**Tool Implementation Instruction**

The PapViz is implemented on Python3.x as well as it is made to run on Python 2.x. The dependent packages are: Matplotlib, Pillow (PIL.Image) and Numpy. (Note: This Tool will not run on Cygwin from Windows. If have already installed python 2.x OR python 3.x as AND configure numpy, matplotlib and Pillow dependencies either in linux or in Windows, you can run the PapViz directly from command line)

The “configure.sh” file covers complete implementation of the tool including. By this, user do not have to worry about calling the PapViz manually. The scrip automatically takes care of:

* Configure Conda
* Configure Python3 with separate environment through the Conda
* Call the code automatically and generate and store outputs in directory wherever user wants.

(See at the end detail of the three steps in Tutorial section).

Note: If the tutorial instruction fails somehow for the configuration of python, please do the manual configuration as explained in “**Manual Configuration section**” section in the end

**Directory Structure**

PapViz.zip

--- PapViz

---- input

---- input\_data.txt

---- reference\_gene\_models.txt

---- PapGene\_SeqCovViz\_Instructions.doc

---- output

----- (All sample-wise output results)

---- script

---- PapViz.py (executable python script)

**Specification of the Output**

* The output of the tool generates individual PDF files; each represent a sample (The file-name is corresponding to the Sample-ID).
* Each transcripts visualization is shown with Gene-ID and Transcript-ID.
* Each blocks of exons on a transcript are represented with rectangle shape with arrow at the end of the blocks. Size of the block is related to exon length. The arrow direction on the exon block indicates orientation of transcript on DNA strand. The 5'->3' is in forward and 3'->5' in reverse direction and order of exon number should also be considered accordingly. Orientation of the transcripts can also be seen by looking at end of transcript with text either 5’ or 3’ on left and right position.
* Position of the reads are marked with fixed size pillar mark over the corresponding exons. The position of reads either at upper-side, if the read is from positive strand (+), or at lower-side, if the read is from negative strand (-).
* Color of the reads indicates above or below to a cutoff for the templates values of the reads. For 0> and <200 template value, the reads are shown with red color and for >=200 template value, the reads are shown with green color.
* Marker for overlapped reads are merged and plotted as single marker. i.e.,

PG-MAT-DNA-0004-01 CDKN2A\_029\_50 CDKN2A NM\_000077.4 CCDS6510.1 chr9 - + 21974652 21974761 1776

PG-MAT-DNA-0004-01 CDKN2A\_001\_21 CDKN2A NM\_000077.4 CCDS6510.1 chr9 - + 21974743 21974859 25

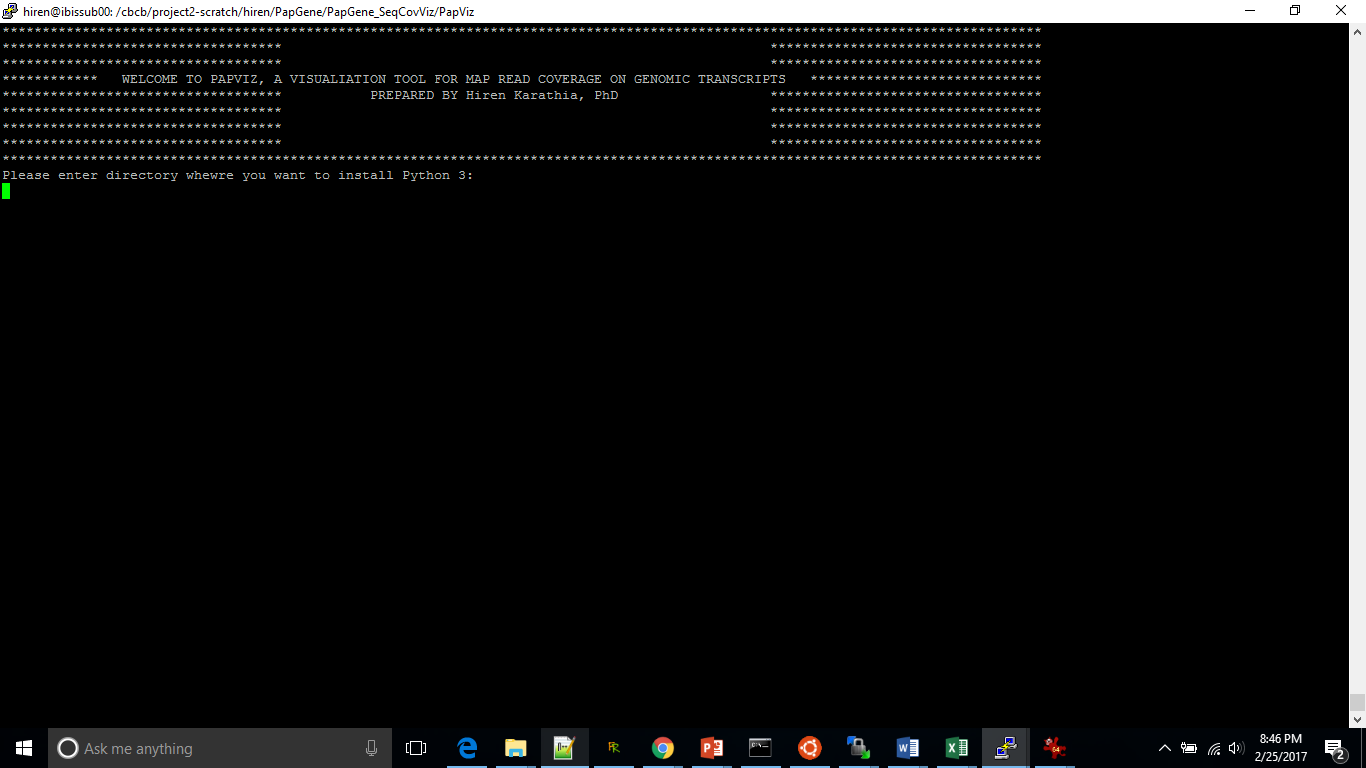
In above example of input file, the two reads whose mapping regions are partially overlapped. In such cases read markers are merged and considered as single mark. Thus reason of having less number of read marks coverage on a transcript as compared to the number of read entries in the input file is certainly because those overlapped reads are merged and plotted as single mark.

**Delivering items**

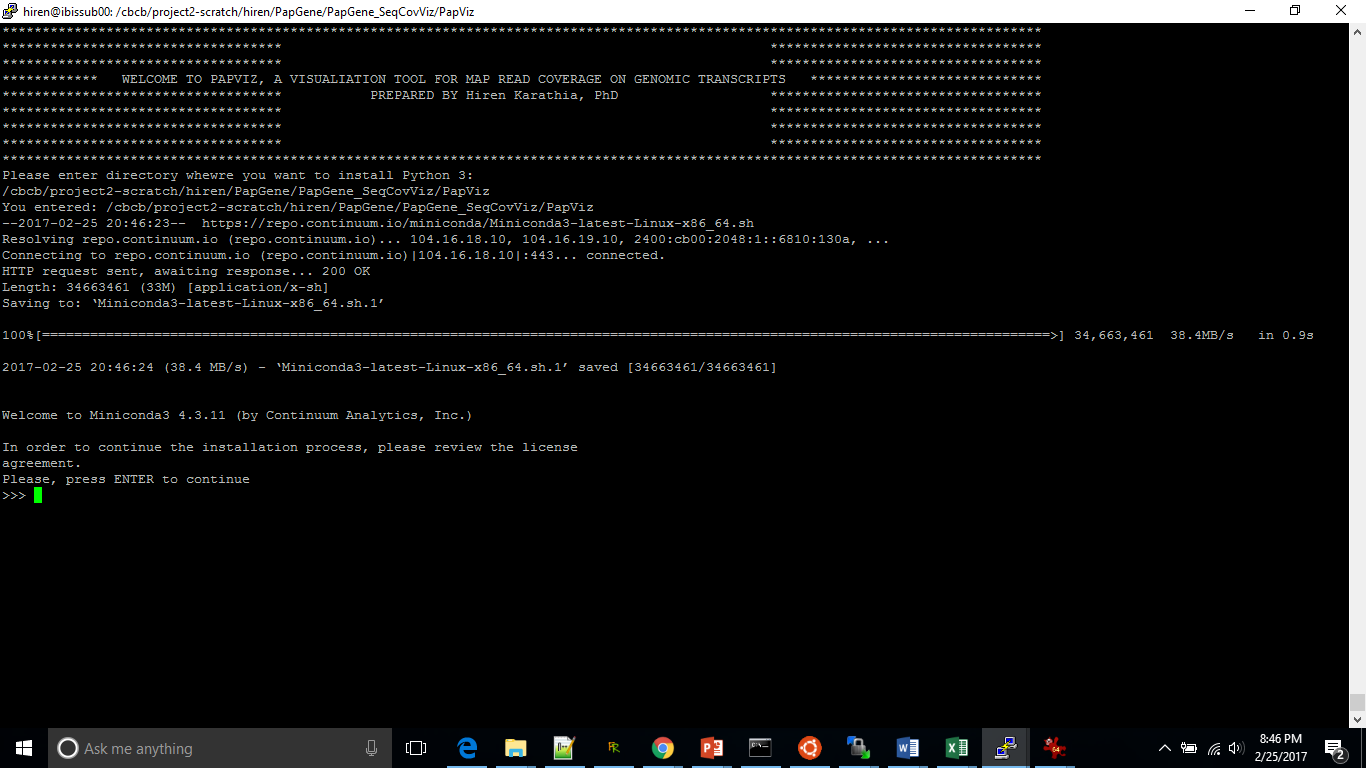
1. Code solution.
2. Instructions on how to execute your code solution (Microsoft Word).
3. The five output documents generated by your solution when applied to the supplied input file.

**Tutorial of the configuration and run the tool**

1. At running the “*congfigure.sh*” script on terminal by typing “*bash configure.sh*”, you will see following screen. Here you will need to provide directory, where you uncompressed the PapViz package i.e., type here “*/home/John/PepViz”* if you uncompressed the *PepViz* package inside the “*/home/John*” directory



1. After you entered your directory of PepViz, first it will download Conda and will configure automatically. Next it will configure python inside your directory. During this, it will ask few questions like accepting terms and conditions etc., you just need to press enter until you reach to a terminal where it will ask you the following:



It will ask you, as follow, for input directory where you wish to install miniconda3.

*Miniconda3 will now be installed into this location:*

*<default directory>*

*- Press ENTER to confirm the location*

*- Press CTRL-C to abort the installation*

*- Or specify a different location below*

Enter here whatever directory you entered in the first step (“*/home/John/PepViz*”) with adding “*miniconda3*” directory name i.e., enter here “*/home/John/PepViz/miniconda3*”

At the end it will ask you followong

Do you wish the installer to prepend the Miniconda3 install location to PATH in your <home>/.bashrc ? [yes|no]

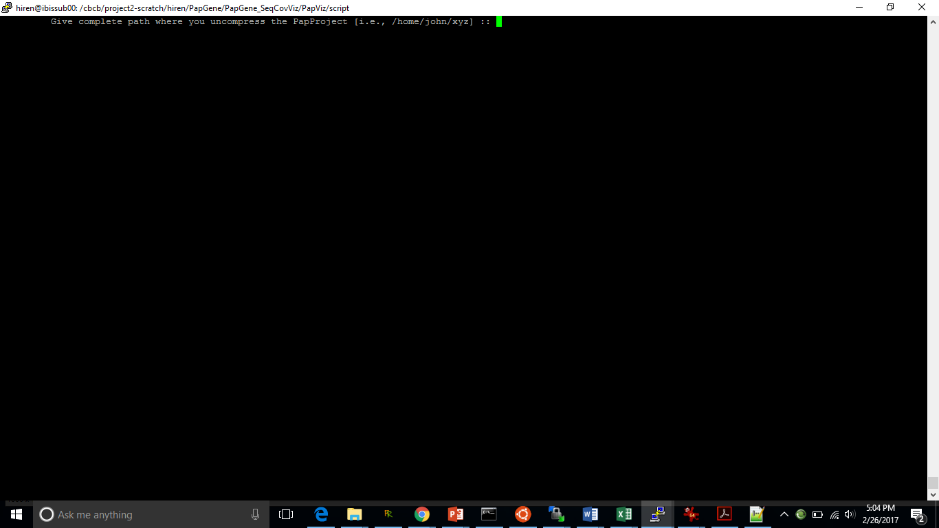
[no] >>>

Enter “*no*” here

Next it will ask you to install required software. Put there “yes”

>>> yes

1. At the end PepViz Starts automatically. First it will clear the screen and will ask you to input directory where ever you store your PepViz package. i.e., here you can enter “*/home/John/PepViz”*



Next it will ask you to insert directory where you want to store output i.e., you can enter “*/home/John/PepViz/output”*.

**That’s it, you will get your output stored in your selected directory**

**Manual Configuration**

If the above process fails, please follow manual configuration as follows. The reason of failing the above configuration may be if

1. You are working with Ubuntu. In Ubuntu, I tested the configuration and came to know that there is permission issue that gives error in the automatic configuration.
2. By mistake, you did not input the “*miniconda3*” directory correctly in step-2 above. If in the above step-2, you provide the “*miniconda3”* directory apart from where you uncompressed “PapViz” folder, in next step it will not find that folder to configure dedicated Python folder, called “Py3” and at the end you will get error of configuration in mid and will not call the PapViz program appropriately.
3. Any other unknown reason!!!

But do not worry, just follow the steps, as given below and you will surely get the Python3 configuration and call the PapViz correctly.

Steps of manual configuration and call the PapViz

1. Go to the folder where you uncompressed the PapViz

i.e., if you uncompressed the PapViz inside “*/home/John”*, go to PapViz by typing

*$ cd /home/John/PapViz*

1. Type following command

*$ wget* [*https://repo.continuum.io/miniconda/Miniconda3-latest-Linux-x86\_64.sh*](https://repo.continuum.io/miniconda/Miniconda3-latest-Linux-x86_64.sh)

1. Next type

*$ bash Miniconda3-latest-Linux-x86\_64.sh*

You will follow self-explanatory instructions like

*Please, press ENTER to continue*

*>>>*

Just enter here, and keep entering until you don’t reach to following

*Do you approve the license terms? [yes|no]*

*>>>*

*Please answer 'yes' or 'no':*

Type here *“yes”*

In following step you will be asked following

*Miniconda3 will now be installed into this location:*

*/home/hiren/miniconda3*

*- Press ENTER to confirm the location*

*- Press CTRL-C to abort the installation*

*- Or specify a different location below*

*[/home/<user>/miniconda3] >>>*

If you press “ENTER“, as it explained, by default the *miniconda3* will get configured in the */home/<user>* directory, if not, you can input your desired directory where you want to configure the “*miniconda3*”. In case of the second option, we recommend to input here your PapViz directory i.e., */home/John/PapViz/miniconda3*

Next, you will be asked following:

*Do you wish the installer to prepend the Miniconda3 install location*

*to PATH in your /home/<user>/.bashrc ? [yes|no]*

*[no] >>>*

You can type either “yes” or “no”, if you type yes, it will configure this python in your environment.

1. If you are inside “*/home/John/PapViz”* and if you installed the miniconda3 from the previous step in the “*/home/John/PapViz*” thentype as follows instructions

If you are not using Ubuntu, you can do as follows

*$ /home/John/PapViz/miniconda3/bin/conda create --name Py3 python=3 numpy matplotlib Pillow*

If you are using ubuntu then

*$ sudo /home/John/PapViz/miniconda3/bin/conda create --name Py3 python=3 numpy matplotlib Pillow*

*$ source activate Py3*

If it doesn’t work, then type

*$ source activate /home/John/PapViz/miniconda3/envs/Py3*

1. Next call the PapViz as follows. Assume you are inside PapViz directory i.e., “*/home/John/PapViz*”

*$ python script/PapViz.py*

In the terminal you will be asked to input the directory where you uncompressed the PapViz. So, give complete directory path i.e., in our case “*/home/John/PapViz”*

Next, it will ask you to input desired output directory, you can put any directory you wish. In this case you can put *“/home/John/PapViz/output”*

AND

That’s it, you will receive a message saying you can see your output inside the directory you mentioned. So easy !!! isn’t it??

Enjoy….

Wait, unfortunately ☹

If through any unknown factor, still you get trouble in running the PapViz or you find any bug in output of it, please write me on [hirkarathia@gmail.com](mailto:hirkarathia@gmail.com), I will be very much happy to help you.

**THANK YOU**