Section II: Visualizing SNPs with RNAMapper

This step allows you to visualize the SNP marker frequency within your genome as well as on chromosomes of interest. RNAMapper does this by creating some variant call files (.vcf) with some special columns and then visualizing these files. Running RNAmapper.py requires all of your chromosomes split into their own .vcf file (see previous section) and running each chromosome file individually or in parallel.

1. Creating the marker files using RNAmapper.py

For those with linux knowledge, please see the "--help" option or the GitHub repository's "Usage" section for argument information.

For those without linux knowledge, a helper script has been created in the `helperscripts/` folder. The helper script runs each chromosome in parallel to minimize time and takes two required arguments and an optional argument.

- Argument 1 (required): The path to your .vcf chromosome files directory (either absolute or relative to the helperscript file)
- Argument 2 (required): The common prefix of your .vcf chromosome files. For example, for a prefix "mappingproject_" you might have files named: "mappingproject_wt_chr1.vcf" and "mappingproject_wt_chr2.vcf" within your directory.
- Argument 3 (optional): A path to an already created output directory. If not specified, outputs in the directory specified in Argument 1.

For example, if you had your split chromosome files in a directory with the path: "~/data/RNAmapper/" and your files used the prefix "mappingproject_", you might use three arguments. The directory in the third argument must already be created. Please notice the inclusion of forward slashes at the end of the first and third arguments.

- Argument 1: ~/data/RNAmapper/
- Argument 2: mappingproject_
- Argument 3: ~/data/RNAmapper/mappedfiles/

You would then run your command with arguments separated by spaces like so:

`./RNAmapper.sh ~/data/RNAmapper/ mappingproject_ ~/data/RNAmapper/mappedfiles/`

When using the helper script, it will use all of your computer's available resources. The time it will take to run is dependent on the specifications of your computer. We recommend not using your laptop while the script completes.

Once the script is complete, output files will be present in either the specified output directory or the same folder as your split .vcf files. These files include a marker file and a stats file for each mutant chromosome.

2. Visualizing the genome with RNAMGenomeGrapher.R

This script visualizes linked regions in a genome in conjunction with the RNAmapper.py script. This script takes three required arguments and an optional argument:

- Argument 1 (required): The directory for the marker files for the genome of interest.
- Argument 2 (required): The common prefix for the mutant marker files that was specified above when creating the files. If you are unsure, this should consist of everything before "mut_chr#_atMarkers.vcf"
- Argument 3 (required): The path and name for the output plot. This will have .jpg automatically appended to end.
- Argument 4 (optional): An optional argument to include extra plot info such as a title, axis labels, and improved tick marks. Can be specified with any string such as: "TRUE."

For example, if you had your marker files with the prefix "mappingproject_" in "~/data/mapping/mappedfiles/" directory and wanted to name your plots "MyFavoriteMutant", you might use all four arguments. Please notice the inclusion of forward slashes at the end of the first argument and the lack of file extension after the third argument.

- Argument 1: ~/data/RNAmapper/mappedfiles/
- Argument 2: mappingproject_
- Argument 3: ~/data/RNAmapper/mappedfiles/MyFavoriteMutant
- Argument 4: TRUE

You would then run your command with arguments separated by spaces like so:

`Rscript ./RNAMGenomeGrapher.R ~/data/RNAmapper/mappedfiles/ mappingproject_ ~/data/RNAmapper/mappedfiles/MyFavoriteMutant TRUE`

This would output a file named: “MyFavoriteMutant.jpg” in the “~/data/RNAmapper/mappedfiles/” with your mapped genome. You can then use this image to decide which chromosomes you would like to map in more detail.

3. Visualizing chromosomes with RNAMChromosomeGrapher.R

This script visualizes linked regions in a chromsome in conjunction with the RNAmapper.py script. This script takes three required arguments and an optional argument:

- Argument 1 (required): The path to the marker file of interest.
- Argument 2 (required): The path to the stats file of which to append the region of linkage.
- Argument 3 (required): The path to the output plot which will have chromosome number and .jpg automatically appended to end.
- Argument 4 (optional): The linkage ratio in a proportion. The default is 0.98.

If you wanted to visualize chromosome 2 in our “mappingproject_” example, you might just use the three required arguments. Please notice the inclusion of forward slashes at the end of the first.

- Argument 1: ~/data/RNAmapper/mappedfiles/mappingproject_mut_chr2_atMarkers.vcf
- Argument 2: ~/data/RNAmapper/mappedfiles/mappingproject_mut_chr2_stats.txt
- Argument 3: ~/data/RNAmapper/mappedfiles/MyFavoriteMutant

You would then run your command with arguments separated by spaces like so:

`Rscript ./RNAMChromosomeGrapher.R ~/data/RNAmapper/mappedfiles/mappingproject_mut_chr2_atMarkers.vcf ~/data/RNAmapper/mappedfiles/mappingproject_mut_chr2_stats.txt ~/data/RNAmapper/mappedfiles/MyFavoriteMutant`

This would output a file named: “MyFavoriteMutant_chr2.jpg” in the “~/data/RNAmapper/mappedfiles/” directory and append some information about linkage to the stats file. You can then use this image to decide which linkage region you would like to move forward with in Section III.