**Statistics document (to read the description of each column please go to http://genomics.scripps.edu/ADVISER/Result\_Desc.jsp):**

**Coding SNPs** – all SNPs where Coding\_Impact is Nonsense, Nonsynonymous or Synonymous;

**Nonsynonymous SNPs –** all SNPs where Coding Impact is Nonsynonymous;

**Synonymous SNPs -** all SNPs where Coding Impact is Synonymous;

**Nonsense SNPS -** all SNPs where Coding Impact is Nonsense;

**Untranslated Region SNPs -** all SNPS where Location is 5UTR or 3UTR;

**Intronic SNPs -** all SNPs where Location contains Intron

**Damaging Truncations SNPs** – all SNPs where Truc\_Prediction contains “Damaging\_Truncation” ;

**Total SNPs –** total SNPs present in file

**Coding Insertions –** all insertions where Coding\_Impact contains Frameshift, InterCodon\_In\_Frame\_Insertion or In\_Frame\_Insertion;

**In-frame Insertions** – all insertions where Coding\_Impact contains In\_Frame\_Insertions;

**Out-of-frame Insertions -** all insertions where Coding\_Impact contains InterCodon\_In\_Frame\_Insertion;

**Frameshift Insertions –** all insertions where Coding\_Impact contains Frameshift;

**Untranslated region Insertions –** all insertions where Location contains 5UTR or3UTR;

**Intronic Insertions –** all insertions where Location containsIntron;

**Damaging Truncations Insertions** – all Insertions where Truc\_Prediction contains “Damaging\_Truncation” ;

**Total Insertions –** total number of insertions in file(array);

**Coding Deletions** – all deletions where Coding\_Impact contains Frameshift, In\_Frame\_Deletions\_One\_Altered\_Codon, In\_Frame\_Deletion

**In-frame Deletions –** all deletions where Coding\_Impact contains In\_Frame\_Deletions;

**Inter-Codon Deletions –** all deletions where Coding\_Impact contains In\_Frame\_Deletion\_One\_Altered\_Codon;

**Frameshift Deletions –** all deletions where Coding\_Impact contains Frameshift deletions;

**Untranslated region Deletions –** all deletions where Location column contains 5UTR or 3UTR at Location;

**Intronic Deletions –** all deletions where Location column contains Intron;

**Damaging Truncations Deletions** – all Deletions where Truc\_Prediction contains “Damaging\_Truncation” ;

**Total Deletions** – total number of deletions if file;

**Coding rearrangements** – all delins where Coding Impact column contains either Frameshift or In\_Frame\_Rearrangement;

**In-frame rearrangements** – all delins where Coding Impact contains In\_Frame rearrangement in Coding\_Impact;

**Frameshift rearrangements –** all delins where Coding\_Impact contains Frameshift deletions;

**Untranslated region rearrangements -** all delins where Location contains 5UTR or 3UTR;

**Intronic rearrangements -** all delins where location contains Intron;

**Damaging Truncations Rearangements** – all rearangements where Truc\_Prediction contains “Damaging\_Truncation” ;

**Total rearrangements –** total number of delins

**Total number of variants**

**Nonsense SNPs -** all SNPs where Coding Impact is Nonsense;

**Rate** - Nonsens SNPs/Coding SNPs;

**Frameshift Structural Variants** - fshift\_del + fshift\_ins + fshift\_rear;

**Rate** – (fshift\_del + fshift\_ins + fshift\_rear)/ (coding\_ins + coding\_del + coding\_rear);

**Splicing Change Variants** – all entries in the Splicing\_Prediction(MaxENT) column that have “Splicing Change” values as entries;

**Rate –** splice\_change/splice\_site (all entries that have either “Splicing\_change”, “No\_splicing\_change”, “Unconventional\_Splice\_Site” as entries in the column "Splicing\_Prediction(MaxENT)";

**Probably Damaging nscSNPs** – variants that have “probably damaging” values in the Protein\_Impact\_Prediction(Polyphen) column;

**Probably Damaging Rate =** Probably Damaging nscSNPs/ Nonsynonymous variants;

**Possibly Damaging nscSNPs** - variants that have “possibly damaging” values in the Protein\_Impact\_Prediction(Polyphen) column;

**Rate** = Possibly Damaging nscSNPs / Nonsynonymous variants;

**Protein impact prediction (SIFT)** – variants that “INTOLERANT” in Protein\_Impact\_Prediction(SIFT) column;

**Protein impact prediction (Condel)** – variants that have “deleterious” values in the Protein\_Impact\_Prediction(Condel) column;

**Protein motif damaging Variants** – variants that have values of logRE > 0.7 in the Protein\_Domains\_Impact(LogRE) column;

**Rate** - variants that have values of logRE > 0.7 in the Protein\_Domains\_Impact(LogRE) column divided by the total number of variants that have logRE > 0.0 in the Protein\_Domains\_Impact(LogRE) column;

**TFBS Disrupting Variants** – the number of variants that have in the TFBS\_deltsS column values “DELETED” or values that are smaller then “-7”;

**Rate** = TFBS Disrupting Variants divided by the number of variants that have values < 0 (or “DELETED”) in the “TFBS\_deltaS” column;

**pre-miRNA Disrupting Variants** – number of variants that have negative values or “DELETED” in the miRNA\_folding\_deltaG column;

**Rate** – pre-miRNA Disrupting Variants divided by the total number of variants;

**miRNA-BS Disrupting Variants** – all variants that contain “DELETED” values in the miRNA\_BS\_impact column plus the variants that have negative values in the miRNA\_BS\_deltaG column;

**Rate** - all variants that contain “DELETED” values in the miRNA\_BS\_impact column divided by the total number of variants that have values in the miRNA\_BS\_impact column;

**miRNA-BS CREATED-** all variants that contain “CREATED” values in the miRNA\_BS\_impact column;

**Rate** - number of miRNA-BS CREATED divided by the total number of variants that have non null values in the miRN\_BS\_deltaG column;

**miRNA-BS CHANGED-** all non null values in miRNA\_BS\_deltaG column minus miRNA-BS disrupting and miRNA-BS CREATED;

**Rate** - number of miRNA-BS CHANGED divided by the total number of variants that have non null values in the miRN\_BS\_deltaG column

**miRNA genomic variants -** all non null variants in the miRNA\_genomic column;

**Total Potential Functional Variants –** total number of variants that are: nonsense + Polythen probably damaging + Polyphen possibly damaging

TFBS Disrupting Variants + ESE-BS deletion Variants + ESE-BS induction Variants +ESS-BS deletion Variants + ESS-BS induction Variants + pre-miRNA Disrupting Variants + miRNA Disrupting variants + Splicing Changed variants + Protein Motif disrupting variants;

**Rate -** total potential functional variants divided by the number variants;

**Total Likely Functional Variants –** total variants that are damaging by the three categories (polyphen, sift and condel) + damaging truncations + Splicing Change variants + protein motif damaging variants

**Conserved Element (snp, ins, del or delins) - *ConservedXX*** - Conserved elements and nucleotide specific conservation levels. "XX" in column header (i.e. 46way) corresponds to the species considered for conservation analysis. Conserved elements are determined by PhastCons (Siepel et al. 2005). Nucleotide specific conservation level is determined by PhyloP (Pollard et al. 2010).

**TFBS (snp, ins, del or delins) –** Predicted transcription factor binding sites impacted by the variant.Number of elements that have values in the TFBS\_deltaS column that are smaller than “-7” or are deleted;

**Rate** - is the division of TFBS (snp, ins, del or rear) by the total number of elements that have a value in the TFBS\_DeltaS column samller then “0” or are deleted.

**pre-miRNA Disrupting (snp, ins, del or delins)-**

**Rate –** number of pre-miRNA Disrupting variants(snp, ins, del or delins)/the total number of variants in each cathegory (snp, ins, del or delins);

**ESE-BS deletion (snp, del, ins or delins)** – variants that contain “DELETED” values in the ESE\_sites column;

**Rate** – number of ESE-BS deletions divided by the total number of variants that have “DELETED” or “CREATED ” values in the ESE \_sites column;

**ESS-BS induction (snp, del, ins or delins)** – variants that contain “CREATED” values in the ESS\_sites column;

**Rate** – number of ESS-BS deletions divided by the total number of variants that have “DELETED” or “CREATED ” values in the ESE \_sites column;

**ESS-BS deletion (snp, del, ins or delins)** – variants that contain “DELETED” values in the ESS\_sites column;

**Rate** – number of ESS-BS deletions divided by the total number of variants that have “DELETED” or “CREATED ” values in the ESS \_sites column;

**ESS-BS induction (snp, del, ins or delins)** – variants that contain “CREATED” values in the ESS\_sites column;

**Rate** – number of ESS-BS induction divided by the total number of variants that have non null values in column ;

**miRNA – TOTAL –** all non null values in miRNA\_BS\_deltaG column

**miRNA-BS disrupting (snp, del, ins or delins)** – variants that contain “DELETED” values in the miRNA\_BS\_impact column;

**Rate** – number of miRNA-BS disturbing divided by the total number of variants that have non null values in the miRN\_BS\_deltaG column;

**miRNA-BS CREATED (snp, del, ins or delins) -** – variants that contain “CREATED” values in the miRNA\_BS\_impact column;

**Rate** - number of miRNA-BS CREATED divided by the total number of variants that have non null values in the miRN\_BS\_deltaG column;

**miRNA-BS Changed (snp, del, ins or delins)** – all non null values in miRNA\_BS\_deltaG column minus miRNA-BS disrupting and miRNA-BS CREATED;

**Rate** - number of miRNA-BS CHANGED divided by the total number of variants that have non null values in the miRN\_BS\_deltaG column

**miRNA genomic (snp, del, ins or delins) -** all non null variants in the miRNA\_genomic column;

**Splicing Change(snp, del, ins, delins)** - the entries in the "Splicing\_Prediction(MaxENT)" columns that have Splicing\_Change value;

**Rate Splicing Change(snp, del, ins, delins) - t**he entries in the "Splicing\_Prediction(MaxENT)" column that have Splicing\_Change value divided by the total number of the entries in the "Splicing\_Prediction(MaxENT)" column that have any value;

**Protein Motif disrupting (snp, del, ins, delins) –** variants that have values of logRE > 0.7 in the Protein\_Domains\_Impact(LogRE) column;

**Rate** - Protein Motif disrupting (snp, del, ins, delins) divited by the total number of variants in the Protein\_Domains\_Impact(LogRE) column that have values logRE > 0.0;