Consiglio Nazionale delle Ricerche

ISTITUTO ITALIANO DI TECNOLOGIA

elixir ITALY

NextGenIT
Consolidation of the Italian Infrastructure for Omics Data and Bioinformatics

MUR
Centro Nazionale di Ricerca Sviluppo di terapia genica e farmaci con tecnologia a RNA

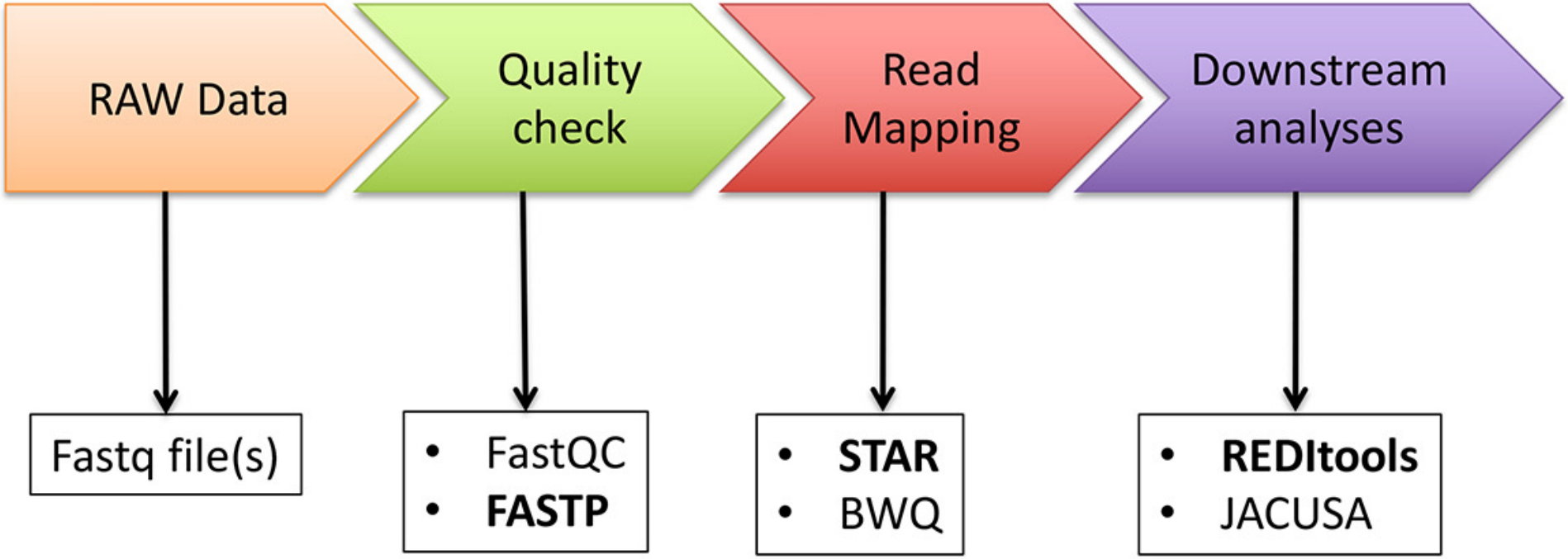
ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

TRAINING COURSE IN Computational Methods for Epitranscriptomics

Bari, 26th-28th April 2023

Profiling known RNA editing events in RNAseq experiments

Alessandro Silvestris PhD – RTDA UNIBA
asilvestris@aliceposta.it



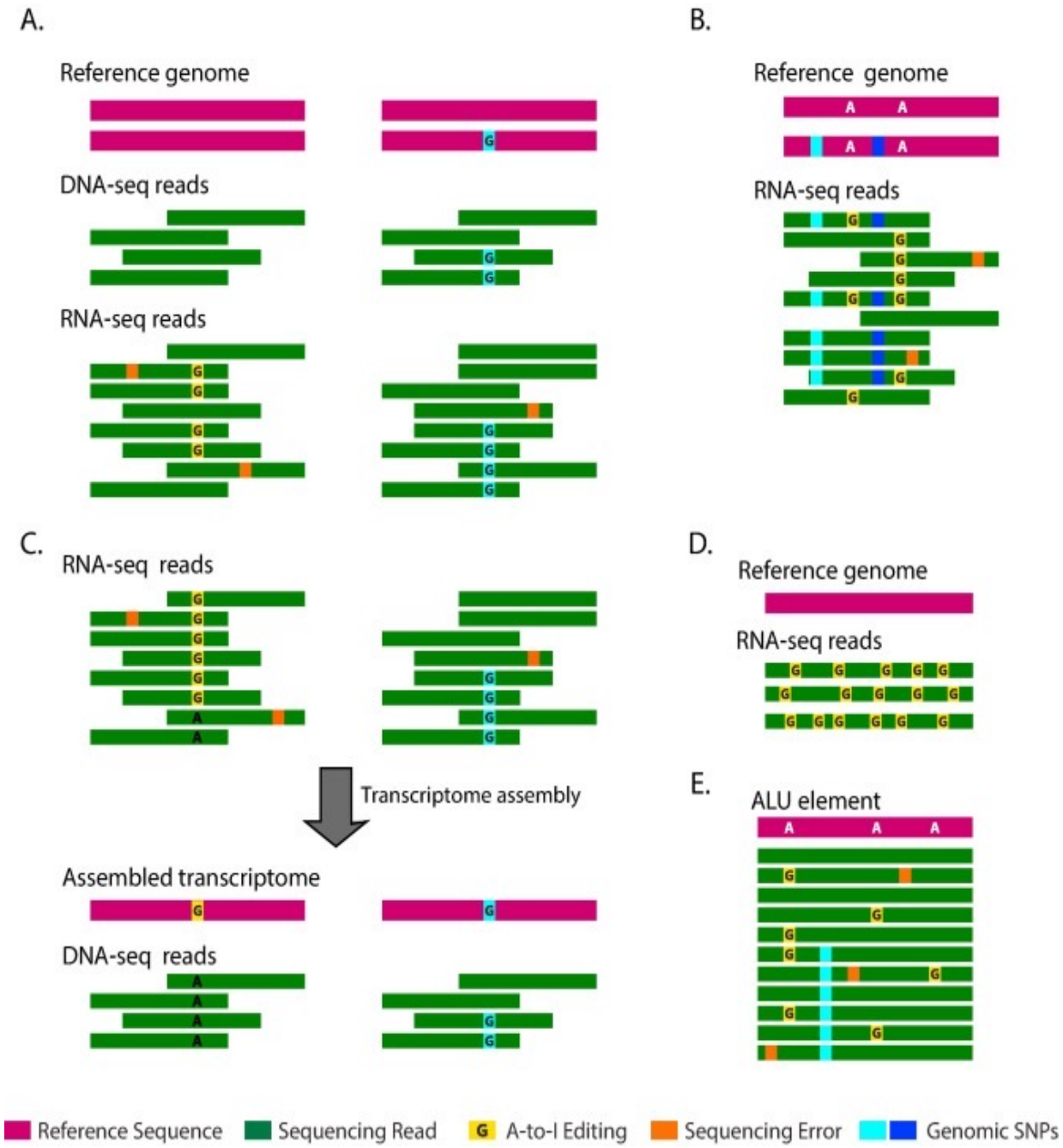
RNA editing detection and quantification

***De novo* approach**

RNA editing candidates can be detected using REDItools. There are two current versions: 1) REDItools 1.3 or 2) REDItools 2.0. REDItools2 is a faster re-implementation of REDItools1 for HPC clusters. Its serial version is about ten times faster than REDItools1.

"Known" approach

While the *de novo* approach provides a list of most likely editing candidates, the "known" approach focuses on a limited pool of known events in order to better investigate RNA editing dynamics in different experimental contexts. The "known" approach can be carried out using the REDItools package and a list of events from own data or from public databases such as DARNED, RADAR and REDIportal.



Quantifying RNA editing in deep transcriptome datasets

The quantification of RNA editing is important to compare values across samples and study its potential role in different experimental conditions or in human disorders.

Determine the fraction of edited transcripts of a site (editing levels) by dividing the number of the ‘G’-containing transcripts that map to the site, by the total number of transcripts mapped to the position.

For example, the editing levels of the leftmost editing site in Fig. E is 4/11 as we found evidence for editing in 4 reads out of 11 reads.

The accuracy of measuring the editing levels of a site depends on the site coverage in the RNA-seq dataset, which in turn is determined mainly by the sequencing depth and the expression levels of the transcript of interest. Unfortunately, sufficient coverage for each editing site is often not available in a typical RNA-seq. In order to overcome this limitation have been developed **metrics for unbiased RNA editing quantification in a sample**:

Overall editing level

The overall editing is defined as the total number of reads with G at all known editing positions over the number of all reads covering the positions without imposing specific sequencing coverage criteria. It can be calculated using REDItools tables obtained imposing loosing parameters.

ALU editing index

The Alu editing index (AEI) is a metric to quantify the global RNA editing activity of sample and is defined as the weighted average of editing events occurring in all Alu elements. The pipeline to calculate AEI is described in Roth et al. (2019) and available here.

Recoding index

The overall editing calculated at recoding positions residing in coding protein genes is named recoding index (REI). It has been initially described in Silvestris et al. (2019). This metric, used to investigate the activity of ADAR2, can be calculated using REDItools tables obtained imposing loosing parameters and a list of recoding sites from REDIportal.

How to calculate the Recoding Editing Index (REI)

Download REDIportal nonsynonymous editing annotation (1585 sites)

```
cd /home/student_X/RNAseq/DnaRna_
wget https://raw.githubusercontent.com/BioinfoUNIBA/QEdit/master/Example_files/nonsynonymous_table_NONREP_2BS.txt
```

Index your REDItools output table by tabix

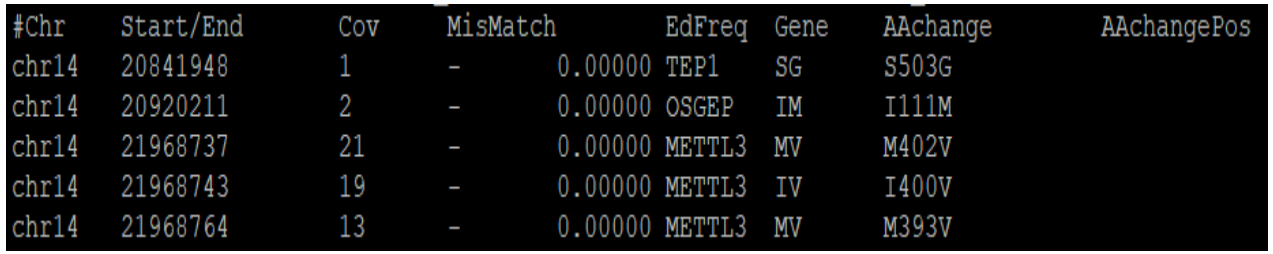
```
bgzip /home/Student_x/RNAseq/DnaRna_/outTable_X
```

<- must be obtained by running REDItools with coverage and minimal editing filters turned off

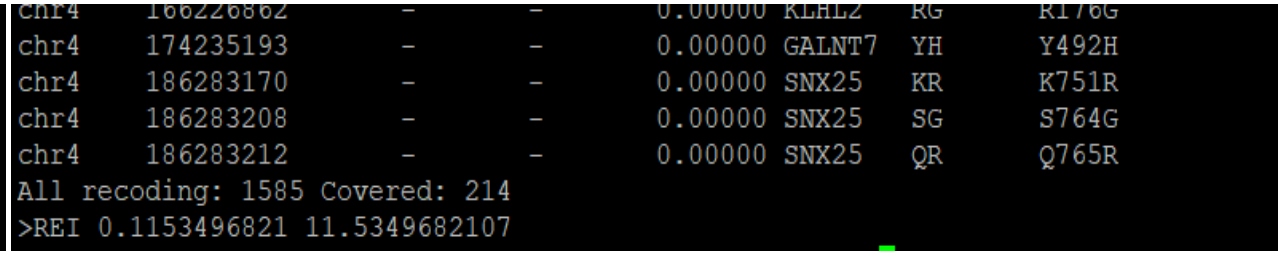
```
tabix -s 1 -b 2 -e 2 -c R outTable_X.gz
```

Run the GetREI.py script on a REDItools table

```
python /data/QEdit/scripts/GetREI.py -i outTable_X.gz -r nonsynonymous_table_NONREP_2BS.txt
```



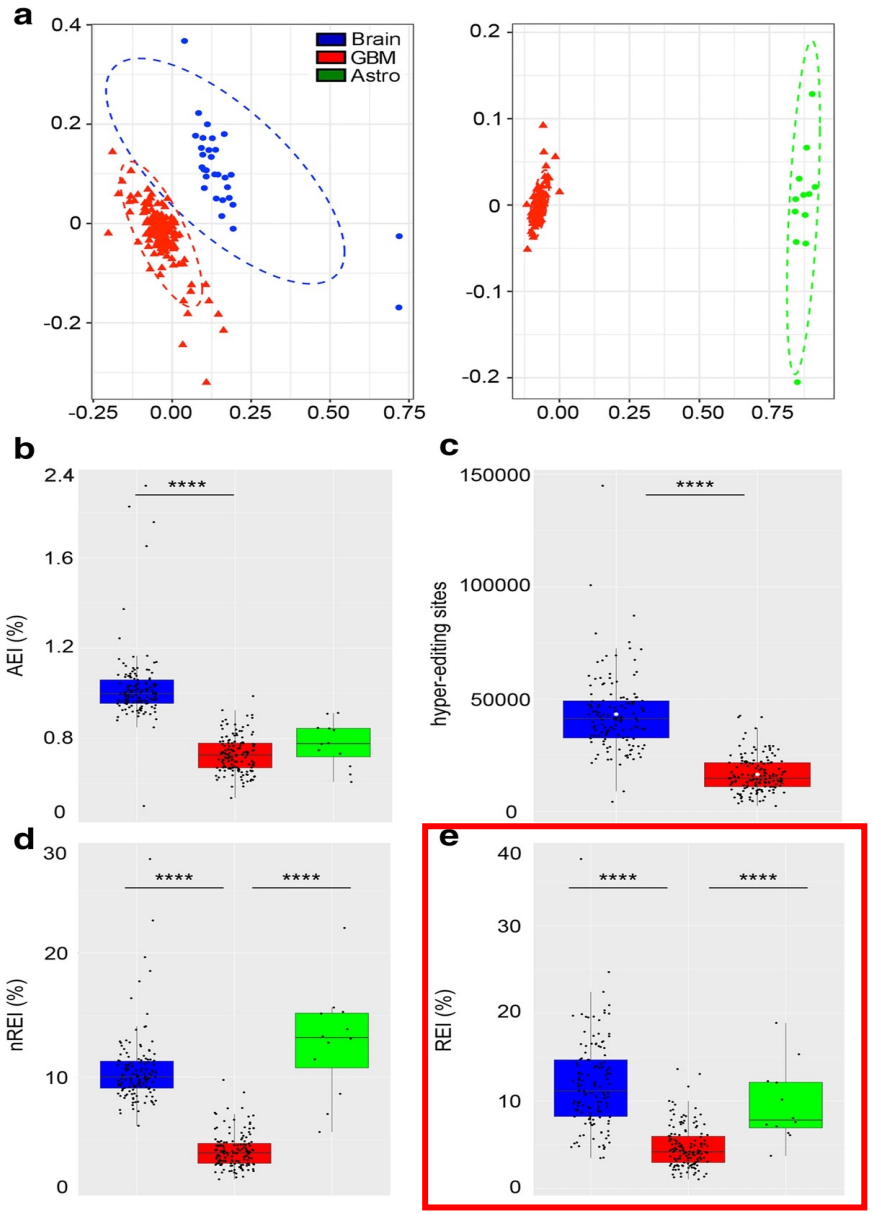
#Chr	Start/End	Cov	MisMatch	EdFreq	Gene	AAchange	AAchangePos
chr14	20841948	1	-	0.00000	TEP1	SG	S503G
chr14	20920211	2	-	0.00000	OSGEP	IM	I111M
chr14	21968737	21	-	0.00000	METTL3	MV	M402V
chr14	21968743	19	-	0.00000	METTL3	IV	I400V
chr14	21968764	13	-	0.00000	METTL3	MV	M393V



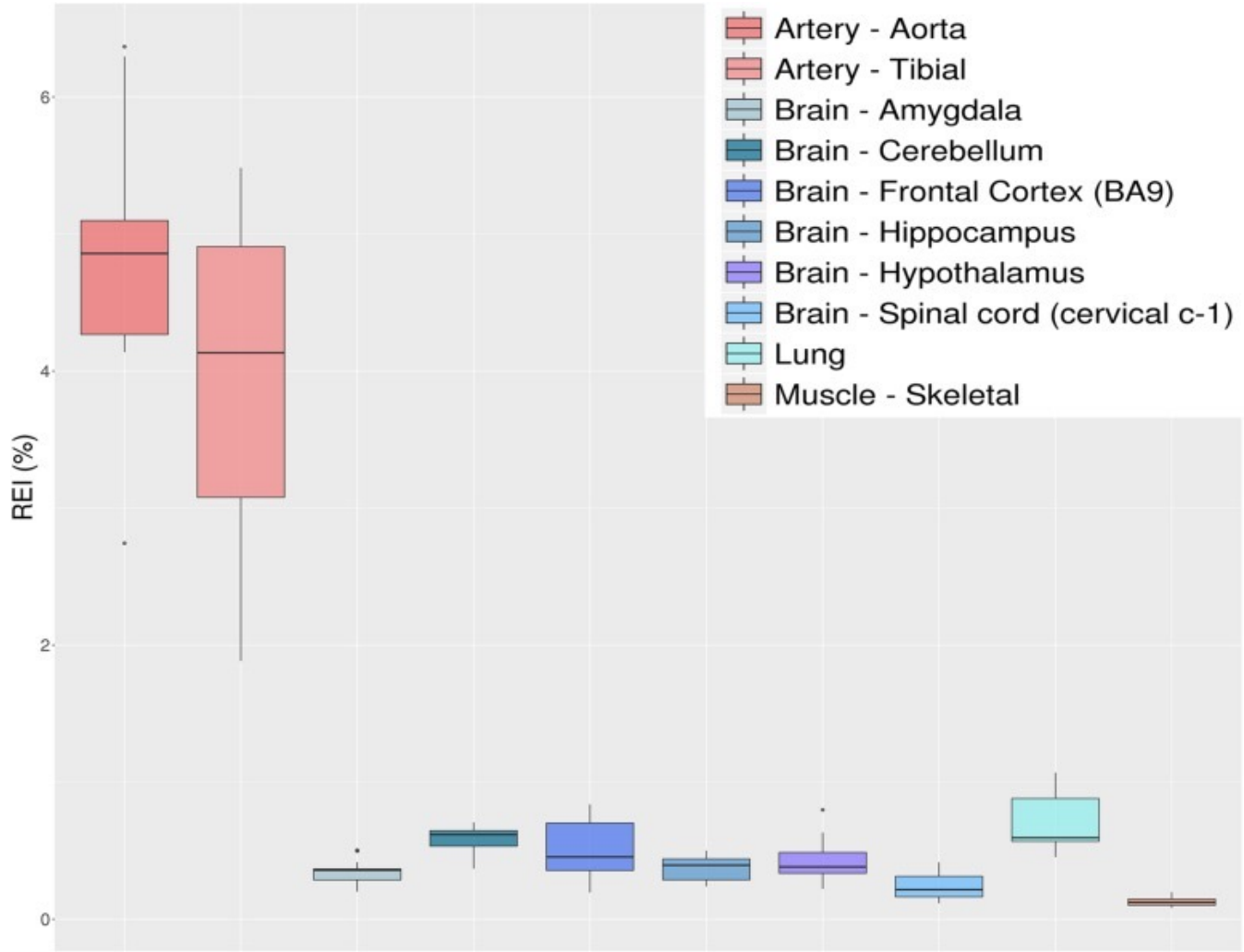
chr4	166226862	-	-	0.00000	KLHL2	RG	R176G
chr4	174235193	-	-	0.00000	GALNT7	YH	Y492H
chr4	186283170	-	-	0.00000	SNX25	KR	K751R
chr4	186283208	-	-	0.00000	SNX25	SG	S764G
chr4	186283212	-	-	0.00000	SNX25	QR	Q765R

All recoding: 1585 Covered: 214
>REI 0.1153496821 11.5349682107

Recoding Editing Index (REI)

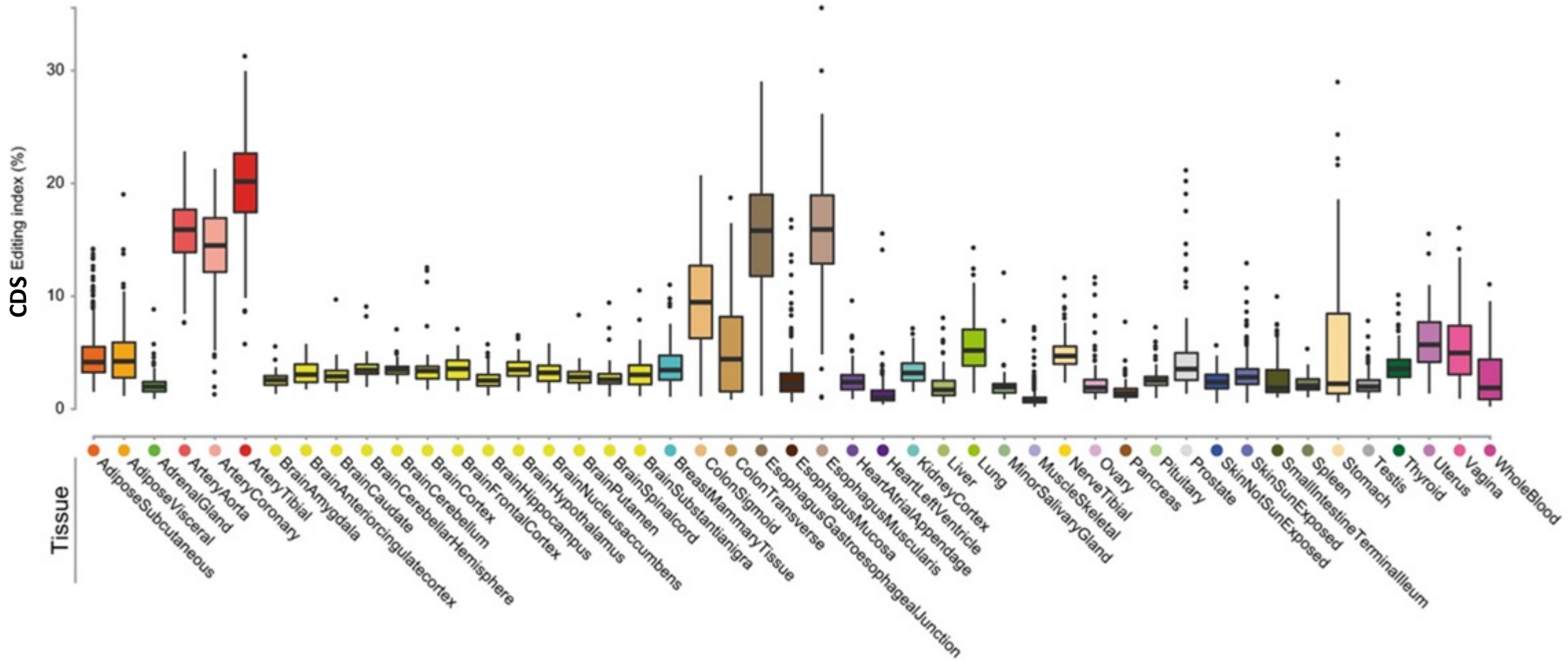


From [Dynamic inosinome profiles reveal novel patient stratification and gender-specific differences in glioblastoma]



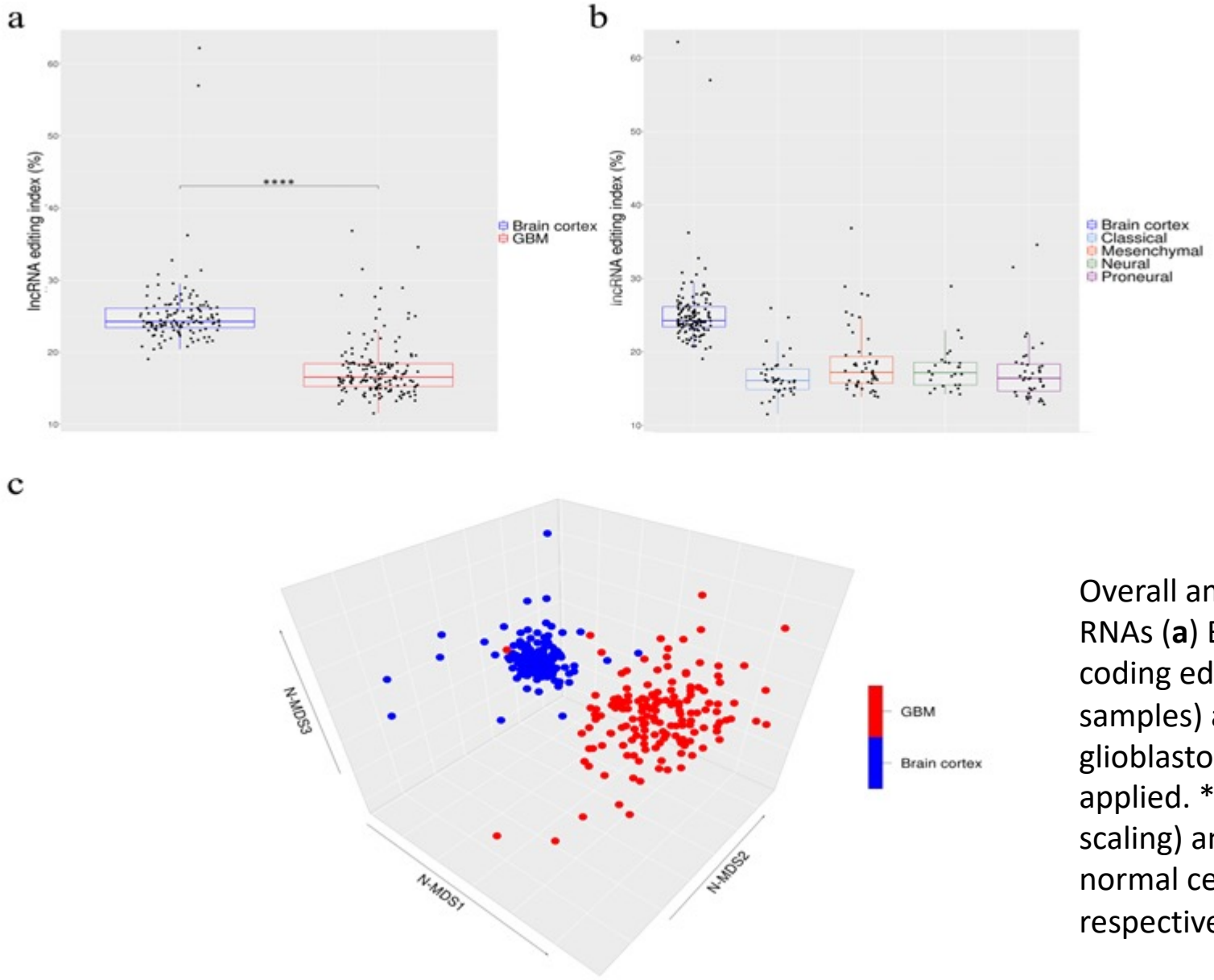
Distributions of recoding editing index (REI) values over 10 selected tissues from the GTEx project reported as box-plots. REI is calculated as the weighted average of editing levels over all known recoding sites from the REDIportal database. Most brain sub-tissues show similar levels of recoding editing. A remarkable exception is represented by the aorta and tibial artery showing a surprisingly high editing level.

From [Quantifying RNA Editing in Deep Transcriptome Datasets]



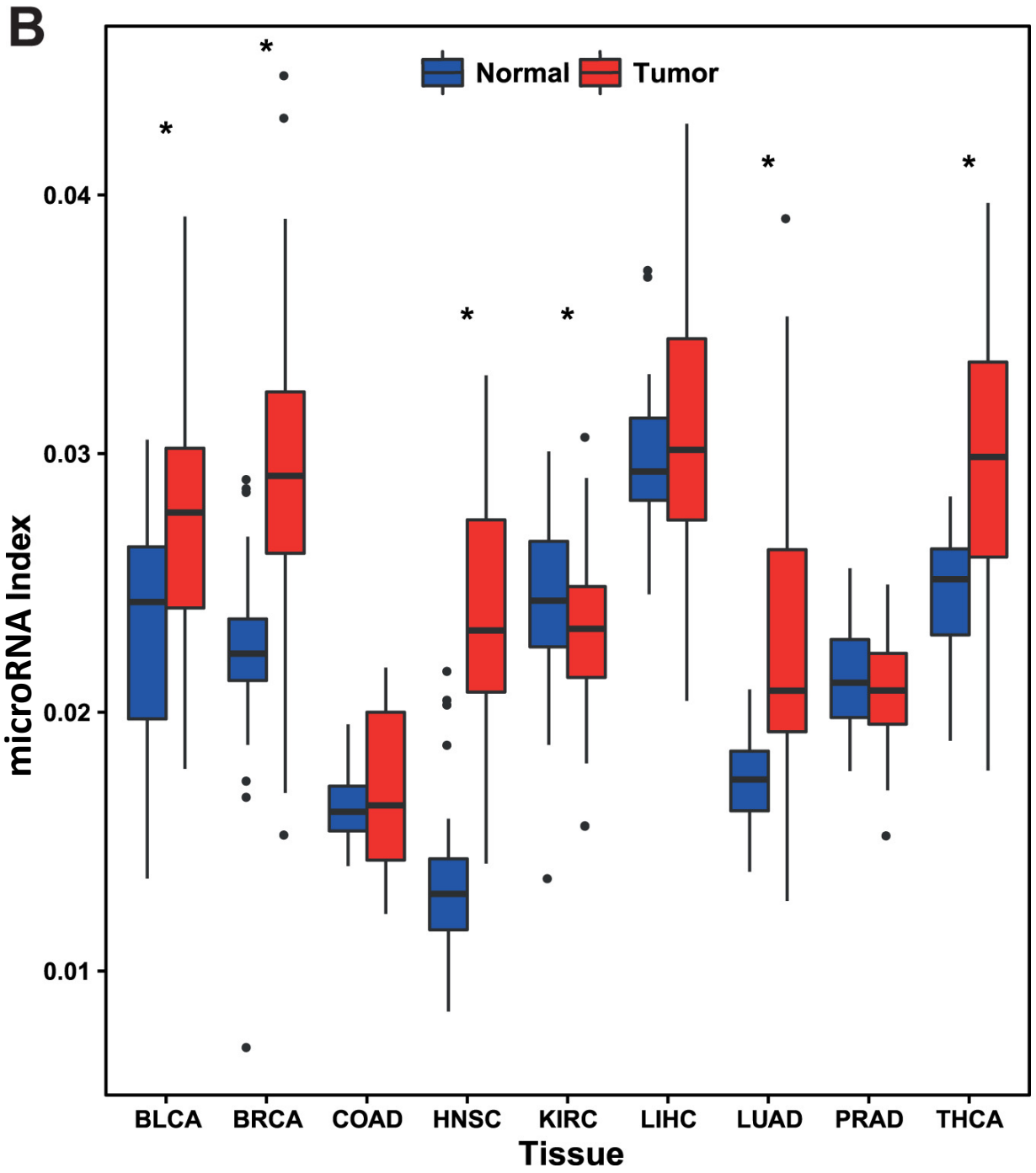
Box-and-whisker plots, depicting the distribution of per-tissue CDS editing index values reveal that although the number of edited sites in the brain is large, CDS editing activity (number of deamination events) is comparable to most other tissues and is much lower than in arteries, colon or esophagus. Box-and-whisker plots show the medians (horizontal lines), upper and lower quartiles (box edges), and 1.5 × the interquartile range (whiskers).

From: Landscape of adenosine-to-inosine RNA recoding across human tissues



Overall amount of A-to-I RNA editing in long non-coding RNAs (**a**) Boxplots showing the distributions of long non-coding editing index values across normal brain cortex (132 samples) and primary glioblastomas (156 samples) and (**b**) glioblastoma subtypes. Two-tailed Mann–Whitney U test was applied. **** $p \leq 0.0001$ (**c**) 3D-MDS (multidimensional scaling) analysis of RNA editing profiles in glioblastoma and normal cerebral cortex. Red and blue points indicate respectively GBMs and normal brains.

From De Novo A-to-I RNA Editing Discovery in lncRNA



Calculation of editing levels of known sites across all samples

We collected a list of 129 A-to-I editing sites within mature human miRNAs from 17 different studies. In order to measure the editing levels at these sites, we used the same alignment of the TCGA data to the human genome and kept only bases with Phred score $Q \geq 30$. We used mpileup for base calling for each of the known editing sites and calculated the editing levels as the ratio between the number of G bases and the sum of A and G bases.

We looked for sites showing editing at a level significantly higher than 1% (using the binomial test followed by FDR multiple testing correction) in at least one of the 55 samples types.

Editing of microRNA targets is elevated in tumors. Boxplot of editing index per sample for matched samples in 9 different cancers. Targets-editing is significantly ($P < 0.05$ Wilcoxon signed-rank test) higher in 5 cancers and lower in 1

From **Human cancer tissues exhibit reduced A-to-I editing of miRNAs coupled with elevated editing of their targets**

RECAP:

Download and unzip REDIportal annotations

```
wget http://srv00.recas.ba.infn.it/webshare/rediportalDownload/table1_full.txt.gz
gunzip table1_full.txt.gz
```

Prepare REDIportal annotations for REDItools

```
awk 'OFS="\t"{sum+=1; print $1,"rediportal","ed",$2,$2,".",$5,".","gene_id \""sum"\"; transcript_id \""sum"\";"}'
table1_full.txt > atlas.gtf
sort -V -k1,1 -k4,4n atlas.gtf > atlas.sorted.gtf
bgzip atlas.sorted.gtf
tabix -p gff atlas.sorted.gtf.gz
```

Annotate ALU, REP NON ALU and NON REP sites using known editing events from REDIportal:

```
python /data/REDItools/accessory/AnnotateTable.py -a /data/atlas.sorted.gtf.gz -n ed -k R -c 1 -i
outTable_613813579_chr4.out.rmsk.snp.alu -o outTable_613813579_chr4.out.rmsk.snp.alu.ed -u
python /data/REDItools/accessory/AnnotateTable.py -a /data/atlas.sorted.gtf.gz -n ed -k R -c 1 -i
outTable_613813579_chr4.out.rmsk.snp.nonalu -o outTable_613813579_chr4.out.rmsk.snp.nonalu.ed -u
python /data/REDItools/accessory/AnnotateTable.py -a /data/atlas.sorted.gtf.gz -n ed -k R -c 1 -i
outTable_613813579_chr4.out.rmsk.snp.nonrep -o outTable_613813579_chr4.out.rmsk.snp.nonrep.ed -u
```

Extract known editing events from ALU, REP NON ALU and NON REP sites and convert in TAB for editing levels quantification with REDItoolKnown:

```
mv outTable_613813579_chr4.out.rmsk.snp.alu.ed alu
mv outTable_613813579_chr4.out.rmsk.snp.nonalu.ed nonalu
mv outTable_613813579_chr4.out.rmsk.snp.nonrep.ed nonrep
cat alu.ed nonalu.ed nonrep.ed > alu-nonalu-nonrep.ed
awk -v FS="\t" -v OFS="\t" '{if ($19=="ed") print}' alu-nonalu-nonrep.ed > knownEditing # from de novo protocol
```

where `$19=="ed"` selects only known RNA editing events.

```
mkdir /home/student_X/artery
cd /home/student_X/artery
mkdir SRR1083076
cd SRR1083076
cp /data/artery/SRR1083076/knownEditing .
sort -k1,1 -k2,2n knownEditing > knownEditing_sorted
bgzip knownEditing_sorted
tabix –p vcf knownEditing_sorted.gz
```

Now repeat each of the above steps for the other 5 samples the changing SRR code accordingly:

```
mkdir SRR1091254
mkdir SRR1368668

mkdir /home/student_X/brain
cd /home/student_X/brain
mkdir SRR1086680
mkdir SRR1311771
```

REDItoolKnown.py

REDItoolKnown.py has been developed to explore the RNA editing potential of RNA-Seq data sets using known editing events. Such events can be downloaded from DARNED database or generated from supplementary materials of a variety of publications. Known RNA editing events have to be stored in TAB files.

```
USAGE: python REDItoolKnown.py [options]
Options:
-i              BAM file
-I              Sort input BAM file
-f              Reference in fasta file
-l              List of known RNA editing events
-C              Base interval to explore [100000]
-k              List of chromosomes to skip separated by comma or file
-t              Number of threads [1]
-o              Output folder [rediFolder_268123878]
-F              Internal folder name [null]
-c              Min. read coverage [10]
-q              Min. quality score [30]
-m              Min. mapping quality score [30]*
-O              Min. homoplymeric length [5]
-s              Infer strand (for strand oriented reads) [1]
-g              Strand inference type 1:maxValue 2:useConfidence [1]
-x              Strand confidence [0.70]
-S              Strand correction
-G              Infer strand by gff annotation (must be sorted, otherwise use -X)
-X              Sort annotation files
-K              File with positions to exclude
-e              Exclude multi hits
-d              Exclude duplicates
-p              Use paired concardant reads only
-u              Consider mapping quality
-T              Trim x bases up and y bases down per read [0-0]
-B              Blat file for correction
-U              Remove substitutions in homopolymeric regions
-v              Min. num. of reads supporting the variation [3]
-n              Min. editing frequency [0.1]
-E              Exclude positions with multiple changes
-P              File containing splice sites annotations
-r              Num. of bases near splice sites to explore [4]
-H              No Table Header
-h              Print this help

*This value may change according to the aligner:
        - For Bowtie use 255
        - For Bowtie2 use 40
        - For BWA use 30
        - For RNA-STAR use 255
        - For HISAT2 use 60
        - For Tophat1 use 255
        - For Tophat2 use 50
        - For GSNAP use 30
```

After generating the tab file sorted containing the list of known sites for each sample run the REDItoolKnown script to quantify the editing levels at these sites:

```
python /data/REDItools/main/REDItoolKnown.py -i /data/artery/SRR1083076/SRR1083076.bam -f
/data/data_reditools/Epitranscriptome_course_2023/refs/GRCh37.primary_assembly.genome_chrs_4_14_19.fa -o
/home/student_X/SRR1083076 -l /home/student_X/artery/SRR1083076/knownEditing_sorted.gz -m 255 -e -p -u -E -c
5 -v 1

python /data/REDItools/main/REDItoolKnown.py -i /data/artery/SRR1091254/SRR1091254.bam -f
/data/data_reditools/Epitranscriptome_course_2023/refs/GRCh37.primary_assembly.genome_chrs_4_14_19.fa -o
/home/student_X/SRR1091254 -l /home/student_X/artery/SRR1091254/knownEditing_sorted.gz -m 255 -e -p -u -E -c
5 -v 1

python /data/REDItools/main/REDItoolKnown.py -i /data/artery/SRR1368668/SRR1368668.bam -f
/data/data_reditools/Epitranscriptome_course_2023/refs/GRCh37.primary_assembly.genome_chrs_4_14_19.fa -o
/home/student_X/SRR1368668 -l /home/student_X/artery/SRR1368668/knownEditing_sorted.gz -m 255 -e -p -u -E -c
5 -v 1

python /data/REDItools/main/REDItoolKnown.py -i /data/brain/SRR1086680/SRR1086680.bam -f
/data/data_reditools/Epitranscriptome_course_2023/refs/GRCh37.primary_assembly.genome_chrs_4_14_19.fa -o
/home/student_X/SRR1086680 -l /home/student_X/brain/SRR1086680/knownEditing_sorted.gz -m 255 -e -p -u -E -c
5 -v 1

python /data/REDItools/main/REDItoolKnown.py -i /data/brain/SRR1311771/SRR1311771.bam -f
/data/data_reditools/Epitranscriptome_course_2023/refs/GRCh37.primary_assembly.genome_chrs_4_14_19.fa -o
/home/student_X/SRR1311771 -l /home/student_X/brain/SRR1311771/knownEditing_sorted.gz -m 255 -e -p -u -E -c
5 -v 1
```

Differential RNA editing

The identification of differential RNA editing is still an open question. Nonetheless, dysregulated RNA editing at recoding events can be calculated employing the Mann-Whitney U-test described in Silvestris et al. (2019) or the statistical pipeline proposed by Tran et al. (2019) Both pipelines are embedded with the get_DE_events.py script.

Prepare a comma separated sample informations file (e.g ArteryAorta_vs_BrainCerebellum.sif) required as input by the get_DE_events.py script.

•A csv sample file containing the main informations about each sample to be used in the experiment.
•A name for Samples group1 (e.g. ArteryAorta)
•A name for Samples group2 (e.g BrainCerebellum)

An example file is provided:

Sample,Group,Type
SRR1083076,GROUPA,ArteryAorta
SRR1086680,GROUPB,BrainCerebellum
SRR1091254,GROUPA,ArteryAorta
SRR1311771,GROUPB,BrainCerebellum
SRR1368668,GROUPA,ArteryAorta

Create a confortable workdir (e.g. ArteryAorta_vs_BrainCerebellum) and enter it

```
mkdir ArteryAorta_vs_BrainCerebellum && cd ArteryAorta_vs_BrainCerebellum
```

Run sample_path_folder_creator.py that will copy the Reditools tables in different directories following the sample/Group subdivisions reported in the sample informations file (.sif).

```
python /data/QEdit/scripts/sample_path_folder_creator.py /home/student_X/ArteryAorta_vs_BrainCerebellum.sif
```

Note. The script assumes that REDItools outputs (e.g. SRR1071289, SRR1101591) are contained in a "tables" folder in your main working directory, otherwise modify the last line of the script accordingly.

Run the get_DE_events.py script (Mann-Whitney U-test) on multiple REDItools tables following the sample/Group subdivions reported in the sample informations file (.sif). The option -sig yes in combination with -cpval 2 (BH correction), returns only significantly edited positions. MtsA and mtsB, represents the minimum threshold of samples per group on which the statistical tests are applied.

```
python /data/QEdit/scripts/get_DE_events.py -input_file /home/student_X/ArteryAorta_vs_BrainCerebellum.sif > DE_res
```

Alternatively, run the get_DE_events.py script on the same samples applying the the statistical pipeline proposed by Tran et al. (2019)

```
python /data/QEdit/scripts/get_DE_events.py -linear -input_file ArteryAorta_vs_BrainCerebellum.sif
```

get_DE_events.py

This scripts and its related files are part of the supplemental material for the paper "Quantifying RNA editing in deep transcriptome datasets"

This script compares REDItools output table arising from multiple samples and returns dysregulated RNA editing by means of the Mann-Whitney U-test described in Silvestris et al. (2019) or the statistical pipeline proposed by Tran et al. (2019).

REDItoools output table are pre-filtered according to these main following criteria.

- RNAseq coverage per position (default **10 reads**)
- Minimum editing frequency per position (default **10%**)

For each editing candidate, the script applies the MannWhitney test to check the significance between the two groups, A and B.

By default the test is carried out only if the number of editing events per position is equal to 50% of the samples per group. This treshold can be manually modified (for both groups) by playing with the -mtsA and -mtsB options respectively.

Returned p-values can be corrected using Benjamini Hochberg or Bonferroni tests.

Usage:

```
usage: get_DE_events.py [-h] [-c MIN_COVERAGE] [-cpval PVALUE_CORRECTION]
                        [-input_file SAMPLES_INFORMATIONS_FILE]
                        [-gene_pos_file GENE_POS_FILE] [-f MIN_EDIT_FREQUENCY]
                        [-mtsA GROUPA_MIN_SAMPLE_TESTING]
                        [-mtsB GROUPB_MIN_SAMPLE_TESTING]
                        [-sig ONLY_SIGNIFICANT]
                        [-siglevel STATISTICAL_SIGNIFICANCE] [-linear]
                        [-graph] [-chr_col CHR_COLUMN] [-rsite RSITE]
```

optional arguments: **-h**, --help show this help message and exit
-c MIN_COVERAGE Coverage-q30
-cpval PVALUE_CORRECTION 1 --> Bonferroni correction / 2 --> Benjamini Hochberg
-input_file SAMPLES_INFORMATIONS_FILE (.sif) Comma separated file e.g: Sample,Group,Type (e.g SRR1093527,GROUPA,BrainCerebellum..., SRR1088437,GROUPB,ArteryTibial... etc) An example file is provided here
-gene_pos_file GENE_POS_FILE nonsynonymous_table_NONREP derived from Rediportal NOTE: A gene_pos file is required by -graph or -rsite. An example file can be found here here.
-f MIN_EDIT_FREQUENCY Editing Frequency
-mtsA GROUPA_MIN_SAMPLE_TESTING min percentage of groupA samples
-mtsB GROUPB_MIN_SAMPLE_TESTING min percentage of groupB samples
-sig ONLY_SIGNIFICANT Return only statistically significant editing events
-siglevel STATISTICAL_SIGNIFICANCE cutoff level to reject H0 hypothesis default 0.05
-linear Enable linear statistical model (Tran et al., 2019).
-graph R graph compatible table containing the following columns: Site|Delta|Mannwhitney|pval|Benjamini Hochberg corrected pvalue|status NOTE: THIS OPTION CAN BE USED ONLY IN COMBINATION with -Gene_pos_file
-chr_col CHR_COLUMN If set to "yes" a chromosome_position column will be added to R graph table. NOTE: THIS OPTION IS SPECIFIC FOR -graph & -Gene_pos_file COMBINATION
-rsite RSITE If set to "yes" all recoding sites will be shown in the output table. NOTE: THIS OPTION ONLY WORKS IN DEFAULT MODE.
e.g. python ../REDItools/accessory/get_DE_events.py -cpval 2 -input_file sample_information.csv -sig yes

Table 4 | Potential dysregulated RNA editing events in HD samples from BioProject PRJNA316625

Position (hg19)	Gene name	AA change	Controls	HDs	*P* value
chr6:102372589	*GRIK2*	Q621R	0.904	0.844	0.017
chr6:102337702	*GRIK2*	Y571C	0.846	0.777	0.020
chr4:158281294	*GRIA2*	R717G	0.591	0.497	0.029
chr14:26917530	*NOVA1*	S265G	0.403	0.307	0.025
chr12:121078907	*CABP1*	Q140R	0.307	0.233	0.008

We report the genomic position (human genome assembly hg19), the gene name, the amino acid (AA) change induced by RNA editing, the mean editing level in controls, the mean editing level in HD samples and the Mann–Whitney *P* value.

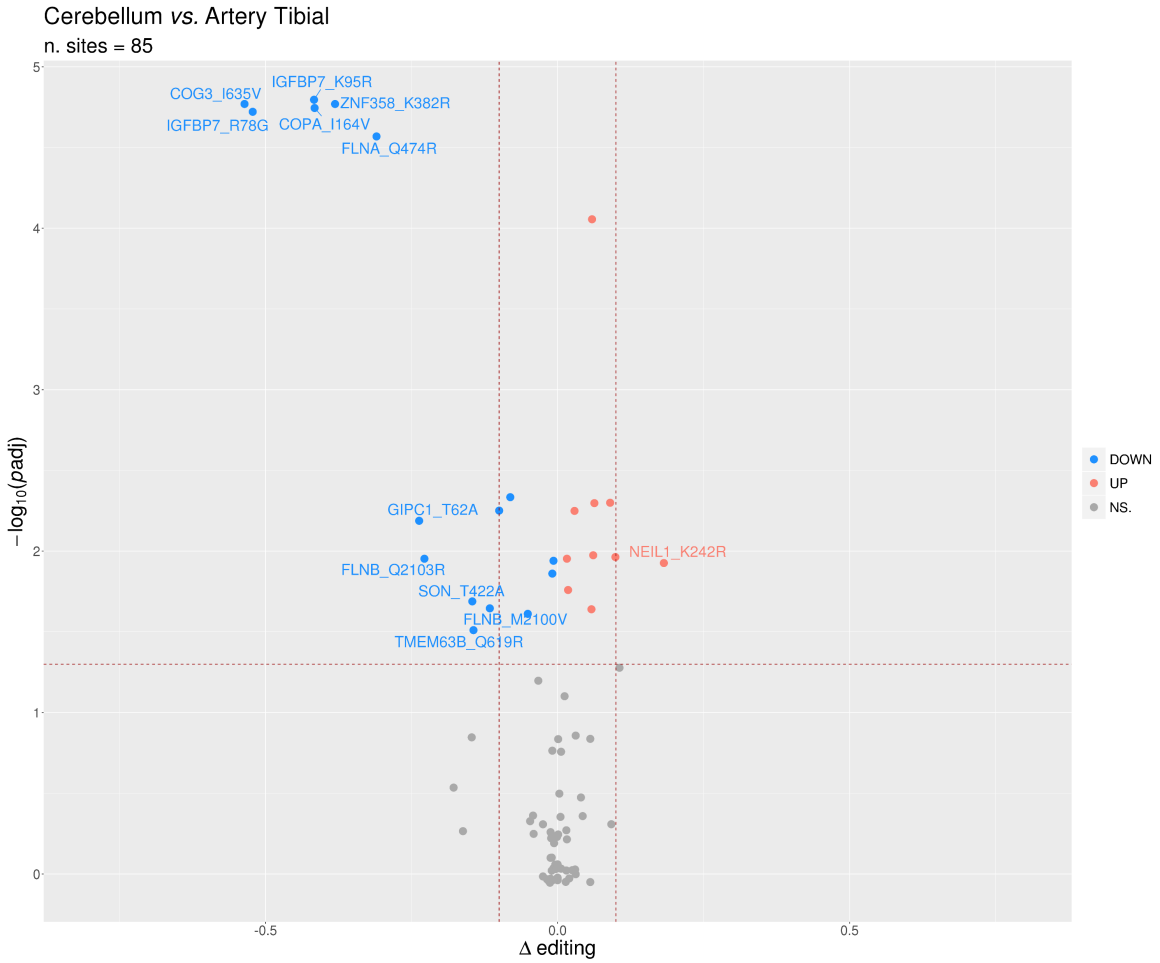
Annotate editing events with the UCSC refGene transcriptome annotation to add gene and region information:

```
python /data/REDItools/accessory/AnnotateTable.py -a
/data/data_reditools/Epitranscriptome_course_2023/refs/refGene_sorted.gtf.gz -i DE_res -o annotated_DE_res -u
-n refgene -k 'chromosome'
```

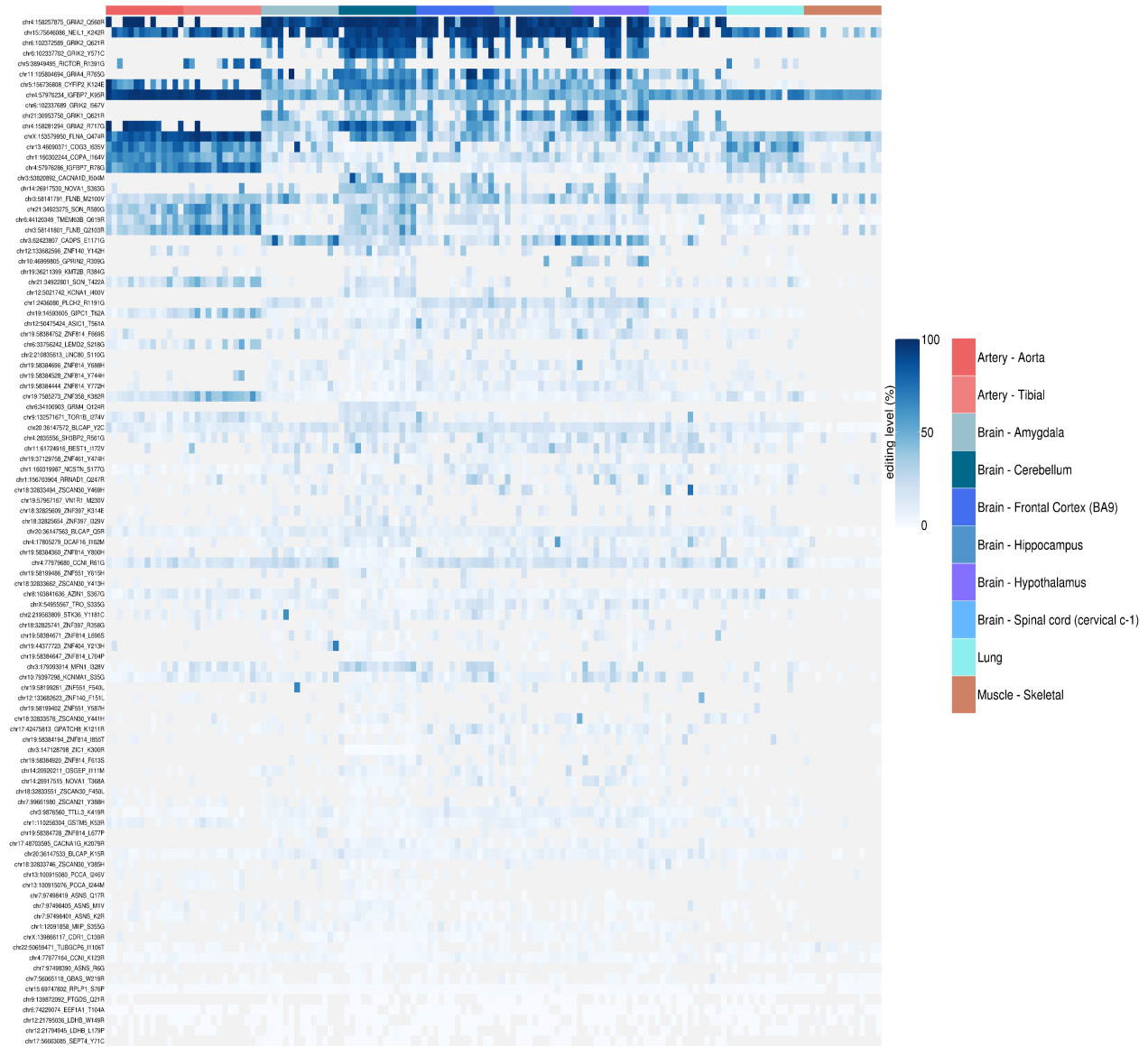
chromosome	position	editing_type	SRR1083076_ArteryAorta	SRR1091254_ArteryAorta	SRR1368668_ArteryAorta	SRR1086680_BrainCerebellum	SRR1311771_BrainCerebellum	[groupA_samples/groupB_samples]	delta_diff	pvalue (Mannwhitney)							
chr4	158257875	AG	-	1.00^AG^15	-	1.00^AG^68	1.00^AG^67	[1, 2]	-	-	transcript,CDS,exon	GRIA2					
chr4	158257879	AG	-	0.33^AG^6	-	0.14^AG^10	0.12^AG^9	[1, 2]	-	-	transcript,CDS,exon	GRIA2					
chr4	158281294	AG	-	1.00^AG^37	-	0.79^AG^60	0.70^AG^71	[1, 2]	-	-	transcript,CDS,exon	GRIA2					
chr4	158258136	AG	-	-	-	-	1.00^AG^12	[0, 1]	-	-	transcript	GRIA2					
chr4	158258137	AG	-	-	-	0.90^AG^9	0.92^AG^11	[0, 2]	-	-	transcript	GRIA2					
chr4	158258138	AG	-	-	-	-	0.50^AG^6	[0, 1]	-	-	transcript	GRIA2					
chr14	101312647	AG	-	-	-	-	0.50^AG^5	[0, 1]	-	-	transcript,exon	MEG3					
chr14	101320975	AG	-	-	-	0.71^AG^15	0.64^AG^14	[0, 2]	-	-	transcript	MEG3					
chr14	20917878	TC	-	-	-	0.70^TC^7	-	[0, 1]	-	-	transcript	OSGEP					
chr14	101318811	AG	-	-	-	0.55^AG^6	-	[0, 1]	-	-	exon,transcript$transcript,exon	MIR770$MEG3					
chr14	101301102	AG	0.16^AG^3	-	-	-	-	[1, 0]	-	-	transcript	MEG3					
chr19	7585273	AG	0.47^AG^8	-	0.23^AG^3	-	-	[2, 0]	-	-	transcript,exon,CDS	ZNF358					
chr19	53121375	TC	-	-	0.33^TC^4	0.73^TC^8	-	[1, 1]	-	-	transcript	ZNF83					
chr19	1988785	AG	-	-	-	0.30^AG^3	-	[0, 1]	-	-	transcript	BTBD2					
chr19	11560483	AG	-	-	-	0.27^AG^3	-	[0, 1]	-	-	transcript	PRKCSH					
chr19	41079920	AG	-	-	-	0.36^AG^4	-	[0, 1]	-	-	transcript	SPTBN4					

Editing levels visualization examples (https://github.com/BioinfoUNIBA/QEdit):

volcano plot



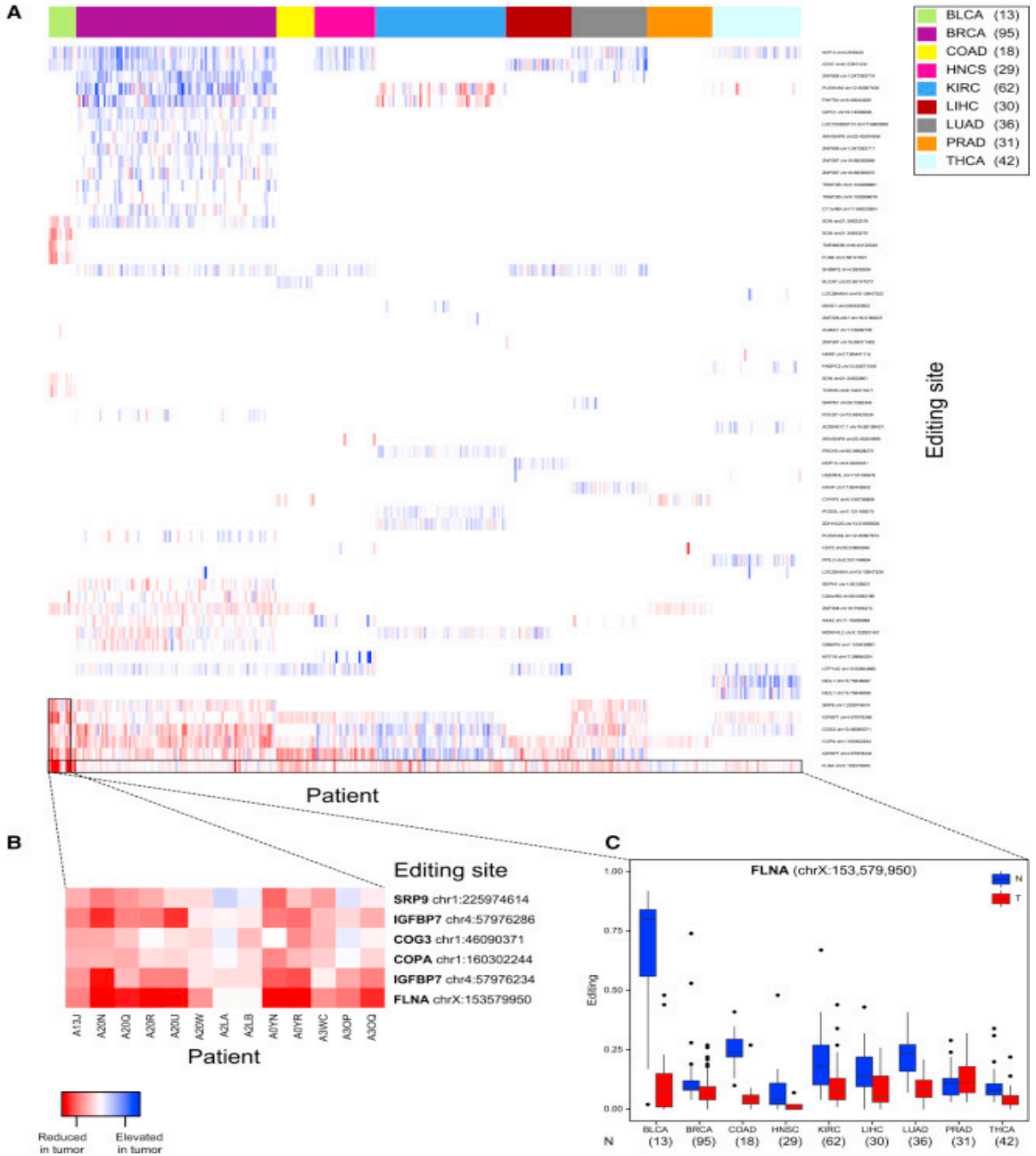
heatmap



Pan-cancer Editing Alteration in Coding Sequences

To analyze editing in the coding sequences, we compiled a list of editing sites to be tested for alteration in cancer. We took all sites documented in the RADAR database (Ramaswami and Li, 2014) in coding regions (UCSC annotation, 47,025 sites), supplemented by novel sites detected using the hyper-editing scheme within coding sequence (8,299 sites). We then excluded synonymous sites within Alu repeats, sites located in immunoglobulins (likely to be due to somatic hyper-mutations) and known SNPs (dbSNP 135). We did keep the following three sites: chr20, 36147572; chr5,156736808; and chrX,153579950 in the BLCAP, CYFIP2, and FLNA sites, respectively, which are known editing sites (Levanon et al., 2005). We further removed all sites annotated by wAnnovar (Chang and Wang, 2012) as residing in non-coding regions, resulting in 9,484 sites. We then quantified the editing levels in these sites using the REDItools script (Picardi and Pesole, 2013), trimming six bases at both ends of the reads.

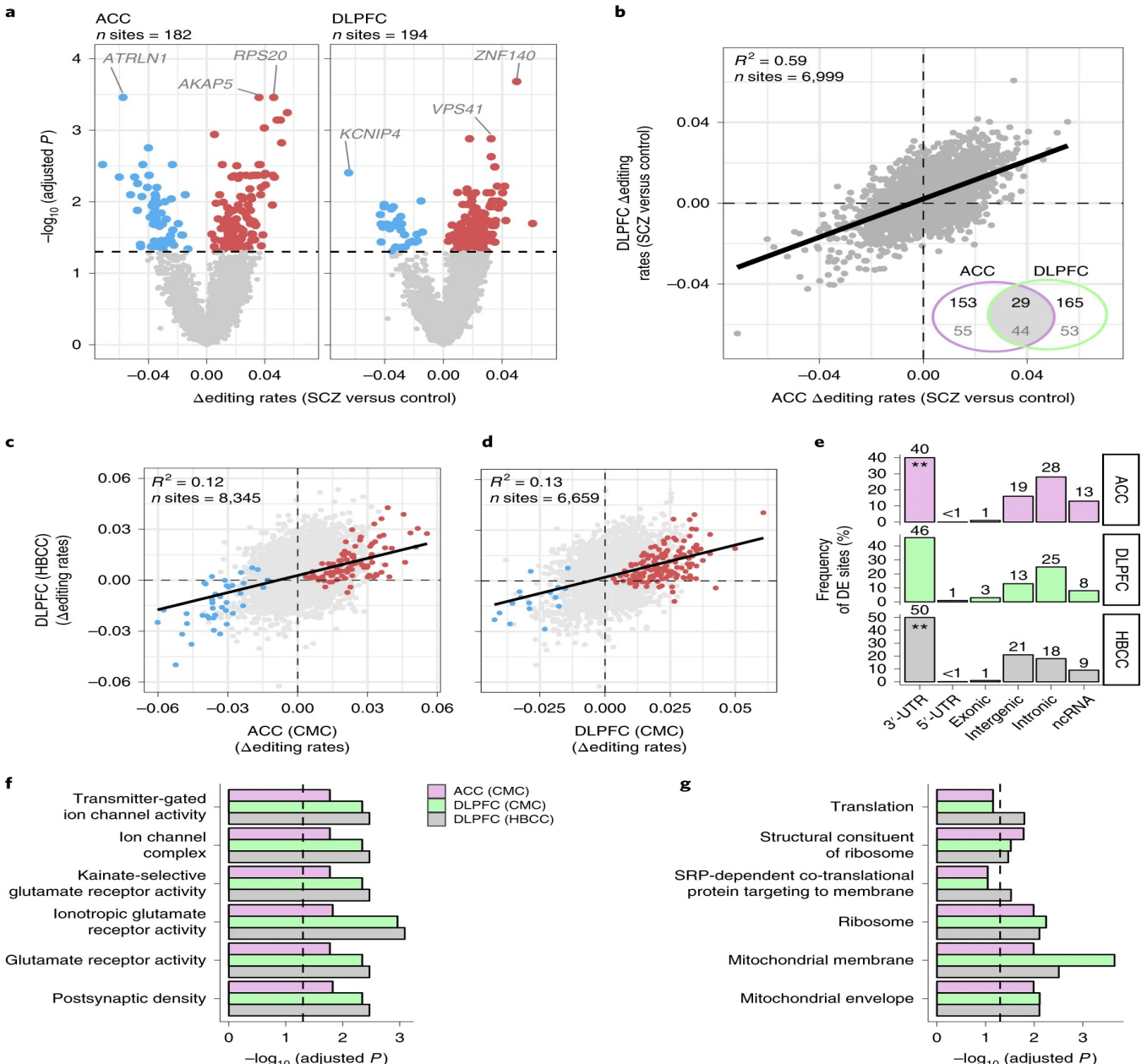
For each of the sites, differences in editing level between normal and cancer tissues were evaluated using the $\chi 2$ test followed by 5% FDR multiple-testing correction. Finally, we discarded sites where the absolute difference in the average editing level between normal and cancer samples was less than 5%



From [Elevated RNA Editing Activity Is a Major Contributor to Transcriptomic Diversity in Tumors]

From: [Global landscape and genetic regulation of RNA editing in cortical samples from individuals with schizophrenia]

Fig. 3: Identification of differentially edited sites in SCZ.



a, Differential editing sites in the ACC ($n_{control}$ = 245, n_{SCZ} = 225) and DLPFC ($n_{control}$ = 286, n_{SCZ} = 254). The dotted line marks a multiple test-corrected level of significance (adjusted $P < 0.05$, limma, linear regression with Benjamini–Hochberg correction). Red points indicate over-edited sites and blue points indicate under-edited sites. For the top three sites, their respective gene bodies have been indicated.