International Standard

ISO 20397-3

First edition
2025-07

Biotechnology — Massively parallel sequencing —

Part 3: General requirements and guidance for metagenomics

Biotechnologie — Séquençage à grande échelle —

Partie 3: Exigences générales et recommandations pour la métagénomique

Reference number
ISO 20397-3:2025(en)

© ISO 2025



COPYRIGHT PROTECTED DOCUMENT

© ISO 2025
All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ... **iv**
Introduction ... **v**
1 Scope ... **1**
2 Normative references ... **1**
3 Terms and definitions ... **1**
4 Principle ... **4**
4.1 General ... 4
5 Sampling strategy ... **5**
5.1 General ... 5
5.2 Primary sample type ... 6
5.3 Primary sample stabilization and storage ... 6
5.4 Primary sample transportation ... 6
5.5 DNA/RNA isolation ... 7
5.6 DNA/RNA sample quality ... 7
6 Nucleic acid library preparation ... **7**
7 Design and review process including sequencing strategy and assessment ... **7**
7.1 General ... 7
7.2 Short read ... 7
7.3 Long read ... 8
7.4 Hybrid assembly ... 8
7.5 Sample preparation and library construction ... 8
7.6 Assessment ... 8
8 Database construction ... **9**
8.1 General ... 9
8.2 Public database ... 9
8.3 Self-build database ... 9
9 Bioinformatics analysis ... **10**
9.1 General ... 10
9.2 Identification list of microorganisms ... 11
10 Validation and verification ... **11**
10.1 General ... 11
10.2 In silico sequence control for bioinformatics pipeline evaluation ... 11
10.3 Real sample control for pipeline evaluation ... 11
11 Evaluation ... **12**
12 Test report ... **12**
12.1 General ... 12
12.2 Test report content ... 12
Annex A (informative) **Checklist for NA sample quality assessment before library construction** ... **13**
Annex B (informative) **Methods for sample stabilization and storage** ... **14**
Annex C (informative) **Bioinformatics pipeline** ... **15**
Bibliography ... **17**

© ISO 2025 – All rights reserved

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC *276 Biotechnology*, Subcommittee SC 1, *Analytical methods.*

A list of all parts in the ISO 20397 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

© ISO 2025 – All rights reserved

Introduction

Massively parallel sequencing (MPS) is a high-throughput analytical approach to nucleic acid sequencing utilizing massively parallel processing, that allows whole genomes, transcriptomes, and specific nucleic acid targets from different organisms to be investigated in a relatively short time.

Metagenomics approaches are an extremely powerful strategy in large-scale genomics applications as a way to study the taxonomic and functional composition of microbial communities from environmental, agricultural, and clinical primary samples/samples. Metagenomics does not require isolation of single bacterium from complex microbial community, but catalogues by sequencing all genes and genomes from total DNA (tDNA). It has great advantages to identify new species, including microorganisms that are difficult to culture under typical laboratory conditions.

Analysing metagenomics data involves a complex and statistically driven process that extends beyond traditional MPS pipelines to include identification, functional and relative abundance analyses. In metagenomics, whole genomic DNA is prepared from primary samples, regardless of its microbial composition and is characterized by whole genome sequencing. The annotation of resulting DNA fragments, individual reads or assembled sequence contigs, to individual taxonomic groups or known genome sequences, is carried out by sophisticated bioinformatic tools. The analysis is not limited to traditional MPS pipeline but also focuses on the identification of functional composition of a microbial community, which include the assignment of protein-coding open reading frames to functional categories, such as protein domain families or gene ontologies. Consequently, the analysis of whole genome sequencing (WGS) metagenomics data sets involves a significant statistical component, as sequence data must be evaluated based on relative abundances rather than on absolute presence/absence data.

© ISO 2025 – All rights reserved

Biotechnology — Massively parallel sequencing —

Part 3:
General requirements and guidance for metagenomics

1 Scope

This document specifies general requirements and guidance for metagenomics-dedicated sample preparation, and generating and analysing metagenomics sequence data obtained from massive parallel sequencing platforms. The specified metagenomics process includes the following stages:

a) sampling strategy and process, including type, storage, transportation, extraction, quality;

b) nucleic acid library preparation

c) design and review process including sequencing strategy and assessment;

d) database construction;

e) bioinformatics analysis and report

f) validation and verification for bioinformatics pipeline, and database

This document applies to laboratories and research organizations.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 20397-1, *Biotechnology — Massively parallel sequencing — Part 1: Nucleic acid and library preparation*

ISO 20397-2, *Biotechnology — Massively parallel sequencing — Part 2: Quality evaluation of sequencing data*

ISO/TS 24420, *Biotechnology — Massive Parallel DNA Sequencing — General requirements for data processing of shotgun metagenomics sequences*

ISO 20395, *Biotechnology — Requirements for evaluating the performance of quantification methods for nucleic acid target sequences — qPCR and dPCR*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 20397-1, ISO 20397-2 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

© ISO 2025 – All rights reserved