ISO/IEC 23092-1:2020 (E)

Information technology — Genomic information representation — Part 1: Transport and storage of genomic information

Contents

Foreword

Introduction

1 Scope

2 Normative references

3 Terms and definitions

4 Mathematical operators

4.1 Arithmetic operators
4.2 Logical operators
4.3 Relational operators
4.4 Bitwise operators
4.5 Assignment
4.6 Unary operators

5 Structure of coded genomic data

5.1 Genomic records
5.2 Data classes
5.3 Access units
5.4 Datasets
5.5 Selective access

6 Data format

6.1 Format structure
6.1.1 General
6.1.2 Box order
6.2 Syntax and semantics
6.2.1 Method of specifying syntax in tabular form
6.2.2 Bit ordering
6.2.3 Specification of syntax functions
6.3 Syntax for representation
6.4 Output data unit
6.5 Data structures common to file format and transport format
6.5.1 File header
6.5.1.1 General
6.5.1.2 Syntax
6.5.1.3 Semantics
6.5.2 Dataset group
6.5.2.1 General
6.5.2.2 Dataset group header
6.5.2.2.1 General
6.5.2.2.2 Syntax
6.5.2.2.3 Semantics
6.5.2.3 Reference
6.5.2.3.1 General
6.5.2.3.2 Syntax
6.5.2.3.3 Semantics
6.5.2.3.4 ref_uri semantics
6.5.2.3.5 Supported FASTA format
6.5.2.3.6 Conversion to raw reference

6.5.2.3.7 Checksum
6.5.2.4 Reference metadata
6.5.2.4.1 General
6.5.2.4.2 Syntax
6.5.2.4.3 Semantics
6.5.2.5 Label List
6.5.2.5.1 General
6.5.2.5.2 Syntax
6.5.2.5.3 Semantics
6.5.2.5.4 Label
6.5.2.5.4.1 General
6.5.2.5.4.2 Syntax
6.5.2.5.4.3 Semantics
6.5.2.6 Dataset group metadata
6.5.2.6.1 General
6.5.2.6.2 Syntax
6.5.2.6.3 Semantics
6.5.2.7 Dataset group protection
6.5.2.7.1 General
6.5.2.7.2 Syntax
6.5.2.7.3 Semantics
6.5.3 Dataset
6.5.3.1 General
6.5.3.2 Dataset header
6.5.3.2.1 General
6.5.3.2.2 Syntax
6.5.3.2.3 Semantics
6.5.3.2.4 Alphabets
6.5.3.2.5 Block contiguity
6.5.3.3 Dataset metadata
6.5.3.3.1 General
6.5.3.3.2 Syntax
6.5.3.3.3 Semantics
6.5.3.4 Dataset protection
6.5.3.4.1 General
6.5.3.4.2 Syntax
6.5.3.4.3 Semantics
6.5.3.5 Dataset parameter set
6.5.3.5.1 General
6.5.3.5.2 Syntax
6.5.3.5.3 Semantics
6.5.4 Access unit
6.5.4.1 General
6.5.4.2 Access unit decapsulation
6.5.4.3 Access unit header
6.5.4.3.1 General
6.5.4.3.2 Syntax
6.5.4.3.3 Semantics
6.5.4.4 Access unit information
6.5.4.4.1 General
6.5.4.4.2 Syntax
6.5.4.4.3 Semantics
6.5.4.5 Access unit protection
6.5.4.5.1 General
6.5.4.5.2 Syntax
6.5.4.5.3 Semantics
6.5.5 Block
6.5.5.1 General
6.5.5.2 Block header
6.5.5.2.1 General
6.5.5.2.2 Syntax
6.5.5.2.3 Semantics
6.6 Data structures specific to file format
6.6.1 General

6.6.2 Indexing
6.6.2.1 Master index table
6.6.2.1.1 General
6.6.2.1.2 Syntax
6.6.2.1.3 Semantics
6.6.3 Descriptor stream
6.6.3.1 General
6.6.3.2 Descriptor stream header
6.6.3.2.1 General
6.6.3.2.2 Syntax
6.6.3.2.3 Semantics
6.6.3.3 Descriptor stream protection
6.6.3.3.1 General
6.6.3.3.2 Syntax
6.6.3.3.3 Semantics
6.6.4 Offset
6.6.4.1 Syntax
6.6.4.1.1 Semantics
6.7 Data structures specific to transport format
6.7.1 General
6.7.2 Data streams
6.7.3 Dataset mapping table list
6.7.3.1 General
6.7.3.2 Syntax
6.7.3.3 Semantics
6.7.4 Dataset mapping table
6.7.4.1 General
6.7.4.2 Syntax
6.7.4.3 Semantics
6.7.5 Packet
6.7.5.1 General
6.7.5.2 Packet header
6.7.5.2.1 General
6.7.5.2.2 Syntax
6.7.5.2.3 Semantics
6.8 Reference procedure to convert transport format to file format

Annex A (informative) IETF RFC 3986 specification summary

Annex B (informative) Selective access strategies

B.1 Access unit assembling
B.2 Paired or multi-segment reads and access units
B.3 Selective access and block-based encoding
B.4 Overlapping AUs ranges
B.5 Non-overlapping AUs ranges

Annex C (informative) Depacketization process

Page count: 56