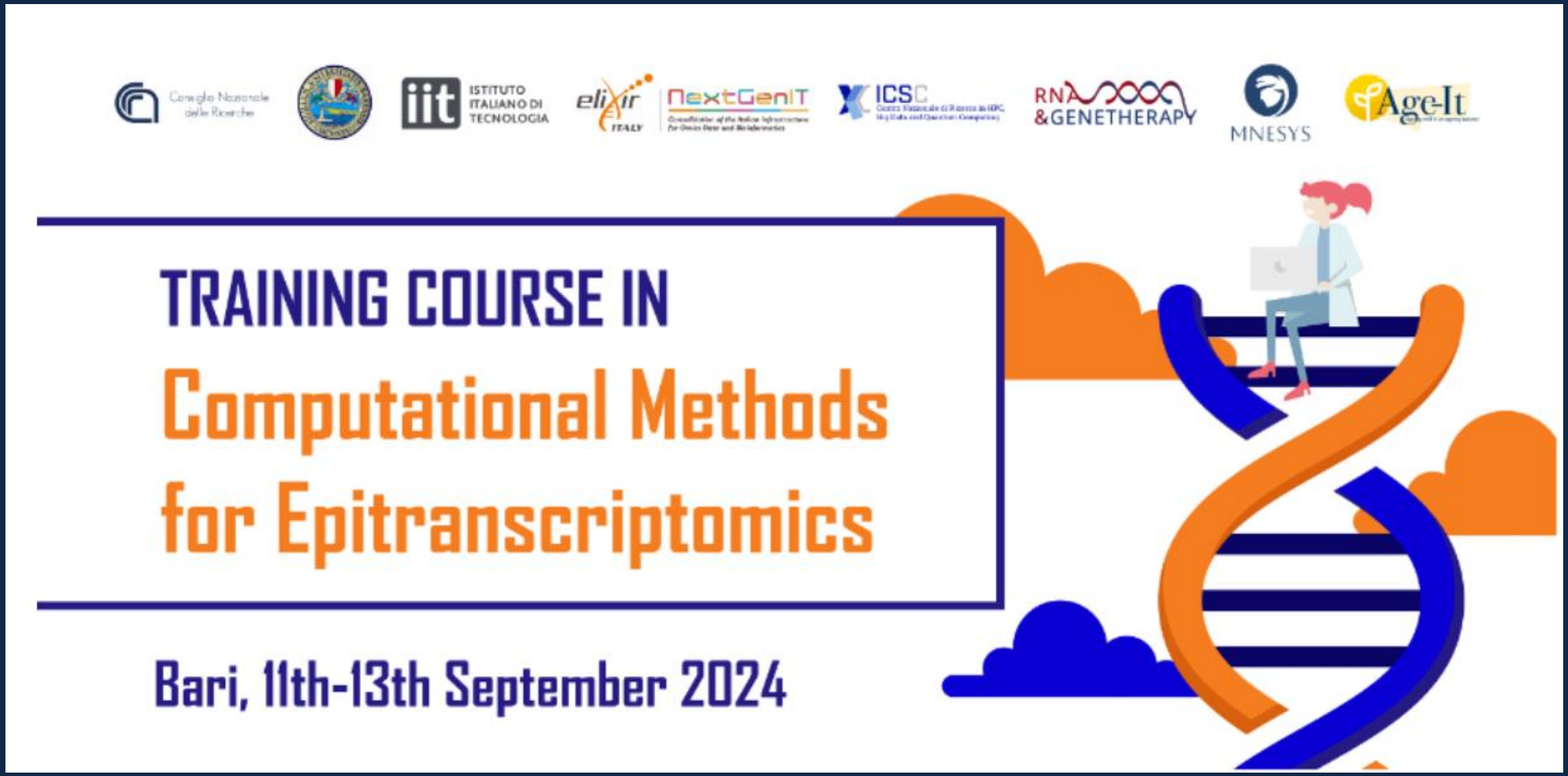
Computational approaches for detection and quantification of A-to-I RNA-editing

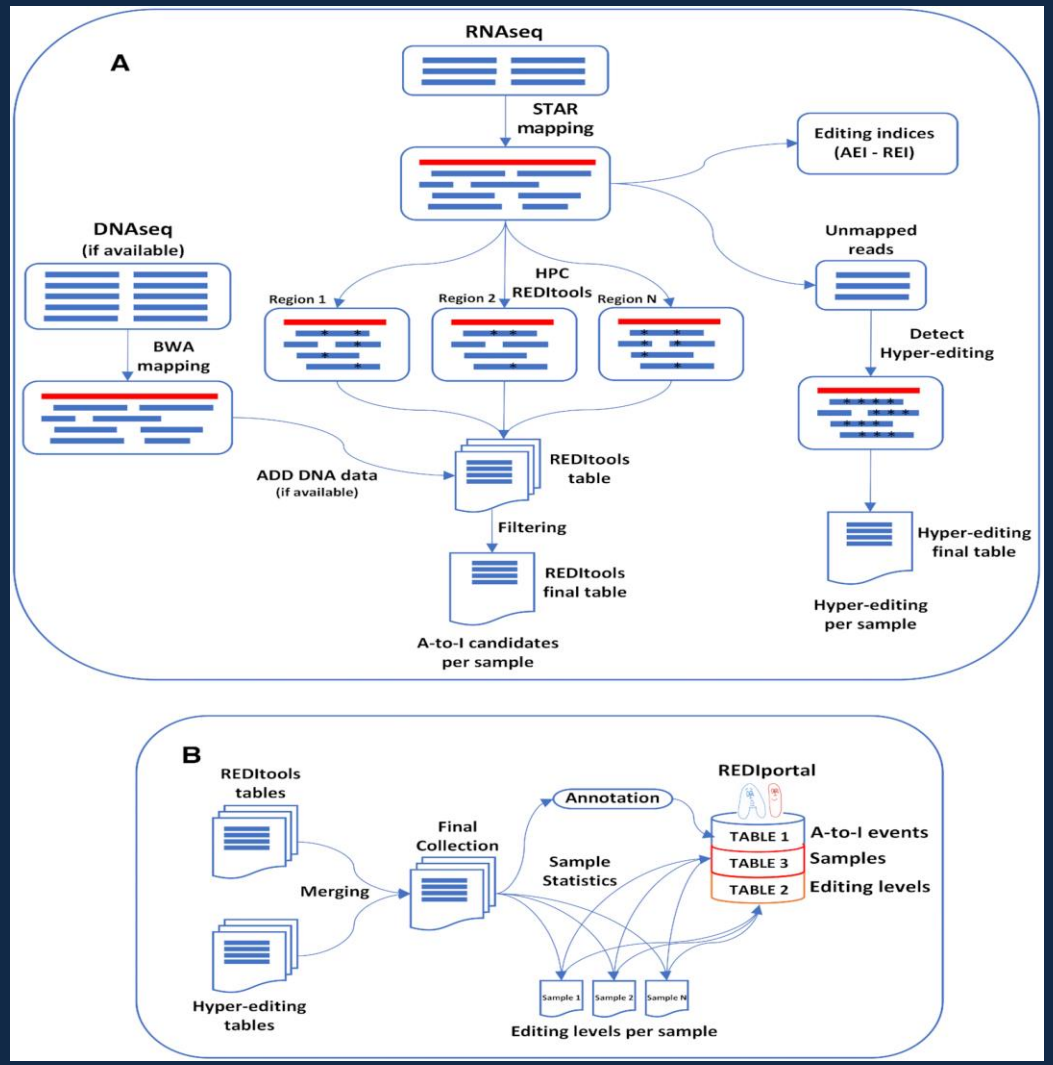


Dr. Domenico Alessandro Silvestris PhD RTD-A
Dipartimento di Bioscienze, Biotecnologie e Ambiente (DBBA), UNIBA email: domenico.silvestris@uniba.it

REDIportal

An ATLAS of A-to-I RNA editing events in human and other organisms

(http://srv00.recas.ba.infn.it/atlas/)



Data processing and database construction. (A) RNAseq data in fastq format are aligned on the human genome by STAR and converted in BAM files. In parallel, if DNAseq reads are available, are aligned on the same genome version by BWA. RNAseq BAM files are analyzed by HPC REDItools and the editing calling is distribute across different computing nodes, each working on a given genomic region. Resulting REDItools tables undergo to further filtering steps before the generation of the final table of A-to-I candidates. RNAseq unmapped reads are re-analyzed to detect hyper-edited reads and provide a list of hyper-editing sites per sample. RNAseq BAM files are further used to calculate the AEI and REI indices. (B) REDItools table and hyper-editing tables are merged in the final REDIportal collection. All events are annotated and stored in the MySQL TABLE1. They are also used to interrogate all RNAseq data to recover RNA editing levels and populate the MySQL TABLE2. Main RNA editing statistics per sample are also computed and collected in the MySQL TABLE3

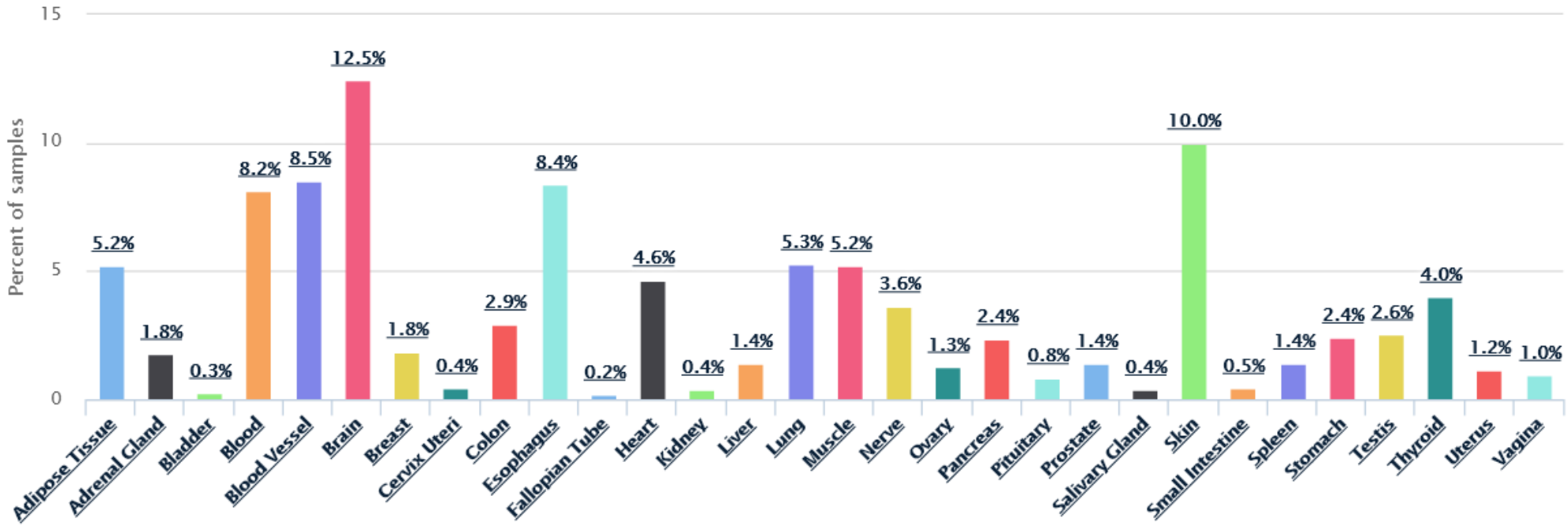
Welcome To REDIportal V2.0

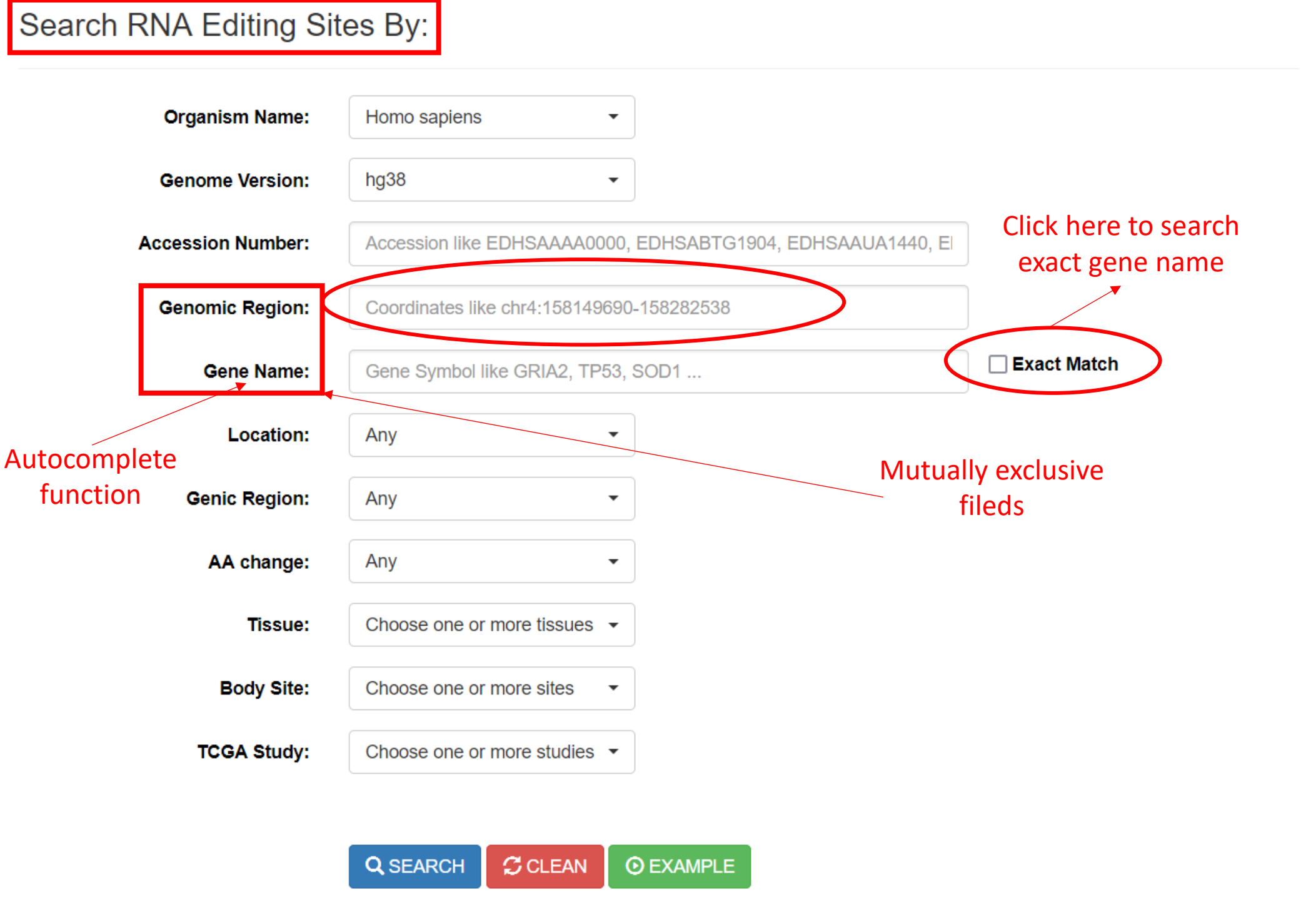
The largest RNA editing resource for human and other organisms V2.0

RNA editing is a relevant epitranscriptomic phenomenon by which primary RNAs are modified by base substitutions, insertions and/or deletions. In humans and other mammals, it mainly involves the deamination of adenosines to inosines by the ADAR family of enzymes acting on double RNA strands. A-to-I RNA editing has a plethora of biological effects and its deregulation has been linked to several human disorders. To better investigate A-to-I editing in eukaryotes, we have updated our already rich REDIportal catalogue (4.5 millions), raising its collection to about **16 millions of events**. REDIportal V2.0 is based on 9642 human RNAseq samples from **549 individuals (31 tissues and 54 body sites)** of the GTEx project. Now users can search at position level (by providing a genomic region or a gene name) and at sample level (by providing at a sample accession name) to have an overview of RNA editing per RNAseq experiment. REDIportal V2.0 implements a Gene View module to look at individual events in their genic context and hosts the CLAIRE resource (Pubmed). REDIportal V2.0 officially starts collecting RNA editing in non-human organisms. It currently stores 107,094 A-to-I mouse events from nascent RNAseq data (Pubmed).

Human Tissues loaded in the ATLAS

Click the columns to view body sites.




Search RNA Editing Sites By:
Organism Name:
Homo sapiens
Genome Version:
hg38
Accession Number:
Accession like EDHSAAAA0000, EDHSABTG1904, EDHSAAUA1440, EI
Genomic Region:
Coordinates like chr4:158149690-158282538
Gene Name:
Gene Symbol like GRIA2, TP53, SOD1 ...
Exact Match
Location:
Any
Genic Region:
Any
AA change:
Any
Tissue:
Choose one or more tissues
Body Site:
Choose one or more sites
TCGA Study:
Choose one or more studies
SEARCH
CLEAN
EXAMPLE
Click here to search exact gene name
Autocomplete function
Mutually exclusive fileds

Show 10 rows | Column visibility | Download

		Accession	Chr	Position	Ref	Ed	Strand	Location	Repeats	Gene	Region	EditedIn	FurtherAnn	ExFun	ProtSupp
	E U	EDHSABTG1827	chr4	157336723	A	G	+	NONREP	-/-	GRIA2	exonic	2280 1282	coding non-coding	Nonsyn	pride
	E U	EDHSABTG1828	chr4	157336726	A	G	+	NONREP	-/-	GRIA2	exonic	655 700	coding non-coding	Nonsyn	pride
	E U	EDHSABTG1829	chr4	157336727	A	G	+	NONREP	-/-	GRIA2	exonic	2014 1111	coding non-coding	Syn	pride
	E U	EDHSABTG1903	chr4	157360141	A	G	+	NONREP	-/-	GRIA2	exonic	881 346	coding non-coding	Syn	pride
	E U	EDHSABTG1904	chr4	157360142	A	G	+	NONREP	-/-	GRIA2	exonic	2491 1364	coding non-coding	Nonsyn	pride
		Accession	Chr	Position	Ref	Ed	Strand	Location	Repeats	Gene	Region	EditedIn	FurtherAnn	ExFun	ProtSupp

Showing 1 to 5 of 5 entries (filtered from 15,680,833 total entries)

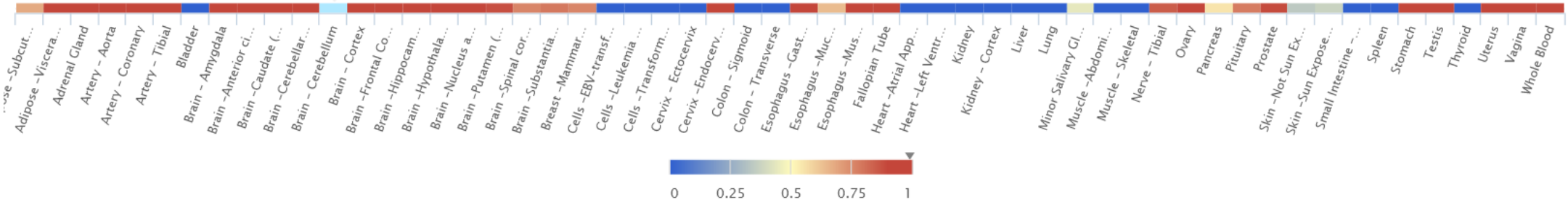
Previous 1 Next

GTEx Heat-Map | GTEx Box Plot | TCGA Heat-Map | TCGA Box Plot | Other info | Alternative Annotations | Editing Details

RNA Editing Heat-Map

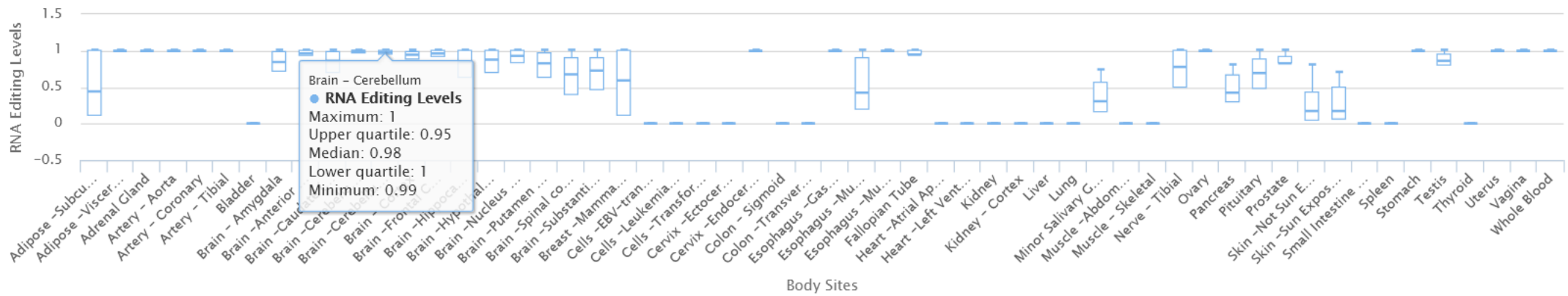
Level Variation Across Tissues

Brain – Cerebellum
Mean Editing: **0.99**



GTEx Heat-Map | GTEx Box Plot | TCGA Heat-Map | TCGA Box Plot | Other info | Alternative Annotations | Editing Details

Box Plot Of RNA Editing values



GTEx Heat-Map | GTEx Box Plot | TCGA Heat-Map | TCGA Box Plot | Other info | Alternative Annotations | Editing Details

Other annotations

Gene Id	Gene Type	Phast	dbSNP	Known in past DBs
ENSG00000120251	protein_coding	803	-	Yes

Site currently under maintenance.

Drawbacks can occur, sorry for that.

Home | Search | CLAIRE | JBrowse | Publications | Download | Contact us

Show 10 rows | Column visibility | Download

		Accession	Chr	Position	Ref	EditedIn	FurtherAnn	ExFun	ProtSupp
	E U	EDHSABTG1827	chr4	157336723	A	2280 1282	coding non-coding	Nonsyn	pride
	E U	EDHSABTG1828	chr4	157336726	A	655 700	coding non-coding	Nonsyn	pride
	E U	EDHSABTG1829	chr4	157336727	A	2014 1111	coding non-coding	Syn	pride

GTEx Heat-Map | GTEx Box Plot | TCGA Heat-Map | TCGA Box Plot | Other info

GTEx			TCGA	
N. Samples	**N. Tissues**	**N. Body Sites**	**N. Samples**	**N. Studies**
2280	20	38	1282	26

View GTEx Editing Details | View TCGA Editi

Gencode Annotation

Gene	Transcript	ExonNumber	cDNAchange	ProteinChange
GRIA2	ENST00000703718.1	exon10	c.A1580G	p.Q527R
GRIA2	ENST00000703767.1	exon10	c.A1439G	p.Q480R
GRIA2	ENST00000264426.14	exon11	c.A1820G	p.Q607R
GRIA2	ENST00000296526.12	exon11	c.A1820G	p.Q607R
GRIA2	ENST00000393815.6	exon11	c.A1679G	p.Q560R
GRIA2	ENST00000507898.5	exon11	c.A1679G	p.Q560R
GRIA2	ENST00000645636.1	exon11	c.A1820G	p.Q607R
GRIA2	ENST00000703717.1	exon11	c.A1679G	p.Q560R
GRIA2	ENST00000703752.1	exon11	c.A1679G	p.Q560R
GRIA2	ENST00000703765.1	exon11	c.A1679G	p.Q560R
GRIA2	ENST00000703760.1	exon12	c.A1847G	p.Q616R

Close

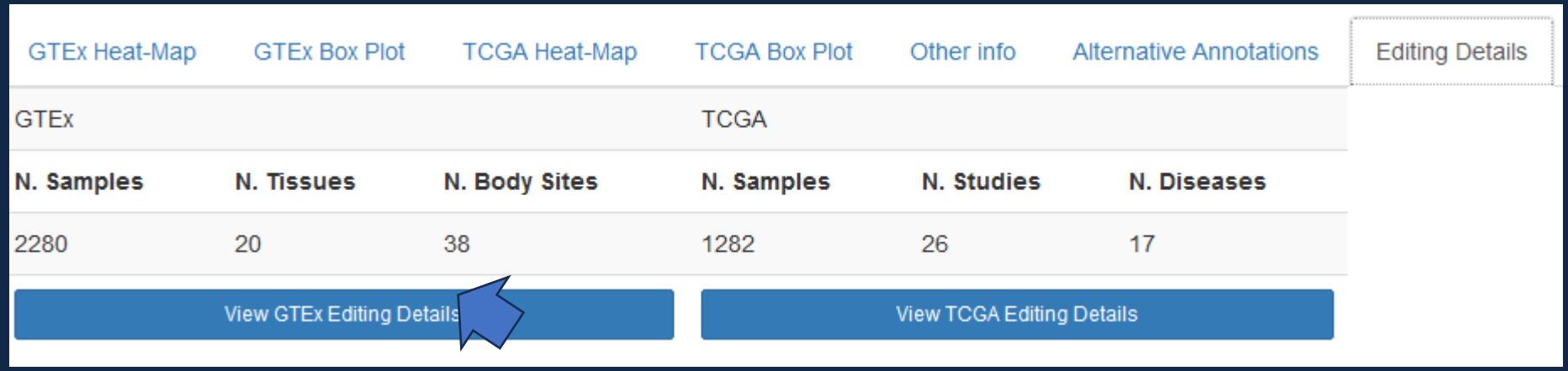
GTEx Heat-Map | GTEx Box Plot | TCGA Heat-Map | TCGA Box Plot | Other info | Alternative Annotations | Editing Details

RefSeq Annotation

Gene	Region	ExFun	Transcript	ExonNumber	cDNAchange	ProteinChange
GRIA2	exonic	nonsynonymous SNV	NM_000826	exon11	c.A1820G	p.Q607R
GRIA2	exonic	nonsynonymous SNV	NM_001083619	exon11	c.A1820G	p.Q607R
GRIA2	exonic	nonsynonymous SNV	NM_001083620	exon11	c.A1679G	p.Q560R
GRIA2	exonic	nonsynonymous SNV	NM_001379000	exon11	c.A1679G	p.Q560R
GRIA2	exonic	nonsynonymous SNV	NM_001379001	exon11	c.A1679G	p.Q560R

UCSC Annotation

Gene	Region	ExFun	Transcript	ExonNumber	cDNAchange	ProteinChange
ENST00000264426.14;GRIA2	exonic	nonsynonymous SNV	ENST00000264426.14	exon11	c.A1820G	p.Q607R
ENST00000264426.14;GRIA2	exonic	nonsynonymous SNV	ENST00000296526.12	exon11	c.A1820G	p.Q607R



Show 10 rows | Export in | Column visibility | Filter Editing Levels | Reset Filters

Search:

RNAseq Run	WGS Run	Tissue	BodySite	n.As	n.Gs	EditingFreq	gCoverage	gFreq
SRR1068687	SRR8233673	Esophagus	Esophagus - Gastroesophageal Junction	0	6	1.00	41	0.00
SRR1069048	SRR8233669	Skin	Skin - Not Sun Exposed (Suprapubic)	8	4	0.33	50	0.00
SRR1069166	SRR8233654	Stomach	Stomach	0	20	1.00	44	0.00
SRR1069188	NA	Brain	Brain - Cerebellum	1	32	0.97	0	0.00
SRR1069376	SRR8233672	Blood Vessel	Artery - Aorta	0	39	1.00	37	0.00
SRR1069714	SRR8233590	Stomach	Stomach	0	14	1.00	41	0.00
SRR1070086	SRR8233479	Nerve	Nerve - Tibial	2	14	0.88	43	0.00
SRR1070111	SRR8233557	Blood Vessel	Artery - Aorta	0	21	1.00	44	0.00
SRR1070159	SRR8233583	Nerve	Nerve - Tibial	2	5	0.71	39	0.00
SRR1070260	SRR8233663	Breast	Breast - Mammary Tissue	6	2	0.25	31	0.00
Select	Select	Select	Select	n.As	n.Gs	EditingFreq	gCoverage	gFreq

Showing 1 to 10 of 2,280 entries

Previous 1 2 3 4 5 … 228 Next

View in Ensembl GRCH38 Genome Browser

Show 10 | Column visibility | Download

		Accession	Chr	Position	Ref	Ed	Strand	Location	Repeats	Gene	Region	EditedIn	FurtherAnn	ExFun	ProtSupp
	E U	EDHSABTG1827	chr4	157336723	A	G	+	NONREP	-/-	GRIA2	exonic	2280 1282	coding non-coding	Nonsyn	pride

Logi

e!Ensembl BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Search all species…

Human (GRCh38.p14) ▾

Location: 4:157,336,723-157,336,723

Location-based displays

- Whole genome
- Chromosome summary
- Region overview
- **Region in detail**
- Comparative Genomics
  - Synteny
  - Alignments (image)
  - Alignments (text)
  - Region Comparison
- Genetic Variation
  - Variant table
  - Resequencing
  - Linkage Data
- Markers
- Other genome browsers
  - UCSC
  - NCBI
  - Ensembl GRCh37

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

Chromosome 4: 157,336,723-157,336,723

Add/remove tracks | Share | Export image | Reset configuration

Region in detail

Add/remove tracks | Share | Resize image | Export image | Reset configuration | Reset track order | Switch image

Location: 4:157336723-157336723 Go Gene: Go

Add/remove tracks | Custom tracks | Share | Resize image | Export image | Reset configuration | Reset track order | Drag/Select:

Show 10 ro… | Column visibility | Download

	View in UCSC Genome Browser	Accession	Chr	Position	Ref	Ed	Strand	Location	Repeats	Gene	Region	EditedIn	FurtherAnn	ExFun	ProtSupp
	E U	EDHSABTG1827	chr4	157336723	A	G	+	NONREP	-/-	GRIA2	exonic	2280 1282	coding non-coding	Nonsyn	pride

Genomes | Genome Browser | Tools | Mirrors | Downloads | My Data | View | Help | About Us

UCSC Genome Browser on Human (GRCh38/hg38)

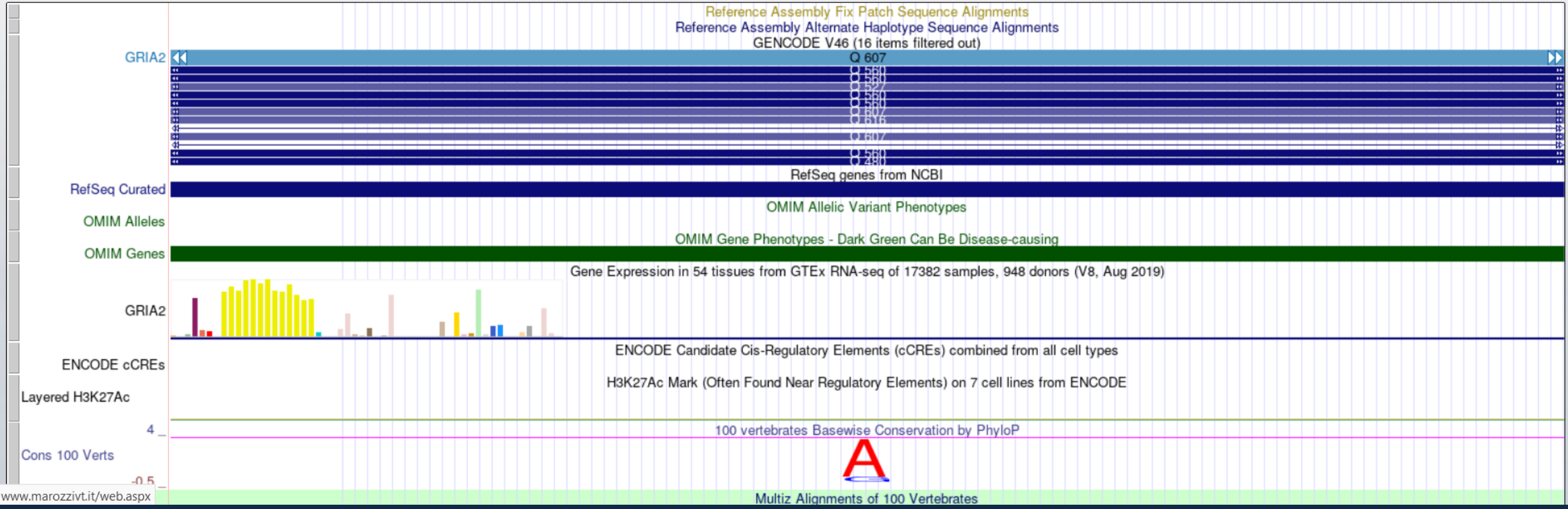
Move <<< << < > >> >>> Zoom in 1.5x 3x 10x Base Zoom out 1.5x 3x 10x 100x

A guided tutorial is available for new users: Start tutorial

Close | Don't show again

Multi-region | chr4:157,336,723-157,336,723 | 1 bp. | gene, chromosome range, search terms, help pages, see exampl | Search | Examples

chr4 (q32.1) 16.3 16.1 p15.2 4p15.1 4p14 13 p12 4q12 4q13.1 q13.3 q22.1 4q24 4q25 4q26 27 28.1 4q28.3 31.21 31.3 q32.1 q32.3 34.1 q34.3



www.marozzivt.it/web.aspx

Search RNA Editing By Sample:

Organism: Homo sapiens

Sample Name: Sample Name like SRR1077183 for GTEx or TCGA-IZ-8195-01A-31R-2404-07 for TCGA ...

Data Source: Any

Data Status: Any

Data Type: Any

GTEx Tissue: Choose one or more tissues

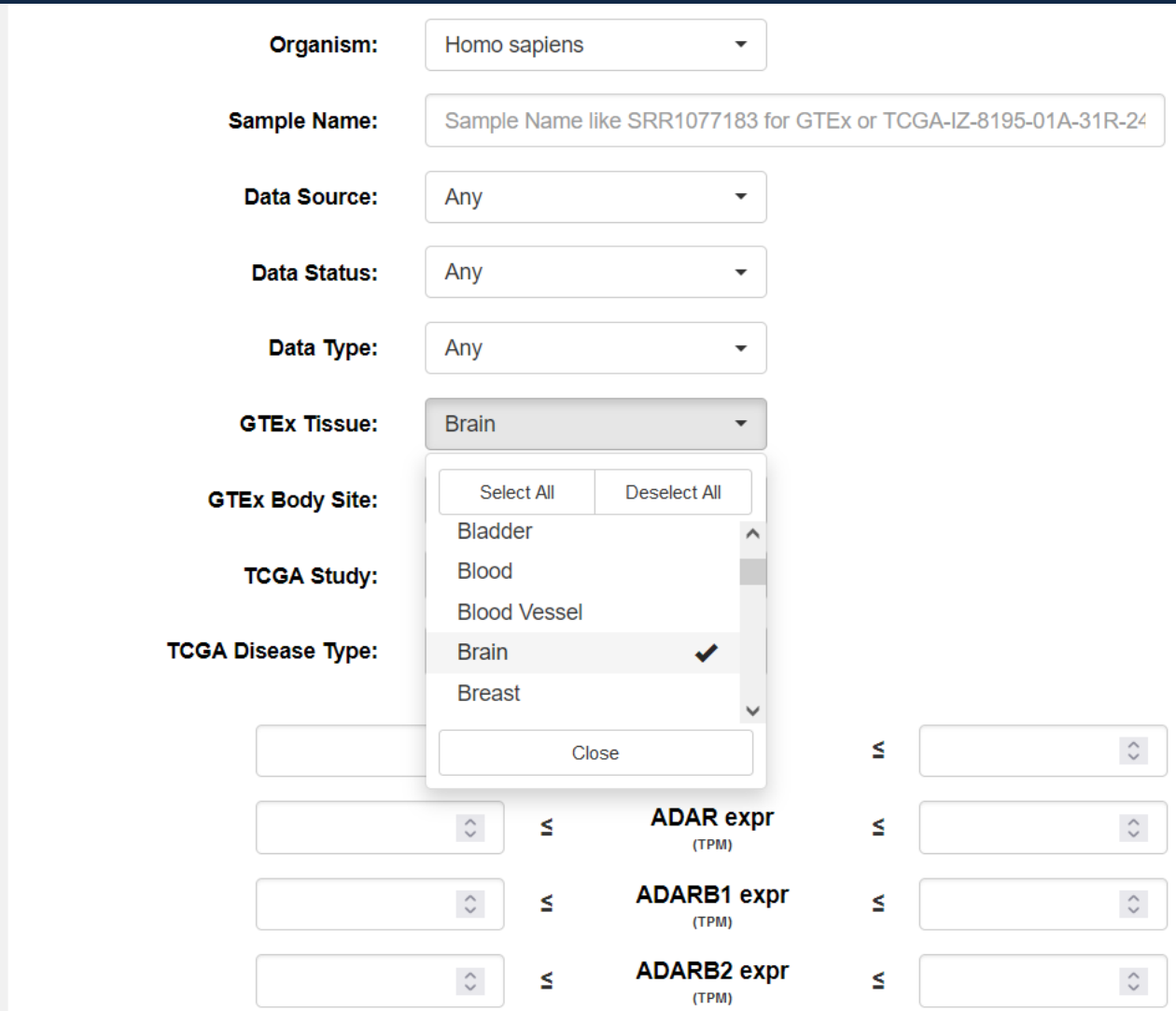
GTEx Body Site: Choose one or more sites

TCGA Study: Choose one or more studies

TCGA Disease Type: Choose one or more studies

	≤	**AEI** (%)	≤	
	≤	**ADAR expr** (TPM)	≤	
	≤	**ADARB1 expr** (TPM)	≤	
	≤	**ADARB2 expr** (TPM)	≤	

SEARCH CLEAN EXAMPLE


Organism:
Homo sapiens
Sample Name:
Sample Name like SRR1077183 for GTEx or TCGA-IZ-8195-01A-31R-24
Data Source:
Any
Data Status:
Any
Data Type:
Any
GTEx Tissue:
Brain
GTEx Body Site:
TCGA Study:
TCGA Disease Type:
Select All
Deselect All
Bladder
Blood
Blood Vessel
Brain
Breast
Close
≤
ADAR expr
(TPM)
ADARB1 expr
(TPM)
ADARB2 expr
(TPM)

Show 10 rows | CSV | Excel | PDF | Select all | Deselect all

Search:

		Sample	WGS/WES	Source	Organism	Events	Hyper	Body Site/Study	Status	Type	AEI	REI	ADAR	ADARB1	ADARB2
		SRR1069188	NA	G	Hs	91267	74935	Brain - Cerebellum	normal	B	2.501	4.222	95.077	59.382	7.696
		SRR1070986	SRR8233641	G	Hs	86257	53929	Brain - Nucleus accumbens (basal ganglia)	normal	B	2.116	2.077	68.416	12.565	10.562
		SRR1071289	SRR8233649	G	Hs	12684	12440	Brain - Hypothalamus	normal	B	1.766	1.326	9.100	1.050	3.966
		SRR1071880	SRR8233643	G	Hs	66042	40439	Brain - Hypothalamus	normal	B	1.942	1.013	66.078	18.685	13.820
		SRR1072178	SRR8233659	G	Hs	86836	0	Brain - Caudate (basal ganglia)	normal	B	2.411	1.929	56.810	7.694	24.295
		SRR1072367	SRR8233659	G	Hs	52840	29987	Brain - Frontal Cortex (BA9)	normal	B	1.870	0.835	51.555	4.681	64.410
		SRR1072504	SRR8233641	G	Hs	92239	53140	Brain - Hypothalamus	normal	B	2.068	2.462	100.380	16.618	11.893
		SRR1072797	SRR8233650	G	Hs	55321	46360	Brain - Hypothalamus	normal	B	2.183	4.615	177.233	23.630	8.741
		SRR1073143	SRR8233650	G	Hs	51426	33528	Brain - Anterior cingulate cortex (BA24)	normal	B	2.064	2.895	79.414	22.118	9.168
		SRR1073755	SRR8233492	G	Hs	51633	42367	Brain - Spinal cord (cervical c-1)	normal	B	2.055	0.728	40.243	2.532	20.203
		Filter	WGS/WES	Source	Organism	Events	Hyper	Body Site/Study	Status	Type	AEI	REI	ADAR	ADARB1	ADARB2

Showing 1 to 10 of 1,409 entries

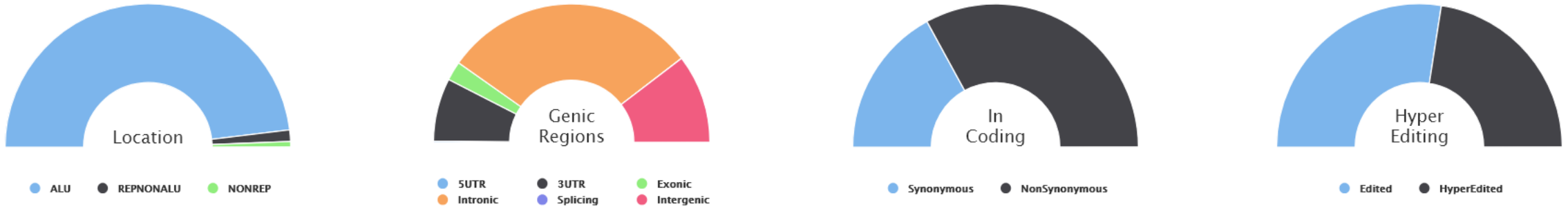
Previous 1 2 3 4 5 … 141 Next

Show 10 rows | CSV | Excel | PDF | Select all | Deselect all

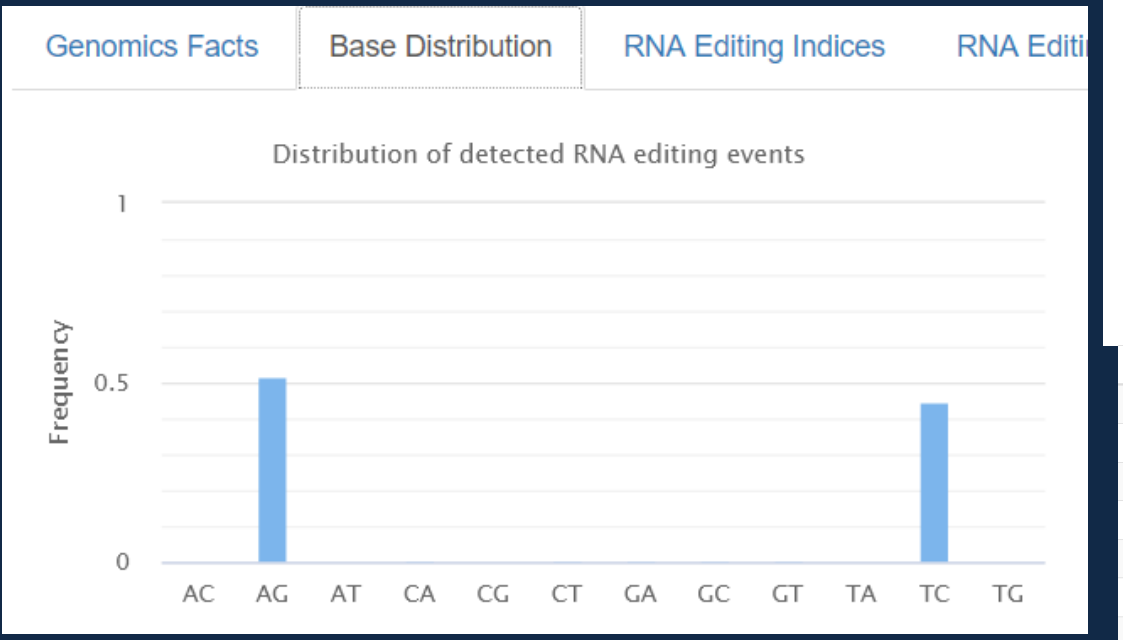
Search:

		Sample	WGS/WES	Source	Organism	Events	Hyper	Body Site/Study	Status	Type	AEI	REI	ADAR	ADARB1	ADARB2
				GTEx Project											
	☑	SRR1069188	NA	G	Hs	91267	74935	Brain - Cerebellum	normal	B	2.501	4.222	95.077	59.382	7.696

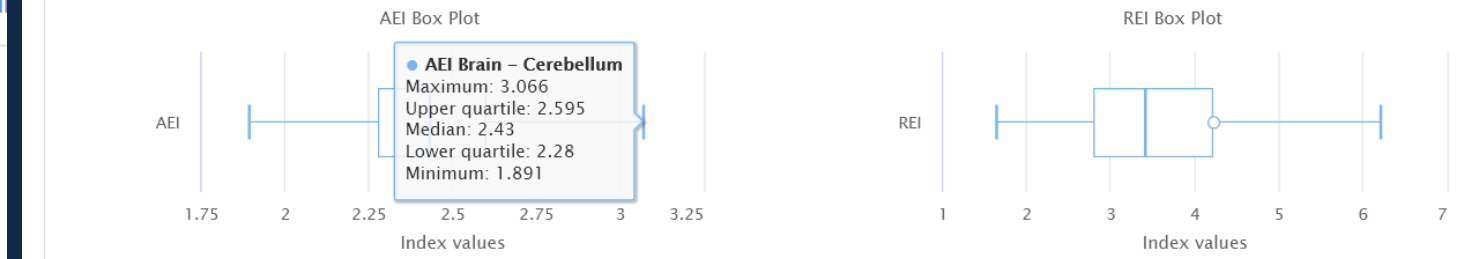
Genomics Facts | Base Distribution | RNA Editing Indices | RNA Editing Levels | Transcriptome Coverage



Genomics Facts | Base Distribution | RNA Editing Indices | RNA Editi

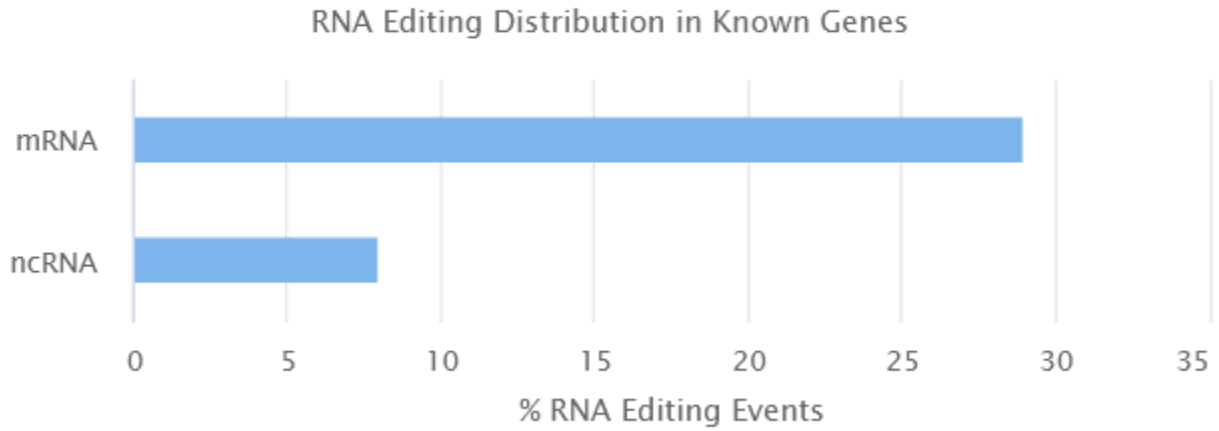


Genomics Facts | Base Distribution | RNA Editing Indices | RNA Editing Levels | Transcriptome Coverage

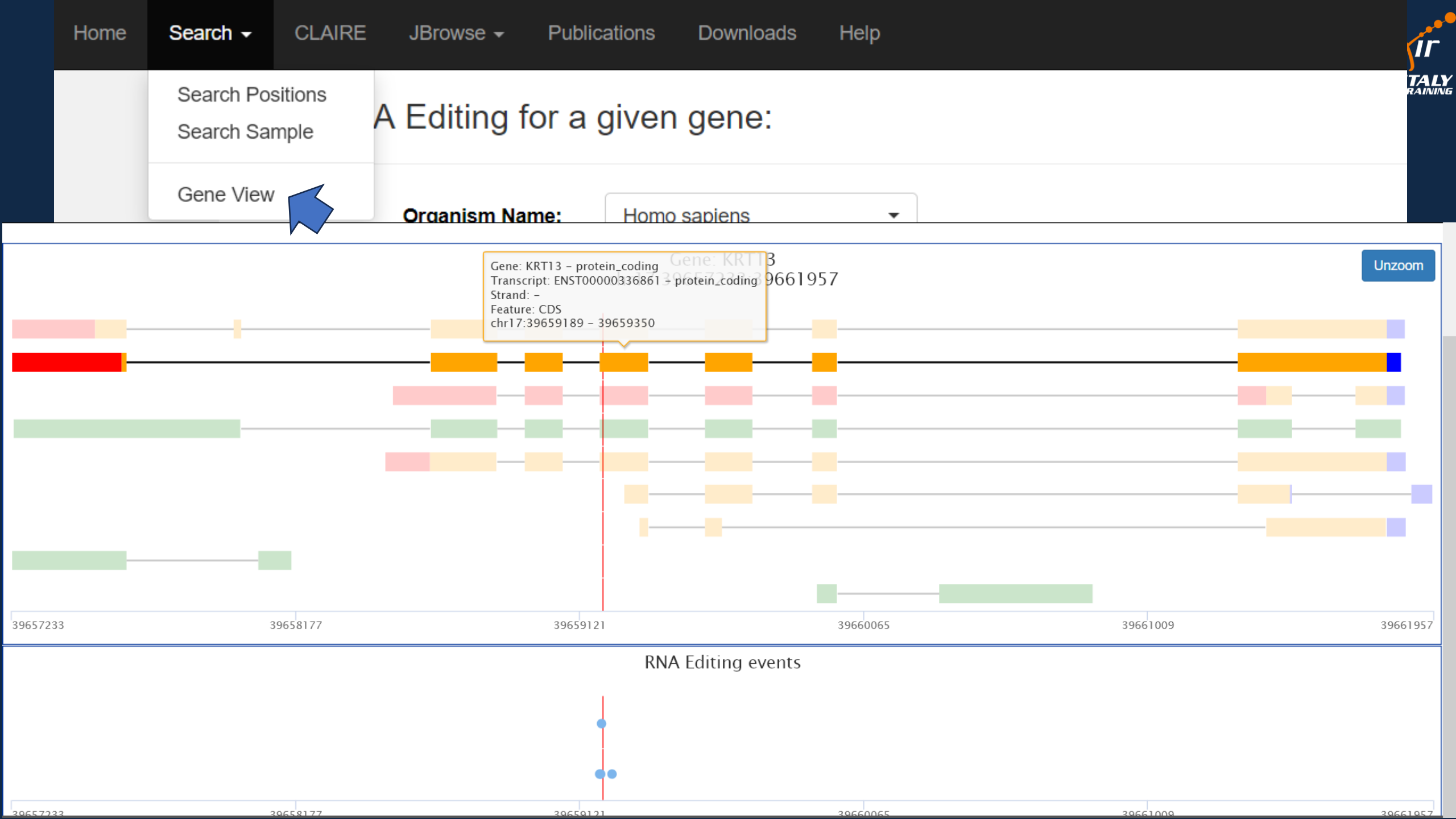


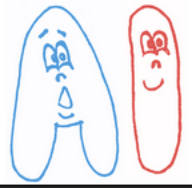
		Chr	Position	Coverage	EditingFreq	Gene	AAchange	PosAAchange	Sample
☐	V	chr1	934659	14	-1	HES4	QR	Q175R	SRR1069188
☐	V	chr1	949464	33	-1	ISG15	KR	K35R	SRR1069188
☐	V	chr1	1216628	22	-1	SCNN1D	TA	T10A	SRR1069188
☐	V	chr1	1219383	51	-1	SCNN1D	QR	Q28R	SRR1069188
☐	V	chr1	1334650	46	-1	CCNL2	SP	S13P	SRR1069188
☐	V	chr1	1425783	102	-1	ATAD3B	KR	K495R	SRR1069188
☐	V	chr1	1431061	74	0.01351	ATAD3B	YC	Y604C	SRR1069188
☐	V	chr1	1690549	2	-1	NADK	QR	Q189R	SRR1069188

Genomics Facts | Base Distribution | RNA Editing Indices | RNA Editing Levels | Transcriptome Coverage



	Gene_ID	Symbol	Type	Chr	Start	End	Strand	TPM	N_alu	N_events	Ed_Pri	Ed_Mat
V	ENSG00000171401	KRT13	protein_coding	chr17	39657233	39661957	-	32491.979	0	4	0.248	0.0
V	ENSG00000163209	SPRR3	protein_coding	chr1	152974223	152976332	+	29704.313	0	28	0.002	0.0
V	ENSG00000170477	KRT4	protein_coding	chr12	53200333	53207900	-	19985.308	1	4	0.008	0.0
V	ENSG00000105427	CNFN	protein_coding	chr19	42891171	42894449	-	6259.265	3	59	0.036	0.0
V	ENSG00000186081	KRT5	protein_coding	chr12	52908359	52914314	-	6029.393	0	3	0.001	0.0
V	ENSG00000140519	RHCG	protein_coding	chr15	90014629	90039820	-	5709.125	6	110	0.045	0.0
V	ENSG00000106211	HSPB1	protein_coding	chr7	75931990	75933612	+	4656.451	1	1	0.009	0.0
V	ENSG00000135046	ANXA1	protein_coding	chr9	75766781	75785309	+	4264.72	1	23	0.083	0.0


Home
Search
CLAIRE
JBrowse
Publications
Downloads
Help
Search Positions
Search Sample
Gene View
A Editing for a given gene:
Organism Name:
Homo sapiens
Gene: KRT13 – protein_coding
Transcript: ENST00000336861 – protein_coding
Strand: –
Feature: CDS
chr17:39659189 – 39659350
Unzoom
39657233
39658177
39659121
39660065
39661009
39661957
RNA Editing events



REDIportal

An ATLAS of A-to-I RNA editing events in human and other organisms

Home | Search | CLAIRE | JBrowse | Publications | **Downloads** | Help | Contact us

Download REDIportal Database

Annotated positions

Organism Name: Homo sapiens

Genome Version: hg19

- hg19
- hg38

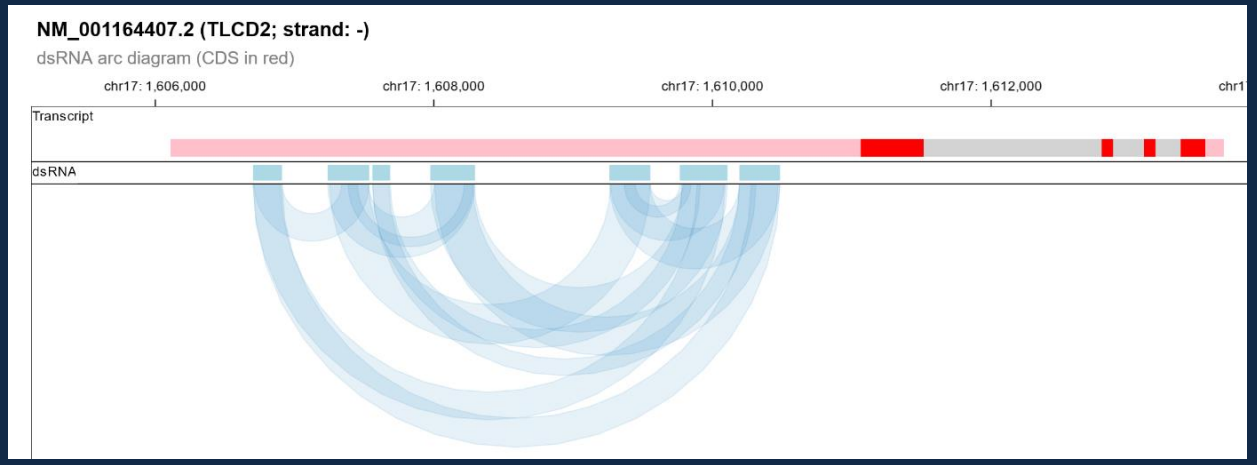
Copy this link for download in backgroud

Direct Download

Note For users searching for the previous version of REDIportal database:

- It can be downloaded from here

New currently on going features:



- TCGA samples editing levels
- Double strand RNA graphical view
- Proteomic support to editing from PRIDE
- probability scores from machine learning to support editing sites

THANK YOU
FOR YOUR ATTENTION!