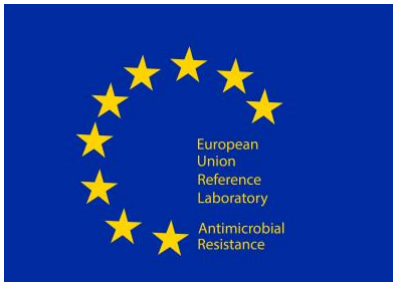
PROTOCOL

DTU Genomic Proficiency Test 2024

EU Reference Laboratory for Antimicrobial Resistance
National Food Institute
Technical University of Denmark



October 2024

Table of Contents

1. Introduction

The DTU Genomic Proficiency Test (PT) 2024 marks the fifth iteration of the DTU Genomic PTs. Its primary aim is to evaluate and compare the technical and analytical capabilities of participating laboratories in Whole Genome Sequencing (WGS) and various bioinformatics analyses of bacterial genomes, including Multilocus Sequence Typing (MLST), plasmid characterization, and identification of antimicrobial resistance (AMR) mechanisms. Engaging in the DTU Genomic PT 2024 will support the development of reliable laboratory results for monitoring and research purposes. The evaluation will not yield a pass or fail outcome; rather, participants are encouraged to assess their own performance and consider any necessary adjustments to their pipelines based on the results.

The DTU Genomic PT 2024 is coordinated by the National Food Institute, Technical University of Denmark (DTU), and is funded by the European Commission (EC), via the European Union Reference Laboratory for Antimicrobial Resistance (EURL-AR) and the Fleming Fund (SEQAFRICA Regional Grant). All members of the EURL-AR network are invited to participate. Additional countries from Africa and Asia are invited via The Fleming Fund, which is a UK aid investment to tackle AMR in low- and middle-income countries around the world. The program is managed by the UK Department of Health and Social Care, in partnership with Mott MacDonald, the Fleming Fund Grants Management Agent. Through the Fleming Fund's SeqAfrica and EQAsia projects, laboratories from the Africa Pathogen Genomics Initiative (Africa PGI), the Institut Pasteur's SARA network, the ACORN network, and the EQAsia laboratory network were also invited to register for participation.

The DTU Genomic PT 2024 focuses on the following three organisms, each represented by two strains (please note that signing up for each organism separately is allowed):

1. *Escherichia coli*
2. *Staphylococcus aureus*
3. *Enterococcus* (*E. faecalis* and/or *E. faecium*)

Two strains from each of the above three organisms are shipped to the participants as live bacterial cultures and are requested to be handled for WGS. Institutes/organizations that signed up to participate will receive the PT-material (live bacterial cultures) according to the registered sign-up information.

Attention! The DTU Genomic PT 2024 covers only **single-end or paired-end short read sequences** (FASTQ files). Assessment of long-read sequences (*e.g.*, Oxford Nanopore Technologies - ONT) is not offered as part of the present PT.

2. Shipping and Receipt of Test Material

In October 2024, laboratories located in Europe (members of the EURL-AR network), in Africa (SEQAFRICA partners, Institut Pasteur SARA network members and Africa PGI network members), and in Asia (ACORN network members and members of the EQAsia laboratory network), will receive a parcel containing two *E. coli*, two *S. aureus* and two *Enterococcus* strains in transport swabs (contents of the parcel

will correspond to the registered sign-up information). Bacterial transport swabs are shipped as UN3373, Biological Substance Category B. Information on the test material and the sample naming is presented in **Table 1**.

Attention! On arrival of the parcel to the laboratory, open the parcel to confirm that the content is as listed in the cover letter. **Please confirm receipt of the parcel through the confirmation form enclosed in the shipment or send by email.**

Table 1. Organisms and test strains codes included in the DTU Genomic PT 2024.

Organism	Test strain code
Escherichia coli	GPT24-01
	GPT24-02
Staphylococcus aureus	GPT24-03
	GPT24-04
Enterococcus faecalis	GPT24-05
Enterococcus faecium	GPT24-06

3. Handling and Storage of Test Material

Live bacterial cultures are shipped as transport swabs. Upon receipt, store the transport swabs at 4-8°C. The PT organizers encourage participants to subculture and prepare the bacterial cultures for storage in their strain collection (*e.g.*, in a -80°C freezer) within 48 hours of receiving the parcel, according to own standard procedures.

4. Analysis of Test Material and Result Submission

The bacterial strains should be cultured on appropriate agar and under appropriate growth conditions, according to each laboratory's routine procedures. Following incubation and assessment of purity of the bacterial cultures, DNA extraction and WGS should be performed, according to the laboratory's standard procedures. Please register information about the methods applied via the **Test forms (Appendix 3)**.

Proceed to the *in-silico* analyses of the obtained WGS data, which includes the identification of the:

1. Sequence Type (ST) by Multilocus Sequence Typing (MLST)
2. Plasmid replicons
3. AMR mechanisms (AMR genes and chromosomal point mutations)
4. AMR phenotype

Participants are requested to submit the raw sequencing data (FASTQ files) as well as the results of the different bioinformatics analyses mentioned above. An overview of the requested data and results to be submitted is provided in **Table 2** .

Table 2. Overview of the requested data and results to be submitted for the DTU Genomic PT 2024.

Item no	Type of data/results to be submitted
1	Raw sequencing data (FASTQ files)
2	Method information
3	Sequence Type (ST) by MLST
4	Plasmid replicons
5	**AMR genes***
6	**AMR chromosomal mutations***
7	**Predicted AMR phenotype***

* Only related to the antimicrobial compounds included in the DTU Genomic PT 2024, for each organism (Table 3).

Attention! Please report AMR genes, AMR chromosomal mutations and predicted AMR phenotype **only for the antimicrobials included in the DTU Genomic PT 2024 for each organism, according to Table 3**. Submitted AMR genes, AMR chromosomal mutations and predicted AMR phenotype related to **antimicrobial compounds other than those listed in Table 3 for each organism are not part of the expected results and will be marked as incorrect.**

Table 3. Antimicrobial agents included in the DTU Genomic PT 2024 for each organism.

Class	Antimicrobial	Abbreviation	*E. coli*	*S. aureus*	*Enterococcus*
Aminoglycosides	Amikacin	AMI	x	-	-
	Gentamicin	GEN	x	x	x
	Kanamycin	KAN	-	x	-
	Streptomycin	STR	-	x	-
Amphenicols	Chloramphenicol	CHL	x	x	x
Beta-lactams	Ampicillin	AMP	x	-	x
	Cefepime	FEP	x	-	-
	Cefotaxime	FOT	x	-	-
	Cefoxitin	FOX	x	x	-
	Ceftazidime	TAZ	x	-	-
	Ertapenem	ETP	x	-	-
	Imipenem	IMI	x	-	-
	Meropenem	MERO	x	-	-
	Penicillin (benzylpenicillin)	PEN	-	x	-
	Temocillin	TRM	x	-	-
Folate pathway antagonists	Sulfamethoxazole	SMX	x	x	-
	Trimethoprim	TMP	x	x	-
Glycopeptides	Teicoplanin	TEI	-	-	x
	Vancomycin	VAN	-	x	x
Lincosamides	Clindamycin	CLI	-	x	-
Lipopeptides	Daptomycin	DAP	-	-	x
Macrolides	Azithromycin	AZI	x	-	-
	Erythromycin	ERY	-	x	x
Oxazolidinones	Linezolid	LZD	-	x	x
Pleuromutilin	Tiamulin	TIA	-	x	-
Polymyxins	Colistin	COL	x	-	-
Pseudomonic acid	Mupirocin	MUP	-	x	-
Quinolones	Ciprofloxacin	CIP	x	x	x
	Nalidixic acid	NAL	x	-	-
Rifamycin	Rifampin	RIF	-	x	-
Steroid antibacterials	Fusidic acid	FUS	-	x	-
Streptogramins	Quinupristin/dalfopristin	SYN	-	x	x
Tetracyclines	Tetracycline	TET	x	x	x
	Tigecycline	TGC	x	-	x

4.1. Submission of raw sequencing data

Participants should upload their raw sequencing data (FASTQ files) via the ScienceData platform (https://sciencedata.dk/) using the unique link provided in the cover letter. This cover letter is included as a hardcopy with the shipment of the PT material and send via email as a PDF to the PT contact persons.

For detailed instructions on how to upload your files to ScienceData, please refer to Appendix 1. Only FASTQ files uploaded in your ScienceData folder by the submission deadline will be considered for evaluation. The FASTQ files must be named according to the guidelines provided in the box below. A pre-screening step will be performed to check the sequence file format (FASTQ) and file name. Files that do not comply with the present submission guidelines will be excluded from further analysis.

Attention! Before submitting your FASTQ files to ScienceData, ensure they are named as follows to match the corresponding samples. Your laboratory ID is located at the top of your cover letter.

- Read 1: [Lab ID]_[Strain code]_R1.fastq.gz (*Example*: 2024-01_GPT24-01_R1.fastq.gz)
- Read 2: [Lab ID]_[Strain code]_R2.fastq.gz (*Example*: 2024-01_GPT24-01_R2.fastq.gz)

When uploading the files to ScienceData, please verify that the **file sizes on your system match the file sizes displayed on the platform** after the upload. Participants are also required to upload the MD5 Checksum for each uploaded FASTQ file. The MD5 Checksum is used to verify the file integrity, as any modification to the file will result in a different MD5 hash. For detailed instructions on uploading files and the MD5 Checksum to ScienceData, please refer to Appendix 1.

4.2. Submission of bioinformatics analyses results

The DTU Webtool is used for submitting both the method information and the results of all *in-silico* analyses mentioned above. The webtool manual (Appendix 2) provides detailed instructions on how to submit the results. The PT organizers strongly recommend that laboratories read it carefully before submitting their data. The person(s) responsible for submitting the results through the DTU Webtool are encouraged to have the completed Test forms (Appendix 3) on hand during the process.

The DTU Webtool can be accessed at https://genomic-pt.dtu.dk. Regarding login information, your **personal login ID** will be sent to you by email, along with instructions on how to generate a password to access the webtool.

Attention! Participants must carefully evaluate the AMR results before submitting them via the DTU Webtool. It is highly recommended that this evaluation involves collaboration between a bioinformatician and a microbiologist with expertise in AMR.

Attention! Before finally submitting your results in the DTU Webtool, please ensure that you have filled in all the relevant fields, as **you can only 'finally submit' once**. After, "Final submit", data entry will be blocked.

4.2.1. Method Information

Details on how to submit method information through the DTU Webtool are provided in Appendix 2. Additionally, test forms that summarize your results prior to entering them into the DTU Webtool can be found in Appendix 3. The following method information must be submitted:

- Information related to the received test material
- Details on the WGS and bioinformatic analysis methods used
- Identification of ST by MLST
- Identification of plasmid replicons
- Identification of AMR genes, chromosomal mutations responsible for AMR, and the predicted AMR phenotype

4.2.2. Multilocus Sequence Typing (MLST)

The ST is submitted via the DTU Webtool in the relevant tab. In the MLST tab, participants are required to submit the allelic numbers for each of the seven chromosomal housekeeping genes included in the MLST scheme for each organism (**Table 4**). Additionally, participants should submit the ST for each sample. If a result does not show a perfect match or if an allele cannot be detected, enter "0".

Table 4. Alleles included in the MLST schemes relevant for the DTU Genomic PT 2024.

	Allele 1	**Allele 2**	**Allele 3**	**Allele 4**	**Allele 5**	**Allele 6**	**Allele 7**
E. coli	*adk*	*fumC*	*gyrB*	*icd*	*mdh*	*purA*	*recA*
S. aureus	*arcC*	*aroE*	*glpF*	*gmk*	*pta*	*tpi*	*yqiL*
E. faecalis	*gdh*	*gyd*	*pstS*	*gki*	*aroE*	*xpt*	*yqiL*
E. faecium	*atpA*	*ddl*	*gdh*	*purK*	*gyd*	*pstS*	*adk*

4.2.3. Plasmid Replicons

Participants are required to submit plasmid replicon data using the dropdown menu provided in the DTU Webtool. This year, to accommodate the use of different plasmid classification tools and their varying outputs, we have simplified what should be reported.

For ***Enterobacteriaceae***, replicons are grouped into larger incompatibility groups based on the available literature[1]. This means that **only this larger incompatibility groups should be reported**, such as FIA, FIB, FIC, FII, HI1, HI2, N, P, or Q. For example, if your tool outputs "IncFII(pAR0022)," you should report only "FII" in the DTU Webtool.

For **Gram-positive bacteria**, a similar grouping approach has been adopted for simplicity, following the plasmid classification frameworks of Lanza *et al*. (2015)[2] and Jensen et al. (2009)[3]. In this case, please report the plasmid family (e.g., **Inc18, NT_Rep, Rep1, Rep2, Rep3, RepA_N, RepL, Rep_trans**) followed by the replicon name, separated by a comma. For example, "Rep_trans, rep14a." We recommend only using **PlasmidFinder** for Gram-positive bacteria, as other tools may use entirely different plasmid classification approaches.

4.2.4. AMR genes, mutations and predicted phenotype

Participants are required to submit AMR genes, mutations and predicted phenotype **for the relevant antimicrobial compounds, according to Table 3**, using the dropdown menu provided in the DTU Webtool. For AMR genes, please evaluate carefully multiple hits for variants of AMR genes that come from the same location in the genome. In such cases, the participants should select the variant to report based on the highest percent identity and percent coverage to the reference gene and submit the best quality hit. Submitting several variants from the same position in the genome will be marked as incorrect.

4.3. Deadline for submission of results

Submission is successful once you tick the 'final submit' option in the DTU Webtool (see webtool manual, Appendix 2). After selecting 'final submit', both the primary and secondary contact persons will receive an email with the submitted results attached. Results must be submitted electronically **no later than 9 December**

[1] Orlek A, Stoesser N, Anjum MF, Doumith M, Ellington MJ, Peto T, Crook D, Woodford N, Walker AS, Phan H, Sheppard AE. Plasmid Classification in an Era of Whole-Genome Sequencing: Application in Studies of Antibiotic Resistance Epidemiology. Front Microbiol. 2017 Feb 9;8:182. doi: 10.3389/fmicb.2017.00182. PMID: 28232822; PMCID: PMC5299020.

[2] Lanza VF, Tedim AP, Martínez JL, Baquero F, Coque TM. The Plasmidome of Firmicutes: Impact on the Emergence and the Spread of Resistance to Antimicrobials. Microbiol Spectr. 2015 Apr;3(2):PLAS-0039-2014. doi: 10.1128/microbiolspec.PLAS-0039-2014. PMID: 26104702.

[3] Jensen LB, Garcia-Migura L, Valenzuela AJ, Løhr M, Hasman H, Aarestrup FM. A classification system for plasmids from enterococci and other Gram-positive bacteria. J Microbiol Methods. 2010 Jan;80(1):25-43. doi: 10.1016/j.mimet.2009.10.012. Epub 2009 Oct 29. PMID: 19879906.

2024 at 16:00. Immediately after this deadline, the webtool will be closed for further edits and submissions. Late submission of results will not be accepted.

5. Evaluation of the submitted results

The six test strains of the DTU Genomic PT 2024 were sequenced using both Illumina and MinION platforms, allowing the genomes to be fully closed. Submitted results are evaluated compared to the reference closed genomes in two parts:

a) The QC of the submitted raw sequencing data and the assemblies

b) The bioinformatic analyses

Participating laboratories will receive an email from the DTU Genomic PT 2024 organizer once the evaluation is complete. Below are details on how each part is evaluated. The submitted information on the methods applied are not evaluated but are used as background information.

Attention! The evaluation will not indicate a pass/fail result. Instead, participants are encouraged to assess their own performance and consider whether any adjustments to their pipeline are necessary based on the results obtained.

5.1. Evaluation of the raw sequence data

The evaluation is based on the submitted sequence data (FASTQ files), which will be assembled using SPAdes (http://bioinf.spbau.ru/spades) and run through a QC pipeline by the PT organizers.

The output from the QC analysis is collected in two tables:

a) A **summarizing scoring table**, which provides an overall performance summary of each sample. It is based on the following criteria:

- Average coverage
- MLST
- Q-score of R1
- Q-score of R2
- Proportion of correct cgMLST genes identified
- Number of contigs
- N50
- Genomic coverage with a minimum depth of 10x
- Proportion of reads mapping to the reference genome
- Size of the assembled genome compared to the reference genome

b) A **QC parameter table**, which contains specific values from the QC analysis, including (but not limited to) those used for scoring.

5.2. Evaluation of the data from the bioinformatics analyses

For the bioinformatics analyses (identification of ST by MLST, plasmid replicons, AMR genes, mutations and predicted phenotype) an individual evaluation report will be generated for each laboratory. Upon login into the DTU Webtool, clicking on 'Download report' will give access to the evaluation report, which includes the obtained results, expected results and scores.

In the DTU Genomic PT 2024, the reported results for the bioinformatics analysis of each strain are evaluated individually, and each submitted result will receive a score of "1" or "0", according to **Table 5**. A submitted result that matches the list of expected results will receive a score of "1", while a mismatch (obtained result not expected) will be scored as "0". Expected results that are not submitted as obtained results will be listed in the evaluation report and also scored as "0".

Table 5. Scoring of bioinformatics analysis results in the DTU Genomic PT 2024.

Category	Obtained score
Submission of expected data	1
Submission of unexpected data	0
Not submitting expected data	0

6. Analysis and publication of results

A guidance document will be sent to the laboratories, summarizing the results of the DTU Genomic PT 2024 and providing the scientific background for the evaluation. The results may later be published in peer-reviewed journals. The authors and co-authors of the publications will include those who contributed to the preparation and execution of the PT. Due to the anonymity of the results, individual participating laboratories will not be acknowledged in the publications.

Individual results will be anonymized using laboratory codes, which are confidential and known only to the individual laboratory and the PT organizers. For laboratories associated with the Fleming Fund grant, SEQAFRICA, the complete list of laboratory codes is known to the project management team, Mott MacDonald, and the Fleming Fund. For laboratories participating as part of the EURL-AR network, the complete list of laboratory codes is known to the EU Commission.

7. Contact

If you have any questions or concerns, please do not hesitate to contact us.

DTU Genomic PT 2024 Coordinator:

Joana Mourão

National Food Institute, Technical University of Denmark
Kemitorvet, Building 204,
DK-2800 Kgs. Lyngby, DENMARK
E-mail: joam@food.dtu.dk