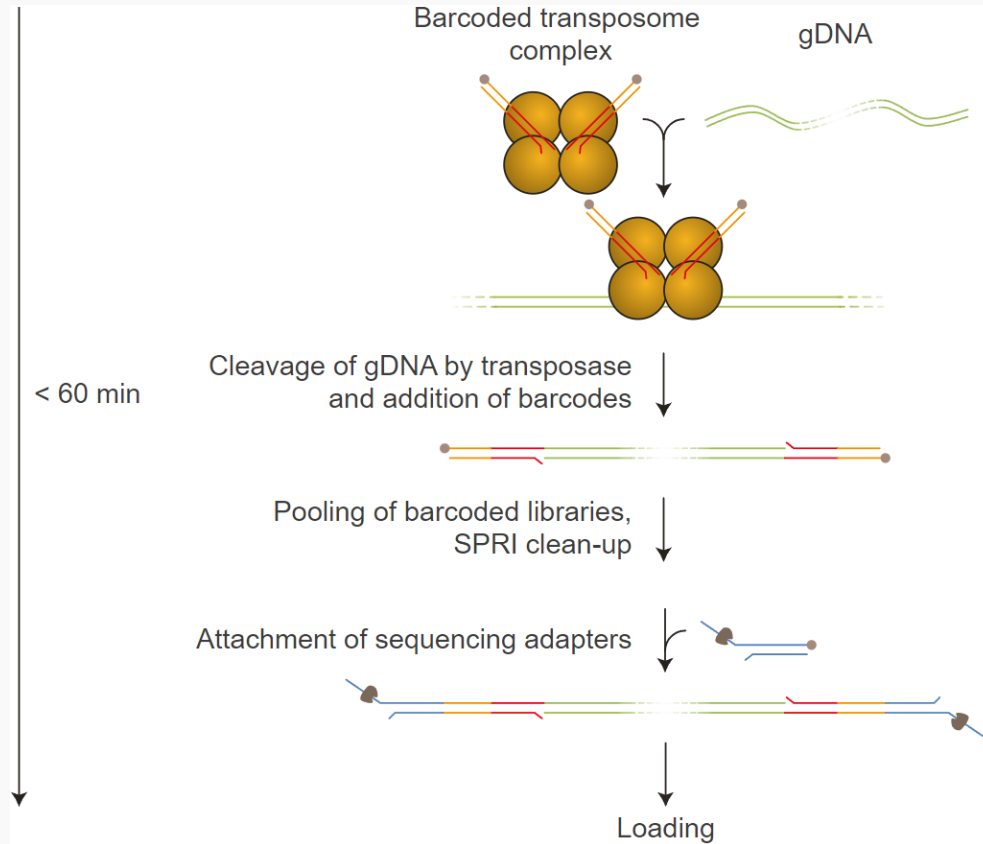
<https://www.qiagen.com/us/products/diagnostics-and-clinical-research/solutions-for-laboratory-developed-tests/ez1-advanced-xl-instrument>

- **Manual Extraction Protocols**
 - Qiagen DNA Blood Mini Kit
 - Epicentre MasterPure Complete
- **Automated Extraction Protocols**
 - Qiagen EZ1 Virus Kit
 - Qiagen EZ1 RNA Tissue Kit
- **Host Depletion**
 - Filtration via Centrifugal Devices
 - Molzym
 - Enzymatic and Mechanical



Department of Health
Wadsworth Center

WORKFLOW DEVELOPMENT AND OPTIMIZATION: POST-EXTRACTION



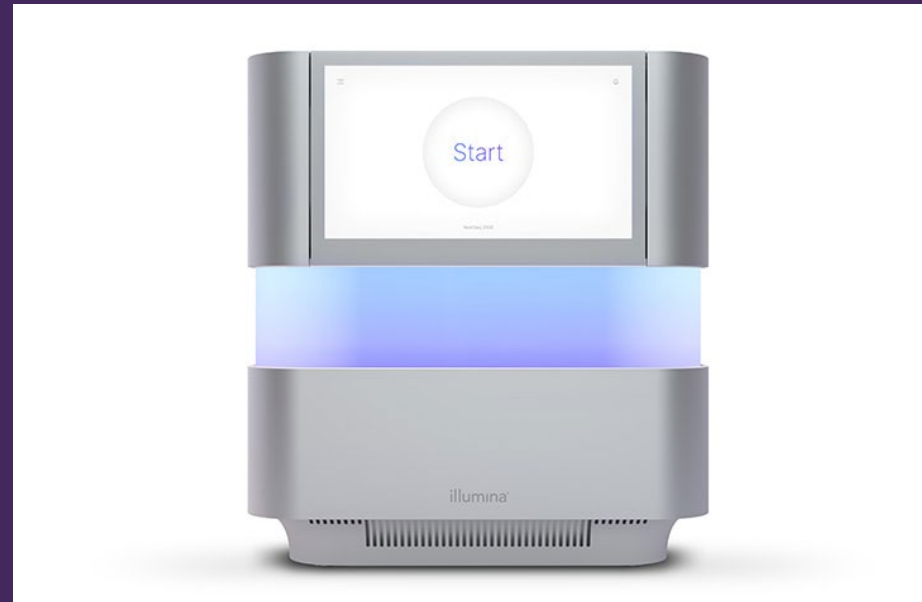
SPRI – Solid-Phase Reversible Immobilization



Department of Health
Wadsworth Center

- **Clean and Concentration**
 - AMPure XP
 - Zymo DNA
 - Zymo RNA
- **Library Preparation**
 - Illumina
 - DNA Prep
 - ONT
 - Ligation Sequencing Kit
 - Native Barcoding Kit
 - Rapid Barcoding Kit
 - Rapid PCR Barcoding Kit
- **Bioinformatic Host Depletion**
 - ONT Adaptive Sequencing

THREAT AGNOSTIC SEQUENCING



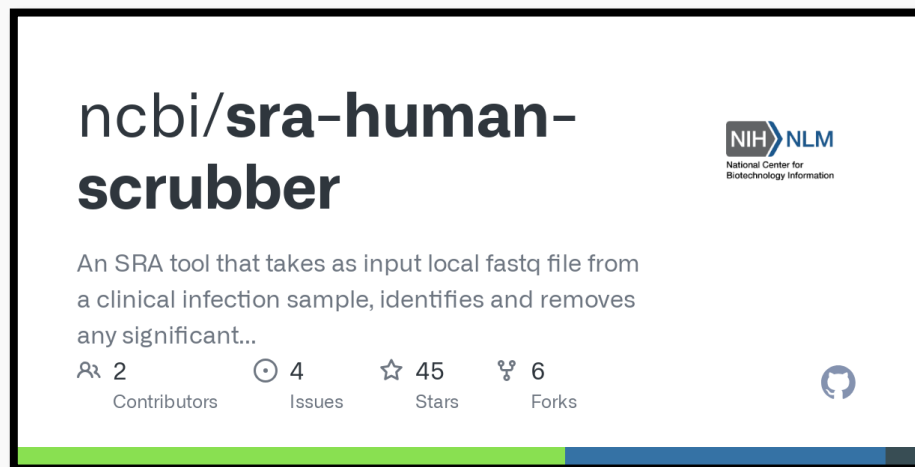
Department of Health
Wadsworth Center

<https://sapac.illumina.com/systems/sequencing-platforms/miseq.html>

<https://support.illumina.com/>

<https://www.labmedica.com/molecular-diagnostics/articles/294793737/pocket-sized-devices-enable-real-time-dna-and-rna-sequencing-anywhere.html>

BIOINFORMATIC ANALYSIS



Department of Health
Wadsworth Center

<https://github.com/ncbi/sra-human-scrubber>
Amazon Web Services (AWS)

BIOINFORMATIC ANALYSIS

seqera

LaunchpadHelp and Support

JHUAPL_Bio | NY_Private

LaunchpadRunsActionsDatasetsData ExplorerData StudiosCompute EnvironmentsCredentialsSecretsParticipantsSettings

Launchpad

Launch a run by selecting one of the available pipelines below. Pipelines contain a configured workflow repository, compute environment, and launch parameters. Skip this and [launch a run without configuration](#) →

Showing 1-4 of 4

Sort by: Recently updated

Search...

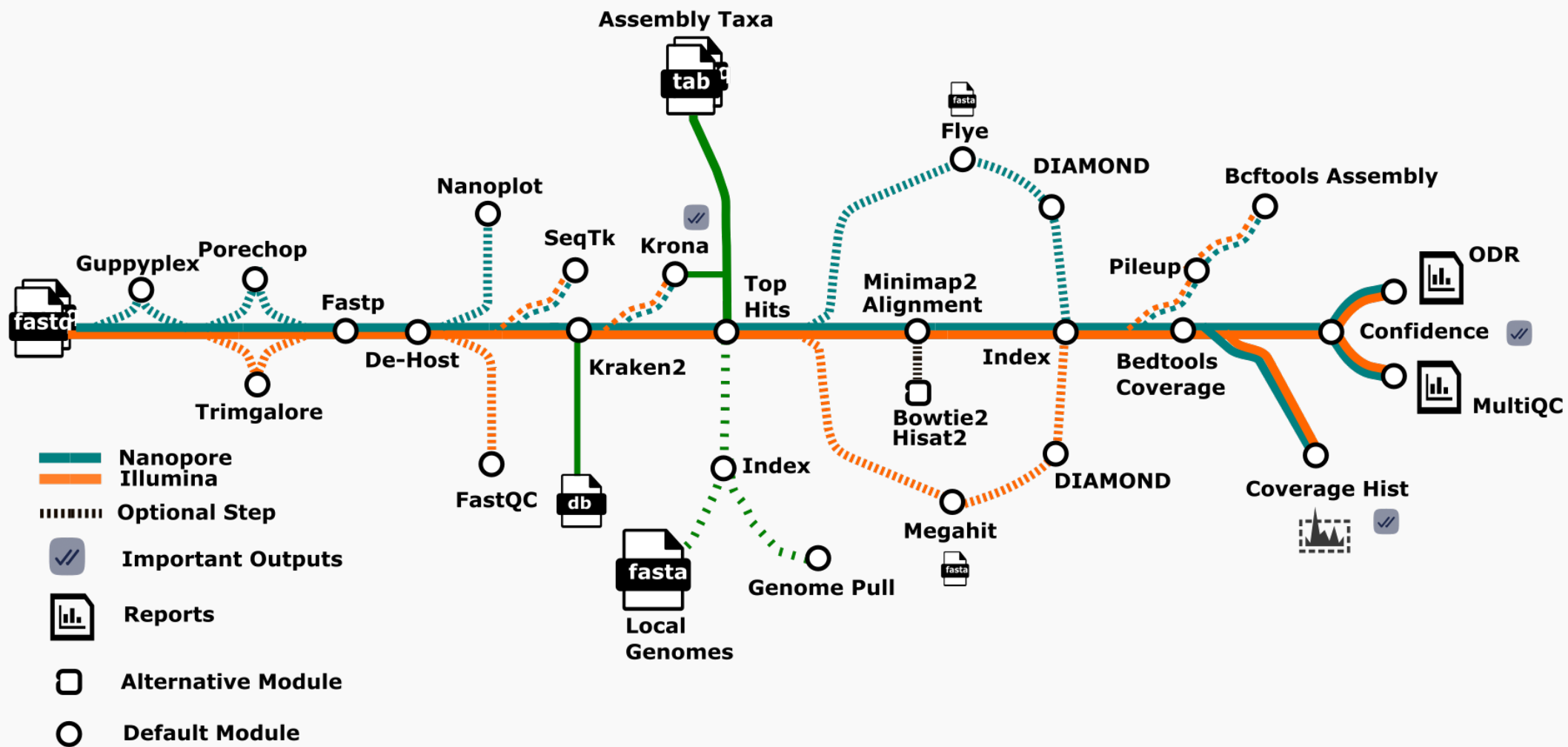
Add pipeline

 TaxTriage https://github.com/jhuapl-bio/taxtr... shared	 Updated Apr 16, 2024, 12:01 PM	 Launch ⋮
 TaxTriage_NY_MainBranch https://github.com/jhuapl-bio/taxtr...	 Updated Mar 6, 2024, 1:07 PM	 Launch ⋮
 TaxTriage_MCE_NY https://github.com/jhuapl-bio/taxtr...	 Updated Mar 6, 2024, 1:04 PM	 Launch ⋮
 TaxTriage_MCE https://github.com/jhuapl-bio/taxtr... shared	 Updated Mar 5, 2024, 9:42 AM	 Launch ⋮

< Prev 1 Next >



BIOINFORMATIC ANALYSIS PIPELINES



Department of Health
Wadsworth Center

<https://github.com/jhuapl-bio/taxtriage>

TAXTRIAGE REPORT



A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.

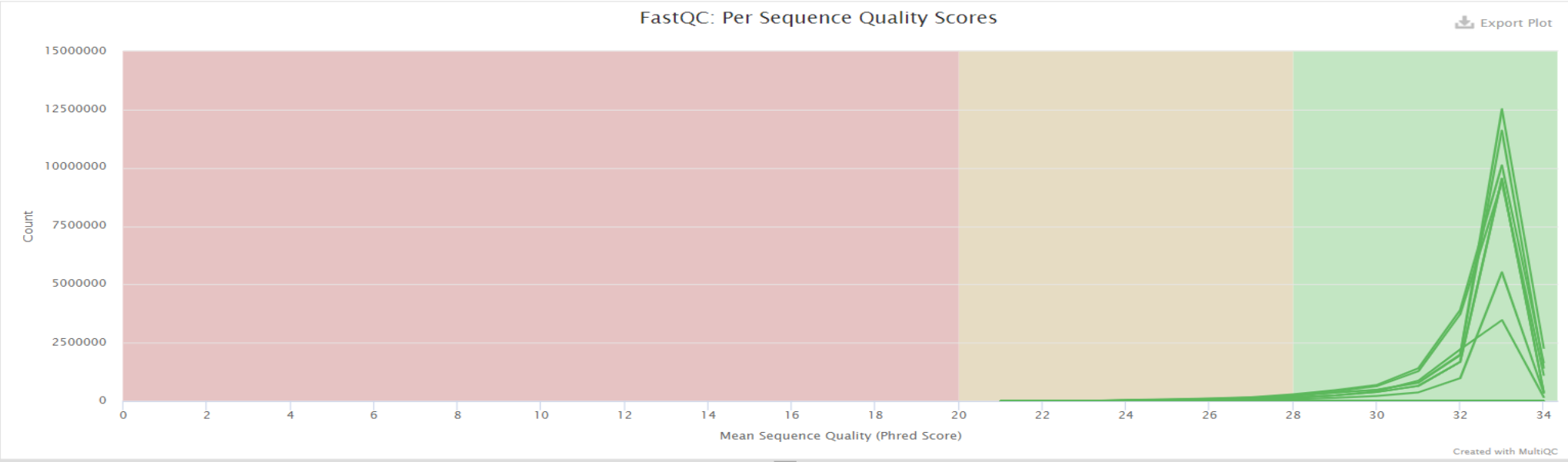
This report has been generated by the [nf-core/taxtriage](#) analysis pipeline. For information about how to interpret these results, please see the [documentation](#).

Per Sequence Quality Scores 14

The number of reads with average quality scores. Shows if a subset of reads has poor quality.

[Help](#)

Y-Limits: ☒ on



Department of Health
Wadsworth Center

TAXTRIAGE REPORT

Kraken

Kraken is a taxonomic classification tool that uses exact k-mer matches to find the lowest common ancestor (LCA) of a given sequence. DOI: 10.1186/gb-2014-15-3-r46.

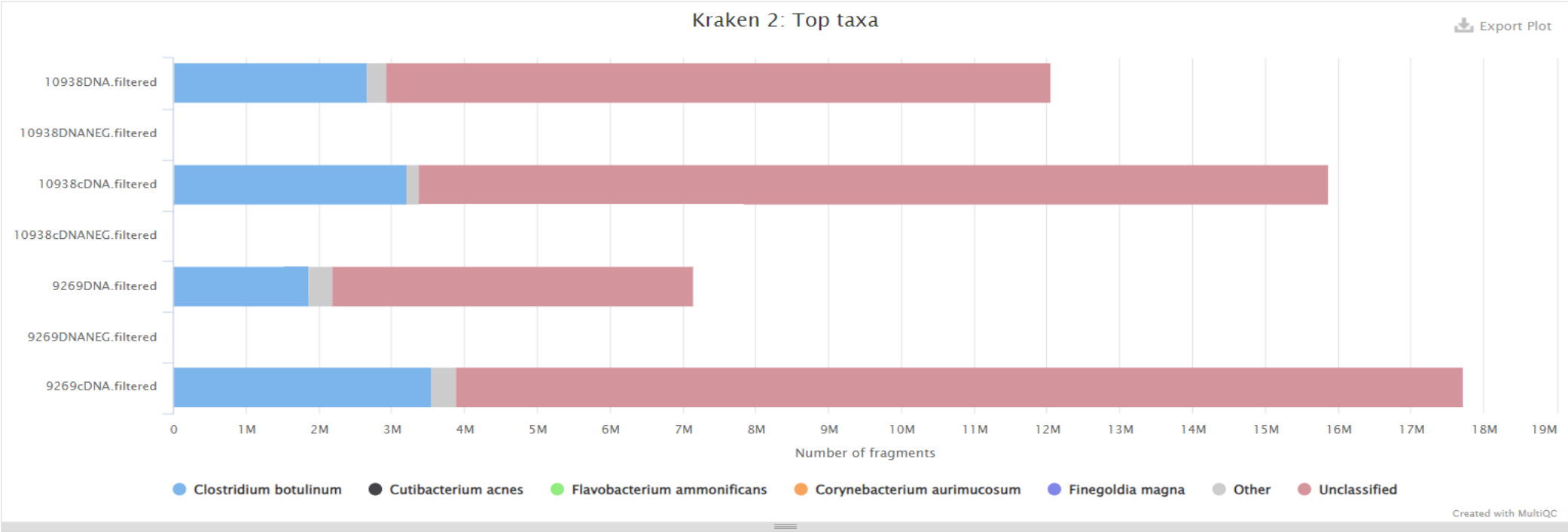
Top taxa

The number of reads falling into the top 5 taxa across different ranks.

CountsPercentages

SpeciesGenusFamilyOrderClassPhylumKingdomDomainRoot

Help



TAXTRIAGE PATHOGEN REPORT

Sample	Sample Type	Organism	% Reads in Sample	# Aligned to Sample
10938DNA	Unknown	Clostridium novyi B str. NCTC 9691	98.87	652567
10938DNA	Unknown	Clostridium haemolyticum NCTC 9693	0.58	3838
10938DNA	Unknown	Clostridium novyi B str. ATCC 27606	0.07	490
10938DNA	Unknown	Clostridium botulinum C/D str. It1	0.05	297
10938DNA	Unknown	Clostridium sp. K25	0.04	273
10938DNA	Unknown	Clostridium tetani E88	0.04	254
10938DNA	Unknown	Clostridium botulinum D str. 16868	0.03	208
10938DNA	Unknown	Clostridium novyi A str. BKT29909	0.03	196
10938DNA	Unknown	Clostridium botulinum BKT015925	0.03	189



Department of Health
Wadsworth Center

PILOT STUDY

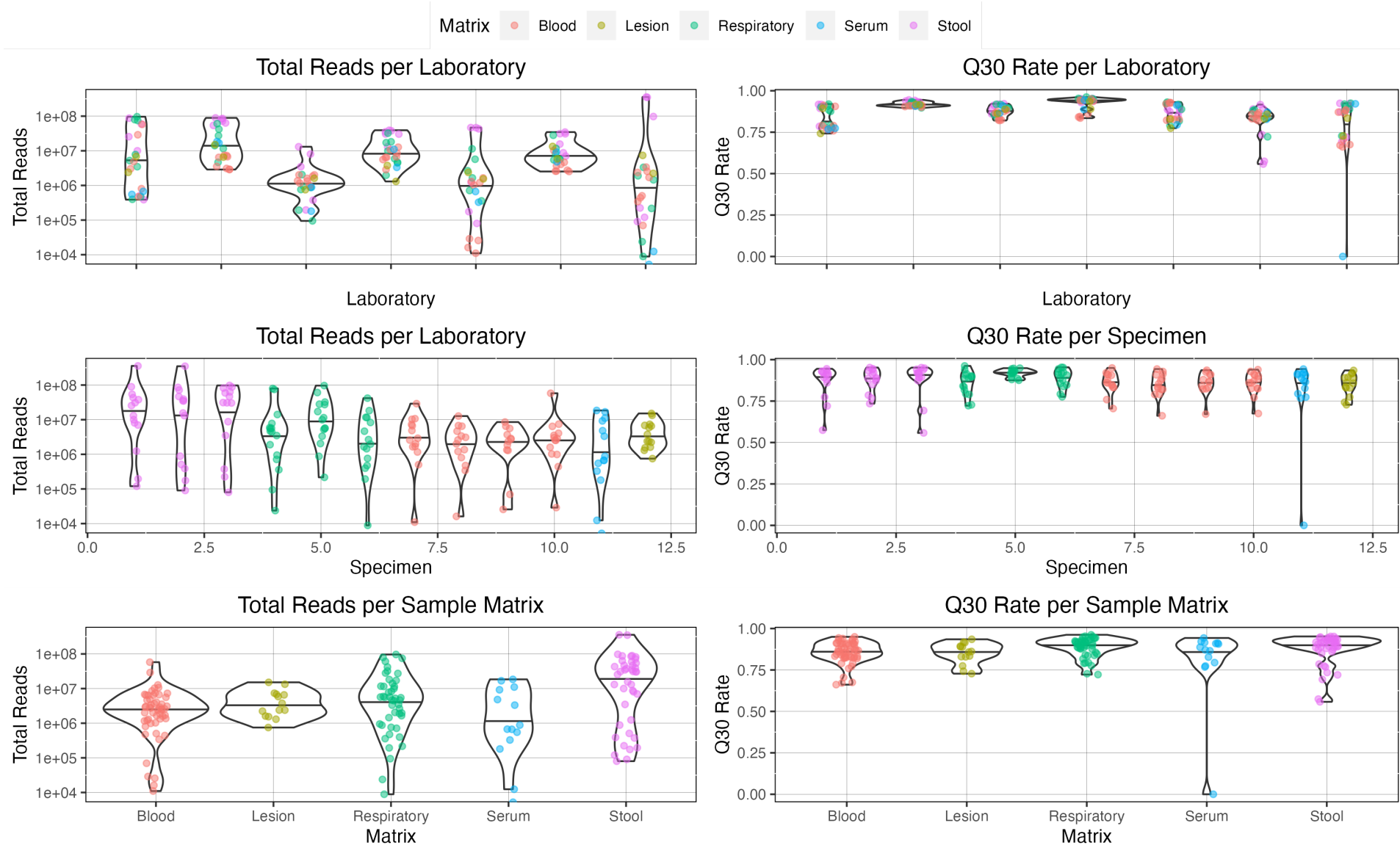


Department of Health
Wadsworth Center

Created with Adobe Firefly

Sequencing Performance (I)

(Slide courtesy of Johns Hopkins Applied Physics Laboratory)



EXPECTED PATHOGEN DETECTION

- Known Organisms in MCE Panel Samples

Specimen	Organism	Preparation
1	<i>Clostridium difficile</i>	DNA, RNA
2	<i>Clostridium difficile</i>	DNA, RNA
	<i>Escherichia coli</i> O157:H7	DNA, RNA
4	Severe acute respiratory syndrome coronavirus 2	RNA
5	Severe acute respiratory syndrome coronavirus 2	RNA
	Human respiratory syncytial virus A	RNA
6	Human respiratory syncytial virus B1	DNA, RNA
7	<i>Bacillus cereus</i>	DNA, RNA
8	<i>Burkholderia thailandensis</i>	DNA, RNA
10	<i>Burkholderia thailandensis</i>	DNA, RNA
11	Zika virus	RNA
12	Orthopoxvirus - Mpox virus	DNA, RNA



EXPECTED PATHOGEN DETECTION

- TaxTriage Confidence Metric: v1 Values

Preparation	Specimen	Organism	NY – v1
DNA	1	<i>Clostridium difficile</i>	0.52
DNA	2	<i>Clostridium difficile</i>	0.50
DNA		<i>Escherichia coli</i> O157:H7	0.29
DNA	7	<i>Bacillus cereus</i>	Not Detected
DNA	8	<i>Burkholderia thailandensis</i>	Not Detected
DNA	10	<i>Burkholderia thailandensis</i>	0.05
DNA	12	Orthopoxvirus - Mpox virus	0.83



EXPECTED PATHOGEN DETECTION

- TaxTriage Confidence Metric: v2 Values

Preparation	Specimen	Organism	NY – v2
DNA	1	<i>Clostridium difficile</i>	0.60
DNA	2	<i>Clostridium difficile</i>	0.60
DNA	2	<i>Escherichia coli</i> O157:H7	0.47
DNA	7	<i>Bacillus cereus</i>	Not Detected
DNA	8	<i>Burkholderia thailandensis</i>	Not Detected
DNA	10	<i>Burkholderia thailandensis</i>	0.14
DNA	12	Orthopoxvirus - Mpox virus	1.00

NY – v1
0.52
0.50
0.29
Not Detected
Not Detected
0.05
0.83



EXPECTED PATHOGEN DETECTION

- **TaxTriage Confidence Metric: v1 Values**

Preparation	Specimen	Organism	NY – v1
RNA	1	<i>Clostridium difficile</i>	0.07
RNA	2	<i>Clostridium difficile</i>	0.05
RNA		<i>Escherichia coli</i> O157:H7	0.06
RNA	4	Severe acute respiratory syndrome coronavirus 2	0.24
RNA	5	Severe acute respiratory syndrome coronavirus 2	0.82
RNA		Human respiratory syncytial virus A	Not Detected
RNA	6	Human respiratory syncytial virus B1	0.83
RNA	7	<i>Bacillus cereus</i>	0.08
RNA	8	<i>Burkholderia thailandensis</i>	0.17
RNA	10	<i>Burkholderia thailandensis</i>	0.17
RNA	11	Zika virus	Not Detected
RNA	12	Orthopoxvirus - Mpox virus	0.83



Department of Health
Wadsworth Center

EXPECTED PATHOGEN DETECTION

- TaxTriage Confidence Metric: v2 Values

Preparation	Specimen	Organism	NY – v2
RNA	1	<i>Clostridium difficile</i>	0.23
RNA	2	<i>Clostridium difficile</i>	Not Applicable
RNA	2	<i>Escherichia coli</i> O157:H7	0.12
RNA	4	Severe acute respiratory syndrome coronavirus 2	0.85
RNA	5	Severe acute respiratory syndrome coronavirus 2	0.98
RNA	5	Human respiratory syncytial virus A	Not Detected
RNA	6	Human respiratory syncytial virus B1	1.00
RNA	7	<i>Bacillus cereus</i>	0.16
RNA	8	<i>Burkholderia thailandensis</i>	0.85
RNA	10	<i>Burkholderia thailandensis</i>	0.85
RNA	11	Zika virus	Not Detected
RNA	12	Orthopoxvirus - Mpox virus	1.00

NY – v1
0.07
0.05
0.06
0.24
0.82
Not Detected
0.83
0.08
0.17
0.17
Not Detected
0.83



Department of Health
Wadsworth Center

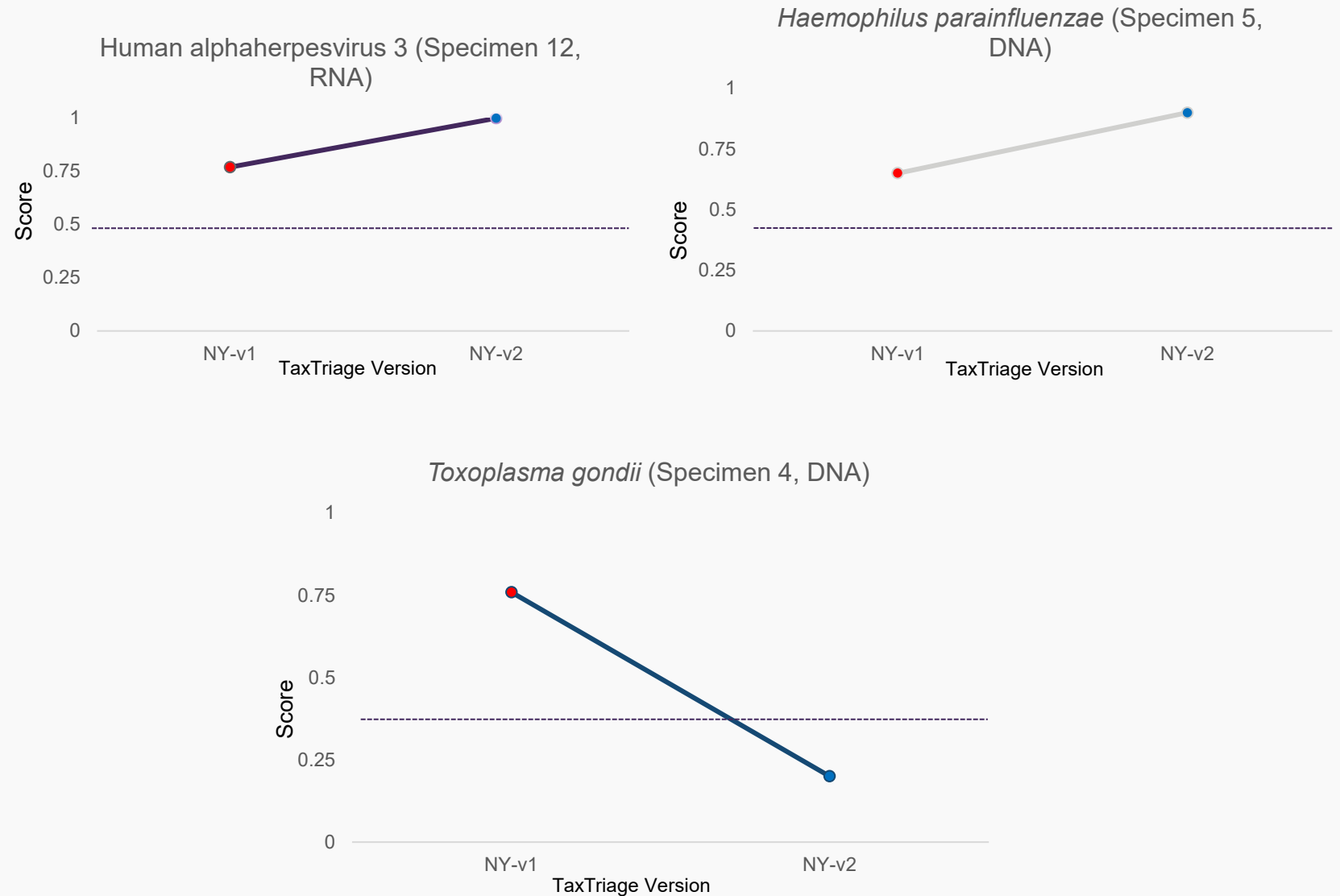
“MISSED” PATHOGENS DEEP DIVE

- Lab-X sample 4 (RNA prep) contained SARS-CoV-2 that was not detected
- Histogram reveals four reads, all mapping to a single location and only covering 46 nucleotides
 - Organism is being screened out of pathogen report due to three non-overlapping reads requirement.
- SARS-CoV-2 is detected, if reads processed with Chan Zuckerberg ID (CZID)
 - However, CZID also detects 91 other potential pathogens in the sample.
 - TaxTriage reports 25 organisms in the complete pathogens report, with none having a score above 0.4.



CO-INFECTION DETECTION

- V2 increases scores of high-likelihood pathogens
- V2 reduces scores of low likelihood pathogens
- Horizontal line depicts cut-off score for pathogen report (0.4)



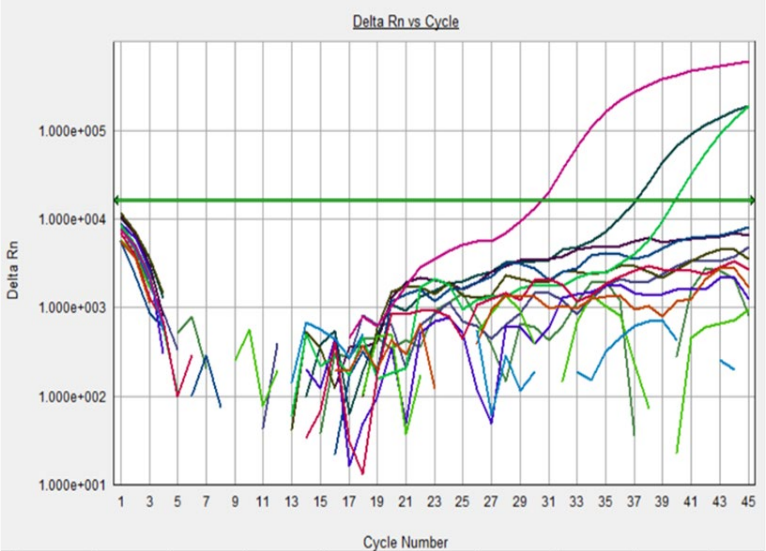
SPECIMEN TESTING



Department of Health
Wadsworth Center

Created with Adobe Firefly

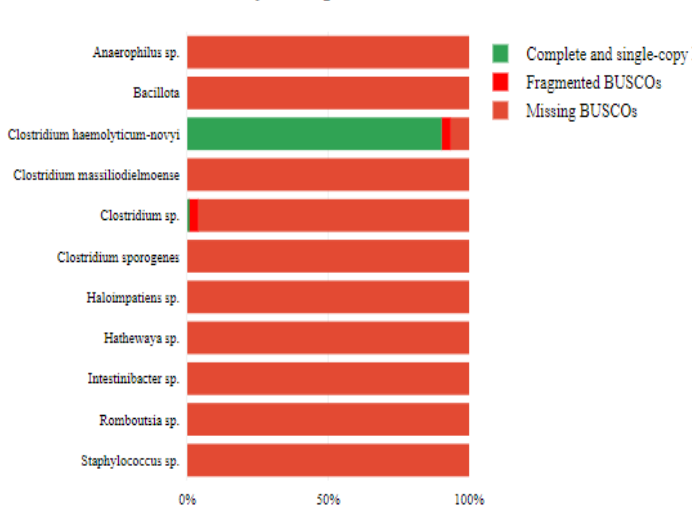
A TALE OF TWO COWS...



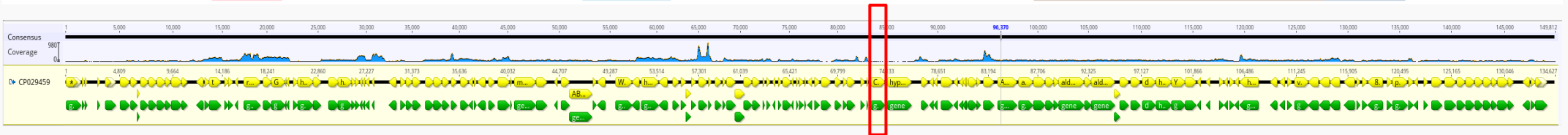
Clostridium haemolyticum-novyi

Virulence Factor	Detection	Confidence
bont_A	Not Detected	High
bont_B	Not Detected	High
bont_C	Not Detected	High
bont_D	Not Detected	High
bont_E	Not Detected	High
bont_F	Not Detected	High
bont_FA(H)	Not Detected	High
bont_G	Not Detected	High
bont_X	Not Detected	High

Assembly Completeness of Bacterial Bins



Cluster ID	Plasmid Subtype	Circular	Detected Resistance Markers	Length (bp)	Number of Contigs	Predicted Mobility	Nearest NCBI Accession	Mash Distance to Nearest NCBI Accession	Coverage	Predicted Host Range
AE802	AO659	No		92 445bp	24	Non-mobilizable	CP029459	0.023	40.2	Clostridium
AG731	-	No		41 095bp	15	Non-mobilizable	CP029460	0.054	42.8	Clostridium



Department of Health
Wadsworth Center

A TALE OF TWO COWS...

Sample	Sample Type	Organism	Class	% Reads in Sample	# Aligned to Sample	Alignment Conf	Locations
10938DNANEG	Unknown	Micrococcus luteus	Pathogen	5.51	8	0.92	Unknown
10938DNANEG	Unknown	Acinetobacter junii	Pathogen	16.55	24	0.91	blood, urine
10938DNANEG	Unknown	Pseudomonas aeruginosa	Pathogen	5.51	8	0.86	abscess, blood, urine
9269cDNA	Unknown	Clostridium septicum	Pathogen	0.039	844	0.82	gut
10938cDNA	Unknown	Clostridium septicum	Pathogen	0.088	1313	0.81	gut
9269cDNA	Unknown	Clostridium tetani	Pathogen	0.00056	12	0.81	skin, blood
9269DNANEG	Unknown	Acinetobacter radioresistens	Pathogen	1.68	6	0.79	blood, resp
9269DNA	Unknown	Clostridium septicum	Pathogen	0.0073	19	0.79	gut
10938cDNA	Unknown	Bacillus licheniformis	Pathogen	0.00073	12	0.79	blood
9269DNANEG	Unknown	Chryseobacterium indologenes	Pathogen	79.55	284	0.74	blood
10938DNANEG	Unknown	Kocuria rhizophila	Pathogen	11.034	16	0.74	Unknown
9269cDNA	Unknown	Bacillus licheniformis	Pathogen	0.00046	10	0.69	blood
10938cDNA	Unknown	Clostridium novyi	Pathogen	0.029	477	0.64	blood
10938cDNA	Unknown	Clostridium botulinum	Pathogen	0.0025	42	0.63	blood, stool
10938DNA	Unknown	Clostridium novyi	Pathogen	0.34	2352	0.62	blood
9269cDNA	Unknown	Clostridium novyi	Pathogen	0.048	1045	0.55	blood



Department of Health
Wadsworth Center

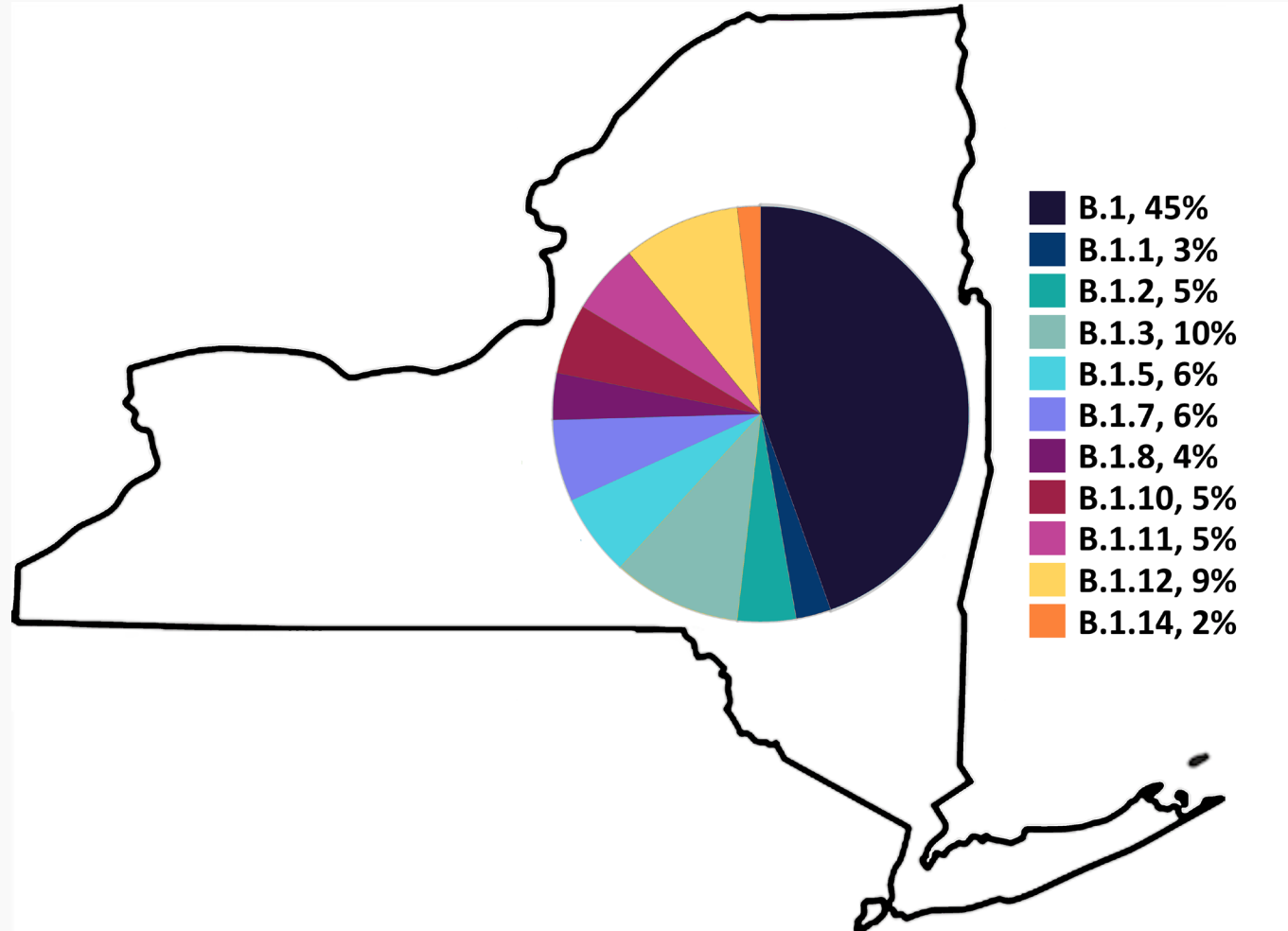
A TALE OF TWO COWS...

Specimen ID (Type)	Detected Organism	# Reads Aligned	Confidence Metric (0-1)	Taxonomic ID #	Pathogenic Subsp/Strains	Coverage	HHS Percentile	K2 Reads
Cow_09269_cDNA (unknown)	Clostridium novyi	191392 (14.47% - 0.48%)	0.99	1542		0.88	100.0	72951
Cow_09269_DNA (unknown)	Clostridium novyi	434846 (38.15% - 2.77%)	0.98	1542		0.77	100.0	278157
Cow_10938_cDNA (unknown)	Clostridium novyi	66035 (5.79% - 0.18%)	0.88	1542		0.8	100.0	25773
Cow_10938_DNA (unknown)	Clostridium novyi	55894 (4.76% - 0.21%)	0.87	1542		0.79	100.0	31755
Cow_10938_cDNA (unknown)	Rhodococcus globerulus*	1413 (0.12% - 0.00%)	0.52	1504		0.04	100.0	1576
Cow_09269_cDNA (unknown)	Clostridium septicum	909 (0.07% - 0.00%)	0.48	1504		0.03	100.0	1086



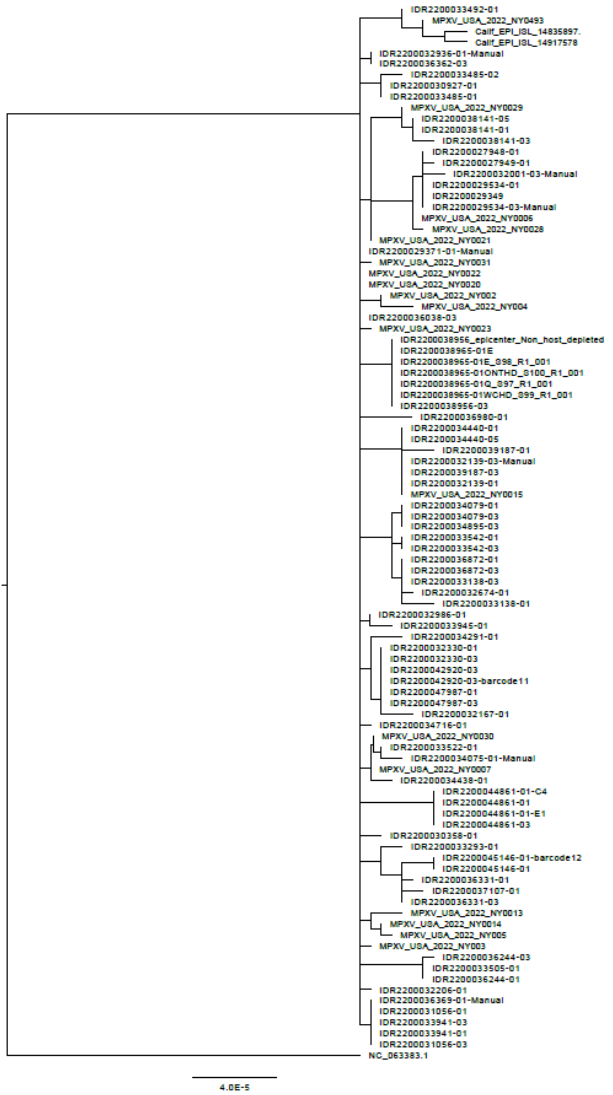
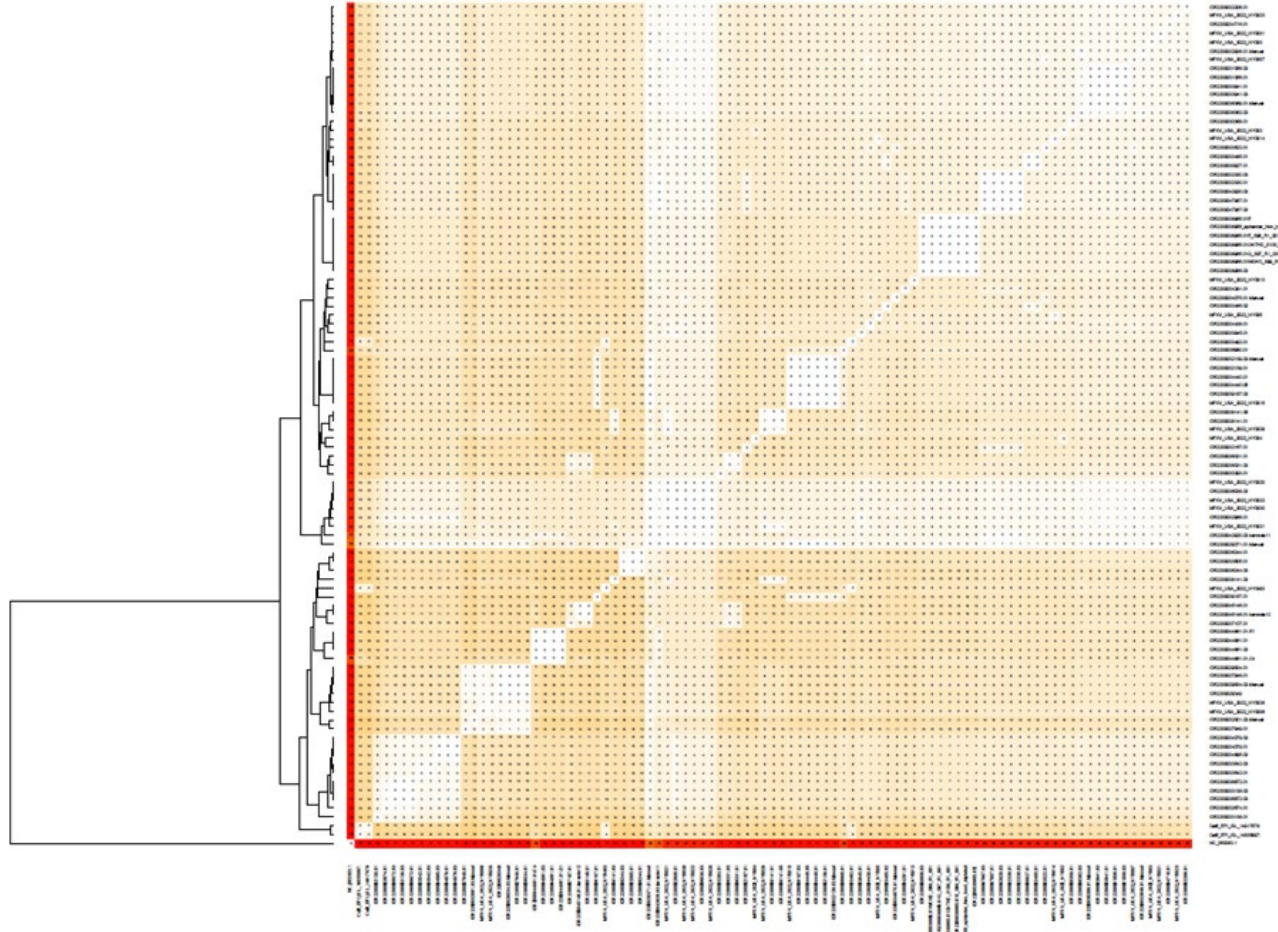
TASS WORKFLOW: 2022 MPOX OUTBREAK

New York State mpox NextClade Lineage Designation



Department of Health
Wadsworth Center

DATA CAN BE USED FOR MORE THAN DETECTION!



Department of Health
Wadsworth Center

SUMMARY

- Improvement in report generation using TaxTriage v2:
 - Decreases chances of false positives
 - Decrease organisms identified on reports
 - Report cutoff threshold of 0.4
- Increase efficiency
 - For multicenter evaluation,
 - 24 samples (12 DNA and 12 RNA)
 - Reduced run time from 22 to 2 hours
- Utility
 - Early warning
 - Reflex testing



Acknowledgements

Biodefense Laboratory

- Dominick Centurioni
- Alexander Diaz
- Hannah Cooper
- Melissa D'Amico
- Zahra LaTerra
- Shannon Maurer

Wadsworth Center AGTC

Wadsworth Center Bioinformatics Core

- Pascal LaPierre
- Erica Lasek-Nesselquist

Johns Hopkins University

- Brian Merritt
- Peter Thielen

Association of Public Health Laboratories

- Tyler Wolford

US Centers for Disease Control and Prevention

- Matt Mauldin
- Jarad Schiffer



Department of Health
Wadsworth Center



Department of Health Wadsworth Center