



中华人民共和国国家计量技术规范

JJF xxxx—202x

高通量基因组测序系统 分析性能确认和验证技术规范

Technical specification for analysis performance validation and verification
of next generation genome sequencing systems

(征求意见稿)

xxxx - xx - xx 发布

xxxx - xx - xx 实施

国家市场监督管理总局 发布

高通量基因组测序系统分析性能确认和验证技术规范

JJF xxxx—202x

**Technical specification for analysis performance validation and verification
of next generation genome sequencing systems**

归口单位:全国生物计量技术委员会

主要起草单位:

复旦大学

中国计量科学研究院

参加起草单位:

本规范委托全国生物计量技术委员会负责解释

本规范主要起草人：

参加起草人：

全国生物计量技术委员会

目 录

引言	II
1 范围	1
2 引用文件	1
3 术语、定义和缩略语	1
3.1 术语和定义	1
3.2 缩略语	3
4 概述	4
4.1 高通量基因测序系统的测量原理	4
4.2 测量系统组成	4
5 确认与验证的指标	4
5.1 确认指标	4
5.2 验证指标	4
6 工作条件	5
6.1 环境条件	5
6.2 验证或确认用仪器及试剂	5
7 确认与验证时机	5
7.1 确认时机	5
7.2 验证时机	5
8 确认程序	6
8.1 文库质量	6
8.2 数据质量控制	6
8.3 基因变异类型准确度	7
8.4 检出限	8
8.5 分析特异性	8
8.6 可报告范围	8
9 验证程序	9
9.1 基因组变异类型准确度	9
9.2 检出限	9
9.3 可报告范围	9
10 性能确认/验证结果表达	9
11 性能确认/验证时间间隔	9
附录 A 性能确认原始记录格式	10
附录 B 性能验证原始记录格式	14
参考文献	16

引言

高通量测序技术（High-throughput sequencing），又称下一代基因测序（Next Generation Sequencing, NGS）已经成为生物医学研究的重要工具。为了确保测序数据的准确性和可靠性，必须对高通量基因组测序系统进行分析性能确认和验证。本规范旨在为高通量基因组测序系统的分析性能确认和验证提供统一的计量技术规范。

本规范的编制参考了 JJF 1071《国家计量校准规范编写规则》和 JJF 1059.1《测量不确定度评定与表示》的基础性系列规范。主要性能指标参考了 WS/T 505《定性测定性能评价指南》、WS/T 514《临床检验方法检出能力的确立和验证》、WS/T 416《干扰实验指南》、以及 CNAS-GL039《分子诊断检验程序性能验证指南》中相关内容。

本规范为首次发布。

高通量基因组测序系统分析性能确认和验证技术规范

1 范围

本规范适用于以人基因组 DNA 样品为测量对象的高通量基因测序（以下简称 NGS）测量系统的分析性能确认和验证。测量结果包括基因变异类型和/或变异等位基因频率等。

高通量基因组测序分析系统，需要根据系统预期用途和最终产生的测量结果进行分析性能确认和验证。

2 引用文件

本规范引用了下列文件：

GB/T 30989 高通量基因测序技术规程

GB/T 35537 高通量基因测序结果评价要求

WS/T 505 定性测定性能评价指南

WS/T 514 临床检验方法检出能力的确立和验证

WS/T 416 干扰实验指南

CNAS-GL039 分子诊断检验程序性能验证指南

ISO 9000 质量管理体系-基本原则和术语

ISO 15189 医学实验室质量和能力认可准则

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，其最新版本适用于本规范。

3 术语、定义和缩略语

3.1 术语和定义

3.1.1 高通量测序 next generation sequencing

能一次并行几十万到几百万条核酸分子序列测定和一般读长较短等为标志，适用于 DNA 的测序技术。

[GB/T 35890-2018, 3.1]

3.1.2 确认 validation

通过提供客观证据对特定的预期使用或应用要求已得到满足的认定。

[ISO 9000, 3.8.5]

3.1.3 验证 verification

通过提供客观证据对规定要求已得到满足的认定。[ISO 9000, 3.8.4]

3.1.4 测量精密度 measurement precision

精密度

在规定条件下,对同一或类似被测对象重复测量所得示值或测得值间的一致程度。

注:

- 1 测量精密度通常在规定测量条件下用不精密程度以数字形式表示。定性测量精密度为在规定测量条件下检测到的真变异结果的个数占全部变异结果个数的百分比的变异系数。
- 2 规定条件可以是重复性测量条件,期间精密度测量条件或复现性测量条件。
- 3 测量精密度用于定义测量重复性、期间精密度或测量复现性。
- 4 “测量精密度”不等同于“测量准确度”。

[JJF 1001, 5.10, 注有修改]

3.1.5 期间测量精密度 intermediate measurement precision

在一组期间精密度测量条件下的测量精密度。

[JJF 1001, 5.12]

3.1.6 测量准确度 measurement accuracy

被测量的测得值与其真值间的一致程度。

注:

- 1 定性测量(基因变异类型)结果的准确度以符合率和精密度表示。

[JJF 1001, 5.8, 注有修改]

3.1.7 标准参考数据 standard reference data

用于评估测序数据结果可靠性和准确性的对比参考数据,其本身需要不同方法的严格验证。

3.1.8 阳性符合率 true positive rate

检出且判断为阳性的变异个数占标准参考数据集中所有阳性变异个数的百分比。

3.1.9 阴性符合率 true negative rate

检出且判断为阴性的样本个数占标准参考数据集中所有阴性样本个数的百分比。

3.1.10 检出限 limit of detection

由给定测量程序获得的测得值，其声称的物质成分不存在的误判概率为 β ，声称的物质成分存在的误判概率为 α 。

注：

- 1 国际理论和应用化学联合会(IUPAC)推荐的 α 和 β 的默认值为 0.05。
- 2 也被称作检测低限、最小可检测浓度(或值)，缩写为 LOD。

[JJF 1001, 7.18]

3.1.11 分析特异性 analytical specificity

一种测量程序只对其旨在测量的可测量进行确认的能力。

[GB/T 37871, 3.1.5]

3.1.12 基因组变异类型 genomic variant types

基于分子特征和功能影响对基因序列变化进行的系统性分类，涵盖单核苷酸变异（SNV）、插入缺失（Indel）、拷贝数变异（CNV）等。

3.1.13 变异等位基因频率 variant allele frequency

在特定基因组位点上，支持某一突变的测序读数（reads）占该位点总有效测序读数（野生型+变异型）的百分比。

3.1.14 可报告范围 reportable range

能够可靠地检测基因变异类型以及特定基因组区域。

3.2 缩略语

下列缩略语适用于本规范。

NGS：高通量测序/下一代基因测序（Next Generation Sequencing）

RM：标准物质（Reference Material）

CRM：有证标准物质（Certified Reference Material）

LOD：检出限（Limit of Detection）

CV：变异系数（Coefficient of Variation）

VAF：变异等位基因频率（Variant Allele Frequency）

TP：真阳性（True Positive）

FP: 假阳性 (False Positive)

TN: 真阴性 (True Negative)

FN: 假阴性 (False Negative)

SRD: 标准参考数据 (Standard Reference Data)

DNA: 脱氧核糖核酸 (Deoxyribonucleic Acid)

SNV: 单核苷酸变异 (Single Nucleotide Variant)

INDEL: 短插入和短缺失变异 (Insertion and Deletion)

CNV: 拷贝数变异 (Copy Number Variant)

SV: 结构变异 (Structure Variant)

4 概述

4.1 高通量基因测序系统的测量原理

基于大规模并行测序技术，其核心是通过将 DNA 分子随机片段化并连接特异性接头后，固定在固体载体（如流动槽或纳米微球）表面形成空间分离的克隆簇或 DNA 纳米球 (DNB)，随后通过边合成边测序的方法实现碱基序列的读取。在测序反应中，DNA 聚合酶将荧光标记的 dNTP 逐个掺入互补链，每轮反应仅延伸一个碱基，通过高分辨率成像系统捕获荧光信号并转化为碱基序列信息，循环重复此过程完成全序列测定。最终，海量短序列数据经生物信息学比对和组装，实现全基因组或靶向区域的高精度覆盖。

4.2 测量系统组成

高通量基因组测序测量系统的组成包括高通量测序建库试剂、高通量测序仪、数据分析软件、关键实验耗材等。该测量系统输出的检测结果包括基因变异类型和/或变异等位基因频率等。

5 确认与验证的指标

5.1 确认指标

包括基因组变异类型准确度（符合率、精密度）、检出限、分析特异性（包含交叉反应和干扰物质）、可报告范围（目标基因区域和变异类型等）。

5.2 验证指标

包括基因组变异类型准确度（符合率、精密度）、检出限、可报告范围（目

标基因区域和变异类型等)。

6 工作条件

6.1 环境条件

6.1.1 温度：(10~30)℃；

6.1.2 相对湿度：≤80%；

6.1.3 供电电源：电压(220±22)V，频率(50±1)Hz；

6.1.4 附近无影响仪器正常工作的电磁场及机械振动；

6.1.5 仪器接地良好；

6.1.6 其他：无冷凝水。

6.2 验证或确认用仪器及试剂

6.2.1 核酸标准物质：应采用经计量行政部门批准发布的有证标准物质，特性量值或标称特性值满足验证或确认的测量系统的预期用途。

6.2.2 微光分光光度计：经过校准，或用有证标准物质进行标定。

6.2.3 芯片电泳：测量范围为≥1 bp，灵敏度≤1ng/μL (DNA)，或用有证标准物质进行标定。

6.2.4 移液器：规格为 10 μL、100 μL、200 μL、1000 μL。

6.2.5 配制标准物质超纯水：【电阻率不小于 18.2 MΩ·cm (25 °C)】。

7 确认与验证时机

7.1 确认时机

7.1.1 测量系统在正式投入使用前；

7.1.2 实验室所使用的测量系统与试剂供应商所要求的测量系统不完全一致时；

7.1.3 任何严重影响测量系统分析性能的情况发生后，应在测量系统重新启用前对受影响的性能部分进行确认。影响测量系统分析性能的情况可包括但不限于更换设备型号、更换试剂批号等。

7.2 验证时机

7.2.1 新测量系统常规应用前，新测量系统也包含现用测量系统的任一要素（仪器、试剂等）变更，如试剂升级、仪器更新、校准品溯源性改变等应按照新系统来进行验证；

7.2.2 任何严重影响测量系统分析性能的情况发生后，应在测量系统重新启用前对受影响的性能部分进行验证。影响测量系统分析性能的情况可包括但不限于仪器主要部件故障、仪器搬迁、设施或环境的严重失控、试剂和关键耗材更换等；

7.2.3 常规使用期间，实验室可基于分析系统的稳定性，利用日常工作产生的检验和质控数据，定期对检验程序的分析性能进行评估，应能满足检验结果预期用途的要求。

8 确认程序

8.1 文库质量

文库浓度应符合质量要求，确保文库构建成功。

8.2 数据质量控制

选择合适的数据质量控制软件，对测序数据进行质量控制。控制流程包括但不限于去除重复序列、去除低质量数据、去除接头序列等。合格的原始测序数据应满足以下条件：

——碱基质量值大于等于 30（即 Q30：碱基识别错误率不大于 0.1%）的占比在 85%以上；

——一条序列中不明确的碱基个数要小于 5 个；

——接头序列污染比例不超过 1%；

——过滤低质量以及接头序列等后的剩余序列平均长度大于 120 bp。

为了确保数据的准确性和可靠性，比对后的数据需要满足特定的质量控制要求，具体要求如表 1 所示。

文件格式：建议使用通用的文件格式，如比对结果使用“.bam”文件或其他，测序结果使用“.fastq”文件或其他。如使用内部的标准格式文件，应建立该标准格式的详细说明，并注明该文件格式与其它格式的兼容性及相互转换的方法。

表 1 比对后的数据质量控制要求

测序类型	平均覆盖度	比对率	重复率	Q30 碱基占比
全基因组测序	$\geq 30\times$	$\geq 95\%$	$\leq 15\%$	$\geq 85\%$
全外显子测序	$\geq 100\times$	$\geq 98\%$	$\leq 20\%$	$\geq 85\%$
靶向测序 (组织)	$\geq 500\times$ (目标区域)	$\geq 99\%$	$\leq 50\%$	$\geq 85\%$

靶向测序 (游离 DNA)	≥10000×(目标区域)	≥99%	≤80%	≥85%
------------------	---------------	------	------	------

8.3 基因变异类型准确度

8.3.1 符合率

使用 CRM 或 RM，采用给定的测量系统进行测量，参照表 2 评估其符合率。

表 2 准确度评估表

待评价系统	标准参考数据		总数
	阳性	阴性	
阳性	真阳性数 (TP)	假阳性数 (FP)	TP+FP
阴性	假阴性数 (FN)	真阴性数 (TN)	FN+TN
总数	TP+FN	FP+TN	N

$$TPR = \frac{TP}{TP+FN} \times 100\% \quad (1)$$

式中：

TPR ——阳性符合率；

TP ——真阳性数；

FN ——假阴性数。

$$FPR = \frac{TN}{FP+TN} \times 100\% \quad (2)$$

式中：

FPR ——阴性符合率；

TN ——真阴性数；

FP ——假阳性数。

8.3.2 精密度

使用 RM，连续测量 3 个建库批次，每个批次进行 3 次重复测量（或每天一个分析批，同一个批次重复测量 3 次），计算阳性符合率的 CV，应不影响最终结果的判读。如果因为质量控制程序或者操作问题判断一批为失控，应剔除数据，并增加执行一个分析批。

8.4 检出限

8.4.1 评估要求

由于 NGS 对不同基因变异类型（SNV、INDEL、CNV）的 LOD 各不相同，因此需分别确认每种变异类型的 LOD。

8.4.2 评估方法

将 CRM 或已知样品梯度稀释不少于 3 个不同突变丰度，每个水平重复测量不少于 3 次，确保数据点满足统计学要求。将测量得到的数据进行 Probit 分析，记录 95%阳性检出率对应的 VAF 值为 C_y 。再将 VAF 值为 C_y 的样品进行重复测量 20 次，95%阳性检出率对应的 VAF 值即为检出限。

8.5 分析特异性

8.5.1 评估要求

应通过交叉反应和干扰实验进行确认。应根据检测系统中包含的试剂类型选择潜在的交叉反应物质进行确认。评估包括内、外源干扰物质和其他交叉反应的序列对测序结果的影响。交叉反应物质应包括不含目标序列的 DNA 样品、具有同源序列的核酸序列、检测范围外的变异，以及等位基因、假基因和其他类型的交叉反应序列。

8.5.2 交叉反应

对高浓度的交叉反应物质和 RM 或已知样品同时测量，重复 3 次，分别统计对交叉反应物质和 RM 或已知样品的目标变异测量结果为阳性的次数。

8.5.3 干扰实验

根据检测系统中包含的试剂类型和预期用途选择相应的干扰物质，确定干扰物质研究浓度。制备至少 2 个分析物水平（包含弱阳性水平在内）的样品，每个样品实验组和对照组均重复测试不少于 2 次。采用配对比对的方式，比较添加（实验组）与未添加（对照组）相应浓度干扰物质的样品检测结果的差异。如变异检测的成功率不低于 95%，则认为该浓度的物质不产生干扰。

8.6 可报告范围

8.6.1 评估要求

可报告范围应明确所覆盖目标基因区域范围和变异类型。详细列出目标基因区域，如全基因组坐标、外显子组或特定基因 panel 的覆盖范围。变异类型包括

但不限于：SNV、INDEL、CNV、SV、基因融合等。

8.6.2 评估方法

使用 CRM 或 RM 连续测量 3 个批次，每个分析批重复测量 3 次，统计可以稳定检出的变异类型，以及对应的基因组区域范围。

9 验证程序

9.1 基因组变异类型准确度

9.1.1 符合率

参见 8.3.1。

9.1.2 精密度

使用 RM，重复测量 3 次，计算阳性符合率的 CV，应不影响最终结果的判读。

9.2 检出限

使用 CRM 稀释至制造商声称的检出限，重复测量 5 次，测量结果应 100% 为阳性。如果不能满足 5 次均为阳性，可在不同批内进行 20 次重复测量（如测量 5 天，每天测定 4 次），检测结果应至少 17 次为阳性。

9.3 可报告范围

参见 8.6.2。

10 性能确认/验证结果表达

经性能确认/验证后的高通量基因测序系统，出具性能确认/验证报告，报告内需包含确认/验证的预期用途、原始数据及结论。

11 性能确认/验证时间间隔

在测量系统未发生重大变化时，性能确认无需再做。每年定期回顾性能参数，无重大变化时，性能验证可以用回顾性分析来替代。

附录 A 性能确认原始记录格式

项目：					
确认依据：					
1.检测系统概述					
检测方法		检测日期			
检测设备					
检测试剂					
2.文库及数据质量控制					
文库质量标准：					
数据质量控制要求：					
3. 基因组变异类型准确度					
3.1 符合率确认					
符合率结果统计表					
标品编号	预期结果	测定结果	是否接受 (Y/N)	阳性符合率 (%)	阴性符合率 (%)
确认结论：					

3.2 精密度确认

精密度结果统计表

检测天数	重复	高/中浓度	低浓度
第 1 天	重复 1		
	重复 2		
	重复 3		
第 2 天	重复 1		
	重复 2		
	重复 3		
第 3 天	重复 1		
	重复 2		
	重复 3		
第 4 天	重复 1		
	重复 2		
	重复 3		
第 5 天	重复 1		
	重复 2		
	重复 3		
平均值			
标准差			
变异系数 (CV%)			

确认结论：

4. 检出限确认

检出限确认结果统计表

变异类型	VAF 水平	水平 1	水平 2	水平 3	水平 4	水平 5	Cy
	重复						
SNV	重复 1						
	重复 2						
	重复 3						
INDEL	重复 1						
	重复 2						
	重复 3						
CNV	重复 1						
	重复 2						
	重复 3						

变异类型	ΔAF 水平 重复	10C _y	3.16C _y	C _y	C _y /3.16	C _y /10	阳性检出率
SNV	重复 1						
	重复 2						
	...						
	重复 10						
INDEL	重复 1						
	...						
	重复 10						
CNV	重复 1						
	...						
	重复 10						

确认结论：

5. 分析特异性确认

5.1 交叉反应确认

交叉反应结果统计表

交叉反应物质名称	样品类型	终浓度	检测结果 (阳性检出次数/总检测次数)		交叉情况
			交叉物质	RM/已知样品	

确认结论：

5.2 干扰实验确认

干扰实验结果统计表

干扰物质	干扰物质浓度	中高浓度						低浓度					
		对照组 1			实验组 1			对照组 2			实验组 2)		
		重复 1	重复 2	重复 3	重复 1	重复 2	重复 3	重复 1	重复 2	重复 3	重复 1	重复 2	重复 3

确认结论：

6. 可报告范围

6.1 目标基因区域和变异类型确认

目标基因区域和变异类型确认结果分析表

验证内容	类型	第 1 批			第 2 批			第 3 批			阳性符合率(%)
		重复 1	重复 2	重复 3	重复 1	重复 2	重复 3	重复 1	重复 2	重复 3	
变异类型	SNV										
	INDEL										
	CNV										
	基因融合										
	SV										
目标基因区域	全基因组坐标										
	外显子组										
	特定基因 panel										

确认结论：

附录 B 性能验证原始记录格式

项目：					
验证依据：					
1. 检测系统概述					
检测方法		检测日期			
检测设备					
检测试剂					
2. 文库及数据质量控制					
文库质量标准：					
数据质量控制要求：					
3. 基因变异类型准确度					
3.1 符合率验证					
符合率结果统计表					
标品编号	预期结果	测定结果	是否接受 (Y/N)	阳性符合 率 (%)	阴性符合 率 (%)
验证结论：					

3.2 精密度验证

精密度验证结果统计表

	高/中浓度	低浓度
重复 1		
重复 2		
重复 3		
平均值 (mean)		
标准差 (SD)		
变异系数 (CV%)		

验证结论:

4. 检出限验证

方法 1:

检出限验证结果统计表

CRM	重复 1	重复 2	重复 3	重复 4	重复 5
预期结果	阳性	阳性	阳性	阳性	阳性
测量结果					
符合率%					

方法 2:

检出限验证结果统计表

测量天数	重复 1	重复 2	重复 3	重复 4
第 1 天				
第 2 天				
第 3 天				
第 4 天				
第 5 天				
检测阳性数 (n)				
总数 (N)				
阳性检出率 (%)				

验证结论:

5.可报告范围

5.1 目标基因区域和变异类型确认

目标基因区域和变异类型确认结果分析表

验证内容	类型	第 1 批			第 2 批			第 3 批			阳性 符合 率(%)
		重 复 1	重 复 2	重 复 3	重 复 1	重 复 2	重 复 3	重 复 1	重 复 2	重 复 3	
变异类型	SNV										
	INDEL										
	CNV										
	基因融合										
	SV										
目标基因区域	全基因组坐标										
	外显子组										
	特定基因 panel										

验证结论：

参考文献

1. GUM (Guide to the Uncertainty in Measurement) ISO/IEC Guide 98-3: 2008
2. EP09-A3 Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline— Third Edition
3. EP17-A2 Evaluation of Detection Capability for Clinical Laboratory Measurement Procedure; Approved Guideline – Second Edition

全国生物计量技术委员会