ISO 20397-2:2021 (E)

Biotechnology — Massively parallel sequencing — Part 2: Quality evaluation of sequencing data

Contents

Foreword

Introduction

1 Scope

2 Normative references

3 Terms and definitions

4 Raw data

- 4.1 General
- 4.2 Raw data file
- 4.3 Quality assessment of raw data
- 4.3.1 General
- 4.3.2 Basic statistics
- 4.3.3 Quality metrics
- 4.4 Raw data pre-processing

5 Sequence alignment and mapping

- 5.1 General
- 5.2 Alignment and mapping file format
- 5.3 Quality control of sequencing alignment and mapping
- 5.3.1 Basic alignment statistics
- 5.3.1.1 General
- 5.3.1.2 Mapping statistics for single-end reads
- 5.3.1.3 Mapping statistics for paired-end reads
- 5.3.1.4 Mapped subread length
- 5.3.2 Quality indicators
- 5.3.3 Methods for alignment and mapping quality assessment
- 5.4 Alignment post-processing

6 Variant calling

- 6.1 General
- 6.2 Data file for variant calling
- 6.3 Quality metrics in the variant calling
- 6.4 Processing of false positive variants
- 6.5 Sequence annotation

7 Validation

- 7.1 General
- 7.2 Validation of quality metrics

8 Documentation

Annex A (informative) Quality metrics for specific example MPS platforms

Annex B (informative) Coverage and read recommendations by applications

Annex C (informative) Software for sequence alignment and mapping

Page count: 21