PD ISO/TS 22692:2020

BSI Standards Publication

Genomics informatics — Quality control metrics for DNA sequencing

bsi.

National foreword

This Published Document is the UK implementation of ISO/TS 22692:2020.

The UK participation in its preparation was entrusted to Technical Committee IST/35, Health informatics.

A list of organizations represented on this committee can be obtained on request to its committee manager.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

Published by BSI Standards Limited 2020

ISBN 978 0 539 00455 7

ICS 35.240.80

Compliance with a British Standard cannot confer immunity from legal obligations.

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 October 2020.

Amendments/corrigenda issued since publication

Date	Text affected

PD ISO/TS 22692:2020

TECHNICAL SPECIFICATION

ISO/TS 22692

First edition
2020-10

Genomics informatic — Quality control metrics for DNA sequencing

Informatique génomique — Mesures de contrôle de la qualité pour le séquençage de l'ADN

Reference number
ISO/TS 22692:2020(E)

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

Contents

Page

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 215, *Health informatics*, Subcommittee SC 1, *Genomics informatics*.

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.

Introduction

The rapid progress in Next Generation Sequencing (NGS) technology has drastically reduced the cost and time for genomic analysis. A number of research institutions, corporations, and government agencies are competitively collecting a large volume of genomic data through multi-national, multi-institutional projects such as "DiscovEHR"[9], "gnomAD"[10] and "UK Biobank"[11]. The demand for sharing of "high quality" genomic data is growing because large-scale reference data is required for reliable detection of mutation for both industrial and clinical applications.

However, the quality of available genomic data is less than desirable. To establish consistent quality control metrics, details of each stage of NGS process need to be recorded, shared and standardized (processes and data elements collected and coded for each stage and sub-stage). These processes include sample preparation, library preparation, sequencing, and data processing, among others, as shown in Figure 1.

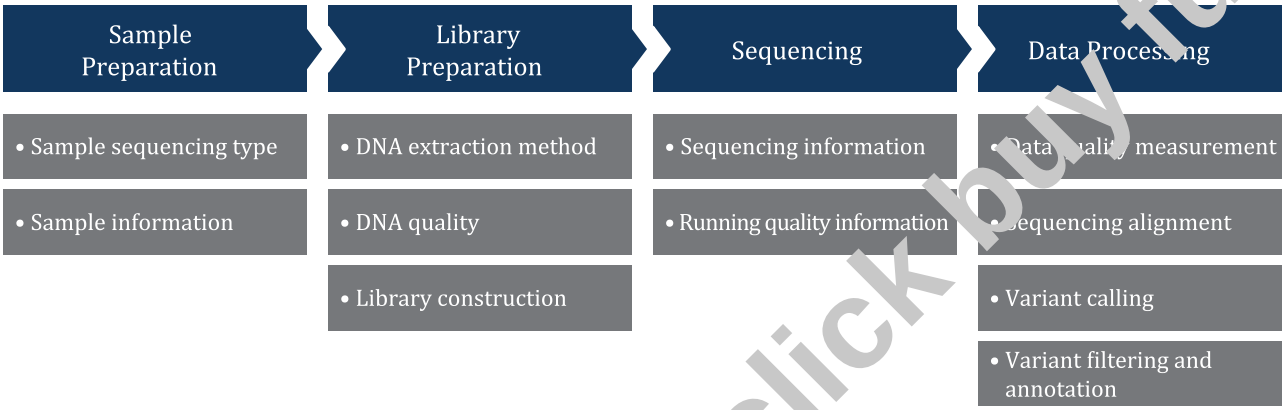

Figure 1 — NGS process

Genomics informatic — Quality control metrics for DNA sequencing

1 Scope

This document identifies quality metrics for the detection of DNA variants using next generation sequencing (NGS) technology. It also defines the data types, relationships, optionality, cardinalities and terminology bindings of the data.

This document provides a basis for sharing and for the application of "high quality" genomic data and contributes to the realization of the precision medicine and the development of relevant industries.

This document is intended to serve as a catalogue of sequencing data elements used to address quality metrics for various clinical, industrial and commercial applications. The exchange of these data allows researchers, commercial entities, and regulatory bodies to assess for the purpose of selective utilization of the data by setting application-specific quality criteria

This document is not intended for

— sequencing methods other than NGS, such as the Sanger sequencing,

— targets other than genome, such as transcriptome or proteome, or

— specimens of species other than humans.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions 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 http://www.electropedia.org/

3.1
copy number variation
CNV
variation (3.18) in the number of copies of one or more sections of the *DNA* (3.3)

[SOURCE: ISO/TS 20428:2017, 3.7]

3.2
deletion
contiguous removal of one or more bases from a genomic sequence

[SOURCE: ISO/IEC 23092-2:2019, 3.4]