Proposal/thoughts for data versions for current and future GTEx releases

For Option 1 Analysis Freeze (intermediate freeze):

GENCODE version: GENCODE 19
Human Genome Build: GRCh37 / hg19
RNA-Seq Aligner: continue with current alignments, TopHat 1.4

For Option 2.1 Release (Jan/Feb 2015):

Same as above

For Option 2.2 Release (Summer 2015):

GENCODE version: Gencode 21
Human Genome Build: GRCh38 / hg20 ??
RNA-Seq Aligner: Star 2 pass

Proposal/thoughts for data versions for current and future GTEx releases

Constructing Variant IDs for all variants, starting with Option 1 Analysis Freeze

For all variants in all VCFs (arrays, WES, WGS), we will assign a constructed variant ID as follows, and will provide a conversion/look-up table to be able to map back to variants with already assigned an ID, such as RS IDs.

SNPs
Chr_Pos_REF_ALT

Indels
Chr_StartPos_REF_ALT

Structural variants (SV)
An uninformative ID, similar to the 1000 Genomes Project SV IDs:
Chr_StartPos_ArbitraryNumber