



中华人民共和国公共安全行业标准

GA/T 2353—2025

法庭科学 线粒体DNA的数据结构

Forensic sciences—Data structure of mitochondrial DNA

2025-10-13发布

2026-02-01实施

中华人民共和国公安部 发布

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	2
5 通用交换信息文件	2
附录A(资料性) IUPAC代码的英文及中文含义	6
附录B(资料性) 修订的剑桥参考序列(rCRS)	7
附录C(资料性) DDEM文件注释行	15
附录D(资料性) mtDNA完整序列DDEM文件示例	17
附录E(资料性) mtDNA部分序列DDEM文件示例	25
附录F(资料性) 异质性碱基序列DDEM文件示例	28
参考文献	37

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。

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。

本文件由全国刑事技术标准化技术委员会(SAC/TC 179)提出并归口。

本文件起草单位：公安部鉴定中心、公安部防范电信网络诈骗信息监控中心、上海市公安局、北京海华鑫安生物信息技术有限责任公司、上海诚明融鑫科技有限公司。

本文件主要起草人：胡兰、刘冰、孙辉、刘亚楠、王彤、李冬涛、郝思静、虞秀华。

法庭科学 线粒体 DNA 的数据结构

1 范围

本文件给出了法庭科学所选用的人类线粒体 DNA 碱基序列类型,规定了系统间进行线粒体 DNA 数据交换的文件格式、数据结构和基本要求。

本文件适用于需要交换法庭科学线粒体 DNA 数据的外部系统(如 DNA 实验室管理信息系统、DNA 数据分析软件等)的设计、开发和测试。

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 2312—1980 信息交换用汉字编码字符集 基本集

GB 18030—2022 信息技术 中文编码字符集

GB/T 41009—2021 法庭科学 DNA 数据库选用的基因座及其数据结构

3 术语和定义

下列术语和定义适用于本文件。

3.1

线粒体 DNA mitochondrial DNA ; mtDNA

存在于人类细胞的线粒体中的闭环双链 DNA。

注:人类线粒体 DNA 约 16.5 Kb;不遵循孟德尔遗传定律,表现为母系遗传,通过卵细胞将其中的遗传信息传递给后代;无有丝分裂和减数分裂的周期变化;遗传物质位于细胞器内,不受核移植的影响;单个细胞中 mtDNA 拷贝数多;存在异质性现象。

[来源:GB/T 21679—2023,3.3]

3.2

高变区 hypervariable region ; HVR

位于 mtDNA 非编码区的 DNA 区域,突变率可达到单拷贝核基因组的 6~17 倍,且无修复系统、不受选择压力的影响,积累了较多的变异,多态性好。

注:也称为控制区(control region),或者 D-环区(D-loop region),通常根据在 mtDNA 链上的位置分为高变区 I (HVR-I, 位于 16 024 bp~16 365 bp)、高变区 II (HVR-II, 位于 73 bp~340 bp)和高变区 III (HVR-III, 位于 340 bp~576 bp)。

[来源:GA/T 1979—2022,3.3]

3.3

碱基序列 base sequence

测序片段中记录碱基排列的字符串,碱基序列中的每个碱基应使用大写字母(A、T、C、G 和 N)或小写字母(a、t、c、g 和 n),其中字母 A 和 a 表示腺嘌呤,字母 T 和 t 表示胸腺嘧啶,字母 C 和 c 表示胞嘧啶,字母 G 和 g 表示鸟嘌呤,字母 N 和 n 表示未测定的碱基。