**DRAGENTM服务器操作手册简要版**

1. putty软件下载

下载链接：<https://www.putty.org/>

1. FileZilla软件下载

下载链接：<https://filezilla-project.org/>

1. PyCharm软件下载

下载链接：<https://www.jetbrains.com/pycharm/download/>

1. 服务器地址

* IP：
* 登录名：
* 登录密码：

1. 登录服务器（通过putty）：

* 首次登录在Host Name(or IP address)填写服务器IP地址（如图1）
* 在Saved Sessions 填入服务器名称例如1010，点击Save按钮保存（如图2），双击Default Settings下对应的服务器名称即可（如图3），输入用户名后回车
* 系统提示输入用户名对应的密码（输入的密码屏幕不显示），输入完成后，即可完成服务器（如图4）

图形用户界面, 文本, 应用程序

描述已自动生成

图表 1 putty登录页面

图形用户界面, 应用程序

描述已自动生成

图表 2 putty设置服务器

图片包含 形状

描述已自动生成

图表 3 putty输入用户名

图形用户界面, 文本

描述已自动生成

图表 4 root用户成功登录服务器

1. 本地电脑登录服务器（通过FileZilla）：

* 首次登录，依次填入下图红框标注的Host、Username、Password、Port，然后点击Quickconnect
* 如果已经登录过，则直接点击Quickconnect右边的▼会出现曾经连接过的地址，点击连接即可
* 在服务器文件目录的Remote site中输入需要下载的文件所在的路径文件，右键点击Download即可

图形用户界面, 应用程序

描述已自动生成

1. 移动硬盘拷贝数据到服务器

* 通过服务器后面的USB连接移动硬盘
* fdisk -l 命令查看挂在的移动硬盘名称
* 在服务器上创建/share/文件夹，创建需要分析对应项目的分析文件夹入：/staging/wes-2021.01.25/
* 通过mount命令实现硬盘与服务器/share/文件夹的连接：
* 通过cp命令讲硬盘数据拷贝到服务器上:cp /share/ /staging/wes-2020.01.25/
* 借助umount卸载硬盘:umount /share/
* 拔出移动硬盘

1. 文件输入

* BAM文件
* germline: -b sample.bam somatic: --tumor-bam-input tumor.bam
* fastq(.gz)
* WGS/WES：-1 sample.R1.fastq -2 sample.R2.fastq
* Tumor only：--tumor-fastq1 tumor.R1.fastq --tumor-fastq2 tumor.R2.fastq
* Tumor-normal：-1 N.R1.fastq -2 N.R2.fastq --tumor-fastq1 T.R1.fastq --tumor-fastq2 T.R2.fastq

1. WGS数据分析实例：

图形用户界面, 文本, 应用程序

描述已自动生成

1. WES数据分析实例：

图形用户界面, 文本, 应用程序

描述已自动生成

1. WGS/WES分析的一些重要结果文件：

* Small Variant（SNV和small INDEL）结果文件：\*hard-filtered.vcf.gz，vcf文件格式，可查看带#的meta-information了解更多。此外还包含gvcf文件\*hard-filtered.gvcf.gz可用于做家系分析。
* CNV（Copy Number Variant）结果文件：\*cnv.vcf.gz，vcf文件格式，可查看带#的meta-information了解更多。
* SV（Structural Variant）结果文件：\*sv.vcf.gz，vcf文件格式，可查看带#的meta-information了解更多。
* STRs（Short Tandem Repeats）结果文件：\*repeats.vcf.gz，Core VCF Fields描述如下（只适用于WGS全基因组分析）：

|  |  |
| --- | --- |
| **Field** | **Description** |
| CHROM | Chromosome identifier |
| POS | Position of the first base before the repeat region in the reference |
| ID | Always . |
| REF | The reference base at position POS |
| ALT | List of repeat alleles in format <STRn> where n is the number of repeatunits |
| QUAL | Always . |
| FILTER | Always PASS |

Additional INFO Fields描述如下：

|  |  |
| --- | --- |
| **Field** | **Description** |
| SVTYPE | Always STR |
| END | Position of the last base of the repeat region in the reference |
| REF | Number of repeat units spanned by the repeat in the reference |
| RL | Reference length in bp |
| RU | Repeat unit in the reference orientation |
| REPID | REPID Repeat id from the repeat-specification file |

* ROH（Regions of Homozygosity）结果文件：\*.roh.bed，每一行代表一个ROH，包含5列，分别是Chromosome Start End Score #Homozygous #Heterozygous；其中Score是基于纯合变异个数和杂合变异个数的打分值，#Homozygous是该区域内纯合变异的个数，#Heterozygous是该区域内杂合变异的个数。此外\*roh\_metrics.csv列出了large ROH (>3 Mb) 的个数和SNVs百分比。
* coverage统计报告：\*wgs\_coverage\_metrics.csv，包含了基因组层面的与覆盖度相关的各种统计指标。如果要做个性化的统计，如大于30X的覆盖度，还可以使用\*wgs\_fine\_hist.csv文件，该文件包含两列，分别是深度和对应的位点数。
* mapping统计报告：\*.mapping\_metrics.csv，包含各种比对指标的统计。