

DIN EN ISO 23418



ICS 07.100.30

**Mikrobiologie der Lebensmittelkette –
Gesamtgenomsequenzierung zur Typisierung und genomischen
Charakterisierung von Bakterien –
Allgemeine Anforderungen und Leitfaden (ISO 23418:2022);
Deutsche Fassung EN ISO 23418:2022**

Microbiology of the food chain –
Whole genome sequencing for typing and genomic characterization of bacteria –
General requirements and guidance (ISO 23418:2022);
German version EN ISO 23418:2022

Microbiologie de la chaîne alimentaire –
Séquençage de génome entier pour le typage et la caractérisation génomique des bactéries –
Exigences générales et recommandations (ISO 23418:2022);
Version allemande EN ISO 23418:2022

Gesamtumfang 58 Seiten

DIN-Normenausschuss Lebensmittel und landwirtschaftliche Produkte (NAL)

Nationales Vorwort

Dieses Dokument (EN ISO 23418:2022) wurde vom Technischen Komitee ISO/TC 34 „Food products“ in Zusammenarbeit mit dem Technischen Komitee CEN/TC 463 „Mikrobiologie der Lebensmittelkette“ erarbeitet, dessen Sekretariat von AFNOR (Frankreich) gehalten wird.

Das zuständige deutsche Normungsgremium ist der Arbeitsausschuss NA 057-01-06 AA „Mikrobiologie der Lebensmittelkette“ im DIN-Normenausschuss Lebensmittel und landwirtschaftliche Produkte (NAL).

Aktuelle Informationen zu diesem Dokument können über die Internetseiten von DIN (www.din.de) durch eine Suche nach der Dokumentennummer aufgerufen werden.

Für die in diesem Dokument zitierten Dokumente wird im Folgenden auf die entsprechenden deutschen Dokumente hingewiesen:

ISO 8601-1:2019	siehe	DIN ISO 8601-1:2020-12
ISO 11238:2018	siehe	DIN EN ISO 11238:2018-11
ISO 16140-1:2016	siehe	DIN EN ISO 16140-1:2016-11
ISO 16140-3:2021	siehe	DIN EN ISO 16140-3:2021-06

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN EN ISO 11238:2018-11, *Medizinische Informatik — Identifikation von Arzneimitteln — Datenelemente und Strukturen zur eindeutigen Identifikation und zum Austausch von vorgeschriebenen Informationen zu Stoffen (ISO 11238:2018); Deutsche Fassung EN ISO 11238:2018*

DIN EN ISO 16140-1:2016-11, *Mikrobiologie der Lebensmittelkette — Verfahrensvalidierung — Teil 1: Terminologie (ISO 16140-1:2016); Deutsche Fassung EN ISO 16140-1:2016*

DIN EN ISO 16140-3:2021-06, *Mikrobiologie der Lebensmittelkette — Verfahrensvalidierung — Teil 3: Arbeitsvorschrift für die Verifizierung von Referenz- und validierten alternativen Verfahren in einem Einzel-Labor (ISO 16140-3:2021); Deutsche Fassung EN ISO 16140-3:2021*

DIN ISO 8601-1:2020-12, *Datum und Uhrzeit — Darstellung für den Informationsaustausch — Teil 1: Grundlegende Regeln (ISO 8601-1:2019)*

Deutsche Fassung

**Mikrobiologie der Lebensmittelkette —
Gesamtgenomsequenzierung zur Typisierung und
genomischen Charakterisierung von Bakterien —
Allgemeine Anforderungen und Leitfaden (ISO 23418:2022)**

Microbiology of the food chain —
Whole genome sequencing for typing and genomic
characterization of bacteria —
General requirements and guidance (ISO 23418:2022)

Microbiologie de la chaîne alimentaire —
Séquençage de génome entier pour le typage et la
caractérisation génomique des bactéries —
Exigences générales et recommandations
(ISO 23418:2022)

Diese Europäische Norm wurde vom CEN am 20. Mai 2022 angenommen.

Die CEN-Mitglieder sind gehalten, die CEN/CENELEC-Geschäftsordnung zu erfüllen, in der die Bedingungen festgelegt sind, unter denen dieser Europäischen Norm ohne jede Änderung der Status einer nationalen Norm zu geben ist. Auf dem letzten Stand befindliche Listen dieser nationalen Normen mit ihren bibliographischen Angaben sind beim CEN-CENELEC-Management-Zentrum oder bei jedem CEN-Mitglied auf Anfrage erhältlich.

Diese Europäische Norm besteht in drei offiziellen Fassungen (Deutsch, Englisch, Französisch). Eine Fassung in einer anderen Sprache, die von einem CEN-Mitglied in eigener Verantwortung durch Übersetzung in seine Landessprache gemacht und dem Management-Zentrum mitgeteilt worden ist, hat den gleichen Status wie die offiziellen Fassungen.

CEN-Mitglieder sind die nationalen Normungsinstitute von Belgien, Bulgarien, Dänemark, Deutschland, Estland, Finnland, Frankreich, Griechenland, Irland, Island, Italien, Kroatien, Lettland, Litauen, Luxemburg, Malta, den Niederlanden, Norwegen, Österreich, Polen, Portugal, der Republik Nordmazedonien, Rumänien, Schweden, der Schweiz, Serbien, der Slowakei, Slowenien, Spanien, der Tschechischen Republik, der Türkei, Ungarn, dem Vereinigten Königreich und Zypern.



EUROPÄISCHES KOMITEE FÜR NORMUNG
EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION

CEN-CENELEC Management-Zentrum: Rue de la Science 23, B-1040 Brüssel

Inhalt

	Seite
Europäisches Vorwort	4
Vorwort	5
Einleitung	6
1 Anwendungsbereich	7
2 Normative Verweisungen	7
3 Begriffe	7
4 Kurzbeschreibung	13
4.1 Allgemeines	13
4.2 Laborbetrieb: Probenvorbereitung und Sequenzierung	13
4.3 Bioinformatische Analyse	14
4.3.1 Allgemeines	14
4.3.2 SNP-Analysen	14
4.3.3 MLST-Analysen	14
4.3.4 Analyse des k-mer-Abstandes	14
4.4 Metadatenformate und Hinterlegung von Sequenzdaten in Repositorien	14
4.5 Validierung und Verifizierung des WGS-Workflows	15
5 Allgemeiner Laborleitfaden	15
5.1 Bakterienisolierung und DNA-Extraktion	15
5.2 Laborumgebung	15
5.3 Standardverfahrensanweisungen und nicht konforme Arbeit	15
5.4 Labor-Informations- und Management-System	15
5.5 Kompetenz im Labor	15
6 Laborbetrieb	16
6.1 Herstellung und Lagerung der Proben	16
6.2 Bakterienisolate	16
6.3 Isolierung der DNA	16
6.4 Vorbereitung der Bibliothek und Sequenzierung	17
6.4.1 Vorbereitung der Bibliothek	17
6.4.2 DNA-Sequenzierung	17
6.4.3 Verwendung von Kontrollen	17
6.4.4 Beurteilung der Qualität von Read-Rohdaten	18
6.4.5 Speicherung und Aufbewahrung von Proben und Daten	18
7 Bioinformatische Datenanalyse	18
7.1 Anforderungen an Software und bioinformatische Pipelines, die für die Datenanalyse verwendet werden	18
7.2 Protokollierung und Dokumentation	19
7.3 Qualitätsbeurteilungen	19
7.4 SNP-Analysen	20
7.5 MLST-Analysen (cgMLST und wgMLST)	20
7.6 Detektion von Zielgenen	21
7.7 Phylogenetische Baum- oder Dendrogramm-Generierung	21
7.8 Metriken und Protokolldateien	21
7.9 Interpretation und Meldung der Ergebnisse bioinformatischer Analysen	21
7.9.1 Interpretation der Ergebnisse aus bioinformatischen Pipelines	21

7.9.2	Angabe der Ergebnisse der Genomanalyse.....	22
8	Metadaten	22
8.1	Allgemeines	22
8.2	Interoperabilität und Zukunftssicherheit von Metadaten	22
8.2.1	Allgemeines	22
8.2.2	Ontologien.....	23
8.2.3	ISO WGS Slim	23
8.3	Formatierung der Metadaten unter Verwendung dieses Dokuments	23
8.4	Metadaten im Zusammenhang mit der Probenahme.....	23
8.5	Metadaten im Zusammenhang mit dem Isolat	25
8.6	Metadaten im Zusammenhang mit der Sequenz.....	26
9	Sequenz-Repositorien.....	29
10	Validierung und Verifizierung.....	29
10.1	Validierung	29
10.1.1	Allgemeines	29
10.1.2	Validierung des Laborbetriebs.....	30
10.1.3	Validierung der bioinformatischen Pipeline	30
10.1.4	Validierung des Gesamt-Workflows.....	32
10.2	Verifizierung	32
10.2.1	Allgemeines	32
10.2.2	Verifizierung des Laborbetriebs.....	32
10.2.3	Verifizierung der bioinformatischen Pipeline.....	32
	Anhang A (informativ) Entwicklung von Qualitätsmetriken und Verwendung von Kontrollen.....	33
	Anhang B (informativ) Felder zur Dokumentation der Labor-Kontaktinformationen.....	39
	Anhang C (informativ) Felder zur Dokumentation des geographischen Standortes der Probenahme	41
	Anhang D (informativ) Felder zur Dokumentation der Isolat-Passage Geschichte.....	42
	Anhang E (informativ) Felder zur Dokumentation des Antibiotogramms und assoziierter Methoden.....	43
	Anhang F (informativ) Felder zur Dokumentation von Virulenzfaktoren und assoziierter Methoden.....	45
	Anhang G (informativ) Metriken zur Qualitätskontrolle der Sequenz.....	46
	Anhang H (informativ) Spezifikation der Metadaten.....	47
	Anhang I (informativ) Anweisungen für die Integration der Ontology Slim durch Softwareentwickler	50
	Literaturhinweise.....	55

Europäisches Vorwort

Dieses Dokument (EN ISO 23418:2022) wurde vom Technischen Komitee ISO/TC 34 „Food products“ in Zusammenarbeit mit dem Technischen Komitee CEN/TC 463 „Mikrobiologie der Lebensmittelkette“ erarbeitet, dessen Sekretariat von AFNOR gehalten wird.

Diese Europäische Norm muss den Status einer nationalen Norm erhalten, entweder durch Veröffentlichung eines identischen Textes oder durch Anerkennung bis Dezember 2022, und etwaige entgegenstehende nationale Normen müssen bis Dezember 2022 zurückgezogen werden.

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. CEN ist nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren.

Dieses Dokument wurde im Rahmen eines Normungsauftrages erarbeitet, den die Europäische Kommission und die Europäische Freihandelsassoziation CEN erteilt haben.

Rückmeldungen oder Fragen zu diesem Dokument sollten an das jeweilige nationale Normungsinstitut des Anwenders gerichtet werden. Eine vollständige Liste dieser Institute ist auf den Internetseiten von CEN abrufbar.

Entsprechend der CEN-CENELEC-Geschäftsordnung sind die nationalen Normungsinstitute der folgenden Länder gehalten, diese Europäische Norm zu übernehmen: Belgien, Bulgarien, Dänemark, Deutschland, die Republik Nordmazedonien, Estland, Finnland, Frankreich, Griechenland, Irland, Island, Italien, Kroatien, Lettland, Litauen, Luxemburg, Malta, Niederlande, Norwegen, Österreich, Polen, Portugal, Rumänien, Schweden, Schweiz, Serbien, Slowakei, Slowenien, Spanien, Tschechische Republik, Türkei, Ungarn, Vereinigtes Königreich und Zypern.

Anerkennungsnotiz

Der Text von ISO 23418:2022 wurde von CEN als EN ISO 23418:2022 ohne irgendeine Abänderung genehmigt.

Vorwort

ISO (die Internationale Organisation für Normung) ist eine weltweite Vereinigung nationaler Normungsinstitute (ISO-Mitgliedsorganisationen). Die Erstellung von Internationalen Normen wird üblicherweise von Technischen Komitees von ISO durchgeführt. Jede Mitgliedsorganisation, die Interesse an einem Thema hat, für welches ein Technisches Komitee gegründet wurde, hat das Recht, in diesem Komitee vertreten zu sein. Internationale staatliche und nichtstaatliche Organisationen, die in engem Kontakt mit ISO stehen, nehmen ebenfalls an der Arbeit teil. ISO arbeitet bei allen elektrotechnischen Normungsthemen eng mit der Internationalen Elektrotechnischen Kommission (IEC) zusammen.

Die Verfahren, die bei der Entwicklung dieses Dokuments angewendet wurden und die für die weitere Pflege vorgesehen sind, werden in den ISO/IEC-Direktiven, Teil 1 beschrieben. Es sollten insbesondere die unterschiedlichen Annahmekriterien für die verschiedenen ISO-Dokumentenarten beachtet werden. Dieses Dokument wurde in Übereinstimmung mit den Gestaltungsregeln der ISO/IEC-Direktiven, Teil 2 erarbeitet (siehe www.iso.org/directives).

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. ISO ist nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren. Details zu allen während der Entwicklung des Dokuments identifizierten Patentrechten finden sich in der Einleitung und/oder in der ISO-Liste der erhaltenen Patenterklärungen (siehe www.iso.org/patents).

Jeder in diesem Dokument verwendete Handelsname dient nur zur Unterrichtung der Anwender und bedeutet keine Anerkennung.

Für eine Erläuterung des freiwilligen Charakters von Normen, der Bedeutung ISO-spezifischer Begriffe und Ausdrücke in Bezug auf Konformitätsbewertungen sowie Informationen darüber, wie ISO die Grundsätze der Welthandelsorganisation (WTO, en: World Trade Organization) hinsichtlich technischer Handelshemmnisse (TBT, en: Technical Barriers to Trade) berücksichtigt, siehe www.iso.org/iso/foreword.html.

Dieses Dokument wurde vom Technischen Komitee ISO/TC 34, *Food products*, Unterkomitee SC 9, *Microbiology*, in Zusammenarbeit mit dem Europäischen Komitee für Normung (CEN), Technisches Komitee CEN/TC 463, *Mikrobiologie der Lebensmittelkette*, in Übereinstimmung mit der Vereinbarung zur technischen Zusammenarbeit zwischen ISO und CEN (Wiener Vereinbarung) erarbeitet.

Rückmeldungen oder Fragen zu diesem Dokument sollten an das jeweilige nationale Normungsinstitut des Anwenders gerichtet werden. Eine vollständige Auflistung dieser Institute ist unter www.iso.org/members.html zu finden.

Einleitung

Next generation sequencing (NGS) bietet schnellen und kostengünstigen Zugang mit hohem Durchsatz zu mikrobiellen Gesamtgenomsequenzen im Hochdurchsatzverfahren und wird bei einer wachsenden Zahl von Problemen in der Lebensmittelmikrobiologie angewandt. Gesamtgenomsequenzen ermöglichen eine Repräsentation des biologischen Potentials eines sequenzierten Organismus in Einzelbasenauflösung. Gesamtgenomsequenzierung (WGS, en: whole genome sequencing) bietet bei vielen Anwendungsmöglichkeiten erhebliche Vorteile gegenüber bestehenden Technologien (z. B. Serotypisierung, Pulsfeld-Gelelektrophorese, phänotypische Antibiotikaresistenz). WGS-basierte Analysen werden von Laboren für öffentliche Gesundheit verwendet, um Ausbrüche zu erkennen und Mutationen, Gene und andere genetische Merkmale zur Charakterisierung der Virulenz und des Überlebenspotentials nachzuweisen. Innerhalb der Lebensmittelindustrie besteht Interesse daran, Gesamtgenomsequenzen einzusetzen, um bakterielle Isolate aus Lebensmittelbestandteilen und Oberflächen zu charakterisieren, ihre Herkunft und Ökologie besser zu verstehen und bestehende Verfahren zur Risikominimierung zu aktualisieren. Einige Unternehmen haben die Kapazitäten zur Erfassung und Analyse von Daten der Gesamtgenomsequenzierung geschaffen oder schaffen diese gerade. Andere wenden sich bezüglich dieser Dienstleistungen an Drittlabore, wie sie es bereits für andere mikrobiologische Analysen getan haben.

Dieses Dokument dient als Leitfaden für die laborlichen und bioinformatischen Aspekte der Gesamtgenomsequenzierung, sowie für die Handhabung der assoziierten Metadaten zu bakteriellen Mikroorganismen in Lebensmitteln, die entlang der gesamten Lebensmittelkette (z. B. Inhaltsstoffe, Lebensmittel, Futtermittel, Produktionsumgebung) beprobt werden. Obwohl die Mikrobiologie der Lebensmittelkette auch Viren und Pilze umfasst, bezieht sich dieses Dokument nur auf Bakterien. Dieses Dokument soll auf alle derzeit verfügbaren Technologien des Next generation sequencing anwendbar sein. Es darf zur Analyse von Daten der Gesamtgenomsequenzierung mit proprietärer, Open Source- oder kundenspezifischer Software eingesetzt werden. Es ist nicht zur Festlegung der Sequenzierungsschemikalien, analytischen Verfahren oder Software vorgesehen. Dieses Dokument definiert Labor-, Daten- und Metadaten-Verwaltungspraktiken, um sicherzustellen, dass Analysen eindeutig angegeben werden, transparent und für Anfragen zugänglich sind. Dieses Dokument soll Labore in die Lage versetzen, ihre Managementsysteme für Qualität und technische Abläufe zu entwickeln. Kunden von Laboren und Aufsichtsbehörden können es auch zur Anerkennung und Bestätigung der Kompetenz von Laboren nutzen. Dieses Dokument kann auch in anderen Bereichen angewendet werden (z. B. Umwelt, Gesundheit von Menschen, Gesundheit von Tieren).

1 Anwendungsbereich

Dieses Dokument legt die Mindestanforderungen für die Generierung und Analyse von Daten der Gesamtgenomsequenzierung (WGS) für Bakterien, die aus der Lebensmittelkette gewonnen wurden, fest. Dieser Prozess kann die folgenden Phasen umfassen:

- a) Umgang mit Bakterienkulturen;
- b) keimfreie Isolierung der genomischen DNA;
- c) Vorbereitung der genomischen DNA-Bibliothek, Sequenzierung und Beurteilung der Read-Qualität von Rohdaten und deren Speicherung;
- d) bioinformatische Analyse zur Bestimmung der genetischen Verwandtschaft, des genetischen Gehalts und zur Vorhersage des Phänotyps sowie Validierung der bioinformatischen Pipeline;
- e) die Erfassung von Metadaten und die Hinterlegung von Sequenzdaten in Repositorien;
- f) Validierung des Gesamt-WGS-Workflows (gebrauchstauglich für die beabsichtigte Anwendung).

Dieses Dokument ist anwendbar auf Bakterien, die isoliert wurden aus:

- Produkten, die für den menschlichen Verzehr vorgesehen sind;
- Produkten, die als Futtermittel vorgesehen sind;
- Umgebungsproben aus Bereichen, in denen Lebensmittel und Futtermittel gehandhabt und hergestellt werden;
- Proben aus dem Bereich der Primärproduktion.

WARNHINWEIS — Zum Schutz der Gesundheit des Laborpersonals ist es unerlässlich, dass der Umgang mit Bakterienkulturen nur in Laboratorien mit geeigneter Ausstattung und unter der Leitung eines qualifizierten Mikrobiologen erfolgt und dass bei der Entsorgung allen inkubierten Materials mit äußerster Vorsicht vorgegangen wird. Personen, die dieses Dokument anwenden, sollten mit der üblichen Laborpraxis vertraut sein. Dieses Dokument erhebt nicht den Anspruch, alle gegebenenfalls zutreffenden Sicherheitsaspekte im Zusammenhang mit seiner Anwendung zu behandeln. Es liegt in der Verantwortung des Anwenders, angemessene Sicherheits- und Schutzmaßnahmen zu treffen.

2 Normative Verweisungen

Es gibt keine normativen Verweisungen in diesem Dokument.

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

ISO und IEC stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

- ISO Online Browsing Platform: verfügbar unter <https://www.iso.org/obp>
- IEC Electropedia: verfügbar unter <https://www.electropedia.org/>

3.1

Adaptersequenz

DNA mit einer bekannten Sequenz, die an das Ende eines DNA-Bibliotheksfragments angefügt wird, um den Sequenzierungsprozess zu erleichtern (z. B. Annealing an eine Fließzelle)

3.2

Annotation

Prozess der Identifizierung von Genen und anderen Merkmalen der Genom-*Assemblierungen* (3.4)

3.3

Antibiogramm

Zusammenfassung der Ergebnisse antimikrobieller Empfindlichkeitsprüfungen, die für einen spezifischen Mikroorganismus durchgeführt wurden, üblicherweise in tabellarischer Form

3.4

Assemblierung

Ergebnis aus dem Prozess des Alignments und der Zusammenführung von Sequenzabschnitten (*Reads* (3.38)) zu größeren zusammenhängenden Sequenzen (*Contigs* (3.10))

3.5

Basenaufruf

Prozess, bei dem Positionen in den Sequenzabschnitten (*Reads* (3.38)) Nukleotide und Qualitätswerte zugewiesen werden

3.6

Bioinformatik

Erfassung, Speicherung und Analyse biologischer Daten einschließlich Sequenzen

3.7

bioinformatische Pipeline

einzelne Programme, Skripte oder miteinander verknüpfte Softwarekomponenten, wobei die Ausgabe eines Programms als Eingabe für den nächsten Schritt der Datenverarbeitung verwendet wird

3.8

carry-over-Kontamination

Probenkontamination im Zusammenhang mit vorhergehenden Experimenten, die auf die aktuelle Analyse übertragen wurde (z. B. carry-over-Kontamination durch Amplifikationsprodukte aus vorhergehenden Polymerase-Kettenreaktion-(PCR-)Experimenten auf die aktuelle PCR-Analyse oder carry-over-Kontamination von bereits sequenzierten Proben von einem Sequenzierungslauf auf den anderen)

3.9

Chemical Entities of Biological Interest Ontology

ChEBI

Ontologie (3.35) zur Beschreibung kleiner chemischer Verbindungen

3.10

Contig

zusammenhängendes Stück einer DNA-Sequenz, die sich aus der *Assemblierung* (3.4) kleinerer, überlappender DNA-*Reads* (3.38) ergibt

3.11

kontrolliertes Vokabular

en **controlled vocabulary**

endlicher Vorrat von Werten, die die einzig zulässigen Werte für ein Datenelement darstellen

[QUELLE: ISO 11238:2018, 3.18, modifiziert — Anmerkung 1 zum Begriff wurde gelöscht.]

3.12

Abdeckung

Häufigkeit, mit der eine bestimmte Basenposition in einem Sequenzierungslauf gelesen wird

Anmerkung 1 zum Begriff: Die Anzahl an *Reads* (3.38), die eine bestimmte Position abdecken.

[QUELLE: ISO 20397-2:2021, 3.6, modifiziert — Die zugelassene Benennung „coverage depth“ wurde gelöscht.]

3.13

Kreuzkontamination

Kontamination einer Probe (Bakterien*isolat* (3.23) oder DNA) mit anderen Proben während der Vorbereitung eines Sequenzierungslaufs

3.14

DNA-Probe

Teil einer DNA, die aus der verarbeiteten Probe extrahiert wird

3.15

Draft-Assemblierung

de novo-Assemblierung (3.4) von Genomen, bestehend aus *Contigs* (3.10) ohne implizite Reihenfolge, die üblicherweise mit Hilfe der Gesamtgenom-shotgun-Sequenzierung mit einer short-read-Technologie erzeugt werden

3.16

Umweltontologie

EnvO, en: Environment Ontology

Ontologie (3.35) zur Beschreibung von Umweltmerkmalen und Lebensräumen

3.17

FoodEx2-Ontologie

FoodEx2

genormtes System zur Klassifizierung und Beschreibung von Lebensmitteln, entwickelt von der Europäischen Behörde für Lebensmittelsicherheit (EFSA)

3.18

Lebensmittel-Ontologie

FoodOn, en: Food Ontology

Ontologie (3.35) zur Beschreibung von Lebensmittelprodukten, Tierfutter und Lebensmittelverarbeitung

3.19

Gazetteer-Ontologie

GAZ

Ontologie (3.35) zur Beschreibung geographischer Orte

3.20

Index

Oligonukleotidsequenzen, die im Prozess der Bibliotheksvorbereitung verwendet werden, um DNA aus spezifischen Proben zu markieren oder mit einem Barcode zu versehen, so dass mehrere Proben (d. h. mehrere *Bibliotheken* (3.25)) in einem Bibliothekenpool kombiniert (gemultiplext) und in einer einzigen Sequenzierungsreaktion analysiert werden können

3.21

International Nucleotide Sequence Database Collaboration

INSDC

Initiative, die von der DNA-Datenbank Japans (DDBJ), dem Europäischen Laboratorium für Molekularbiologie, Europäischen Institut für Bioinformatik (EMBL-EBI) und dem National Center for Biotechnology Information (NCBI) betrieben wird

3.22

International Organization for Standardization whole genome sequencing slim

ISO WGS Slim

Ontology Slim (*Ontologie*, 3.35) mit interoperablen Feldern und Begriffen bezüglich der Verwendung von WGS (3.49) für die Mikrobiologie der Lebensmittelkette

3.23

Isolat

Population von Bakterienzellen in Reinkultur, die aus einem einzigen *Stamm* (3.45) stammen

3.24

k-mer

mögliche Sequenz der Länge k , die in einer Gesamtgenomsequenz enthalten ist

3.25

Bibliothek

Sammlung von genomischen DNA-Fragmenten aus einem einzigen *Isolat* (3.23) zur Bestimmung der Genomsequenz(en)

Anmerkung 1 zum Begriff: Eine Sammlung von Bibliotheken, von denen jede aus einem einzigen Isolat stammt, wird als „Bibliothekenpool“ bezeichnet und zur Analyse in ein Sequenziergerät gegeben. Dieses Multiplexen von Bibliotheken würde auch für ein einzelnes Isolat das Ergebnis liefern, vorausgesetzt es werden bei der Vorbereitung der Bibliothek eines jeden einzelnen Isolats eindeutige Indizes verwendet.

Anmerkung 2 zum Begriff: Es kann auch eine Bibliothek gemischter DNA, d. h. aus einer Mischung verschiedener Spezies stammender DNA, erstellt werden. Dies fällt jedoch nicht in den Anwendungsbereich dieses Dokuments, da es dabei um metagenomische Sequenzierung geht.

3.26

Managementsystem

Qualitäts-, Verwaltungs- und technische Systeme, die den Betrieb einer Organisation regeln

Anmerkung 1 zum Begriff: In diesem Dokument bezieht sich „Organisation“ auf das Labor.

3.27

Mapping

Verwendung von Software für das Alignment von Sequenzabschnitten (*Reads* (3.38)) mit Referenzsequenzen

3.28

Metadaten

Daten, die andere Daten definieren und beschreiben

[QUELLE: ISO/IEC 11179-1:2015, 3.2.16]

3.29

minimale Daten zum Abgleich

MDM, en: minimal data for matching

Informationen, die zur Beschreibung der Probenquelle und -herkunft einer genomischen Sequenz erforderlich sind, wie durch den Global Microbial Identifier [1] definiert und durch die *International Nucleotide Sequence Database Collaboration* (3.21) implementiert

3.30

Multilocus-Sequenztypisierung

MLST, en: multi-locus sequence typing

Verfahren der Genomanalyse, bei der Nukleotidvarianten innerhalb vordefinierter Sätze von Loci identifiziert werden

Anmerkung 1 zum Begriff: Das Verfahren wurde ursprünglich bei sieben Loci angewandt, mittlerweile auch bei entweder Kerngenomloci für cgMLST oder Gesamtgenom-Loci für wgMLST.

3.31

N50

Länge (N), so dass *Contigs* (3.10) von N oder länger die Hälfte der Basen in der *Assemblierung* (3.4) umfassen

3.32

NCBITaxon

automatische Übersetzung der Taxonomiedatenbank des National Center for Biotechnology Information (NCBI) in OBO/OWL

3.33

NG50

Länge (N) der DNA, so dass *Contigs* (3.10) von N oder länger die Hälfte der Basen im Genom umfassen

3.34

Open Biological and Biomedical Ontology Foundry

OBO Foundry

Sammlung von *Ontologien* (3.35), die von einer Gemeinschaft von Ontologie-Entwicklern erstellt wurden, die sich zur Zusammenarbeit und zur Einhaltung gemeinsamer Grundsätze verpflichtet haben

3.35

Ontologie

kontrolliertes Vokabular (3.11), das in einer Hierarchie angeordnet ist, wobei die Begriffe durch logische Beziehungen miteinander verbunden sind

3.36

Ontology Slim

Reihe von Ontologiefeldern und -begriffen, die als Teil einer bestimmten Sammlung kommentiert werden, oft zu einem bestimmten Zweck, und die extrahiert werden dürfen, um eine Datei zu erstellen, die sich von der ursprünglichen *Ontologie* (3.35) unterscheidet

3.37

Phred-Qualitätswert der Sequenz

Q

Maß für die Wahrscheinlichkeit (*P*), dass eine Base an einer bestimmten Position der Sequenz falsch zugeordnet wird, ausgedrückt durch:

$$Q = -10 \lg P$$

Anmerkung 1 zum Begriff: Ein Wert von Q30 zeigt an, dass eine Wahrscheinlichkeit von 1 zu 1 000 besteht, dass eine Base falsch zugeordnet ist (d. h. der Basenaufbau ist zu 99,9 % richtig).

3.38

Read

aus einem DNA- oder RNA-Fragment abgeleitete Nukleotidsequenz

3.39

Sequenz-Repository

Datenbank, in der Datensätze aus der *Gesamtgenomsequenzierung* (3.49) gespeichert und verwaltet werden

Anmerkung 1 zum Begriff: Ein öffentliches Repository ermöglicht den uneingeschränkten Zugang zu den Daten, während ein privates Repository oder eines im Verbund den Zugang zu den Daten einschränkt.

3.40

Sequenzierungsreplik

<biologisch> Sequenzierung einer anderen Kolonie aus demselben *Isolat* (3.23), das aus demselben Probenmaterial gewonnen wurde, um die biologische Variation zu bewerten

3.41

Sequenzierungsreplik

<technisch> Resequenzierung derselben biologischen Probe oder *Bibliothek* (3.25) zur Bewertung der Sequenzvariation aufgrund der Instrumentierung und des Protokolls

3.42

Serotyp

Klassifizierungsschema auf der Grundlage des Antigenproteinnachweises oder des sequenzbasierten Nachweises von Genen, die bakterielle Oberflächenmoleküle kodieren

3.43

Einzelnukleotid-Polymorphismus

SNP, en: single nucleotide polymorphism

Einzelnukleotid-Variante (3.44), die eine bestimmte Qualitäts- oder Frequenzschwelle überschreitet

3.44

Einzelnukleotid-Variante

SNV, en: single nucleotide variant

Unterschiede zwischen den Nukleotiden an der gleichen genomischen Position von zwei oder mehr *Isolaten* (3.23)

3.45

Stamm

Nachkommen einer einzelnen Isolierung in Reinkultur, die in der Regel von einer einzigen Ausgangskolonie auf einem festen Wachstumsmedium abstammen

Anmerkung 1 zum Begriff: Ein Stamm darf als ein *Isolat* (3.23) oder eine Gruppe von Isolaten angesehen werden, die sich von anderen Isolaten derselben Gattung und Art durch phänotypische und genotypische Merkmale unterscheiden lassen können.

Anmerkung 2 zum Begriff: Siehe Literaturhinweis [2].

3.46

Validierung

Ermittlung der Leistungskenngrößen eines Verfahrens und Bereitstellung eines objektiven Nachweises, dass die Leistungsanforderungen für eine festgelegte vorgesehene Anwendung erfüllt sind

[QUELLE: ISO 16140-1:2016, 2.81]

3.47

validierte Dateneingabe

automatisierter Prozess, der sicherstellt, dass die in ein Repository eingegebenen Daten korrekt sind

3.48

Verifizierung

Nachweis, dass ein validiertes Verfahren bei der Anwendung durch den Anwender die in der Validierungsuntersuchung bestimmten Verfahrensspezifikationen erfüllt und für den vorgesehenen Zweck geeignet ist

[QUELLE: ISO 16140-3:2021, 3.21, modifiziert — Anmerkung 1 zum Begriff wurde gelöscht.]

3.49

Gesamtgenomsequenzierung

WGS, en: whole genome sequencing

Prozess zur Bestimmung der DNA-Sequenz des Genoms eines Organismus unter Verwendung der gesamten genomischen DNA als Eingabe

4 Kurzbeschreibung

4.1 Allgemeines

WGS-Analysen von Bakterien in der Lebensmittel- und Futtermittelkette bestehen aus einer Kultivierung des reinen Bakterienisolats, einer DNA-Isolierung in einem mikrobiologischen Labor, Sequenzierungsschritten in einer geeigneten Sequenzierungsumgebung und bioinformatischen Analysen, die in einer speziellen Computerumgebung durchgeführt werden.

ANMERKUNG Bei dem mikrobiologischen Labor, der Sequenzierungseinrichtung und der bioinformatischen Einrichtung kann es sich um die gleiche Organisation handeln.

4.2 Laborbetrieb: Probenvorbereitung und Sequenzierung

Die Probenvorbereitung und Sequenzierung sollten die folgenden Schritte umfassen:

- a) Informationen über die zu sequenzierenden Isolate, einschließlich Barcodes für Multiplex-Proben, werden in die entsprechenden Aufzeichnungssysteme, wie z. B. ein Laborinformationsmanagementsystem (LIMS) oder Arbeitsblätter zur Probenbeschreibung oder beides, eingegeben.
- b) Reine Isolate (von denen mindestens die Gattung und idealerweise auch die Art bekannt ist) werden kultiviert und die genomische DNA wird extrahiert.
- c) Aus der qualitätsgeprüften genomischen DNA (siehe Tabelle A.1 für Leitlinien zu DNA-Quantitäts- und Qualitätsmetriken) werden DNA-Sequenzierungsbibliotheken hergestellt. Dieser Prozess sollte Folgendes umfassen:
 - 1) DNA-Fragmentierung, falls für die angewandte Sequenzierungstechnologie erforderlich;
 - 2) Ligation von Indizes und Adaptern, in Übereinstimmung mit den Protokollen der angewandten Sequenzierungstechnologie;
 - 3) Quantifizierung, Normalisierung und Qualitätskontrolle der resultierenden Bibliothek;
 - 4) Pooling von Bibliotheken im Falle von Multiplex-Sequenzierungsläufen.
- d) Bibliotheken (d. h. Bibliothekenpools) werden sequenziert.
- e) Die vom Sequenzierungsinstrument erzeugten Qualitätsmetriken werden idealerweise für jeden Lauf aufgezeichnet, um eine Überwachung der Leistung zu ermöglichen.

4.3 Bioinformatische Analyse

4.3.1 Allgemeines

Pipelines für bioinformatische Analysen dürfen sich auf *In-silico*-Vorhersagen des Phänotyps (z. B. Virulenz) oder auf den Nachweis von Clustern genetisch ähnlicher Isolate (d. h. gleicher Stamm, Sequenztyp oder Serotyp) konzentrieren. Pipelines, die auf vergleichenden Ansätzen basieren, dürfen verwendet werden, um das Vorhandensein und die Zustände von Markern in Rohdaten und assemblierten Sequenzen zu erkennen, um *In-silico*-Stamm- (z. B. Sequenztyp) und Phänotypvorhersagen zu machen.

Sequenzdaten für mehrere Isolate dürfen mit SNP-, MLST- oder k-mer-Abstands-Analyseverfahren analysiert werden, um Cluster eng verwandter Bakterien zu identifizieren. Die Ergebnisse dieser Analysen dürfen verwendet werden, um auf Beziehungen zwischen Isolaten zu schließen, die mit phylogenetischen Bäumen und Dendrogrammen veranschaulicht werden dürfen.

4.3.2 SNP-Analysen

Bei SNP-Analysen werden Reads auf eine Referenzsequenz gemapped oder Reads zu Contigs assembliert, die verglichen werden. Zur Bestimmung von SNPs werden SNVs qualitätsgefiltert, um SNP-Positionen zu identifizieren.

4.3.3 MLST-Analysen

Für MLST-Analysen werden die Reads assembliert oder gemapped. Allele werden identifiziert, qualitätsgefiltert und mit einer cgMLST- oder wgMLST-Datenbank verglichen.

4.3.4 Analyse des k-mer-Abstandes

Sequenzdaten für mehrere Isolate dürfen mit k-mer-Abstandsverfahren analysiert werden, um Cluster verwandter Bakterien zu identifizieren. K-mer-Analysen haben den Vorteil, dass sie sehr schnell sind, weisen aber einige Einschränkungen auf, insbesondere in Bezug auf die Präzision (d. h. sie sind bei der Artbestimmung anwendbar, werden aber nicht für detaillierte Quellenrückverfolgung von eng verwandten Stämmen empfohlen).

4.4 Metadatenformate und Hinterlegung von Sequenzdaten in Repositorien

Für alle Sequenzen müssen Metadatenätze erstellt und sicher gespeichert werden. Sequenzdaten und entsprechende Metadaten sollten konsistent formatiert und dokumentiert werden. Diese Metadaten dürfen nach alleinigem Ermessen des Eigentümers der Metadaten freigegeben werden. Sequenzierungsdaten und die entsprechenden Metadaten müssen Sicherheitserwägungen, Kosten und Nutzen, geistigen Eigentumsrechten, vertraulichen Geschäftsinformationen, Vertragsbeschränkungen oder anderen verbindlichen schriftlichen Vereinbarungen unterliegen.

ANMERKUNG Zum Schutz privater oder eigentumsbezogener Informationen können für Metadaten oder Sequenzierungsdaten, oder beides, Lizenz- oder Datenschutzrichtlinien, oder beides, geltend gemacht werden.

Zur Förderung von bewährten Verfahren für die Datenverwaltung [3] bietet dieses Dokument optionale Metadaten-Berichtsformate, die auf einen gemeinschaftlichen Datenstandard abgestimmt sind (z. B. Ontologien für MDM oder OBO Foundry). Diese Formate und Standards erleichtern die Reproduzierbarkeit und das gemeinsame Verständnis der Terminologie. Zur Formatierung und Bereitstellung von Werten für die empfohlenen Metadatenfelder wurde eine ISO WGS Slim erstellt. WGS und ausgewählte Metadaten dürfen in eine öffentlich zugängliche Datenbank übertragen (hochgeladen) werden.

4.5 Validierung und Verifizierung des WGS-Workflows

Der gesamte WGS-Workflow muss validiert werden, um sicherzustellen, dass die Verfahren für den bestimmungsgemäßen Gebrauch geeignet sind.

ANMERKUNG Weitere Einzelheiten zur Validierung und Verifizierung des WGS-Workflows finden sich in Abschnitt 10 und in Tabelle 4.

5 Allgemeiner Laborleitfaden

5.1 Bakterienisolierung und DNA-Extraktion

Die Bakterienisolierung und DNA-Extraktion sollte in einem allgemeinen mikrobiologischen Labor durchgeführt werden, das für die Arbeit mit den spezifischen Bakterien, einschließlich Krankheitserregern, geeignet ist. Bei der Vorbereitung einer Sequenzierungsbibliothek, die eine DNA-Amplifikation mittels Polymerase-Kettenreaktion (PCR, en: polymerase chain reaction) beinhaltet, sollten prä- und post-PCR-Schritte in verschiedenen oder getrennten Bereichen des Labors durchgeführt werden, um eine carry-over-Kontamination zu vermeiden.

5.2 Laborumgebung

Luftbewegungen, Vibrationen, Temperatur und Luftfeuchte können die Leistung vieler Sequenzer beeinträchtigen und sollten bei der Platzierung der Geräte im Labor berücksichtigt werden. Labore sollten den Leitfaden des Sequenziergeräteherstellers zur Vorbereitung des Standorts als Orientierungshilfe konsultieren.

5.3 Standardverfahrensanweisungen und nicht konforme Arbeit

Labore sollten Standardverfahrensanweisungen (SOP, en: standard operating procedures), Arbeitsablaufdokumente, Bestandskontrollen von Reagenzien und Gerätewartungsprotokolle pflegen und einhalten. SOPs sollten Verfahren für die Verwendung von Positiv- und Negativkontrollen für die DNA-Extraktion, die Vorbereitung der Sequenzbibliothek und die Sequenzierungsschritte enthalten. SOPs sollten Verfahren zur Überwachung von Arbeitsvorgängen auf Ausführungsqualität und Fehler (falsche Kennzeichnung von Proben oder Kreuzkontamination) enthalten.

Im Falle einer falschen Kennzeichnung oder Kontamination von Proben muss die Grundursache für Fehler in der Sequenzierung untersucht werden:

- a) Sicherstellung, dass Läufe, die falsch gekennzeichnete Proben oder mit mehreren Stämmen kontaminierte Proben enthalten, nicht für bioinformatische Analysen zur Probeninterpretation verwendet oder in Datenbanken hochgeladen werden;
- b) Umsetzung von Maßnahmen zur Aufrechterhaltung der Qualität und zur Verhinderung des erneuten Auftretens von Fehlern.

5.4 Labor-Informations- und Management-System

Die Informationen zu Proben müssen mit einem LIMS oder einem ähnlichen System zur Dokumentation und Verfolgung von Informationen erfasst werden.

5.5 Kompetenz im Labor

Labore sollten Aufzeichnungen führen, die die Aus- und Fortbildung sowie die Eignung der Personen dokumentieren, die Sequenzierungs- und bioinformatische Analysen durchführen, sowie Richtlinien zur Aufbewahrung von Proben.

Das Labor sollte seine Leistung bei der WGS-Analyse durch Vergleich mit den Ergebnissen anderer Labore überwachen, sofern dies möglich und angemessen ist. Diese Überwachung sollte geplant und überprüft werden, idealerweise jährlich, und u. a. mindestens eine der folgenden Maßnahmen umfassen:

- a) Teilnahme an einem Eignungsprüfungsprogramm;
- b) Teilnahme an Vergleichsprüfungen zwischen Laboren mit Ausnahme von Eignungsprüfungen;
- c) Verifizierung des Analyseprozesses durch Einführung von Blindproben oder Proben, deren Merkmale dem Bearbeiter nicht bekannt sind.

Die Daten (z. B. Sequenzierungsdaten, Laufmetriken, von der organisierenden Einrichtung bereitgestellte Ergebnisberichte) aus diesen Überwachungstätigkeiten sollten analysiert, zur Steuerung und gegebenenfalls zur Verbesserung der Tätigkeit des Labors verwendet werden. Wenn sich herausstellt, dass die Ergebnisse der Analyse von Daten aus diesen Überwachungstätigkeiten außerhalb der vordefinierten Kriterien liegen, sollten geeignete Maßnahmen ergriffen werden, um zu verhindern, dass falsche Ergebnisse bei der Probenanalyse verwendet werden.

6 Laborbetrieb

6.1 Herstellung und Lagerung der Proben

Jegliches zu sequenzierende Material (bakterielles Isolat oder extrahierte genomische DNA) sollte so gehandhabt und gelagert werden, dass das Risiko des Probenabbaus, der falschen Kennzeichnung und Kreuzkontamination minimiert wird.

6.2 Bakterienisolate

Bakterienisolate sollten mit Verfahren gelagert und kultiviert werden, die das Potential für die Einführung genetischer Veränderungen (z. B. Verlust von Plasmiden oder Einführung von Polymorphismen durch Kultivierung und Überimpfung) minimieren. Wenn das Labor ein Bakterienisolat erhält, muss das Labor die Reinheit des Isolats sicherstellen und idealerweise die Art bestätigen, bevor weitere Schritte durchgeführt werden. Wenn die Sorge besteht, dass potentiell instabile Elemente (z. B. Plasmide) während der Passage aus einer Probe verloren gehen können, sollten Sequenzen idealerweise von mindestens zwei biologischen Replikaten gesammelt werden. Die Anzahl der einzelnen Koloniepässagen, die nach Erhalt der Isolate durchgeführt werden, sollte in den Metadaten vermerkt werden. Bakterienisolate sollten mit Verfahren wie dem Einfrieren als Glycerin-Stock bei -80 °C archiviert werden.

6.3 Isolierung der DNA

Für die bakterielle DNA-Isolierung sollte ein Extraktionsverfahren gewählt werden, das für die jeweiligen Organismus geeignet ist und DNA von ausreichender Qualität im Hinblick auf die verwendete Sequenzierplattform liefert. Die Isolierung bakterieller DNA wird durch eine Reihe von Faktoren beeinflusst, darunter Zellart (grampositiv oder -negativ), Wachstumsphase (früh-, mittel-, spät-logarithmisch oder -stationär) und Nährmedium. Die Quantität und Qualität der DNA sollte beurteilt und dokumentiert werden. Die Lagerungsbedingungen werden die DNA-Integrität und die Vorbereitung der Bibliothek für bestimmte Sequenziertechnologien beeinflussen.

ANMERKUNG Für die Wiedergewinnung von Plasmiden sind manche DNA-Extraktionsverfahren besser geeignet als andere. Wenn Plasmide für eine spezifische Anwendung wichtig sind, kann ein geeignetes Verfahren verwendet werden.