DIN ISO 20397-2:2026-10 (E)

Biotechnology - Massively parallel sequencing - Part 2: Quality evaluation of sequencing data (ISO 20397-2:2021)

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **1**

4 Raw data ... **6**
- 4.1 General ... 6
- 4.2 Raw data file ... 6
- 4.3 Quality assessment of raw data ... 6
  - 4.3.1 General ... 6
  - 4.3.2 Basic statistics ... 7
  - 4.3.3 Quality metrics ... 7
- 4.4 Raw data pre-processing ... 8

5 Sequence alignment and mapping ... **8**
- 5.1 General ... 8
- 5.2 Alignment and mapping file format ... 9
- 5.3 Quality control of sequencing alignment and mapping ... 9
  - 5.3.1 Basic alignment statistics ... 9
  - 5.3.2 Quality indicators ... 10
  - 5.3.3 Methods for alignment and mapping quality assessment ... 11
- 5.4 Alignment post-processing ... 11

6 Variant calling ... **11**
- 6.1 General ... 11
- 6.2 Data file for variant calling ... 11
- 6.3 Quality metrics in the variant calling ... 12
- 6.4 Processing of false positive variants ... 12
- 6.5 Sequence annotation ... 12

7 Validation ... **12**
- 7.1 General ... 12
- 7.2 Validation of quality metrics ... 13

8 Documentation ... **14**

Annex A (informative) **Quality metrics for specific example MPS platforms** ... **15**

Annex B (informative) **Coverage and read recommendations by applications** ... **16**

Annex C (informative) **Software for sequence alignment and mapping** ... **18**

Bibliography ... **19**