# RowAnalytics

·         Distribution of the number of 'hits' in the upstream/inside/downstream components of each gene

·         Within a gene, are the 'hit' SNPs equally represented or are there specific genotypes that are over/under-represented?

·         How many simple networks do each of the 'hit' SNPs contribute to?

·         Distribution of how many 'hit' SNPs for the same gene the simple networks contain

·         How many unique cases are there for each SNP ('hit' or 'non-hit'? Are there any patterns?

·         Can we identify driver mutations? Can we identify other mutations that drive/affect disease phenotype pleitropy? Are there any commonalities in the pathways of the driver/secondary mutations?

·         Are any of the 'hit' SNPs druggable? (this may require Annotator and can't be answered yet)

·         Which SNPs were associated with the protective effects? Which simple & merged networks do they occur in? How many other SNPs were associated with those genes and were any of them 'hits' in the disease risk analysis?

There were 91 hits in the Risk-associated SNPs that were within 50kb of a transcription start site. Of these 91 hits, 46 of them were within the protein coding sequence. 2 were after the transcription start site but before the coding sequence, and 7 were downstream of the transcription end site. 23 were within 50kb upstream, and 13 were within 50kb downstream.

For the protective variants, a similar pattern was observed. 21 were within the coding sequence, 0 were were after the tx start and 4 were between the coding end and transcription end. 3 were 50kb downstream and 5 were 50kb upstream.

The script used to tally these counts was biased toward coding sequences for the tally as sometimes alternative reading frames led to repeats of tallies. For example, when there was a duplicate where one reading frame counted as a coding sequence and another counted it as downstream of the coding sequence before the transcription end site, the coding sequence hit was taken and the hit for the transcription end was removed.