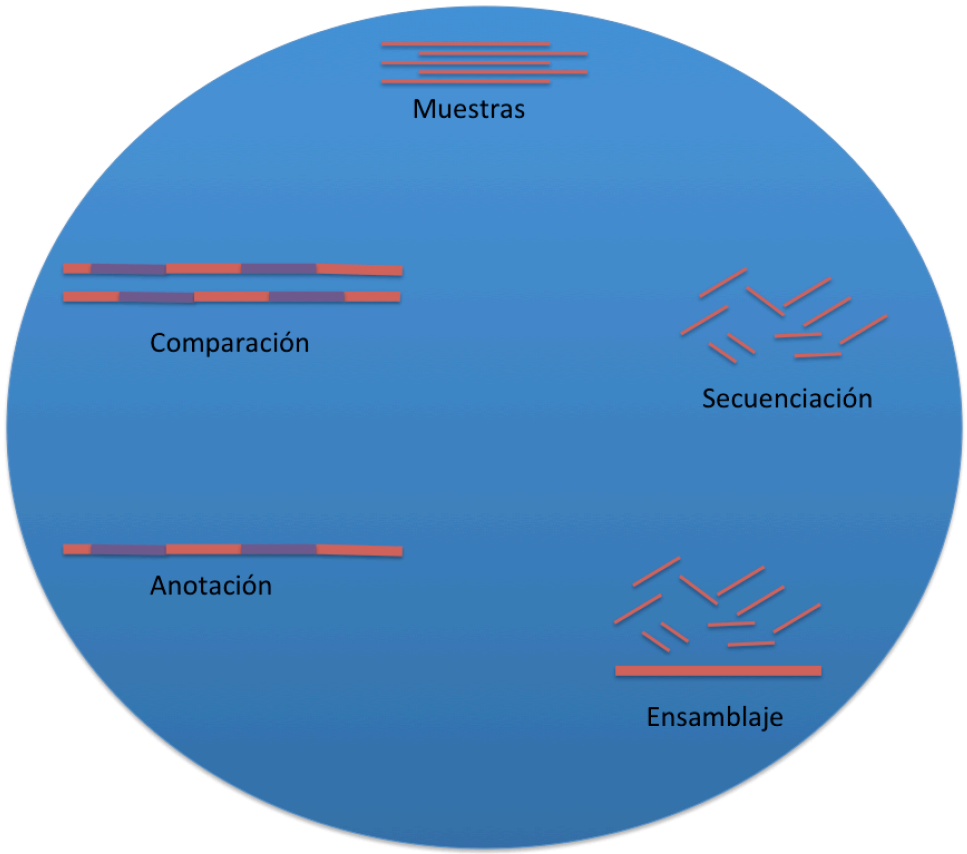


Herramientas de Ensamblaje de Genomas

Ensamblaje de Genomas

La genómica se caracteriza por el estudio de los genomas y esto se logra con un gran número de genes obtenidos con herramientas que coleccionan estos datos.

El ensamblado de genomas es parte de estas herramientas lo cual busca encontrar una representanción del genoma en estudio.

Esta tarea no es nada fácil, por lo que se han creado diferentes algoritmos para la reconstrucción de las secuencias de genes.

Estos algoritmos automatizados son los encargados de realizar lecturas cortas de una secuencia. Actualmente existen herramientas automatizadas que tenemos la facilidad de instalar en nuestras computadoras para poder realizar diferentes estudios entre ellas el ensamblaje de genomas.

Se mostrará paso a paso la implementación de ciertas herramientas que son necesarias para realizar el ensamblaje de genomas desde nuestras computadoras.

Herramientas a implementar

La implementación de las herramientas de este tutorial será en el sistema operativo ubuntu.

Las herramientas que se mostrará como instalar son:

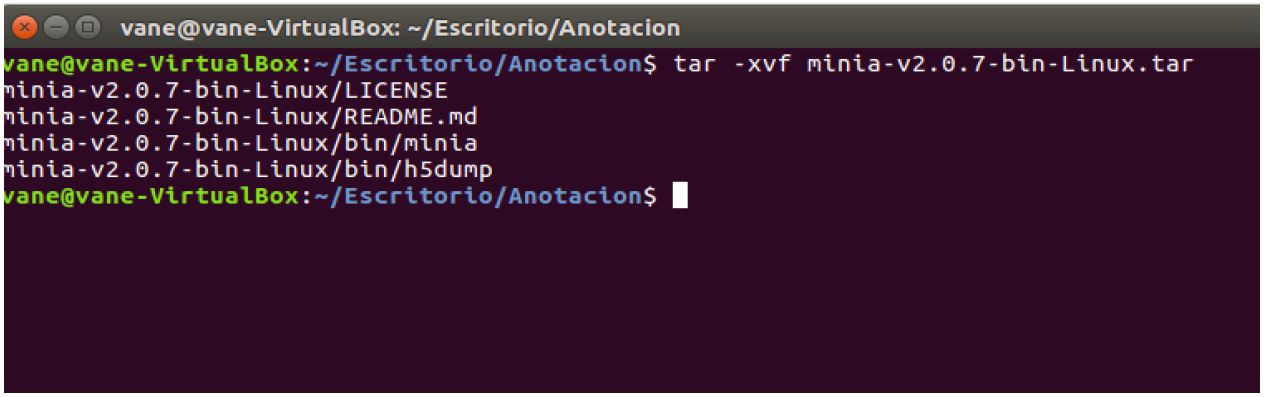
- Minia.
- KmerGenie
- QUAST
- SSPACE.

Hay pre-requisitos para estas herramientas que generalmente se encuentran instaladas de forma predeterminada en los sistemas operativos como ubuntu o biolinux que son:

- python
- R
- RStudio
- perl

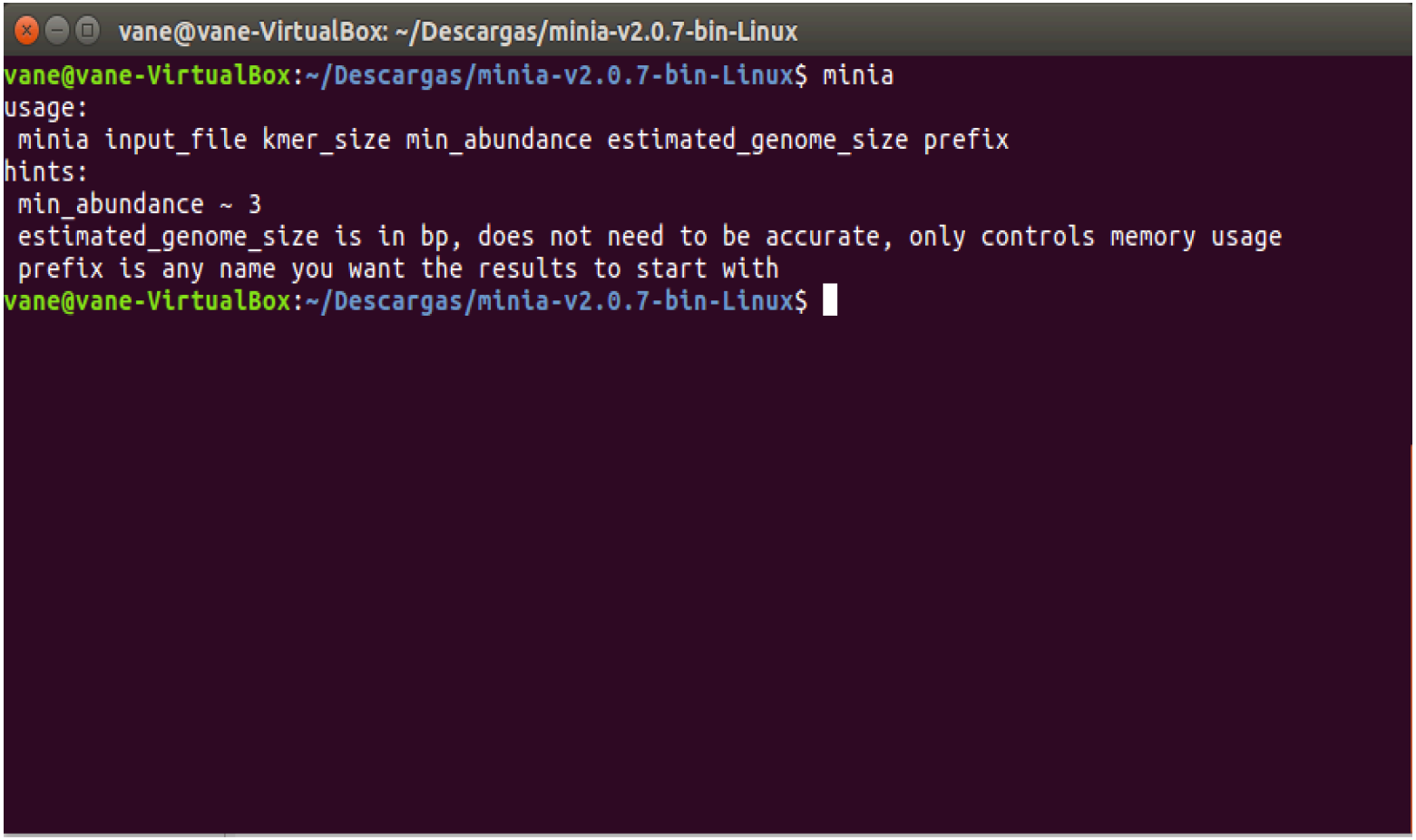
MINIA

Minia es un ensamblador muy potente, basado en el algoritmo del gráfico de Bruijn. Minia produce resultados de contiguidad y precisión.



En la siguiente dirección http://minia.genouest.org/ se podrá descargar el paquete que contiene el programa.
Una vez descargado se desempaquetará el programa con el siguiente comando mostrado en la imagen.

Una vez desempaquetado, vamos a compilar Minia con el siguiente comando visto en la imagen, donde "**sudo**" pide permiso administrador por lo que es importante poner la contraseña del usuario de sesión.



Una vez realizada la instalación se digitará en la terminal “**minia**” y el nos mostrará , una pequeña descripción de uso de la herramienta.

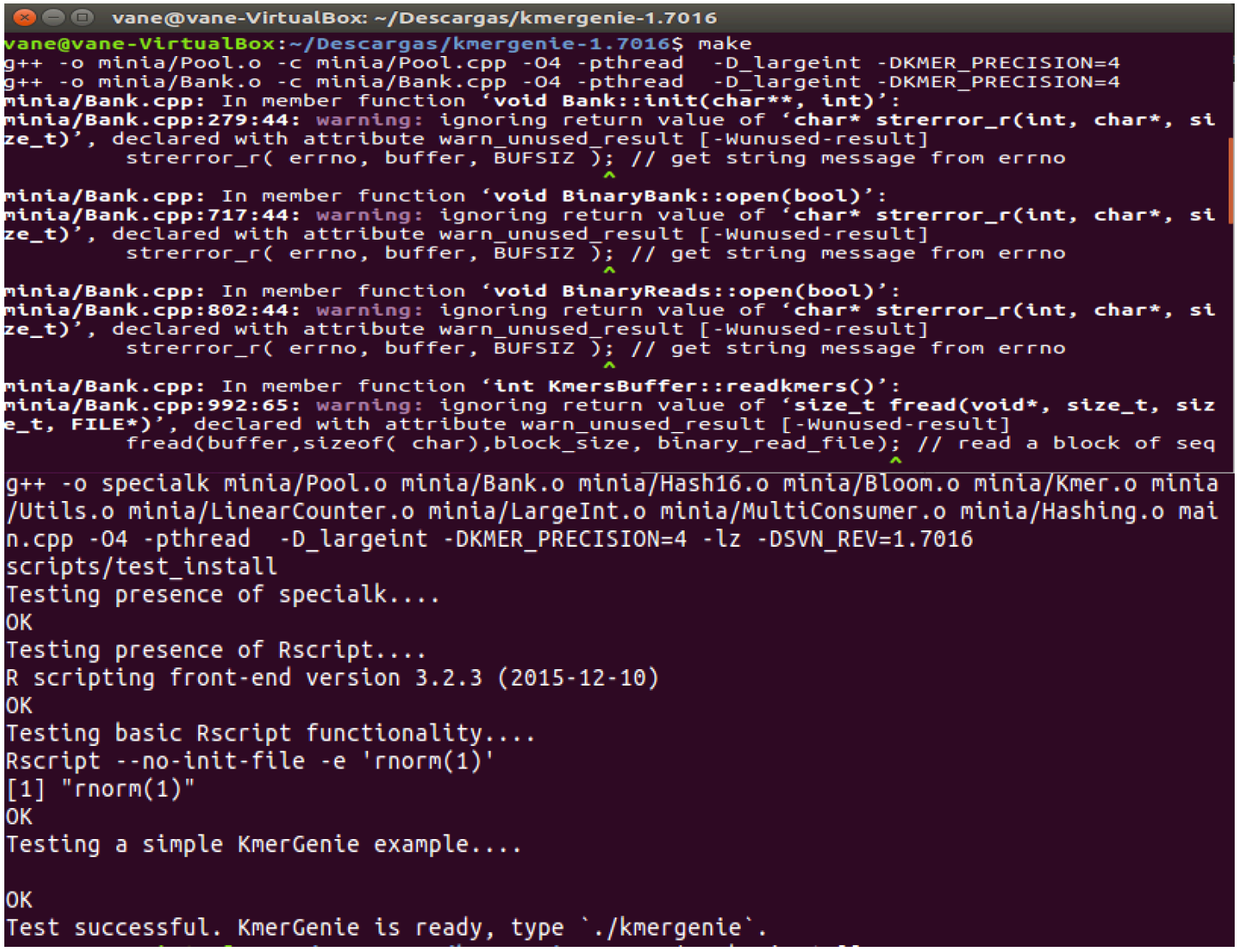
KMERGENIE

Es una herramienta complementaria de Minia que permite realizar predicciones de genomas con la longitud del tamaño del genoma que se necesite denominado K, también permite optimizar más fácilmente los parámetros K utilizados y generar histogramas de abundancia que permite al usuario tomar decisiones.

```
vane@vane-VirtualBox: ~/Descargas
vane@vane-VirtualBox:~/Descargas$ tar -xzvf kmergenie-1.7016.tar.gz
kmergenie-1.7016/
kmergenie-1.7016/third_party/
kmergenie-1.7016/third_party/docopt.pyc
kmergenie-1.7016/third_party/__init__.pyc
kmergenie-1.7016/third_party/docopt.py
kmergenie-1.7016/third_party/__init__.py
kmergenie-1.7016/minia/
kmergenie-1.7016/minia/MultiConsumer.cpp
kmergenie-1.7016/minia/LargeInt.cpp
kmergenie-1.7016/minia/Pool.cpp
kmergenie-1.7016/minia/Hashing.h
kmergenie-1.7016/minia/Kmer.h
kmergenie-1.7016/minia/Utils.cpp
kmergenie-1.7016/minia/Hash16.h
kmergenie-1.7016/minia/Utils.h
kmergenie-1.7016/minia/LinearCounter.h
kmergenie-1.7016/minia/Kmer.cpp
kmergenie-1.7016/minia/Bank.h
kmergenie-1.7016/minia/Pool.h
kmergenie-1.7016/minia/Bloom.cpp
kmergenie-1.7016/minia/LinearCounter.cpp
kmergenie-1.7016/minia/Hash16.cpp
kmergenie-1.7016/minia/Bank.cpp
```

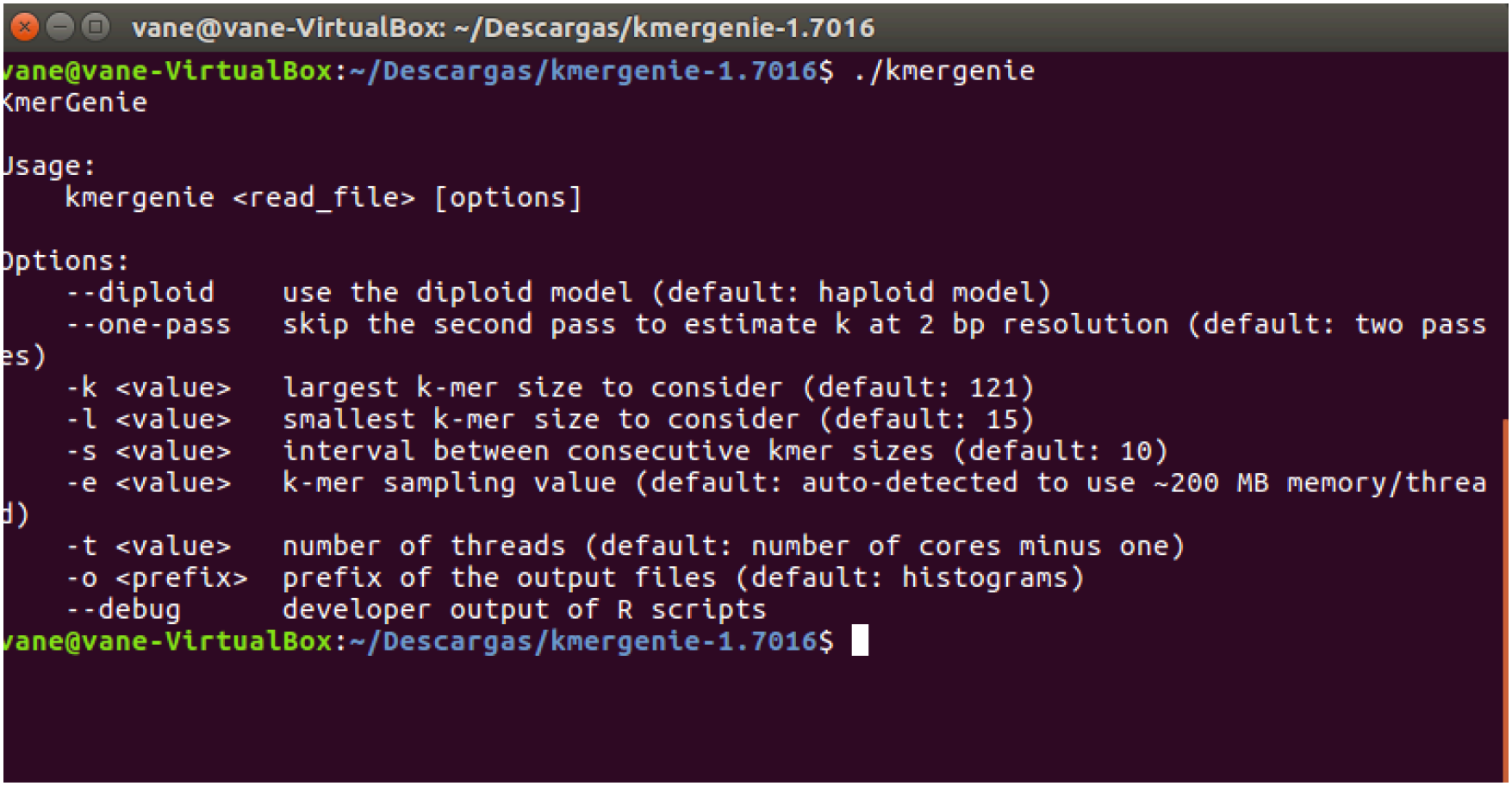
En la siguiente dirección se podrá descargar la herramienta http://kmergenie.bx.psu.edu/.

Se navegará a la carpeta donde esta ubicada la herramienta con el comando "**cd**".y se descomprime para poder compilar KmerGenie.



Una vez descomprimido, se ingresa a la carpeta de Kmergenie con el comando "**cd**", se coloca el comando "**make**" que es el que compilará la herramienta y se deberá ver como en la pantalla.

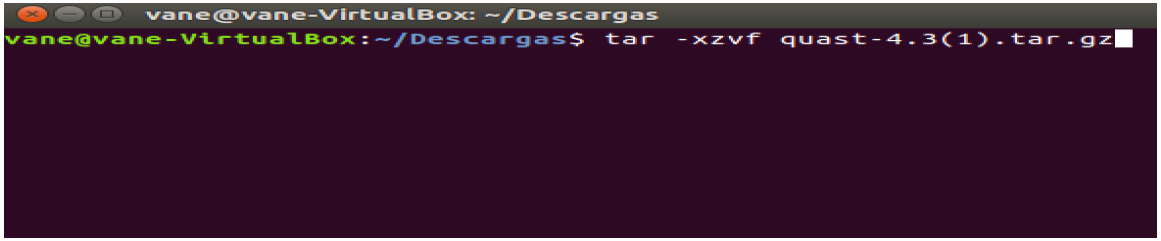
Al digitar en la terminal “./**kmergenie**" se despliega diferentes opciones de uso de la herramienta kmergenie. Y estará lista para funcionar.



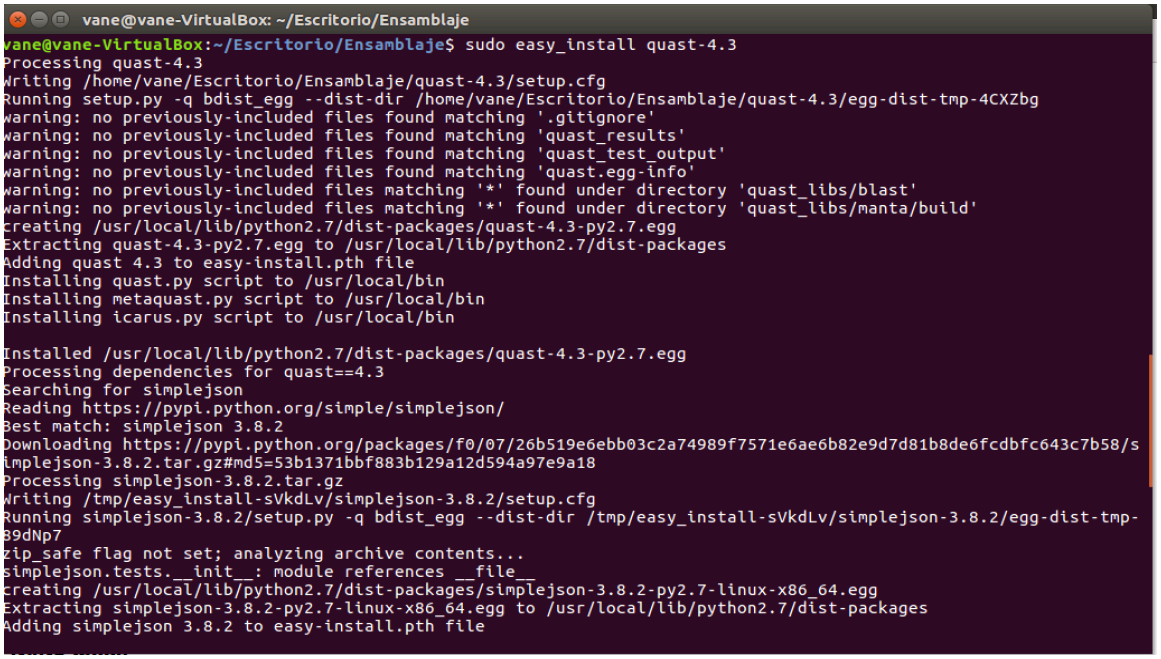
QUAST

QUAST es una herramienta para evaluar y comparar el ensamblaje del genomas ya sea con o sin genomas de referencia. QUAST produce tablas y gráficos interactivos y fácil de interpretar.

En la siguiente dirección http://bioinf.spbau.ru/quast se descargar el paquete que contiene el programa y se descomprime como se ve en la imagen.



Se ingresa con el comando "**cd**" a la carpeta QUAST y se escribirá en la terminal "**./setup.py install**" que es la instalación básica, o ya bien se podría la instalación completa "**./setup.py install_full**" pero hay que tener presente que esta instalación completa pesa cerca de 800MB por lo que deben asegurarse de que tienen la memoria suficiente.



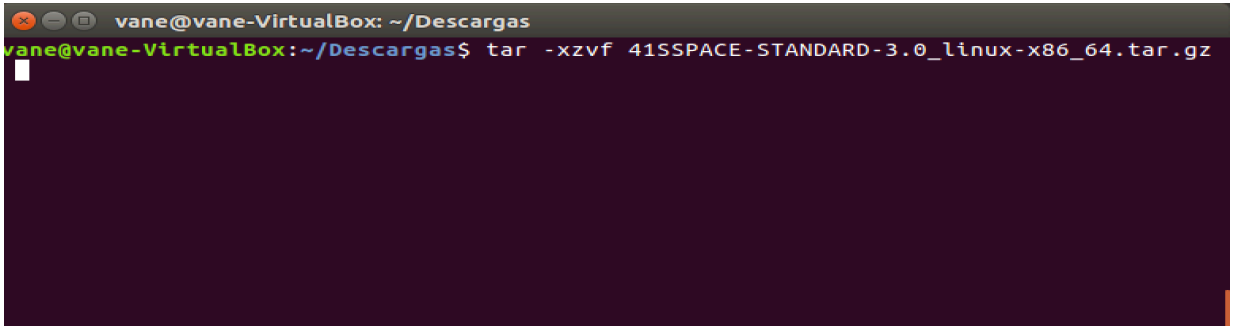
SSPACE

Es un programa para scaffold (super conjuntos de clones pre ensamblados) que mejora el tiempo de ejecución y muestra resultados prometedores donde la cantidad de conjuntos de clones iniciales se reducen considerablemente.

Sspace es capaz de evaluar el orden, la distancia y la orientación de los conjuntos de clones y combinar los super conjunto de clones, además utiliza formato FASTA y NGS (secuenciación de próxima generación).

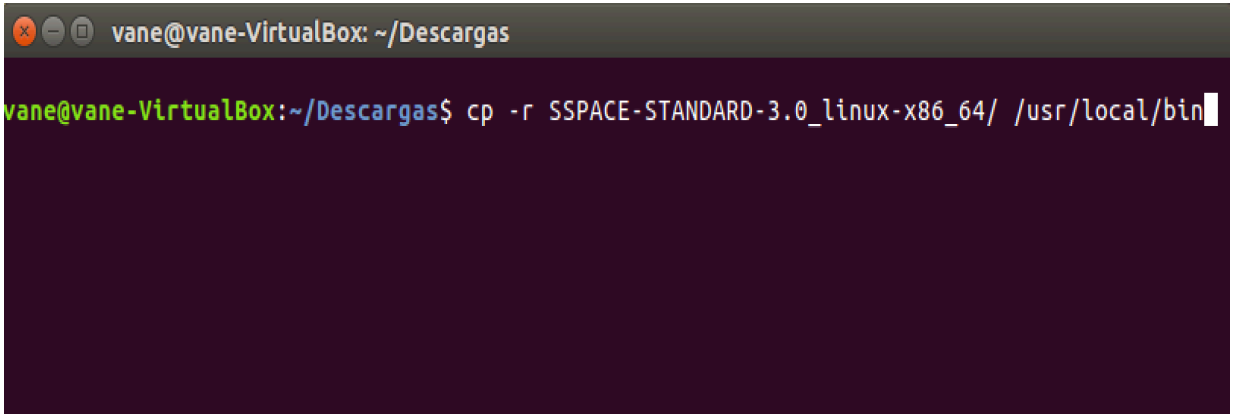
http://www.baseclear.com/genomics/bioinformatics/basetools/SSPACE es la página oficial para descargar SSPACE hay que llenar un pequeño formulario, una vez completando se podrá descargar el paquete.

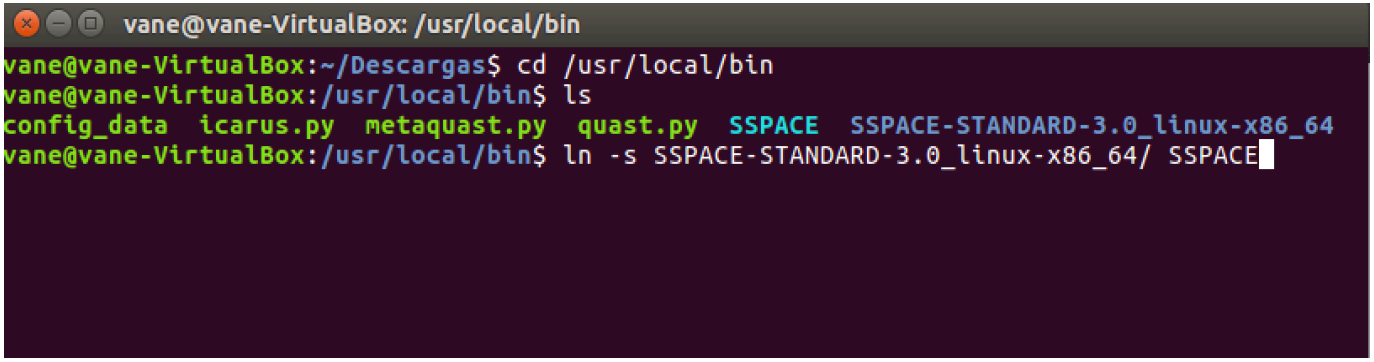
Como se ha venido realizando se desempaquetará el programa con el comando "**tar –xzvf**".



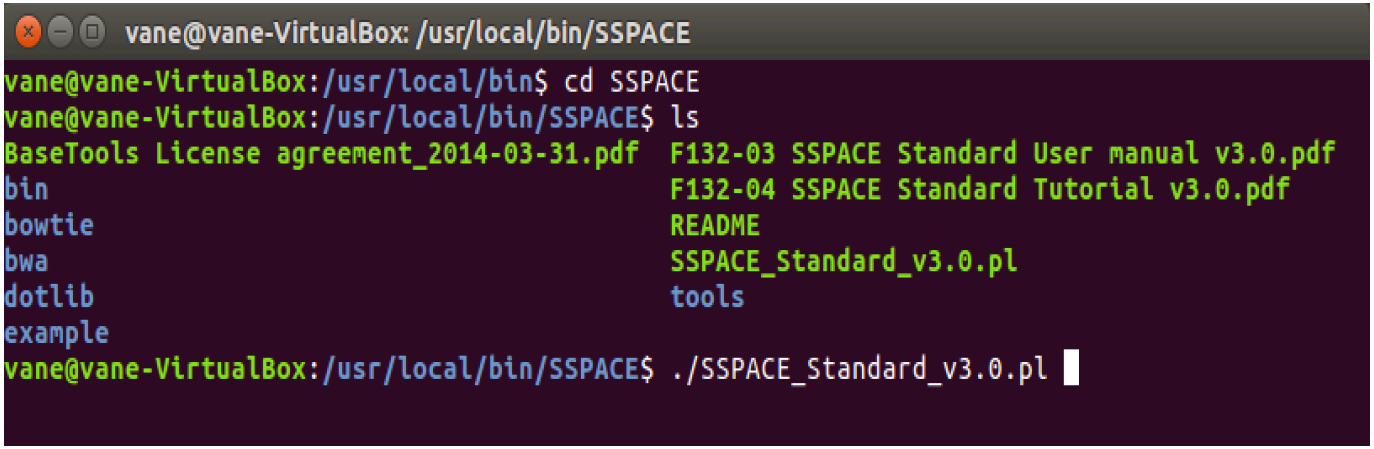
Está herramienta necesita una librería que en algunos casos no se encuentra instalada por lo que se recomienda ejecutar el siguiente comando "**sudo cpan Perl4::CoreLibs**" antes de iniciar la instalación de SSPACE.

Ahora, se utiliza el comando "**cp -r**" que servirá para copiar la carpeta SSPACE a /usr/local/bin. El sistema de archivos /usr es donde normalmente se encuentran instalados todos los programas.



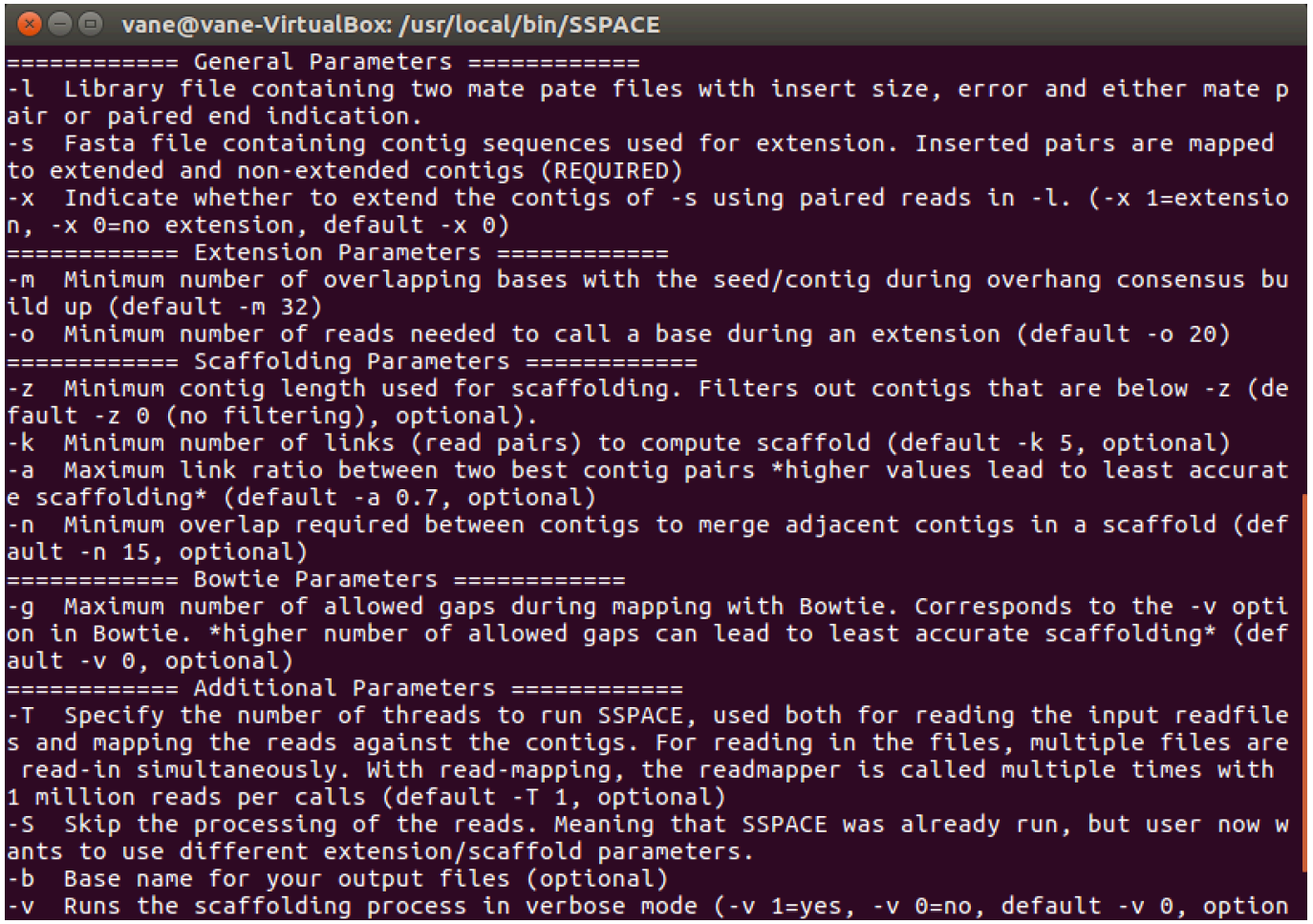


El primer comando "**cd**" navegará a la carpeta donde se encuentra SSPACE.
"**ln –s**" realizará una version independiente de la carpeta original llamada "SSPACE-STANDARD-3.0_linux-x86_64" y la llamé SSPACE.



Con el comando "**cd**" se ingresó a la nueva carpeta llamada SSPACE, y con "**ls**" se desplegará el listado con todos los paquetes e información que contiene SSPACE.

El comando “**SSPACE_Standard_v3.0.pl**” ejecutará el programa y deberá aparecer una pantalla muy similar a la que se muestra.



Las herramientas están listas para funcionar.!