

团 体 标 准

T/SZAS 90—2024

临床全外显子组测序性能确认及室内质控 规程

Specification for validation and internal quality control of clinical whole exome
sequencing

2024 - 11 - 26 发布

2024 - 12 - 05 实施

深圳市标准化协会 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	3
5 检测性能要求	3
6 性能确认要求	4
7 室内质控要求	5
8 室间质量评价	6
附录 A（资料性） 临床全外显子组测序实验相关建议	7
参考文献	8

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。

本文件由深圳华大基因股份有限公司提出。

本文件由深圳市标准化协会归口。

本文件起草单位：北京协和医院、浙江省人民医院、大连市妇女儿童医疗中心（集团）、吉林大学第一医院、中国医科大学附属盛京医院、重庆医科大学附属儿童医院、云南省第一人民医院、深圳华大基因股份有限公司、深圳华大医学检验实验室、安诺优达基因科技（北京）有限公司、北京贝瑞和康医学检验实验室有限公司、深圳华大基因科技有限公司。

本文件主要起草人：张抒扬、姜丹、唐少华、刘鹏、唐美芳、邱玲、徐秀华、刘睿智、刘雅萍、何蓉、何晓燕、吕涛、郭丹、吴平、李云、张嘉文、吴祺航、唐玉婧、肖红豆、邱娇、蒋婷、郑乔松、沈鑫翌、李倩一、刘杨杨。

临床全外显子组测序性能确认及室内质控规程

1 范围

本文件规定了临床全外显子组测序性能确认及室内质量控制（室内质控）的过程，涵盖检测性能和室内质控的要求、性能确认的过程、方法和评估要求等。

本文件适用于以大规模并行测序技术为技术路线的临床全外显子组测序，适用于该类型检测服务提供商及提供相关检测的各类检测机构。

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 22576.1—2018 医学实验室 质量和能力的要求 第1部分：通用要求

3 术语和定义

下列术语和定义适用于本文件。

3.1

高通量测序 high-throughput sequencing

能够一次并行对大量核酸分子进行平行序列测定的技术。

[来源：GB/T 42751—2023, 3.1]

3.2

外显子组 exome

个体基因组中能够编码产生蛋白质氨基酸序列的DNA区域的总和，在人类基因组中往往呈不连续的片段分布，占整个人类基因组的1~2%。

3.3

全外显子组测序 whole exome sequencing; WES

利用探针杂交富集外显子区域的DNA序列，并通过大规模并行测序技术进行全外显子区域的测序。

[来源：T/CHIA 21.2—2021, 3.3, 有修改]

3.4

临床全外显子组测序 clinical whole exome sequencing; cWES

采用全外显子组测序（本标准第3.3条），针对公共数据库中明确致病的基因，对获得的全外显子组数据进行变异校正及解读信息注释，从而获得与受检者相关的遗传病变异信息的检测方法。

3.5

性能确认 validation

对未经确认或自建的检验程序，以及明显修改过的检验程序，通过提供客观证据对特定的预期用途或应用要求已得到满足的认定。

3.6

标签 Barcode

一段特征性的脱氧核苷酸段片段，在多样本混合测序时，充当识别特定样本来源的唯一编号。

[来源：GA/T 1693—2020, 3.10]

3.7

Q20

测序数据中，碱基识别质量值大于20的碱基占有所有碱基的比例。

示例：碱基识别质量值为 20 时，表示碱基的正确率为 99%以上， $Q20 \geq 95\%$ ，则表示测序数据中 95%以上的碱基识别质量值大于 20。

[来源：T/SZAS 13—2019, 3.1.12]

3.8

Q30

测序数据中，碱基识别质量值大于30的碱基占有所有碱基的比例。

示例：碱基识别质量值为 30 时，表示碱基的正确率为 99.9%以上， $Q30 \geq 85\%$ ，则表示测序数据中 85%以上的碱基识别质量值大于 30。

[来源：T/SZAS 13—2019, 3.1.13]

3.9

平均测序深度 average sequencing depth

测序得到的落在目标区域的碱基数与目标区域长度的比值。

3.10

覆盖度 coverage ratio

将测序序列比对到参考序列上时，所有被比对到的区域占捕获芯片范围总区域的百分比。

3.11

单核苷酸变异 single nucleotide variation; SNV

在基因组水平，由单个核苷酸位点的变异（替代、插入或缺失）所引起的脱氧核糖核苷酸序列变异。

[来源：T/SZAS 13—2019, 3.1.2, 有修改]

3.12

插入缺失型变异 insertion and deletion; InDel

在基因组的某个位置上所发生的小片段序列的插入或者缺失，插入或缺失片段的长度在50 bp以下。

[来源：T/SZAS 13—2019, 3.1.3]

3.13

拷贝数变异 copy number variation; CNV

长度为1 kb以上的基因组大片段的拷贝数增加或者减少。

3.14

线粒体变异 mitochondrial mutation

线粒体DNA (mtDNA) 发生的变异，这些变异可以导致线粒体功能障碍或疾病。

3.15

杂合性缺失 loss of heterozygosity; LOH

在一对同源染色体上，位于相同基因座的两个等位基因中的一个发生缺失，导致的杂合状态丧失。

3.16

短串联重复序列 short tandem repeats; STR

短串联重复序列是由2到6个核苷酸组成的序列，在基因组中串联重复出现。

3.17

染色体非整倍体 aneuploidy

在正常的染色体组中，丢失或增加了一条或几条染色体。

3.18

灵敏度 sensitivity

实际为阳性的结果，被正确检测到的概率。

3.19

精确度 precision

在所有检出为阳性的结果中，实际为阳性的比例。

3.20

参考品 reference material; RM

在一项或多项特性上具有足够的均匀性和稳定性，用于校准、给其他物质赋值或提供质量保证的物质。

[来源：GB/T 19703—2020/ ISO 15194:2009, 3.4, 有修改]

3.21

读序 reads

高通量平台产生的含有碱基序列和质量值的序列片段。

[来源: GB/T 35890—2018, 3.1, 有修改]

3.22

基因型质量值 genotype quality; GQ

Phred格式的质量值, 表示在该位点该基因型存在的可能性。

注: 该值越高, 则Genotype结果正确的可能性越大。

4 缩略语

下列缩略语适用于本文件。

DNA: 脱氧核糖核酸 (deoxyribonucleic acid)

Indel: 插入缺失型变异 (insertion and deletion)

SNV: 单核苷酸变异 (single nucleotide variation)

5 检测性能要求

5.1 样本质量控制要求

5.1.1 确定样本类型

各实验室应根据检测范围及目的, 确定应包含的样本类型, 如外周血、羊水、流产组织及基因组 DNA 等。

5.1.2 样本运输要求

样本运输过程除满足 GB/T 22576.1—2018 中 5.4.5 的要求外, 还应包含清晰完整的受检者签署的知情同意书、样本编码及送检信息, 确保无混样风险。

5.1.3 样本质量控制要求

根据不同类型样本的特性, 进行充分测试, 评估各样本类型的采样量, 以及各类型样本在检测流程中的结果准确性及稳定性。采样量宜为文库构建环节 DNA 投入量的 3~4 倍, 各样本类型的采样量见附录 A。

5.2 DNA 提取质量要求

5.2.1 提取的样本类型应符合提取试剂的说明书规定范围, 提取后的 DNA 总量应至少满足一次文库构建的需要, OD (A260/A280) 值宜在 1.8~2.0 之间, 电泳检测 DNA 主带明显, 无明显降解。

5.2.2 自动化提取时应包含污染质控, 即空白对照提取浓度 $< 1 \text{ ng}/\mu\text{L}$ 。

5.2.3 产前样本提取完毕后应进行母源污染检测。

5.3 文库构建及杂交捕获质量要求

5.3.1 文库构建及杂交捕获试剂的批号、名称、外观应符合说明书标示, 有效期应在规定范围内。杂交后的文库 DNA 总量应为测序上机环节投入量的 2 倍以上, 文库的片段大小、体积、浓度等均应符合相应平台的上机标准。文库片段大小宜在 200 bp~500 bp。

5.3.2 室内质控参考品应在建库环节加入, 与临床样本一同参与建库及后续的全部实验流程。每批次实验应包含提取空白对照、阳性质控品及阴性质控品各一个, 空白对照提取浓度应 $< 1 \text{ ng}/\mu\text{L}$, 阴阳性质控品质量要求同 5.3.1。

5.4 测序质量要求

不同的测序平台有不同的合格阈值, 不同平台应参考测序试剂盒或测序平台说明书, 制定能保证分析成功率的测序数据合格指标, 以保证测序质量和靶区覆盖深度。测序的Q30应不低于85%。

5.5 信息分析质量要求

测序下机后的数据建议关注以下关键指标：平均测序深度、20×覆盖度、Q30和污染比例，这些指标的实测值应不低于各平台设置的最低要求且非胎儿样本性别应与送检信息一致。其中，目标区域的平均测序深度应根据检测策略对基因变异判定所需的数据深度的需求而定。

其余指标的建议合格标准为：测序下机原始数据经过滤后，Q30应不低于85%，20×覆盖度应不低于97.5%，生信环节提示无污染。

5.6 遗传解读质量要求

5.6.1 建议选用的变异注释数据见附录 A。所有数据库均应定期更新，公共数据库应至少一年更新一次。

5.6.2 变异注释的信息应至少包括：变异基本信息（染色体坐标、合子状态、基因名称、转录本、HGVS命名等）、群体数据（该位点在正常人群和本地样本中的频率）、基因元件与功能影响、基因相关的疾病信息、疾病数据库收录情况、软件计算与预测结果等。

5.6.3 位点筛选主要考虑三要素：所涉及疾病的相关特异性表型与临床主诉吻合、符合疾病遗传模式，以及致病性达到报告筛选标准，解读参考信息参见附录 A。

5.6.4 针对不同的变异类型，各实验室应在临床大样本量验证数据的基础上，对拟报告位点建立低质量位点判断阈值，并对低质量位点进行额外的金标准方法学验证补充实验。低质量位点阈值示例见附录 A。

6 性能确认要求

6.1 性能确认时机

6.1.1 实验室建立起临床全外显子组测序检测程序后，检测程序常规应用前应先进行性能确认。开展检测的实验室宜每年至少进行一次性能确认。

6.1.2 任何可能影响临床全外显子组测序检测程序性能的情况发生前，如发生试剂升级、试剂替换、实验流程变更、生信流程变更、解读流程变更等情况，应在检测程序重新启用前进行性能确认。

6.2 性能确认样本要求

6.2.1 性能确认样本来源

按照以下原则选择用于临床全外显子组测序性能确认的样本：

- a) 性能确认样本可使用过往临床样本。选择的临床样本应有可信参比方法提供的位点数据，即采用金标准方法、行业公认方法或经确认性能满足临床预期用途的方法检测过；
- b) 除临床样本外，性能确认亦可使用位点准确可信的独立第三方参考品，即由与检测程序提供方或试剂盒制造商无关的第三方制造，独立开发，不受仪器或试剂制造商的影响的第三方参考物质。参考品应足够均匀和稳定，且结果可信。

6.2.2 性能确认样本种类

用于临床全外显子组测序性能确认的样本应符合以下要求：

- a) 性能确认样本应包含参考序列参考品、特定位点阴性及阳性参考品或临床明确诊断阳性样本。参考序列参考品需使用有公开参考序列信息的通用参考品（如 NA12878 和炎黄一号），或经多个二代测序平台交叉验证过，有可信外显子组数据可供参考比对的参考品；
- b) 阳性参考品或临床明确诊断阳性样本选取时，变异类型应覆盖制造商说明书声明的检测范围内的全部变异类型；
- c) 特定位点阴性指挑选临床已知致病或疑似致病位点至少 100 个，选取经多个二代测序平台交叉验证或金标准方法验证过，在挑选出的位点上无致病变异的参考品或临床样本。挑选的已知位点宜覆盖所有染色体，且宜覆盖涉及疾病的不同遗传模式。

6.3 性能确认评估要求

6.3.1 准确性评估

选取参考序列参考品至少1种、特定位点阴性参考品或临床样本至少5种、阳性参考品或临床样本至少每种变异类型各1种，且阳性参考品或临床样本总数不少于10种（其余要求见6.2）。性能确认样本使用待确认的检测程序进行检测，准确性评估合格要求为：

- a) 各实验环节均需达到检测程序的质控指标；
- b) 参考序列参考品：SNV 灵敏度 $\geq 98\%$ ，SNV 精确率 $\geq 95\%$ ；InDel 灵敏度 $\geq 84\%$ ，InDel 精确率 $\geq 74\%$ ；
- c) 阳性参考品的阳性变异全部检出，阳性变异为制造商说明书声明的检测范围内的变异类型；
- d) 特定位点阴性参考品或临床样本，经验证过的位点上全部无阳性变异检出。

6.3.2 重复性评估（批内稳定性）

选取参考序列参考品至少1种、阳性参考品或临床样本至少2种，批次内重复建库三次，合格要求为：

- a) 各实验环节均应达到检测程序的质控指标；
- b) 将所有 SNV 及 InDel 检出结果按图 1 汇总，计算一致性比值作为重复性评价参数，该参数不得低于 90%。

6.3.3 重现性评估（批间稳定性）

选取参考序列参考品至少1种、阳性参考品或临床样本至少2种，分三个批次进行重复建库，合格要求为：

- a) 各实验环节均应达到检测程序的质控指标；
- b) 将所有 SNV 及 InDel 检出结果按图 1 汇总，计算一致性比值作为重现性评价参数，该参数不得低于 90%。

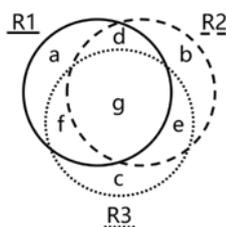


图1 一致性比值验证

^a 注：R1、R2 和 R3 分别为同一参考品的 3 个重复检出的阳性变异位点数。一致性比值= $g / (a+b+c+d+e+f+g) \times 100\%$ ，即交集数除以并集数。

6.3.4 结果判定

准确性评估、重复性评估及重现性评估均合格，则该实验室的cWES检测流程及其程序可被认定为性能合格。

6.3.5 不合格措施

如性能确认不合格，应排查实验及分析流程，修正检测程序后重新进行性能确认。

7 室内质控要求

7.1 室内质控时机与室内质控品选择

实验室应建立室内质控体系，定期监控检测程序的稳定性。室内质控推荐使用第三方参考品或临床明确诊断阳性样本，与临床样本一同进行文库构建及后续的全流程检测实验与分析。室内质控品应至少包含一种参考序列参考品及一种阳性变异参考品。阳性变异参考品建议每半年至一年进行一次更换，其余要求同6.2.2。

7.2 室内质控评估要求

7.2.1 结果判定

室内质控参考品需与临床样本一并进行文库构建及后续全部实验流程。合格要求为：

- a) 各实验环节均应达到检测程序的质控指标；
- b) 参考序列参考品：SNV 灵敏度 $\geq 98\%$ ，SNV 精确率 $\geq 95\%$ ；InDel 灵敏度 $\geq 84\%$ ，InDel 精确率 $\geq 74\%$ ；
- c) 准确检出阳性参考品的全部阳性位点。

7.2.2 不合格纠正措施

参考序列参考品或阳性变异参考品不满足7.2.1的合格要求时，室内质控不合格，应排查实验及分析流程，修正检测程序后重新实验。

8 室间质量评价

实验室应制定室间质评的政策和措施，计划性地参与外部质量控制活动，如国内外实验室认可机构组织的能力验证活动，实验室主管机构组织的比对活动，国际间、国内同行间的实验室比对试验等，并在实验结束后提交相应实验记录、实验结果等相关文件。根据室间质评结果评估实验室的检测准确性，并采取相应的改进措施。

附录 A (资料性) 临床全外显子组测序实验相关建议

A.1 cWES 建议采样量

以300 ng的建库投入量为例，建议采样量如下：外周血 ≥ 2 mL，如为婴幼儿取外周血或脐带血 ≥ 1 mL；流产组织取 ≥ 100 mg的胚胎、胎儿组织或绒毛；羊水取 ≥ 10 mL；基因组DNA总量 ≥ 1 μg ；口腔拭子 > 4 支；干血斑 > 3 个，单个直径 > 8 mm。

A.2 变异注释数据库及注释信息选择

变异注释所选用的数据库应至少包括人群频率数据库、疾病数据库、表型数据库，以及预测结果数据库。染色体坐标建议使用hg19参考基因组，转录本建议使用MANE转录本。人群频率数据库应包括常见的公用数据库，如千人基因组、dbSNP、gnomAD等，建议包含区域细分频率库（如ExAC东亚数据库East Asian）以及实验室基于自身样本构建的本地频率数据库。疾病数据库，应至少包括ACMG指南推荐的ClinVar和HGMD。表型数据库建议使用OMIM。预测结果数据库建议使用dbNSFP，建议使用Revel值综合反映预测结果。

A.3 致病性解读参考信息

致病性解读主要参考ACMG变异解读指南、ClinGen序列变异解读工作组和英国临床基因组科学学会等对该指南细化的规则，以及变异位点相关文献。

A.4 低质量位点阈值

不同测序平台、不同测序深度下，低质量位点阈值存在差异，各实验室应在临床大样本量验证数据的基础上，对拟报告位点建立起低质量位点判断阈值。

示例1：基于MGI测序平台，在 $200\times$ 测序深度下的低质量位点建议阈值为：SNV位点测序深度 $< 40\times$ 或SNV变异率 $< 40\%$ ；InDel位点测序深度 $< 60\times$ 或InDel变异率 $< 45\%$ ；纯合子或半合子位点测序深度 $< 40\times$ 或变异率 $< 85\%$ 。

示例2：基于Illumina测序平台，在 $60\times$ 测序深度下的低质量位点建议阈值为：位点突变率 $< 35\%$ 或突变深度 $< 10\times$ 且GQ值(genotype quality) < 20 ；纯合子或半合子位点突变率 $< 85\%$ 或突变深度 $< 10\times$ 且GQ值(genotype quality) < 20 。

参 考 文 献

- [1] GB/T 35890—2018 高通量测序数据序列格式规范
 - [2] GB/T 19703—2020/ISO 15194:2009 体外诊断医疗器械 生物源性样品中量的测量 有证参考物质及支持文件内容的要求
 - [3] GB/T 42751—2023 信息技术 生物特征识别 高通量测序基因分型系统规范
 - [4] GA/T 1693—2020 法庭科学 DNA二代测序检验规范
 - [5] T/SZAS 13—2019 基因组学数据集
 - [6] T/CHIA 21.2—2021 组学样本处理与数据分析标准 第2部分：全外显子组测序数据分析
-