

团 体 标 准

T/ZGYCXH XXX—XXXX

STR 遗传标记高通量测序产前诊断样本质量控制技术指南

Technical Guideline for Quality Control of Prenatal Diagnosis Samples Based on
STR Markers and High-Throughput Sequencing

(征求意见稿)

v20260430

XXXX - XX - XX 发布

XXXX - XX - XX 实施

中国遗传学会 发布

目 录

| | |
|------------------------------|-----|
| 目 录 | I |
| 前 言 | II |
| 引 言 | III |
| 1 范围 | 1 |
| 2 规范性引用文件 | 1 |
| 3 术语和定义 | 1 |
| 4 符号和缩略语 | 2 |
| 5 伦理要求 | 2 |
| 6 设备、耗材和试剂 | 2 |
| 7 样本采集、保存与运输 | 2 |
| 8 操作步骤 | 3 |
| 8.1 主要实验流程如下图 | 3 |
| 8.2 样本预处理与核酸提取 | 3 |
| 8.3 文库制备 | 3 |
| 8.4 上机测序 | 4 |
| 9 数据分析 | 4 |
| 9.1 主要数据分析流程如下图 | 4 |
| 9.2 数据质控 | 4 |
| 9.3 单胎/同卵双胞胎的结果判定 | 5 |
| 9.4 异卵双胞胎的结果判定 | 5 |
| 10 质量控制 | 6 |
| 10.1 起始 DNA 样本的定量 | 6 |
| 10.2 二代测序文库的片段分析、定量 | 6 |
| 10.3 二代测序数据的质控 | 6 |
| 10.4 胎儿游离 DNA 比例 | 6 |
| 附 录 A (资料性) 家系基因座数统计结果 | 7 |

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。

本文件由中国遗传学会提出并归口。

本文件起草单位：复旦大学附属妇产科医院，云南省第一人民医院，华西第二医院，贵州省人民医院，南京市妇幼保健院，上海韦翰斯生物医药科技有限公司。

本文件主要起草人：

引 言

产前诊断是保障妊娠安全、实现优生优育的关键技术手段，其在早期识别胎儿异常、指导临床决策、降低遗传病与出生缺陷风险等方面具有重要价值。目前常用的产前诊断方法包括无创产前检测、羊水穿刺、绒毛取样及超声检查等。其中，羊水穿刺与绒毛取样作为有创介入手段，存在一定的样本污染风险，主要包括母体细胞污染（maternal cell contamination, MCC）、操作过程中引入的污染以及环境DNA污染等。这些污染可导致母体或外源DNA混入，严重影响检测结果的准确性。研究显示，羊水样本中MCC发生率可高达68%，其中显著污染（MCC>5%）占比约24%。即使经离心处理的无血性羊水标本仍可能存在MCC。一旦发生MCC，误诊、漏诊的可能性将大大增加。此外，特殊亲子关系（如供卵妊娠），以及实验过程中发生样本混淆、标识错误等情况，也可导致检测结果的误判，造成严重的临床后果。

目前，临床中常采用短串联重复序列（short tandem repeat, STR）检测等方法分析母体和胎儿来源样本，进行MCC的监测，排除非生物学亲子关系或样本来源错误等，发挥多方面的质控作用。然而，上述方法通常未涵盖母体血浆样本，当STR检测结果显示多个基因座上存在等位基因不匹配时，无法有效甄别其根源是“样本混淆”还是“母子亲子关系不符”。母体血浆中含有胎儿来源的游离DNA（cfDNA），可作为天然混合样本用于母子亲子关系分析，并可以作为母体与胎儿的参照，有效鉴别样本混淆或实验过程发生标识错误等情况。然而血浆样本中cfDNA片段较短，常规STR检测技术往往无法适用。尽管已有研究探索采用不同遗传标记策略对血浆样本进行亲子鉴定，目前尚未形成标准化的适用于产前诊断样本污染质控监测的指导方案。

为弥补上述空白，本文件建立了一套基于母体外周血浆样本的STR检测与数据分析方法。该方法通过高通量测序技术分析血浆中母体与胎儿来源cfDNA的STR图谱，可实现亲子关系验证与胎儿DNA比例估算，从而准确判断样本污染状况。同时，通过与母体及胎儿样本的STR结果进行比对，可进一步排除样本混淆的可能性。该技术具有无创、快速、准确、可靠等特点，在操作安全性与检测便捷性方面具备显著优势。

本文件的制定与实施，旨在为产前诊断样本污染的质控监测提供标准化技术依据，提升相关操作的规范性与检测结果的可靠性，对推动产前诊断质量管理体系建设具有积极意义。

STR 遗传标记高通量测序产前诊断样本质量控制技术指南

1 范围

本文件规定了产前诊断样本污染质控监测的伦理要求、设备、耗材和试剂、样本采集、保存与运输、操作步骤和质量控制等要求。

本文件适用于医学和研究用途的产前诊断样本污染质控监测。

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 37223—2018 亲权鉴定技术规范

GB/T 43642—2024 法医学个体识别技术规范

GB/T 44393-2024 序列多态STR等位基因命名规则

GA/T1972-2021 法医物证检验术语

JJG 646 移液器检定规程

YY/T 0657 医用离心机

3 术语和定义

下列术语和定义适用于本文件。

3.1

短串联重复STR Short tandem repeat

由 2 到 6 个核苷酸组成的重复序列，广泛存在于真核生物基因组中。由核心单位重复数目在个体间呈高度变异性并且数量丰富，构成 STR 基因座的遗传多态性。

3.2

高通量测序技术 High-throughput sequencing technology

又称第二代测序技术(Next generation sequencing)，通过对核酸序列进行片段化、连接标签序列等处理，实现并行检测数百万条 DNA 序列，极大提高检测通量，降低检测成本。

3.3

胎儿游离DNA Cell free fetal DNA

存在母体循环中的胎儿来源 DNA，主要由胎盘组织中细胞滋养层和合体滋养层细胞的凋亡产生。

3.4

胎儿游离 DNA 比例 Cell-free fetal DNA fraction (FF)

孕妇血浆样本中胎儿来源的游离 DNA 占全部细胞游离 DNA 总量的比例，通常以百分数表示。

3.5

胎儿特异性等位基因 Fetus-specific allele

在孕妇外周血游离DNA中检出的、源自胎儿且与母体基因型不同的等位基因。

3.6

痕量等位基因 Minority STR alleles

对于 gDNA 样本被其他样本污染，若污染来源样本和被污染样本的 STR 基因型不同，这些 STR 基因座的数据上会体现出至少一个来源于污染样本的变异比例较低的等位基因，变异比例与污染的比例相关。

4 符号和缩略语

下列符号和缩略语适用于本文件。

STR: 短串联重复 (short tandem repeat)

PCR: 聚合酶链式反应 (polymerase chain reaction)

gDNA: 基因组DNA (genomic DNA)

cfDNA: 游离DNA (cell free DNA)

bp: 碱基对 (base pair)

Mb: 兆字节 (megabyte)

FSA: 胎儿特异性等位基因 (fetus-specific allele)

5 伦理要求

人源组织取材的研究应遵守政府法律和所有相关机构的规定，在取材检测前应获得患者的知情同意，并仅可在已通过项目伦理审查的医院和科研机构进行，获取胎儿生物样本产前诊断仅可在已经获得产前诊断资质的医院进行。

6 设备、耗材和试剂

6.1 仪器设备

6.1.1 PCR 仪。

6.1.2 高通量测序仪，可满足双端 150bp 测序。

6.1.3 基于荧光染料的核酸定量仪。

6.1.4 300bp 以下分辨率 ≥ 3 bp 的核酸片段分析仪。

6.1.5 微量可调移液器: 0.1 μ L~10 μ L, 20 μ L~200 μ L, 100 μ L~1000 μ L, 符合 JJG 646 要求。

6.1.6 台式低温低速离心机: 转速 ≥ 600 g, 符合 YY/T 0657 的要求。

6.1.7 台式低温高速离心机: 转速 ≥ 16000 g, 符合 YY/T 0657 的要求。

6.1.8 掌上离心机: 转速 ≥ 2000 g, 符合 YY/T 0657 的要求。

6.1.9 涡旋混合器。

6.1.10 磁力架。

6.2 耗材

6.2.1 无核酸酶离心管，规格为 1.5mL、2.0mL、0.5mL、0.2mL。

6.2.2 无核酸酶移液器吸头，规格为 10 μ L、200 μ L、1mL。

6.3 试剂

6.3.1 无核酸酶水。

6.3.2 血浆 cfDNA、血液、组织 gDNA 提取试剂。

6.3.3 多重 STR 扩增引物，带条形码和二代测序接头序列的扩增引物。

6.3.4 PCR 反应液。

6.3.5 纯化磁珠。

6.3.6 核酸定量荧光染料。

6.3.7 二代测序芯片和配套测序试剂。

7 样本采集、保存与运输

7.1 样本采集

母体外周血样本在 12 孕周后采集，使用游离 DNA 专用采血管，采集体积 5mL，胎儿来源样本（脐血、羊水、绒毛）根据临床标准采集方法进行采集，其中脐血体积 400 μ L，使用 EDTA 抗凝管。羊水体积 2mL，保存于无菌带盖玻璃或塑料管中。绒毛组织约 1.5cm \times 3cm \times 3cm 或 3-5mg，浸没于组织保存液中。

7.2 样本保存与运输

游离 DNA 专用采血管采集后应于室温（15 $^{\circ}$ C \sim 25 $^{\circ}$ C）保存和运输，并在 96 h 内完成血浆分离；分离后的血浆如 2 周内使用可于 2 $^{\circ}$ C \sim 8 $^{\circ}$ C 保存，如需长期保存应置于-70 $^{\circ}$ C 及以下，避免反复冻融，运输时应采用干冰运输。母体外周血全血样本采集后应于 2 $^{\circ}$ C \sim 8 $^{\circ}$ C 保存和运输，并在 24 h 内分离血浆或提取 gDNA；羊水及绒毛样本采集后应根据组织保存液类型选择室温或 2 $^{\circ}$ C \sim 8 $^{\circ}$ C 保存和运输，建议 2 周内完成 gDNA 提取。提取的 gDNA 可于-20 $^{\circ}$ C 保存，长期保存建议置于-70 $^{\circ}$ C 及以下。

8 操作步骤

以下操作的所在实验室应满足《医疗机构临床检验实验室管理办法》（卫医发〔2006〕73号），《医疗机构临床基因扩增检验实验室管理办法》（卫办医政发〔2010〕194号）。

8.1 主要实验流程如下图

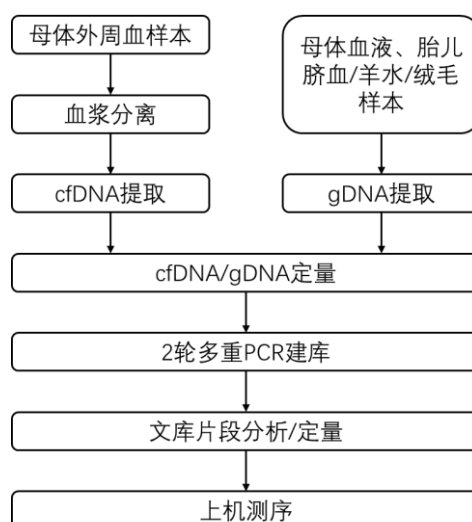


图 1 主要实验流程图

8.2 样本预处理与核酸提取

8.2.1 对于母体外周血样本，使用低温低速离心机 1600g，4 $^{\circ}$ C，离心采血管 10min，收集上清液至新的离心管。继续使用低温高速离心机 16000g，4 $^{\circ}$ C，离心上清液 10min，收集上清液至新的离心管，获得血浆。分离上清液后的血液样本用于母体 gDNA 提取。

8.2.2 使用 cfDNA 提取试剂提取 1mL 血浆样本，获得 cfDNA。使用 gDNA 提取试剂提取母体外周血全血样本（体积 200 μ L）、羊水样本（体积 2mL）或绒毛样本（质量 3~5mg），获得 gDNA。

8.2.3 使用荧光染料法对提取的 cfDNA 和 gDNA 样本进行定量。1mL 血浆样本提取的 cfDNA，40 μ L 洗脱，浓度范围在 0.1~0.4ng/ μ L 之间，总量在 4-16ng 之间。200 μ L 血液样本提取的 gDNA，总量应不低于 1 μ g。2mL 羊水、3-5mg 绒毛组织提取的 gDNA，总量应不低于 200ng。使用微量分光光度计对提取的 gDNA 样本进行纯度检测，260/280 比值在 1.8~2.0 之间，260/230 在 1.8~2.2 之间。

8.3 文库制备

8.3.1 配制第一轮多重 PCR 反应体系（27 μ L 体系：STR 引物池，3 μ L；PCR 反应液，9 μ L；gDNA \times μ L，水，补至 27 μ L），gDNA 起始量 20~200ng，cfDNA 起始量 \geq 4ng。

8.3.2 设置 PCR 程序并运行（预变性，96°C，3 min；（96°C，30 s；60°C，4 min）22 cycles；终延伸，72°C，4 min；低温保存，4°C，∞）。

8.3.3 使用纯化磁珠对 PCR 产物进行纯化。

8.3.4 配制第二轮多重 PCR 反应体系（10 μL 体系：PCR 反应液，5 μL；带有条形码序列的引物，0.5 μL，水，4.5 μL）。将反应体系加入上一步的纯化后磁珠内，温和混匀并短暂离心。

8.3.5 设置 PCR 程序并运行（预变性，96°C，3 min；（96°C，15 s；58°C，15 s；72°C，30 s）8 cycles；终延伸，72°C，5 min；低温保存，4°C，∞）。

8.3.6 PCR 产物纯化，使用纯化磁珠对 PCR 产物进行纯化，使用 30 μL 水洗脱获得 DNA 文库。

8.3.7 使用荧光染料法对 DNA 文库进行定量，文库总量 ≥ 50 ng；使用核酸片段分析仪确认文库片段主峰介于 250~300 bp 之间，无杂带。

8.4 上机测序

使用双端 150 bp 测序模式进行测序，gDNA 文库测序数据量 200 Mb，cfDNA 测序数据量 600 Mb。

9 数据分析

9.1 主要数据分析流程如下图

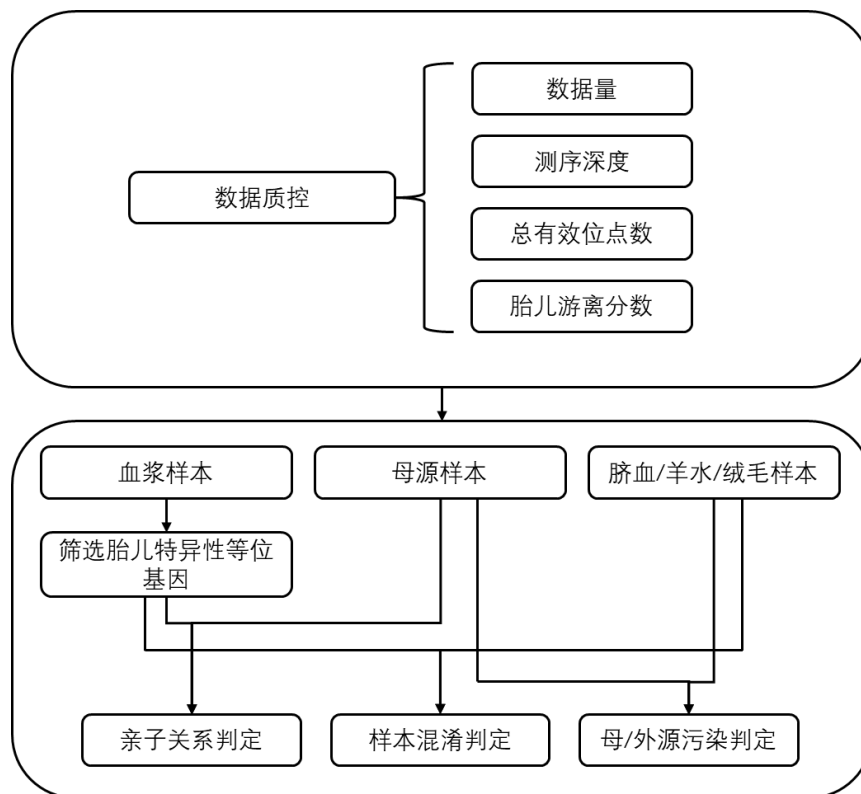


图 2 主要数据分析流程图

9.2 数据质控

9.2.1 使用 FASTX-Toolkit 0.0.14 软件 FASTQ-Quality-Filter 工具，参数选择 -p 80 -q 20，使用去除低质量 reads 软件默认参数，去除低质量的 bases。

9.2.2 筛选总有效基因座：使用 MISA 软件进行 STR 分型，保留 gDNA 和 cfDNA 共有的基因座，且 cfDNA 测序深度 >500×，gDNA 测序深度 >100×，统计出的总有效基因座记为 A，A ≥ 250。计算

胎儿游离分数，确认游离分数 $\geq 5\%$ 。

9.2.3 筛选 cfDNA 中 FSA：孕妇游离 DNA 中母体背景比较高，约 90%，而胎儿比例常低于 10%，根据占比高低，可以将孕妇游离 DNA 等位基因分为主要等位基因和次要等位基因；主要等位基因即母体等位基因，而次要等位基因往往为胎儿特有的等位基因，即 FSA；针对每个 STR 基因座，保留包含次要等位基因（变异等位基因频率为 0.02~0.15）的 STR 基因座，为降低 STR 滑移峰的干扰，去除次要等位基因中与主要等位基因重复数相差 ± 1 个的等位基因，最后保留下来的次要等位基因即为 FSA，这些含 FSA 的 STR 基因座成为有效 STR 基因座，把有效 STR 基因座的数量记为 S，把包含 1 个 FSA 的有效 STR 基因座数记为 S1，包含 2 个 FSA 的有效 STR 基因座数记为 S2，包含 3 个及以上 FSA 的有效 STR 基因座数记为 S3，则 $S=S1+S2+S3$ 。

9.3 单胎/同卵双胞胎的结果判定

9.3.1 亲子关系判定：理论上，若孕妇为胎儿生物学母亲，对于单胎或同卵双胞胎下的孕妇游离 DNA，每个 STR 基因座的 FSA 数不超过 1 个，即 $S2=0$ ，否则，该基因座不符合孟德尔遗传定律，即 S2 为不符合孟德尔遗传定律的总基因座数，而 S1 为符合孟德尔遗传定律的基因座个数。根据 GB/T 37223—2018 中对排除标准的相关定义，即“累积亲权指数小于 0.0001”时支持排除亲权关系，结合实践过程当中出现 3 个以上矛盾 STR 基因座时，累积亲权指数必然趋近于 0，将 S2 的数量是否 ≥ 3 作为排除胎儿与孕妇的亲子关系的依据。

9.3.2 样本混淆判定：将孕妇游离 DNA 与胎儿来源样本（脐血、羊水或绒毛）的基因座进行比较，根据 GB/T 43642—2024 判断胎儿（脐血、羊水或绒毛）与游离 DNA 中的胎儿是否为同一个体，如不为同一个体，则判定为样本混淆。

9.3.3 母源污染判定：若胎儿来源样本（脐血、羊水或绒毛）中存在痕量等位基因座，且所有痕量等位基因座与母本基因座符合孟德尔遗传，判定为母源污染。以母本特有等位基因座的测序深度占比计算污染比例。

9.3.4 外源污染判定：若胎儿来源样本（脐血、羊水或绒毛）中存在痕量等位基因座，且痕量等位基因座与母本基因座存在不符合孟德尔遗传等情况，判定为外源污染。以非胎儿等位基因座测序深度占比计算污染比例。

9.4 异卵双胞胎的结果判定

9.4.1 亲子关系判定：理论上，对于异卵双胞胎下的孕妇游离 DNA，每个 STR 基因座的 FSA 数不超过 2 个，否则，该基因座不符合孟德尔遗传定律，即 $S3=0$ ，S3 为不符合孟德尔遗传定律的总基因座数，而 S1 和 S2 为符合孟德尔遗传定律的基因座数。根据 GB/T 37223—2018 中对排除标准的相关定义，即“累积亲权指数小于 0.0001”时支持排除亲权关系，结合实践过程当中出现 3 个以上矛盾 STR 基因座时，累积亲权指数必然趋近于 0，将 S3 的数量是否 ≥ 3 作为排除胎儿与孕妇的亲子关系的依据。

9.4.2 样本混淆判定：将孕妇游离 DNA 与 2 个胎儿来源样本（脐血、羊水或绒毛）的基因型座分别进行比较，根据 GB/T 43642—2024 判断胎儿（脐血、羊水或绒毛）与游离 DNA 中的胎儿是否为同一个体，如不为同一个体，则判定为样本混淆。

9.4.3 母源污染判定：若 2 个胎儿来源样本（脐血、羊水或绒毛）中某个存在痕量等位基因座，且所有痕量等位基因座与母本基因座符合孟德尔遗传，判定为母源污染。以母本特有等位基因座的测序深度占比计算污染比例。

9.4.4 外源污染判定：若 2 个胎儿来源样本（脐血、羊水或绒毛）中某个存在痕量等位基因座，且痕量等位基因座与母本基因座存在不符合符合孟德尔遗传等情况，判定为外源污染。以非胎儿等位基因座测序深度占比计算污染比例。

胎儿来源样本以羊水为例，亲子关系和样本混淆的判定方法如下表所示：

表 1 亲子关系和样本混淆的判定方法

| | 羊水包含FSA | 羊水不包含FSA | 羊水包含FSA | 羊水不包含FSA |
|-------------|---------|----------|---------|----------|
| $S2 \leq 3$ | 质控合格 | 样本混淆 | / | / |

| | | | | |
|----------------------------------|----------|----------|-------|---------|
| S2>3 | 供卵 | 供卵+样本混淆 | / | / |
| S3≤3 | / | / | 质控合格 | 样本混淆 |
| S3>3 | / | / | 供卵 | 供卵+样本混淆 |
| | 单胎/同卵双胞胎 | 单胎/同卵双胞胎 | 异卵双胞胎 | 异卵双胞胎 |
| S2: 含 2 个胎儿特异性等位基因的 STR 基因座数。 | | | | |
| S3: 含 3 个及以上胎儿特异性等位基因的 STR 基因座数。 | | | | |

胎儿来源样本以羊水为例，母源、外源污染的判定方法如下表所示：

表 2 母源、外源污染的判定方法

| | 羊水存在痕量等位基因 | |
|---------------------------|------------|--|
| 羊水痕量等位基因座与母亲基因座符合孟德尔遗传定律 | 母源污染 | |
| 羊水痕量等位基因座与母亲基因座不符合孟德尔遗传定律 | 外源污染 | |
| | | |

10 质量控制

10.1 起始 DNA 样本的定量

健康人 1mL 血浆样本提取的 cfDNA，以 40 μL 洗脱体积为例，浓度范围在 0.1~0.4ng/μL 之间，投入检测的总量应大于 4ng。若浓度大于 1ng/μL，应考虑存在 gDNA 污染，可使用琼脂糖凝胶电泳鉴定是否存在大片段 gDNA 条带宜对血浆样本重新进行高速离心并分离上清液，或重新采样。其它样本类型提取的 gDNA，投入检测的总量应大于 20ng。

10.2 二代测序文库的片段分析、定量

成功构建的二代测序文库，使用核酸片段分析仪确认片段长度介于 250~300bp 之间，无其它杂带。使用荧光染料定量的文库总量不低于 50ng。

10.3 二代测序数据的质控

gDNA 样本的平均测序深度应不小于 100×，cfDNA 的平均测序深度应不小于 500×，总有效 STR 基因座个数应不少于 250 个。

10.4 胎儿游离 DNA 比例

胎儿游离 DNA 比例应 ≥5%，若不满足 5% 的要求，应在距上一次采样的至少一周后重新采样。

附 录 A
(资料性)
家系基因座数统计结果

选择 12 例羊水判定亲子关系不支持，15 例羊水判定亲子关系支持，合计 27 例家系。每个家系的总有效基因座和胎儿特异性等位基因种类 >1 的基因座数统计结果如下：

表 A.1 家系基因座数统计结果

| 家系序号 | 总有效基因座数 | >1基因座数 | 判定结果 |
|------|---------|--------|------|
| 1 | 265 | 21 | 供卵 |
| 2 | 276 | 16 | 供卵 |
| 3 | 265 | 21 | 供卵 |
| 4 | 273 | 13 | 供卵 |
| 5 | 265 | 21 | 供卵 |
| 6 | 278 | 10 | 供卵 |
| 7 | 280 | 8 | 供卵 |
| 8 | 278 | 10 | 供卵 |
| 9 | 265 | 21 | 供卵 |
| 10 | 280 | 8 | 供卵 |
| 11 | 265 | 21 | 供卵 |
| 12 | 250 | 20 | 供卵 |
| 13 | 256 | 0 | 支持 |
| 14 | 265 | 0 | 支持 |
| 15 | 266 | 0 | 支持 |
| 16 | 270 | 0 | 支持 |
| 17 | 269 | 0 | 支持 |
| 18 | 261 | 0 | 支持 |
| 19 | 250 | 0 | 支持 |
| 20 | 261 | 0 | 支持 |
| 21 | 256 | 0 | 支持 |
| 22 | 271 | 0 | 支持 |
| 23 | 260 | 0 | 支持 |
| 24 | 252 | 0 | 支持 |
| 25 | 275 | 0 | 支持 |
| 26 | 272 | 0 | 支持 |
| 27 | 271 | 0 | 支持 |