ISO 23418:2022-06 (E)

Microbiology of the food chain - Whole genome sequencing for typing and genomic characterization of bacteria - General requirements and guidance

Contents

Page

Foreword ..... **v**

Introduction ..... **vi**

1 Scope **1**

2 Normative references **1**

3 Terms and definitions **1**

4 Principle **6**
4.1 General 6
4.2 Laboratory operation: sample preparation and sequencing 6
4.3 Bioinformatics analysis 7
4.3.1 General 7
4.3.2 SNP analyses 7
4.3.3 MLST analyses 7
4.3.4 Kmer distance analysis 7
4.4 Metadata formats and sequence repository deposition 7
4.5 Validation and verification of WGS workflow 8

5 General laboratory guidance **8**
5.1 Bacterial isolation and DNA extraction 8
5.2 Laboratory environment 8
5.3 Standard operating procedures and nonconforming work 8
5.4 Laboratory information management system 8
5.5 Laboratory competence 8

6 Laboratory operations **9**
6.1 Sample preparation and storage 9
6.2 Bacterial isolates 9
6.3 DNA isolation 9
6.4 Library preparation and sequencing 9
6.4.1 Library preparation 9
6.4.2 DNA sequencing 10
6.4.3 Use of controls 10
6.4.4 Assessing raw read data quality 10
6.4.5 Sample and data storage and retention 10

7 Bioinformatic data analysis **11**
7.1 Requirements for software and bioinformatic pipelines used for data analysis 11
7.2 Logging and documentation 11
7.3 Quality assessments 11
7.4 SNP analyses 12
7.5 MLST analyses (cgMLST and wgMLST) 12
7.6 Target gene detection 13
7.7 Phylogenetic tree or dendrogram generation 13
7.8 Metrics and log files 13
7.9 Interpreting and reporting the results of bioinformatics analyses 13
7.9.1 Interpreting results from bioinformatics pipelines 13
7.9.2 Reporting genomic analysis results 14

8 Metadata **14**
8.1 General 14
8.2 Metadata interoperability and future-proofing 14

8.2.1 General 14
8.2.2 Ontologies 14
8.2.3 ISO WGS Slim 14
8.3 Formatting metadata using this document 15
8.4 Metadata associated with sample collection 15
8.5 Metadata associated with the isolate 16
8.6 Metadata associated with the sequence 17
9 Sequence repositories 19
10 Validation and verification 20
10.1 Validation 20
10.1.1 General 20
10.1.2 Validation of laboratory operations 21
10.1.3 Validation of the bioinformatics pipeline 21
10.1.4 Validation of the end-to-end workflow 22
10.2 Verification 22
10.2.1 General 22
10.2.2 Verification of laboratory operations 22
10.2.3 Verification of the bioinformatics pipeline 22
Annex A (informative) **Development of quality metrics and use of controls 24**
Annex B (informative) **Laboratory contact information fields 29**
Annex C (informative) **Geographic location of sample collection fields 31**
Annex D (informative) **Isolate passage history fields 32**
Annex E (informative) **Antibiogram results and methods fields 33**
Annex F (informative) **Virulence factor detection and methods fields 35**
Annex G (informative) **Sequence quality control metrics 36**
Annex H (informative) **Metadata specification 37**
Annex I (informative) **Instructions for ontology slim integration by software developers 40**
Bibliography 44