

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL			
Doc. No. PNIQ01	Ver. No. 01	Effective Date:	Page 1 of 13

1. **PURPOSE:** To describe the procedure for certifying personnel to enable full participation in PulseNet International WGS activities.
2. **SCOPE:** This procedure applies to all PulseNet personnel generating sequence data for surveillance activities within the PulseNet International network and is limited to wet lab processes only.
3. **DEFINITIONS/TERMS:**
 - 3.1 **Allele:** One of two or more alternative forms of a gene that arise by mutation and are found at the same place on a chromosome
 - 3.2 **APHL:** Association of Public Health Laboratories
 - 3.3 **ATCC:** American Type Culture Collection. A non-profit organization which collects, stores, and distributes standard reference microorganisms, cell lines, and other materials for research and development.
 - 3.4 **BaseSpace:** Illumina cloud-based computing environment for next-generation sequencing data analysis and management, including data sharing
 - 3.5 **CDC:** Centers for Disease Control and Prevention, Atlanta, GA, USA
 - 3.6 **cgMLST:** Core Genome Multi-Locus Sequencing Typing
 - 3.7 **Certification file evaluator:** An individual who evaluates certification files
 - 3.8 **Certification files:** Fastq files submitted by PulseNet participants for WGS certification evaluation
 - 3.9 **Core genome:** Genes shared by a vast majority (e.g., >98% of all strains of *Listeria monocytogenes* strains) of strains of the same species
 - 3.10 **Coverage:** The average number of reads that includes a given nucleotide in the reconstructed sequence
 - 3.11 **Critical Quality metrics:** Average de novo coverage, average quality (Q score), assembly length, secondary species abundance (contamination), percent core genome present, allele and hqSNP difference compared to the reference, and average read length. Failing to meet any one of these metrics will result in the rejection of the certification submission.
 - 3.12 **De Novo Assembly:** A sequence assembly generated from the short raw reads without the use of a reference genome
 - 3.13 **DNA:** Deoxyribonucleic Acid
 - 3.14 **FASTQ:** a text-based format for storing both a biological sequence and its corresponding quality scores
 - 3.15 **Fastq-Certified:** An individual who is certified in WGS laboratory methods for generating sequencing data
 - 3.16 **FTP:** File Transfer Protocol; a standard network protocol used to transfer computer files from one host to another host via Internet
 - 3.17 **Gzip:** A file format and a software application used for file compression and decompression to transfer data quickly over the Internet
 - 3.18 **hqSNP:** High Quality Single Nucleotide Polymorphism
 - 3.19 **Lyve-SET:** hqSNP pipeline developed by Katz *et al.* (2017), used by the Enteric Diseases Laboratory Branch, CDC, and imbedded as a workflow in Terra.Bio (Theiagen, Highlands Ranch, CO, USA)

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL			
Doc. No. PNIQ01	Ver. No. 01	Effective Date:	Page 2 of 13

3.20 MB: Mega Base

3.21 PNI: PulseNet International

3.22 PT: Proficiency Test

3.23 QA/QC: Quality assurance/quality control

3.24 Q score: The quality score for each individual base position in a sequence, indicating the accuracy of the base call. Phred scores are used, where $Q = -10\log(\text{Error Probability})$. The higher the quality score, the more reliable the base call. A Q30 means a 1 in 1000 likelihood of an incorrect base call at that position.

3.25 Read: A unit of continuous DNA sequence (based pairs) derived by sequencing a part of the fragmented target DNA

3.26 SOP: Standard Operating Procedure

3.27 Terra.Bio: Cloud-based sequence analysis platform developed by the Broad Institute of the Massachusetts Institute of Technology and the Harvard University and utilized by Theiagen Genomics (Highlands Ranch, CO, USA) to offer a for-fee common platform for public health laboratories to host, analyze and share data (Libuit *et al.*, 2023)

3.28 TSV: Tab Separated Values

3.29 WGS Certification Report: A report that contains the evaluation and results of the participant's WGS certification files

3.30 WGS: Whole Genome Sequencing

4. RESPONSIBILITIES:

4.1 Individuals performing PulseNet International WGS-related work must submit certification file(s) and have them approved before being able to generate sequence data for PulseNet surveillance

4.1.1 Certification serves as an initial competency assessment of newly trained personnel.

Submitted certification files must document the submitter's highest level of competence in producing raw sequence data.

4.1.2 Individuals can be certified for *Salmonella*, *Escherichia/Shigella*, *Listeria*, *Campylobacter*, and *Vibrio*

4.2 A laboratory is considered PulseNet-certified as long as at least one individual within the laboratory has passed the certification for at least one organism

4.2.1 In order to maintain its PulseNet certification status, a laboratory is required to participate and successfully pass the annual proficiency test (see SOP PNIQ02 for details)

5. PROCEDURE:

5.1 PulseNet International participants purchase the WGS laboratory certification strain set for organisms they wish to become certified (if they do not already have them) from ATCC (<https://www.atcc.org>)

5.1.1 The WGS Laboratory Certification set includes *Campylobacter*, *Escherichia*, *Listeria*, *Salmonella*, and *Vibrio*

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL			
Doc. No. PNIQ01	Ver. No. 01	Effective Date:	Page 3 of 13

Genus/Species	Serotype	ATCC Number	Strain designation	Price (US\$) ²
<i>Campylobacter jejuni</i>	NA	ATCC 33560	CIP-702	\$269.00
<i>Escherichia coli</i> ¹	O6:H1	ATCC 25922	FDA strain Seattle 1946	\$77.00
<i>Listeria monocytogenes</i>	Lineage II	ATCC BAA-679	EGDe	\$431.00
<i>Salmonella enterica</i>	Typhimurium	ATCC 51812	U24	\$431.00
<i>Vibrio parahaemolyticus</i>	O3:K6	ATCC 17802	EB101 (P. Baumann 113)	\$77.00
Total				\$1285.00

Table 1. ATCC strains included in the PNI certification panel

¹ Covers also certification for the *Shigella* spp.

² Freeze-dried preparation; list pricing for United States (Oct 2023). Pricing will depend on the shipping destination and will, in certain cases, be significantly higher than what is quoted in the table.

- 5.1.2 Go to the ATCC website to create an account/profile to get the exact pricing and place the order: <https://www.atcc.org/web-profile/create-a-web-profile>
- 5.2 Freeze (-70°C) or stock the certification strains according to your laboratory's policy within one week of receiving them to ensure that your laboratory will have stock cultures of this PNI certification set for future use, including the PulseNet certification of additional personnel. Avoid multiple passes and freeze/thaw cycles for frozen cultures.
- 5.3 Sequence the strains according to methods accepted by PulseNet:
 - 5.3.1 Any DNA extraction method producing DNA that meets the minimum quality requirements outlined in the PulseNet SOP PNL33 is acceptable
 - 5.3.2 Acceptable library preparation kits are the following: Illumina Nextera XT (PNL34), Illumina DNA Prep (PNL35, PNL44), Qiagen QIAseq FX (PNL36), KAPA HyperPlus (PNL37), NEBNext Ultra II FS (PNL42), and Illumina DNA PCR-Free (PNL43).
NOTE: library preparations from automated liquid handlers are not accepted.
 - 5.3.3 Acceptable sequencing platforms are the following: Illumina MiSeq (PNL38), Illumina MiniSeq (PNL39), Illumina iSeq (PNL40) and Illumina NextSeq.
NOTE: sequencing chemistries (read lengths) below 2x150 bp are not accepted.
 - 5.3.3.1 Please follow the recommended genome loads outlined in the Appendix PNIQ01-1. It is acceptable to include routine isolates in the same sequencing run with the certification strains.
 - 5.3.3.2 Make sure that the output fastq files are in the **.fastq.gz** format
- 5.4 Share the certification fastq.gz files with the certification file evaluator
 - 5.4.1 **Laboratories using Terra.Bio for sequence analysis:**
 - 5.4.1.1 Contact Theiagen (michelle.scribner@theiagen.com; frank.ambrosio@theiagen.com) to receive access to the PNI QA Evaluation Workspace:
https://app.terra.bio/#workspaces/theiagen-validations/PNI_Eval_Workspace/data
 - 5.4.1.2 Prepare the files for the upload to the PNI QA Evaluation Workspace:

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL			
Doc. No. PNIQ01	Ver. No. 01	Effective Date:	Page 4 of 13

5.4.1.2.1 Save the certification fastq.gz files to a local drive folder

5.4.1.2.2 Prepare the metadata sheet .tsv file for the certification samples using the following naming conventions:

5.4.1.2.2.1 Data table naming convention:

entity:*NameOfInstitution_illuminaPE_id*

e.g.: entity:USA_CDC_illuminaPE_id

5.4.1.2.2.1.1 The data table name should include the name of your institution in place of the italics in the example above. The name may not include any spaces or special characters, except for underscores (“_”). Please include this entire string (from the e in entity through the d in _id) in cell 1A of the metadata spreadsheet.

NOTE: *place an identifier for the country in front of the name of the institution if the institution name is the same in two countries, e.g., CDC USA and CDC Taiwan (USA_CDC; TW_CDC)*

5.4.1.2.2.2 Sample naming convention:

SampleIdentifier_NameOfInstitution_InitialsOfTechnician

e.g.: ATCC_33560_USA_CDC_et

NOTE: *place an identifier for the country in front of the name of the institution if the institution name is the same in two countries, e.g., CDC USA and CDC Taiwan (USA_CDC; TW_CDC)*

	A	B	C	D
1	entity:USA_CDC_illuminaPE_id	read1	read2	
2	ATCC_10708_USA_CDC_et	10708-C2-M3235-23-002_S17_L001_R1_001.fastq.gz	10708-C2-M3235-23-002_S17_L001_R2_001.fastq.gz	
3	ATCC_33560_USA_CDC_et	ATCC-33560A-M947-22-040_S5_L001_R1_001.fastq.gz	ATCC-33560A-M947-22-040_S5_L001_R2_001.fastq.gz	
4				
5				
6				

Figure 1. Screenshot of the metadata tsv file opened in Excel

5.4.1.3 Upload the fastq.gz files and the metadata .tsv file to the PNI QA Evaluation Workspace

5.4.1.4 Optional: Start the “TheiaProk_Illumina_PE_PHB” workflow on the certification files

5.4.2 Laboratories not using Terra.Bio but using Illumina BaseSpace:

5.4.2.1 The fastq.gz files must be named using the PulseNet standard file naming convention: *SampleIdentifier-InitialsOfTechnician-NameOfInstitution-IlluminaInstrumentID-RunStartDate* e.g., ATCC-33560-et-USA-CDC-M3235-231031

NOTE: *place an identifier for the country in front of the name of the institution if the institution name is the same in two countries, e.g., CDC USA and CDC Taiwan (USA-CDC; TW-CDC)*

5.4.2.1.1 Illumina instrument IDs are typically in the following formats:

5.4.2.1.1.1 MiSeq MXXXXX

5.4.2.1.1.2 MiniSeq MNXXXXX

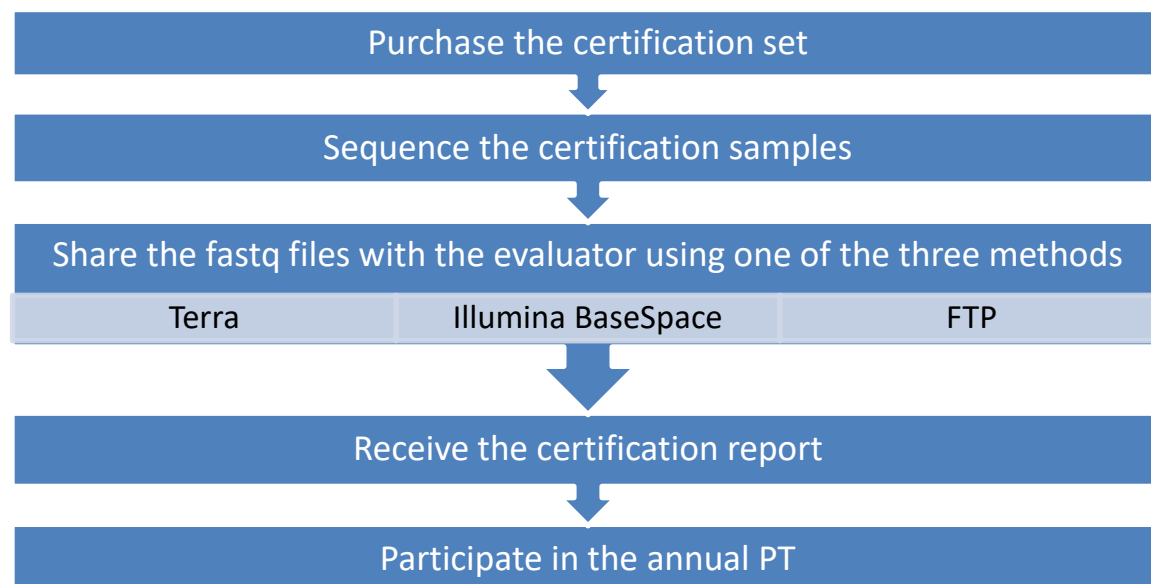
5.4.2.1.1.3 iSeq FSXXXXX

5.4.2.1.1.4 NextSeq VHXXXXXX/VLXXXXXX/NBXXXXXX/NSXXXXXX

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL			
Doc. No. PNIQ01	Ver. No. 01	Effective Date:	Page 5 of 13

- 5.4.2.1.2 Run start date format: YearMonthDate (YYMMDD)
- 5.4.2.2 Share the **Project** containing the certification fastq files with pulsenetngslab@cdc.gov listed as a collaborator in a separate BaseSpace project dedicated for certification
 - 5.4.2.2.1 Name the project using the following naming convention:
NameOfInstitution_Certification e.g., USA_CDC_Certification
NOTE: *place an identifier for the country in front of the name of the institution if the institution name is the same in two countries, e.g., CDC USA and CDC Taiwan (USA_CDC; TW_CDC)*
- 5.4.2.3 Share the **Run** containing the certification samples with pulsenetngslab@cdc.gov listed as a collaborator
- 5.4.3 **Laboratories not using Terra.Bio or BaseSpace:**
 - 5.4.3.1 Contact pulsenetngslab@cdc.gov to receive instructions on how to upload the fastq.gz files to the PulseNet QA ftp site
- 5.4.4 Notify the certification file evaluator on the availability of the certification files by emailing pulsenetngslab@cdc.gov
 - 5.4.4.1 Use the subject line: PNI certification – *NameOfInstitution*
 - 5.4.4.2 Include the following information in the body of the email:
 - 5.4.4.2.1 Name and email address of the analyst who performed the library prep and sequencing
 - 5.4.4.2.2 Library prep kit used to generate the certification files
 - 5.4.4.2.3 Sequencing platform used generate the certification files
 - 5.4.4.2.4 If sharing in BaseSpace: the list of filenames submitted and the name of the project in which the files were placed
- 5.5 Submitters will receive a WGS Certification Report summarizing the results of their certification file(s) evaluation (See Appendix PNIQ01-2 for an example of a WGS Certification Report).
 - 5.5.1 Certifications fastq files are evaluated based on the criteria outlined in Appendix PNIQ01-3.
 - 5.5.2 If the submitted certification files pass the certification evaluation, the submitter is considered certified as long as they remain in their current laboratory and that laboratory successfully completes the annual proficiency testing (SOP PNIQ02). All fastq-certified individuals should perform routine sequencing at least every six months.
 - 5.5.3 If the submitted certification files do not pass the certification evaluation, the individual will need to review the troubleshooting comments received from the evaluator and resubmit once results have improved
 - 5.5.3.1 If the submitter fails certification three times, the individual will not be allowed to submit again for six months. Before resubmitting, the individual will be expected to work with their coordinating laboratory until satisfactory results are achieved. This includes but is not limited to troubleshooting and training in the PulseNet WGS protocols.

6. FLOW CHART:



7. RELATED DOCUMENTS:

7.1. Available at https://www.aphl.org/programs/food_safety/Pages/PulseNet-International-SOPs.aspx

Document Number	Title
PNL33	DNA Extraction and QC SOP
PNL34	Illumina Nextera XT Library Preparation SOP
PNL35	Illumina DNA Prep Library Preparation SOP
PNL36	Qiagen QIAseq FX Library Preparation SOP
PNL37	KAPA HyperPlus Library Preparation SOP
PNL38	Sequencing on the MiSeq SOP
PNL39	Sequencing on the MiniSeq SOP
PNL40	Sequencing on the iSeq SOP
PNL42	NEBNext Ultra II FS Library Preparation SOP
PNL43	Illumina DNA PCR-Free Library Preparation SOP
PNL44	PulseNet Rapid Library Preparation SOP
PNQ07	Illumina Sequence Data QC SOP

8. REFERENCES:

- 8.1 Katz L.S., Griswold T., Williams-Newkirk A.J., Wager, D., Petkau, A., Sieffert, C., Van Domselaar, G., Deng, X., Carleton, H.A. (2017). A Comparative Analysis of the Lyve-SET Phylogenomics Pipeline for Genomic Epidemiology of Foodborne Pathogens. *Frontiers Microbiol.* 8:375.

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL			
Doc. No. PNIQ01	Ver. No. 01	Effective Date:	Page 7 of 13

8.2 Libuit K.G., Doughty E.L., Otieno J.R., Ambrosio F., Kapsak C.J., Smith E.A., Wright S.M., Scribner M.R., Petit III R.A., Mendes C.I., Huergo M., Legacki G., Loreth C., Park D.J., Sevinsky J.R. (2023) Accelerating bioinformatics implementation in public health. Microbial Genomics 9:001051.

9. CONTACTS:

9.1 CDC USA PulseNet NGS Laboratory: pulsenetngslab@cdc.gov

9.2 PulseNet International Quality Assurance Coordinator Eija Trees: ehyytia-trees@cdc.gov

9.3 Theiagen:

9.3.1 Michelle Scribner: michelle.scribner@theiagen.com

9.3.2 Frank Ambrosio: frank.ambrosio@theiagen.com

10. AMENDMENTS: NA

11. APPROVAL SIGNATURES:

Approved By: _____ Date: _____
PulseNet QA/QC Personnel

Approved By: _____ Date: _____
PulseNet WGS Technical Lead

Approved By: _____ Date: _____
PulseNet International Coordinator

Approved By: _____ Date: _____
PulseNet Response and Outbreak Management Team Lead

Approved By: _____ Date: _____
Enteric Diseases Laboratory Branch Chief

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL			
Doc. No. PNIQ01	Ver. No. 01	Effective Date:	Page 8 of 13

Appendix PNIQ01-1: Acceptable Library Preparation Kits, Sequencing Platforms, Sequencing Chemistries and Validated DNA Loading Capacities

Nextera XT¹, Illumina (PNL34)

Illumina Sequencing Kit	DNA Load (Mb)
MiSeq v2, 300	60
MiSeq v2, 500	80
MiSeq v3, 600 ² no <i>E. coli/Shigella/Vibrio</i>	200
MiSeq v3, 600 ² <i>E. coli/Shigella/Vibrio</i>	175
MiSeq Micro, 300	35
MiSeq Nano, 500	13
MiniSeq Mid Output	60
MiniSeq High Output	100

¹300 cycle chemistry not acceptable for *Escherichia*, *Shigella* or *Vibrio*

²Recommended to run 500 cycles for better R2 quality scores

Illumina DNA Prep (DNA Flex) (PNL35)

Illumina Sequencing Kit	DNA Load (Mb)
MiSeq v2 300	90
MiSeq v2 500	100
MiSeq v3 600 ¹ no <i>E. coli/Shigella/Vibrio</i>	200
MiSeq v3 600 ¹ <i>E. coli/Shigella/Vibrio</i>	175
Micro 300 no <i>E. coli/Shigella/Vibrio</i>	35
Micro 300 <i>E. coli/Shigella/Vibrio</i>	30
Nano 500	13
Nano 300 no <i>E. coli/Shigella/Vibrio</i>	13
Nano 300 <i>E. coli/Shigella/Vibrio</i>	10
MiniSeq Mid Output	60
MiniSeq High Output	100
iSeq <i>E. coli/Shigella/Vibrio</i>	25
iSeq no <i>E. coli/Shigella/Vibrio</i>	35
NextSeq 1000/2000 P1 300	400 ²

¹Recommended to run 500 cycles for better R2 quality scores

²Validation not completed

QIAseq FX, Qiagen (PNL36)

Illumina Sequencing Kit	DNA Load (Mb)
MiSeq v2 300	80
MiSeq v2 500	100

KAPA HyperPlus, Roche (PNL37)

Illumina Sequencing Kit	DNA Load (Mb)
MiSeq v2 300, no <i>E. coli/Shigella/Vibrio</i>	80
MiSeq v2 300, <i>E. coli/Shigella/Vibrio</i>	75
MiSeq v2 500	100

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL

Doc. No. PNIQ01

Ver. No. 01

Effective Date:

Page 9 of 13

NEBNext Ultra II FS (PNL42)

Illumina Sequencing Kit	DNA Load (Mb)
MiSeq v2 300, <i>no E. coli/Shigella/Vibrio</i>	90
MiSeq v2 300, <i>E. coli/Shigella/Vibrio</i>	75
MiSeq v2 500	100

Illumina DNA PCR-Free (PNL43)

Illumina Sequencing Kit	DNA Load (Mb)
MiSeq v2 300	90
MiSeq v2 500	100

PulseNet Rapid Prep (PNL44)

Illumina Sequencing Kit	DNA Load (Mb)
MiSeq v2 300	80
MiSeq v2 500	100

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL

Doc. No. PNIQ01

Ver. No. 01

Effective Date:

Page 10 of 13

Appendix PNIQ01-2: Example WGS Fastq Certification Report

Lab ID – WGS

Enteric Diseases Laboratory Branch

CDC/NCEZID/DFWED

WGS Fastq File Certification Report

PulseNet International Certification Strains for QA/QC



Certification Set:

PNI WGS Fastq Certification Set:

Escherichia, *Salmonella*, *Listeria*,

Campylobacter, and *Vibrio*

Laboratory Submitting Files:

Submitted by:

Library Prep Kit:

Sequencer Type:

Date Fastq Files Shared with Evaluator:

Fastq File Names:

<i>Escherichia coli</i> strain ATCC-25922		
QC statistic	QC value (acceptable value)	Comments
Average quality	(≥ 30.0)	
Average coverage	($\geq 40x$)	
Genus/species/serotype	(<i>Escherichia coli</i> O6:H1)	
Average Read Length	(300 cycle chemistry: ≥ 135 bp; 500 cycle chemistry: ≥ 225 bp)	
Secondary species abundance	(≤ 0.01)	
No. of hqSNPs compared to reference	(< 1/MB)	
Assembled genome size	(5.14 MB, 5% error margin 4.88-5.40)	

<i>Salmonella enterica</i> sp. <i>enterica</i> serovar Typhimurium strain ATCC-51812		
QC statistic	QC value (acceptable value)	Comments
Average quality	(≥ 30.0)	
Average coverage	($\geq 30x$)	
Genus/species/serovar	(<i>Salmonella enterica</i> sp. <i>enterica</i> serovar Typhimurium)	
Average Read Length	(300 cycle chemistry: ≥ 135 bp; 500 cycle chemistry: ≥ 225 bp)	
Secondary species abundance	(≤ 0.01)	
No. of hqSNPs compared to reference	(< 1/ MB)	

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL

Doc. No. PNIQ01

Ver. No. 01

Effective Date:

Page 11 of 13

Assembled genome size

(4.94 MB, 5% error margin 4.69-5.19)

***Listeria monocytogenes* strain ATCC-BAA679**

QC statistic	QC value (acceptable value)	Comments
Average quality	(≥ 30.0)	
Average coverage	($\geq 20x$)	
Genus/species	(<i>Listeria monocytogenes</i>)	
Average Read Length	(300 cycle chemistry: ≥ 135 bp; 500 cycle chemistry: ≥ 225 bp)	
Secondary species abundance	(≤ 0.01)	
No. of hqSNPs compared to reference	(< 1/MB)	
Assembled genome size	(2.91 MB, 5% error margin 2.76-3.06)	

***Campylobacter jejuni* strain ATCC-33560**

QC statistic	QC value (acceptable value)	Comments
Average quality	(≥ 30.0)	
Average coverage	($\geq 20x$)	
Genus/species	(<i>Campylobacter jejuni</i>)	
Average Read Length	(300 cycle chemistry: ≥ 135 bp; 500 cycle chemistry: ≥ 225 bp)	
Secondary species abundance	(≤ 0.01)	
No. of hqSNPs compared to reference	(< 1/MB)	
Assembled genome size	(1.73 MB, 5% error margin 1.64-1.82)	

***Vibrio parahaemolyticus* strain ATCC-17802**

QC statistic	QC value (acceptable value)	Comments
Average quality	(≥ 30.0)	
Average coverage	($\geq 40x$)	
Genus/species	(<i>Vibrio parahaemolyticus</i>)	
Average Read Length	(300 cycle chemistry: ≥ 135 bp; 500 cycle chemistry: ≥ 225 bp)	
Secondary species abundance	(≤ 0.01)	
No. of hqSNPs compared to reference	(< 1/MB)	
Assembled genome size	(5.06 MB, 5% error margin 4.81-5.31)	

PULSENET INTERNATIONAL STANDARD OPERATING PROCEDURE FOR WGS CERTIFICATION OF PERSONNEL

Doc. No. PNIQ01

Ver. No. 01

Effective Date:

Page 12 of 13

QA/QC Certification Result:

Escherichia **Fastq files:** ~~Satisfactory~~ Needs Improvement

Salmonella **Fastq files:** ~~Satisfactory~~ Needs Improvement

Listeria **Fastq files:** ~~Satisfactory~~ Needs Improvement

Campy **Fastq files:** ~~Satisfactory~~ Needs Improvement

Vibrio **Fastq files:** ~~Satisfactory~~ Needs Improvement

Certification is valid as long as laboratory successfully completes annual proficiency tests.

Appendix PNIQ01-3: Critical Quality Metric Thresholds/Acceptable Ranges for PulseNet Certification

Organism	Average read length (bp) ¹	hqSNP difference compared to reference	Average denovo coverage ²	Average quality ³	Assembly length (MB) ⁴	Secondary species abundance	% Core present ⁵	Allele difference compared to reference ⁵
<i>Listeria monocytogenes</i>	135/225	< 1/MB	≥ 20x	≥ 30	2.8-3.2	≤ 1.0	≥ 95	≤ 3
<i>Escherichia coli</i>	135/225	< 1/MB	≥ 40x	≥ 30	4.9-6.0	≤ 1.0	≥ 85	≤ 3
<i>Salmonella enterica</i>	135/225	< 1/MB	≥ 30x	≥ 30	4.4-5.7	≤ 1.0	≥ 85	≤ 3
<i>Campylobacter jejuni</i>	135/225	< 1/MB	≥ 20x	≥ 30	1.4-2.2	≤ 1.0	≥ 85	≤ 3
<i>Vibrio parahaemolyticus</i>	135/225	< 1/MB	≥ 40x	≥ 30	4.9-5.5	≤ 1.0	NA ⁶	NA ⁶

¹Average for R1 and R2 read lengths before trimming

²After quality-based trimming

³Average for R1 and R2 Q scores before trimming

⁴The sequence assembly lengths listed in the table apply to the entire species. For fastq certification, the assembly lengths are strain-specific based on the reference sequence generated by CDC with ± 5% margin of error.

⁵Does not apply to sequences analyzed through Terra.Bio.

⁶A cgMLST scheme for *Vibrio* is still under development.