Practical

Public resources for learning about the function of sequence variants

Variant Effect Predictor (VEP)

Variant Effect Predictor

VEP for Human GRCh37

If you are looking for VEP for Human GRCh37, please go to GRCh37 website.

Species: Human (Homo sapiens)

Assembly: GRCh38.p7

Name for this job (optional):

Either paste data:

Examples: Ensembl default, VCF, Variant identifiers, HGVS notations,
NB: pileup format no longer supported

Or upload file: Choose File No file chosen

Or provide file URL:

Transcript database to use:

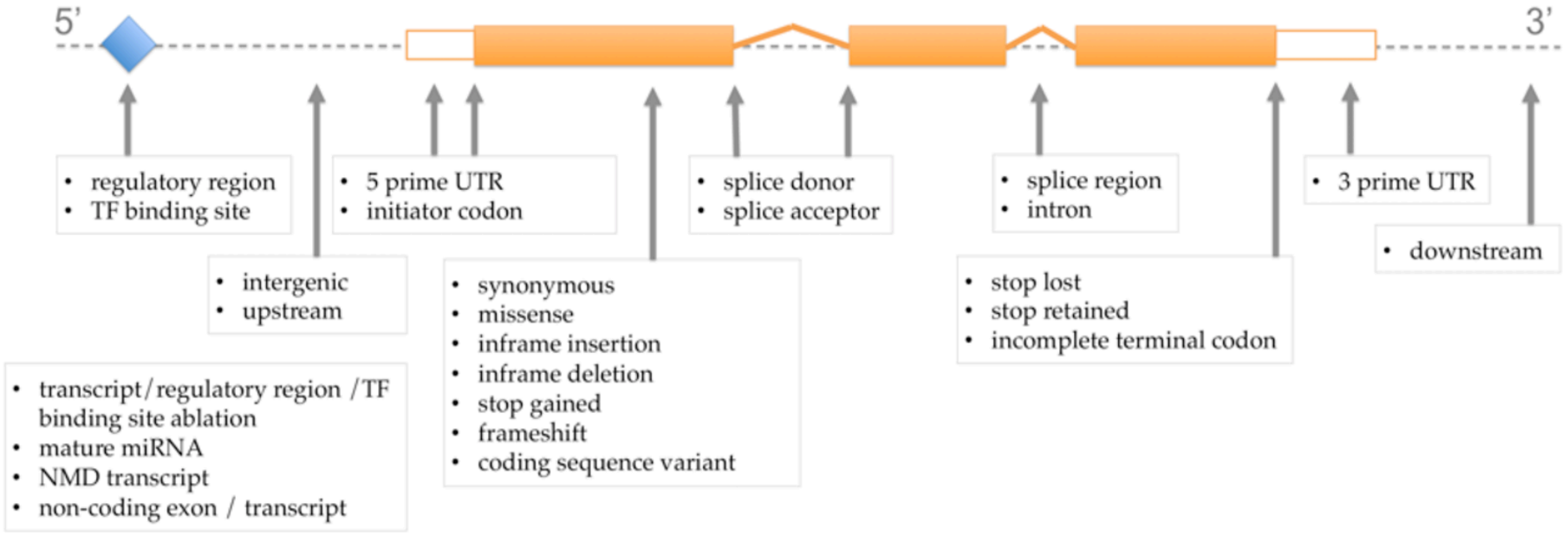
- (•) Ensembl transcripts
- ( ) Gencode basic transcripts
- ( ) RefSeq transcripts
- ( ) Ensembl and RefSeq transcripts

Identifiers and frequency data *Additional identifiers for genes, transcripts and variants; frequency data*

Extra options *e.g. SIFT, PolyPhen and regulatory data*

Filtering options *Pre-filter results by frequency or consequence type*

See below a diagram showing the location of each display term relative to the transcript structure:



*	SO term	SO description	SO accession	Display term	IMPACT
	transcript_ablation	A feature ablation whereby the deleted region includes a transcript feature	SO:0001893	Transcript ablation	HIGH
	splice_acceptor_variant	A splice variant that changes the 2 base region at the 3' end of an intron	SO:0001574	Splice acceptor variant	HIGH
	splice_donor_variant	A splice variant that changes the 2 base region at the 5' end of an intron	SO:0001575	Splice donor variant	HIGH
	stop_gained	A sequence variant whereby at least one base of a codon is changed, resulting in a premature stop codon, leading to a shortened transcript	SO:0001587	Stop gained	HIGH
	frameshift_variant	A sequence variant which causes a disruption of the translational reading frame, because the number of nucleotides inserted or deleted is not a multiple of three	SO:0001589	Frameshift variant	HIGH
	stop_lost	A sequence variant where at least one base of the terminator codon (stop) is changed, resulting in an elongated transcript	SO:0001578	Stop lost	HIGH
	start_lost	A codon variant that changes at least one base of the canonical start codo	SO:0002012	Start lost	HIGH
	transcript_amplification	A feature amplification of a region containing a transcript	SO:0001889	Transcript amplification	HIGH
	inframe_insertion	An inframe non synonymous variant that inserts bases into in the coding sequenc	SO:0001821	Inframe insertion	MODERATE
	inframe_deletion	An inframe non synonymous variant that deletes bases from the coding sequenc	SO:0001822	Inframe deletion	MODERATE
	missense_variant	A sequence variant, that changes one or more bases, resulting in a different amino acid sequence but where the length is preserved	SO:0001583	Missense variant	MODERATE
	protein_altering_variant	A sequence_variant which is predicted to change the protein encoded in the coding sequence	SO:0001818	Protein altering variant	MODERATE
	splice_region_variant	A sequence variant in which a change has occurred within the region of the splice site, either within 1-3 bases of the exon or 3-8 bases of the intron	SO:0001630	Splice region variant	LOW
	incomplete_terminal_codon_variant	A sequence variant where at least one base of the final codon of an incompletely annotated transcript is changed	SO:0001626	Incomplete terminal codon variant	LOW
	stop_retained_variant	A sequence variant where at least one base in the terminator codon is changed, but the terminator remains	SO:0001567	Stop retained variant	LOW
	synonymous_variant	A sequence variant where there is no resulting change to the encoded amino acid	SO:0001819	Synonymous variant	LOW
	coding_sequence_variant	A sequence variant that changes the coding sequence	SO:0001580	Coding sequence variant	MODIFIER

coding_sequence_variant	A sequence variant that changes the coding sequence	SO:0001580	Coding sequence variant	MODIFIER
mature_miRNA_variant	A transcript variant located with the sequence of the mature miRNA	SO:0001620	Mature miRNA variant	MODIFIER
5_prime_UTR_variant	A UTR variant of the 5' UTR	SO:0001623	5 prime UTR variant	MODIFIER
3_prime_UTR_variant	A UTR variant of the 3' UTR	SO:0001624	3 prime UTR variant	MODIFIER
non_coding_transcript_exon_variant	A sequence variant that changes non-coding exon sequence in a non-coding transcript	SO:0001792	Non coding transcript exon variant	MODIFIER
intron_variant	A transcript variant occurring within an intron	SO:0001627	Intron variant	MODIFIER
NMD_transcript_variant	A variant in a transcript that is the target of NMD	SO:0001621	NMD transcript variant	MODIFIER
non_coding_transcript_variant	A transcript variant of a non coding RNA gene	SO:0001619	Non coding transcript variant	MODIFIER
upstream_gene_variant	A sequence variant located 5' of a gene	SO:0001631	Upstream gene variant	MODIFIER
downstream_gene_variant	A sequence variant located 3' of a gene	SO:0001632	Downstream gene variant	MODIFIER
TFBS_ablation	A feature ablation whereby the deleted region includes a transcription factor binding site	SO:0001892	TFBS ablation	MODIFIER
TFBS_amplification	A feature amplification of a region containing a transcription factor binding site	SO:0001892	TFBS amplification	MODIFIER
TF_binding_site_variant	A sequence variant located within a transcription factor binding site	SO:0001782	TF binding site variant	MODIFIER
regulatory_region_ablation	A feature ablation whereby the deleted region includes a regulatory region	SO:0001894	Regulatory region ablation	MODERATE
regulatory_region_amplification	A feature amplification of a region containing a regulatory region	SO:0001891	Regulatory region amplification	MODIFIER
feature_elongation	A sequence variant located within a regulatory region	SO:0001907	Feature elongation	MODIFIER
regulatory_region_variant	A sequence variant located within a regulatory region	SO:0001566	Regulatory region variant	MODIFIER
feature_truncation	A sequence variant that causes the reduction of a genomic feature, with regard to the reference sequence	SO:0001906	Feature truncation	MODIFIER
intergenic_variant	A sequence variant located in the intergenic region, between genes	SO:0001628	Intergenic variant	MODIFIER

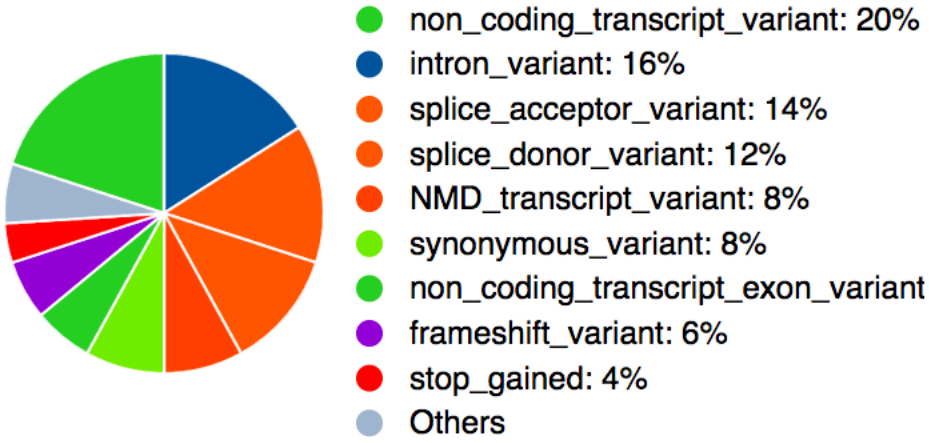
Variant Effect Predictor results

Job details

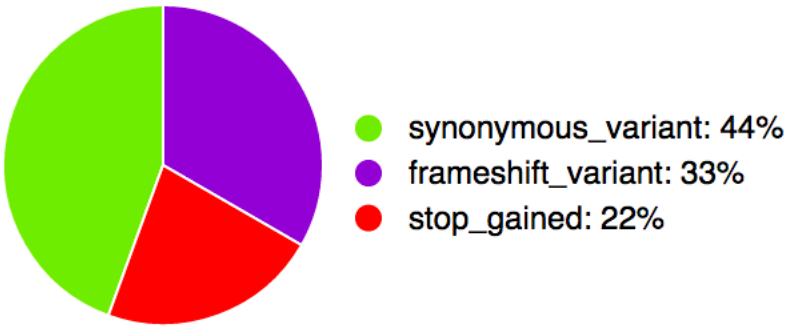
Summary statistics

Category	Count
Variants processed	6
Variants filtered out	0
Novel / existing variants	1 (16.7) / 5 (83.3)
Overlapped genes	5
Overlapped transcripts	28
Overlapped regulatory features	-

Consequences (all)



Coding consequences



Results preview

Navigation | **Filters** | **Download**

Page: 1 of 1 | **Show:** All variants

Uploaded variant | is | defined | Add

All: VCF VEP TXT

BioMart: Variants Genes

Show/hide columns

Uploaded variant	Location	Allele	Consequence	Impact	Symbol	Gene	Feature type	Feature	Biotype
.	1:202724482-202724483	A	frameshift_variant	HIGH	KDM5B	ENSG00000117139	Transcript	ENST00000235790	protein_coding
.	1:202724482-202724483	GG	synonymous_variant	LOW	KDM5B	ENSG00000117139	Transcript	ENST00000235790	protein_coding
.	1:202724482-202724483	A	frameshift_variant	HIGH	KDM5B	ENSG00000117139	Transcript	ENST00000367264	protein_coding
.	1:202724482-202724483	GG	synonymous_variant	LOW	KDM5B	ENSG00000117139	Transcript	ENST00000367264	protein_coding

ExAC

Interested in working on the development of this resource? Apply here.

ExAC Browser (Beta) | Exome Aggregation Consortium

Search for a gene or variant or region

Examples - Gene: PCSK9, Transcript: ENST00000407236, Variant: 22-46615880-T-C, Multi-allelic variant: rs1800234, Region: 22:46615715-46615880

About ExAC

The Exome Aggregation Consortium (ExAC) is a coalition of investigators seeking to aggregate and harmonize exome sequencing data from a wide variety of large-scale sequencing projects, and to make summary data available for the wider scientific community.

The data set provided on this website spans 60,706 unrelated individuals sequenced as part of various disease-specific and population genetic studies. The ExAC Principal Investigators and groups that have contributed data to the current release are listed here.

All data here are released under a Fort Lauderdale Agreement for the benefit of the wider biomedical community - see the terms of use here.

Sign up for our mailing list for future release announcements here

Recent News

August 8, 2016

- CNV calls are now available on the ExAC browser

March 14, 2016

- Version 0.3.1 ExAC data and browser (beta) is released! (Release notes)

January 13, 2015

- Version 0.3 ExAC data and browser (beta) is

Gene: KMT2A

KMT2A	lysine (K)-specific methyltransferase 2A
Number of variants	1865 (Including filtered: 2022)
Number of CNVs	65 (Including filtered: 81)
UCSC Browser	11:118307205-118397539
GeneCards	KMT2A
OMIM	KMT2A
Other	External References

Transcripts

Constraint from ExAC	Expected no. variants	Observed no. variants	Constraint Metric
Synonymous	505.6	466	z = 1.09
Missense	1242.6	764	z = [illegible]
LoF	115.7	4	pLI = 1.00
CNV	10.5	65	z = -2.47

Gene summary

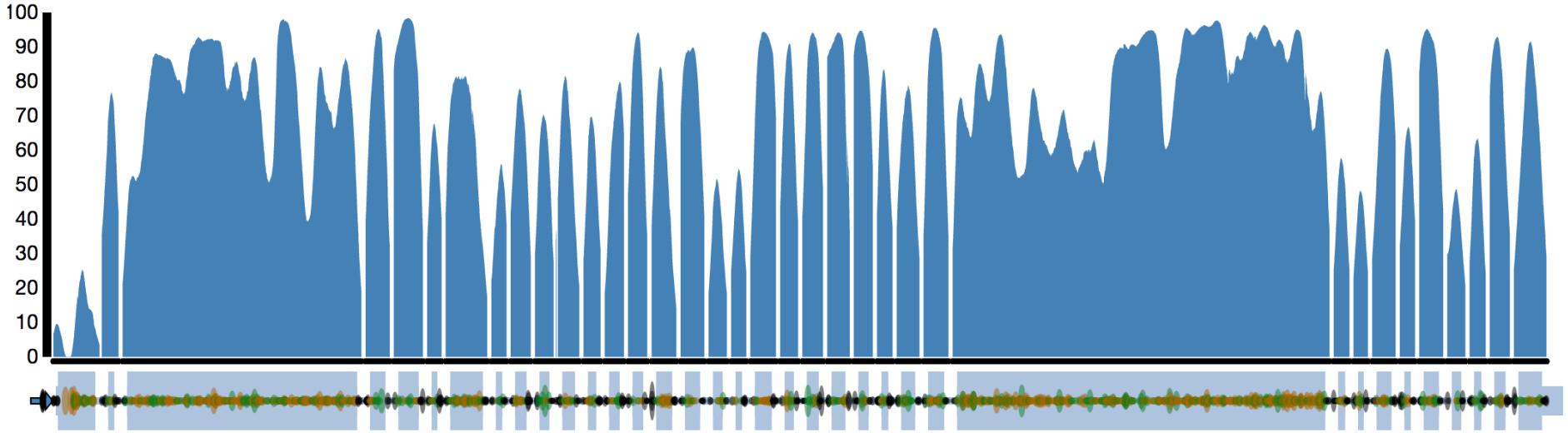
(Coverage shown for canonical transcript: ENST00000534358)

Mean coverage 68.95

Display: Overview | Detail | Include UTRs in plot

Coverage metric: Average | Individuals over X

Metric: mean



All | Missense + LoF | LoF ☐ Include filtered (non-PASS) variants ☐ Invert (highlight rare variants)

Export table to CSV

† denotes a consequence that is for a non-canonical transcript

Variant	Chrom	Position	Consequence	Filter	Annotation	Flags	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
11:118360506 G / T	11	118360506	c.4480-1G>T	PASS	splice acceptor	LoF flag	1	120924	0	0.000008270
11:118366413 A / G	11	118366413	c.5364-2A>G	PASS	splice acceptor		1	115724	0	0.000008641
11:118378325 G / C (rs141515578)	11	118378325	c.10835+1G>C	PASS	splice donor		2	119598	0	0.00001672
11:118380662 G / C	11	118380662	c.10901-1G>C	PASS	splice acceptor		1	116456	0	0.000008587

Gene: ATAD3C

ATAD3C ATPase family, AAA domain containing 3C

Number of variants 640 (Including filtered: 853)

Number of CNVs N/A

UCSC Browser 1:1385069-1405538

GeneCards ATAD3C

Other External References

Transcripts

Constraint from ExAC	Expected no. variants	Observed no. variants	Constraint Metric
Synonymous	87.6	86	z = 0.11
Missense	178.0	185	z = -0.26
LoF	14.8	9	pLI = 0.00
CNV	nan	nan	z = nan

Variant	Chrom	Position	Consequence	Filter	Annotation	Flags	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
1:1386091 C / T	1	1386091	p.Gln10Ter	PASS	stop gained		1	114078	0	0.000008766
1:1387445 T / TA	1	1387445	p.Arg33ThrfsTer87	PASS	frameshift		1	118504	0	0.000008439
1:1387453 C / T	1	1387453	p.Gln35Ter	PASS	stop gained		1	118460	0	0.000008442
1:1387503 G / A	1	1387503	c.152+1G>A	PASS	splice donor		2	114472	0	0.00001747
1:1387769 AG / A	1	1387769	p.Gly60AspfsTer6	PASS	frameshift		1	121228	0	0.000008249
1:1387774 TC / T	1	1387774	p.Arg62ValfsTer4	PASS	frameshift		1	121228	0	0.000008249
1:1389783 TC / T	1	1389783	p.Thr95LeufsTer14	PASS	frameshift		1	67648	0	0.00001478
1:1389839 TC / T	1	1389839	p.Arg114AlafsTer45	PASS	frameshift		2	92736	0	0.00002157
1:1389857 G / T	1	1389857	p.Glu119Ter	PASS	stop gained		1	92176	0	0.00001085
1:1390845 CAG / C	1	1390845	p.Ser129ProfsTer17	PASS	frameshift		1	111912	0	0.000008936
1:1391169 A / C	1	1391169	c.439-2A>C	PASS	splice acceptor		2	118728	0	0.00001685
1:1391177 C / CT	1	1391177	p.Glu150GlyfsTer73	PASS	frameshift		1	119124	0	0.000008395
1:1391206 AAT / A	1	1391206	p.Met159AspfsTer63	PASS	frameshift		7	119860	0	0.00005840
1:1391209 G / GCA	1	1391209	p.Thr160GlnfsTer43	PASS	frameshift		7	119902	0	0.00005838
1:1391231 CG / C	1	1391231	p.Gly168AlafsTer34	PASS	frameshift		1	120004	0	0.000008333
1:1391263 GC / G	1	1391263	p.Pro178HisfsTer24	PASS	frameshift		1	120132	0	0.000008324
1:1391702 TTGAC / T	1	1391702	p.Asp222GlyfsTer28	PASS	frameshift		1	13308	0	0.00007514
1:1392560 TGTGA / T	1	1392560	c.741+4_741+7delAGTG	PASS	splice donor		2	121098	0	0.00001652
1:1394555 G / GGACC	1	1394555	p.Leu254ArgfsTer63	PASS	frameshift		1	82698	0	0.00001209
1:1394593 CG / C	1	1394593	p.Gly266AlafsTer9	PASS	frameshift		3	95316	0	0.00003147
1:1396174 G / A	1	1396174	p.Trp286Ter	PASS	stop gained		3	113124	0	0.00002652
1:1396183 A / ATGCC	1	1396183	p.Ile292LeufsTer25	PASS	frameshift		1	115046	0	0.000008692
1:1397996 C / T	1	1397996	p.Gln333Ter	PASS	stop gained		1	59712	0	0.00001675
1:1403771 C / CCTCTCTCCCCA	1	1403771	p.Tyr367SerfsTer12	PASS	frameshift		1	114454	0	0.000008737
1:1403820 C / A (rs111835061)	1	1403820	p.Cys382Ter	PASS	stop gained		1	117398	0	0.000008518
1:1403848 C / T	1	1403848	p.Gln392Ter	PASS	stop gained		1	114526	0	0.000008732

Variant: 11:118366413 A / G

Filter Status	PASS
dbSNP	Not found in dbSNP
Allele Frequency	8.641e-06
Allele Count	1 / 115724
UCSC	11-118366413-A-G
ClinVar	Click to search for variant in Clinvar

Genotype Quality Metrics

Site Quality Metrics

Annotations

This variant falls on 4 transcripts in 1 genes:

splice acceptor

- KMT2A Transcripts

Note: This list may not include additional transcripts in the same gene that the variant does not overlap.

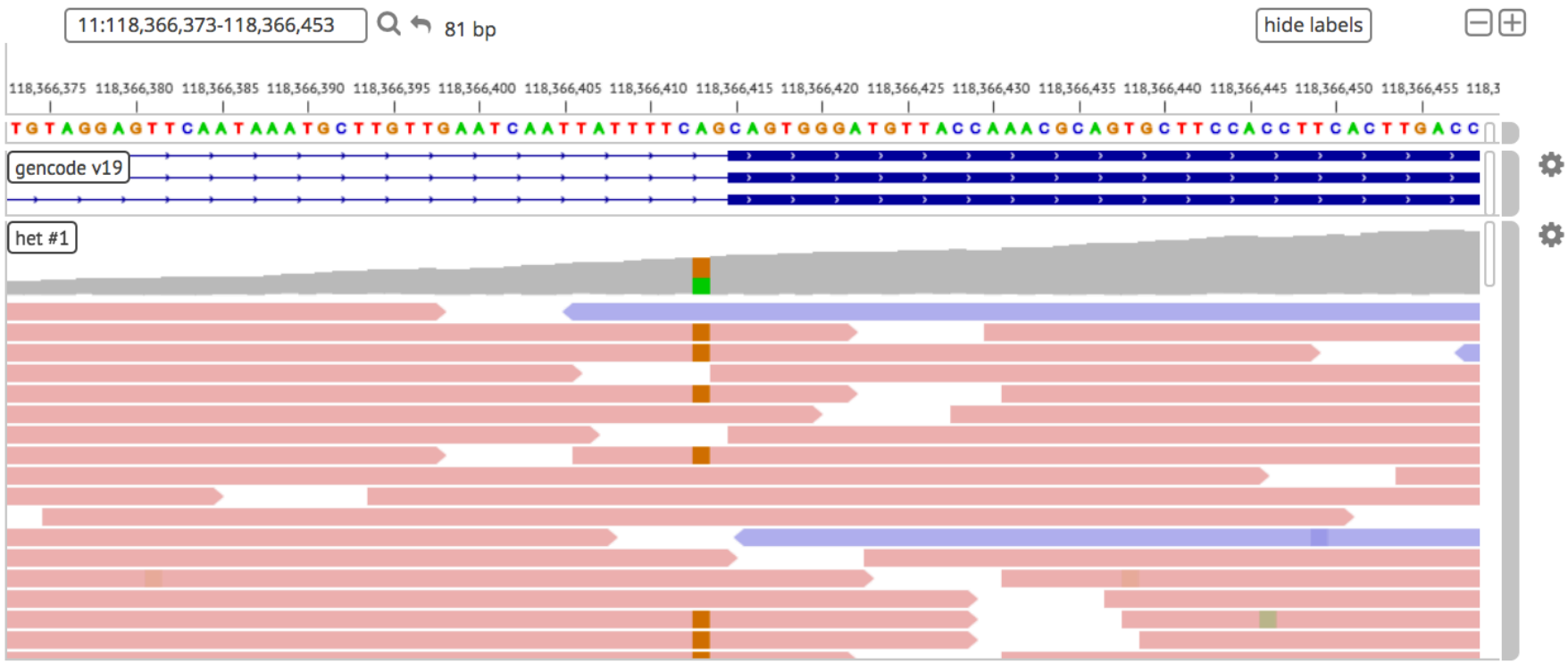
Population Frequencies

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
European (Non-Finnish)	1	65284	0	1.532e-05
African	0	10340	0	0
East Asian	0	8632	0	0
European (Finnish)	0	6614	0	0
Latino	0	11506	0	0
Other	0	846	0	0
South Asian	0	12502	0	0
Total	**1**	**115724**	**0**	**8.641e-06**

Read Data

This interactive IGV.js visualization shows reads that went into calling this variant.

Note: These are reassembled reads produced by GATK HaplotypeCaller --bamOutput so they accurately represent what HaplotypeCaller was seeing when it called this variant.



Variants

All | Missense + LoF | LoF ☐ Include filtered (non-PASS) variants ☐ Invert (highlight rare variants)

Export table to CSV

Variant	Chrom	Position	Consequence	Filter	Annotation	Flags	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
9:127661644 C / T	9	127661644	p.Thr425Thr	PASS	synonymous		1	20962	0	0.00004771
9:127661645 G / A (rs634710)	9	127661645	p.Thr425Met	PASS	missense		8852	21338	1854	0.4148
9:127661653 C / T (rs192663298)	9	127661653	p.Leu422Leu	PASS	synonymous		5	21062	0	0.0002374

Gene: ZNF788

ZNF788	zinc finger family member 788
Number of variants	570 (Including filtered: 678)
Number of CNVs	5 (Including filtered: 11)
UCSC Browser	19:12203078-12248050
GeneCards	ZNF788
Other	External References

Transcripts

Constraint from ExAC	Expected no. variants	Observed no. variants	Constraint Metric
Synonymous	6.9	7	z = -0.02
Missense	15.3	13	z = 0.29
LoF	3.2	2	pLI = 0.05
CNV	2.0	5	z = -0.54

Gene summary

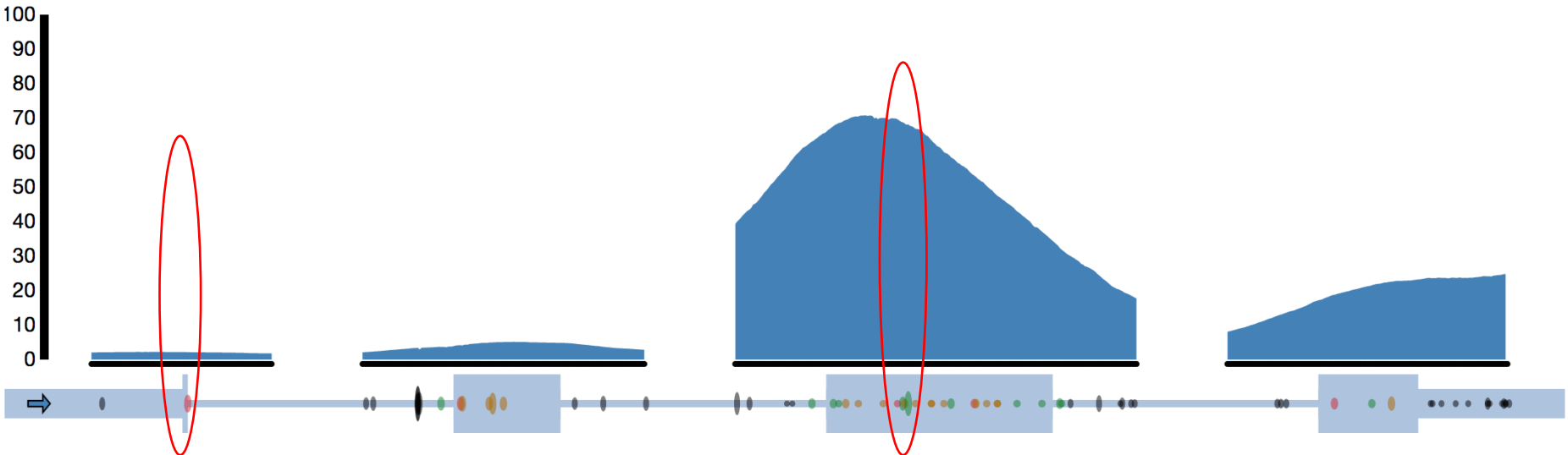
(Coverage shown for canonical transcript: ENST00000596883)

Mean coverage 19.97

Display: Overview | Detail ☐ **Include UTRs in plot**

Coverage metric: Average | Individuals over X

Metric: mean



Ensembl

Variant displays
- Explore this variant
- Genomic context
  - Genes and regulation
  - Flanking sequence
- **Population genetics**
- Phenotype data
- Sample genotypes
- Linkage disequilibrium
- Phylogenetic context
- Citations

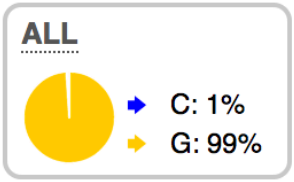
Configure this page
Custom tracks
Export data
Share this page
Bookmark this page

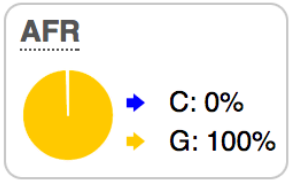
rs34536443 SNP

Most severe consequence	**Missense variant** I See all predicted consequences
Alleles	G/C I Ancestral: **G** I Ambiguity code: **S** I MAF: **0.01** (C)
Location	Chromosome **19:10352442** (forward strand) I View in location tab
Co-located variant	**HGMD-PUBLIC** CM070300
Evidence status	
Clinical significance	
HGVS names	This variant has **20** HGVS names - Show
Synonyms	This variant has **2** synonyms - Hide • **ClinVar** RCV000126195 • **Uniprot** VAR_041874
Genotyping chips	This variant has assays on **4** chips - Show
Original source	Variants (including SNPs and indels) imported from dbSNP (release 147) I View in dbSNP
About this variant	This variant overlaps 15 transcripts, 1 regulatory feature, has 2505 sample genotypes, is associated with 2 phenotypes and is mentioned in 29 citations.
Description from SNPedia	Description not available [More information from SNPedia]

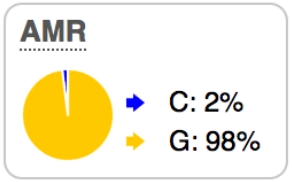
Population genetics

1000 Genomes Project Phase 3 allele frequencies

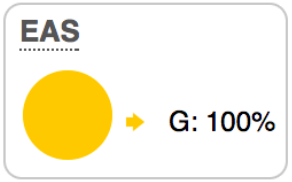




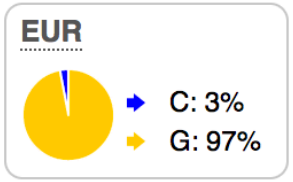
Sub-populations



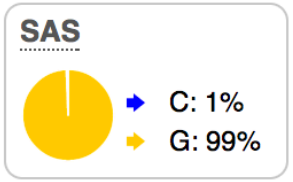
Sub-populations



Sub-populations



Sub-populations



Sub-populations

1000 Genomes Project Phase 3 (32)

Show All entries Show/hide columns

Population	Allele: frequency (count)		Genotype: frequency (count)	
ALL	G: 0.990 (4957)	C: 0.010 (51)	GIG: 0.980 (2453)	CIG: 0.020 (51)
AFR	G: 0.998 (1320)	C: 0.002 (2)	GIG: 0.997 (659)	CIG: 0.003 (2)
ACB	G: 0.995 (191)	C: 0.005 (1)	GIG: 0.990 (95)	CIG: 0.010 (1)
ASW	G: 0.992 (121)	C: 0.008 (1)	GIG: 0.984 (60)	CIG: 0.016 (1)
ESN	G: 1.000 (198)		GIG: 1.000 (99)	
GWD	G: 1.000 (226)		GIG: 1.000 (113)	
LWK	G: 1.000 (198)		GIG: 1.000 (99)	
MSL	G: 1.000 (170)		GIG: 1.000 (85)	
YRI	G: 1.000 (216)		GIG: 1.000 (108)	
AMR	G: 0.980 (680)	C: 0.020 (14)	GIG: 0.960 (333)	CIG: 0.040 (14)
CLM	G: 0.979 (184)	C: 0.021 (4)	GIG: 0.957 (90)	CIG: 0.043 (4)
MXL	G: 0.984 (126)	C: 0.016 (2)	GIG: 0.969 (62)	CIG: 0.031 (2)
PEL	G: 0.988 (168)	C: 0.012 (2)	GIG: 0.976 (83)	CIG: 0.024 (2)
PUR	G: 0.971 (202)	C: 0.029 (6)	GIG: 0.942 (98)	CIG: 0.058 (6)
EAS	G: 1.000 (1008)		GIG: 1.000 (504)	

rs34536443 SNP

Most severe consequence	**Missense variant** | See all predicted consequences
Alleles	**G/C** | Ancestral: **G** | Ambiguity code: **S** | MAF: **0.01** (C)
Location	Chromosome **19:10352442** (forward strand) | View in location tab
Co-located variant	**HGMD-PUBLIC** CM070300
Evidence status	
Clinical significance	
HGVS names	This variant has **20** HGVS names - Show
Synonyms	This variant has **2** synonyms - Hide • **ClinVar** RCV000126195 • **Uniprot** VAR_041874
Genotyping chips	This variant has assays on **4** chips - Show
Original source	Variants (including SNPs and indels) imported from dbSNP (release 147) | View in dbSNP
About this variant	This variant overlaps 15 transcripts, 1 regulatory feature, has 2505 sample genotypes, is associated with 2 phenotypes and is mentioned in 29 citations.
Description from SNPedia	Description not available [More information from SNPedia]

Phenotype Data

Significant association(s)

Show/hide columns | Filter

Phenotype, disease and trait	Source(s)	Mapped Terms	Ontology Accessions	Study	Clinical significance	Reported gene(s)	Associated allele	Statistics	Genomic Locations
ClinVar: phenotype not specified	ClinVar	-	-	-	★★☆☆	TYK2	C	-	View on Karyotype
RHEUMATOID ARTHRITIS	NHGRI-EBI GWAS catalog	7 terms	7 accessions	PMID:24390342	-	TYK2	G	**p-value:** $5.00e^{-16}$ **odds ratio:** 1.46	View on Karyotype

Genes and regulation

Gene and Transcript consequences

Show All entries | Show/hide columns | Filter

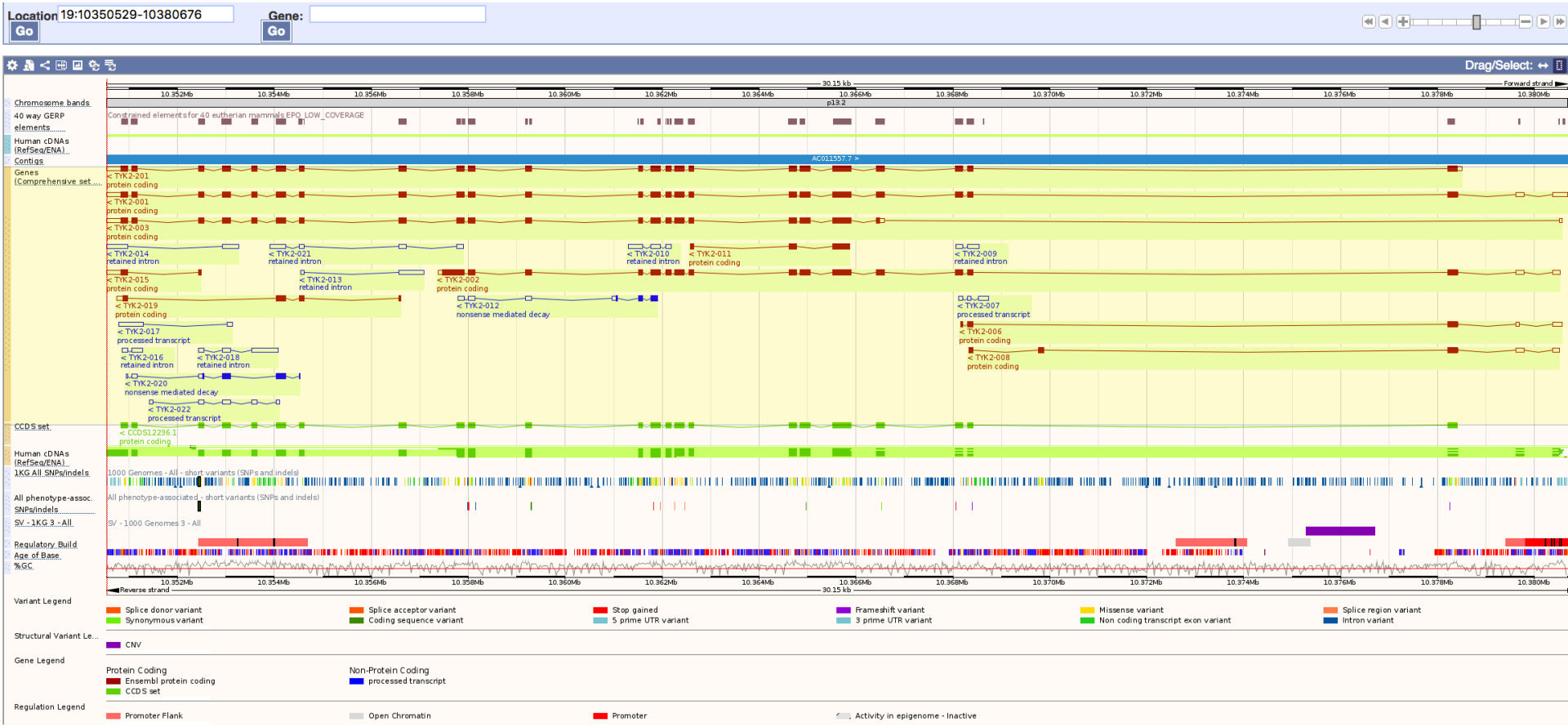
Gene	Transcript (strand)	Allele (transcript allele)	Consequence Type	Position in transcript	Position in CDS	Position in protein	Amino acid	Codons	SIFT	PolyPhen	Detail
ENSG00000105397 HGNC: TYK2	ENST00000264818.10 (-) biotype: protein_coding	C (G)	Missense variant	3412 (out of 3971)	3310 (out of 3564)	1104 (out of 1187)	P/A	CCC/GCC	0	0.979	Show
ENSG00000105397 HGNC: TYK2	ENST00000592137.1 (-) biotype: processed_transcript	C (G)	Non coding transcript exon variant; Non coding transcript variant	464 (out of 536)	-	-	-	-	-	-	Show
ENSG00000105397 HGNC: TYK2	ENST00000525621.5 (-) biotype: protein_coding	C (G)	Missense variant	3792 (out of 4347)	3310 (out of 3564)	1104 (out of 1187)	P/A	CCC/GCC	0	0.979	Show
ENSG00000105397 HGNC: TYK2	ENST00000525976.5 (-) biotype: protein_coding	C (G)	Missense variant	51 (out of 495)	52 (out of 194)	18 (out of 64)	P/A	CCC/GCC	0	0.999	Show
ENSG00000105397 HGNC: TYK2	ENST00000529739.1 (-) biotype: retained_intron	C (G)	Non coding transcript exon	834 (out of 842)	-	-	-	-	-	-	Show
ENSG00000105397 HGNC: TYK2	ENST00000524462.5 (-) biotype: protein_coding	C (G)	Missense variant	2901 (out of 3456)	2755 (out of 3009)	919 (out of 1002)	P/A	CCC/GCC	0	0.979	Show
LRG_121 HGNC: TYK2	LRG_121t1.1 (+) biotype: LRG_gene	C (C)	Missense variant	3688 (out of 4248)	3310 (out of 3564)	1104 (out of 1187)	P/A	CCC/GCC	0	0.979	Show
ENSG00000105397 HGNC: TYK2	ENST00000527481.2 (-) biotype: nonsense_mediated_decay	C (G)	3 prime UTR variant; NMD transcript variant	487 (out of 628)	-	-	-	-	-	-	Show
ENSG00000105397 HGNC: TYK2	ENST00000529422.1 (-) biotype: processed_transcript	C (G)	Intron variant; Non coding transcript variant	-	-	-	-	-	-	-	Show
ENSG00000105397 HGNC: TYK2	ENST00000530220.1 (-) biotype: retained_intron	C (G)	Intron variant; Non coding transcript variant	-	-	-	-	-	-	-	Show
ENSG00000105397 HGNC: TYK2	ENST00000530560.5 (-) biotype: protein_coding	C (G)	Intron variant	-	-	-	-	-	-	-	Show
ENSG00000105397 HGNC: TYK2	ENST00000524470.1 (-) biotype: retained_intron	C (G)	Upstream gene variant	-	-	-	-	-	-	-	Show
ENSG00000105397 HGNC: TYK2	ENST00000529370.5 (-) biotype: protein_coding	C (G)	Downstream gene variant	-	-	-	-	-	-	-	Show

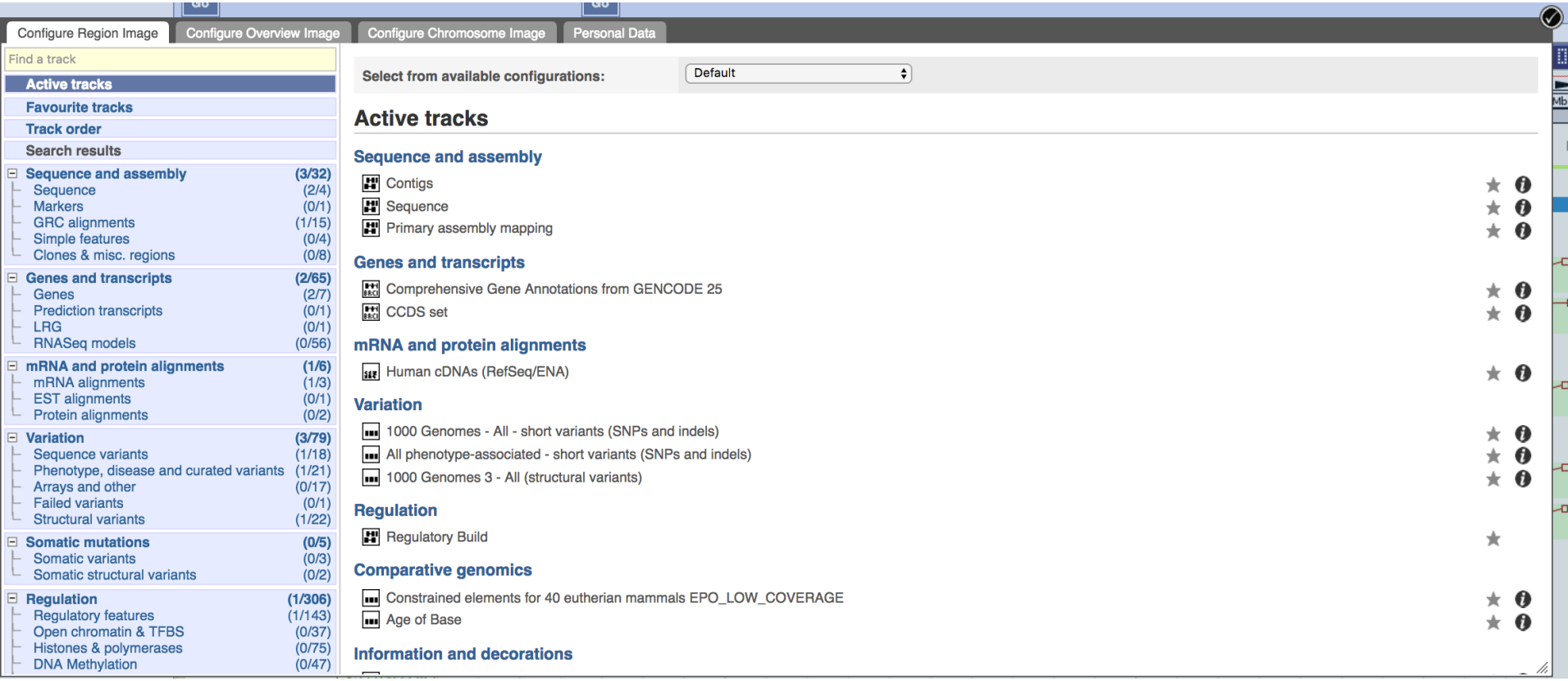
Gene: TYK2 ENSG00000105397

Description	tyrosine kinase 2 [Source:HGNC Symbol;Acc:HGNC:12440]
Synonyms	IMD35, JTK1
Location	Chromosome 19: 10,350,529-10,380,676 reverse strand. GRCh38:CM000681.2
About this gene	This gene has 21 transcripts (splice variants), 55 orthologues, 13 paralogues, is a member of 1 Ensembl protein family and is associated with 5 phenotypes.
Transcripts	Show transcript table

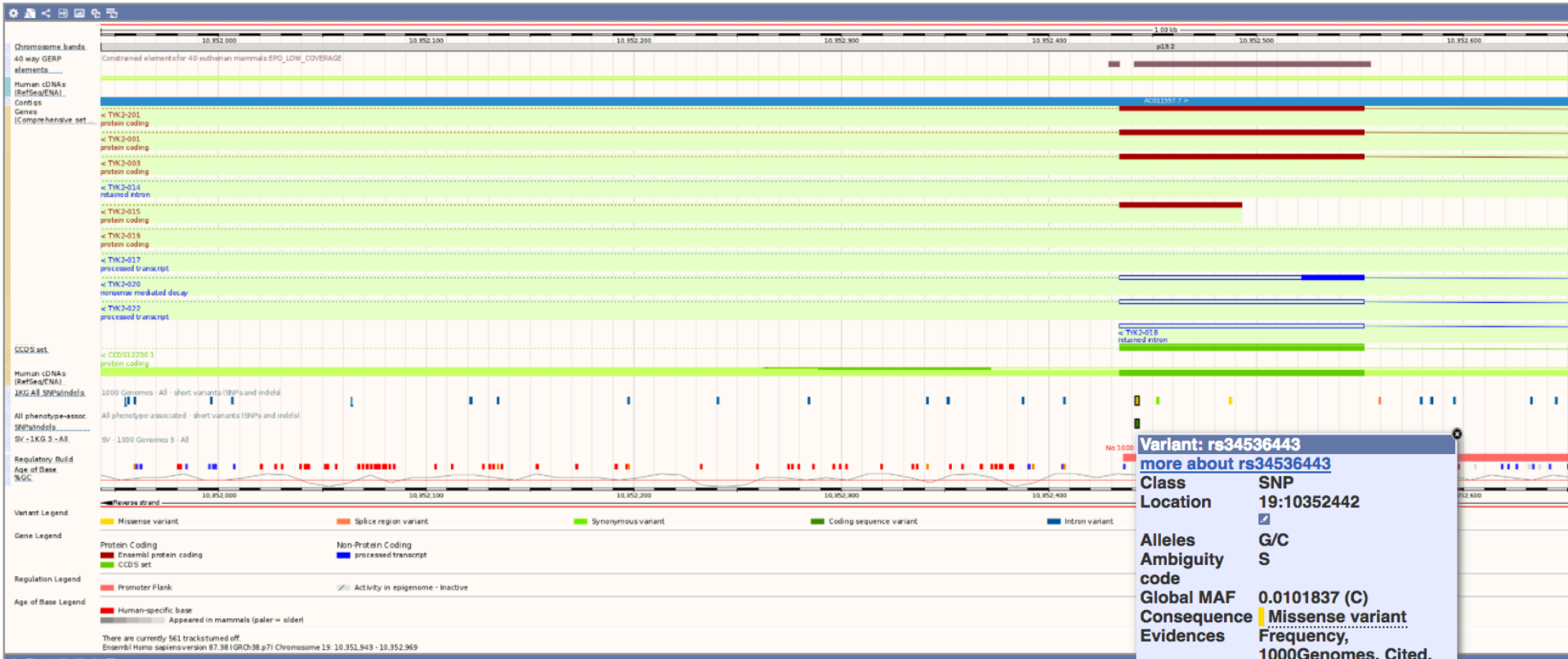
Summary

Name	TYK2 (HGNC Symbol)
CCDS	This gene is a member of the Human CCDS set: CCDS12236.1
UniProtKB	This gene has proteins that correspond to the following UniProtKB identifiers: P29597
RefSeq	Overlapping RefSeq Gene ID 7297 matches and has similar biotype of protein_coding
LRG	LRG_121 provides a stable genomic reference framework for describing sequence variants for this gene
Ensembl version	ENSG00000105397.13
Other assemblies	This gene maps to 10,461,205-10,491,352 in GRCh37 coordinates. View this locus in the GRCh37 archive: ENSG00000105397


Location 19:10350529-10380676
Go
Gene:
Go
Drag/Select:
30.15 kb
Forward strand
Chromosome bands
40 way GERP elements
Constrained elements for 40 eutherian mammals EPO_LOW_COVERAGE
Human cDNAs (RefSeq/ENA)
Contigs
AC011557.7 >
Genes (Comprehensive set ...
< TYK2-201 protein coding
< TYK2-001 protein coding
< TYK2-003 protein coding
< TYK2-014 retained intron
< TYK2-021 retained intron
< TYK2-010 retained intron
< TYK2-011 protein coding
< TYK2-009 retained intron
< TYK2-015 protein coding
< TYK2-013 retained intron
< TYK2-002 protein coding
< TYK2-019 protein coding
< TYK2-012 nonsense mediated decay
< TYK2-007 processed transcript
< TYK2-017 processed transcript
< TYK2-006 protein coding
< TYK2-016 retained intron
< TYK2-018 retained intron
< TYK2-008 protein coding
< TYK2-020 nonsense mediated decay
< TYK2-022 processed transcript
CCDS set
< CCDS12236.1 protein coding
Human cDNAs (RefSeq/ENA)
1KG All SNPs/indels
1000 Genomes - All - short variants (SNPs and indels)
All phenotype-assoc. SNPs/indels
All phenotype-associated - short variants (SNPs and indels)
SV - 1KG 3 - All
SV - 1000 Genomes 3 - All
Regulatory Build
Age of Base
%GC
Reverse strand
Variant Legend
Splice donor variant
Splice acceptor variant
Stop gained
Frameshift variant
Missense variant
Splice region variant
Synonymous variant
Coding sequence variant
5 prime UTR variant
3 prime UTR variant
Non coding transcript exon variant
Intron variant
Structural Variant Le...
CNV
Gene Legend
Protein Coding
Ensembl protein coding
CCDS set
Non-Protein Coding
processed transcript
Regulation Legend
Promoter Flank
Open Chromatin
Promoter
Activity in epigenome - Inactive


Configure Region Image
Configure Overview Image
Configure Chromosome Image
Personal Data
Find a track
Active tracks
Favourite tracks
Track order
Search results
Sequence and assembly (3/32)
Sequence (2/4)
Markers (0/1)
GRC alignments (1/15)
Simple features (0/4)
Clones & misc. regions (0/8)
Genes and transcripts (2/65)
Genes (2/7)
Prediction transcripts (0/1)
LRG (0/1)
RNASeq models (0/56)
mRNA and protein alignments (1/6)
mRNA alignments (1/3)
EST alignments (0/1)
Protein alignments (0/2)
Variation (3/79)
Sequence variants (1/18)
Phenotype, disease and curated variants (1/21)
Arrays and other (0/17)
Failed variants (0/1)
Structural variants (1/22)
Somatic mutations (0/5)
Somatic variants (0/3)
Somatic structural variants (0/2)
Regulation (1/306)
Regulatory features (1/143)
Open chromatin & TFBS (0/37)
Histones & polymerases (0/75)
DNA Methylation (0/47)
Select from available configurations:
Default
Active tracks
Sequence and assembly
Contigs
Sequence
Primary assembly mapping
Genes and transcripts
Comprehensive Gene Annotations from GENCODE 25
CCDS set
mRNA and protein alignments
Human cDNAs (RefSeq/ENA)
Variation
1000 Genomes - All - short variants (SNPs and indels)
All phenotype-associated - short variants (SNPs and indels)
1000 Genomes 3 - All (structural variants)
Regulation
Regulatory Build
Comparative genomics
Constrained elements for 40 eutherian mammals EPO_LOW_COVERAGE
Age of Base
Information and decorations

Locat 19:10351943-10352 Go Gene: Go



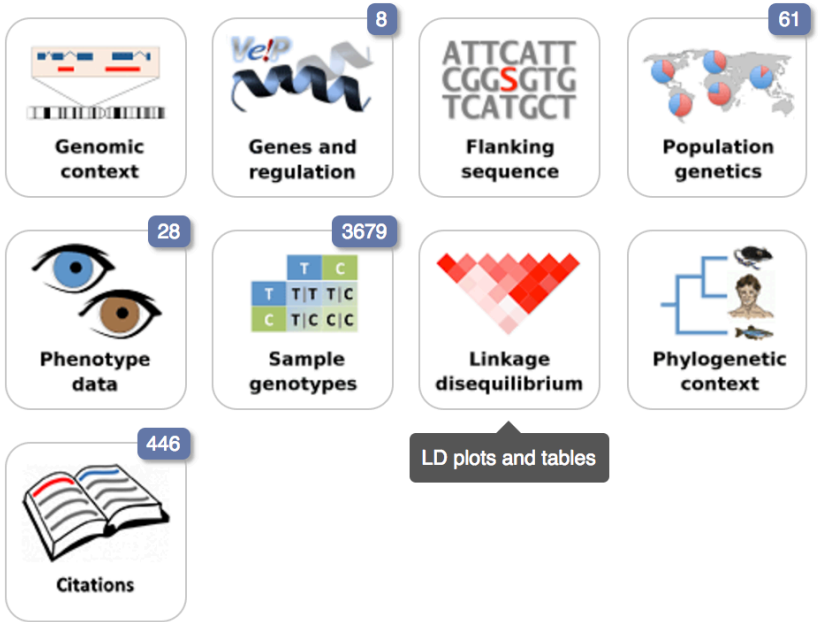
Variant displays
- Explore this variant
- Genomic context
  - Genes and regulation
  - Flanking sequence
- Population genetics
- Phenotype data
- Sample genotypes
- Linkage disequilibrium
- Phylogenetic context
- Citations

Configure this page
Custom tracks
Export data
Share this page
Bookmark this page

rs2476601 SNP

Most severe consequence	Missense variant | See all predicted consequences
Alleles	A/G | Ancestral: **G** | Ambiguity code: **R** | MAF: **0.03** (A)
Location	Chromosome **1:113834946** (forward strand) | View in location tab
Co-located variant	**HGMD-PUBLIC** CM041826
Evidence status	
Clinical significance	
HGVS names	This variant has **31** HGVS names - Show
Synonyms	This variant has **11** synonyms - Hide • **Archive dbSNP** rs117063937, rs60104027, rs52834763 • **ClinVar** RCV000009464, RCV000009462, RCV000009460, RCV000009463, RCV000009461 • **LSDB** 9440, 2009_August_001_234_PTPN22_600716_0001 • **Uniprot** VAR_022605
Genotyping chips	This variant has assays on **12** chips - Show
Original source	Variants (including SNPs and indels) imported from dbSNP (release 147) | View in dbSNP
About this variant	This variant overlaps 8 transcripts, has 3679 sample genotypes, is associated with 28 phenotypes and is mentioned in 446 citations.
Description from SNPedia	This SNP, located in the PTPN22 gene and also known as R620W, or 1858C>T, may influence Rheumatoid Arthritis and other autoimmune diseases, including but not limited to, multiple sclerosis, Crohn's disease, celiac disease and type-1 diabetes.... Show

Explore this variant



Phenotype Data

Significant association(s)

Show All entries | Show/hide columns | Filter

Phenotype, disease and trait	Source(s)	Mapped Terms	Ontology Accessions	Supporting evidence	Study	Clinical significance	Reported gene(s)	Associated allele	Statistics	Genomic Locations
Addison disease, susceptibility to	ClinVar	Primary adrenal insufficiency	HP:0008207	-	-	R ☆☆☆☆	AP4B1-AS1, PTPN22	-	-	View on Karyotype
Crohn's disease (time to surgery)	NHGRI-EBI GWAS catalog	Crohn's disease	EFO:0000384	-	PMID:26192919	-	NR	-	p-value: $4.00e^{-17}$	View on Karyotype
Crohn's disease (time to surgery)	NHGRI-EBI GWAS catalog	Crohn's disease	EFO:0000384	-	PMID:21102463	-	PTPN22	G	p-value: $4.00e^{-9}$ odds ratio: 1.26	View on Karyotype
Crohn's disease (time to surgery)	NHGRI-EBI GWAS catalog	Crohn's disease	EFO:0000384	-	PMID:18587394	-	PTPN22	G	p-value: $1.00e^{-8}$ odds ratio: 1.31	View on Karyotype
Crohns Disease	IIBDGC	Crohn's disease	EFO:0000384	-	-	-	-	A	p-value: $4.50e^{-9}$	View on Karyotype
DIABETES MELLITUS, INSULIN-DEPENDENT, 19	Uniprot	type I diabetes mellitus	EFO:0001359	-	MIM:222100	-	PTPN22	-	-	View on Karyotype
DIABETES MELLITUS, INSULIN-DEPENDENT, SUSCEPTIBILITY TO	ClinVar	-	Orphanet:317445	-	-	R ☆☆☆☆	AP4B1-AS1, PTPN22	-	-	View on Karyotype
DIABETES MELLITUS, INSULIN-DEPENDENT, SUSCEPTIBILITY TO	OMIM	-	Orphanet:317445	-	MIM:600716	-	PTPN22	0001	-	View on Karyotype
HASHIMOTO THYROIDITIS, SUSCEPTIBILITY TO	ClinVar	Hashimoto thyroiditis, Hashimoto's thyroiditis	HP:0000872, EFO:0003779	-	-	R ☆☆☆☆	AP4B1-AS1, PTPN22	-	-	View on Karyotype
Myasthenia gravis	NHGRI-EBI GWAS catalog	Myasthenia gravis, immune system disease	EFO:0000540, EFO:0004991, Orphanet:589	-	PMID:23055271	-	PTPN22	A	p-value: $8.00e^{-10}$ odds ratio: 1.71	View on Karyotype
RHEUMATOID ARTHRITIS	ClinVar	7 terms ⊞	7 accessions ⊞	-	-	R ☆☆☆☆	AP4B1-AS1, PTPN22	-	-	View on Karyotype
RHEUMATOID ARTHRITIS	NHGRI-EBI GWAS catalog	7 terms ⊞	7 accessions ⊞	-	PMID:20453842	-	PTPN22	A	p-value: $9.00e^{-74}$ odds ratio: 1.94	View on Karyotype
RHEUMATOID ARTHRITIS	NHGRI-EBI GWAS catalog	7 terms ⊞	7 accessions ⊞	-	PMID:24390342	-	PTPN22	A	p-value: $9.00e^{-170}$ odds 1.8	View on Karyotype

rs4988235 SNP

Most severe consequence	Intron variant | See all predicted consequences
Alleles	G/A/C | Ancestral: **G** | Ambiguity code: **V** | MAF: **0.16** (A)
Location	Chromosome **2:135851076** (forward strand) | View in location tab
Co-located variant	**HGMD-PUBLIC** CR024269
Evidence status	
Clinical significance	
HGVS names	This variant has **9** HGVS names - Show
Synonyms	**ClinVar** RCV000008124
Genotyping chips	This variant has assays on **7** chips - Show
Original source	Variants (including SNPs and indels) imported from dbSNP (release 147) | View in dbSNP
About this variant	This variant overlaps 3 transcripts, has 3271 sample genotypes, is associated with 3 phenotypes and is mentioned in 80 citations.
Description from SNPedia	Also known as "C/T(-13910)", and located in the MCM6 gene but with influence on the lactase LCT gene, **rs4988235** is one of two SNPs that is associated with the primary haplotype associated with hypolactasia, more commonly known as lactose intolerance in European Caucasian populations. [PMID:11788828], [PMID:15114531]... Show

Phenotype Data

Significant association(s)

Show/hide columns | Filter

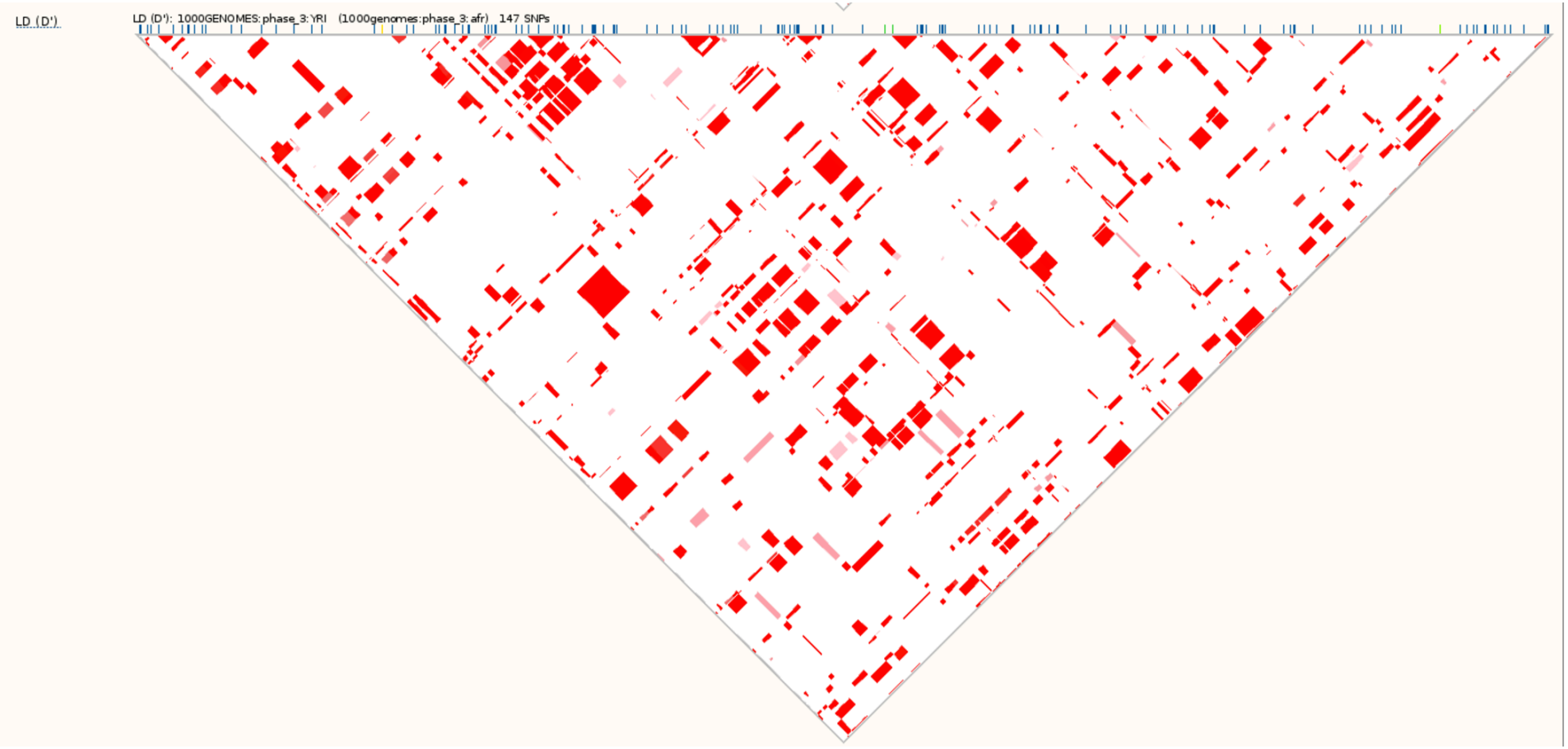
Phenotype, disease and trait	Source(s)	Mapped Terms	Ontology Accessions	Study	Clinical significance	Reported gene(s)	Associated allele	Statistics	Genomic Locations
BODY MASS INDEX	NHGRI-EBI GWAS catalog	body mass index, longitudinal BMI measurement	EFO:0004340, EFO:0005937	PMID:25673413	-	MCM6	A	**p-value:** $5.00e^{-6}$ **beta coefficient:** 0.016 kg/m2 increase	View on Karyotype
LACTASE PERSISTENCE	ClinVar	Autosomal recessive inheritance, Lactose intolerance, abnormality of metabolism/homeostasis, diarrhea	HP:0000007, HP:0001939, HP:0002014, HP:0004789	-		MCM6	A	-	View on Karyotype
LACTASE PERSISTENCE	OMIM	Autosomal recessive inheritance, Lactose intolerance, abnormality of metabolism/homeostasis, diarrhea	HP:0000007, HP:0001939, HP:0002014, HP:0004789	MIM:601806	-	MCM6	0001	-	View on Karyotype

Links to linkage disequilibrium data by population

Show All entries | Show/hide columns | Filter

Population	Description	LD Manhattan plot	Variants in high LD	LD plot	
African					
1000GENOMES:phase_3:**ACB**	African Caribbean in Barbados	View plot	Show	View plot	View table
1000GENOMES:phase_3:**ASW**	African Ancestry in Southwest US	View plot	Show	View plot	View table
1000GENOMES:phase_3:**ESN**	Esan in Nigeria	View plot	Show	View plot	View table
1000GENOMES:phase_3:**GWD**	Gambian in Western Division, The... (more)	View plot	Show	View plot	View table
1000GENOMES:phase_3:**LWK**	Luhya in Webuye, Kenya	View plot	Show	View plot	View table
1000GENOMES:phase_3:**MSL**	Mende in Sierra Leone	View plot	Show	View plot	View table
1000GENOMES:phase_3:**YRI**	Yoruba in Ibadan, Nigeria	View plot	Show	View plot	View table
American					
1000GENOMES:phase_3:**CLM**	Colombian in Medellin, Colombia	View plot	Show	View plot	View table
1000GENOMES:phase_3:**MXL**	Mexican Ancestry in Los Angeles... (more)	View plot	Show	View plot	View table
1000GENOMES:phase_3:**PEL**	Peruvian in Lima, Peru	View plot	Show	View plot	View table
1000GENOMES:phase_3:**PUR**	Puerto Rican in Puerto Rico	View plot	Show	View plot	View table
East Asian					
1000GENOMES:phase_3:**CDX**	Chinese Dai in Xishuangbanna, China	View plot	Show	View plot	View table
1000GENOMES:phase_3:**CHB**	Han Chinese in Bejing, China	View plot	Show	View plot	View table
1000GENOMES:phase_3:**CHS**	Southern Han Chinese, China	View plot	Show	View plot	View table
1000GENOMES:phase_3:**JPT**	Japanese in Tokyo, Japan	View plot	Show	View plot	View table
1000GENOMES:phase_3:**KHV**	Kinh in Ho Chi Minh City, Vietnam	View plot	Show	View plot	View table
European					
1000GENOMES:phase_3:**CEU**	Utah residents with Northern and... (more)	View plot	Show	View plot	View table

LD in Yorubans over a 20kb region around the LCT SNP



LD in British people over a 20kb region around the LCT SNP