



中华人民共和国国家标准

GB/T 26237.14—2023

代替 GB/T 26237.14—2019

信息技术 生物特征识别数据交换格式 第 14 部分：DNA 数据

Information technology—Biometric data interchange formats—
Part 14: DNA data

(ISO/IEC 19794-14:2013, MOD)

2023-05-23 发布

2023-12-01 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	I
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语和定义	2
4 缩略语	3
5 符合性	3
6 DNA 数据格式	4
附录 A (规范性) DNA XML 模式	30
附录 B (资料性) 家系图构建示例	50
附录 C (资料性) DNA 分型试剂盒	58
附录 D (规范性) DNA 位点	87
附录 E (资料性) 数据记录示例	92
附录 F (规范性) 符合性测试方法	96

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

本文件是 GB/T 26237《信息技术 生物特征识别数据交换格式》的第 14 部分，GB/T 26237 已发布如下部分：

- 第 1 部分：框架；
- 第 2 部分：指纹细节点数据；
- 第 3 部分：指纹型谱数据；
- 第 4 部分：指纹图像数据；
- 第 5 部分：人脸图像数据；
- 第 6 部分：虹膜图像数据；
- 第 7 部分：签名/签字时间序列数据；
- 第 8 部分：指纹骨架数据；
- 第 9 部分：血管图像数据；
- 第 10 部分：手形轮廓数据；
- 第 11 部分：签名/签名处理的动态数据；
- 第 13 部分：声音数据；
- 第 14 部分：DNA 数据。

本文件代替 GB/T 26237.14—2019《信息技术 生物特征识别数据交换格式 第 14 部分：DNA 数据》，与 GB/T 26237.14—2019 相比，除结构调整和编辑性改动外，主要技术变化如下：

- a) 删除了“脱氧核糖核酸”“基因组”“线粒体”“个体识别率”“祖源”“表型”“单倍型”“新一代测序技术”“流动 DNA 处置单元”和“快速 DNA 处置单元”10 个术语(见 2019 年版的第 4 章)；
- b) 删除了“iSNPs”“NGS”“PISNPs”“UTC”4 个缩略语，增加了“ICS”“ID”和“IUT”3 个缩略语(见 2019 年版的第 5 章)；
- c) 增加了关于符合性测试的要求(见第 5 章)；
- d) 删除了 D.2.4 常染色体 SNP 位点、D.2.5 Y-SNP 位点、D.2.6 X-SNP 位点、D.2.7 线粒体 SNP 位点(见 2019 年版的附录 D)；
- e) 增加了符合性测试方法(见附录 F)。

本文件修改采用 ISO/IEC 19794-14:2013《信息技术 生物特征识别数据交换格式 第 14 部分：DNA 数据》。

本文件与 ISO/IEC 19794-14:2013 相比做了下述结构调整：

- a) 第 2 章对应 ISO/IEC 19794-14:2013 中的第 3 章；
- b) 第 3 章对应 ISO/IEC 19794-14:2013 中的第 4 章；
- c) 第 4 章对应 ISO/IEC 19794-14:2013 中的第 5 章；
- d) 第 5 章对应 ISO/IEC 19794-14:2013 中的第 2 章；
- e) 将原 6.4.1 DNA 记录通用数据头结构的悬置段改为 6.4.1.1 DNA 记录通用数据头组成，将原 6.4.2 表述元数据的悬置段改为 6.4.2.1 表述元数据结构，将原 6.4.3.1 DNA 信息的悬置段改为 6.4.3.1.1 DNA 信息结构，将原 6.4.3.2 DNA 分型数据的悬置段改为 6.4.3.2.1 DNA 分型数

据类型；

- f) 增加了表 29~表 36,表 37~表 51 对应 ISO/IEC 19794-14:2013 的表 29~表 43,增加了表 52~表 63,表 64 对应 ISO/IEC 19794-14:2013 的表 44；
- g) 附录 A 对应 ISO/IEC 19794-14:2013 中的附录 B。附录 B 和附录 F 为根据修改件内容增加,附录 E 为自主增加。

本文件与 ISO/IEC 19794-14:2013 的技术差异及其原因如下：

- a) 增加了规范性引用文件 GB/T 1988—1998,以便在我国范围内对本文件更好地理解与应用(见 6.4.3.2.11.2)；
- b) 增加了规范性引用文件 GB/T 2659—2000,以便在我国范围内对本文件更好地理解与应用(见表 F.1 和表 F.2)；
- c) 用规范性引用的 GB/T 26237.1—2022 替换了 ISO/IEC 19794—1:2011 和 ISO/IEC 19794—1:2011/Amd 2,两者的一致性程度为修改(见第 3 章、第 6 章和附录 F)；
- d) 增加了规范性引用文件 GB/T 29859—2013,因为引用了该文件的术语(见第 3 章)；
- a) 增加了术语中的“单核苷酸多态性”,以便于读者理解后文中“SNP”的相关内容。删除了“脱氧核糖核酸”“个体识别力”因后文中未提到该两术语。删除了“快速 DNA 处置单元”“流动 DNA 处置单元”,因该两术语在文中仅出现一次,故以注的形式体现；
- b) 增加了表 1、表 4、表 25、表 26、表 31、表 32、表 35、表 36、表 39、表 40、表 41、表 42、表 43、表 44、表 45、表 46、表 47、表 48、表 51、表 53、表 54、表 58、表 59、表 60、表 61、表 62、表 63 中“长度(Byte)”列,便于更好地理解与应用本文件；
- c) 删除了表 3 中“国家代码”行,以便在我国范围内更好地理解与应用本文件；
- d) 更改“GPS”为“BDS”,以便在我国范围内对本文件更好地理解与应用(见 6.4.2.9)；
- e) 增加了表 7 中“X-STR”“SNP”“Y-SNP”“X-SNP”行,删除了“电泳图谱”以符合行业实际情况。增加了表 15 中的“中国合格评定国家认可委员会实验室认可”“中国计量认证”行,删除了“ISO/IEC 17025 认证”“GLP 验证”“AABB 认证”“ISO/ILAC Guild 19 认证”,以便在我国范围内对本文件更好地理解与应用。将表 7 中的“核 DNA 检测”行拆分为“血型检验”“人类血(斑)确证试验”“人类精液(斑)确证试验”“常染色体 STR 及性别检测”“X 染色体 STR 检测”行,以便与实际行业使用情况匹配；
- f) 增加了表 22 中“X-STR 分型图谱”“SNP 分型”“Y- SNP 分型”“X- SNP 分型”“单倍型分型”“NGS 测序数据”,以便与实际行业使用情况匹配；
- g) 增加了 X-STR 分型图谱、SNP 分型、Y-SNP 分型、X-SNP 分型、单倍型分型、NGS 测序数据,以便与实际行业使用情况匹配(见 6.4.3.2.4、6.4.3.2.6、6.4.3.2.7、6.4.3.2.8、6.4.3.2.9 和 6.4.3.2.11)；
- h) 增加 D.1.3 X-STR 位点、D.2 高通量测序 DNA 分型技术检测位点,以便与实际行业使用情况匹配；
- i) 更改了附录 D 中的 DNA 分型试剂盒和 DNA 位点,添加我国国内主流 DNA 分型试剂盒包含的 DNA 位点,以便在我国范围内对本文件更好地理解与应用。

本文件纳入了 ISO/IEC 19794-14:2013/Amd 1:2016 的修正内容,所涉及条款的外侧页边空白位置用垂直双线(∥)进行了标示。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本文件起草单位:深圳华大法医科技有限公司、中国电子技术标准化研究院、西安交通大学、深圳华

大基因股份有限公司、厦门市熠成信息技术有限公司、山西医科大学、华南理工大学、深圳市生命科技产学研资联盟、深圳华大基因科技有限公司、广东省公安厅刑事技术中心、天复(东莞)标准技术有限公司、临汾市公安局。

本文件主要起草人：郭小森、王文峰、赖江华、张洪波、高升杰、沈悦生、吴昊、刘倩颖、苏立伟、宋继伟、杜红丽、严江伟、李泽琴、李倩一、李海燕、郑飞雪、杨越、王秋娟、庞坤杰、高昀露、石红岩、李陶莎、武庆超、魏曙光、穆豪放、姜华艳、王成、张雷、尹焯。

引 言

GB/T 26237《信息技术 生物特征识别数据交换格式》规定了各模态生物特征数据进行交换的格式要求,是提高系统的互操作性的核心。GB/T 33842《信息技术 GB/T 26237 中定义的生物特征识别数据交换格式的符合性测试方法》则针对 GB/T 26237《信息技术 生物特征识别数据交换格式》给出符合性测试方法。自 2016 年开始,生物特征识别数据交换格式的符合性测试方法已经开始逐步并入《信息技术 生物特征识别数据交换格式》,不再额外立标准,因此修订本文件时,将符合性测试方法作为附录。GB/T 26237《信息技术 生物特征识别数据交换格式》的其他部分,会在修订时与本文件保持一致,增添符合性测试方法作为附录内容。GB/T 26237《信息技术 生物特征识别数据交换格式》拟由 15 部分组成。

- 第 1 部分:框架。目的在于描述定义生物特征识别数据交换格式的总体概念和要求。
- 第 2 部分:指纹细节点数据。目的在于规定基于细节点的指纹表示中所用到的概念和数据格式。可以在涉及自动指纹识别的领域内使用。
- 第 3 部分:指纹型谱数据。目的在于规范指纹型谱的术语和定义、指纹型谱数据、指纹型谱数据记录、指纹型谱数据卡格式指纹型谱数据的存储形式,以及基于频谱的指纹数据的交换格式,以指导自动指纹识别的各种领域的应用。
- 第 4 部分:指纹图像数据。目的在于规定一个用来从一个或多个符合 ISO/IEC19785-1 数据结构的指/掌纹图像区域进行存储、记录和传输信息的数据记录交换格式,可用来进行指纹图像数据的交换和比较。
- 第 5 部分:人脸图像数据。目的在于为要求交换人脸图像数据的人脸识别应用提供一种人脸图像格式。
- 第 6 部分:虹膜图像数据。目的在于定义虹膜信息交换所采用的紧凑表示形式,用于虹膜图像这一生物特征的注册、验证和识别系统。
- 第 7 部分:签名/签字时间序列数据。目的在于为签名/签字信息提供了一种数据交换格式,规定了可以互联互通的签名/签字记录格式(时间序列数据格式)。
- 第 8 部分:指纹骨架数据。目的在于规定基于指纹型骨架模式的指纹识别数据交换格式,适用于自动指纹识别的各种应用领域。
- 第 9 部分:血管图像数据。目的在于规定血管生物特征图像的身份识别或身份验证技术所用的图像交换格式,可用于血管图像数据的交换和压缩。
- 第 10 部分:手形轮廓数据。目的在于规定一种手型记录数据交换格式,以支持各种基于手型信息辨识或验证的自动化设备和系统的信息交换。
- 第 11 部分:签名/签名处理的动态数据。目的在于为从时间序列提取的处理后的签名/标志行为数据指定一种数据交换格式,该数据是使用数字化数位板,基于笔的计算设备或高级笔系统等设备捕获的。
- 第 12 部分:脸型特性数据。目的在于规定用于生物特征识别或验证技术的脸型特性数据的交换格式。
- 第 13 部分:声音数据。目的在于指定一种数据交换格式,该数据交换格式可用于存储,记录和

传输假定来自单个会话中记录的单个扬声器的人的数字化语音数据,即声音数据。

- 第 14 部分:DNA 数据。目的在于规定一种用于 DNA 分型数据的交换格式,以支持来进行个体识别的 DNA 分型数据交换。
- 第 15 部分:掌纹图像数据。目的在于规定用于生物特征识别或验证技术的图像交换格式,该技术利用人类掌纹生物特征图像,并且可以用于掌纹图像数据的交换和比较。

信息技术 生物特征识别数据交换格式

第 14 部分:DNA 数据

1 范围

本文件规定了:

- 一种用于 DNA 分型数据的交换格式;
- DNA 数据交换格式的符合性测试方法的要素、测试断言和测试规程。

本文件适用于:

- 使用人类 DNA 来进行个体识别的 DNA 分型数据交换;
- 本文件第 6 章中规定的 DNA 数据交换格式结构的断言测试(对应 ISO/IEC 29109-1:2009 中定义的类型 A 的级别 1);

通过检测 DNA 数据每个字段的值的类型来检查内部一致性的断言测试(对应 ISO/IEC 29109-1:2009 中定义的类型 A 的级别 2)。

本文件不适用于:

- 本文件要求的 CBEFF 结构的符合性测试;
- 生物特征识别数据交换格式符合性测试级别 3 的测试;
- 生物特征识别产品其他特性的测试或其他类型的测试(如接受率、性能、鲁棒性和安全性等);
- 不产生符合本文件第 6 章要求的记录的系统的符合性测试;
- 除此之外的其他任何目的,特别是医疗和其他健康相关的信息交换。

本文件将实现对 DNA 分型数据和基因组测序数据的交换或相互比对,其中包括在符合相关隐私保护法规的基础上,由任何其他系统遵循本文件的数据格式而输出兼容的 DNA 分型数据和基因组测序数据。

本文件涵盖法医 DNA 分型技术、高通量测序技术,或其他基于短串联重复(STRs,包括常染色体、X 染色体、Y 染色体)、线粒体 DNA、单核苷酸多态性(SNP)等遗传标记分析。

一个对象的单一 DNA 分型数据记录可能需要使用多种 DNA 分析技术产生的数据。本文件确保了一个对象由多个 DNA 分析技术产生的数据能够集成在一个记录中。

这种数据格式是为了减少 DNA 处理过程中(登记和比对)的人工参与。为了满足 DNA 技术自动化、智能化对数据交换格式的要求,本文件同时描述了原始数据和分型数据的记录格式。数据的可扩展标记语言(XML)编码用于详细说明 DNA 分型数据交换。规范的 XML 模式(XSD)在附录 A 中给出。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

- GB/T 1988—1998 信息技术 信息交换用七位编码字符集(eqv ISO/IEC 646:1991)
- GB/T 2659—2000 世界各国和地区名称代码(eqv ISO 3166-1:1997)