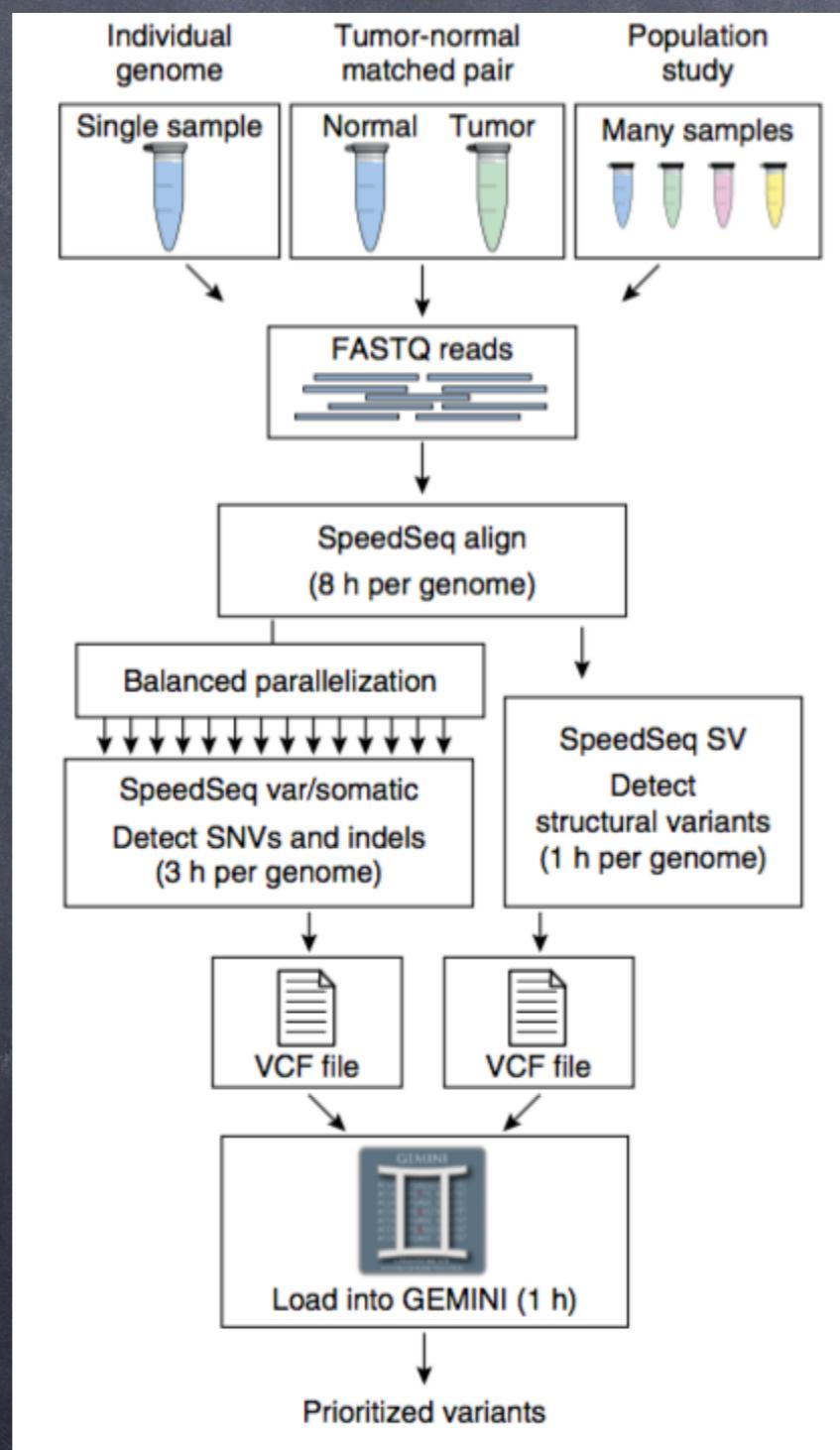


Pooled and Normal Resequencing data analysis

2016.8.23
Linhua Sun

SpeedSeq workflow

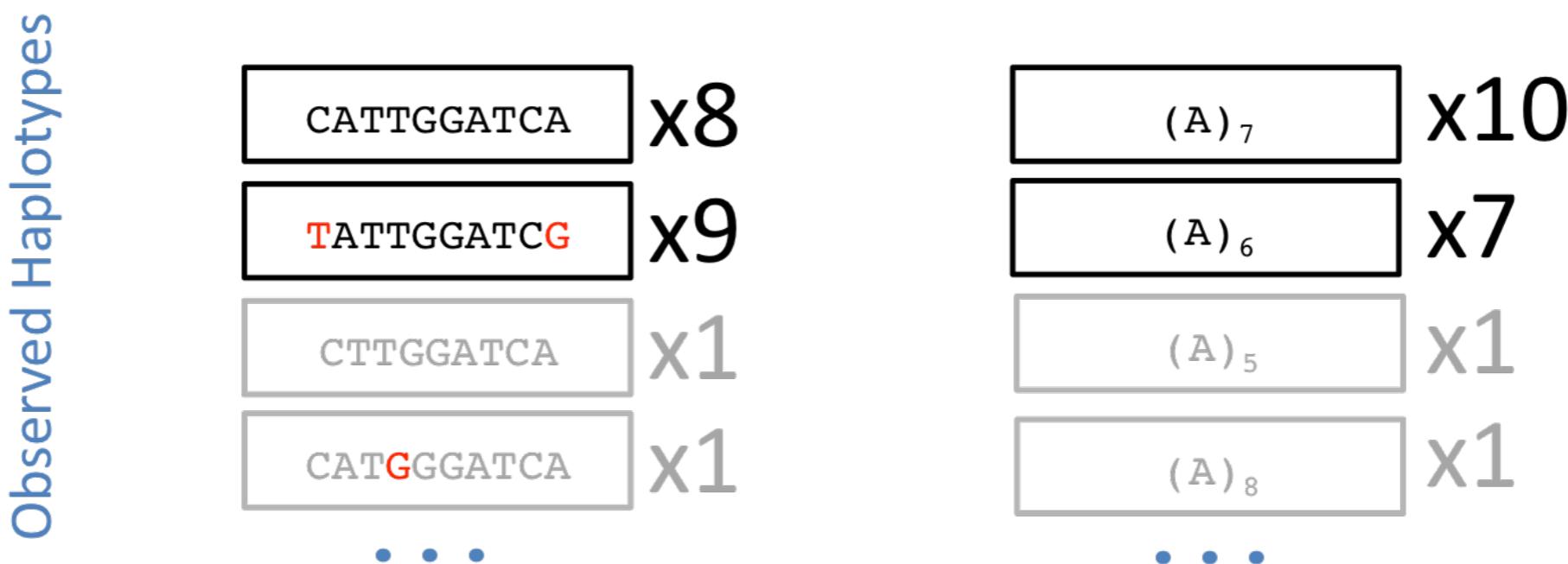


Chiang, C. et al. SpeedSeq: ultra-fast personal genome analysis and interpretation. Nat Meth 12, 966–968 (2015).

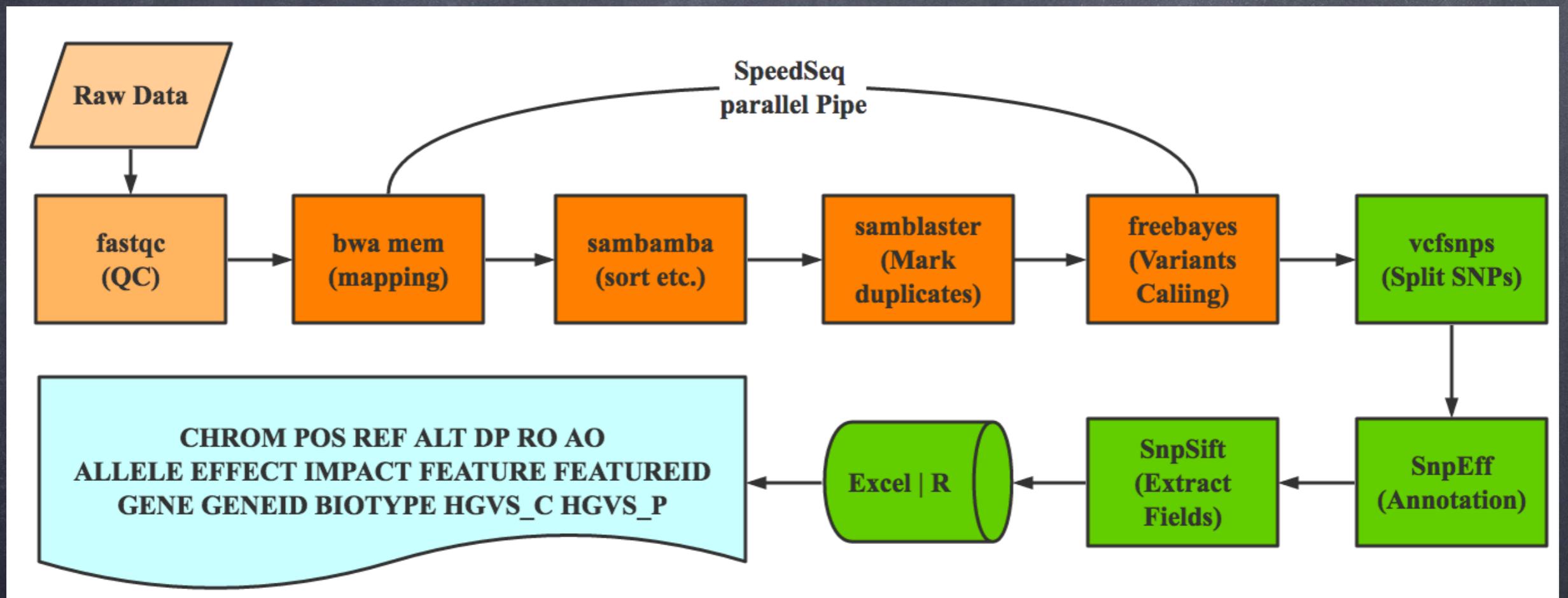
FASTQ alignment and BAM processing

- ① **SpeedSeq** aligns paired-end FASTQ files to the reference genome with **BWA-MEM** 0.7.8 using the “-M” flag to mark shorter alignments as secondary.
- ② Aligned reads are streamed directly into **SAMBLASTER**, which seizes idle CPU cycles that are periodically liberated each time BWA reads a FASTQ data chunk into the buffer.
- ③ Marking duplicates on the presorted BAM file allows simultaneous extraction of discordant read pairs and split-read alignments, followed by rapid sorting and BAM compression with **Sambamba**.

		Variant Region		Variant Region
Reads	Ref			
TACCGAT		CATTGGATCA	CGATTCC...GCATTGC	AAAAAAA-
TACCGAT		CATTGGATCA	CGATTCC...GCATTGC	-AAAAAA-
ACCGAT		TATTGCCATCG	CGATTCC...GCATTGC	-AAAAAA-
ACCGAT		CATTGGATCA	CGATTCC...GCATTGC	AAAAAA-A
ACCGAT		TATTGGATCG	CGATTCC...GCATTGC	-AAAAAAA
CCGAT		C-TTGGATCA	CGATTCC...GCATTGC	AAAAAAA-
CCGAT		CATGGGATCA	CGATTCC...GCATTGC	AAAAAAA A
• • •		• • •	• • •	• • •



New Resequencing Analysis Pipeline



① Pipe

② Parallel

③ New C++ Softwares

④ Better Interface

Abbreviation	Detailed explanations
CHROM	Chromosome.
POS	The reference position, with the 1st base having position 1.
REF	Reference base(s).
ALT	Alternate base(s).
DP	Total read depth at the locus.
RO	Reference allele observation count, with partial observations recorded fractionally.
AO	Alternate allele observations, with partial observations recorded fractionally.
ALLEL	Allele (or ALT): In case of multiple ALT fields, this helps to identify which ALT we are referring to.
EFFECT	Annotation (a.k.a. effect): Annotated using Sequence Ontology terms. Multiple effects can be concatenated using '&'.
IMPACT	Putative_impact: A simple estimation of putative impact / deleteriousness : {HIGH, MODERATE, LOW, MODIFIER}.
FEATURE	Which type of feature is in the next field (e.g. transcript, motif, miRNA, etc.).
FEATUREID	Depending on the annotation, this may be: Transcript ID (preferably using version number), Motif ID, miRNA, ChipSeq peak, Histone mark, etc.
GENE	Common gene name (HGNC). Optional: use closest gene when the variant is "intergenic".
GENEID	Gene ID.
BIOTYPE	Transcript biotype: The bare minimum is at least a description on whether the transcript is {Coding, Noncoding}. Whenever possible, use ENSEMBL biotypes.
HGVS_C	Variant using HGVS notation (DNA level).
HGVS_P	If variant is coding, this field describes the variant using HGVS notation (Protein level). Since transcript ID is already mentioned in feature ID, it may be omitted here.

Refer to these weblinks:

http://snpeff.sourceforge.net/SnpEff_manual.html#input

<https://samtools.github.io/hts-specs/VCFv4.2.pdf>

```

REFGENOME="/sdd1/users/linhua/QIAN_LAB/ATH_SHORTSTAR/Athaliana_167.fa"
ID=$(basename $1 _1.fq.gz)
SnpSift="/sdd1/users/linhua/QIAN_LAB/PROJECT-2016/resequencing_data_process/snpEff/SnpSift.jar"
snpEff="/sdd1/users/linhua/QIAN_LAB/PROJECT-2016/resequencing_data_process/snpEff/snpEff.jar"

# 1.align
speedseq align -t 10 \
-o ${ID}_speedseq_aln_out \
-R "@RG\tID:${ID}\tSM:${ID}\tLB:ILLUMINA" \
$REFGENOME \
${ID}_1.fq.gz \
${ID}_2.fq.gz

tput setaf 3; echo -n "|||||| Align is finished! |||||"

# 2.variants calling
freebayes -f $REFGENOME -F 0.01 -C 1 --pooled-continuous ${ID}_speedseq_aln_out.bam |bgzip > ${ID}_freq.vcf.gz
tabix -p vcf ${ID}_freq.vcf.gz

tput setaf 3; echo -n "|||||| calling is finished! |||||"

# 3.extract useful info
zcat ${ID}_freq.vcf.gz|vcfsnps|vcfbiallelic|vcfintersect -b extend_3K_methylation_genes_ranges_results.bed |bgzip > biallelic_snps_intervaled_${ID}_freq.vcf.gz
tabix -p vcf biallelic_snps_intervaled_${ID}_freq.vcf.gz

tput setaf 3; echo -n "|||||| extract is finished! |||||"

# 4.annotation
java -jar ${snpEff} TAIR10.29 biallelic_snps_intervaled_${ID}_freq.vcf.gz | bgzip > annotated_biallelic_snps_intervaled_${ID}_freq.vcf.gz
tabix -p vcf annotated_biallelic_snps_intervaled_${ID}_freq.vcf.gz

tput setaf 3; echo -n "|||||| annotation is finished! |||||"

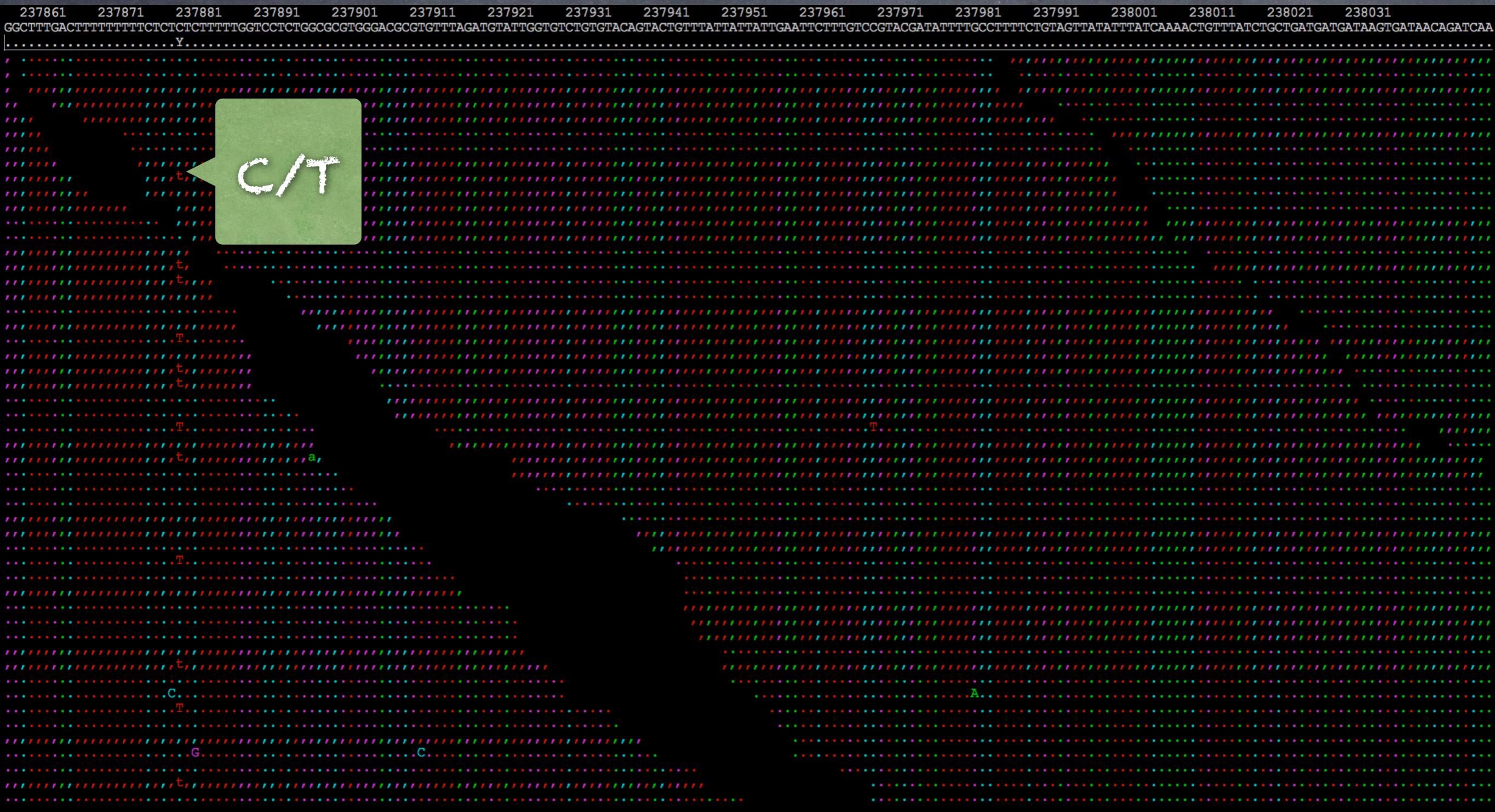
# 5.extract nice fields from VCF file

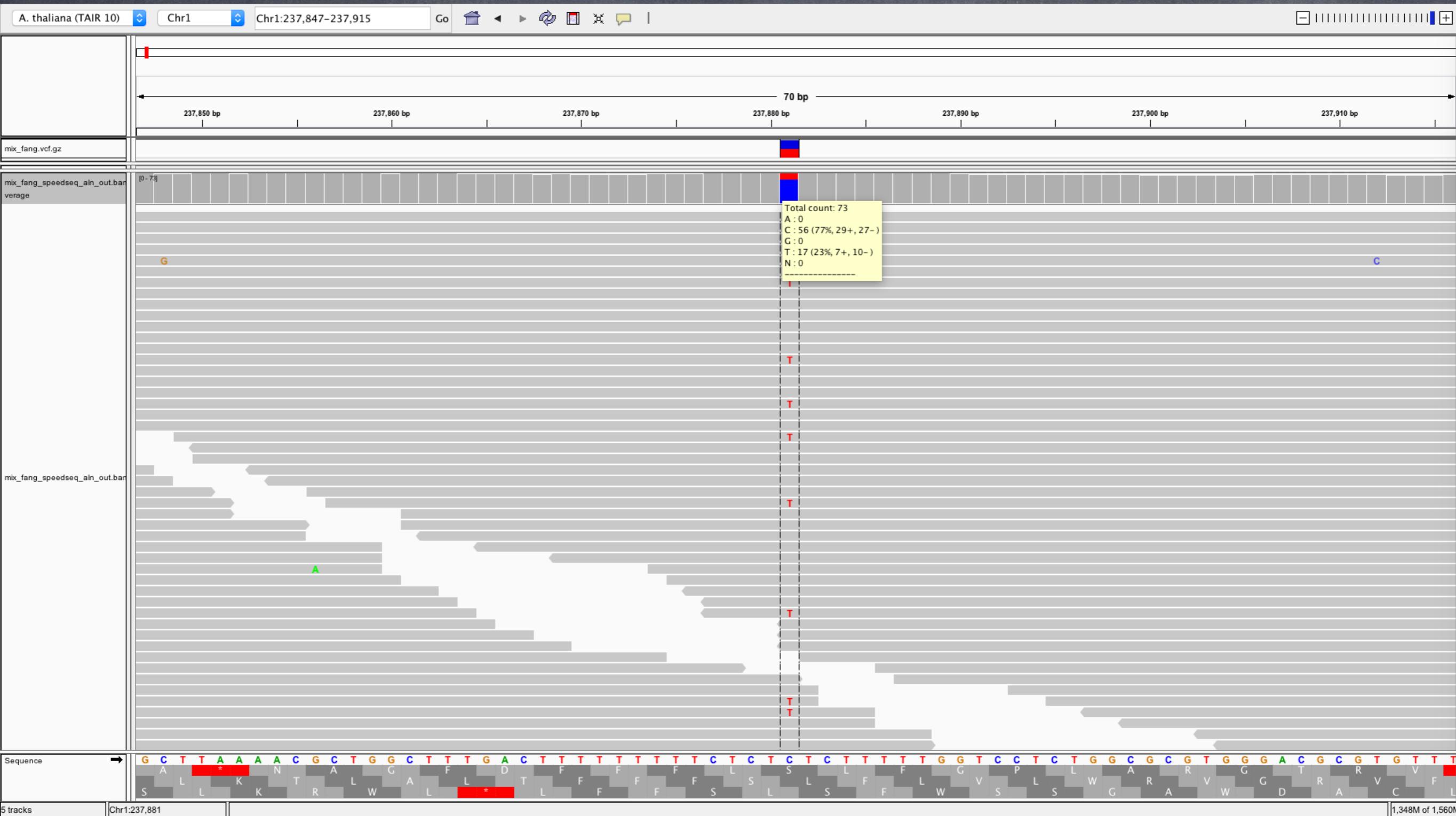
zcat annotated_biallelic_snps_intervaled_${ID}_freq.vcf.gz| \
/sdd1/users/linhua/QIAN_LAB/PROJECT-2016/resequencing_data_process/snpEff/scripts/vcfEffOnePerLine.pl | \
java -jar ${SnpSift} extractFields - \
CHROM POS REF ALT DP RO AO \
"ANN[*].ALLELE" \
"ANN[*].EFFECT" \
"ANN[*].IMPACT" \
"ANN[*].FEATURE" \
"ANN[*].FEATUREID" \
"ANN[*].GENE" \
"ANN[*].GENEID" \
"ANN[*].BIOTYPE" \
"ANN[*].HGVS_C" \
"ANN[*].HGVS_P" |sed "s/#//g" > ${ID}_multilines_results.tsv

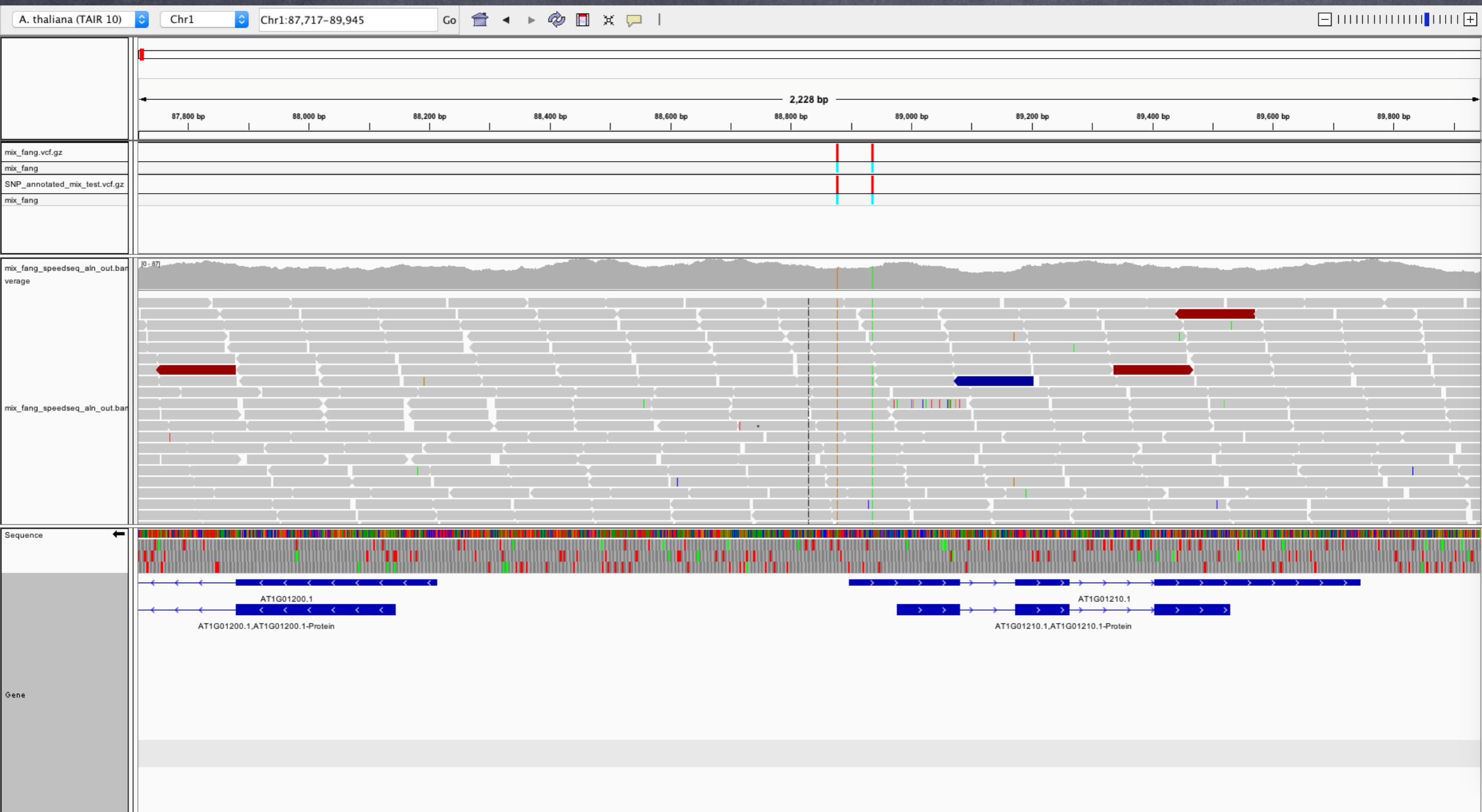
tput setaf 3; echo -n "|||||| extract nice fields is finished! |||||"

```

<http://lomereiter.github.io/sambamba/>
<https://github.com/GregoryFaust/samblaster>
<https://github.com/ekg/freebayes>
<http://snpeff.sourceforge.net/SnpSift.html#Extract>
http://snpeff.sourceforge.net/SnpEff_manual.html#input
<https://github.com/vcflib/vcflib>





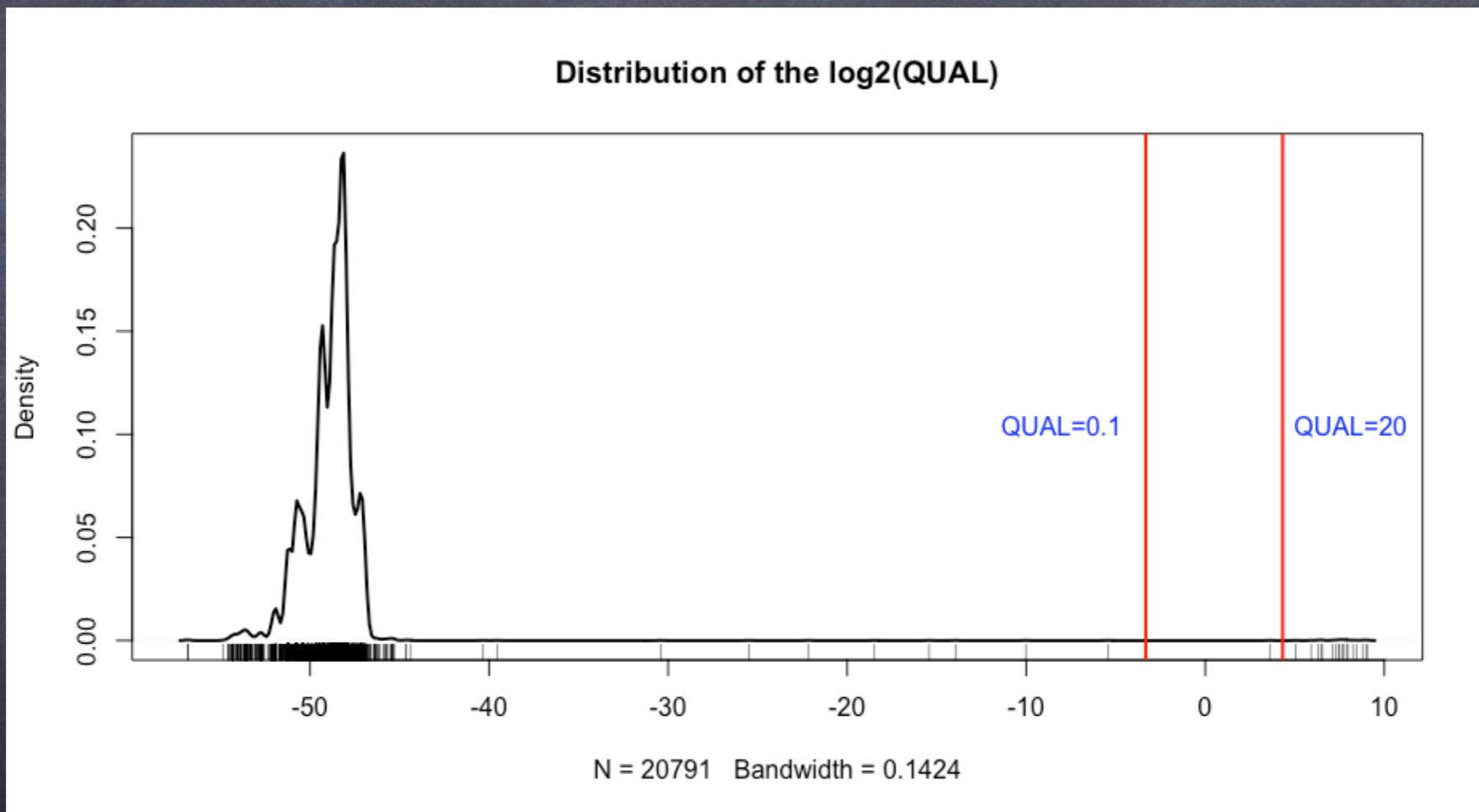


CHROM	POS	REF	ALT	DP	RO	AO	ANN[*].ALLELE	ANN[*].EFFECT	ANN[*].IMPACT	ANN[*].FEATURE	ANN[*].FEATUREID	ANN[*]
Chr1	88879	C	G	60	0	60	G	upstream_gene_variant	MODIFIER	transcript	AT1G01190.1.1	CYP78A
Chr1	88879	C	G	60	0	60	G	upstream_gene_variant	MODIFIER	transcript	AT1G01200.1.1	RABA3
Chr1	88879	C	G	60	0	60	G	upstream_gene_variant	MODIFIER	transcript	AT1G01210.1.1	AT1G01
Chr1	88879	C	G	60	0	60	G	upstream_gene_variant	MODIFIER	transcript	AT1G01220.1.1	FKGP
Chr1	88879	C	G	60	0	60	G	intergenic_region	MODIFIER	intergenic_region	AT1G01200-AT1G01210	RABA3-
Chr1	88937	G	A	64	0	64	A	5_prime_UTR_variant	MODIFIER	transcript	AT1G01210.1.1	AT1G01
Chr1	88937	G	A	64	0	64	A	upstream_gene_variant	MODIFIER	transcript	AT1G01190.1.1	CYP78A
Chr1	88937	G	A	64	0	64	A	upstream_gene_variant	MODIFIER	transcript	AT1G01200.1.1	RABA3
Chr1	88937	G	A	64	0	64	A	upstream_gene_variant	MODIFIER	transcript	AT1G01220.1.1	FKGP
Chr1	237881	C	T	73	56	17	T	upstream_gene_variant	MODIFIER	transcript	AT1G01650.1.1	SPPL4
Chr1	237881	C	T	73	56	17	T	upstream_gene_variant	MODIFIER	transcript	AT1G01650.2.1	SPPL4
Chr1	237881	C	T	73	56	17	T	downstream_gene_variant	MODIFIER	transcript	AT1G01660.1.1	PUB55
Chr1	237881	C	T	73	56	17	T	downstream_gene_variant	MODIFIER	transcript	AT1G01670.1.1	PUB56
Chr1	237881	C	T	73	56	17	T	intergenic_region	MODIFIER	intergenic_region	AT1G01650-AT1G01660	SPPL4-
Chr1	245706	G	A	65	48	17	A	upstream_gene_variant	MODIFIER	transcript	AT1G01660.1.1	PUB55
Chr1	245706	G	A	65	48	17	A	upstream_gene_variant	MODIFIER	transcript	AT1G01690.1.1	ATPRD3
Chr1	245706	G	A	65	48	17	A	downstream_gene_variant	MODIFIER	transcript	AT1G01680.1.1	PUB54
Chr1	245706	G	A	65	48	17	A	intron_variant	MODIFIER	transcript	AT1G01670.1.1	PUB56
Chr1	262745	C	T	67	46	21	T	upstream_gene_variant	MODIFIER	transcript	AT1G01700.1.1	ROPGEF
Chr1	262745	C	T	67	46	21	T	upstream_gene_variant	MODIFIER	transcript	AT1G01710.1.1	AT1G01
Chr1	262745	C	T	67	46	21	T	intergenic_region	MODIFIER	intergenic_region	AT1G01700-AT1G01710	ROPGEF
Chr1	389606	C	T	58	42	16	T	5_prime_UTR_variant	MODIFIER	transcript	AT1G02090.1.1	CSN7
Chr1	389606	C	T	58	42	16	T	5_prime_UTR_variant	MODIFIER	transcript	AT1G02090.2.1	CSN7
Chr1	389606	C	T	58	42	16	T	5_prime_UTR_variant	MODIFIER	transcript	AT1G02090.3.1	CSN7
Chr1	389606	C	T	58	42	16	T	upstream_gene_variant	MODIFIER	transcript	AT1G02100.1.1	AT1G02
Chr1	389606	C	T	58	42	16	T	upstream_gene_variant	MODIFIER	transcript	AT1G02110.1.1	AT1G02
Chr1	389606	C	T	58	42	16	T	upstream_gene_variant	MODIFIER	transcript	AT1G02100.2.1	AT1G02
Chr1	389606	C	T	58	42	16	T	upstream_gene_variant	MODIFIER	transcript	AT1G02100.3.1	AT1G02
Chr1	389606	C	T	58	42	16	T	downstream_gene_variant	MODIFIER	transcript	AT1G02080.1.1	AT1G02
Chr1	389606	C	T	58	42	16	T	downstream_gene_variant	MODIFIER	transcript	AT1G02080.2.1	AT1G02
Chr1	433009	G	T	14	10	4	T	upstream_gene_variant	MODIFIER	transcript	AT1G02220.1.1	NAC003
Chr1	433009	G	T	14	10	4	T	downstream_gene_variant	MODIFIER	transcript	AT1G02228.1.1	AT1G02
Chr1	433009	G	T	14	10	4	T	downstream_gene_variant	MODIFIER	transcript	AT1G02230.1.1	NAC004
Chr1	433009	G	T	14	10	4	T	downstream_gene_variant	MODIFIER	transcript	AT1G02250.1.1	NAC005
Chr1	433009	G	T	14	10	4	T	intergenic_region	MODIFIER	intergenic_region	AT1G02220-AT1G02228	NAC003
Chr1	439785	C	T	25	19	6	T	upstream_gene_variant	MODIFIER	transcript	AT1G02230.1.1	NAC004
Chr1	439785	C	T	25	19	6	T	upstream_gene_variant	MODIFIER	transcript	AT1G02250.1.1	NAC005
Chr1	439785	C	T	25	19	6	T	downstream_gene_variant	MODIFIER	transcript	AT1G02260.1.1	AT1G02
Chr1	439785	C	T	25	19	6	T	downstream_gene_variant	MODIFIER	transcript	AT1G02270.1.1	AT1G02
Chr1	439785	C	T	25	19	6	T	intergenic_region	MODIFIER	intergenic_region	AT1G02250-AT1G02260	NAC005
Chr1	445595	C	A	42	33	9	A	upstream_gene_variant	MODIFIER	transcript	AT1G02260.1.1	AT1G02
Chr1	445595	C	A	42	33	9	A	downstream_gene_variant	MODIFIER	transcript	AT1G02280.1.1	TOC33
Chr1	445595	C	A	42	33	9	A	downstream_gene_variant	MODIFIER	transcript	AT1G02290.1.1	AT1G02
Chr1	445595	C	A	42	33	9	A	downstream_gene_variant	MODIFIER	transcript	AT1G02280.2.1	TOC33
Chr1	445595	C	A	42	33	9	A	intron_variant	MODIFIER	transcript	AT1G02270.1.1	AT1G02
Chr1	466157	C	T	65	45	20	T	missense_variant	MODERATE	transcript	AT1G02340.1.1	HFR1
Chr1	466157	C	T	65	45	20	T	upstream_gene_variant	MODIFIER	transcript	AT1G02330.1.1	AT1G02
Chr1	466157	C	T	65	45	20	T	upstream_gene_variant	MODIFIER	transcript	AT1G02335.1.1	GL22
Chr1	466157	C	T	65	45	20	T	upstream_gene_variant	MODIFIER	transcript	AT1G02350.1.1	AT1G02
Chr1	466157	C	T	65	45	20	T	downstream_gene_variant	MODIFIER	transcript	AT1G02320.1.1	AT1G02
Chr1	484211	G	A	69	50	19	A	upstream_gene_variant	MODIFIER	transcript	AT1G02391.1.1	AT1G02

CHROM	POS	REF	ALT	DP	RO	AO	ANN[*].ALLELE	ANN[*].EFFECT
Chr1	88879	C	G	60	0	60	G_G_G_G_G	
Chr1	88937	G	A	64	0	64	A_A_A_A	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_intergenic_5_prime_UTR_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant
Chr1	237881	C	T	73	56	17	T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_interge
Chr1	245706	G	A	65	48	17	A_A_A_A	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_intron_variant
Chr1	262745	C	T	67	46	21	T_T_T	upstream_gene_variant_upstream_gene_variant_intergenic_region
Chr1	389606	C	T	58	42	16	T_T_T_T_T_T_T_T	5_prime_UTR_variant_5_prime_UTR_variant_5_prime_UTR_variant_upstream_gene_variant_upstream_gene_var
Chr1	433009	G	T	14	10	4	T_T_T_T_T	upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_inter
Chr1	439785	C	T	25	19	6	T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_interge
Chr1	445595	C	A	42	33	9	A_A_A_A_A	upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_intro
Chr1	466157	C	T	65	45	20	T_T_T_T_T	missense_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_
Chr1	484211	G	A	69	50	19	A_A_A_A	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_intergenic_region
Chr1	488998	G	A	72	54	18	A_A_A_A	missense_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_variant
Chr1	701752	C	T	69	50	19	T_T_T_T_T	upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_intro
Chr1	863625	C	A	44	33	11	A_A_A_A_A	upstream_gene_variant_downstream_gene_variant_intron_variant_intron_variant
Chr1	864361	A	G	77	0	77	G_G_G_G_G_G_G	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstrea
Chr1	868720	C	A	30	23	7	A_A_A_A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_ge
Chr1	926694	C	T	65	0	64	T_T_T_T_T_T	upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downs
Chr1	993202	C	A	45	35	10	A_A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstrea
Chr1	1035355	C	T	74	57	17	T_T_T	upstream_gene_variant_upstream_gene_variant_intron_variant
Chr1	1098510	G	A	57	41	16	A_A_A	missense_variant_downstream_gene_variant_downstream_gene_variant
Chr1	1107174	C	T	68	45	23	T_T_T	missense_variant_upstream_gene_variant_downstream_gene_variant
Chr1	1122547	T	G	48	37	10	G_G_G_G	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_intergenic_region
Chr1	1184116	G	A	56	38	18	A_A_A_A_A	upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_inter
Chr1	1187373	C	T	67	46	21	T_T_T_T_T	missense_variant_missense_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_varia
Chr1	1219083	C	T	69	40	29	T_T_T_T_T	synonymous_variant_synonymous_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene
Chr1	1245547	G	A	62	48	14	A_A_A_A_A	downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_int
Chr1	1293842	G	A	66	42	23	A_A_A_A_A_A_A	3_prime_UTR_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_
Chr1	1310066	G	T	55	43	12	T_T_T_T	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_intergenic_region
Chr1	1345160	C	T	70	55	15	T_T_T_T	upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_intron_variant
Chr1	1363159	G	A	48	34	13	A_A_A_A_A	missense_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene
Chr1	1404038	G	A	46	33	13	A_A_A_A_A_A_A_A_A_A	synonymous_variant_synonymous_variant_synonymous_variant_upstream_gene_variant_upstream_gene_varian
Chr1	1488659	G	A	61	46	15	A_A_A_A_A	synonymous_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstream
Chr1	1520201	C	T	40	27	13	T_T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_ge
Chr1	1520232	C	T	29	21	8	T_T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_ge
Chr1	1550239	G	A	69	49	20	A_A_A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_ge
Chr1	1575055	G	T	25	20	5	T_T_T_T	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_intergenic_region
Chr1	1582653	G	T	20	15	5	T_T_T_T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstr
Chr1	1583508	G	A	40	30	10	A_A_A_A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstr
Chr1	1583515	C	A	37	22	15	A_A_A_A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstr
Chr1	1593802	G	T	43	33	10	T_T_T_T	downstream_gene_variant_downstream_gene_variant_downstream_gene_variant_intergenic_region
Chr1	1608846	C	T	58	42	16	T_T	upstream_gene_variant_intron_variant
Chr1	1674982	C	T	64	49	15	T_T_T_T_T_T_T	missense_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gen
Chr1	1676186	G	T	28	21	7	T_T_T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstrea
Chr1	1694070	G	T	29	23	6	T_T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstr
Chr1	1706240	G	T	39	0	39	T_T_T_T_T	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_intergenic_
Chr1	1921582	G	A	60	45	15	A_A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_
Chr1	1995107	G	A	64	47	17	A_A_A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_downstr
Chr1	2055431	A	T	59	1	57	T_T_T_T	5_prime_UTR_variant_upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene
Chr1	2075166	G	A	57	45	12	A_A_A_A_A_A_A	synonymous_variant_synonymous_variant_downstream_gene_variant_downstream_gene_variant_downstream_ge
Chr1	2101066	G	A	58	43	15	A_A_A_A_A	upstream_gene_variant_upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_intron va
Chr1	2176767	T	C	21	7	13	C_C_C_C_C	upstream_gene_variant_upstream_gene_variant_downstream_gene_variant_downstream_gene_variant_interge

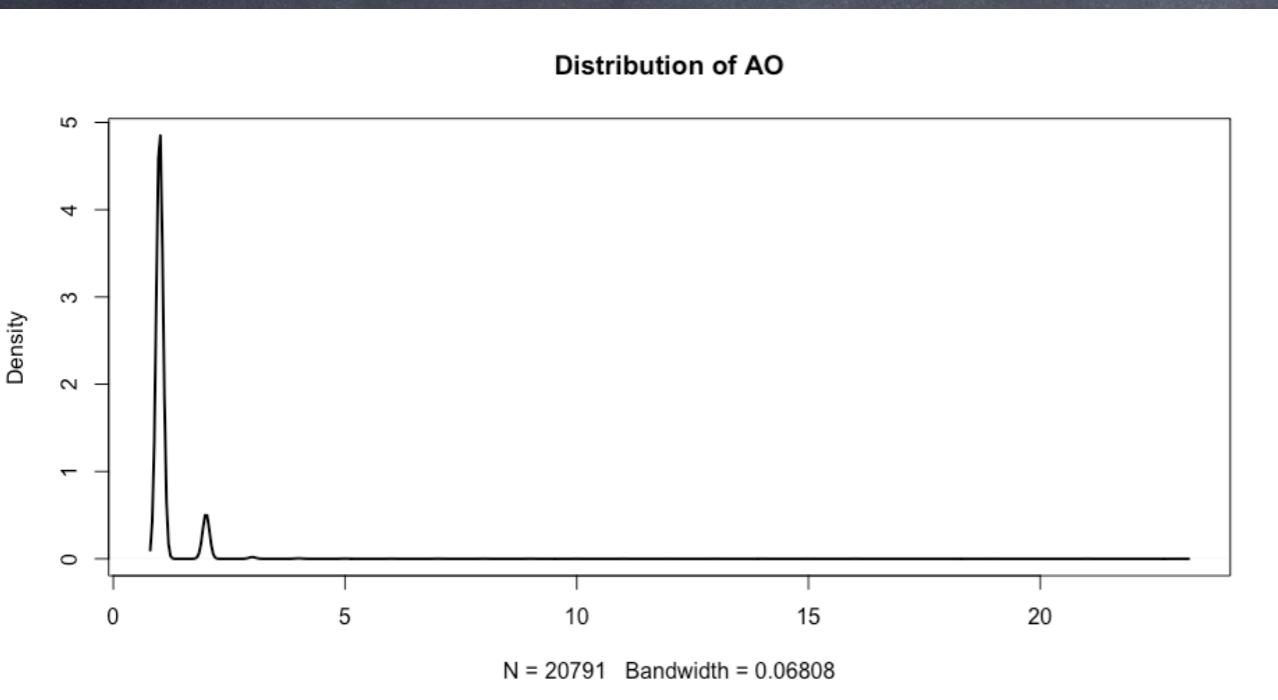
CHROM	POS	REF	AL1	DP	RO	AO	AL	EFFECT	IMPACT	FEATURE	FEATUREID	GENE	GENEID	BIOTYPE	HGVS_C	HGVS_P
Chr1	466157	C	T	65	45	20	T	missense_variant	MODERATE	transcript	AT1G02340. 1. 1	HFR1	AT1G02340	protein_coding	c. 655G>A	p. Ala219Thr
Chr1	488998	G	A	72	54	18	A	missense_variant	MODERATE	transcript	AT1G02400. 1. 1	GA20X6	AT1G02400	protein_coding	c. 680G>A	p. Gly227Asp
Chr1	1098510	G	A	57	41	16	A	missense_variant	MODERATE	transcript	AT1G04170. 1. 1	EIF2_GAMMA	AT1G04170	protein_coding	c. 519G>A	p. Met173Ile
Chr1	1107174	C	T	68	45	23	T	missense_variant	MODERATE	transcript	AT1G04190. 1. 1	TPR3	AT1G04190	protein_coding	c. 673G>A	p. Asp225Asn
Chr1	1187373	C	T	67	46	21	T	missense_variant	MODERATE	transcript	AT1G04400. 1. 1	CRY2	AT1G04400	protein_coding	c. 425G>A	p. Cys142Tyr
Chr1	1187373	C	T	67	46	21	T	missense_variant	MODERATE	transcript	AT1G04400. 2. 1	CRY2	AT1G04400	protein_coding	c. 425G>A	p. Cys142Tyr
Chr1	1219083	C	T	69	40	29	T	synonymous_variant	LOW	transcript	AT1G04490. 1. 1	AT1G04490	AT1G04490	protein_coding	c. 243G>A	p. Lys81Lys
Chr1	1219083	C	T	69	40	29	T	synonymous_variant	LOW	transcript	AT1G04490. 2. 1	AT1G04490	AT1G04490	protein_coding	c. 243G>A	p. Lys81Lys
Chr1	1363159	G	A	48	34	13	A	missense_variant	MODERATE	transcript	AT1G04840. 1. 1	PCMP-H64	AT1G04840	protein_coding	c. 1706C>T	p. Ala569Val
Chr1	1404038	G	A	46	33	13	A	synonymous_variant	LOW	transcript	AT1G04950. 1. 1	TAF6	AT1G04950	protein_coding	c. 1218C>T	p. Leu406Leu
Chr1	1404038	G	A	46	33	13	A	synonymous_variant	LOW	transcript	AT1G04950. 3. 1	TAF6	AT1G04950	protein_coding	c. 1218C>T	p. Leu406Leu
Chr1	1404038	G	A	46	33	13	A	synonymous_variant	LOW	transcript	AT1G04950. 2. 1	TAF6	AT1G04950	protein_coding	c. 1218C>T	p. Leu406Leu
Chr1	1488659	G	A	61	46	15	A	synonymous_variant	LOW	transcript	AT1G05160. 1. 1	KA01	AT1G05160	protein_coding	c. 883C>T	p. Leu295Leu
Chr1	1674982	C	T	64	49	15	T	missense_variant	MODERATE	transcript	AT1G05610. 1. 1	APS2	AT1G05610	protein_coding	c. 793G>A	p. Glu265Lys
Chr1	2075166	G	A	57	45	12	A	synonymous_variant	LOW	transcript	AT1G06750. 2. 1	AT1G06750	AT1G06750	protein_coding	c. 834G>A	p. Gln278Gln
Chr1	2075166	G	A	57	45	12	A	synonymous_variant	LOW	transcript	AT1G06750. 1. 1	AT1G06750	AT1G06750	protein_coding	c. 813G>A	p. Gln271Gln
Chr1	2218845	C	T	45	34	11	T	stop_gained	HIGH	transcript	AT1G07220. 1. 1	AT1G07220	AT1G07220	protein_coding	c. 183G>A	p. Trp61*
Chr1	2429052	C	T	66	52	14	T	missense_variant	MODERATE	transcript	AT1G07870. 2. 1	AT1G07870	AT1G07870	protein_coding	c. 1507G>A	p. Val503Ile
Chr1	2429052	C	T	66	52	14	T	missense_variant	MODERATE	transcript	AT1G07860. 1. 1	AT1G07860	AT1G07860	protein_coding	c. 298G>A	p. Val100Ile
Chr1	2563408	C	T	44	35	9	T	missense_variant	MODERATE	transcript	AT1G08170. 1. 1	AT1G08170	AT1G08170	protein_coding	c. 265G>A	p. Asp89Asn
Chr1	2616285	G	A	86	66	20	A	missense_variant	MODERATE	transcript	AT1G08300. 1. 1	NVL	AT1G08300	protein_coding	c. 1144G>A	p. Glu382Lys
Chr1	2959096	G	A	63	45	18	A	synonymous_variant	LOW	transcript	AT1G09170. 1. 1	AT1G09170	AT1G09170	protein_coding	c. 1501C>T	p. Leu501Leu
Chr1	2995552	G	A	57	42	15	A	missense_variant	MODERATE	transcript	AT1G09270. 1. 1	IMPA4	AT1G09270	protein_coding	c. 436G>A	p. Ala146Thr
Chr1	2995552	G	A	57	42	15	A	missense_variant	MODERATE	transcript	AT1G09270. 2. 1	IMPA4	AT1G09270	protein_coding	c. 436G>A	p. Ala146Thr
Chr1	2995552	G	A	57	42	15	A	missense_variant	MODERATE	transcript	AT1G09270. 3. 1	IMPA4	AT1G09270	protein_coding	c. 190G>A	p. Ala64Thr
Chr1	3102485	G	A	68	50	18	A	synonymous_variant	LOW	transcript	AT1G09575. 1. 1	AT1G09575	AT1G09575	protein_coding	c. 348C>T	p. Ile116Ile

Basic description of QUAL of the variants

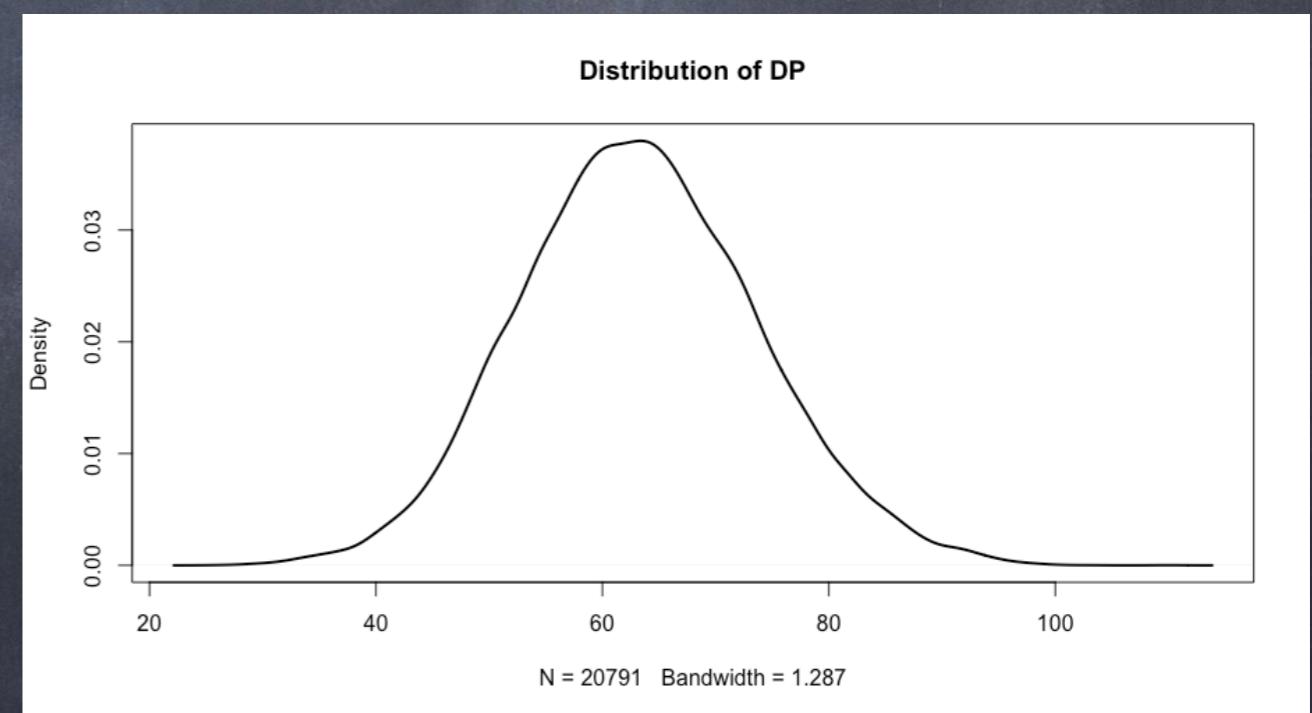


In selected region (interested genes)

Basic description of AO and DP of the variants

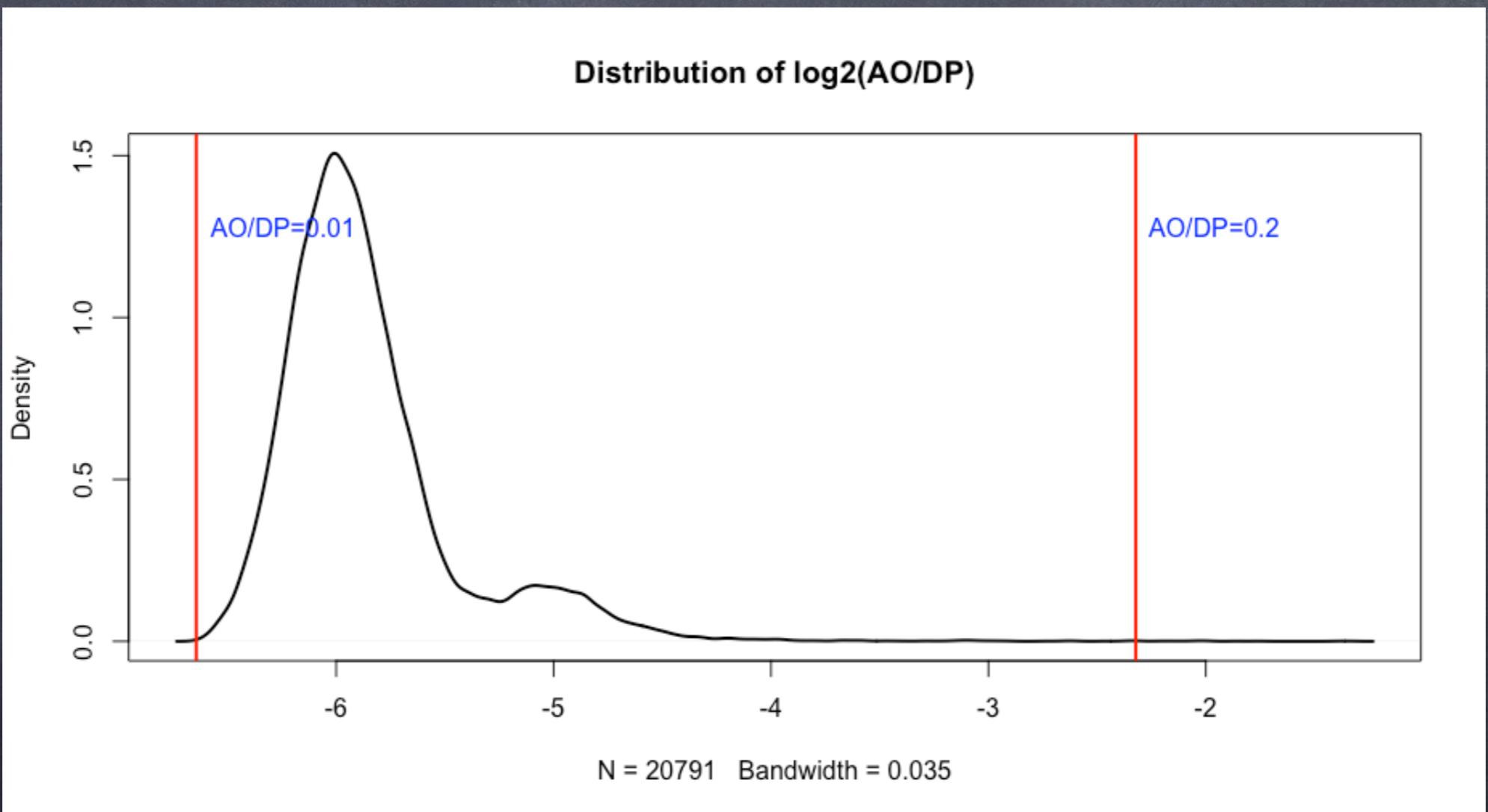


DP: Total read depth at the locus
AO: Alternate allele observations



In selected region (interested genes)

Basic description of AO and DP of the variants



In selected region (interested genes)

Distribution of the QUAL and AO/DP in the Whole Genome Scale

