ISO/TS 22690:2021-10 (E)

Genomics informatics - Reliability assessment criteria for high-throughput gene-expression data

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Normative references 1

3 Termsanddefinitions 1

4 From sample to RNA 3
4.1 General 3
4.2 RNA integrity 3
4.3 RNA concentration 4
4.4 RNA purity 4

5 Expressionprofiling 4

6 Quality control metrics in RNA-seq analysis 4
6.1 General 4
6.2 Sequencing read 4
6.2.1 Total number of reads 4
6.2.2 Read length 5
6.2.3 Base call quality 5
6.2.4 GC content 5
6.2.5 Overrepresented sequence 5
6.2.6 Adapter residue 5
6.3 Alignment 5
6.3.1 Alignment ratio 5
6.3.2 Gene body coverage uniformity 5
6.3.3 Strand specificity 6
6.3.4 Insert size 6
6.3.5 Mismatch 6
6.3.6 Contamination from other sources 6
6.4 Expression 6
6.4.1 Expression distribution 6
6.4.2 Expressed genes 6
6.4.3 Saturation 7
6.4.4 Reproducibility 7
6.5 Differentially expressed genes 7
6.6 Biological interpretation of differentially expressed genes 7
6.7 Sample certificate of origin 8
6.8 Quality control of batch effects 8

7 Spike-in controls 8

8 Proficiencytesting 8
8.1 General 8
8.2 RNA sources 9
8.3 Experimental design 9

9 Process management 9

Bibliography 10