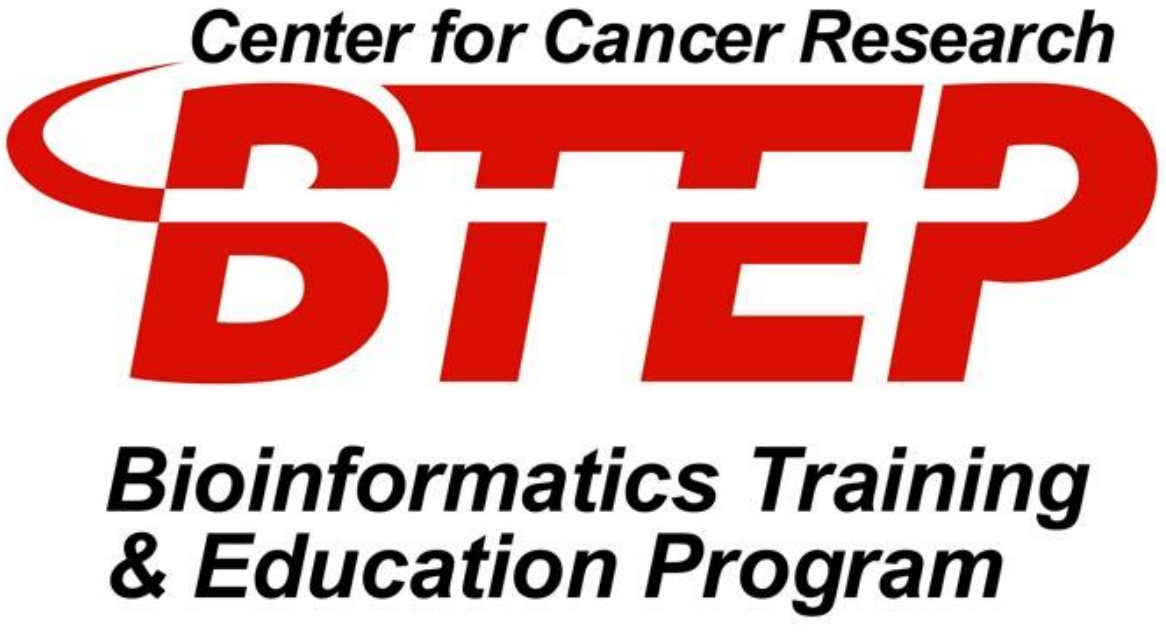
Unix Commands for Beginners



Alexandra L Emmons Ph.D. & Joe Wu Ph.D.
BTEP/GAU/CCR/NCI/NIH - email ncibtep@mail.nih.gov
Bioinformatics Training and Education Program

Table of Contents

Class overview

Class overview slides

Unix commands for beginners

Unix Linux Command line High Performance Computing HPC Biowulf Bioinformatics Data science

Course Overview and Learning Objectives

This class will introduce **novices** to essential Unix commands for getting started in bioinformatics. Unix is an operating system like Windows and MacOS. However, in Unix, users interact with the computer by issuing commands rather than through a point-and-click interface. The ability to use Unix is important as many bioinformatics software are written to work on Unix and Unix-like operating systems. For instance, see the list of software available on the Bioconda *(https://bioconda.github.io/#)* repository. Further, bioinformatics often deals with large and complex datasets such as those derived from Next Generation Sequencing (NGS), which are too cumbersome to analyze on a personal computer with limited computation power. Thus, bioinformatics is commonly performed on high performance computing systmes such as the one at NIH known as Biowulf *(https://hpc.nih.gov/systems/)*. Biowulf runs on Linux (a Unix-like operating system) and has around 1000 scientific applications installed. Biowulf staff maintain and update the system as well as the installed software.

Learning Objectives

After this class, participants will know the essential commands needed for getting started with bioinformatics. These commands will enable participants to sign onto the Biowulf cluster, navigate through its directories, and work with files. Participants will not become experts in Unix or bioinformatics after this class, however the material presented form the basis for participants to continue learning advanced Unix skills for bioinformatics.

Course Prerequisites

There are no prerequisites for attending this class although it is targeted towards the novices. A Biowulf account is not required for participation.

Class overview slides

Unix for Bioinformatics Beginners

- `ssh` (connect to remote computer such as a high performance computing cluster)
- `pwd` (print working directory)
- `cd` (change directory)
- `ls` (list directory content)
- `mkdir` (make a directory)
- `touch` (creates an empty file)
- `nano` (basic editor for editing text files)
- `cp` (make a copy of files or folders)
- `mv` (rename files/folder or move files or folders)
- `rm` (to remove files or folders)
- `cat` (print file content to screen)
- `less` (page through file content)
- `column` (display tabular data nicely aligned on screen)
- `wc` (word count, line count and character count)
- `grep` (pattern search)
- `sinteractive` (request compute and time resources on Biowulf's compute node to perform data analysis and file transfers)
- `module` (find, load, and unload applications installed on high performance computing clusters)
- `man` (view manual for a command)

> **Tip**
>
> All Unix commands follow the syntax of `command option(s) input`. Think of the `command` as a verb in natural language. The `input` is what users would like the command to act on and `option(s)` changes the default way in which a command runs.

Connecting to Remote Computer

As bioinformatics often deal with dataset that are too large to be analyzed using compute resources on a personal computer, a motive for learning Unix command line is to enable scientists to perform bioinformatics analyses on high performance computing systems such as Biowulf at NIH. To connect to a remote computer, the `ssh` command construct below can be used and the breakdown is as follows.

- `ssh`: The command for connecting to a remote computer.
- `username`: The user name that is used to sign onto a remote computer such as high performance computing (for Biowulf, `username` will be the user's NIH user name). Username is followed by `@` (ie. the "at" symbol).

- `remote`: This is the name of the remote computer (Biowulf would be `biowulf.nih.gov`) to connect to.

```
ssh username@remote.computer
```

The command above could be annunciated as `ssh username at remote.computer`.

To sign onto Biowulf, the NIH high performance computing system, use the following. Again, replace username with the user's NIH user name.

```
ssh username@biowulf.nih.gov
```

Next, enter the user specific password for the remote computer. Note that nothing shows up when the password is being type, but keep typing and hit enter when ready to sign on.

The Command Prompt

In Unix, the command prompt is where users issue commands to interact with the computer. For example, upon connecting to Biowulf, users will see the prompt shown below. Replace username with the user's NIH user name. The prompt below tells users that they are connected to Biowulf and are currently in the `home` directory (denoted by ~).

```
[username@biowulf ~]$
```

Caution

The prompt will look different depending on the computer system. It can also be customized, so do not panic if a prompt does not look exactly like that above.

Print the Working Directory

Definition

The working directory is the folder in a computer system in which a user is currently in.

It is a good idea to know which directory a user is currently in as data analysis progresses. To check, use the following.

```
pwd
```

For instance, upon signing on to Biowulf users will land in their `home` directory (also denoted by ~). The `pwd` command will give the directory path below (replace username with the user's specific user name).

```
/home/username
```

> **Note**
>
> In Unix, a directory path indicates where in the file system hierarchy a user is. Directory paths that start with `/` are known as absolute. Each part of the directory path is separated by `/` and followed by the name of the folder in the file system hierarchy.

Change Directory

> **Note**
>
> The user's `home` directory on Biowulf is not the place to store analysis input and output as it only has 16gb of storage space and users cannot request an increase in quota. The user's `data` directory can be used for this as storage quota can be increased at the request of the user.

To change directory use the `cd` command follow by the directory to change into. Upon signing onto Biowulf, users should change into their `data` directory to performing analyses. To do this use the following. Replace username with the user's specific user name.

```
cd /data/username
```

The `pwd` command can be used to check whether the directory change is successful.

```
pwd
```

If the directory change is successful, the following path will be shown, again replace username with the user's NIH user name.

```
/data/username
```

To go back to the user's home directory use `cd`, `cd ~`, or `cd /home/username`. Use `cd ..` to go back one directory. To go back two directories, use `cd ../../`. To go back three directories use `cd ../../../` and so on.

> **Note**

> The examples below will be done using a folder called `unix_demonstration` in the instructor's `/data` folder. There is no need to follow along. To change into `unix_demonstration` from the `/data` folder, just do `cd unix_demonstration`.

Listing Directory Content

To view the files and subfolder within a directory, the `ls` command can be used. For instance, `ls` will retrieve the following items in the `unix_demonstration` directory when issued.

```
file.txt
```

Recall that a Unix command come with options that alter its default behavior. The `-l` option of `ls` gives a detailed view of each item in a directory. For instance, on the left hand column of the table below, lines that start with `-` are files while those starting with `d` are folders. The fifth column list the file or folder size. Note file.txt has a file size of 0 bytes since it is empty. Also, `ls -l` by default displays the files size in bytes.

```
-rw-r-----. 1 owner group    0 Jun 15 16:20 file.txt
```

Creating a New Directory

To make a new folder, use the `mkdir` command follow by the name of the folder.

```
mkdir folder
```

> **Recall**
>
> As mentioned previously, if the first column in the `ls -l` results starts with `d` then this indicates a directory, which is what `folder` should be.

```
ls -l
```

```
-rw-r-----. 1 owner group    0 Jun 16 10:09 file.txt
drwxr-x---. 2 owner group 4096 Jun 16 10:18 folder
```

Editing Files

Nano is a basic file editor that is built into Unix and enables editing of plain text files including txt, csv, fasta, genbank, gtf, and fastq. To edit a file using Nano, just use `nano` followed by the file name.

```
Unix is an operating system, just like Windows or MacOS.
Linux is a variety of Unix, and sometimes the names are used interch
```

For instance, to add the text above to file.txt, use:

```
nano file.txt
```

Then copy and paste the text into the editor. Hit control-x and then save to go back to the command prompt.

Note

If a file does not exist, then the `nano` command will create it as well as open the editor to enable editing. To create a blank file without opening the Nano editor, use the `touch` command instead.

```
touch bioinformatics.txt
```

```
ls -l
```

```
-rw-r-----. 1 owner group    0 Jun 16 18:13 bioinformatics.txt
-rw-r-----. 1 owner group  135 Jun 16 18:13 file.txt
drwxr-x---. 2 owner group 4096 Jun 16 10:18 folder
```

Copying Files and Folders

To copy a file use `cp` followed by the file name and the name of the duplicate.

For instance, to make a copy of bioinformatics.txt called btep_bioinformatics.txt, do:

```
cp bioinformatics.txt btep_bioinformatics.txt
```

```
ls
```

```
bioinformatics.txt  btep_bioinformatics.txt  file.txt  folder
```

The `cp` command can be used to copy a file into a folder. Just provide the file and the path to the folder in which to copy it. If a new name is desired for the duplicate, then append it to the destination folder path separated by a `/`. For instance, the command below will make a duplicate of bioinformatics.txt in `folder` called bioinformatics_for_beginners.txt.

```
cp bioinformatics.txt folder/bioinformatics_for_beginners.txt
```

Tip

To see contents of a folder other than those in the working directory, supply the path to the folder to the `ls` command.

```
ls folder
```

```
bioinformatics_for_beginners.txt
```

To copy a folder include the `-r` option in `cp`. The `-r` option recursively copies everything in a folder. The arguments are the folder to copy and the name of the duplicate folder.

```
cp -r folder bioinformatics_folder
```

The `-1` option of `ls` prints directory content one item per line.

```
ls -1
```

```
bioinformatics_folder
bioinformatics.txt
btep_bioinformatics.txt
file.txt
folder
```

```
ls bioinformatics_folder
```

```
bioinformatics_for_beginners.txt
```

Renaming Files and Folders

The `mv` command can be used to rename files and folders.

To rename a file, use the `mv` command then provide the name of the file to rename and the new name of the file.

```
mv btep_bioinformatics.txt bioinformatics_for_noobies.txt
```

```
ls -1
```

```
bioinformatics_folder
bioinformatics_for_noobies.txt
bioinformatics.txt
file.txt
folder
```

To rename a folder, use the `mv` command and then provide the name of the folder to rename and the new name of the folder.

```
mv folder project_folder
```

```
ls -1
```

```
bioinformatics_folder
bioinformatics_for_noobies.txt
bioinformatics.txt
file.txt
project_folder
```

Moving Files and Folders

To move a file from one folder to another, the `mv` command can be used. The arguments in this application of the `mv` command is the file to be moved and the name of the folder in which the file will be moved to.

```
mv bioinformatics.txt bioinformatics_folder
```

```
ls bioinformatics_folder
```

```
bioinformatics_for_beginners.txt  bioinformatics.txt
```

Tip

`mv` can also move folders. Just supply the name of the folder to be moved as the first argument and the path of the destination folder.

Deleting Files and Folders

To delete a file use the `rm` command followed by the file name.

For instance, to delete bioinformatics_for_noobies.txt do:

```
rm bioinformatics_for_noobies.txt
```

Warning

There is no trash can or recycling bin in Unix. Once a file is deleted, it cannot be recovered. Use the `-i` option with `rm` to confirm deletion.

```
rm -i bioinformatics_for_noobies.txt
```

Type `n` for no and `y` for yes (to delete).

```
rm: remove regular empty file 'bioinformatics_for_noobies.txt'? n
```

Empty folders can be deleted using `rmdir`. On the other hand, folders with content can be removed using `rm -r`, where the `-r` option tells `rm` to delete the folder and recursively delete everything inside that folder.

```
rm -r -i project_folder
```

The command above will ask if the user wants to delete the project_folder and everything in it. Type `y` to confirm deletion. Users will be asked to confirm the deletion of each item in the folder.

```
rm: descend into directory 'project_folder'? y
rm: remove regular empty file 'project_folder/bioinformatics_for_beg
rm: remove directory 'project_folder'? y
```

```
ls -1
```

```
bioinformatics_folder
bioinformatics_for_noobies.txt
file.txt
```

Viewing File Content

The `cat` command will print the entire content of a file to the terminal screen.

For instance, to view the content of file.txt in the terminal screen do:

```
cat file.txt
```

```
Unix is an operating system, just like Windows or MacOS.
Linux is a variety of Unix, and sometimes the names are used interch
```

The `cat` command can also be used to view fastq or fq files, which contain sequences from high throughput sequencers. For long files that take a while to completely print, hit control-c to exit the `cat` command and return to the prompt.

```
cat example_fastq.fq
```

Tabular data in the form of csv files can be displayed by `cat` as well.

```
cat example_rna_sequencing_counts.csv
```

Paging through Files

Rather than printing file content in its entirety to the terminal, users can page through files using `less`.

```
less example_rna_sequencing_counts.csv
```

Users can use the up and down arrow on the keyboard to scroll up/down the file to view content. The up/down arrows enable scrolling line by line.

Printing Tabular Data to Terminal Nicely Aligned

For tabular data in the form of csv files, which could contain multiple columns, the columns do not print to the terminal nicely aligned. The `column` command can fix this.

The options and arguments in the column command include:

- Option `-t`: Creates a table.
- Option `-s`: Prompts users to provide the column separator (ie. comma for csv files)
- Argument: Name of the file which users would like to view (ie. example_rna_sequencing_counts.csv)
- `|` is the pipe operator, which sends output from one command as input for another (ie.`less` to enable paging through the data table)

```
column -t -s ',' example_rna_sequencing_counts.csv | less
```

{{Sdet}}{{Ssum}}`column` command output{{Esum}}

```
Geneid              HBR_1.bam  HBR_2.bam  HBR_3.bam  UHR_1.bam  UHR_1
U2                  0          0          0          0          0
CU459211.1          0          0          0          0          0
CU104787.1          0          0          0          0          0
BAGE5               0          0          0          0          0
ACTR3BP6            0          0          0          0          0
5_8S_rRNA           0          0          0          0          0
AC137488.1          0          0          0          0          0
AC137488.2          0          0          0          0          0
CU013544.1          0          0          0          0          0
CT867976.1          0          0          0          0          0
```

```
CT867977.1          0       0       0       0       0
CT978678.1          0       0       0       0       0
CU459202.1          0       0       0       0       0
AC116618.1          0       0       0       0       0
CU463998.1          0       0       0       0       0
CU463998.3          0       0       0       0       0
CU463998.2          0       0       0       0       0
U6                  0       0       0       0       0
LA16c-60D12.1       0       0       0       3       2
LA16c-13E4.3        0       0       0       0       0
LA16c-60D12.2       0       0       0       0       4
ZNF72P              0       0       0       0       1
BNIP3P2             0       0       0       0       2
YME1L1P1            0       0       0       0       0
OR11H1              0       0       0       0       0
LA16c-60G3.6        0       0       0       0       3
ARHGAP42P3          0       0       0       0       2
LA16c-23H5.4        0       0       0       0       1
LA16c-60G3.8        0       0       0       0       2
LA16c-60G3.7        0       0       0       0       0
LA16c-60G3.5        0       0       0       0       0
LA16c-2F2.8         0       0       0       0       3
NEK2P2              0       0       0       0       1
LA16c-2F2.5         0       0       0       0       0
NF1P6               0       0       0       1       4
MED15P7             0       0       0       0       2
POTEH               0       0       1       7       7
POTEH-AS1           0       0       0       0       2
RNU6-816P           0       0       0       0       0
LA16c-3G11.7        0       0       0       2       1
LA16c-83F12.6       0       0       0       3       1
GRAMD4P2            0       0       0       0       1
DUXAP8              10      4       9       250     199
LL22NC03-N64E9.1    0       1       0       16      9
BMS1P22             0       1       3       25      24
LL22NC03-N14H11.1   4       2       0       98      38
NBEAP3              0       0       0       5       3
LA16c-60H5.7        0       0       0       11      6
LA16c-4G1.4         0       0       0       0       0
LA16c-4G1.3         0       0       0       0       3
AC092854.1          0       0       0       0       0
ABCD1P4             0       0       0       0       2
LA16c-17H1.3        0       0       0       0       0
PABPC1P9            0       0       0       0       1
SLC9B1P4            0       0       0       0       1
ACTR3BP2            0       0       0       0       0
CHEK2P4             0       0       0       0       0
```

KB-67B5.17	0	0	0	0	0
KB-67B5.12	0	0	0	0	0
KCNMB3P1	0	0	0	2	1
CCT8L2	0	0	0	1	2
FABP5P11	0	0	0	0	0
AP000547.1	0	0	0	0	1
TPTEP1	19	23	20	141	94
KB-7G2.9	0	0	1	3	4
SLC25A15P5	0	0	0	2	2
PARP4P3	0	0	0	1	1
ANKRD62P1-PARP4P3	0	1	0	3	3
ANKRD62P1	0	0	0	2	0
KB-7G2.8	0	0	0	0	0
VWFP1	2	2	4	1	2
AC005301.8	0	0	0	1	0
XKR3	0	0	0	3	4
HSFY1P1	0	0	0	0	0
GPM6BP3	0	0	0	0	0
AC007064.22	0	0	0	0	0
ZNF402P	0	0	0	1	1
AC007064.24	1	3	2	3	0
MTND1P17	0	0	0	0	0
AC007064.25	0	0	0	0	0
AC006548.19	0	0	0	0	0
IGKV1OR22-5	0	0	0	0	0
IGKV2OR22-4	0	0	0	0	0
IGKV2OR22-3	0	0	0	0	0
IGKV3OR22-2	0	0	0	0	1
IGKV1OR22-1	0	0	0	0	0
AC006548.28	1	1	1	4	40
GAB4	0	0	0	0	0
AC006548.26	0	0	0	0	0
VN1R9P	0	0	0	0	0
CECR7	22	27	32	31	27
AC006946.16	7	6	4	3	3
AC006946.17	0	2	2	0	3
IL17RA	58	65	58	95	71
AC006946.12	0	1	1	1	1
CECR6	26	45	32	9	6
AC006946.15	3	3	1	0	2
CECR5	30	44	33	127	82
CECR5-AS1	0	2	2	1	2
CECR1	30	28	32	39	35
AC005300.5	0	0	0	1	0
FAM32B	0	1	0	0	0
CECR3	1	0	1	0	1
CECR9	0	0	0	0	0

RN7SL843P	0	0	0	0	0
CECR2	115	88	119	119	81
AC004019.10	1	5	1	0	1
CLCP1	1	2	0	0	0
DNAJA1P6	0	0	0	0	0
SLC25A18	100	111	74	6	8
AC004019.13	1	2	1	1	2
ATP6V1E1	392	454	391	244	181
BCL2L13	212	284	264	244	166
BID	70	85	84	112	76
MIR3198-1	0	0	0	0	0
LINC00528	1	0	0	1	2
MICAL3	339	429	425	211	145
XXbac-B461K10.4	6	6	9	1	2
MIR648	0	0	6	0	0
XXbac-B476C20.14	2	7	2	2	1
XXbac-B476C20.13	2	9	4	1	2
RHEBP3	1	2	3	0	0
FLJ41941	0	0	0	0	2
XXbac-B476C20.9	3	5	4	2	3
PEX26	135	174	137	205	135
XXbac-B476C20.11	0	0	1	3	0
XXbac-B476C20.10	0	0	0	0	0
ARL2BPP10	0	0	0	0	0
TUBA8	13	13	15	2	2
AC008079.10	55	68	53	14	14
USP18	3	6	2	28	14
AC008079.9	0	0	0	0	0
GGTLC5P	0	0	0	0	0
AC011718.2	1	0	0	0	0
AC011718.3	0	0	0	0	0
PPP1R26P3	0	0	0	0	0
XXbac-B33L19.6	0	0	0	0	0
FAM230A	0	0	0	2	4
XXbac-B33L19.12	0	0	0	0	0
GGTLC3	0	0	0	0	0
Metazoa_SRP	0	0	1	0	2
XXbac-B33L19.15	0	0	0	0	0
TMEM191B	6	1	3	0	2
PI4KAP1	18	11	36	65	33
SCARNA17	0	0	0	0	0
RN7SKP131	0	0	0	0	0
RIMBP3	1	2	0	2	3
XXbac-B33L19.10	0	0	0	0	0
CA15P2	0	0	0	0	0
PPP1R26P2	0	0	0	0	0
PPP1R26P4	0	0	0	0	0

AC008132.15	0	0	0	0	1
GGT3P	0	0	0	0	0
AC008132.14	0	0	0	0	0
E2F6P1	0	0	0	3	0
AC008132.13	0	0	1	1	3
BCRP7	0	0	0	0	0
AC008103.3	0	0	0	0	0
AC008103.4	0	0	0	0	0
DGCR6	73	91	76	24	13
PRODH	102	127	123	44	34
AC007326.1	0	0	0	3	1
AC007326.10	0	1	0	0	3
DGCR5	196	239	201	8	9
AC007326.9	11	9	12	1	0
DGCR9	30	30	27	5	2
AC000095.9	0	0	0	0	0
AC000095.11	9	5	9	0	0
CA15P1	4	3	5	1	2
DGCR2	304	354	314	308	178
Y_RNA	4	2	2	5	2
DGCR11	4	10	13	19	6
AC004461.4	1	1	0	0	2
DGCR12	2	1	7	2	2
AC004471.9	5	4	5	4	3
AC004471.10	1	1	0	0	2
TSSK1A	0	0	0	0	2
DGCR14	46	50	37	77	39
TSSK2	0	0	0	0	0
GSC2	0	0	0	0	0
LINC01311	3	1	3	11	8
SLC25A1	32	50	41	226	138
CLTCL1	33	40	36	102	71
AC000081.2	0	0	1	0	1
SNORA15	0	0	0	0	1
KRT18P62	1	1	0	1	0
HIRA	20	15	10	59	27
C22orf39	64	74	76	100	88
RN7SL168P	0	1	0	4	2
MRPL40	29	36	31	113	70
AC000068.5	5	4	1	6	8
UFD1L	38	58	53	228	148
AC000068.9	0	0	0	0	0
AC000068.10	0	0	0	0	2
CDC45	2	1	0	224	154
CLDN5	60	86	57	2	2
LINC00895	0	0	0	0	0
AC000077.2	0	0	0	0	1

AC000067.1	0	0	0	0	0
SEPT5	481	635	547	145	97
GP1BB	0	0	0	0	0
TBX1	1	0	2	7	9
GNB1L	14	17	8	39	36
AC000089.3	1	0	0	2	1
C22orf29	83	123	95	211	149
TXNRD2	55	52	41	84	68
AC000078.5	2	9	2	8	4
COMT	96	113	87	205	111
MIR4761	0	0	1	1	0
ARVCF	143	157	135	86	35
TANGO2	36	47	38	68	48
MIR185	0	0	0	0	0
AC006547.13	7	17	15	13	9
AC006547.15	0	0	0	0	0
DGCR8	80	96	67	238	162
MIR3618	0	0	0	0	0
MIR1306	0	0	0	0	0
AC006547.8	0	0	0	0	0
TRMT2A	42	62	49	164	82
MIR6816	0	0	0	0	0
RANBP1	52	74	70	593	399
SNORA77	0	0	0	0	0
ZDHHC8	95	97	91	114	84
AC006547.14	6	5	4	3	3
XXbac-B444P24.8	0	1	3	0	0
LINC00896	0	3	0	7	4
RTN4R	74	93	84	19	11
MIR1286	1	3	0	0	0
XXbac-B444P24.10	1	1	0	1	4
DGCR6L	47	42	38	96	46
XXbac-B444P24.13	1	2	3	2	1
XXbac-B444P24.14	3	3	3	3	0
XXbac-B33L19.4	0	0	0	1	0
USP41	0	0	0	3	2
ZNF74	38	40	41	110	79
RNU6-225P	0	0	0	0	0
SCARF2	11	7	2	34	23
XXbac-B562F10.12	0	0	0	0	0
KLHL22	63	80	83	81	52
XXbac-B562F10.11	4	8	13	10	5
RNY1P9	0	0	0	0	0
RN7SL812P	0	0	3	1	0
KRT18P5	4	3	1	4	2
AC007731.1	0	0	0	0	0
MED15	94	110	87	258	156

AC007050.17	0	0	0	3	2
AC007050.18	0	0	1	0	0
SMPD4P1	2	4	2	1	3
IGLL4P	0	0	0	0	1
SLC9A3P2	0	0	0	0	0
AC007050.1	0	0	0	0	0
ABHD17AP4	0	1	3	0	2
POM121L4P	0	0	0	0	1
BCRP5	0	0	0	0	0
TMEM191A	5	7	5	0	4
PI4KA	1102	1211	1079	433	343
SERPIND1	13	11	8	158	109
SNAP29	33	48	44	114	82
AC007308.7	0	0	0	0	0
CRKL	343	423	367	758	482
XXbac-B135H6.15	5	3	4	1	6
AIFM3	85	82	100	7	13
LZTR1	140	179	126	297	188
XXbac-B135H6.18	4	5	1	1	2
THAP7	36	56	39	95	77
THAP7-AS1	4	7	6	9	2
TUBA3FP	2	1	2	6	6
P2RX6	35	33	31	23	19
SLC7A4	19	25	14	9	4
AC002472.11	0	0	1	0	0
MIR649	0	0	0	0	0
P2RX6P	5	2	2	0	1
LRRC74B	0	2	2	1	1
TUBA3GP	0	0	0	0	0
BCRP2	1	0	0	0	1
KB-1592A4.15	0	0	0	0	0
POM121L7	1	0	1	0	0
E2F6P2	0	1	0	0	0
FAM230B	0	0	0	0	1
KB-1592A4.14	0	0	0	0	0
GGT2	0	0	1	3	2
E2F6P3	0	0	0	0	0
POM121L8P	0	0	0	0	0
BCRP6	0	0	0	0	0
KB-1183D5.13	0	0	0	0	0
KB-1183D5.14	0	0	0	0	0
PPP1R26P5	0	0	0	0	1
KB-1183D5.18	0	0	0	0	0
KB-1183D5.16	0	0	0	0	0
RIMBP3B	0	1	0	0	2
RN7SKP63	0	0	0	0	0
HIC2	29	25	28	209	106

TMEM191C	1	1	6	1	5
PI4KAP2	28	26	20	96	60
RN7SKP221	0	0	0	0	0
RIMBP3C	0	0	0	0	1
UBE2L3	207	241	198	470	310
YDJC	18	28	24	155	83
CCDC116	2	0	1	5	3
KB-1440D3.14	0	0	0	2	1
SDF2L1	5	14	3	51	34
PPIL2	60	65	60	243	146
MIR301B	0	0	0	0	0
MIR130B	0	0	0	0	0
KB-1440D3.13	0	0	0	4	5
YPEL1	25	45	31	17	9
RN7SL280P	1	0	1	1	1
MAPK1	741	955	901	840	502
RNA5SP493	0	0	1	3	4
KB-1027C11.4	0	1	1	1	0
PPM1F	80	106	108	217	140
LL22NC03-86G7.1	26	50	28	69	74
TOP3B	2	0	1	12	10
PRAMENP	0	0	0	0	0
IGLVI-70	0	0	0	0	1
IGLV4-69	0	0	0	0	0
IGLVI-68	0	0	0	0	0
LL22NC03-23C6.13	0	0	0	0	1
LL22NC03-23C6.12	0	0	0	0	1
IGLV10-67	0	0	0	0	0
LL22NC03-23C6.15	0	0	0	0	0
IGLVIV-66-1	0	0	0	0	0
IGLVV-66	0	0	0	0	0
IGLVIV-65	0	0	0	0	0
IGLVIV-64	0	0	0	0	0
IGLVI-63	0	0	0	0	1
IGLV1-62	0	0	0	0	2
LL22NC03-88E1.18	0	0	0	0	0
IGLV8-61	0	0	0	0	0
LL22NC03-88E1.17	0	0	0	0	0
ABHD17AP5	0	0	0	0	0
LL22NC03-30E12.10	0	0	0	0	0
LL22NC03-30E12.11	0	0	0	0	0
IGLV4-60	0	0	0	0	0
LL22NC03-30E12.13	0	0	0	0	0
SOCS2P2	0	0	0	0	1
IGLVIV-59	0	0	0	0	0
IGLVV-58	0	0	0	0	0
BMP6P1	0	0	0	0	0

IGLV6-57	0	0	0	1	0
IGLVI-56	0	0	0	0	0
IGLV11-55	0	0	0	0	0
IGLV10-54	0	0	0	0	0
IGLVIV-53	0	0	0	0	1
TOP3BP1	0	0	0	1	0
LL22NC03-123E1.5	0	0	0	0	1
VPREB1	0	0	0	1	0
CH17-264L24.1	12	12	18	16	15
LL22NC03-2H8.5	2	0	1	2	6
LL22NC03-2H8.4	0	0	0	0	0
LL22NC03-80A10.6	9	11	3	10	3
BMS1P20	3	1	7	6	1
IGLV5-52	1	2	3	0	0
IGLV1-51	1	4	1	0	1
IGLV1-50	1	3	0	0	0
LL22NC03-80A10.11	0	0	0	0	1
IGLV9-49	0	0	0	0	1
LL22NC03-75A1.9	0	0	0	0	0
IGLV5-48	0	0	0	0	0
IGLV1-47	0	0	0	0	0
LL22NC03-22A12.9	0	0	0	0	1
IGLV7-46	0	0	0	0	0
LL22NC03-22A12.12	0	0	0	0	0
IGLV5-45	0	0	0	0	0
IGLV1-44	0	0	0	0	0
IGLV7-43	0	0	0	0	0
IGLVI-42	0	0	0	0	0
IGLVVII-41-1	0	0	0	0	0
IGLV1-41	0	0	0	0	0
IGLV1-40	0	0	0	0	1
ASH2LP1	0	0	0	0	0
IGLVI-38	0	0	0	0	1
IGLV5-37	0	0	0	0	0
IGLV1-36	0	0	0	0	1
KB-288A10.17	0	1	0	0	1
IGLV7-35	0	0	0	0	0
KB-288A10.19	0	0	1	0	1
ZNF280B	28	37	39	65	55
ZNF280A	0	0	0	11	11
PRAME	0	0	0	822	472
LL22NC03-63E9.3	0	0	0	1	1
IGLV2-34	0	0	0	0	0
IGLV2-33	0	0	0	0	0
IGLV3-32	0	0	0	0	0
IGLV3-31	0	0	0	0	1
IGLV3-30	0	0	0	0	0

BCRP4	0	0	0	0	0
POM121L1P	0	0	0	0	1
GGTLC2	0	0	0	0	1
IGLV3-29	0	0	0	0	1
AC244250.2	0	0	0	0	0
IGLV2-28	0	0	0	0	0
IGLV3-27	0	0	0	2	0
IGLV3-26	0	0	0	0	0
IGLVVI-25-1	0	0	0	0	0
IGLV3-25	0	0	0	0	2
LL22NC03-102D1.16	0	0	0	0	0
LL22NC03-102D1.18	0	0	0	0	0
IGLV3-24	0	0	0	0	0
IGLV2-23	0	0	0	0	0
AC244250.1	0	0	0	0	0
IGLVVI-22-1	0	0	0	0	0
IGLV3-22	0	0	0	0	0
IGLV3-21	0	0	0	0	0
IGLVI-20	0	0	0	0	0
IGLV3-19	0	0	0	0	1
LL22NC03-48A11.14	0	0	0	0	0
AC244250.4	0	0	0	0	0
IGLV2-18	0	0	0	0	0
IGLL5	0	0	0	1	0
IGLV3-17	0	0	0	0	0
IGLV3-16	0	0	0	0	0
IGLV3-15	0	0	0	0	0
AC244250.3	0	0	0	0	0
IGLV2-14	0	0	0	26	12
IGLV3-13	0	0	0	0	0
IGLV3-12	0	0	0	0	0
LL22NC03-24A12.8	0	0	0	0	0
LL22NC03-24A12.9	0	0	0	0	0
AC244157.1	0	0	0	0	0
IGLV2-11	0	0	0	0	0
LL22NC03-84E4.13	0	0	0	0	0
IGLV3-10	0	0	0	0	0
LL22NC03-84E4.11	0	0	0	0	0
IGLV3-9	0	0	0	0	1
IGLV2-8	0	0	0	0	1
MIR650	0	0	0	0	0
LL22NC03-84E4.8	0	0	0	0	1
IGLV3-7	0	0	0	0	0
IGLV3-6	0	0	0	0	0
AC245028.1	0	0	0	0	0
IGLV2-5	0	0	0	0	0
IGLV3-4	0	0	0	0	0

IGLV4-3	0	0	0	0	0
IGLV3-2	0	0	0	0	0
IGLV3-1	0	0	0	0	0
MIR5571	0	0	0	0	0
IGLJ1	0	0	0	0	0
IGLC1	0	0	0	0	0
IGLJ2	0	0	0	0	0
IGLC2	0	0	0	706	503
IGLJ3	0	0	0	0	1
IGLC3	0	0	0	1170	317
IGLJ4	0	0	0	0	0
IGLC4	0	0	0	0	0
IGLJ5	0	0	0	0	1
IGLC5	0	0	0	0	0
IGLJ6	0	0	0	0	0
IGLC6	0	0	0	2	1
IGLJ7	0	0	0	0	0
IGLC7	0	0	0	8	5
AP000361.2	0	0	0	0	0
AP000362.1	0	0	0	1	0
RSPH14	4	8	5	1	0
GNAZ	223	247	214	32	17
AC000029.1	0	1	0	0	0
U7	0	0	0	0	0
RAB36	29	30	33	20	15
BCR	228	357	302	332	204
BCRP8	0	1	0	0	0
RN7SL263P	1	1	0	1	0
FBXW4P1	2	4	5	8	6
AP000343.1	0	2	0	0	0
AP000343.2	0	0	1	0	0
CES5AP1	2	2	0	0	1
ZDHHC8P1	21	16	14	1	2
KB-1269D1.8	1	1	0	0	0
AP000344.3	0	0	0	0	1
AP000344.4	0	1	0	1	3
AP000345.1	0	0	1	0	1
AP000345.4	0	0	0	4	4
PCAT14	0	0	1	202	156
AP000345.2	0	0	0	0	2
IGLL1	0	0	0	0	0
KB-208E9.1	1	3	1	0	1
KB-1572G7.5	0	1	0	1	1
DRICH1	2	2	1	0	4
KB-1572G7.4	0	0	1	0	0
KB-1572G7.3	0	3	0	0	0
GUSBP11	1	0	1	1	0

AP000347.4	0	0	0	1	5
ASLP1	1	1	2	5	3
KB-1572G7.2	0	0	1	0	0
AP000347.2	4	4	8	12	8
RGL4	16	27	18	17	15
ZNF70	26	46	32	80	48
VPREB3	0	0	0	0	0
C22orf15	2	7	7	4	3
CHCHD10	44	53	39	46	36
MMP11	3	1	1	15	17
AP000349.2	0	0	0	0	0
SMARCB1	70	78	62	169	131
DERL3	4	8	7	81	43
KB-1125A3.11	0	3	0	1	0
SLC2A11	54	63	46	28	34
AP000350.10	0	0	0	0	0
KB-1125A3.12	4	4	1	8	5
RN7SL268P	1	1	1	4	1
MIF	3	3	3	3	1
MIF-AS1	5	1	5	8	11
AP000350.5	11	14	12	30	24
AP000350.1	0	0	0	0	0
AP000350.6	2	2	1	3	1
AP000350.7	0	0	0	0	0
AP000350.8	0	0	0	0	0
GSTT2B	3	4	3	10	4
KB-1125A3.10	0	0	0	0	0
DDTL	3	3	3	2	17
KB-226F1.2	1	2	1	1	4
DDT	4	5	4	18	8
AP000351.3	0	0	0	0	0
GSTT2	0	0	0	0	0
AP000351.4	0	0	0	1	0
GSTTP1	0	0	0	0	1
AP000351.13	0	0	0	0	1
CABIN1	280	320	309	509	296
KB-318B8.7	13	11	7	14	10
SUSD2	8	8	7	11	17
GGT5	22	31	23	10	4
AP000354.4	0	0	1	0	0
POM121L9P	1	1	2	0	1
BCRP1	0	0	0	0	0
AP000354.2	0	0	0	0	1
SPECC1L	64	80	70	156	82
SPECC1L-ADORA2A	0	1	0	0	1
ADORA2A	2	4	2	4	0
ADORA2A-AS1	2	5	5	6	12

UPB1	0	1	2	1	10
AP000355.2	0	0	0	0	1
GUCD1	62	62	62	384	214
SNRPD3	71	87	98	293	211
GGT1	9	14	13	86	42
LRRC75B	20	30	26	21	24
AP000356.2	0	0	0	0	0
BCRP3	2	1	2	5	5
POM121L10P	0	0	0	0	0
ARL5AP4	0	0	0	0	0
AP000357.4	0	0	0	0	1
LL22NC03-N95F10.1	0	0	0	0	0
AP000358.5	0	0	0	0	0
CRIP1P4	0	0	0	0	1
PIWIL3	0	0	0	1	3
SGSM1	151	187	161	5	8
SNORD56	0	0	1	0	0
TMEM211	0	0	1	0	1
KIAA1671	126	148	131	158	97
CTA-243E7.4	3	3	0	1	3
CTA-243E7.3	0	1	1	1	1
CTA-243E7.2	0	2	1	4	4
CTA-243E7.1	5	6	3	12	19
CTA-221G9.12	7	7	6	4	1
CTA-221G9.10	0	0	0	0	0
CRYBB3	0	1	1	1	0
CTA-221G9.7	0	1	0	0	0
CRYBB2	0	1	0	0	4
Z99916.1	0	0	0	0	0
CTA-221G9.11	0	0	0	0	0
RP3-462D8.2	0	0	0	0	1
IGLL3P	0	0	0	0	0
CTA-246H3.8	0	1	0	2	2
CTA-246H3.11	0	2	2	4	2
LRP5L	7	5	5	41	24
CTA-246H3.12	0	0	0	1	2
CTA-390C10.9	0	1	0	0	0
IGLVIVOR22-1	0	0	0	0	0
CRYBB2P1	24	27	23	54	37
MIR6817	0	0	0	0	0
CTA-390C10.10	10	18	11	12	8
AL008721.1	0	0	0	0	0
CTA-407F11.8	1	1	0	1	2
ADRBK2	206	283	260	170	134
CTA-407F11.7	0	0	1	2	3
YES1P1	3	5	4	9	4
RNA5SP494	0	0	0	0	0

MYO18B	0	0	0	86	85
CTA-407F11.9	0	0	0	2	1
CTA-125H2.2	0	0	0	1	6
Z98949.1	0	0	0	1	1
CTA-125H2.1	0	0	0	0	0
RN7SKP169	0	0	0	0	1
CTA-125H2.3	0	0	0	0	0
CTA-796E4.3	0	0	0	0	1
CTA-796E4.4	1	0	0	0	1
SEZ6L	399	508	463	16	17
RP11-259P1.1	3	5	2	0	0
RNA5SP495	0	0	0	0	0
CTB-1048E9.7	0	0	0	0	1
ASPHD2	44	56	80	18	13
HPS4	114	176	183	318	202
SRRD	35	37	43	62	47
TFIP11	60	93	52	141	77
CTA-445C9.14	14	23	13	27	13
TPST2	38	39	38	85	46
MIR548J	0	0	0	2	0
HMGB1P10	2	3	2	9	10
CRYBB1	2	0	1	3	3
CRYBA4	1	0	0	0	1
ISCA2P1	0	0	0	0	1
MIAT	429	453	443	43	34
MIAT_exon1	0	0	0	0	0
CTA-373H7.7	0	4	0	0	0
MIAT_exon5_1	0	0	0	0	0
MIAT_exon5_2	0	0	0	0	0
MIATNB	4	6	8	11	14
MIAT_exon5_3	0	0	0	0	0
LINC01422	2	4	6	3	9
RP1-40G4P.1	0	0	0	0	0
CTA-992D9.11	0	0	0	0	8
RNU6-1066P	0	0	0	0	0
CTA-992D9.10	0	0	0	0	0
CTA-992D9.9	0	0	0	0	0
CTA-992D9.6	1	0	0	0	2
CTA-992D9.8	0	0	0	1	4
CTA-992D9.7	0	0	0	0	1
CTA-503F6.2	0	0	0	0	1
RP5-1172A22.1	1	0	0	2	1
CTA-929C8.5	0	0	0	0	0
CTA-929C8.7	0	0	0	0	0
CTA-929C8.6	0	0	0	0	3
CTA-929C8.8	1	3	1	0	0
RP1-205F14P.1	0	0	0	1	0

RP11-46E17.6	0	0	0	0	0
RP1-231P7P.1	0	0	0	0	2
RP1-213J1P__B.1	0	0	0	2	1
RP1-213J1P__B.2	0	0	0	0	0
RP11-375H17.1	0	0	0	0	1
MN1	44	34	31	123	93
PITPNB	80	92	64	111	87
TTC28-AS1	42	51	66	48	46
TTC28-AS1_1	0	0	0	0	0
MIR3199-2	1	0	0	1	0
TTC28-AS1_2	0	0	0	0	0
TTC28	87	123	79	191	108
TTC28-AS1_3	0	0	0	0	0
TTC28-AS1_4	0	0	0	0	0
RN7SL757P	1	0	0	0	0
SNORD42	0	0	0	0	0
MIR5739	0	0	1	3	2
RN7SL162P	0	0	0	0	0
CHEK2	1	5	3	98	82
HSCB	7	9	11	35	20
CCDC117	68	75	58	259	176
XBP1	67	102	79	1082	636
Z93930.1	0	0	0	1	0
CTA-292E10.6	4	7	2	10	12
CTA-292E10.8	1	0	0	4	3
CTA-292E10.7	0	0	0	0	0
ZNRF3	97	103	102	99	89
ZNRF3-IT1	1	6	3	4	3
ZNRF3-AS1	6	3	3	8	2
C22orf31	0	2	0	0	1
KREMEN1	50	62	36	75	59
CTA-747E2.10	1	0	2	0	0
Z95116.1	0	0	0	0	0
RNU6-810P	1	0	0	0	0
RNU6-1219P	0	0	0	0	0
EMID1	11	19	29	12	6
RHBDD3	23	40	33	79	57
CTA-984G1.5	2	1	3	9	3
EWSR1	346	356	357	920	675
GAS2L1	73	77	66	32	23
RASL10A	9	27	17	0	2
AC002059.10	0	0	1	1	1
AP1B1	230	272	256	358	244
MIR3653	0	0	0	0	0
SNORD125	0	0	0	0	0
AC000041.8	0	0	0	0	0
RFPL4AP6	0	0	0	0	1

RFPL1S	116	107	115	6	2
AC000041.10	0	0	0	0	0
RFPL1	0	0	0	0	0
AC000035.3	0	0	0	0	0
NEFH	194	265	253	27	24
THOC5	73	123	65	218	137
CTA-256D12.11	0	0	1	1	2
NIPSNAP1	105	131	119	291	188
NF2	206	270	235	445	307
RPEP4	2	4	1	2	0
RP1-76B20.11	0	1	0	0	1
RP1-76B20.12	0	0	0	0	0
CABP7	18	21	25	2	1
ZMAT5	16	12	6	20	20
UQCR10	83	94	104	185	98
ASCC2	35	42	39	144	91
MTMR3	36	46	35	29	15
AC004819.1	1	0	1	0	1
RNU6-331P	0	0	0	0	0
AC003681.1	0	0	1	0	0
RP3-394A18.1	81	94	75	86	63
MIR6818	0	0	0	0	0
HORMAD2-AS1	2	0	0	1	1
CNN2P1	0	1	0	0	1
HORMAD2	0	0	0	1	5
CTA-85E5.7	0	0	0	0	0
CTA-85E5.6	0	0	0	0	0
RP3-438O4.4	0	0	0	0	2
RP1-102K2.6	0	0	0	2	0
LIF	2	3	3	154	97
RP1-102K2.8	0	0	0	3	1
OSM	0	0	1	0	0
RP1-102K2.9	2	3	0	0	1
GATSL3	5	6	3	2	0
RP1-130H16.18	0	0	0	0	0
TBC1D10A	13	15	14	39	30
SF3A1	184	232	179	388	259
CCDC157	3	15	9	9	8
RNF215	54	65	42	57	40
SEC14L2	53	81	68	42	33
RP4-539M6.19	0	0	0	0	0
AC004832.1	0	0	1	0	0
RP4-539M6.20	3	6	4	4	6
RNU6-564P	1	1	0	0	0
RP4-539M6.21	0	2	1	4	4
MTFP1	0	0	0	3	5
RP4-539M6.22	0	0	0	1	1

RP4-539M6.18	0	0	0	1	0
SEC14L3	1	0	0	0	2
RP4-539M6.14	0	0	0	0	0
SDC4P	0	0	0	0	0
SEC14L4	0	0	0	3	2
SEC14L6	4	12	3	3	7
SIRPAP1	0	0	0	0	0
GAL3ST1	19	18	21	4	6
PES1	65	80	70	271	202
RP1-56J10.8	0	0	0	0	1
TCN2	19	26	23	56	49
SLC35E4	18	32	26	21	12
DUSP18	8	13	9	12	10
CTA-963H5.5	1	2	0	3	0
OSBP2	119	137	98	113	57
AL022336.1	2	0	0	0	0
EIF4HP2	1	4	9	3	6
MORC2-AS1	1	2	3	5	2
MORC2	75	82	50	146	100
TUG1	188	227	207	545	349
TUG1_1	0	0	0	0	0
TUG1_2	0	0	0	0	0
TUG1_3	0	0	0	0	0
TUG1_4	0	0	0	0	0
RP3-430N8.11	0	0	0	0	0
RP3-430N8.10	0	0	0	0	0
RP3-412A9.17	0	0	0	2	5
RN7SL633P	0	0	0	0	0
SMTN	42	51	47	142	91
RP3-412A9.16	1	2	2	11	6
SELM	83	86	77	25	18
RP3-412A9.12	0	0	0	0	0
INPP5J	59	77	58	7	8
PLA2G3	1	5	2	13	12
RP3-412A9.15	0	0	0	0	0
MIR3928	0	0	0	0	0
RNF185	57	105	74	106	82
RNF185-AS1	0	0	0	0	0
LIMK2	96	119	103	113	70
RNU6-1128P	0	0	0	0	1
PIK3IP1	112	165	128	52	32
PIK3IP1-AS1	1	1	3	2	0
RNA5SP496	0	0	0	0	0
PATZ1	72	74	61	191	116
LINC01521	12	14	13	8	9
RNU6-338P	0	0	0	0	0
DRG1	83	105	96	191	143

EIF4ENIF1	95	107	96	94	77
RP11-247I13.11	0	0	2	1	1
RNU6-28P	1	0	0	0	1
SFI1	112	125	107	142	102
H2AFZP6	1	0	1	0	0
RP11-247I13.6	0	0	0	0	0
RP11-247I13.3	0	1	1	4	1
PISD	195	250	233	156	95
MIR7109	0	0	0	0	0
PRR14L	219	269	234	418	294
DEPDC5	51	79	76	77	50
RN7SL20P	0	0	0	1	2
CTA-440B3.1	0	0	1	1	2
RNU6-201P	0	0	0	0	0
RP1-180M12.1	0	1	2	0	0
C22orf24	0	6	6	4	2
YWHAH	1029	1363	1146	336	220
CTA-342B11.1	0	0	0	1	0
CTA-342B11.2	0	0	0	0	2
RN7SL305P	0	0	0	0	0
SC22CB-1E7.1	0	0	0	0	0
SLC5A1	0	0	0	0	6
AP1B1P1	0	0	0	0	1
RP1-127L4.10	0	0	0	0	0
AP1B1P2	0	0	0	0	0
RP1-127L4.7	0	0	0	1	0
C22orf42	0	2	0	0	1
RP1-90G24.8	0	0	0	0	0
RFPL3-AS1_1	0	0	0	0	0
RFPL2	5	3	10	1	1
AL008723.1	0	0	0	1	0
IGLCOR22-1	0	0	0	0	0
RP1-90G24.10	4	7	3	2	3
SLC5A4	7	12	5	13	9
CPSF1P1	41	43	42	6	12
RP1-90G24.6	1	1	3	0	0
RP1-90G24.11	0	3	0	1	0
RP1-149A16.12	0	1	1	1	0
RFPL3	0	0	0	0	0
IGLCOR22-2	0	0	0	0	0
RFPL3S	1	1	1	0	0
RFPL3-AS1_2	0	0	0	0	0
RP1-149A16.3	1	3	2	1	1
IGLVIVOR22-2	0	0	0	0	0
RP1-149A16.17	0	1	0	4	2
RP1-149A16.16	2	0	0	2	0
RTCB	109	150	124	339	182

BPIFC	0	0	0	0	2
FBXO7	163	181	194	426	375
SYN3	74	110	78	11	12
LL22NC03-104C7.1	1	4	5	1	1
RNA5SP497	0	0	0	0	0
LL22NC01-116C6.1	0	0	1	0	1
TIMP3	214	271	243	805	491
CTA-415G2.2	0	0	0	0	0
RP1-302D9.1	0	0	0	0	0
RP1-302D9.2	0	0	0	0	0
RP1-302D9.5	0	0	0	0	1
LARGE	173	213	199	83	58
RP1-302D9.3	0	0	0	0	0
SC22CB-1D7.1	1	1	2	3	1
MIR4764	1	0	0	0	0
SNORA76	0	0	0	0	0
SNORA50	0	0	0	0	0
LARGE-IT1	0	0	0	0	0
LARGE-AS1	7	6	5	5	7
LL22NC03-32F9.1	0	0	0	1	0
LL22NC03-86D4.1	0	0	0	0	1
LL22NC03-13G6.2	0	1	2	0	0
RP1-101G11.2	0	0	0	0	0
RP1-101G11.3	0	0	0	3	4
RP1-288L1.5	2	3	0	0	1
RP1-288L1.4	0	0	1	0	0
ISX-AS1	1	0	0	2	1
ISX	0	0	0	6	4
LINC01399	0	0	0	1	0
CTA-714B7.7	0	0	0	0	0
RP3-323A16.1	0	0	1	75	77
CTA-714B7.4	0	0	0	0	0
COX7BP1	0	0	0	0	0
RNU7-167P	0	0	0	0	0
HMGXB4	32	45	49	124	103
RP3-510H16.3	2	0	1	0	0
TOM1	99	119	103	84	56
MIR3909	1	0	0	1	1
MIR6069	0	0	0	0	0
CTA-286B10.7	0	0	0	1	0
CTA-286B10.8	0	0	0	0	0
HMOX1	13	19	13	83	51
MCM5	29	32	34	544	294
RP4-569D19.5	0	0	0	4	1
RP4-569D19.8	0	0	0	0	0
RASD2	55	81	60	15	15
MB	0	1	0	9	8

CITF22-62D4.1	0	0	0	0	1
APOL6	64	54	56	331	197
RP1-41P2.7	0	0	0	0	0
MRPS16P3	0	0	0	0	0
APOL5	0	0	1	0	0
RBFOX2	621	738	711	515	367
NDUFA9P1	3	3	2	4	4
RP1-78B3.1	10	8	4	21	23
CTA-212A2.4	0	0	0	0	0
CTA-212A2.3	0	0	0	0	9
CTA-212A2.2	0	0	0	0	0
CTA-212A2.1	1	0	1	0	1
APOL3	8	6	12	39	33
Z95114.4	0	0	0	0	1
Z95114.3	0	0	0	0	0
Z95114.5	0	0	0	0	1
Z95114.6	0	0	0	0	1
Z95114.7	0	0	0	0	0
MTND1P10	0	0	0	0	1
APOL4	10	6	8	9	5
APOL2	113	151	135	148	112
APOL1	4	9	6	87	65
MYH9	449	512	485	2685	1608
MIR6819	0	0	0	1	1
RP4-633O19__A.1	1	2	1	2	0
RPS15AP38	0	0	0	0	1
BX470187.1	0	0	0	0	0
RP5-1119A7.14	0	1	3	0	0
TXN2	27	41	39	82	49
FOXRED2	74	108	107	218	104
EIF3D	103	119	131	745	494
RP5-1119A7.11	0	0	0	0	0
RP5-1119A7.10	0	0	0	0	0
RP5-1119A7.17	41	56	44	0	0
CACNG2	33	34	48	0	1
RP1-293L6.1	3	2	4	0	0
IFT27	53	58	56	77	57
PVALB	100	129	102	25	17
CITF22-24E5.1	2	0	0	0	0
CTA-833B7.2	0	0	1	1	4
NCF4	3	7	1	24	17
CSF2RB	2	4	4	50	42
CSF2RBP1	0	0	0	0	1
LL22NC01-81G9.3	0	0	0	0	1
TEX33	0	0	0	0	0
TST	18	21	18	88	61
MPST	34	29	33	131	76

KCTD17	47	74	73	32	21
RN7SKP214	0	0	0	1	0
TMPRSS6	1	3	0	5	5
RP5-1170K4.7	0	0	0	0	1
IL2RB	1	0	0	6	6
AL022314.1	0	0	0	0	1
RP1-151B14.6	0	0	0	1	0
C1QTNF6	11	8	10	85	65
SSTR3	30	36	48	14	12
RP1-151B14.9	0	0	0	0	0
RAC2	8	8	5	111	79
CYTH4	13	18	16	22	19
ELFN2	106	137	114	19	17
RP1-63G5.7	1	2	1	1	3
RP1-63G5.8	0	0	0	0	0
MFNG	4	5	6	75	30
CARD10	2	5	5	97	76
RP5-1177I5.3	1	1	0	0	0
CDC42EP1	68	100	65	144	81
LGALS2	0	0	0	7	5
GGA1	119	155	128	307	193
SH3BP1	20	13	16	55	34
Z83844.1	1	3	1	7	2
PDXP	5	6	5	2	1
RN7SL385P	0	1	0	0	0
LGALS1	75	93	78	729	447
NOL12	9	10	15	28	17
RP1-37E16.12	0	0	0	0	0
TRIOBP	36	28	25	164	103
H1F0	106	146	112	252	167
GCAT	34	27	33	76	56
GALR3	1	0	2	1	1
ANKRD54	33	48	55	91	50
MIR658	0	0	0	0	0
MIR659	0	0	0	1	0
EIF3L	191	222	206	704	454
RP5-1014D13.2	12	23	15	48	32
RNU6-900P	0	0	0	0	0
MICALL1	52	64	62	149	96
RP5-1039K5.12	1	8	1	4	2
C22orf23	2	1	5	9	6
AL031587.1	0	2	0	0	0
RP5-1039K5.19	167	192	150	105	83
RP5-1039K5.17	0	0	0	0	0
POLR2F	15	28	17	12	8
MIR6820	0	0	0	0	0
SOX10	0	0	0	0	0

MIR4534	0	0	0	0	0
RP5-1039K5.16	6	3	10	4	7
RP5-1039K5.18	0	0	1	0	1
PICK1	79	81	85	46	33
RP5-1039K5.13	0	0	0	0	0
SLC16A8	9	13	11	11	5
BAIAP2L2	5	17	8	18	8
CTA-228A9.3	10	3	3	4	2
PLA2G6	61	57	62	86	48
CTA-228A9.4	2	3	4	13	7
MAFF	31	27	20	97	63
TMEM184B	207	235	215	191	140
RP1-5O6.5	0	0	0	1	0
RN7SL704P	0	0	0	0	1
AL020993.1	0	0	0	0	0
CSNK1E	232	229	222	592	338
RP3-449O17.1	10	15	10	30	28
RP1-5O6.4	2	0	1	6	4
RP3-434P1.6	0	0	0	0	0
KCNJ4	49	70	50	2	1
KDELR3	6	7	7	106	92
DDX17	1532	1797	1637	2814	1818
DMC1	2	4	0	15	11
RP1-199H16.6	1	0	0	0	0
FAM227A	16	12	17	40	48
CBY1	17	19	13	25	22
RP3-508I15.10	3	1	0	0	4
RP3-508I15.9	12	7	15	7	9
TOMM22	52	67	51	178	112
JOSD1	149	161	133	261	204
GTPBP1	146	187	152	335	182
PRDX3P1	3	5	2	11	3
RP3-508I15.18	0	0	0	0	2
SUN2	289	369	314	266	151
RP3-508I15.19	0	2	0	4	1
RP3-508I15.20	1	1	1	2	3
RP3-508I15.14	1	0	2	3	0
RP3-508I15.21	1	5	2	4	9
RP3-508I15.22	0	1	1	4	0
DNAL4	19	42	34	46	18
NPTXR	390	446	394	69	59
CBX6	738	873	830	362	240
CTA-150C2.13	0	0	0	0	0
APOBEC3A	0	0	0	0	1
APOBEC3B	1	0	0	19	22
APOBEC3B-AS1	0	0	0	0	0
APOBEC3C	4	3	3	80	48

APOBEC3D	1	1	2	16	8
APOBEC3F	1	1	0	11	6
APOBEC3G	4	9	3	41	25
CTA-150C2.20	0	0	0	0	0
RP4-742C19.12	0	2	1	0	0
APOBEC3H	4	3	0	0	0
CBX7	185	257	208	25	21
COX5BP7	0	0	0	0	0
RP4-742C19.13	6	19	10	8	6
FUNDC2P4	0	0	0	0	0
PDGFB	34	45	32	40	28
RP3-333H23.8	0	0	0	0	0
RPL3	659	761	721	4027	2406
SNORD83B	0	0	0	0	0
SNORD83A	2	0	0	7	5
AL022326.1	0	0	0	1	1
SNORD43	0	0	0	0	0
SYNGR1	763	997	870	54	51
RP3-333H23.9	1	2	2	0	0
TAB1	99	91	70	82	75
MGAT3	257	328	274	35	15
MGAT3-AS1	5	9	8	1	2
MIEF1	95	122	114	312	164
ATF4	291	358	307	1354	819
RPS19BP1	72	88	72	129	95
CACNA1I	106	139	136	2	7
ENTHD1	0	0	0	2	3
RP1-172B20.6	0	0	0	0	1
RN7SKP210	0	0	0	0	0
UQCRFS1P1	1	5	6	4	8
GRAP2	1	0	1	64	45
RP3-370M22.8	0	0	0	8	1
FAM83F	0	1	6	28	20
RP3-496C20.1	0	2	0	1	0
TNRC6B	582	686	592	528	327
RPL7P52	2	4	2	3	2
ADSL	40	41	48	302	176
RP5-1042K10.14	0	0	1	1	0
SGSM3	109	136	114	147	104
RP5-1042K10.10	1	2	1	4	4
MKL1	139	132	104	137	81
RP5-1042K10.13	0	0	1	1	1
RP5-1042K10.12	2	0	2	0	4
RP4-591N18.2	4	10	8	11	7
AL031594.1	1	1	1	2	2
COX6B1P3	1	0	0	1	0
RPL4P6	1	2	1	8	5

GAPDHP37	0	1	0	0	1
MCHR1	12	21	17	2	3
SLC25A17	39	39	40	119	96
RP3-408N23.4	1	1	3	5	1
JTBP1	1	1	0	6	1
MIR4766	0	0	0	2	0
ST13	389	470	463	880	640
XPNPEP3	31	43	39	149	105
DNAJB7	2	2	0	2	1
RNU6-379P	0	0	0	0	0
RBX1	50	68	60	122	77
AL080243.1	0	0	0	0	0
RP11-12M9.3	0	0	0	0	0
RP11-12M9.4	0	0	0	0	0
EP300	302	403	405	569	328
MIR1281	0	0	0	0	0
RNU6-375P	0	0	0	0	1
RP1-85F18.6	0	3	0	3	3
EP300-AS1	1	2	0	1	1
LRRC37A14P	0	1	0	0	0
L3MBTL2	91	96	83	108	82
RP4-756G23.5	6	8	9	2	5
CHADL	32	40	39	2	3
RANGAP1	256	316	267	464	296
MIR6889	0	0	0	0	0
ZC3H7B	417	531	475	453	288
TEF	335	384	341	43	34
RNU6-495P	0	2	0	0	0
CTA-223H9.9	0	0	2	4	1
TOB2	203	279	197	173	116
PHF5A	34	36	17	134	79
ACO2	301	366	347	254	168
POLR3H	106	93	91	131	77
CSDC2	72	94	65	6	5
PMM1	34	40	55	51	31
DESI1	129	188	159	190	134
XRCC6	365	415	378	1431	973
NHP2L1	168	201	178	366	219
RNU6-476P	0	1	0	3	1
C22orf46	34	36	32	56	46
MEI1	1	5	1	8	6
HMGN2P10	0	1	0	0	0
RNU6ATAC22P	0	0	0	0	0
CCDC134	9	24	19	39	24
RP5-821D11.7	12	14	13	34	18
SREBF2	497	669	575	627	384
MIR33A	0	1	2	0	0

SHISA8	33	31	18	0	0
TNFRSF13C	4	2	2	11	9
MIR378I	0	0	0	0	0
CENPM	0	3	1	65	35
LINC00634	28	16	21	0	1
SEPT3	721	907	863	54	38
CTA-250D10.19	4	3	4	0	1
WBP2NL	14	11	8	5	9
SLC25A5P1	0	0	1	0	1
NAGA	30	43	53	144	107
FAM109B	9	3	4	6	5
SNORD13P1	0	0	0	0	1
SMDT1	25	42	34	28	15
NDUFA6	67	101	81	167	109
RP1-257I20.14	2	1	2	8	4
NDUFA6-AS1	10	18	21	15	6
OLA1P1	1	0	3	3	0
CYP2D6	15	23	16	17	9
RP4-669P10.19	2	1	0	4	1
RP4-669P10.16	1	3	0	2	0
RP4-669P10.20	4	6	5	2	2
CYP2D7	6	9	8	2	3
CYP2D8P	2	3	4	6	7
TCF20	205	259	209	455	281
OGFRP1	0	3	1	13	6
Z83851.4	1	2	1	8	5
CTA-989H11.1	1	3	1	8	1
LINC01315	1	3	5	2	4
NFAM1	5	5	11	8	9
CTA-126B4.7	0	0	0	1	1
SERHL	4	3	5	2	2
RRP7A	59	89	80	157	86
Z93241.1	0	0	0	0	0
SERHL2	6	11	7	18	14
RRP7BP	20	30	18	52	42
RN7SKP80	284	319	228	250	413
RNU6-513P	0	0	0	0	0
POLDIP3	139	155	150	263	169
RNU12	6	10	7	17	4
CYB5R3	126	169	136	372	250
ATP5L2	0	3	1	2	0
A4GALT	8	5	2	36	23
RPL5P34	0	0	0	0	0
GOLGA2P4	0	1	0	0	1
RP1-47A17.1	0	0	1	1	2
ARFGAP3	53	65	44	187	145
PACSIN2	114	127	111	261	174

AL049758.2	0	1	0	1	0
AL022476.2	5	3	1	0	4
TTLL1	16	29	26	10	12
AL022237.3	0	1	0	0	0
BIK	1	3	6	4	5
MCAT	19	27	25	43	27
TSPO	12	10	13	136	88
TTLL12	57	94	59	278	212
SCUBE1	1	5	3	11	15
Z82214.3	0	0	0	0	0
Z82214.2	0	0	0	0	0
Z99756.1	0	0	0	1	0
RP4-754E20__A.5	1	0	0	1	3
MPPED1	122	154	138	1	3
EFCAB6-AS1	2	3	0	0	2
EFCAB6	36	43	35	6	4
RP3-388M5.8	1	0	1	0	0
HMGN2P9	0	0	0	0	0
RP3-388M5.9	3	3	6	2	4
SULT4A1	290	334	288	5	7
PNPLA5	3	2	2	0	3
Z97055.1	0	0	0	0	0
PNPLA3	6	5	5	45	37
SAMM50	45	46	50	132	118
RP4-796I17.5	0	1	0	0	0
PARVB	69	89	68	162	134
CTA-414D7.1	3	6	4	10	10
RP4-671O14.6	3	7	5	18	15
PARVG	7	10	8	18	19
RP4-671O14.5	0	0	1	3	1
RP4-671O14.7	0	3	0	3	7
KIAA1644	149	205	176	162	111
RP1-32I10.10	0	0	0	0	0
RP1-32I10.11	0	0	0	0	0
CTA-397C4.2	0	0	0	0	0
MRPS18CP6	0	0	0	0	0
RP5-1033E15.3	0	0	0	0	0
LDOC1L	174	261	176	237	149
KRT18P23	0	0	0	0	2
LINC00207	0	0	0	0	4
LINC00229	0	0	0	0	0
ANP32BP2	0	0	0	0	0
PRR5	10	5	6	26	23
PRR5-ARHGAP8	0	1	0	1	1
ARHGAP8	2	1	2	6	6
PHF21B	15	14	9	16	16
RP1-127B20.4	1	1	0	1	1

RP4-753M9.1	1	5	0	5	4
NUP50-AS1	27	24	31	32	29
CTA-217C2.2	1	3	0	6	2
NUP50	131	172	154	413	289
CTA-268H5.12	0	0	0	0	1
KIAA0930	312	364	315	260	166
MIR1249	1	2	4	0	1
CTA-268H5.14	1	3	1	0	0
CTA-268H5.9	0	0	0	0	0
UPK3A	0	0	0	2	1
FAM118A	53	50	50	113	66
SMC1B	3	0	4	6	11
RIBC2	0	4	1	12	11
RP1-102D24.5	0	0	0	0	1
FBLN1	43	47	33	122	96
LINC01589	0	0	0	3	1
RNU6-1161P	0	0	0	1	0
CTA-941F9.10	16	22	15	31	17
ATXN10	193	253	246	199	154
MIR4762	1	0	0	0	1
RP1-37M3.8	0	3	2	5	1
WI2-85898F10.2	0	4	3	0	1
WI2-85898F10.1	1	0	0	0	0
WNT7B	7	11	6	9	9
CR536603.1	0	0	0	0	0
CITF22-92A6.2	3	2	2	4	5
CITF22-92A6.1	0	0	1	1	6
LINC00899	7	14	9	7	5
PRR34	7	12	15	7	2
RP6-109B7.5	4	8	3	0	0
PRR34-AS1	6	13	5	15	5
MIRLET7BHG	82	137	87	61	56
RP6-109B7.2	10	14	11	2	1
RP6-109B7.4	17	15	12	6	6
MIR3619	2	2	1	5	1
MIRLET7A3	0	0	0	0	0
MIR4763	0	0	0	0	0
MIRLET7B	0	0	0	0	0
PPARA	72	101	84	193	147
RP4-695O20.1	0	1	1	3	5
CDPF1	15	13	11	28	11
PKDREJ	6	6	7	8	12
TTC38	29	28	20	90	71
GTSE1-AS1	2	2	3	2	2
GTSE1	2	2	1	145	111
TRMU	64	78	75	191	111
CELSR1	0	6	4	250	136

RP5-996D20.3	0	0	0	0	0
RP3-439F8.1	0	0	0	0	9
GRAMD4	97	106	75	156	96
CERK	209	249	268	298	188
CTA-29F11.1	2	6	12	14	13
TBC1D22A	39	52	47	78	50
U51561.1	1	1	2	1	4
FP325331.1	2	4	3	2	1
RP1-111J24.1	1	1	0	1	3
CITF22-49D8.1	0	1	0	0	2
LL22NC03-75H12.2	0	0	0	0	1
LINC00898	7	6	1	5	3
RP11-191L9.4	8	8	8	32	28
RP11-191L9.5	1	0	2	0	2
RP13-455A7.1	1	0	0	2	3
CTA-280A3.2	4	3	3	7	12
CTA-280A3__B.2	0	0	0	0	0
LL22NC03-27C5.1	0	0	0	0	2
LL22NC03-121E8.4	0	0	0	0	1
LL22NC03-121E8.3	0	0	0	0	0
MIR3201	0	0	0	0	0
RP11-536P6.3	0	0	0	0	0
FAM19A5	140	163	151	18	16
CTA-299D3.8	17	16	18	0	1
MIR4535	0	0	0	0	1
AL954742.1	0	0	0	0	0
LINC01310	8	7	6	0	9
RPL35P8	0	0	0	0	1
RP1-34P24.3	0	0	0	0	3
C22orf34	2	4	2	36	25
CTA-722E9.1	0	0	0	0	3
RP4-566L20.1	0	0	0	0	1
MIR3667	0	0	0	0	0
RP1-29C18.10	0	0	1	13	19
RP1-29C18.9	0	0	0	7	1
RP1-29C18.8	1	0	1	17	14
RP5-983L19.2	0	0	0	0	4
RN7SKP252	0	0	0	0	0
RPL5P35	0	0	0	0	1
BRD1	92	116	103	186	121
RP3-522J7.7	3	2	8	5	8
RP3-522J7.5	2	3	2	3	3
RP3-522J7.6	1	2	1	1	2
RP11-494O16.3	0	1	0	0	0
RP11-494O16.4	0	0	0	2	0
ZBED4	51	68	77	196	153
ALG12	21	30	40	53	27

CITF22-1A6.3	0	0	0	1	0
CRELD2	30	40	43	74	64
CITF22-49E9.3	0	0	0	1	4
PIM3	47	59	55	159	103
MIR6821	0	0	0	0	0
IL17REL	0	1	2	0	5
TTLL8	0	0	0	0	0
MLC1	297	324	307	13	12
MOV10L1	10	8	8	0	7
RP5-898I4.1	1	0	0	0	1
PANX2	59	71	74	9	7
TRABD	33	40	37	103	79
RP3-402G11.25	2	7	3	4	5
RP3-402G11.26	9	17	7	2	0
SELO	52	72	33	77	37
RP3-402G11.27	7	2	0	5	1
RP3-402G11.28	2	2	3	4	2
TUBGCP6	248	289	237	279	174
HDAC10	4	9	12	9	9
MAPK12	60	57	56	116	73
MAPK11	64	98	82	25	23
PLXNB2	241	301	248	475	285
DENND6B	61	101	67	36	20
XX-C283C717.1	0	1	0	0	0
XX-C00717C00720L.1	0	0	0	0	0
PPP6R2	170	249	259	227	150
RN7SL500P	0	0	1	0	0
SBF1	682	840	696	530	339
ADM2	0	2	1	27	21
MIOX	0	0	0	1	1
LMF2	57	73	69	174	113
NCAPH2	86	93	78	219	131
SCO2	4	24	10	19	14
CTA-384D8.36	8	12	8	7	1
TYMP	23	24	30	58	22
ODF3B	9	11	9	11	12
CTA-384D8.35	0	1	0	0	1
CTA-384D8.34	0	0	0	0	0
CTA-384D8.31	0	0	0	6	2
KLHDC7B	0	0	1	3	8
SYCE3	0	0	1	0	0
CPT1B	5	3	5	9	8
CHKB-CPT1B	2	4	2	6	5
CHKB	6	8	4	11	8
CHKB-AS1	5	1	3	1	2
CTA-384D8.33	0	2	4	1	2
MAPK8IP2	410	545	482	13	17

```
ARSA            53      55      65      63      47
SHANK3          363     456     375     43      25
RNU6-409P       0       0       0       0       0
AC000036.4      0       2       1       0       1
ACR             0       0       0       0       2
AC002056.5      0       0       0       0       0
AC002056.3      0       0       0       0       0
RPL23AP82       41      59      54      32      23
RABL2B          74      62      54      68      50
```

{{Edet}}

Word, Character, and Line Count in a File

The `wc` command is used to obtain word, character, and line count in a file.

```
wc file.txt
```

The output (from left to right) for `wc` can be interpreted as follows.

- 2 indicates that the file has two lines.
- 23 indicates that the file has 23 words.
- 135 indicates that the file has 135 characters.
- The file in which statistics were generated (ie. file.txt).

```
  2  23 135 file.txt
```

The `wc` results above can be obtained separately. For instance, to just get the number of lines in a file include the `-l` option.

```
wc -l file.txt
```

```
2 file.txt
```

Word count can be obtained using the `-w` option.

```
wc -w file.txt
```

```
23 file.txt
```

Character count can be obtained uisng the `-m` option.

```
wc -m file.txt
```

```
135 file.txt
```

Pattern Searching

Sometimes users may want to search for a keyword in a file. The `grep` command can be used to do this. The `grep` command prints every line in a file that contains the search pattern. The arguments in `grep` are as follows.

- Search pattern (ie. Linux)
- File to search (ie. file.txt)

```
grep Linux file.txt
```

```
Linux is a variety of Unix, and sometimes the names are used interch
```

To search for lines that do not contain a pattern include the `-v` option.

```
grep -v Linux file.txt
```

```
Unix is an operating system, just like Windows or MacOS.
```

What would happen if "linux" with a lower case "l" is used as the search pattern instead of an upper case "L" as shown in the file.txt? Nothing is returned because `grep` is case sensitive. To ignore case, include the `-i` option.

```
grep linux file.txt
```

```
grep -i linux file.txt
```

```
Linux is a variety of Unix, and sometimes the names are used interch
```

Working with Software Installed on Biowulf

The `module` command is important for anyone who wishe to work with software that are installed on Biowulf. This `avail` option enables users to browse the software that are available on the cluster.

```
module avail
```

Glimpse of `module avail` results. Users can scroll up/down using the arrow keys to learn what software are available. Hit Q to exit the `module avail` results and return to the prompt.

```
------------------------------------------------------- Global Aliases
   bowtie1        -> bowtie/1.3.1                  deeptrio/1.6.0 ->
   bowtie2        -> bowtie/2-2.5.3                deeptrio/1.6.1 ->
   deeptrio/1.5.0 -> deepvariant/1.5.0-deeptrio

---------------------------------------- /data/classes/BTEP/apps/
   biostars/1.0

------------------------------------------------ /usr/local/lmod/module
   3DSlicer/4.8.1                                          hwloc/2.9
   3DSlicer/5.2.2                                   (D)    hwloc/2.9
```

To find if a particular software is available include the software name with the `avail` option.

```
module avail fastqc
```

```
------------------------------------------------------- Global Aliases


------------------------------------------------ /usr/local/lmod/module
   fastqc/0.11.8    fastqc/0.11.9    fastqc/0.12.1 (D)

  Where:
   D:  Default Module

Module defaults are chosen based on Find First Rules due to Name/Vers
See https://lmod.readthedocs.io/en/latest/060_locating.html for detai

If the avail list is too long consider trying:

"module --default avail" or "ml -d av" to just list the default modul
```

```
"module overview" or "ml ov" to display the number of modules for eac

Use "module spider" to find all possible modules and extensions.
Use "module keyword key1 key2 ..." to search for all possible modules
```

To get a description of a software use `module whatis`.

```
module whatis fastqc
```

```
fastqc/0.12.1       : fastqc: It provide quality control functions to
fastqc/0.12.1       : Version: 0.12.1
```

Warning

Upon signing onto Biowulf, users will land in the "log-in" node, which is not meant for analyzing data or transferring files. To perform these task, request time and resources on a compute node using the `sinteractive` command. To learn more about the `sinteractive` command, see https://bioinformatics.ccr.cancer.gov/docs/unix-on-biowulf-2024/Lesson4/#requesting-an-interactive-session *(https://bioinformatics.ccr.cancer.gov/docs/unix-on-biowulf-2024/Lesson4/#requesting-an-interactive-session)*

To load a software use `module load` followed by the name of the package.

```
module load samtools
```

```
[+] Loading samtools 1.19  ...
```

Getting Help

To get help, use the `man` command which pulls up the manual for a command. Use the up and down arrow keys to scroll through the manual and learn about the different options. Hit q to return to exit `man` and return to the prompt.

For instance, to pull up the manual for `grep`, do the following.

```
man grep
```

Glimpse of `man` command output for `grep`.

```

```

```
GREP(1)                          General Commands Manual

NAME
       grep, egrep, fgrep - print lines matching a pattern

SYNOPSIS
       grep [OPTIONS] PATTERN [FILE...]
       grep [OPTIONS] -e PATTERN ... [FILE...]
       grep [OPTIONS] -f FILE ... [FILE...]

DESCRIPTION
       grep searches for PATTERN in each FILE.  A FILE of “-” stands
       FILE is given, recursive  searches  examine  the  working  di
       searches read standard input.  By default, grep prints the ma

       In  addition,  the variant programs egrep and fgrep are the s
       respectively.   These  variants  are  deprecated,  but   are
       compatibility.

OPTIONS
   Generic Program Information
       --help Output a usage message and exit.

       -V, --version
              Output the version number of grep and exit.
```

Continual Learning

- See https://bioinformatics.ccr.cancer.gov/docs/unix-on-biowulf-2024/ *(https://bioinformatics.ccr.cancer.gov/docs/unix-on-biowulf-2024/)* for a more detailed Unix and Biowulf class. Check the BTEP Calendar *(https://bioinformatics.ccr.cancer.gov/btep/)* for upcoming advanced Unix classes. Email ncibtep@nih.gov or drop by BTEP's Virtual Office Hours *(https://bioinformatics.ccr.cancer.gov/btep/introducing-btep-office-hours/)* with questions.
- Coursera and Dataquest offers self learning classes for Unix and other data science related topics such as R and Python programming. See https://bioinformatics.ccr.cancer.gov/btep/self-learning/ *(https://bioinformatics.ccr.cancer.gov/btep/self-learning/)* to learn how to request a Coursera or Dataquest license.
    - Example Coursera classes:
        - Hands-on Introduction to Linux Commands and Shell Scripting *(https://www.coursera.org/learn/hands-on-introduction-to-linux-commands-and-shell-scripting#modules)*
        - Command Line Tools for Genomic Data Science *(https://www.coursera.org/learn/genomic-tools#modules)*

        - Unix and Bash for Beginners Specialization *(https://www.coursera.org/specializations/unix-and-bash-for-beginners)*
    - Example Dataquest classes:
        - https://www.dataquest.io/course/command-line-elements/ *(https://www.dataquest.io/course/command-line-elements/)*
        - https://www.dataquest.io/course/command-line-intermediate/ *(https://www.dataquest.io/course/command-line-intermediate/)*