|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | |  |
| [pasa v2.5.2](https://github.com/shenwei356/seqkit/releases/tag/v2.4.0) | | |  |
| **应用移植指南** | | |
| **文档版本** | **01** | |
| **发布日期** | **2023-03-24** | |
| 华为网格系统---方案4-032.png | | | | |
|  | 华为技术有限公司 | | 附件1-16K |  |

|  |
| --- |
| 版权所有 © 华为技术有限公司2022。 保留一切权利。  非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。  商标声明  附件3-版权声明页图 和其他华为商标均为华为技术有限公司的商标。  本文档提及的其他所有商标或注册商标，由各自的所有人拥有。  注意  您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。  由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。 |

|  |  |
| --- | --- |
| 华为技术有限公司 | |
| 地址： | 深圳市龙岗区坂田华为总部办公楼 邮编：518129 |
| 网址： | <https://www.huawei.com> |
| 客户服务邮箱： | [support@huawei.com](mailto:support@huawei.com) |
| 客户服务电话： | 4008302118 |

目 录

[目 录 ii](#_Toc18378)

[1 介绍 1](#_Toc9499)

[2 环境要求 2](#_Toc17343)

[3 移植规划数据 4](#_Toc18558)

[4 配置编译环境 5](#_Toc19178)

[4.1安装Samtools 6](#_Toc14748)

[4.2安装gmap 6](#_Toc21465)

[4.3安装minimap2 7](#_Toc14743)

[4.4安装blat 7](#_Toc29865)

[4.5安装pblat 8](#_Toc8577)

[4.6安装URI 9](#_Toc22531)

[4.7安装PASA 10](#_Toc14282)

[5 获取源码 12](#_Toc7999)

[6 运行和验证 13](#_Toc25201)

[7 修订记录 14](#_Toc16988)

# 介绍

PASA是一种真核基因组注释工具，利用表达的转录序列的剪接比对来自动建模基因结构，并保持基因结构注释与最新的实验序列数据一致。PASA还可以识别和分类所有转录本比对支持的剪接变异。

关于PASA的更多信息请访问<https://github.com/PASApipeline/PASApipeline/wiki>。

一句话描述：一种真核基因组注释工具

建议的版本

建议使用版本为“[pasa v2.5.2](https://github.com/shenwei356/seqkit/releases/tag/v2.4.0)”。

# 环境要求

硬件要求

硬件要求如表2-1所示。

硬件要求

| 名称 | 说明 |
| --- | --- |
| 架构 | ARM |

软件要求

软件要求如表2-2所示。

注意

* 不同HPC应用的依赖软件不同，建议按照如下步骤判断其依赖软件：
* 1、查看其上游社区是否提供安装指导文档；
* 2、搜索网络上是否已经有社区提供相关安装文档或博客；
* 3、尝试安装该软件，根据报错情况，决定安装哪些依赖软件；
* 4、咨询华为工程师是否有相关经验。

软件要求

| 项目 | 版本 | 下载地址 |
| --- | --- | --- |
| Samtools | V1.10 | <https://github.com/samtools/samtools/releases/download/1.10/samtools-1.10.tar.bz2> |
| gmap | 2021-07-23 | <http://research-pub.gene.com/gmap/src/gmap-gsnap-2021-07-23.tar.gz> |
| minimap2 | V2.22 | <https://github.com/lh3/minimap2/archive/refs/tags/v2.22.tar.gz> |
| blat | V35.1 | [https://codeload.github.com/djhshih/blat/tar.gz/v35.1]( https:/codeload.github.com/djhshih/blat/tar.gz/v35.1) |
| pblat | V2.5 | <https://github.com/icebert/pblat/archive/refs/tags/2.5.tar.gz> |
| URI | 1.35 | <https://cpan.metacpan.org/authors/id/G/GA/GAAS/URI-1.35.tar.gz> |
| Pasa | 2.5.2 | <https://github.com/PASApipeline/PASApipeline/archive/refs/tags/pasa-v2.5.2.tar.gz> |

操作系统要求

操作系统要求如表2-3所示。

操作系统要求

| 项目 | 版本 | 下载地址 |
| --- | --- | --- |
| Kernel | 4.18.0 | Linux 4.18.0-147.el8.aarch64 |
| openEuler | 21.03 |  |

# 移植规划数据

本章节给出BUSCO软件在移植过程中涉及到的相关软件安装规划路径的用途及详细说明。

移植规划数据

| 序号 | 软件安装规划路径 | 用途 | 说明 |
| --- | --- | --- | --- |
| 1 | /usr/local/SAMTOOLS/ | Samtools的安装规划路径。 | 这里的安装规划路径只是一个举例说明，建议部署在共享路径中。需要根据实际情况调整，**后续章节凡是遇到安装路径的命令，都以现网实际规划的安装路径为准进行替换，不再单独说明。** |
| 2 | /usr/local/GMAP/ | gmap的安装规划路径。 |
| 3 | /usr/local/MINIMAP2/ | minimap2的安装规划路径。 |
| 4 | /usr/local/BLAT/ | blat的安装规划路径。 |
| 5 | /usr/local/PBLAT/ | pblat的安装规划路径。 |
| 6 | /usr/local/URI/ | URI的安装规划路径。 |
| 7 | /usr/local/PASA/ | pasa的安装规划路径。 |

# 配置编译环境

前提条件

使用SFTP工具将各安装包上传至服务器对应目录下。

配置流程

配置流程

| 序号 | 配置项 | 说明 |
| --- | --- | --- |
| 1 | 安装Samtools | 参考4.1 **安装**Samtools**软件**。 |
| 2 | 安装gmap | 参考4.2 **安装**gmap**软件**。 |
| 3 | 安装minimap2 | 参考4.3 **安装**minimap2**软件**。 |
| 4 | 安装blat | 参考4.4 **安装**blat**软件**。 |
| 5 | 安装pblat | 参考4.5 **安装**pblat**软件**。 |
| 6 | 安装URI | 参考4.6 **安装**URI**软件**。 |
| 7 | 安装pasa | 参考4.7 **安装**pasa**软件**。 |

## **4.1安装Samtools**

操作步骤

使用SSH工具，登录服务器。

下载并解压Samtools安装包。

**mkdir -p /usr/local/SMATOOLS/**

**cd /usr/local/SMATOOLS/**

**wget https://github.com/samtools/samtools/releases/download/1.10/samtools-1.10.tar.bz2**

**tar -xvf samtools-1.10.tar.bz2**

*注：请以压缩包实际名称进行解压*

编译安装。

**cd /usr/local/SMATOOLS/samtools-1.10**

**./configure --prefix=/usr/local/SMATOOLS/samtools-1.10 --without-curses CC=clang CXX=clang++ FC=flang**

**make -j8 && make install**

添加环境变量。

**export PATH=/usr/local/SMATOOLS/samtools-1.10/bin:$PATH**

**----结束**

## **4.2安装gmap**

操作步骤

使用SSH工具，登录服务器。

下载并解压gmap安装包。

**mkdir -p /usr/local/GMAP/**

**cd /usr/local/GMAP/**

**wget http://research-pub.gene.com/gmap/src/****gmap-gsnap-2021-07-23.tar.gz**

**tar -xvf gmap-gsnap-2021-07-23.tar.gz**

*注：请以压缩包实际名称进行解压*

编译安装。

**cd /usr/local/GMAP/gmap-gsnap-2021-07-23**

**sed -i 's/\_\_asm\_\_ ("cpuid"/\/\/\_\_asm\_\_ ("cpuid"/g' src/cpuid.c**

**sed -i 's/\_\_asm\_\_ ("xgetbv"/\/\/\_\_asm\_\_ ("xgetbv"/g' src/cpuid.c**

**./configure --prefix=/usr/local/GMAP/gmap-gsnap-2021-07-23 CC=clang CXX=clang++ FC=flang**

**make -j8 && make install**

添加环境变量。

**export PATH=/usr/local/GMAP/gmap-gsnap-2021-07-23/bin:$PATH**

**----结束**

## **4.3安装minimap2**

操作步骤

使用SSH工具，登录服务器。

下载并解压**minimap2**安装包。

**mkdir -p /usr/local/MINIMAP2/**

**cd /usr/local/MINIMAP2/**

**wget https://github.com/lh3/minimap2/archive/refs/tags/v2.22.tar.gz**

**tar -xvf minimap2-2.22.tar.gz**

*注：请以压缩包实际名称进行解压*

编译安装。

**cd minimap2-2.22/**

**sed -i '1 i CC= gcc' Makefile**

**make arm\_neon=1 aarch64=1**

**mkdir -p /usr/local/MINIMAP2/minimap2-2.22**

**cp minimap2 /usr/local/MINIMAP2/minimap2-2.22**

添加环境变量。

**export PATH=/usr/local/MINIMAP2/minimap2-2.22/:$PATH**

**----结束**

## **4.4安装blat**

操作步骤

使用SSH工具，登录服务器。

下载并解压blat安装包。

**mkdir -p /usr/local/BLAT/**

**cd /usr/local/BLAT/**

**wget https://codeload.github.com/djhshih/blat/tar.gz/v35.1**

**tar -xvf blat-35.1.tar.gz**

*注：请以压缩包实际名称进行解压*

编译安装。

**cd blat-35.1**

**export MACHTYPE=aarch64**

**make -j8**

**mkdir -p /usr/local/BLAT/blat-35.1**

**cp -r bin /usr/local/BLAT/blat-35.1**

添加环境变量。

**export PATH=/usr/local/BLAT/blat-35.1/bin:$PATH**

**----结束**

## **4.5安装pblat**

操作步骤

使用SSH工具，登录服务器。

下载并解压pblat安装包。

**mkdir -p /usr/local/PBLAT/**

**cd /usr/local/PBLAT/**

**wget https://github.com/icebert/pblat/archive/refs/tags/2.5.tar.gz**

**tar -xvf pblat-2.5.tar.gz**

*注：请以压缩包实际名称进行解压*

安装RPM包。

**cd pblat-2.5/**

**sed -i 's/x86\_64/aarch64/g' Makefile**

**sed -i 's/CFLAGS=-O -Wall/CFLAGS=-O -Wall -L.\/usr\/lib64 /g' Makefile**

**sed -i 's/HG\_INC=-I.\/inc -I.\/htslib/HG\_INC=-I.\/inc -I.\/htslib -**

**I.\/usr\/include/g' Makefile**

**sed -i 's/\&\& args != NULL//g' lib/htmshell.c**

**yumdownloader openssl-devel**

**yumdownloader krb5-devel**

**yumdownloader libcom\_err-devel**

**yumdownloader openssl-libs**

**rpm2cpio openssl-devel-1.0.2k-25.el7\_9.aarch64.rpm | cpio -ivdm**

**rpm2cpio krb5-devel-1.15.1-54.el7\_9.aarch64.rpm | cpio -ivdm**

**rpm2cpio libcom\_err-devel-1.42.9-19.el7.aarch64.rpm | cpio -ivdm**

**rpm2cpio openssl-libs-1.0.2k-25.el7\_9.aarch64.rpm | cpio -ivdm**

编译安装。

**make -j8**

**mkdir -p /usr/local/PBLAT/pblat-2.5/pblat**

**cp pblat /usr/local/PBLAT/pblat-2.5/pblat**

添加环境变量。

**export PATH=/usr/local/PBLAT/pblat-2.5/pblat/:$PATH**

**----结束**

## **4.6安装URI**

操作步骤

使用SSH工具，登录服务器。

下载并解压URI安装包。

**mkdir -p /usr/local/URI/**

**cd /usr/local/URI/**

**wget https://cpan.metacpan.org/authors/id/G/GA/GAAS/URI-1.35.tar.gz**

**tar -xvf URI-1.35.tar.gz**

*注：请以压缩包实际名称进行解压*

编译安装。

**cd URI-1.35/**

**perl Makefile.PL PREFIX=/usr/local/URI/URI-1.35/perl5**

**LIB=/usr/local/URI/URI-1.35/usr/lib64/perl5**

**make && make install**

下载RPM包。

**cd ..**

**yumdownloader perl-DBD-SQLite**

**yumdownloader perl-DBI**

**rpm2cpio perl-DBD-SQLite-1.39-3.el7.aarch64.rpm | cpio -ivdm**

**rpm2cpio perl-DBI-1.627-4.el7.aarch64.rpm | cpio -ivdm**

添加环境变量。

**export**

**PERL5LIB=/home/usr/local/URI/URI-1.35/perl5/:/home/usr/local/URI/URI-1.35/perl5**

**URI/:/home/usr/local/URI/URI-1.35/perl5/vendor\_perl/:/home/lipzh21/test/usr/lib6**

**4/perl5/vendor\_perl/auto/:/home/usr/local/URI/URI-1.35/perl5/vendor\_perl/DBD/:/h**

**ome/usr/local/URI/URI-1.35/perl5/vendor\_perl/DBI/:/home/lipzh21/test/usr/lib64/p**

**erl5/vendor\_perl/Bundle**

*注：请以自身实际目录为准*

**----结束**

## **4.7安装PASA**

操作步骤

使用SSH工具，登录服务器。

下载并解压pblat安装包。

**mkdir -p /usr/local/PASA/**

**cd /usr/local/PASA/**

**wget https://github.com/PASApipeline/PASApipeline/archive/refs/tags/pasa-v2.5.2.tar.gz**

**tar -xvf pasa-v2.5.2.tar.gz**

*注：请以压缩包实际名称进行解压*

编译安装并修改数据输出位置。

**cd PASApipeline-pasa-v2.5.2/**

**make -j8**

**cd sample\_data/**

**cp mysql.confs/alignAssembly.config ./**

**sed -i**

**'s/DATABASE=sample\_mydb\_pasa/DATABASE=\/usr\/local\/PASA\/PASApipeline-pasa-v2.5.2\/database.sqlite/g'**

**alignAssembly.config**

解压数据文件。

**gunzip -d genome\_sample.fasta.gz**

添加环境变量。

**export PATH=/usr/local/PASA/PASApipeline-pasa-v2.5.2/**

**----结束**

# 获取源码

操作步骤

下载pasa安装包“pasa-v2.5.2.tar.gz”。

下载地址：<https://github.com/PASApipeline/PASApipeline/archive/refs/tags/pasa-v2.5.2.tar.gz>

----结束

# 运行和验证

操作步骤

使用SSH工具，登录服务器。

进入安装路径。

**cd /usr/local/PASA/PASApipeline-pasa-v2.5.2/sample\_data/**

执行命令

**{ time Launch\_PASA\_pipeline.pl -c alignAssembly.config -C -R -g**

**genome\_sample.fasta -t all\_transcripts.fasta.clean -T -u all\_transcripts.fasta -**

**f FL\_accs.txt --ALIGNERS blat,gmap --CPU 96; } 2>&1 |tee PASA.log**

# 修订记录

| 发布日期 | 修订记录 |
| --- | --- |
| 2023-03-24 | 第一次正式发布。 |