Description Structural Variants Annotations in WGS

Department of Human Genetics
Radboud University Medical Center,
Nijmegen, The Netherlands

June 2018

For questions please contact:
Maartje.vandeVorst@radboudumc.nl
Jordi.CorominasGalbany@radboudumc.nl

SegmentID
Individual ID for each structural variant.

Chromosome
Chromosome where the identified variant is located on. Chromosomes can be chr1 – chr22, chrX, chrY and chrM (mitochondrial DNA).

Start position
Start position of the identified variant (genomic position on the respective chromosome). Positions are 1 based and we count the bases.

End position
End position of the identified variant (genomic position on the respective chromosome). In case of a substitution or insertion the start and end position are identical, in case of a deletion the end position is the start position + the length of the deletion.

Length
Total size of the structural variant. In case of breakend and insertion the Length =1.

Gain
Only available for copy number variants indicating gain of copy or loss of copy.

MATEID
ID of mate breakend. *Standard argument in SV detection, not relevant for most of researchers.*

CIGAR
CIGAR alignment for each alternate indel allele. CIGAR is a string describing how the reads aligns with the reference. *Standard argument in SV detection, not relevant for most of researchers.*

RIGHT_SVINSSEQ and LEFT_SVINSSEQ
Known right and left side of insertion for an insertion of unknown length.

HOMLEN and **HOMSEQ**
Length and Sequence of base pair identical homology at event breakpoints. *Standard argument in SV detection, not relevant for most of researchers.*

CIPOS and **CIEND**
Confidence interval around POS (start) and END. It gives an estimation of how reliable are the starting and end position of the structural variants. *Standard argument in SV detection, not relevant for most of researchers.*

SVINSSEQ and **SVINSLEN**
Sequence and Length of insertion.

SVLEN
Difference in length between REF and ALT alleles.

MATE_BND_DEPTH
Read depth at remote translocation mate breakend.

JUNCTION_QUAL
If the SV junction is part of an EVENT (ie. a multi-adjacency variant), this field provides the QUAL value for the adjacency in question only. *Standard argument in SV detection, not relevant for most of researchers.*

EVENT
ID of event associated to breakend. *Additional ID not relevant for most of researchers.*

SVTYPE
Type of structural variant. Possible annotations: BND (breakend --- associated with translocations), DEL (Deletion), DUP (Duplication), INS (Insertion) and INV (Inversion).

END
End position of the variant described in this record.

BND_DEPTH
Read depth at local translocation breakend.

Quality
SV calling quality score.

Filter
Filter indicating reliable variants (PASS) or the statistic that failed the quality threshold.

GT
Genotype. Where: 0/0 is homozygous for the reference allele; 0/1 heterozygous and 1/1 homozygous for the alternative allele.

FT
Sample filter, 'PASS' indicates that all filters have passed for this sample.

GQ
Genotype Quality.

PL
Normalized, Phred-scaled likelihoods for genotypes as defined in the VCF specification. *Standard argument in SV variant calling files, not relevant for most of researchers.*

PR
Spanning paired-read support for the ref and alt alleles in the order listed. *Standard argument in SV variant calling files, not relevant for most of researchers.*

SR
Split reads for the ref and alt alleles in the order listed, for reads where P(allele|read)>0.999.
Standard argument in SV detection, not relevant for most of researchers.

Start_cytoBand and **End_cytoBand**
Start and end cytoband of the variant described in this record.

ANNOVAR_GENE_COMPONENT
The gene component tells whether a variant is located within a gene or outside the gene. Possible annotations: exonic, splicing, intergenic, intronic, ncRNA_exonic, ncRNA_intronic, ncRNA_splicing, upstream, downstream, UTR3 and UTR5.

ANNOVAR_GENE_NAME
Name of the gene in which the variant was identified. In case of intergenic variants, the closest genes are annotated (indicating the distance to the variant).

DISEASE
Gene panel where the gene or region has been previously observe/associated.

Decipher_duplications_AF and **Decipher_deletions_AF**
Allele frequency from the Decipher database of the structural variant (duplication or deletion) described in this record. Data retrieved from https://decipher.sanger.ac.uk/

Decipher_duplications_OVERLAP and **Decipher_deletions_OVERLAP**
Reciprocal overlap between the structural variant and the SV (duplication or deletion) in this database.

Wellderly_413-CNV_AF
Allele frequency from the Wellderly database (focused on copy number variants) described in this record. Data retrieved from Complete Genomics.

Wellderly_413-CNV_OVERLAP
Reciprocal overlap between the structural variant and the copy number variants in this database.

Wellderly_413-Events_AF
Allele frequency from the Wellderly database (general events) described in this record.

Wellderly_413-Events_OVERLAP
Reciprocal overlap between the structural variant and the events present in this database.

GoNL_AF
Allele frequency from the GoNL database of the structural variant described in this record. Data retrieved from http://www.nlgenome.nl/

GoNL_OVERLAP

Reciprocal overlap between the structural variant and the SV in this database.

1000G_AF
Allele frequency from the 1000G database of the structural variant described in this record. Data retrieved from http://www.internationalgenome.org/

1000G_OVERLAP
Reciprocal overlap between the structural variant and the SV in this database.

Database_of_Genomic_Variants_ID
ID from the database of genomic variants of the structural variants described in this record. Data retrieved from http://dgv.tcag.ca/dgv/app/home

Database_of_Genomic_Variants_OVERLAP
Reciprocal overlap between the structural variant and the SV in this database.

EXAC_ID
ID from the EXAC database of the structural variant described in this record. Data retrieved from http://exac.broadinstitute.org/

EXAC_OVERLAP
Reciprocal overlap between the structural variant and the SV in this database.