***TSC2* MHPA pipeline documentation**

1. Include in one folder:

Python files:

* Analyzer.py
* v12-q50.py
* merger.py
* zeroscreator.py

Text files:

* genomeADalts.txt
* genomeADfreqs.txt
* genomeADposns.txt
* genomewithinst.txt
* exonCoordData.txt

Matlab files:

* Matlab\_commands\_3\_16\_GenomADupdate.m
* matlab\_input\_genomeADcoord.mat

Other files:

* MHPA\_TSC2.bed
* dataWithaa.xlsm
* indelmacro.xlsm
* all .bam/.bai files to be analyzed

1. Use python Analyzer.py in Terminal. In the Analyzer window (see below) choose the TSC2 gene location (chr16:2097990-2138713), the minimum variant allele frequency, and the minimum read count for the variant allele. Proceed with ‘Begin Analysis’.

Graphical user interface, text, application, chat or text message

Description automatically generated

1. When the python analysis is complete (mergez.txt and namelist.txt files generated), run the ‘Matlab\_commands\_3\_16\_GenomADupdate.m’ script in Matlab.
2. Open the Matlab analysis output files ( ‘snvdata.csv’ and ‘formatfilterindels.csv’) in excel. Open ‘dataWithaa.xlsm’ and ‘indelmacro.xlsm’, and use macros (Developer -> Macros) for annotation and formatting of the output files.