[gd44@hardac-login allenlab]$ head -100 ExAC.r0.3.1.sites.vep.vcf

##fileformat=VCFv4.2

##ALT=<ID=NON_REF,Description="Represents any possible alternative allele at this location">

##FILTER=<ID=InbreedingCoeff_Filter,Description="InbreedingCoeff <= -0.8">

##FILTER=<ID=LowQual,Description="Low quality">

##FILTER=<ID=NewCut_Filter,Description="VQSLOD > -2.632 && InbreedingCoeff >-0.8">

##FILTER=<ID=VQSRTrancheINDEL95.00to96.00,Description="Truth sensitivity tranche level for INDEL model at VQS Lod: 0.9503 <= x < 1.2168">

##FILTER=<ID=VQSRTrancheINDEL96.00to97.00,Description="Truth sensitivity tranche level for INDEL model at VQS Lod: 0.7622 <= x < 0.9503">

##FILTER=<ID=VQSRTrancheINDEL97.00to99.00,Description="Truth sensitivity tranche level for INDEL model at VQS Lod: 0.0426 <= x < 0.7622">

##FILTER=<ID=VQSRTrancheINDEL99.00to99.50,Description="Truth sensitivity tranche level for INDEL model at VQS Lod: -0.8363 <= x < 0.0426">

##FILTER=<ID=VQSRTrancheINDEL99.50to99.90,Description="Truth sensitivity tranche level for INDEL model at VQS Lod: -8.5421 <= x < -0.8363">

##FILTER=<ID=VQSRTrancheINDEL99.90to99.95,Description="Truth sensitivity tranche level for INDEL model at VQS Lod: -18.4482 <= x < -8.5421">

##FILTER=<ID=VQSRTrancheINDEL99.95to100.00+,Description="Truth sensitivity tranche level for INDEL model at VQS Lod < -37254.4742">

##FILTER=<ID=VQSRTrancheINDEL99.95to100.00,Description="Truth sensitivity tranche level for INDEL model at VQS Lod: -37254.4742 <= x < -18.4482">

##FILTER=<ID=VQSRTrancheSNP99.60to99.80,Description="Truth sensitivity tranche level for SNP model at VQS Lod: -4.9627 <= x < -1.8251">

##FILTER=<ID=VQSRTrancheSNP99.80to99.90,Description="Truth sensitivity tranche level for SNP model at VQS Lod: -31.4709 <= x < -4.9627">

##FILTER=<ID=VQSRTrancheSNP99.90to99.95,Description="Truth sensitivity tranche level for SNP model at VQS Lod: -170.3725 <= x < -31.4709">

##FILTER=<ID=VQSRTrancheSNP99.95to100.00+,Description="Truth sensitivity tranche level for SNP model at VQS Lod < -39645.8352">

##FILTER=<ID=VQSRTrancheSNP99.95to100.00,Description="Truth sensitivity tranche level for SNP model at VQS Lod: -39645.8352 <= x < -170.3725">

##FORMAT=<ID=AD,Number=.,Type=Integer,Description="Allelic depths for the ref and alt alleles in the order listed">

##FORMAT=<ID=DP,Number=1,Type=Integer,Description="Approximate read depth (reads with MQ=255 or with bad mates are filtered)">

##FORMAT=<ID=GQ,Number=1,Type=Integer,Description="Genotype Quality">

##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">

##FORMAT=<ID=MIN_DP,Number=1,Type=Integer,Description="Minimum DP observed within the GVCF block">

##FORMAT=<ID=PL,Number=G,Type=Integer,Description="Normalized, Phred-scaled likelihoods for genotypes as defined in the VCF specification">

##FORMAT=<ID=SB,Number=4,Type=Integer,Description="Per-sample component statistics which comprise the Fisher's Exact Test to detect strand bias.">

##GATKCommandLine=<ID=ApplyRecalibration,Version=3.1-163-g4284d7a,Date="Fri Jun 06 05:42:47 EDT 2014",Epoch=1402047767783,CommandLineOptions="analysis_type=ApplyRecalibration input_file=[] showFullBamList=false read_buffer_size=null phone_home=AWS gatk_key=null tag=NA read_filter=[] intervals=[/seq/dax/macarthur_joint_calling/v2/scattered/temp_0001_of_1000/scattered.intervals] excludeIntervals=null interval_set_rule=UNION interval_merging=ALL interval_padding=0 reference_sequence=/seq/references/Homo_sapiens_assembly19/v1/Homo_sapiens_assembly19.fasta nonDeterministicRandomSeed=false disableDithering=false maxRuntime=-1 maxRuntimeUnits=MINUTES downsampling_type=BY_SAMPLE downsample_to_fraction=null downsample_to_coverage=1000 baq=OFF baqGapOpenPenalty=40.0 refactor_NDN_cigar_string=false fix_misencoded_quality_scores=false allow_potentially_misencoded_quality_scores=false useOriginalQualities=false defaultBaseQualities=-1 performanceLog=null BQSR=null quantize_quals=0 disable_indel_quals=false emit_original_quals=false preserve_qscores_less_than=6 globalQScorePrior=-1.0 validation_strictness=SILENT remove_program_records=false keep_program_records=false sample_rename_mapping_file=null unsafe=null disable_auto_index_creation_and_locking_when_reading_rods=true num_threads=1 num_cpu_threads_per_data_thread=1 num_io_threads=0 monitorThreadEfficiency=false num_bam_file_handles=null read_group_black_list=null pedigree=[] pedigreeString=[] pedigreeValidationType=STRICT allow_intervals_with_unindexed_bam=false generateShadowBCF=false variant_index_type=DYNAMIC_SEEK variant_index_parameter=-1 logging_level=INFO log_to_file=null help=false version=false input=[(RodBinding name=input source=/seq/dax/macarthur_joint_calling/v2/scattered/temp_0001_of_1000/genotypes.unfiltered.vcf.gz)] recal_file=(RodBinding name=recal_file source=/seq/dax/macarthur_joint_calling/v2/macarthur_joint_calling.indels.recal) tranches_file=/seq/dax/macarthur_joint_calling/v2/macarthur_joint_calling.indels.tranches out=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub no_cmdline_in_header=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub sites_only=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub

bcf=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub ts_filter_level=95.0 lodCutoff=null ignore_filter=null excludeFiltered=false mode=INDEL filter_reads_with_N_cigar=false filter_mismatching_base_and_quals=false filter_bases_not_stored=false">

##GATKCommandLine=<ID=SelectVariants,Version=3.3-33-g58cfab1,Date="Fri Nov 21 21:05:54 EST 2014",Epoch=1416621954228,CommandLineOptions="analysis_type=SelectVariants input_file=[] showFullBamList=false read_buffer_size=null phone_home=AWS gatk_key=null tag=NA read_filter=[] intervals=null excludeIntervals=null interval_set_rule=UNION interval_merging=ALL interval_padding=0 reference_sequence=/seq/references/Homo_sapiens_assembly19/v1/Homo_sapiens_assembly19.fasta nonDeterministicRandomSeed=false disableDithering=false maxRuntime=-1 maxRuntimeUnits=MINUTES downsampling_type=BY_SAMPLE downsample_to_fraction=null downsample_to_coverage=1000 baq=OFF baqGapOpenPenalty=40.0 refactor_NDN_cigar_string=false fix_misencoded_quality_scores=false allow_potentially_misencoded_quality_scores=false useOriginalQualities=false defaultBaseQualities=-1 performanceLog=null BQSR=null quantize_quals=0 disable_indel_quals=false emit_original_quals=false preserve_qscores_less_than=6 globalQScorePrior=-1.0 validation_strictness=SILENT remove_program_records=false keep_program_records=false sample_rename_mapping_file=null unsafe=null disable_auto_index_creation_and_locking_when_reading_rods=false no_cmdline_in_header=false sites_only=false never_trim_vcf_format_field=false bcf=false bam_compression=null simplifyBAM=false disable_bam_indexing=false generate_md5=false num_threads=4 num_cpu_threads_per_data_thread=1 num_io_threads=0 monitorThreadEfficiency=false num_bam_file_handles=null read_group_black_list=null pedigree=[] pedigreeString=[] pedigreeValidationType=STRICT allow_intervals_with_unindexed_bam=false generateShadowBCF=false variant_index_type=DYNAMIC_SEEK variant_index_parameter=-1 logging_level=INFO log_to_file=null help=false version=false variant=(RodBinding name=variant source=/humgen/gsa-firehose/ExAC_GATKV2.5/MacArthur_HC/v2/fullset/ExAC_HC.chrX.00.vcf.gz) discordance=(RodBinding name= source=UNBOUND) concordance=(RodBinding name= source=UNBOUND) out=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub sample_name=[] sample_expressions=null sample_file=[exac_panel.samples] exclude_sample_name=[] exclude_sample_file=[] select_expressions=[] excludeNonVariants=true excludeFiltered=false preserveAlleles=false restrictAllelesTo=ALL keepOriginalAC=false mendelianViolation=false mendelianViolationQualThreshold=0.0 select_random_fraction=0.0 remove_fraction_genotypes=0.0 selectTypeToInclude=[] keepIDs=null fullyDecode=false forceGenotypesDecode=false justRead=false maxIndelSize=2147483647 ALLOW_NONOVERLAPPING_COMMAND_LINE_SAMPLES=false filter_reads_with_N_cigar=false filter_mismatching_base_and_quals=false filter_bases_not_stored=false">

##GATKCommandLine=<ID=VariantAnnotator,Version=3.2-2-gec30cee,Date="Mon Nov 24 21:39:16 EST 2014",Epoch=1416883156611,CommandLineOptions="analysis_type=VariantAnnotator input_file=[] showFullBamList=false read_buffer_size=null phone_home=AWS gatk_key=null tag=NA read_filter=[] intervals=[X:1-7763528] excludeIntervals=null interval_set_rule=UNION interval_merging=ALL interval_padding=0 reference_sequence=/seq/references/Homo_sapiens_assembly19/v1/Homo_sapiens_assembly19.fasta nonDeterministicRandomSeed=false disableDithering=false maxRuntime=-1 maxRuntimeUnits=MINUTES downsampling_type=BY_SAMPLE downsample_to_fraction=null

downsample_to_coverage=250 baq=OFF baqGapOpenPenalty=40.0 refactor_NDN_cigar_string=false fix_misencoded_quality_scores=false allow_potentially_misencoded_quality_scores=false useOriginalQualities=false defaultBaseQualities=-1 performanceLog=null BQSR=null quantize_quals=0 disable_indel_quals=false emit_original_quals=false preserve_qscores_less_than=6 globalQScorePrior=-1.0 validation_strictness=SILENT remove_program_records=false keep_program_records=false sample_rename_mapping_file=null unsafe=null disable_auto_index_creation_and_locking_when_reading_rods=false num_threads=1 num_cpu_threads_per_data_thread=1 num_io_threads=0 monitorThreadEfficiency=false num_bam_file_handles=null read_group_black_list=null pedigree=[] pedigreeString=[] pedigreeValidationType=STRICT allow_intervals_with_unindexed_bam=false generateShadowBCF=false variant_index_type=DYNAMIC_SEEK variant_index_parameter=-1 logging_level=INFO log_to_file=null help=false version=false variant=(RodBinding name=variant source=../filtered/ExAC_HC.chrX.00.filtered.vcf.gz) snpEffFile=(RodBinding name= source=UNBOUND) dbsnp=(RodBinding name= source=UNBOUND) comp=[] resource=[] out=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub no_cmdline_in_header=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub sites_only=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub bcf=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub annotation=[SexAlleleCounts] excludeAnnotation=[] group=[] expression=[] useAllAnnotations=false list=false alwaysAppendDbsnpId=false MendelViolationGenotypeQualityThreshold=0.0 sampleMappingFile=samples_pop_sex.tsv filter_reads_with_N_cigar=false filter_mismatching_base_and_quals=false filter_bases_not_stored=false">

##GATKCommandLine=<ID=VariantFiltration,Version=3.2-2-gec30cee,Date="Sun Nov 23 08:50:16 EST 2014",Epoch=1416750616237,CommandLineOptions="analysis_type=VariantFiltration input_file=[] showFullBamList=false read_buffer_size=null phone_home=AWS gatk_key=null tag=NA read_filter=[] intervals=[X:1-7763528] excludeIntervals=null interval_set_rule=UNION interval_merging=ALL interval_padding=0 reference_sequence=/seq/references/Homo_sapiens_assembly19/v1/Homo_sapiens_assembly19.fasta nonDeterministicRandomSeed=false disableDithering=false maxRuntime=-1 maxRuntimeUnits=MINUTES downsampling_type=BY_SAMPLE downsample_to_fraction=null downsample_to_coverage=1000 baq=OFF baqGapOpenPenalty=40.0 refactor_NDN_cigar_string=false fix_misencoded_quality_scores=false allow_potentially_misencoded_quality_scores=false useOriginalQualities=false defaultBaseQualities=-1 performanceLog=null BQSR=null quantize_quals=0 disable_indel_quals=false emit_original_quals=false preserve_qscores_less_than=6 globalQScorePrior=-1.0 validation_strictness=SILENT remove_program_records=false keep_program_records=false sample_rename_mapping_file=null unsafe=null disable_auto_index_creation_and_locking_when_reading_rods=false num_threads=1 num_cpu_threads_per_data_thread=1 num_io_threads=0 monitorThreadEfficiency=false num_bam_file_handles=null read_group_black_list=null pedigree=[] pedigreeString=[] pedigreeValidationType=STRICT allow_intervals_with_unindexed_bam=false generateShadowBCF=false variant_index_type=DYNAMIC_SEEK variant_index_parameter=-1 logging_level=INFO log_to_file=null help=false version=false variant=(RodBinding name=variant source=../fullset/ExAC_HC.chrX.00.vcf.gz) mask=(RodBinding name= source=UNBOUND) out=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub

no_cmdline_in_header=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub sites_only=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub bcf=org.broadinstitute.gatk.engine.io.stubs.VariantContextWriterStub filterExpression=[VQSLOD > -2.632 && InbreedingCoeff >-0.8, InbreedingCoeff <= -0.8] filterName=[NewCut_Filter, InbreedingCoeff_Filter] genotypeFilterExpression=[] genotypeFilterName=[] clusterSize=3 clusterWindowSize=0 maskExtension=0 maskName=Mask filterNotInMask=false missingValuesInExpressionsShouldEvaluateAsFailing=false invalidatePreviousFilters=false filter_reads_with_N_cigar=false filter_mismatching_base_and_quals=false filter_bases_not_stored=false">

##GVCFBlock=minGQ=0(inclusive),maxGQ=5(exclusive)

##INFO=<ID=AC,Number=A,Type=Integer,Description="Allele count in genotypes, for each ALT allele, in the same order as listed">

##INFO=<ID=AC_AFR,Number=A,Type=Integer,Description="African/African American Allele Counts">

##INFO=<ID=AC_AMR,Number=A,Type=Integer,Description="American Allele Counts">

##INFO=<ID=AC_Adj,Number=A,Type=Integer,Description="Adjusted Allele Counts">

##INFO=<ID=AC_EAS,Number=A,Type=Integer,Description="East Asian Allele Counts">

##INFO=<ID=AC_FIN,Number=A,Type=Integer,Description="Finnish Allele Counts">

##INFO=<ID=AC_Hemi,Number=A,Type=Integer,Description="Adjusted Hemizygous Counts">

##INFO=<ID=AC_Het,Number=A,Type=Integer,Description="Adjusted Heterozygous Counts">

##INFO=<ID=AC_Hom,Number=A,Type=Integer,Description="Adjusted Homozygous Counts">

##INFO=<ID=AC_NFE,Number=A,Type=Integer,Description="Non-Finnish European Allele Counts">

##INFO=<ID=AC_OTH,Number=A,Type=Integer,Description="Other Allele Counts">

##INFO=<ID=AC_SAS,Number=A,Type=Integer,Description="South Asian Allele Counts">

##INFO=<ID=AF,Number=A,Type=Float,Description="Allele Frequency, for each ALT allele, in the same order as listed">

##INFO=<ID=AN,Number=1,Type=Integer,Description="Total number of alleles in called genotypes">

##INFO=<ID=AN_AFR,Number=1,Type=Integer,Description="African/African American Chromosome Count">

##INFO=<ID=AN_AMR,Number=1,Type=Integer,Description="American Chromosome Count">

##INFO=<ID=AN_Adj,Number=1,Type=Integer,Description="Adjusted Chromosome Count">

##INFO=<ID=AN_EAS,Number=1,Type=Integer,Description="East Asian Chromosome Count">

##INFO=<ID=AN_FIN,Number=1,Type=Integer,Description="Finnish Chromosome Count">

##INFO=<ID=AN_NFE,Number=1,Type=Integer,Description="Non-Finnish European Chromosome Count">

##INFO=<ID=AN_OTH,Number=1,Type=Integer,Description="Other Chromosome Count">

##INFO=<ID=AN_SAS,Number=1,Type=Integer,Description="South Asian Chromosome Count">

##INFO=<ID=BaseQRankSum,Number=1,Type=Float,Description="Z-score from Wilcoxon rank sum test of Alt Vs. Ref base qualities">

##INFO=<ID=CCC,Number=1,Type=Integer,Description="Number of called chromosomes">

##INFO=<ID=ClippingRankSum,Number=1,Type=Float,Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref number of hard clipped bases">

##INFO=<ID=DB,Number=0,Type=Flag,Description="dbSNP Membership">

##INFO=<ID=DP,Number=1,Type=Integer,Description="Approximate read depth; some reads may have been filtered">

##INFO=<ID=DS,Number=0,Type=Flag,Description="Were any of the samples downsampled?">

##INFO=<ID=END,Number=1,Type=Integer,Description="Stop position of the interval">

##INFO=<ID=FS,Number=1,Type=Float,Description="Phred-scaled p-value using Fisher's exact test to detect strand bias">

##INFO=<ID=GQ_MEAN,Number=1,Type=Float,Description="Mean of all GQ values">

##INFO=<ID=GQ_STDDEV,Number=1,Type=Float,Description="Standard deviation of all GQ values">

##INFO=<ID=HWP,Number=1,Type=Float,Description="P value from test of Hardy Weinberg Equilibrium">

##INFO=<ID=HaplotypeScore,Number=1,Type=Float,Description="Consistency of the site with at most two segregating haplotypes">

##INFO=<ID=Hemi_AFR,Number=A,Type=Integer,Description="African/African American Hemizygous Counts">

##INFO=<ID=Hemi_AMR,Number=A,Type=Integer,Description="American Hemizygous Counts">

##INFO=<ID=Hemi_EAS,Number=A,Type=Integer,Description="East Asian Hemizygous Counts">

##INFO=<ID=Hemi_FIN,Number=A,Type=Integer,Description="Finnish Hemizygous Counts">

##INFO=<ID=Hemi_NFE,Number=A,Type=Integer,Description="Non-Finnish European Hemizygous Counts">

##INFO=<ID=Hemi_OTH,Number=A,Type=Integer,Description="Other Hemizygous Counts">

##INFO=<ID=Hemi_SAS,Number=A,Type=Integer,Description="South Asian Hemizygous Counts">

##INFO=<ID=Het_AFR,Number=.,Type=Integer,Description="African/African American Heterozygous Counts">

##INFO=<ID=Het_AMR,Number=.,Type=Integer,Description="American Heterozygous Counts">

##INFO=<ID=Het_EAS,Number=.,Type=Integer,Description="East Asian Heterozygous Counts">

##INFO=<ID=Het_FIN,Number=.,Type=Integer,Description="Finnish Heterozygous Counts">

##INFO=<ID=Het_NFE,Number=.,Type=Integer,Description="Non-Finnish European Heterozygous Counts">

##INFO=<ID=Het_OTH,Number=.,Type=Integer,Description="Other Heterozygous Counts">

##INFO=<ID=Het_SAS,Number=.,Type=Integer,Description="South Asian Heterozygous Counts">

##INFO=<ID=Hom_AFR,Number=A,Type=Integer,Description="African/African American Homozygous Counts">

##INFO=<ID=Hom_AMR,Number=A,Type=Integer,Description="American Homozygous Counts">

##INFO=<ID=Hom_EAS,Number=A,Type=Integer,Description="East Asian Homozygous Counts">

##INFO=<ID=Hom_FIN,Number=A,Type=Integer,Description="Finnish Homozygous Counts">

##INFO=<ID=Hom_NFE,Number=A,Type=Integer,Description="Non-Finnish European Homozygous Counts">

##INFO=<ID=Hom_OTH,Number=A,Type=Integer,Description="Other Homozygous Counts">

##INFO=<ID=Hom_SAS,Number=A,Type=Integer,Description="South Asian Homozygous Counts">

##INFO=<ID=InbreedingCoeff,Number=1,Type=Float,Description="Inbreeding coefficient as estimated from the genotype likelihoods per-sample when compared against the Hardy-Weinberg expectation">

##INFO=<ID=MLEAC,Number=A,Type=Integer,Description="Maximum likelihood expectation (MLE) for the allele counts (not necessarily the same as the AC), for each ALT allele, in the same order as listed">

##INFO=<ID=MLEAF,Number=A,Type=Float,Description="Maximum likelihood expectation (MLE) for the allele frequency (not necessarily the same as the AF), for each ALT allele, in the same order as listed">

##INFO=<ID=MQ,Number=1,Type=Float,Description="RMS Mapping Quality">

##INFO=<ID=MQ0,Number=1,Type=Integer,Description="Total Mapping Quality Zero Reads">

##INFO=<ID=MQRankSum,Number=1,Type=Float,Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref read mapping qualities">

##INFO=<ID=NCC,Number=1,Type=Integer,Description="Number of no-called samples">

##INFO=<ID=NEGATIVE_TRAIN_SITE,Number=0,Type=Flag,Description="This variant was used to build the negative training set of bad variants">

##INFO=<ID=POSITIVE_TRAIN_SITE,Number=0,Type=Flag,Description="This variant was used to build the positive training set of good variants">

##INFO=<ID=QD,Number=1,Type=Float,Description="Variant Confidence/Quality by Depth">

##INFO=<ID=ReadPosRankSum,Number=1,Type=Float,Description="Z-score from Wilcoxon rank sum test of Alt vs. Ref read position bias">

##INFO=<ID=VQSLOD,Number=1,Type=Float,Description="Log odds ratio of being a true variant versus being false under the trained gaussian mixture model">

##INFO=<ID=culprit,Number=1,Type=String,Description="The annotation which was the worst performing in the Gaussian mixture model, likely the reason why the variant was filtered out">

##contig=<ID=1,length=249250621>

##contig=<ID=2,length=243199373>

##contig=<ID=3,length=198022430>

##contig=<ID=4,length=191154276>

##contig=<ID=5,length=180915260>

##contig=<ID=6,length=171115067>

##contig=<ID=7,length=159138663>

##contig=<ID=8,length=146364022>

##contig=<ID=9,length=141213431>

##contig=<ID=10,length=135534747>

##contig=<ID=11,length=135006516>

##contig=<ID=12,length=133851895>

##contig=<ID=13,length=115169878>

##contig=<ID=14,length=107349540>

##contig=<ID=15,length=102531392>

##contig=<ID=16,length=90354753>

##contig=<ID=17,length=81195210>

##contig=<ID=18,length=78077248>

##contig=<ID=19,length=59128983>

##contig=<ID=20,length=63025520>

##contig=<ID=21,length=48129895>

##contig=<ID=22,length=51304566>

##contig=<ID=X,length=155270560>

##contig=<ID=Y,length=59373566>

##contig=<ID=MT,length=16569>

##contig=<ID=GL000207.1,length=4262>

##contig=<ID=GL000226.1,length=15008>

##contig=<ID=GL000229.1,length=19913>

##contig=<ID=GL000231.1,length=27386>

##contig=<ID=GL000210.1,length=27682>

##contig=<ID=GL000239.1,length=33824>

##contig=<ID=GL000235.1,length=34474>

##contig=<ID=GL000201.1,length=36148>

##contig=<ID=GL000247.1,length=36422>

##contig=<ID=GL000245.1,length=36651>

##contig=<ID=GL000197.1,length=37175>

##contig=<ID=GL000203.1,length=37498>

##contig=<ID=GL000246.1,length=38154>

##contig=<ID=GL000249.1,length=38502>

##contig=<ID=GL000196.1,length=38914>

##contig=<ID=GL000248.1,length=39786>

##contig=<ID=GL000244.1,length=39929>

##contig=<ID=GL000238.1,length=39939>

##contig=<ID=GL000202.1,length=40103>

##contig=<ID=GL000234.1,length=40531>

##contig=<ID=GL000232.1,length=40652>

##contig=<ID=GL000206.1,length=41001>

##contig=<ID=GL000240.1,length=41933>

##contig=<ID=GL000236.1,length=41934>

##contig=<ID=GL000241.1,length=42152>

##contig=<ID=GL000243.1,length=43341>

##contig=<ID=GL000242.1,length=43523>

##contig=<ID=GL000230.1,length=43691>

##contig=<ID=GL000237.1,length=45867>

##contig=<ID=GL000233.1,length=45941>

##contig=<ID=GL000204.1,length=81310>

##contig=<ID=GL000198.1,length=90085>

##contig=<ID=GL000208.1,length=92689>

##contig=<ID=GL000191.1,length=106433>

##contig=<ID=GL000227.1,length=128374>

##contig=<ID=GL000228.1,length=129120>

##contig=<ID=GL000214.1,length=137718>

##contig=<ID=GL000221.1,length=155397>

##contig=<ID=GL000209.1,length=159169>

##contig=<ID=GL000218.1,length=161147>

##contig=<ID=GL000220.1,length=161802>

##contig=<ID=GL000213.1,length=164239>

##contig=<ID=GL000211.1,length=166566>

##contig=<ID=GL000199.1,length=169874>

##contig=<ID=GL000217.1,length=172149>

##contig=<ID=GL000216.1,length=172294>

##contig=<ID=GL000215.1,length=172545>

##contig=<ID=GL000205.1,length=174588>

##contig=<ID=GL000219.1,length=179198>

##contig=<ID=GL000224.1,length=179693>

##contig=<ID=GL000223.1,length=180455>

##contig=<ID=GL000195.1,length=182896>

##contig=<ID=GL000212.1,length=186858>

##contig=<ID=GL000222.1,length=186861>

##contig=<ID=GL000200.1,length=187035>

##contig=<ID=GL000193.1,length=189789>

##contig=<ID=GL000194.1,length=191469>

##contig=<ID=GL000225.1,length=211173>

##contig=<ID=GL000192.1,length=547496>

##contig=<ID=NC_007605,length=171823>

##reference=file:///seq/references/Homo_sapiens_assembly19/v1/Homo_sapiens_assembly19.fasta

##source=SelectVariants

##INFO=<ID=DP_HIST,Number=R,Type=String,Description="Histogram for DP; Mids: 2.5|7.5|12.5|17.5|22.5|27.5|32.5|37.5|42.5|47.5|52.5|57.5|62.5|67.5|72.5|77.5|82.5|87.5|92.5|97.5">

##INFO=<ID=GQ_HIST,Number=R,Type=String,Description="Histogram for GQ; Mids: 2.5|7.5|12.5|17.5|22.5|27.5|32.5|37.5|42.5|47.5|52.5|57.5|62.5|67.5|72.5|77.5|82.5|87.5|92.5|97.5">

##INFO=<ID=DOUBLETON_DIST,Number=A,Type=String,Description="Euclidean distance of carriers of doubletons">

##INFO=<ID=AC_MALE,Number=A,Type=String,Description="Allele count among males">

##INFO=<ID=AC_FEMALE,Number=A,Type=String,Description="Allele count among females">

##INFO=<ID=AN_MALE,Number=1,Type=String,Description="Allele number among males">

##INFO=<ID=AN_FEMALE,Number=1,Type=String,Description="Allele number among females">

##INFO=<ID=AC_CONSANGUINEOUS,Number=A,Type=String,Description="Allele count among individuals with F > 0.05">

##INFO=<ID=AN_CONSANGUINEOUS,Number=1,Type=String,Description="Allele number among individuals with F > 0.05">

##INFO=<ID=Hom_CONSANGUINEOUS,Number=A,Type=String,Description="Homozygote count among individuals with F > 0.05">

##VEP=v81 cache=/humgen/atgu1/fs03/konradk/vep//homo_sapiens/81_GRCh37 db=. polyphen=2.2.2 sift=sift5.2.2 COSMIC=71 ESP=20141103 gencode=GENCODE 19 HGMD-PUBLIC=20144 genebuild=2011-04 regbuild=13 assembly=GRCh37.p13 dbSNP=142 ClinVar=201501

##LoF_info=Info used for LoF annotation

##LoF_flags=Possible warning flags for LoF

##LoF_filter=Reason for LoF not being HC

##LoF=Loss-of-function annotation (HC = High Confidence; LC = Low Confidence)

[gd44@hardac-login allenlab]$ sed -n '200,220p' ExAC.r0.3.1.sites.vep.vcf

##LoF=Loss-of-function annotation (HC = High Confidence; LC = Low Confidence)

##context=1 base context around the variant

##ancestral=ancestral allele

##INFO=<ID=CSQ,Number=.,Type=String,Description="Consequence annotations from Ensembl VEP. Format:
Allele|Consequence|IMPACT|SYMBOL|Gene|Feature_type|Feature|BIOTYPE|EXON|INTRON|HGVSc|HGVSp|cDNA_position|CDS_position|Protein_position|Amino_acids|Codons|Existing_variation|ALLELE_NUM|DISTANCE|STRAND|VARIANT_CLASS|MINIMISED|SYMBOL_SOURCE|HGNC_ID|CANONICAL|TSL|CCDS|ENSP|SWISSPROT|TREMBL|UNIPARC|SIFT|PolyPhen|DOMAINS|HGVS_OFFSET|GMAF|AFR_MAF|AMR_MAF|ASN_MAF|EAS_MAF|EUR_MAF|SAS_MAF|AA_MAF|EA_MAF|CLIN_SIG|SOMATIC|PHENO|PUBMED|MOTIF_NAME|MOTIF_POS|HIGH_INF_POS|MOTIF_SCORE_CHANGE|LoF_info|LoF_flags|LoF_filter|LoF|context|ancestral">

##INFO=<ID=AC_POPMAX,Number=A,Type=String,Description="AC in the population with the max AF">

##INFO=<ID=AN_POPMAX,Number=A,Type=String,Description="AN in the population with the max AF">

##INFO=<ID=POPMAX,Number=A,Type=String,Description="Population with max AF">

##INFO=<ID=clinvar_measureset_id,Number=A,Type=String,Description="Clinvar Measureset ID">

##INFO=<ID=clinvar_conflicted,Number=A,Type=String,Description="Clinvar Conflicted Status">

##INFO=<ID=clinvar_pathogenic,Number=A,Type=String,Description="Clinvar Pathogenic Status">

##INFO=<ID=clinvar_mut,Number=A,Type=String,Description="Clinvar MUT Flag (is disease allele REF?)">

##INFO=<ID=K1_RUN,Number=A,Type=String,Description="Number of ensuing single nucleotide repeats.">

##INFO=<ID=K2_RUN,Number=A,Type=String,Description="Number of ensuing di-nucleotide repeats.">

##INFO=<ID=K3_RUN,Number=A,Type=String,Description="Number of ensuing tri-nucleotide repeats.">

##INFO=<ID=ESP_AF_POPMAX,Number=A,Type=String,Description="Max allele frequency across populations in ESP">

##INFO=<ID=ESP_AF_GLOBAL,Number=A,Type=String,Description="Overall allele frequency in ESP">

##INFO=<ID=ESP_AC,Number=A,Type=String,Description="Allele Count in ESP">

##INFO=<ID=KG_AF_POPMAX,Number=A,Type=String,Description="Max allele frequency across populations in 1000 Genomes">

##INFO=<ID=KG_AF_GLOBAL,Number=A,Type=String,Description="Overall allele frequency in 1000 Genomes">

##INFO=<ID=KG_AC,Number=A,Type=String,Description="Allele Count in 1000 Genomes">

#CHROM POS ID REF ALT QUAL FILTER INFO

#CHROM POS ID REF ALT QUAL FILTER INFO

1 13372 . G C 608.91 PASS
AC=3;AC_AFR=0;AC_AMR=0;AC_Adj=2;AC_EAS=0;AC_FIN=0;AC_H
et=0;AC_Hom=1;AC_NFE=0;AC_OTH=0;AC_SAS=2;AF=6.998e-
05;AN=42870;AN_AFR=770;AN_AMR=134;AN_Adj=8432;AN_EAS=254;AN
_FIN=16;AN_NFE=2116;AN_OTH=90;AN_SAS=5052;BaseQRankSum=0.727;ClippingRankSum=1.15;DP=1
39843;FS=0.000;GQ_MEAN=1
2.48;GQ_STDDEV=15.18;Het_AFR=0;Het_AMR=0;Het_EAS=0;Het_FIN=0;Het_NFE=0;Het_OTH=0;Het_SA
S=0;Hom_AFR=0;Hom_AMR=0 ;Hom_EAS=0;Hom_FIN=0;Hom_NFE=0;Hom_OTH=0;Hom_SAS=1;
InbreedingCoeff=-0.0844;MQ=35.72;MQ0=0;MQRankSum=0.727;NCC=
60853;QD=23.42;ReadPosRankSum=0.727;VQSLOD=-
1.687e+00;culprit=MQ;DP_HIST=14728|2455|2120|518|121|499|534|314|1
11|21|10|2|2|0|0|0|0|0|0|0,1|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0;GQ_HIST=1012|14971|1
72|100|3161|259|127|30|
8|9|5|16|1162|274|59|45|17|2|3|3,0|0|0|0|0|1|0|0|0|0|0|0|0|1|0|0|0|0|0|0;DOUBLETON_DIS
T=.;AC_MALE=2;AC_FEMALE=
0;AN_MALE=5518;AN_FEMALE=2914;AC_CONSANGUINEOUS=0;AN_CONSANGUINEOUS=926;Hom_CON
SANGUINEOUS=0;CSQ=C|non_coding_
transcript_exon_variant&non_coding_transcript_variant|MODIFIER|DDX11L1|ENSG00000223972|Tran
script|ENST00000456
328|processed_transcript|3/3||ENST00000456328.2:n.620G>C||620||||||1||1|SNV|1|HGNC|37102
|YES||||||||||||||||||
|||||||||||||GGA|.,C|downstream_gene_variant|MODIFIER|WASH7P|ENSG00000227232|Transcrip
t|ENST00000488147|unpro cessed_pseudogene||||||||||1|1032|-
1|SNV|1|HGNC|38034|||||||||||||||||||||||||||||||||GGA|.,C|downstream_gene
_variant|MODIFIER|WASH7P|ENSG00000227232|Transcript|ENST00000541675|unprocessed_pseudog
ene||||||||||1|991|-1|
SNV|1|HGNC|38034|||||||||||||||||||||||||||||||||GGA|.,C|splice_region_variant&non_coding_
transcript_exon_vari
ant&non_coding_transcript_variant|LOW|DDX11L1|ENSG00000223972|Transcript|ENST00000450305
|transcribed_unprocess
ed_pseudogene|5/6||ENST00000450305.2:n.412G>C||412||||||1||1|SNV|1|HGNC|37102|||||||||

||||||||||||||||||||||||
GGA|.,C|non_coding_transcript_exon_variant&non_coding_transcript_variant|MODIFIER|DDX11L1|EN
SG00000223972|Tran
script|ENST00000515242|transcribed_unprocessed_pseudogene|3/3||ENST00000515242.2:n.613G>C|
|613||||||1||1|SNV|1
|HGNC|37102|||||||||||||||||||||||||||||||||GGA|.,C|downstream_gene_variant|MODIFIER|WA
SH7P|ENSG00000227232|Tr
anscript|ENST00000538476|unprocessed_pseudogene||||||||||1|1039|-
1|SNV|1|HGNC|38034|||||||||||||||||||||||||
|||||||GGA|.,C|intron_variant&non_coding_transcript_variant|MODIFIER|DDX11L1|ENSG000002239
72|Transcript|ENST00
000518655|transcribed_unprocessed_pseudogene||2/3|ENST00000518655.2:n.482-
31G>C|||||||1||1|SNV|1|HGNC|37102||
|||||||||||||||||||||||||||||GGA|.,C|downstream_gene_variant|MODIFIER|WASH7P|ENSG0000
0227232|Transcript|ENST 00000438504|unprocessed_pseudogene||||||||||1|991|-
1|SNV|1|HGNC|38034|YES|||||||||||||||||||||||||||||||GGA| .,C|downstream_gene_va
riant|MODIFIER|WASH7P|ENSG00000227232|Transcript|ENST00000423562|unprocessed_pseudogen
e| |||||||||1|991|-
1|SNV|1|HGNC|38034|||||||||||||||||||||||||||||||GGA|.,C|regulatory_region_variant|MODIF
IER
|||RegulatoryFeature|ENSR00001576075|CTCF_binding_site||||||||||1|||SNV|1|||||||||||||||
|||||||||||||||||G
GA|.;AC_POPMAX=2;AN_POPMAX=5052;POPMAX=SAS;K1_RUN=G:0;K2_RUN=GA:0;K3_RUN=GAA:0;ES
P_AF_POPMAX=0;ESP_AF_GLOBAL=0 ;ESP_AC=0;KG_AF_POPMAX=0;KG_AF_GLOBAL=0;KG_AC=
0