



中华人民共和国国家计量技术规范

JJF ××××—202×

单分子基因测序仪校准规范

Calibration Specification for Single-Molecule Gene Sequencers

(征求意见稿)

202×-××-××发布

202×-××-××实施

国家市场监督管理总局发布

单分子基因测序仪 校准规范

**Calibration Specification for
Single-Molecule Gene Sequencers**

JJF ××××—××××

本规范委托全国生物计量技术委员会负责解释

目 录

引 言	1
1 范围	4
2 引用文件	4
3 术语和计量单位	4
3.1 单分子基因测序 single-molecule gene sequencing	4
3.2 碱基 base	5
3.3 单分子测序文库 single-molecule sequencing library	5
3.4 条形码 barcode	5
3.5 纳米孔 nanopore	5
3.6 离子电流信号 ionic current signal	5
3.7 马达蛋白 motor protein	5
3.8 零模波导孔 zero-mode wave aperture	6
3.9 脉冲荧光信号 pulsed fluorescent signal	6
3.10 实时测序 real-time sequencing	6
3.11 直接测序 direct sequencing	6
3.12 测序通量 throughput of gene sequencing	6
3.13 碱基识别 base calling	7
3.14 碱基识别质量 quality of base calling	7
3.15 碱基识别错误率 inaccuracy of base calling	7
3.16 Q10	7
3.17 碱基测序准确率 base sequencing accuracy	7
3.18 测序覆盖率 coverage rate of sequencing	7
3.19 测序深度 depth of coverage	8
3.20 平均覆盖深度 average depth of coverage	8
3.21 读长 N50 read length N50	8
3.22 参考数据库 reference database	8
3.23 比对率 mapping rate	8
3.24 序列相对丰度 sequence relative abundance	8
3.25 归一化读段数 normalized read count	9
3.26 等长参考序列相对丰度比 relative abundance ratio of equal-length reference sequences	9
3.27 一致性序列 consensus sequence	9
4 概述	9
5 计量特性	10
6 校准条件	11
6.1 环境条件	11
6.2 校准用标准物质试剂	11
6.3 校准用设备及分析软件	12
7 校准项目和校准方法	12
7.1 reads 数重复性	12

7.2 比对率重复性	12
7.3 序列覆盖率重复性	13
7.4 序列相对丰度偏差	14
7.5 等长参考序列相对丰度比偏差	14
7.6 读长 N50 偏差	15
7.7 GC 含量偏差	15
7.8 测序一致序列准确度	16
7.9 单次测序众数准确度	16
7.10 序列识别质量百分比 (Q10)	17
8 校准结果表达	17
9 复校时间间隔	17
附录 A 标准物质的选择原则	18
附录 B 校准原始记录参考格式	17
附录 C 校准证书(内页) 参考格式	20
附录 D 序列相对丰度偏差测量结果的不确定度评定示例	21
附录 E 等长参考序列相对丰度比偏差测量结果的不确定度评定示例	24
附录 F GC 含量偏差测量结果的不确定度评定示例	27

引 言

JJF 1071《国家计量校准规范编写规则》、JJF 1001《通用计量术语及定义》和 JJF 1059.1《测量不确定度评定与表示》共同构成支撑本规范制定工作的基础性系列规范。

本规范校准方法及计量特性等主要参考了 GB/T 30989-2014《高通量基因测序技术规程》、GB/T 40226-2021《环境微生物宏基因组检测 高通量测序法》、GB/T 46942.1-2025《单分子基因测序 第1部分：术语》、JJF1265-2022《生物计量术语及定义》、JJF2396-2026《高通量基因测序仪校准规范》及 YY/T 1723-2020《高通量基因测序仪》。

本规范为首次发布。

单分子基因测序仪校准规范

1 范围

本规范适用于单分子基因测序仪（测序读长 ≥ 1000 bp）的校准，包括但不限于基于纳米孔链测序、单分子实时荧光测序等原理的单分子基因测序仪。

本规范不适用于依赖核酸模板扩增的高通量基因测序仪，也不适用于以桑格测序技术为原理的基因测序仪。

2 引用文件

本规范引用了下列文件：

GB/T 30989-2014 《高通量基因测序技术规程》

GB/T 35537-2017 《高通量基因测序结果评价要求》

GB/T 40226-2021 《环境微生物宏基因组检测 高通量测序法》

GB/T 46942.1-2025 《单分子基因测序 第1部分：术语》

JJF1265-2022 《生物计量术语及定义》

JJF2396-2026 《高通量基因测序仪校准规范》

YY/T 1723-2020 《高通量基因测序仪》

凡是注日期的引用文件，仅注日期的版本适用于本规范；凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本规范。

3 术语和计量单位

GB/T 30989-2014、GB/T 40226-2021、GB/T 35537-2017、GB/T 46942.1-2025、JJF1265-2022、JJF2396-2026 及 YY/T 1723-2020 界定的以及下列术语和定义适用于本规范。

3.1 单分子基因测序 single-molecule gene sequencing

区别于桑格测序、大规模平行高通量测序及利用共标签进行短序列连接的单分子长片段测序方法，在单个分子水平上对核酸分子进行连续碱基序列测定通常基于光学或电

学信号等。

[来源：GB/T 46942.1-2025 3.17]

3.2 碱基 base

一类含氮原子的有机杂环化合物，是组成嘌呤和嘧啶的主要成分，是拼出遗传密码的“字母”。

注：A（adenine）代表脱氧腺嘌呤核苷酸（简称脱氧腺苷）；T（thymine）代表脱氧胸腺嘧啶核苷酸（简称脱氧胸苷）；C（cytosine）代表脱氧胞嘧啶核苷酸（简称脱氧胞苷）；G（guanine）代表脱氧鸟嘌呤核苷酸（简称脱氧鸟苷）。

[来源：GB/T 30989-2014, 3.16]

3.3 单分子测序文库 single-molecule sequencing library

通过核酸工具酶或其它方法将待测核酸分子与单分子基因测序平台适配的接头进行偶联而获得。

[来源：GB/T 46942.1-2025 3.27]

3.4 条形码 barcode

连接在 DNA 片段两端的短序列，用于区分不同技术重复组。

注：通常由 10~24 个碱基组成，各条形码之间具有足够的序列差异，以便在数据分析时准确拆分。

3.5 纳米孔 nanopore

具有贯穿通道且最窄处尺寸为纳米量级的生物孔道蛋白或固态孔道，是将碱基序列编码为电信号的转换器。

[来源：GB/T 46942.1-2025 5.19]

3.6 离子电流信号 ionic current signal

溶液中的带电离子在电场力作用下穿越纳米孔产生离子电流，当核酸分子单链穿过纳米孔或标签分子与纳米孔互动时则会对该离子电流产生扰动，形成了对应碱基序列信息的连续电流时序信号。

3.7 马达蛋白 motor protein

单分子纳米孔链测序中，用于对待测核酸分子进行控速的蛋白分子。

注：通常为解旋酶，也可以是聚合酶或其他类型蛋白。

[来源：GB/T 46942.1-2025 5.24]

3.8 零模波导孔 zero-mode wave aperture

直径为数十纳米的孔道结构，由于直径小于激发光波长而将激发范围限制在孔底部的聚合酶附近，从而实现对单个分子合成反应荧光信号的观测，有效避免背景游离荧光分子的干扰。

[来源：GB/T 46942.1-2025 5.8]

3.9 脉冲荧光信号 pulsed fluorescent signal

单分子实时荧光测序时，荧光染料标记的核苷酸进入检测区域并在聚合酶介导下发生连接反应延伸引物链，在此过程中染料分子被激发而发射的短时间持续的荧光信号。

[来源：GB/T 46942.1-2025 5.7]

3.10 实时测序 real-time sequencing

在连续测序的过程中，每一条待测核酸分子的测序信号被即时进行碱基识别和序列输出。

注：为单分子基因测序的技术特征。

[来源：GB/T 46942.1-2025 3.23]

3.11 直接测序 direct sequencing

原始模板核酸分子可不经任何扩增与转化处理，直接构建文库进行测序，此过程中核酸分子上的碱基修饰得以保留而被后续直接测定。

注：为单分子基因测序的技术特征。

[来源：GB/T 46942.1-2025 3.21]

3.12 测序通量 throughput of gene sequencing

单次测序可获得序列信息的基因片段数量或可测定的脱氧核糖核酸和核糖核酸（以碱基表示）数量。

[来源：GB/T 30989-2014, 3.21]

3.13 碱基识别 base calling

测序过程中从荧光信号或电学信号转换成序列信息的过程。

[来源：GB/T 30989-2014, 3.26]

3.14 碱基识别质量 quality of base calling

衡量碱基被正确识别的概率。通常以数字值直接表示。

碱基识别质量与碱基识别错误率之间的关系可用式（1）表示：

$$Q=-10 \lg P \quad (1)$$

式中：

Q ——碱基识别质量；

P ——碱基识别错误率。

[来源：GB/T 30989-2014, 3.29]

3.15 碱基识别错误率 inaccuracy of base calling

单次测序错误碱基数占碱基总数的比例。

[来源：GB/T 30989-2014, 3.28]

3.16 Q10

测序数据中，序列识别准确率或碱基识别准确率为 90%的测序识别质量值表示为 Q10。

[来源：GB/T 40226—2021, 3.5, 有修改]

3.17 碱基测序准确率 base sequencing accuracy

对已知序列的参考品进行测序,经过碱基识别后比对到已知序列上,统计比对正确的碱基数占测序获得的总碱基数比例。

[来源：GB/T 35537-2017, 3.1.4]

3.18 测序覆盖率 coverage rate of sequencing

待测样本的核苷酸序列检测结果覆盖于参考序列上的比例（测序覆盖率=覆盖区域长度÷参考序列总长度）。

[来源：GB/T 30989-2014, 3.30]

3.19 测序深度 depth of coverage

待测样本中某个指定的核苷酸被检测的次数。

[来源：GB/T 30989-2014, 3.31]

3.20 平均覆盖深度 average depth of coverage

测序深度的平均值。

[来源：GB/T 30989-2014, 3.32]

3.21 读长 N50 read length N50

将单次运行获得的质量合格的序列片段由长至短进行排序并依次相加，当相加的长度达到总长度一半时加上的最后一条片段的长度，以碱基数表示。

[来源：GB/T 46942.1-2025 4.3]

3.22 参考数据库 reference database

已经测序、分析过或完全已知的序列数据的集合，通常用于比对、分析其他个体的序列信息等。

[来源：JJF 2396—2026, 3.20]

3.23 比对率 mapping rate

比对到参考序列上的 reads 数目除以总测序数据的 reads 数目。

[来源：JJF 2396—2026, 3.18]

3.24 序列相对丰度 sequence relative abundance

某一指定序列在总样本中所占的相对比例。

序列相对丰度可用式（2）表示：

$$N_{\text{相对丰度}} = \frac{n_i}{\sum_{i=1}^n n_i} \times 100\% \quad (2)$$

式中：

n_i ——比对到 i 序列的 reads 数与 i 序列的长度的比；

N ——序列相对丰度。

[来源：JJF 2396—2026，3.19]

3.25 归一化读段数 normalized read count

比对到某参考序列的读段数除以该参考序列长度，用于消除序列长度差异对丰度计算的影响。

3.26 等长参考序列相对丰度比 relative abundance ratio of equal-length reference sequences

两条长度相同的参考序列，其归一化读段数的比值。

3.27 一致性序列 consensus sequence

将同一参考序列的多条测序读段（reads）进行比对后，在每个位置选取出现频率最高的碱基所组成的序列。

注：一致性序列用于与参考序列比对，以计算测序一致序列准确率。

4 概述

单分子测序技术能够直接读取单个 DNA 分子的序列，无需传统测序中常见的 PCR 扩增步骤，从而有效避免因扩增偏差导致的测序误差。这项技术通过检测单分子的电信号（如单分子纳米孔链测序法中的离子电流变化）或光信号（如单分子实时荧光测序法中脉冲荧光信号），实现对 DNA 分子的逐碱基解读。相比于传统短读长测序技术，单分子测序能够以每个 DNA 分子为单位单独测序，因此不仅提供更长的读长（经过酶工程和算法的不断迭代，目前通常达到 10kb~20 kb 或更长），还显著提高了在复杂基因组中解析结构变异和重复序列的能力。

单分子基因测序仪（以下简称测序仪）由多个高精度组件协同工作，构成一个完整且复杂的系统。主机部分包括样本制备模块、测序芯片和平台、光学或电学信号检测系

统，以及数据采集与处理系统；后续的信号解码和分析依赖计算与算法软件支持。此外，测序仪还配备基因测序仪控制软件，用于整合实验设计、信号处理和结果分析的流程。测序仪结构示意图如图 1 所示。

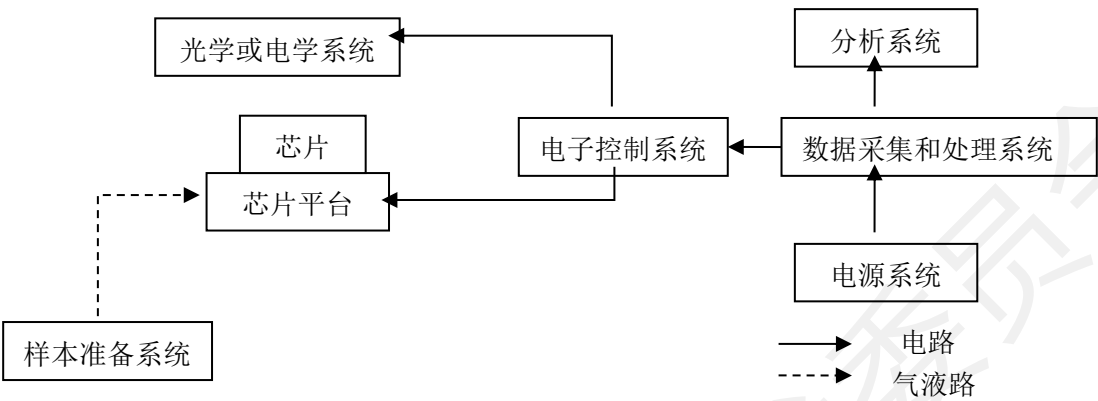


图 1 单分子测序仪结构示意图

5 计量特性

测序仪的计量特性，如表 1 所示。

表 1 测序仪的校准项目

序号	计量特性	计量指标
1	reads 数重复性	≤5%
2	比对率重复性	≤2%
3	序列覆盖率重复性	≤5%
4	序列相对丰度偏差	±3%
5	等长参考序列相对丰度比偏差	±10%
6	读长 N50 偏差	±5%
7	GC 含量偏差	±3%

8	测序一致序列准确率	≥99.9%
9	单次测序众数准确度	≥98%
10	序列识别质量百分比 Q10	>90%

注：

- 1、以上参考指标不适用于合格判定，仅供参考。
- 2、测序数据量应确保每个条形码重复组中，每条参考序列的平均覆盖深度不低于 10,000×。若标准物质说明书对数据量有更严格的要求，以说明书为准。
- 3、序列相对丰度偏差和等长参考序列相对丰度比偏差的判定对象分别为每条参考序列和每个长度梯度。

6 校准条件

6.1 环境条件

校准前需将测序仪开机预热 30 min，在下述适用的环境条件下，确保仪器达到正常工作状态。

6.1.1 环境温度：（18~25）℃。

6.1.2 相对湿度：10%~80%。

6.1.3 电源：100VAC~240 VAC，频率为 50/60 Hz。

注：6.1.1、6.1.2、6.1.3 中条件与制造商的产品规定不一致时，以产品规定为准。

如有特殊条件保证仪器正常工作，参考仪器设备安装要求。

6.2 校准用标准物质试剂

6.2.1 标准物质：校准时应采用来源稳定、溯源信息完备、参考序列信息已知的片段化 DNA 序列有证标准物质。该标准物质应至少包含两个长度梯度的线性双链 DNA 参考序列，每个长度梯度应至少包含两条序列不同的片段，最短参考序列长度宜不小于 1 kb。各片段应按已知且可溯源的拷贝数比例混合，并提供各参考序列的拷贝数浓度标准值及其不确定度。标准物质应设计为可直接用于测序建库，以尽量减少建库步骤引入的偏差。标准物质的浓度、纯度及完整性应满足测序建库的一般要求。标准物质的选择原则参照附录 A。

6.2.2 试剂：配套测序仪相关试剂盒。

6.2.3 配制校准用标准物质所需超纯水：电阻率不小于 18.2 MΩ·cm（25 °C）。

6.3 校准用设备及分析软件

6.3.1 移液器：取样范围为（1~1000）μL，经计量检定合格。

6.3.2 分析软件：推荐使用 Linux 操作系统、Minimap2、DeepVariant、MinKNOW、Dorado、EPI2ME 等标准统计分析软件。建议使用最新、通用或已验证版本。

注：6.3.1 移液器技术要求根据不同量程移液器要求执行。

7 校准项目和校准方法

将测序仪开机预热 30 min，确保仪器达到正常工作状态后，按照测序仪上机流程对标准物质进行测序。当标准物质采用内置条形码设计时，可直接取三个条形码组的数据作为 3 次重复测量结果；否则应重复测定 3 次。将下机数据按条形码拆分后，使用分析软件处理，与参考序列比对并进行统计分析，按照下面的校准方法开展校准。

7.1 reads 数重复性

对每个重复组（或每次测序），统计总 reads 数。根据公式（1）计算测序 reads 数的相对标准偏差。

$$RSD_{reads} = \frac{R_{max} - R_{min}}{C_n} \times \frac{1}{\bar{R}} \times 100\% \quad (1)$$

式中：

RSD_{reads} ——测序总 reads 数的相对标准偏差，以百分数表示；

R_{max} ——3 次测定结果的最大值；

R_{min} ——3 次测定结果的最小值；

$\bar{R}$ ——3 次测定结果的平均值；

C_n ——极差系数， $n=3$ 时， $C_n=1.69$ 。

7.2 比对率重复性

对每个重复组（或每次测序），按公式（2）计算比对率。

$$R_{\text{比对率}} = \frac{n_{\text{map}}}{N_{\text{总reads}}} \times 100\% \quad (2)$$

式中：

$R_{\text{比对率}}$ ——测序比对率，以百分数表示；

n_{map} ——比对到参考序列上的 reads 数目；

$N_{\text{总 reads}}$ ——总测序数据的 reads 数目。

根据公式（3）计算比对率的重复性。

$$RSD_{\text{比对率}} = \frac{R_{\text{比对率max}} - R_{\text{比对率min}}}{C_n} \times \frac{1}{\bar{R}_{\text{比对率}}} \times 100\% \quad (3)$$

式中：

$RSD_{\text{比对率}}$ ——测序比对率的重复性，以百分数表示；

$R_{\text{比对率max}}$ ——3 次测定比对率结果的最大值；

$R_{\text{比对率min}}$ ——3 次测定比对率结果的最小值；

$\bar{R}_{\text{比对率}}$ ——3 次测定比对率结果的平均值；

C_n ——极差系数， $n=3$ 时， $C_n=1.69$ 。

7.3 序列覆盖率重复性

对每个参考序列，分别计算每个重复组（或每次测序）的覆盖率，按公式（4）计算。

$$L_{\text{覆盖率}} = \frac{L_{\text{覆盖长度}}}{L_{\text{总长}}} \times 100\% \quad (4)$$

式中：

$L_{\text{覆盖率}}$ ——序列覆盖率，以百分数表示；

$L_{\text{覆盖长度}}$ ——覆盖区域长度；

$L_{\text{总长}}$ ——参考序列总长度。

根据公式（5）计算序列覆盖率的重复性。

$$RSD_{\text{覆盖率}} = \frac{L_{\text{覆盖率max}} - L_{\text{覆盖率min}}}{C_n} \times \frac{1}{\bar{L}_{\text{覆盖率}}} \times 100\% \quad (5)$$

式中：

$RSD_{\text{覆盖率}}$ ——序列覆盖率的重复性，以百分数表示；

$L_{\text{覆盖率 max}}$ ——3 次测定覆盖率结果的最大值；

$L_{\text{覆盖率 min}}$ ——3 次测定覆盖率结果的最小值；

$\bar{L}_{\text{覆盖率}}$ ——3 次测定覆盖率结果的平均值；

C_n ——极差系数， $n=3$ 时， $C_n=1.69$ 。

7.4 序列相对丰度偏差

对每个重复组，统计比对到各参考序列的 reads 数。按公式（6）计算各参考序列的归一化 reads 数。

$$n_i = \frac{N_i}{L_i} \quad (6)$$

式中：

n_i ——第 i 条参考序列的归一化 reads 数；

N_i ——比对到第 i 条参考序列的 reads 数；

L_i ——第 i 条参考序列的长度。

按公式（7）计算各参考序列的相对丰度。

$$N_{\text{相对丰度}} = \frac{n_i}{\sum_{i=1}^n n_i} \times 100\% \quad (7)$$

式中：

$N_{\text{相对丰度}}$ ——第 i 条参考序列的相对丰度，以百分数表示。

按公式（8）计算序列相对丰度偏差。

$$\Delta N = \bar{N}_{\text{相对丰度}} - N_S \quad (8)$$

式中：

ΔN ——目标序列相对丰度偏差，以百分数表示；

$\bar{N}_{\text{相对丰度}}$ ——3 次测量目标序列相对丰度的平均值；

N_S ——目标序列相对丰度的标准值。

7.5 等长参考序列相对丰度比偏差

对每个长度梯度（如 5 kb 组、10 kb 组），分别计算两条等长参考序列的归一化 reads 数比值。按公式（9）计算。

$$R_{\text{等长}} = \frac{n_i}{n_j} \quad (9)$$

式中：

$R_{\text{等长}}$ ——等长参考序列相对丰度比；

n_i 、 n_j ——分别为同一长度梯度内两条等长参考序列的归一化 reads 数。

按公式（10）计算等长参考序列相对丰度比偏差。

$$\Delta R = |R_{\text{等长}} - 1| \times 100\% \quad (10)$$

式中：

ΔR ——等长参考序列相对丰度比偏差，以百分数表示。

取三个重复组的平均值作为最终结果。

7.6 读长 N50 偏差

对每条参考序列，分别提取比对到该序列的所有 reads。将 reads 按长度从大到小排序，累加长度，当累计长度达到该组 reads 总长度的 50% 时，对应的 read 长度即为 N50。按公式（11）计算读长 N50 偏差。

$$\Delta N50 = \frac{|N50_{\text{实测}} - N50_{\text{标称}}|}{N50_{\text{标称}}} \times 100\% \quad (11)$$

式中：

$\Delta N50$ ——读长 N50 偏差，以百分数表示；

$N50_{\text{实测}}$ ——实测 N50 值，单位为 bp；

$N50_{\text{标称}}$ ——参考序列的标称长度，单位为 bp。

取三个重复组的平均值作为最终结果。

7.7 GC 含量偏差

对每个重复组，统计所有测序碱基中 G 和 C 的数量，按公式（12）计算测序 GC 含量。

$$GC\% = \frac{n_{GC}}{N_{\text{碱基}}} \times 100\% \quad (12)$$

式中：

$GC\%$ ——测序 GC 含量，以百分数表示；

n_{GC} ——测序碱基中 G 和 C 的碱基数；

$N_{\text{碱基}}$ ——测序获得的总碱基数。

按公式（13）计算 GC 含量偏差。

$$\Delta GC\% = \overline{GC\%} - GC\%_S \quad (13)$$

式中：

$\Delta GC\%$ ——GC 含量偏差，以百分数表示；

$\overline{GC\%}$ ——3 次测序 GC 含量的平均值；

$GC\%_S$ ——标准物质的 GC 含量标准值。

7.8 测序一致序列准确度

使用标准物质进行测序，通过比对软件与参考序列比对，在达到足够覆盖度（如 100×）下，生成一致性序列。统计一致性序列与参考序列比对时的匹配碱基数和错配碱基数。按公式（14）计算测序一致序列准确率。

$$AC = \left(1 - \frac{MS}{MB}\right) \times 100\% \quad (14)$$

式中：

AC ——测序一致序列准确率，以百分数表示；

MS ——指定目标区域错配碱基数（包含置换、插入和缺失）；

MB ——指定目标区域总碱基数。

取三个重复组的平均值作为最终结果。

7.9 单次测序众数准确度

使用标准物质进行测序，将每条 read 与参考序列进行比对，按公式（15）计算每条 read 的准确率。

$$Accuracy = \frac{Matches}{Matches+Sub+Ins+Del} \times 100\% \quad (15)$$

式中：

$Accuracy$ ——单条 read 准确率，以百分数表示；

$Matches$ ——正确匹配的碱基总数；

Sub 、 Ins 、 Del ——分别为替换、插入和缺失的碱基数。

统计所有 read 准确率的直方图，取最高峰对应的准确率值作为众数准确度。取三个重复组的平均值作为最终结果。

7.10 序列识别质量百分比（Q10）

统计所有碱基的质量值，按公式（16）计算 Q10 比例。

$$Q10 = \frac{n_{Q10 \text{ 碱基}}}{N_{\text{总碱基}}} \times 100\% \quad (16)$$

式中：

Q10——质量值≥10 的碱基占比，以百分数表示；

$n_{Q10 \text{ 碱基}}$ ——质量值≥10 的碱基数；

$N_{\text{总碱基}}$ ——测序获得的总碱基数。

取三个重复组的平均值作为最终结果。

8 校准结果表达

校准记录应尽可能详尽地记载测量数据和计算结果，推荐的校准记录格式见附录 B。各测量结果的测量不确定度应按 JJF 1059.1 的要求评定，不确定度评定示例见附录 D、E、F。经校准的测序仪应出具校准证书，校准证书应符合 JJF 1071—2010 中 5.12 的要求。

9 复校时间间隔

由于复校时间间隔的长短是由测序仪的使用情况、使用者、仪器本身质量等诸因素所决定的，因此，送校单位可根据实际使用情况自主决定复校时间间隔，建议不超过 1 年。

附录 A

标准物质的选择原则

A.1 标准物质的设计与选择

A.1.1 应采用来源稳定、溯源信息完备、参考序列信息已知的片段化线性双链 DNA 有证标准物质。标准物质应包含至少两个长度梯度的参考序列，每个长度梯度应至少包含两条序列不同的片段。

A.1.2 所有参考序列应避开影响测序准确率评估的特征区域，包括但不限于连续同聚物、串联重复、散在重复、可形成稳定二级结构的区域以及低复杂度区域。

A.2 序列与丰度要求

A.2.1 参考序列的 GC 含量应处于适宜范围，避免极端 GC 含量对测序效率的干扰。

A.2.2 同一长度梯度的双序列应具有足够的序列差异度，确保比对分析时 reads 能够无歧义地归属到正确的参考序列。

A.2.3 样本内各序列的相对丰度标准值应由绝对定量方法测定，其扩展不确定度应满足校准需求。

A.2.4 同长度双序列的标称丰度比应几乎一致，用于评价仪器对序列本身的偏好性。

A.3 浓度、纯度与完整度

A.3.1 标准物质的 DNA 浓度应满足测序平台的建库和上样要求。

A.3.2 DNA 纯度应满足测序建库的一般要求，无蛋白质、RNA、盐离子及有机溶剂污染。

A.3.3 片段的完整度应满足校准要求，不应有明显的降解或断裂。

A.4 储存与操作

A.4.1 标准物质应分装保存，避免反复冻融。

A.4.2 使用过程中应轻柔操作，禁止涡旋振荡。

A.5 数据质量要求

A.5.1 测序数据量应满足标准物质最低分析数据量的要求，确保各重复组均有足够的 reads 数支撑统计计算。

A.5.2 拆分后的各重复组数据应分别进行比对和统计分析，比对软件和参数应事先经过验证。

附录 B

校准原始记录参考格式

客户名称	
器具名称	
型号/规格	
出厂编号	
生产厂商	
客户地址	
校准日期	
校准环境条件及地点	
地点	
温度	
湿度	
其他	

基础信息	
设备最大通量	
测序试剂型号	
数据量	

一、读长总数 (reads) 重复性

	测定值			平均值	重复性
	1	2	3		
reads					

二、比对率重复性

	测定值			平均值	重复性
	1	2	3		
比对率					

三、序列覆盖率重复性

	测定值			平均值	重复性
	1	2	3		

覆盖率					
-----	--	--	--	--	--

四、序列相对丰度偏差

参考序列	测定值			平均值	相对偏差	不确定度 $U(k=2)$
	1	2	3			
梯度 1-序列 1						
梯度 1-序列 2						
梯度 2-序列 1						
梯度 2-序列 2						

五、等长参考序列相对丰度比偏差

长度梯度	测定值			平均值	相对偏差	不 确 定 度 $U(k=2)$
	1	2	3			
梯度 1						
梯度 2						

六、读长 N50 偏差

参考序列	测定值			平均值	相对偏差
	1	2	3		
梯度 1-序列 1					
梯度 1-序列 2					
梯度 2-序列 1					
梯度 2-序列 2					

七、GC 含量偏差

	测定值			平均值	相对偏差	不确定度 $U(k=2)$
	1	2	3			
GC%						

八、测序一致序列准确率

参考序列	测定值			平均值	相对偏差
	1	2	3		
梯度 1-序列 1					

梯度 1-序列 2					
梯度 2-序列 1					
梯度 2-序列 2					

九、单次测序众数准确度

	测定值			平均值
	1	2	3	
众数准确度				

十、序列识别质量百分比（Q10）

	测定值			平均值
	1	2	3	
Q10 比例				

附录 C

校准证书(内页) 参考格式

共 页，第 页

序号	校准项目	校准结果
1	reads 数重复性	
2	比对率重复性	
3	序列覆盖率重复性	
4	序列相对丰度偏差	，不确定度：（ $k=2$ ）
5	等长参考序列相对丰度比偏差	，不确定度：（ $k=2$ ）
6	读长 N50 偏差	
7	GC 含量偏差	，不确定度：（ $k=2$ ）
8	测序一致序列准确率	
9	单次测序众数准确度	
10	序列识别质量百分比（Q10）	

校准员：_____

核验员：_____

附录 D 序列相对丰度偏差测量结果的不确定度评定示例

D.1 测量模型

序列相对丰度偏差按公式 (D.1) 计算：

$$\Delta N = \bar{N} - N_S \quad (\text{D.1})$$

式中：

ΔN ——序列相对丰度偏差，以百分数表示；

$\bar{N}$ ——3 次序列相对丰度测定结果的算术平均值，以百分数表示；

N_S ——标准物质的序列相对丰度标准值，以百分数表示。

D.2 不确定度来源

不确定度来源包括：

a) 测序仪测量重复性引入的标准不确定度 $u(\bar{N})$ ；

b) 标准物质引入的标准不确定度 $u(N_S)$ 。

D.3 标准不确定度分量的评定

D.3.1 测量重复性引入的标准不确定度分量 $u(\bar{N})$

在重复性条件下连续测量 3 次，以 3 次测量的算术平均值作为结果。实验标准偏差 s_N 按公式 (D.2) 计算：

$$s_N = \frac{N_{\max} - N_{\min}}{C_n} \quad (\text{D.2})$$

式中：

N_{max} ——3 次测定结果的最大值，以百分数表示；

N_{min} ——3 次测定结果的最小值，以百分数表示；

C_n ——极差系数， $n=3$ 时， $C_n=1.69$ 。

标准不确定度分量按公式 (D.3) 计算：

$$u(\bar{N}) = \frac{s_N}{\sqrt{3}} \quad (\text{D.3})$$

数据见表 D.1。

表 D.1 测量重复性引入的标准不确定度分量数据

测量次数	N_i (%)	$\bar{N}$ (%)	s_N (%)	$u(\bar{N})$ (%)
1	25.10			
2	24.98	25.04	0.07	0.04
3	25.04			

D.3.2 标准物质引入的标准不确定度分量 $u(N_S)$

由标准物质证书提供的扩展不确定度 $U(N_S)$ 和包含因子 k ，按公式 (D.4) 计算：

$$u(N_S) = \frac{U(N_S)}{k} \quad (\text{D.4})$$

本例中，标准物质证书给出该序列相对丰度标准值 $N_S=25.0\%$ ，扩展不确定度 $U(N_S)=2.0\%$ ， $k=2$ ，则

$$u(N_S) = \frac{2.0\%}{2} = 1.0\%$$

D.4 合成标准不确定度

按公式 (D.5) 计算合成标准不确定度：

$$u_{\Delta N} = \sqrt{u(\bar{N})^2 + u(N_S)^2} \quad (\text{D.5})$$

代入数据：

$$u_{\Delta N} = \sqrt{0.04^2 + 1.0^2} \approx 1.0\%$$

D.5 扩展不确定度

取包含因子 $k=2$ ，扩展不确定度按公式 (D.6) 计算：

$$U_{\Delta N} = 2 \times u_{\Delta N} = 2 \times 1.0\% = 2.0\% \quad (\text{D.6})$$

结果见表 D.2。

表 D.2 合成标准不确定度和扩展不确定度

项目	$u_{\Delta N}$ (%)	$U_{\Delta N}$ (%)	k
序列相对丰度偏差	1.0	2.0	2

附录 E 等长参考序列相对丰度比偏差测量结果的不确定度评定示例

E.1 测量模型

对某一长度梯度，等长参考序列相对丰度比偏差按公式（E.1）计算：

$$\Delta R = |R - 1| \times 100\% \quad (\text{E.1})$$

其中，比值 R 按公式（E.2）计算：

$$R = \frac{n_i}{n_j} \quad (\text{E.2})$$

式中：

ΔR ——等长参考序列相对丰度比偏差，以百分数表示；

R ——两条等长参考序列的归一化 reads 数比值；

n_i 、 n_j ——分别为同一长度梯度内两条等长参考序列的归一化 reads 数。

E.2 不确定度来源

不确定度来源包括：

- a) 测序仪测量重复性引入的标准不确定度 $u(R)$ ；
- b) 标称比值（1:1）引入的标准不确定度 $u(R_S)$ ，通常可忽略。

E.3 标准不确定度分量的评定

E.3.1 测量重复性引入的标准不确定度分量 $u(R)$

在重复性条件下连续测量 3 次，分别计算每次的比值 R ，以 3 次测量的算术平均值作为结果。实验标准偏差 s_R 按公式（E.3）计算：

$$s_R = \frac{R_{\max} - R_{\min}}{C_n} \quad (\text{E.3})$$

式中：

$R_{\max}$ ——3 次比值结果的最大值；

$R_{\min}$ ——3 次比值结果的最小值；

C_n ——极差系数， $n=3$ 时， $C_n=1.69$ 。

标准不确定度分量按公式 (E.4) 计算：

$$u(R) = \frac{s_R}{\sqrt{3}} \quad (\text{E.4})$$

数据见表 E.1。

表 E.1 测量重复性引入的标准不确定度分量数据

测量次数	R_i	$\bar{R}$	s_R	$u(R)$
1	1.0085			
2	1.0080	1.0081	0.0005	0.0003
3	1.0078			

将 $u(R)$ 转换为百分数表示：

$$u(R) = 0.0003 \times 100\% = 0.03\%$$

E.3.2 标称比值引入的标准不确定度分量 $u(R_S)$

标称比值为 1:1，其不确定度可忽略，故 $u(R_S) \approx 0$ 。

E.4 合成标准不确定度

按公式 (E.5) 计算合成标准不确定度：

$$u_{\Delta R} = \sqrt{u(R)^2 + u(R_S)^2} \approx u(R) \quad (\text{E.5})$$

代入数据：

$$u_{\Delta R} \approx 0.03\%$$

E.5 扩展不确定度

取包含因子 $k=2$ ，扩展不确定度按公式（E.6）计算：

$$U_{\Delta R} = 2 \times u_{\Delta R} = 2 \times 0.03\% = 0.06\% \quad (\text{E.6})$$

同时，实测比值偏差为：

$$\Delta R = |1.0081 - 1| \times 100\% = 0.81\%$$

结果见表 E.2。

表 E.2 合成标准不确定度和扩展不确定度

项目	ΔR (%)	$u_{\Delta R}$ (%)	$U_{\Delta R}$ (%)	k
等长参考序列相对 丰度比偏差	0.81	0.03	0.06	2

附录 F GC 含量偏差测量结果的不确定度评定示例

F.1 测量模型

GC 含量偏差按公式 (F.1) 计算：

$$\Delta GC\% = \overline{GC\%} - GC\%_S \quad (F.1)$$

式中：

$\Delta GC\%$ ——GC 含量偏差，以百分数表示；

$\overline{GC\%}$ ——3 次 GC 含量测定结果的算术平均值，以百分数表示；

$GC\%_S$ ——标准物质的 GC 含量标准值，以百分数表示。

F.2 不确定度来源

不确定度来源包括：

a) 测序仪测量重复性引入的标准不确定度 $u(\overline{GC\%})$ ；

b) 标准物质引入的标准不确定度 $u(GC\%_S)$ 。

F.3 标准不确定度分量的评定

F.3.1 测量重复性引入的标准不确定度分量 $u(\overline{GC\%})$

在重复性条件下连续测量 3 次，以 3 次测量的算术平均值作为结果。实验标准偏差 $s_{GC\%}$ 按公式 (F.2) 计算：

$$s_{GC\%} = \frac{GC\%_{max} - GC\%_{min}}{C_n} \quad (F.2)$$

式中：

$GC\%_{max}$ ——3 次测定结果的最大值，以百分数表示；

$GC\%_{min}$ ——3 次测定结果的最小值，以百分数表示；

C_n ——极差系数， $n=3$ 时， $C_n=1.69$ 。

标准不确定度分量按公式 (F.3) 计算：

$$u(\overline{GC\%}) = \frac{s_{GC\%}}{\sqrt{3}} \quad (F.3)$$

数据见表 F.1。

表 F.1 测量重复性引入的标准不确定度分量数据

测量次数	$GC\%_i$ (%)	$\overline{GC\%}$ (%)	$s_{GC\%}$ (%)	$u(\overline{GC\%})$ (%)
1	45.7			
2	45.8	45.8	0.1	0.06
3	45.9			

F.3.2 标准物质引入的标准不确定度分量 $u(GC\%_S)$

由标准物质证书提供的扩展不确定度 $U(GC\%_S)$ 和包含因子 k ，按公式 (F.4) 计算：

$$u(GC\%_S) = \frac{U(GC\%_S)}{k} \quad (F.4)$$

本例中，标准物质证书给出 GC 含量标准值 $GC\%_S=45.8\%$ ，扩展不确定度 $U(GC\%_S)=0.5\%$ ， $k=2$ ，则

$$u(GC\%_S) = \frac{0.5\%}{2} = 0.25\%$$

F.4 合成标准不确定度

按公式 (F.5) 计算合成标准不确定度：

$$u_{\Delta GC\%} = \sqrt{u(\overline{GC\%})^2 + u(GC\%_S)^2} \quad (\text{F.5})$$

代入数据：

$$u_{\Delta GC\%} = \sqrt{0.06^2 + 0.25^2} \approx 0.26\%$$

F.5 扩展不确定度

取包含因子 $k=2$ ，扩展不确定度按公式 (F.6) 计算：

$$U_{\Delta GC\%} = 2 \times u_{\Delta GC\%} = 2 \times 0.26\% = 0.52\% \quad (\text{F.6})$$

结果见表 F.2。

表 F.2 合成标准不确定度和扩展不确定度

项目	$u_{\Delta GC\%}$ (%)	$U_{\Delta GC\%}$ (%)	k
GC 含量偏差	0.26	0.52	2