

病原微生物基因组分析工作站采购参数

1 设备用途

1.1 该设备可用于细菌全基因组、病毒全基因组、宏基因组测序数据分析，支持二代、三代测序数据，可适用于所有主流测序平台（例如：华大、illumina、Ion torrent、真迈、安图、齐碳、Nanopore、Pacbio 等）。每个系统均可以图形界面分析，通过任意选中目标测序数据，点击一个按钮即可完成分析，无需任何代码编写及参数设置。

1.2 该设备可用于病原微生物相关的信息管理，样本信息管理包括样本采集信息、人类/动物/环境样本的详细信息、临床诊断信息、保藏信息，菌株信息包括菌株编号、菌株类型、不同物种的重要耐药基因、毒力基因等信息，测序数据包括测序平台、测序仪型号、测序文件大小、文件格式、存储等管理功能；并且所有数据支持批量导入，已上传的测序数据可批量导出。

1.3 该设备可用于病原微生物数据检出的长期监测，根据不同物种类型，如细菌、病毒等构建不同的可视化图谱，统计监测图谱包括地图（国家/省/市）、柱形图、饼图、曲线图、热图、桑基图、进化树图等。

2 病原微生物基因组分析工作站配置参数如下：

2.1 系统 1：细菌数据管理及基因组全流程分析系统

2.1.1 样本数据管理模块

- 1) 支持细菌样本数据的统一维护，包括样本相关基本信息的录入、编辑、删除，并且支持批量样本导入；
- 2) 支持基因组文件的批量导入，包含原始下机数据及组装基因组文件，不限测序平台类型，并与样本信息自动化建立关联关系；
- 3) 根据不同模块筛选条件，批量导出样本基本数据及分析结果信息；
- 4) 支持菌株数据管理，并与样本自动化建立关联关系；
- 5) 自动化构建菌株数据耐药基因、毒力基因检出热力图谱；
- 6) 提供高质量本地参考数据库，应与国家级病原鉴定基因组数据库相同，能够批量导出.cpin 格式的批量样本加密文件以及.cpin 格式的基因组序列加密文件，

这些文件可直接上传到传染病病原监测网络中。

2.1.2 细菌基因组分析模块

- 1) 基因组全流程分析（无参拼接）
 - a) 一站式完成数据质控、无参拼装、物种鉴定、病原学特征数据分析（包含耐药基因、毒力因子、MLST 型、cgMLST、血清型）、SNP 分析，并同步完成常用功能数据库的基因注释；
 - b) 支持数据分析任务手动终止及重新开启分析；
 - c) 提供本地化中文知识库对病原特征及耐药基因特征进行解读。
- 2) 基因组全流程分析（有参拼接）
 - a) 一站式完成数据质控、有参拼装、物种鉴定、病原学特征数据分析（包含耐药基因、毒力因子、MLST 型、cgMLST、血清型）、SNP 分析，并同步完成常用功能数据库的基因注释；
- 3) 测序数据质控分析
 - a) 分别支持二代、三代测序数据质控分析；
 - b) 自动生成并保留质控后的数据文件；
 - c) 质控标准符合传染病病原监测项目参数规范。
- 4) 细菌基因组拼接分析（无参拼接）
 - a) 自动化完成测序数据质控；
 - b) 支持二代、三代、二代+三代下机数据，不限制测序平台；
 - c) 自动生成基因组成环信息。
- 5) 细菌基因组拼接分析（有参拼接）
 - a) 自动化完成测序数据质控；
 - b) 根据不同的物种，选择对应参考基因组进行基因组拼接；
 - c) 支持二代、三代、二代+三代下机数据，不限制测序平台；
 - d) 自动生成基因组成环信息。
- 6) 物种鉴定（提供操作系统截图、每个物种默认参考基因组清单及系统截图、样本及基因组序列的.cpin 示例加密文件）
 - a) 支持已发现的细菌物种鉴定，需覆盖 500 种以上的物种（Species）类型，且每个物种的参考基因组与国家疾控保持统一；
 - b) 能够识别出大肠杆菌、志贺氏菌等高度相似的菌株。

- 7) MLST 分析
 - a) MLST 分型覆盖 200 种以上的物种;
 - b) 数据库定期更新。
- 8) cgMLST 分析
 - a) cgMLST 分析要求覆盖 110 种以上的物种 (Species) 类型;
 - b) 多样本分析自动构建系统发育进化树。
- 9) 血清型分型
 - a) 血清型预测分型覆盖到: 沙门氏菌、肺炎克雷伯菌、大肠杆菌、肺炎链球菌、副溶血性弧菌、霍乱弧菌、志贺菌、猪链球菌。(其他菌株分子血清型根据传染病病原监测工作需要及时更新数据库)
- 10) 物种鉴定及分型
 - a) 自动化完成物种鉴定、MLST 分型、cgMLST 分型;
 - b) 大肠杆菌支持鉴定五种致泻大肠埃希氏菌分类;
 - c) 血清型覆盖 8 种及以上物种, 选用不同的分子血清型数据库。
- 11) 耐药及毒力基因分析
 - a) 耐药基因覆盖 CARD、ResFinder 两种数据库;
 - b) 毒力基因使用 VFDB 数据库;
 - c) 根据耐药、毒力大类构建统计图;
 - d) 提供耐药元件鉴定、耐药基因和表型相关系数、耐药机理和耐药药物鉴定等分析 以及毒力元件鉴定和毒力机理等分析;
 - e) 自动识别重点耐药基因、毒力基因;
 - f) 耐药基因、毒力基因数据库定期更新。
- 12) 基因组预测及注释分析
 - a) 使用基因组序列完成多个数据库的预测及注释分析;
 - b) 注释数据库包括但不限于: KEGG、COG、NR、MetaCyc、CAZY、PHI、SwissProt、EggNOG、GO 等。
- 13) 细菌基因组快速鉴定及分型
 - a) 利用原始下机数据直接与参考基因组数据库比对, 省略拼接组装的过程, 快速实现物种鉴定及 MLST 分型。
- 14) 系统发育进化树

- a) 实现对同一物种的多个基因组组装数据进行系统发育进化树构建;
- b) 包含至少 3 种树图可视化方式;
- c) 支持自定义进化树图可视化显示的内容, 以及相关样式;
- d) 支持差异基因矩阵可视化及原始数据导出。

15) SNP 分析

- a) 支持多样本、单样本进行 SNP 分析;
- b) 多样本进行 CallSNP 矩阵分析, 并以图形方式展示不同样本间的碱基差异;
- c) 多样本自动化构建进化树, 包含至少 2 种树图可视化方式。

16) 移动元件及相关抗性基因和毒力因子分析

- a) 可以对病原细菌基因组上的 insertion sequences (IS), integrative conjugative elements (ICE), integrons (IN), plasmid 和 transposons (Tn) 进行同步比对和注释。
- b) 判断移动元件是否存在耐药基因及毒力基因。

17) BLAST 分析

- a) 将本地的菌株基因组数据与同物种的公共数据库进行比对, 完成统计学的相似性比对;
- b) 公共数据库中必须包含国家、分离源、分离时间等信息。

18) BLAST 进化树分析

- a) 通过 BLAST 分析的结果挑选相似度高的一批菌株, 构建进化树分析。

19) 自建 BLAST 数据库

- a) 将本地已有的基因组添加到 BLAST 数据库中, 在构建任务的时候可选择是否使用自建库的数据;
- b) 系统自动校验选中的基因组是否可以添加到 BLAST 数据库中。

2.1.3 数据监测模块

- 1) 针对全国各省份细菌基因组数据的实时监测, 从空间、时间方面对菌株进行可视化追踪;
- 2) 按全国、全省、全市的维度展示不同统计地图。

2.1.4 预警分析模块

- 1) 根据样本及分析数据自动化监测聚集性疫情预警, 并给出明显的提示;

- 2) 监测每个分析结果中的重点耐药基因、重点毒力基因，并给出明显的预警提示；
- 3) 提供耐药基因及抗生素类别、毒力基因及毒力因子统计，以及耐药基因及物种等相关聚类图谱。

2.1.5 任务监测管理

- 1) 支持生物信息分析任务创建，以及一站式分析任务的进度等报告查看，并且支持分步骤可视化查看分析结果；
- 2) 支持优先级较高的数据优先分析，并不影响其他已创建的分析任务执行。
- 3) 包含单样本、多样本任务，支持检索、分页查看管理；
- 4) 提供任务终止功能，没有分析结束的任务支持终止运行，可删除历史任务；
- 5) 提供进度日志查看功能，单样本、多样本分析任务均支持查看对应的进度日志；
- 6) 创建分析任务，根据所选测序数据类型判断可选择的分析应用，批量执行，自动排队执行。

2.1.6 可视化统计及监测

根据重点统计数据生成可视化图形，并生成深色自适应的大屏可视化页面。

2.1.7 系统运行环境

配套本地化系统安装包，支持以桌面应用模式打开系统。

2.2 系统 2： 病毒数据管理及基因组全流程分析系统

2.2.1 样本数据管理模块

- 1) 支持病毒样本数据的统一维护，包括样本相关基本信息的录入、编辑、删除，并且支持批量样本导入；
- 2) 支持基因组文件的批量导入，包含原始下机数据及组装基因组文件，不限测序平台类型，并与样本信息自动化建立关联关系；
- 3) 根据不同模块筛选条件，批量导出样本基本数据及分析结果信息；
- 4) 支持毒株数据管理，并与样本自动化建立关联关系；

2.2.2 数据分析监测管理

- 1) 支持生物信息分析任务创建，以及一站式分析任务的进度等报告查看，并且

支持分步骤可视化查看分析结果；

- 2) 支持优先级较高的数据优先分析，并不影响其他已创建的分析任务执行。
- 3) 包含单样本、多样本任务，支持检索、分页查看管理；
- 4) 提供任务终止功能，没有分析结束的任务支持终止运行，可删除历史任务；
- 5) 提供进度日志查看功能，单样本、多样本分析任务均支持查看对应的进度日志；
- 6) 创建分析任务，根据所选测序数据类型判断可选择的分析应用，批量执行，自动排队执行。

2.2.3 病毒 Isolate 分析

- 1) 一站式完成数据质控、拼装、分型、基因组功能注释、新冠病毒、流感病毒、合胞病毒、猴痘病毒、基孔肯雅病毒等的支系分析、SNP 分析；
- 2) 基因组拼装流程支持二代、三代、二代+三代下机数据，不限制测序平台；
- 3) 包含 220 种以上病毒基因组数据，包含 1000 种以上亚型数据；
- 4) 自动化完成病毒数据库注释，包括 KEGG、uniprot、refseq、Genbank、Emapper、pFam 等。
- 5) 支持自定义参考基因组完成基因组拼接及 SNP 分析。

2.2.4 病毒进化分析

- 1) 实现对同一物种的多个基因组组装数据进行系统发育进化树构建。
- 2) 完成基因组两两之间的同源性比对；
- 3) 包含至少两种树图可视化方式。标注特定 lineage 颜色等客户自定义树图修饰功能。

2.2.5 BLAST 比对分析

- 1) 实现对已知物种的基因组进行 BLAST 比对；
- 2) 支持对甲流、乙流、新冠病毒、副流感、腺病毒、博卡病毒、呼吸道合胞病毒、猴痘病毒、基孔肯雅病毒、HIV 病毒等高质量病毒基因组进行全库比对，并生成序列进化树。

2.2.6 SNP 多序列比对分析

- 1) 相同物种的基因组序列进行共有、独有 SNP 的分析，并构建碱基比对可视化热图；
- 2) 支持所有位点数据列表的导出；

- 3) 单基因组序列支持单核苷酸变异、缺失、插入、复合变异等分析及可视化展示。

2.2.7 病毒基因组数据监测：

- 1) 针对全国各省份病毒基因组数据的实时监测，从空间、时间方面对毒株进行可视化追踪；
- 2) 支持将统计图形数据导出并下载在本地；
- 3) 按全国、全省、全市的维度展示不同统计地图。

2.2.8 预警分析

- 1) 聚集性疫情预警：从科学的角度设置聚集性预警统计模型，实时识别系统数据，当有符合统计模型规则出现时主动报出预警提示，并支持查看预警数据详细情况；

2.2.9 系统运行环境

配套本地化系统安装包，支持以桌面应用模式打开系统。

2.3 系统 3： 宏基因组全流程分析系统

2.3.1 样本数据管理模块

- 1) 支持宏基因组样本数据的统一维护，包括样本相关基本信息的录入、编辑、删除，并且支持批量样本导入；
- 2) 支持 16S 扩增子测序、宏基因组测序文件的批量导入；
- 3) 不限测序平台类型，并与样本信息自动化建立关联关系；
- 4) 根据不同模块筛选条件，批量导出样本基本数据及分析结果信息。

2.3.2 数据分析监测管理

- 1) 支持生物信息分析任务创建，以及一站式分析任务的进度等报告查看，并且支持分步骤可视化查看分析结果；
- 2) 支持优先级较高的数据优先分析，并不影响其他已创建的分析任务执行。
- 3) 包含单样本、多样本任务，支持检索、分页查看管理；
- 4) 提供任务终止功能，没有分析结束的任务支持终止运行，可删除历史任务；
- 5) 提供进度日志查看功能，单样本、多样本分析任务均支持查看对应的进度日志；
- 6) 创建分析任务，根据所选测序数据类型判断可选择分析应用，批量执行，

自动排队执行。

2.3.3 宏基因组-16S 扩增子测序数据分析

- 1) 20 分钟分析完成一个样本；
- 2) 对测序数据进行评估，包括 reads 数、reads 长度、GC 含量、整体测序质量的评估等常见评估指标，并过滤低质量 reads；
- 3) 使用质控完成的测序数据进行双端拼接后聚类为 ASV，并对 ASV 代表序列进行物种注释，计算物种丰度；
- 4) 利用 ASV 和物种丰度计算样本的 alpha 多样性（包括 Shannon, Observed Features, Faith's Phylogenetic Diversity 和 Evenness），做 alpha 多样性稀释性曲线；多样本间计算 beta 多样性（包括 Jaccard, Bray-Curtis, unweighted UniFrac 和 weighted UniFrac）；多样本间做 ANCOM 差异分析；并完成可视化图形渲染展示。

2.3.4 宏基因组-全基因组测序数据分析

- 1) 对测序数据进行评估，包括 reads 数、reads 长度、GC 含量、整体测序质量的评估等常见评估指标，并过滤低质量 reads；使用质控完成的测序数据组装，获得叠连群(contigs)；
- 2) 在 reads 和 contig 层面进行物种注释；基于 contig 和三种方法进行分箱 (binning)，并对结果进行整合和优化，以获得更好的 binning 结果；并基于 binning 结果进行重组装；最后获取每个 contig 和 bin 的丰度，及 bin 的物种注释和功能注释。
- 3) 对 VF、Pfam、Swissprot、COG、KEGG-KO、CARD 等数据库进行功能注释。
- 4) 支持质控、组装、物种分类等分析结果下载。

3 硬件配置

- 5) 3.1 服务器：CPU： 2 颗 Xeon SR 4210R 2.4G 10 核 100W（20 核，40 线程）；
内存：DDR4 2933 ECC RDIMM $\geq$ 512G；GPU:显存容量 $\geq$ 16G；硬盘容量：SSD $\geq$ 8TB；数据盘：服务器专用企业级 SATA 硬盘 $\geq$ 20TB；
3.2 其它：显示器 $\geq$ 30 英寸、打印机、键鼠套装等

4 设备软硬件配置清单如下：

- 4.1 满足招标参数的病原微生物分析服务器：1 台；

4.2 满足招标参数的细菌数据管理及基因组全流程分析系统：1 套；

4.3 满足招标参数的病毒数据管理及基因组全流程分析系统:1 套

4.4 满足招标参数的宏基因组全流程分析系统：1 套；

5 技术服务要求

5.1 投标人应提供全套、完整的技术资料，包括产品应用说明书、操作手册、维护保养说明等；

5.2 设备安装、调试和验收要求：设备到达最终用户现场并且实验室条件合格后，在接到用户通知后需安排有经验的工程技术人员到用户现场安装、调试设备，设备安装调试需在 15 个工作日内完成。安装前，提供安装条件、用水、用电条件说明；

5.3 培训要求：安装验收后，厂家在用户所在地对用户进行设备操作和日常维护的现场培训。包括设备使用方法和维护方法等；

5.4 设备生产厂家需在国内直接设有销售服务公司（需提供资质证明，包括厂家服务中心营业执照和工程师名单、联系方法及厂家盖章的售后服务承诺书等）。

★5.5 针对未知病原数据分析与监测，由于需要开展病原体快速甄别、微生物分型解析，感染溯源与流行特征分析，需免费提供 1 套未知病原宏基因组（mNGS）分析系统给予客户；

6 质量保证要求

保修期：主机及相关配套设备保修期 3 年，软件及数据库终身免费维护及升级。自设备验收合格之日起计算。保修期内应提供免费保修。

备注：供货期：15 个工作日完成。