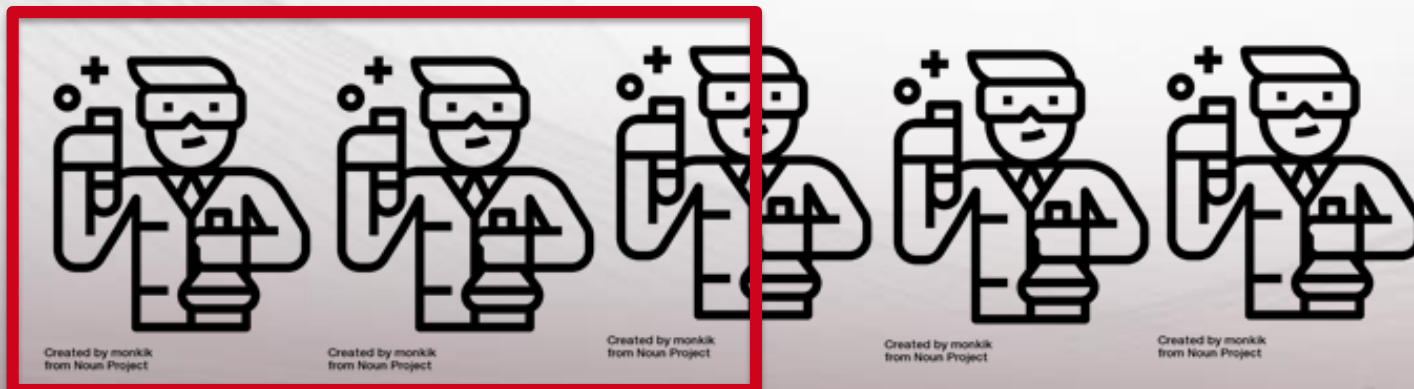
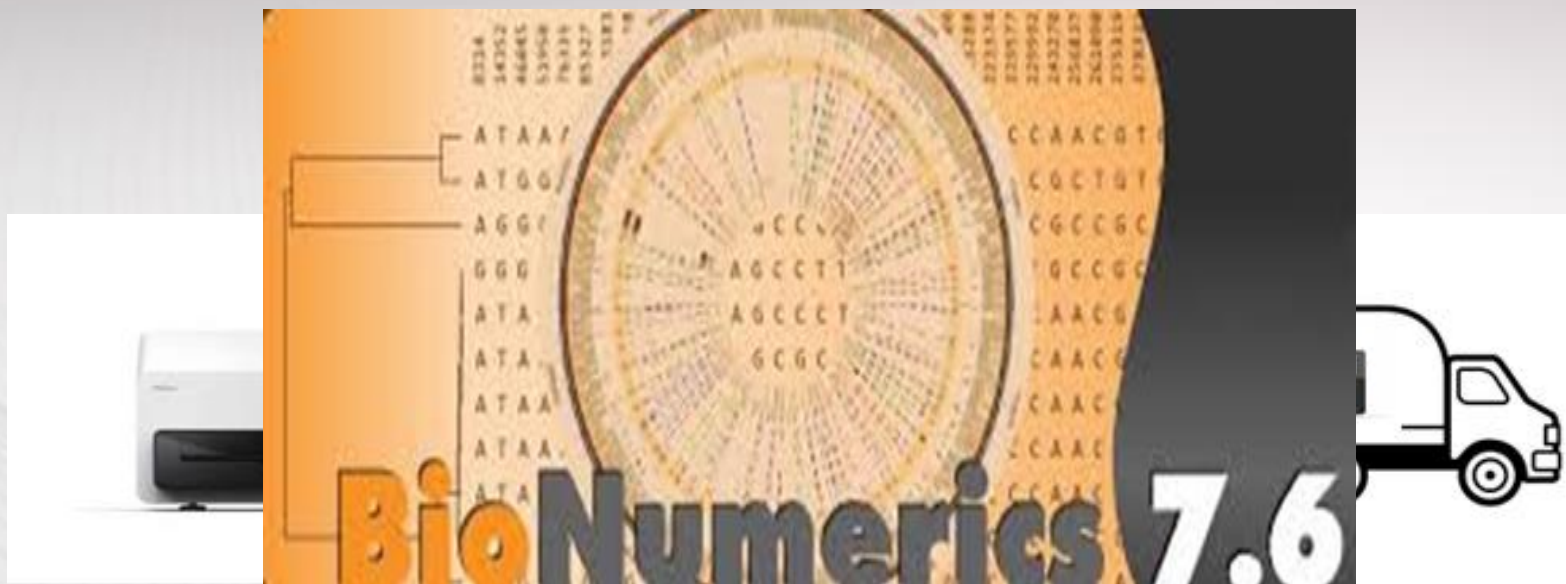




ODH MiSeq Instrument Validation (Limited Measures)

**MidWest Advanced Molecular
Detection (AMD) Training and
Symposium**

ODH Lab



Testing

- PulseNet Organisms
- Non PulseNet organism sequenced:
 - *Mycobacterium tuberculosis*
 - *Clostridium botulinum*
 - *Legionella pneumophila*
 - *Stenotrophomonas maltophilia*

Validation Timeline



2014

2015

2016

2017

APPROVED



Evaluated Metrics

Metric	Pass Criteria
Q30 Score	>30, assessed separately for R1 and R2
Average Coverage	<i>Listeria</i> & <i>Campy</i> 20x, <i>Salmonella</i> 30x, <i>E. coli</i> & <i>Shigella</i> 40x
Average Insert Size	>300 bp
Number of hq SNP differences	< 1 / MB
Number of allele differences	<i>Listeria</i> ≤ 3 , not validated yet for others

Accuracy and Precision



High precision

High accuracy



High precision

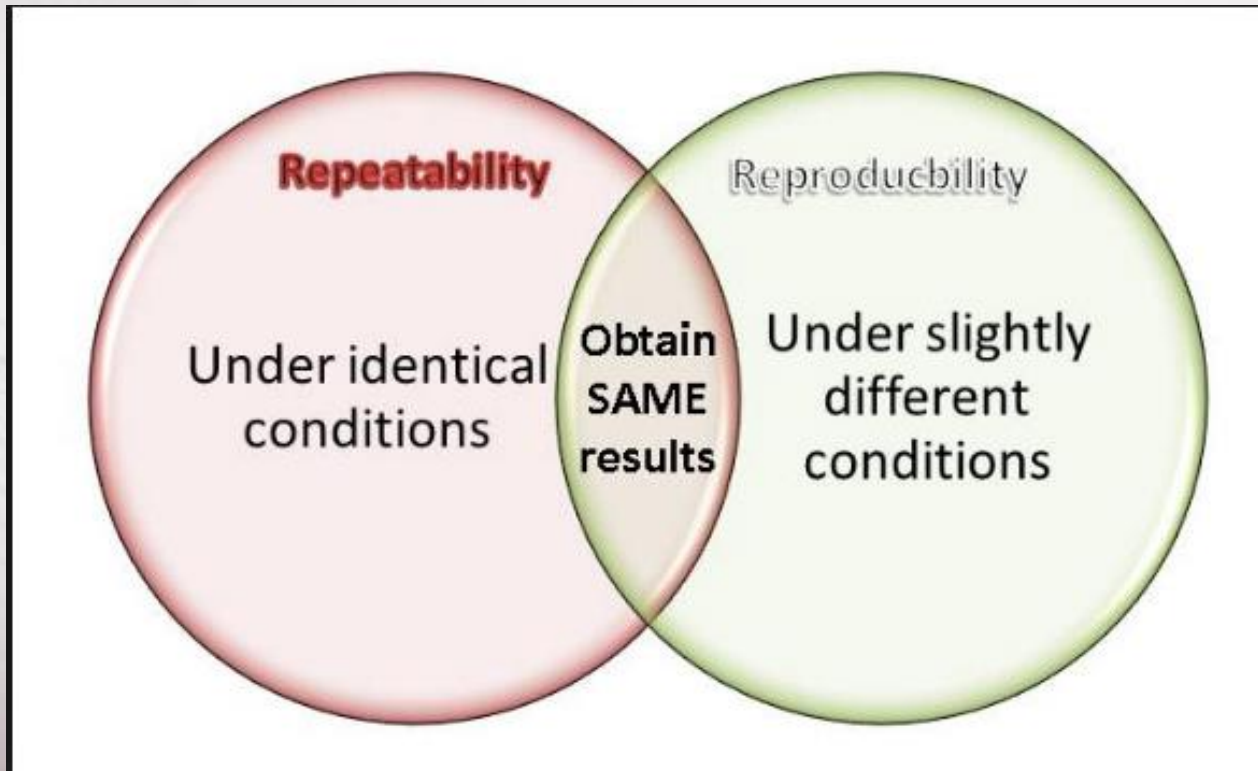
Low accuracy



Low precision

Low accuracy

Repeatability and Reproducibility



Under Review by CDC

Analytical **sensitivity**

- Detection of sequence variation(s) when present within the analyzed genomic region
 - True positive
- For this validation it is the same as accuracy determination.



Under Review by CDC

Analytical **specificity**

- Assay will not detect sequence variation(s) when none are present within the analyzed genomic region.
- True negative



Measures Not Determined

- Reportable Range
 - Regions of gene/genome that are included in the analysis.
- Reference Interval
 - Target regions that the assay can detect and are expected to occur in a reference population.

1st MiSeq Validation

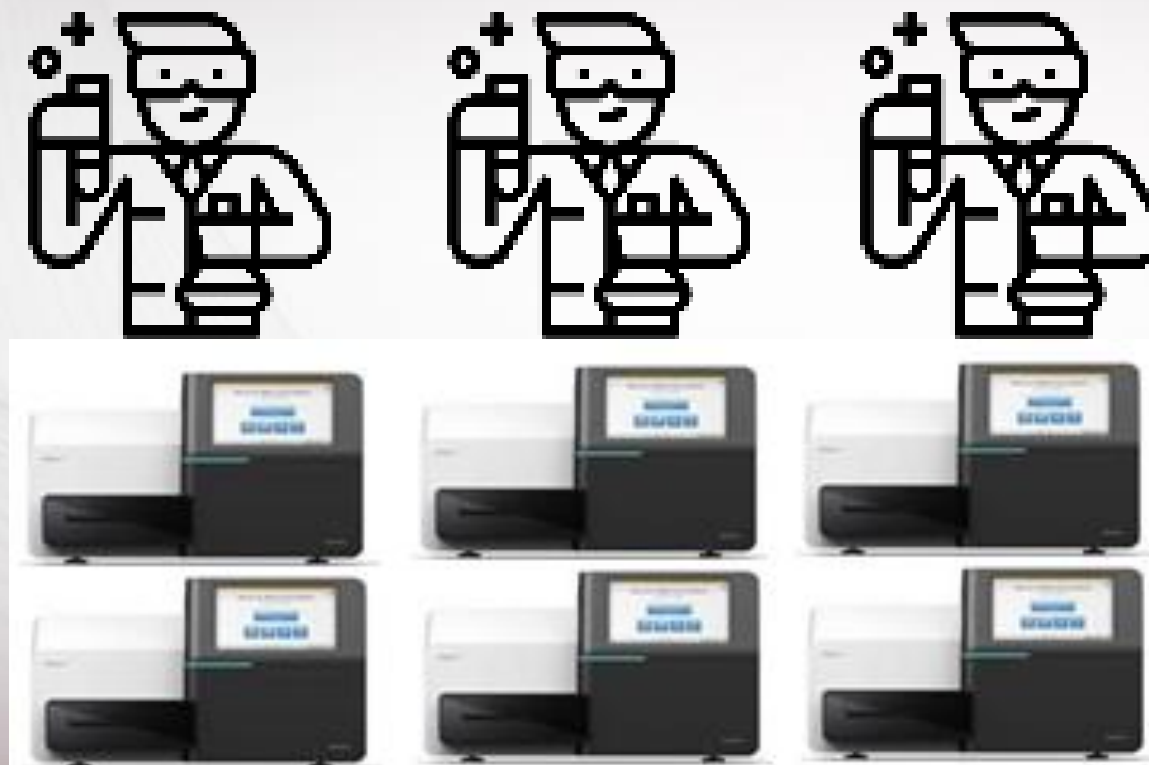
37-PulseNet Organisms (singlet/validation)

- 11 *Salmonella enterica*
- 9 *Listeria monocytogenes*
- 12 Shiga-toxin producing *Escherichia coli* (STEC)
- 5 *Campylobacter*

4-PulseNet Certification strains (singlet/analyst)

- *Salmonella enterica* serovar Heidelberg (CDC# C6472)
- *Listeria monocytogenes* (CDC# H8394)
- *Escherichia coli* (STEC)(CDC# 03-98)
- *Campylobacter jejuni* (CDC# 5480)

1st MiSeq Validation



1st MiSeq Validation Results

PNUSAL000701-2		Pass									
		Quality (Q33)	Coverage								
Reads R1 Quality	PNUSAL000701-2_S3_L001		37.14	30.55							
Reads R2 Quality	PNUSAL000701-2_S3_L001		35.68	30.51							
		average depth	average read length	average mapped read	average insert length	median insert length	number of contigs	N50	longest contig	total genome length	
EDLB Seq Assembly	PNUSAL000701_CDC	114.46	148.48	147.01	455.02	433	28	301683	610341	2950095	
OH Seq Assembly	PNUSAL000701-2	61.59	148.72	146.98	436.87	415	70	115752	285326	2940872	
SNPs differ. from EDLB (1 < 1 Mbp)		PNUSAL000701-2_fastx5.fa PNUSAL000701_CDC_fastx5.fastq.gz-clean_PNUSAL000701									
		0									
PNUSAL000724-2		Fail – 10 SNP differences from EDLB reference									
		Quality (Q30)	Coverage								
Reads R1 Quality	PNUSAL000724-2_S5_L001		37.38	83.65							
Reads R2 Quality	PNUSAL000724-2_S5_L001		35.75	83.66							
		average depth	average read length	average mapped read	average insert length	median insert length	number of contigs	N50	longest contig	total genome length	
EDLB Seq Assembly	PNUSAL000724_CDC	69.97	142.64	140.65	406.64	322	19	477941	784047	2922498	
OH Seq Assembly	PNUSAL000724-2	169.95	144.03	142.77	326	287	21	244830	708855	2923375	
SNPs differ. from EDLB (1 < 1 Mbp)		PNUSAL000724-2_fastx5.fa PNUSAL000724_CDC_fastx5.fastq.gz-clean_PNUSAL000724									
		10									

1st MiSeq Validation Results

Accuracy	Reproducibility (inter-run imprecision)	Repeatability (intra-run imprecision)
# Correct 47 # Incorrect 2 Total Test 49	# Tested 4 # Replicates 3 Total test 12 # Discrepant 0	# Tested 1 # Replicates 2 Total Test 2 # Discrepant 0
Overall Accuracy 95.92%	Score 100%	Score 100%

Incorrect	Reason	Ohio	EDLB	Ohio-Repeat
<i>E.coli non-O157</i>	Low Coverage	Average Depth: 34.93	Average Depth: 67.09	Average Depth: 41.3
<i>Listeria</i>	10 SNP Difference	EDLB SNP Difference: 10		EDLB SNP Difference: 1

2nd MiSeq Validation

4-Certification strains (duplicated)

- *Salmonella enterica* serovar Heidelberg (CDC# C6472)
- *Listeria monocytogenes* (CDC# H8394)
- *Escherichia coli* (STEC)(CDC# 03-98)
- *Campylobacter jejuni* (CDC# 5480)



Created by monkik
from Noun Project



Created by monkik
from Noun Project



2nd MiSeq Validation Results



Enteric Diseases Laboratory Branch
CDC/NCEZID/DFWED
WGS Fastq File Report
Certification Strain for QA/QC



OH – WGS

Certification Set:
Laboratory Submitting Fastq File:
Submitted by:

Fastq File Name:
OH-CDC03-98GM-2-OH-M5185-170207_S2_L001_R1_001.fastq.gz
OH-C6472GM-2-OH-M5185-170207_S4_L001_R1_001.fastq.gz
OH-CDCH8394GM-2-OH-M5185-170207_S6_L001_R1_001.fastq.gz
OH-D5480GM-2-OH-M5185-170207_S8_L001_R1_001.fastq.gz

Sequencer type:
Date Files Received by CDC:

2017 WGS
Ohio (duplicate)
Glen McGillivray

OH-CDC03-98GM-2-OH-M5185-170207_S2_L001_R2_001.fastq.gz
OH-C6472GM-2-OH-M5185-170207_S4_L001_R2_001.fastq.gz
OH-CDCH8394GM-2-OH-M5185-170207_S6_L001_R2_001.fastq.gz
OH-D5480GM-2-OH-M5185-170207_S8_L001_R2_001.fastq.gz
Illumina MiSeq
2/15/2017

STEC O157:H7 strain 03-98

QC statistic	QC value (acceptable value)	Comments
Q-score R1	35.58 (min 30)	
Q-score R2	31.81 (min 30)	
Average coverage	131x (min 40x)	
Average insert size	410 (min 300)	
No. of hqSNPs compared to reference	0 (< 1/ MB)	0 hqSNP out of 5.0 Mbp
No. of allele differences compared to reference	NA	

Salmonella serovar Heidelberg strain C6472

QC statistic	QC value (acceptable value)	Comments
Q-score R1	35.34 (min 30)	
Q-score R2	31.04 (min 30)	
Average coverage	148x (min 40x)	
Average insert size	430 (min 300)	
No. of hqSNPs compared to reference	3 (< 1/ MB)	3 hqSNP out of 4.5 Mbp - passing
No. of allele differences compared to reference	NA	

PulseNet WGS Validation Results Summary Table**

	ID	Strain	PASS/FAIL	Comment
1	CDC03-98GM-OH-M5185-170207	<i>E. coli</i> O157:H7	Pass	0 hqSNP out of 5.0 Mbp
2	CDC03-98GM-2-OH-M5185-170207	<i>E. coli</i> O157:H7	Pass	0 hqSNP out of 5.0 Mbp
3	CDC03-98JM-OH-M5185-170203	<i>E. coli</i> O157:H7	Pass	0 hqSNP out of 5.0 Mbp
4	CDC03-98JM-2-OH-M5185-170203	<i>E. coli</i> O157:H7	Pass	0 hqSNP out of 5.0 Mbp
5	C6472GM-OH-M5185-170207	<i>Salmonella</i> Heidelberg	Pass	3 hqSNP out of 4.5 Mbp
6	C6472GM-2-OH-M5185-170207	<i>Salmonella</i> Heidelberg	Pass	3 hqSNP out of 4.5 Mbp
7	C6472JM-OH-M5185-170203	<i>Salmonella</i> Heidelberg	Pass	3 hqSNP out of 4.5 Mbp
8	C6472JM-2-OH-M5185-170203	<i>Salmonella</i> Heidelberg	Pass	3 hqSNP out of 4.5 Mbp
9	CDCH8394GM-OH-M5185-170207	<i>Listeria monocytogenes</i>	Pass	1 hqSNP out of 3.0 Mbp
10	CDCH8394GM-2-OH-M5185-170207	<i>Listeria monocytogenes</i>	Pass	1 hqSNP out of 3.0 Mbp
11	CDCH8394JM-OH-M5185-170203	<i>Listeria monocytogenes</i>	Pass	0 hqSNP out of 3.0 Mbp
12	CDCH8394JM-2-OH-M5185-170203	<i>Listeria monocytogenes</i>	Pass	1 hqSNP out of 3.0 Mbp
13	D5480GM-OH-M5185-170207	<i>Campylobacter jejuni</i>	Pass	0 hqSNP out of 1.7 Mbp
14	D5480GM-2-OH-M5185-170207	<i>Campylobacter jejuni</i>	Pass	0 hqSNP out of 1.7 Mbp
15	D5480JM-OH-M5185-170207	<i>Campylobacter jejuni</i>	Pass	0 hqSNP out of 1.7 Mbp
16	D5480JM-2-OH-M5185-170207	<i>Campylobacter jejuni</i>	Pass	0 hqSNP out of 1.7 Mbp

2nd MiSeq Validation Results

Accuracy	Reproducibility (inter-run imprecision)	Repeatability (intra-run imprecision) x2
# Correct 16 # Incorrect 0 Total Test 16	# Tested 8 # Replicates 2 Total test 16 # Discrepant 0	# Tested 4 # Replicates 2 Total Test 8 # Discrepant 0
Overall Accuracy 100%	Score 100%	Score 100%

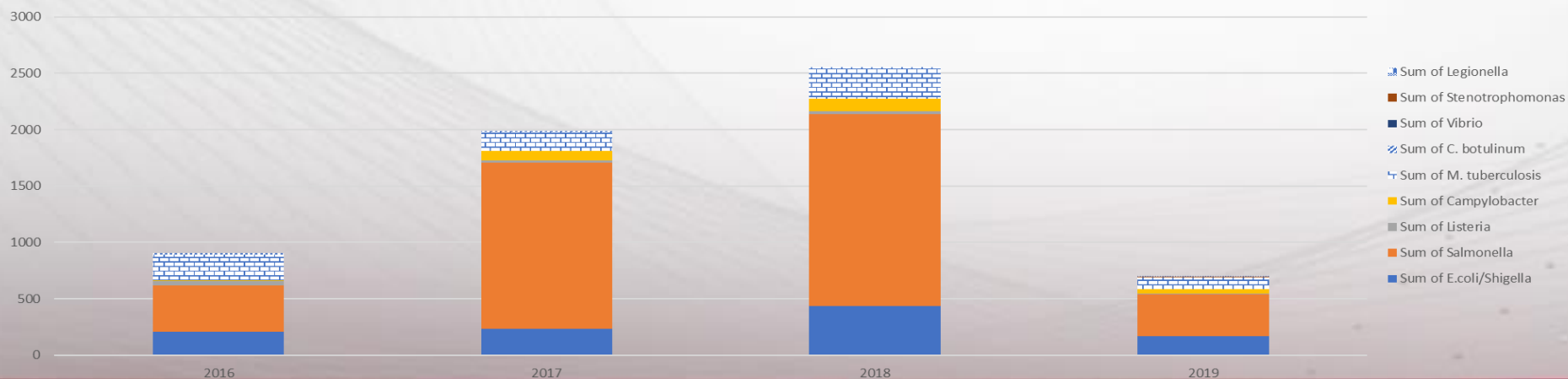
	0 hqSNP Difference	1 hqSNP Difference	2 hqSNP Difference	3 hqSNP Difference
16 Samples tested	9	3	0	4
	100 % Pass			

ODH Challenges

- Lack of Bioinformatic capability (software/personnel)
- Limit QC capability: Nanodrop, Qubit

WGS Growth

Year	<i>E.coli/Shigella</i>	<i>Salmonella</i>	<i>Listeria</i>	<i>Campylobacter</i>	<i>M. tuberculosis</i>	<i>C. botulinum</i>	<i>Vibrio</i>	<i>Stenotrophomonas</i>	<i>Legionella</i>	Total
2016	207	410	39	7	227	16	0	0	0	906
2017	232	1,472	21	86	172	2	0	0	0	1,985
2018	436	1,701	25	109	281	0	0	0	0	2,552
2019	168	374	7	38	100	0	4	6	4	701
Grand Total	1,043	3,957	92	240	780	18	4	6	4	6,144



Sequencing Future?

- *Hepatitis* (GHOST)
- *Cryptosporidium* (CryptoNet)
- *Norovirus* (CaliciNet)
- Carbapenem resistance?

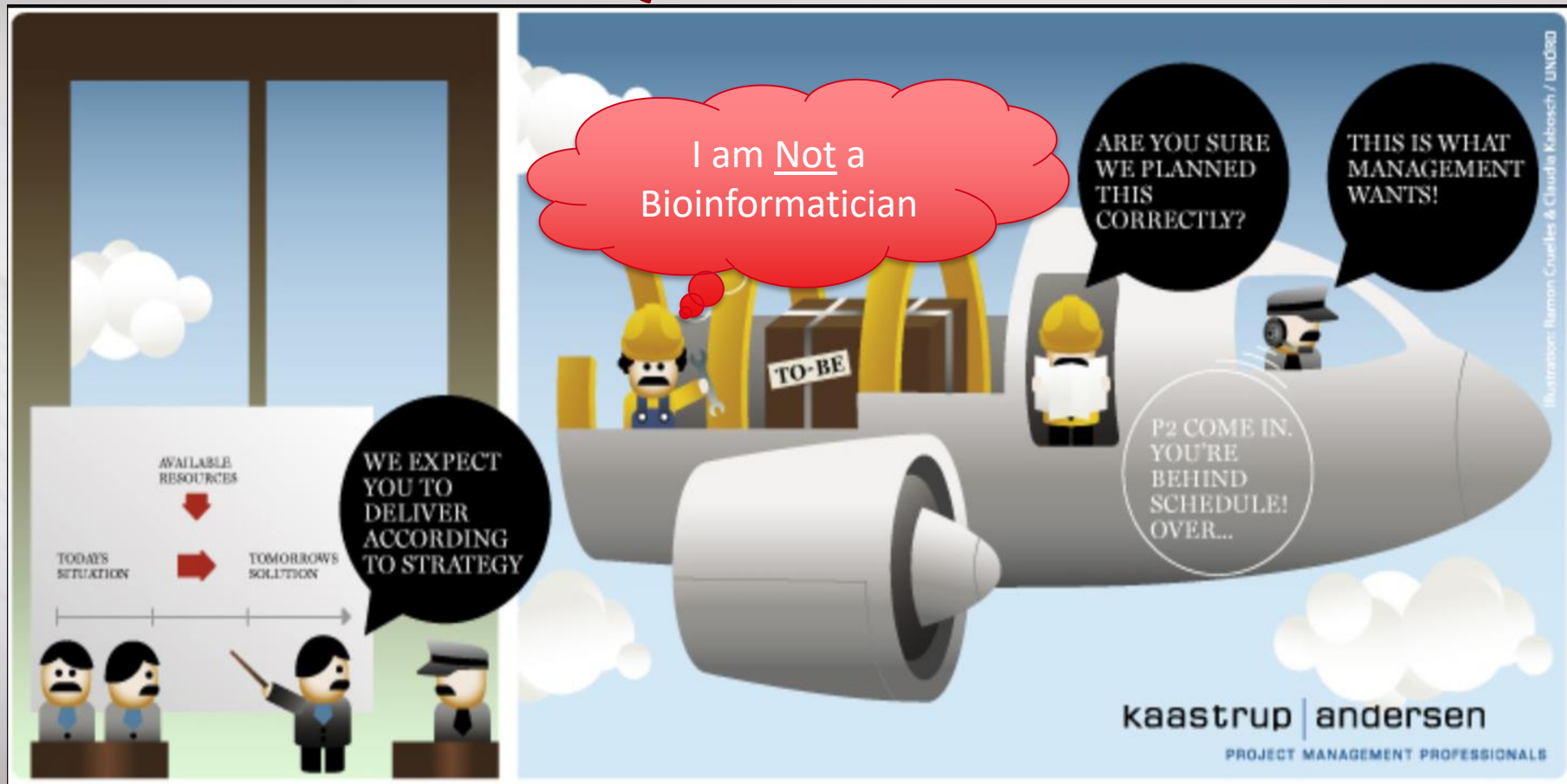
Pending ELC
Funding 2019-2020

Bioinformatics Future ?

- BioNumerics 7.x
- Galaxy
- AMD Regional Bioinformatician
- OAMD Platforms
- Colorado Virtual Machine

Knowledge &
Experience

Questions?



source: www.kaastrupandersen.dk

Thanks

Keoni Omura, MS

Laboratory Scientist

Ohio Department of Health

(614) 728-1729

Keoni.Omura@odh.ohio.gov

Thanks to CDC PulseNet group for all of the analysis and assistance.

Thanks ODH Laboratory Scientist who participated with the validation:

Eric Brandt, Glen McGillivary and Jade Mowery.

Images and Icons from Illuminia, Bionumerics, Kaastrup/Anderson & Noun Project