INTERNATIONAL STANDARD

ISO 20397-1

First edition
2022-03

Biotechnology — Massively parallel sequencing —

Part 1: Nucleic acid and library preparation

Biotechnologie — Séquençage parallèle massif —

Partie 1: Acides nucléiques et préparation des collections

Reference number
ISO 20397-1:2022(E)

© ISO 2022



COPYRIGHT PROTECTED DOCUMENT

© ISO 2022

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

© ISO 2022 – All rights reserved

Contents

Page

Foreword ..... **v**
Introduction ..... **vi**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
4 Nucleic acid sample quality evaluation **2**
4.1 General 2
4.2 Sample quantification 3
4.3 Sample purity 3
4.4 Sample integrity 3
4.4.1 General 3
4.4.2 Agarose gel electrophoresis 3
4.4.3 Capillary gel electrophoresis 3
4.4.4 Microfluidic analysis system 3
4.4.5 PCR method 3
5 Nucleic acid library preparation **4**
5.1 General 4
5.2 Fragmentation 4
5.2.1 General 4
5.2.2 Mechanical fragmentation 4
5.2.3 Enzymatic fragmentation 4
5.2.4 Chemical fragmentation 5
5.2.5 Fragmented nucleic acid sample quantity 5
5.2.6 Fragmented nucleic acid sample purity 5
5.2.7 Fragmented nucleic acid size distribution 5
5.2.8 Fragmented nucleic acid purification using gel electrophoresis 5
5.3 Addition of universal sequences 5
5.3.1 Repair 5
5.3.2 Ligation of adapter 5
5.3.3 Barcoding/indexing 6
5.4 Size selection 6
5.5 Amplification 6
5.6 Purification and clean up procedures 6
5.7 Library quantification 7
5.7.1 Library quantification method 7
5.7.2 Selection of quantification method 7
5.8 Library qualification 7
5.8.1 General 7
5.8.2 Methods 7
6 Validation **7**
7 Reference materials or controls **8**
7.1 General 8
7.2 Control samples 8
7.3 Positive control 8
7.4 Negative control 9
7.5 No-template control 9
7.6 Spike-in control 9
7.7 Reference materials 9
8 Contaminations **9**
8.1 General 9
8.2 Primary sample evaluation 9

© ISO 2022 – All rights reserved

8.3 Protocol and operation procedure 10
Annex A (informative) **Checklist for sample quality assessment before library construction** **11**
Annex B (informative) **Examples of quality criteria for selected MPS platforms and applications** **12**
Annex C (informative) **Reference material list** **14**
Bibliography **15**

© ISO 2022 – 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 documents 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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

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.*

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 2022 – All rights reserved

Introduction

Massively parallel sequencing (MPS) is a high throughput analytical technology for nucleic acid sequencing. MPS methods can process thousands to billions of nucleotide sequence reads simultaneously in a single run, allowing whole genomes, transcriptomes and specific nucleic acid targets from different organisms to be analysed in a relatively short time.

MPS is used in many life science disciplines permitting determination and high throughput analysis of millions of nucleotide bases. The biological variability of deoxyribonucleic and ribonucleic acid polymers from living organisms provides challenges in accurately determining their sequences. The quality of sequence determination by MPS depends on many factors including, but not limited to, sample quality, library preparation, and sequencing data quality.

The quality of nucleic acids and libraries prepared for MPS is critical to obtaining high quality sequence data. Controlling the upstream processing steps of MPS and evaluating nucleic acid samples and libraries for their suitability for sequencing significantly improves MPS results, downstream analyses and ultimately conclusions dependent upon the MPS data.

© ISO 2022 – All rights reserved

Biotechnology — Massively parallel sequencing —

Part 1: Nucleic acid and library preparation

1 Scope

This document specifies the general requirements for and gives guidance on quality assessments of nucleic acid samples. It specifies general guidelines for library preparations and library quality assessments prior to sequencing and data generation.

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 20395:2019, *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 20395:2019 and the following apply.

ISO and IEC maintain terminological 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/

3.1
adapter
oligonucleotides of known sequence that are enzymatically added (e.g. ligase or polymerase chain reaction) to the end(s) of a DNA/cDNA fragment

3.2
barcode
index
short sequence of typically six or more nucleotides that serve as a way to identify or label individual samples when they are sequenced in parallel on a single sequencing lane, chip or both

Note 1 to entry: Barcodes are typically located within the sequencing *adapters* (3.1).

3.3
barcoding
indexing
unique DNA sequence identification
method that enables multiple samples to be pooled for sequencing

Note 1 to entry: Each sample is identified by a unique *barcode* (3.2), which enables identification of results during the parallel analysis.

© ISO 2022 – All rights reserved

3.4
GC content
GC
percentage of guanine and cytosine in one or more nucleic acid sequence(s)

Note 1 to entry: The amount of guanine and cytosine in a polynucleotide is usually expressed as fraction (or percentage) of total nitrogenous bases. Total nitrogenous bases comprise the total number of nucleotide bases of reads from one or more MPS run.

[SOURCE: ISO 20397-2:2021, 3.15]

3.5
library
sequencing library
DNA, cDNA or RNA that has been prepared for massively parallel sequencing within a specific size range and typically containing *adapters* (3.1) and/or identifiers recognised for sequence specific priming, sequence capture, and/or identification of specific extracts

Note 1 to entry: Libraries can be DNA or cDNA. cDNA libraries are prepared for RNA sequencing on most sequencers. Some instruments can directly sequence RNA.

3.6
library preparation
sequencing library preparation
set of procedures used to prepare DNA or RNA fragments containing tags, and sequencing primer binding regions for massively parallel sequencing (MPS)

3.7
spike-in control
spike-in process control
target sequence often of defined sequence identity and concentration that are spiked in the sample at various steps of the massively parallel sequencing protocol

Note 1 to entry: Process controls can be used to evaluate any protocol step but are typically applied as nucleic acids controls prior to *library preparation* (3.6).

3.8
Q score
measure of the sequencing quality of a given nucleotide base

[SOURCE: ISO 20397-2:2021, 3.32, modified — The notes to entry have been deleted.]

4 Nucleic acid sample quality evaluation

4.1 General

The laboratory shall establish, implement and document a workflow for nucleic acid quantification that ensures accurate and reproducible results. Requirements for nucleic acid sample quantity and quality can vary between MPS methods. Nucleic acid purification methods can also affect the quality of nucleic acids used for library preparation.

A quality control procedure shall be developed to clearly define nucleic acid quality and library composition. This procedure shall be verified, implemented and documented, and permit accurate quantification of nucleic acid at the minimum amount of nucleic acid required for the MPS performed. The measurement uncertainty and sensitivity of the procedure used for this determination shall be determined. The quantification allows appropriate adjustment of the nucleic acid concentration for input into the MPS sequencer.

© ISO 2022 – All rights reserved