

ICS 07.080
CCS A 40



**NATIONAL STANDARD OF THE
PEOPLE'S REPUBLIC OF CHINA**

GB/T 43584.1-2025

**Biotechnology - Massively parallel sequencing - Part 1: Nucleic
acid and library preparation**

生物技术 大规模并行测序 第1部分：核酸与文库制备

(English Translation)

Issued on: October 31, 2025

Implemented on: May 1, 2026

Issued by: State Administration for Market Regulation

Table of Contents

Preface	
Introduction	
1 Scope	
2 Normative References	
3 Terms and Definitions	
4 Nucleic acid sample quality evaluation	
4.1 General Rules	
4.2 Sample Quantification	
4.3 Sample purity	
4.4 Sample Integrity	
5 Preparation of nucleic acid libraries	
5.1 General Rules	
5.2 Fragmentation	
5.3 Addition of Universal Sequences	
5.4 Size Selection	
5.5 Amplification	
5.6 Purification and Cleaning Procedures	
5.7 Quantitative Library Analysis	
5.8 Qualitative Analysis of Documents	
6 Verification	
7 References or Comparisons	
7.1 General Rules	
7.2 Control Sample	
7.3 Positive Control	
7.4 Negative Control	
7.5 No template reference	
7.6 Internal Reference Comparison	
7.7 References	
8 Pollution	
8.1 General Rules	
8.2 Initial Sample Evaluation	
8.3 Agreements and Operating Procedures	

Appendix A (Informative) Sample Quality Assessment Checklist Before Library Construction

Appendix B (Informative) Example 10 of Quality Standards for Selected MPS Platforms and Applications

Appendix C (Informative) List of References

Reference

Foreword

This document complies with the provisions of GB/T 1.1-2020 "Standardization Work Guidelines Part 1: Structure and Drafting Rules of Standardization Documents".

Drafting.

This document is Part 1 of GB/T 43584 "Large-scale parallel sequencing in biotechnology". GB/T 43584 has been published as follows.

part.

---Part 1: Nucleic Acids and Library Preparation;

---Part 2: Quality assessment of sequencing data.

This document is equivalent to ISO 20397-1:2022 "Biotechnology - Massive parallel sequencing - Part 1: Nucleic acids and library preparation".

Please note that some content in this document may involve patents. The issuing organization of this document assumes no responsibility for identifying patents.

This document was proposed and is under the jurisdiction of the National Technical Committee on Standardization of Biochemical Testing (SAC/TC387).

Introduction

Massive parallel sequencing (MPS) is a high-throughput analysis method that utilizes massively parallel processing for nucleic acid sequencing. This method can...

To conduct research on the whole genome, transcriptome and specific target nucleic acids of different organisms in a relatively short period of time.

MPS has been used in many life science fields, enabling the determination and high-throughput analysis of millions or even tens of millions of nucleotide bases.

Biological variation in deoxyribonucleic acid (DNA) and RNA polymers within organisms presents challenges to accurate sequence determination. MPS determination can be used to determine the sequence...

Sequencing quality depends on many factors, including but not limited to sample quality, library preparation, platform selection, and sequencing data quality.

GB/T 43584 is proposed to consist of the following parts.

---Part 1: Nucleic Acids and Library Preparation. The purpose is to define the procedures for library preparation and quality assessment before sequencing and data generation.

General guidelines and precautions.

---Part 2: Quality Assessment of Sequencing Data. The aim is to build upon Part 1 to implement specific operational procedures and data quality control, and to provide...

Part 3 provides the foundation for the research.

---Part 3: General Requirements and Guidelines for Metagenomics. The aim is to build upon Parts 1 and 2, specifying the requirements for metagenomics.

Genomics guidelines for sample preparation, generation, and analysis of sequencing data.

The quality of nucleic acids and libraries prepared by MPS is crucial for obtaining high-quality sequence data. This is achieved through upstream processing steps of MPS.

By controlling the sequencing process and systematically evaluating the sequencing suitability of nucleic acid samples and libraries, the quality of MPS results can be significantly improved and downstream analysis optimized.

This process ensures the reliability of conclusions drawn from MPS data.

Biotechnology Large-scale parallel sequencing Part 1: Nucleic Acids and Library Preparation

1 Scope

This document specifies the general requirements and guidelines for nucleic acid sample quality assessment.

This document provides general guidelines for library preparation and library quality assessment prior to sequencing and data generation.

2 Normative references

The contents of the following documents, through normative references within the text, constitute essential provisions of this document. Dated citations are not included.

For references to documents, only the version corresponding to that date applies to this document; for undated references, the latest version (including all amendments) applies.

This document.

GB/T 42077-2022 Performance Evaluation Requirements for Biotechnology Nucleic Acid Target Sequence Quantification Methods. qPCR and dPCR (ISO 20395:2019, IDT)

3 Terms and Definitions

The terms and definitions defined in GB/T 42077-2022, as well as the following terms and definitions, apply to this document.

adapter

An oligonucleotide of a known sequence is enzymatically added (e.g., by ligase or polymerase chain reaction) to the end of a DNA/cDNA fragment.

barcode

index

Typically, these are short sequences consisting of six or more nucleotides. When sequencing these sequences in parallel on a single sequencing channel or chip, it is crucial to ensure accurate differentiation.

Data from samples of the same source.

Note: The barcode is usually located inside the sequencing adapter (3.1).

barcoding

Creating an index

Unique DNA sequence identification.

A method that allows multiple samples to be sequenced.

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

GC content; GC

The percentage (s) of guanine and cytosine in one or more nucleic acid sequences.

Note: The content of guanine and cytosine in polynucleotides is usually expressed as a percentage (or proportion) of the total nitrogenous bases. Total nitrogenous bases refer to the number of bases derived from a single source.

The total number of nucleotide bases read in one or more massively parallel sequencing (MPS) runs.

[Source. GB/T 43584.2-2023, 3.15, with modifications]