trackViewer

Lollipop/dandelion plots for methylation status and mutation data

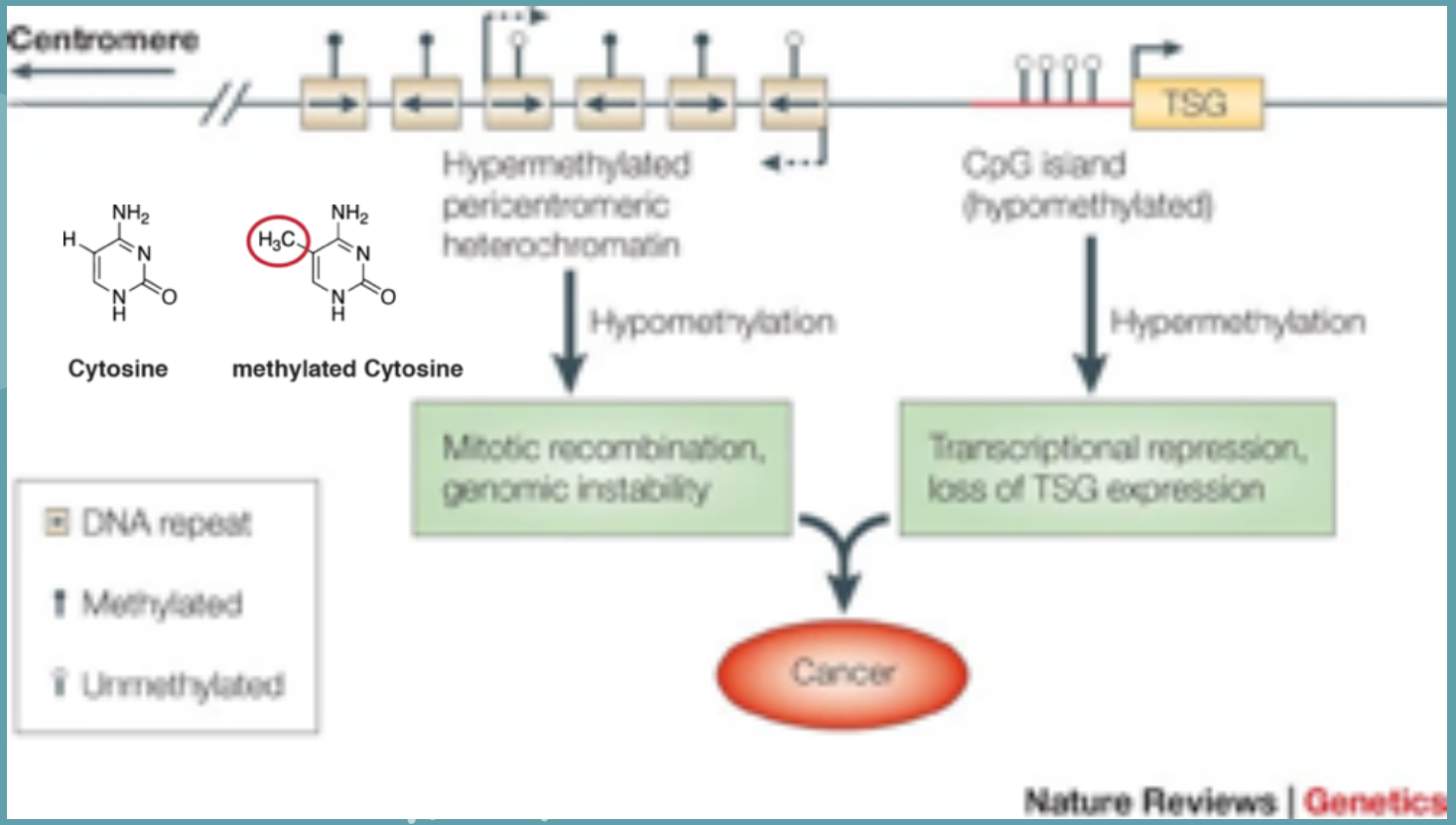
Jianhong Ou

Install the workshop pkg

```
## set the working directory,
## replace "~/Downloads/workshop2020" by your path
wd <- "~/Downloads/workshop2020"
dir.create(wd)
setwd(wd)
library(BiocManager)
install("jianhong/trackViewerBiocAsia2020Workshop",
        build_vignettes = TRUE)
vignette("trackViewer", package="trackViewerBiocAsia2020Workshop")
```

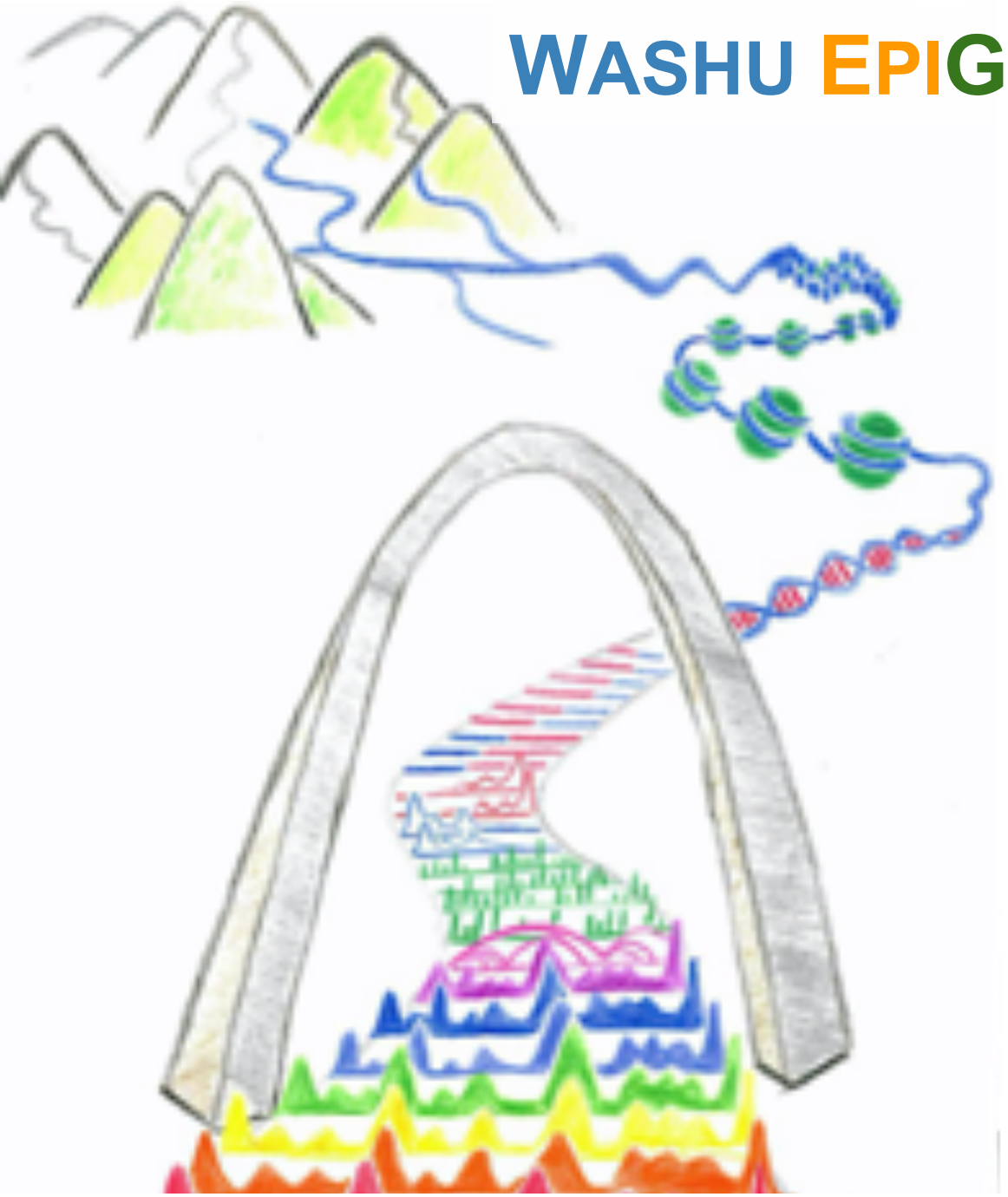
https://github.com/jianhong/trackViewerBiocAsia2020Workshop
https://jianhong.github.io/trackViewerBiocAsia2020Workshop/articles/trackViewer.html
https://bioconductor.org/packages/trackViewer
https://www.nature.com/articles/s41592-019-0430-y
Slides:
https://github.com/jianhong/trackViewerBiocAsia2020Workshop/blob/master/inst/extdata/slides.pdf

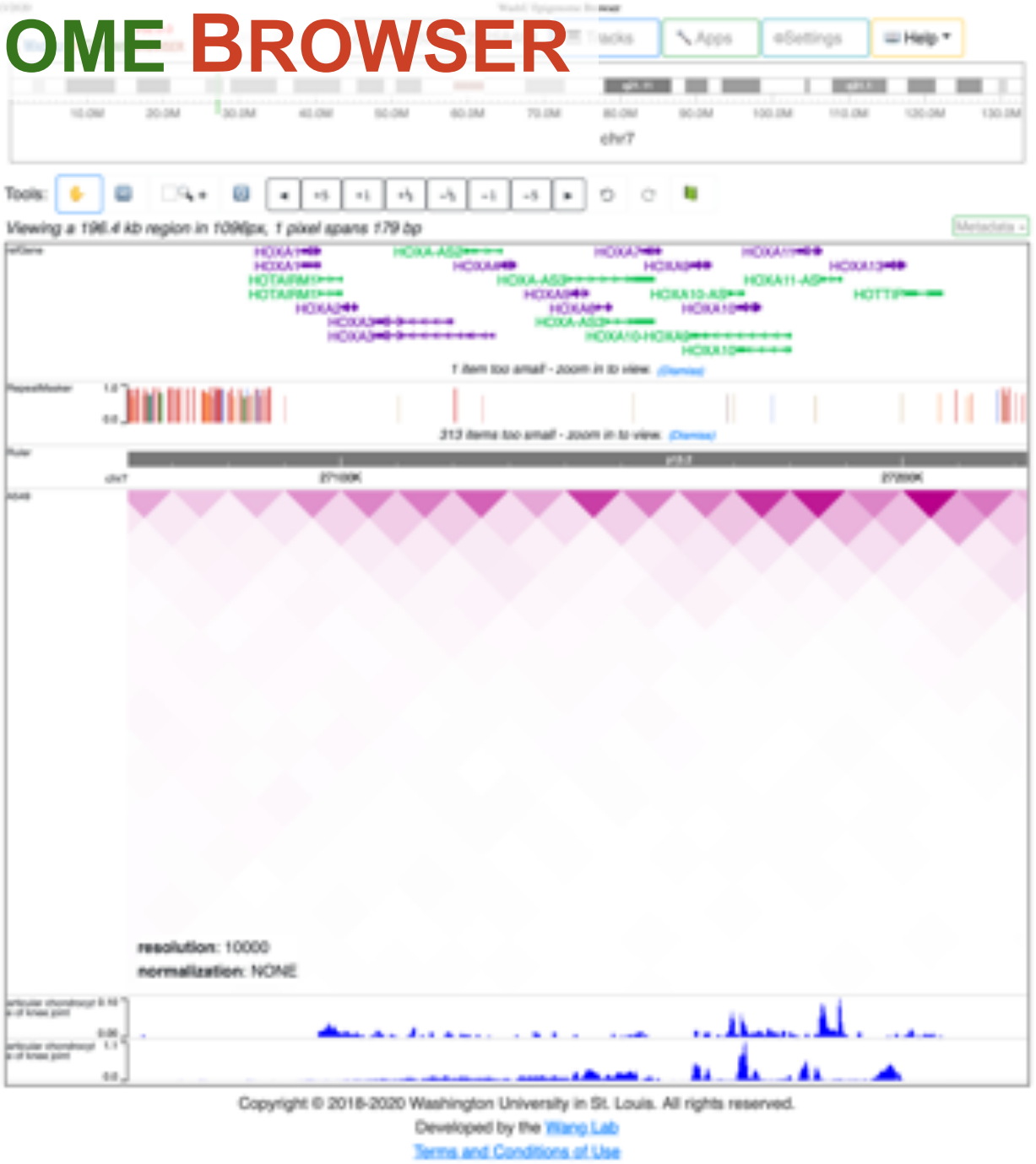
DNA methylation



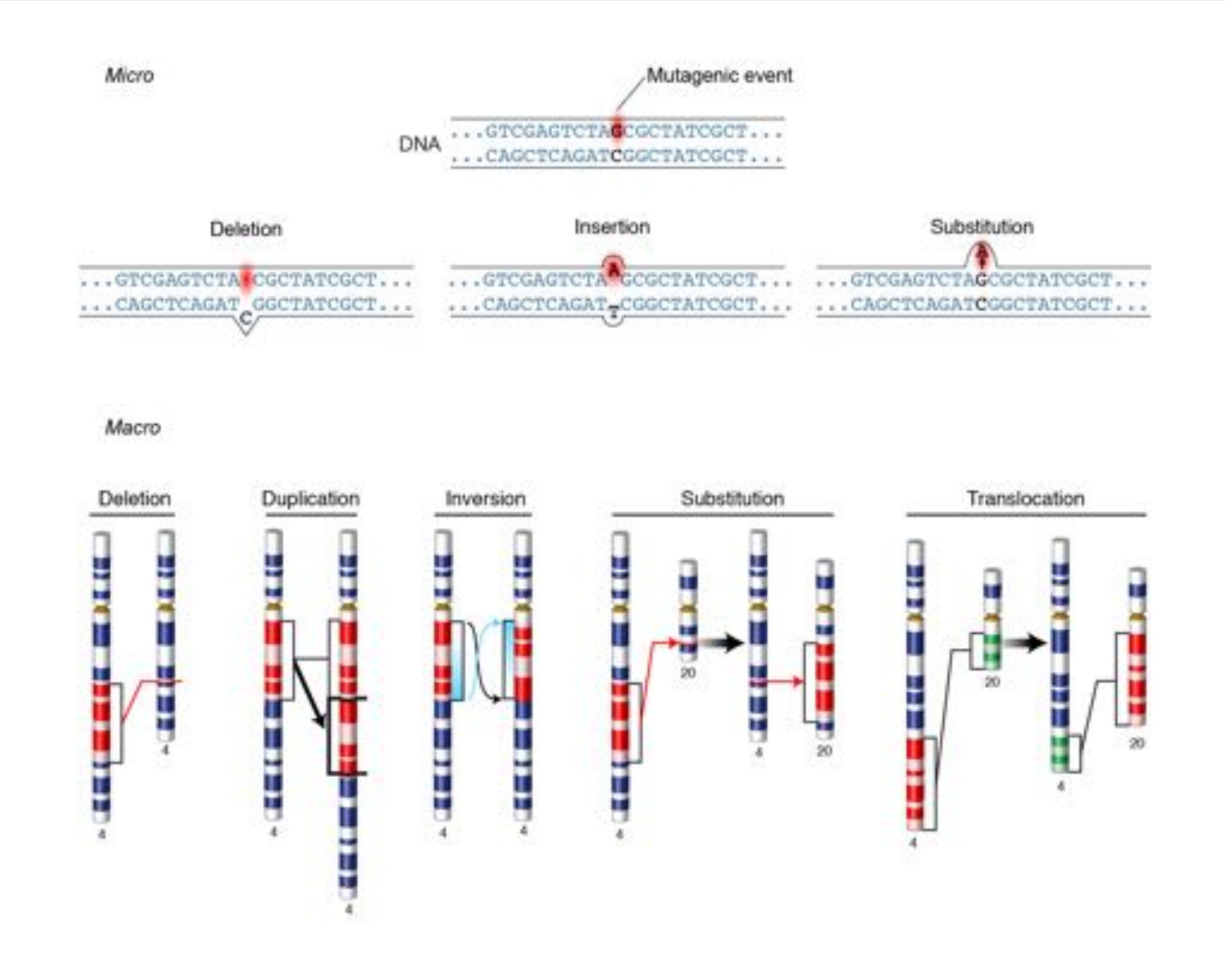
Robertson, K. 2005. doi: 10.1038/nrg1655
http://www.roadmapepigenomics.org/

WASHU EPIGENOME BROWSER



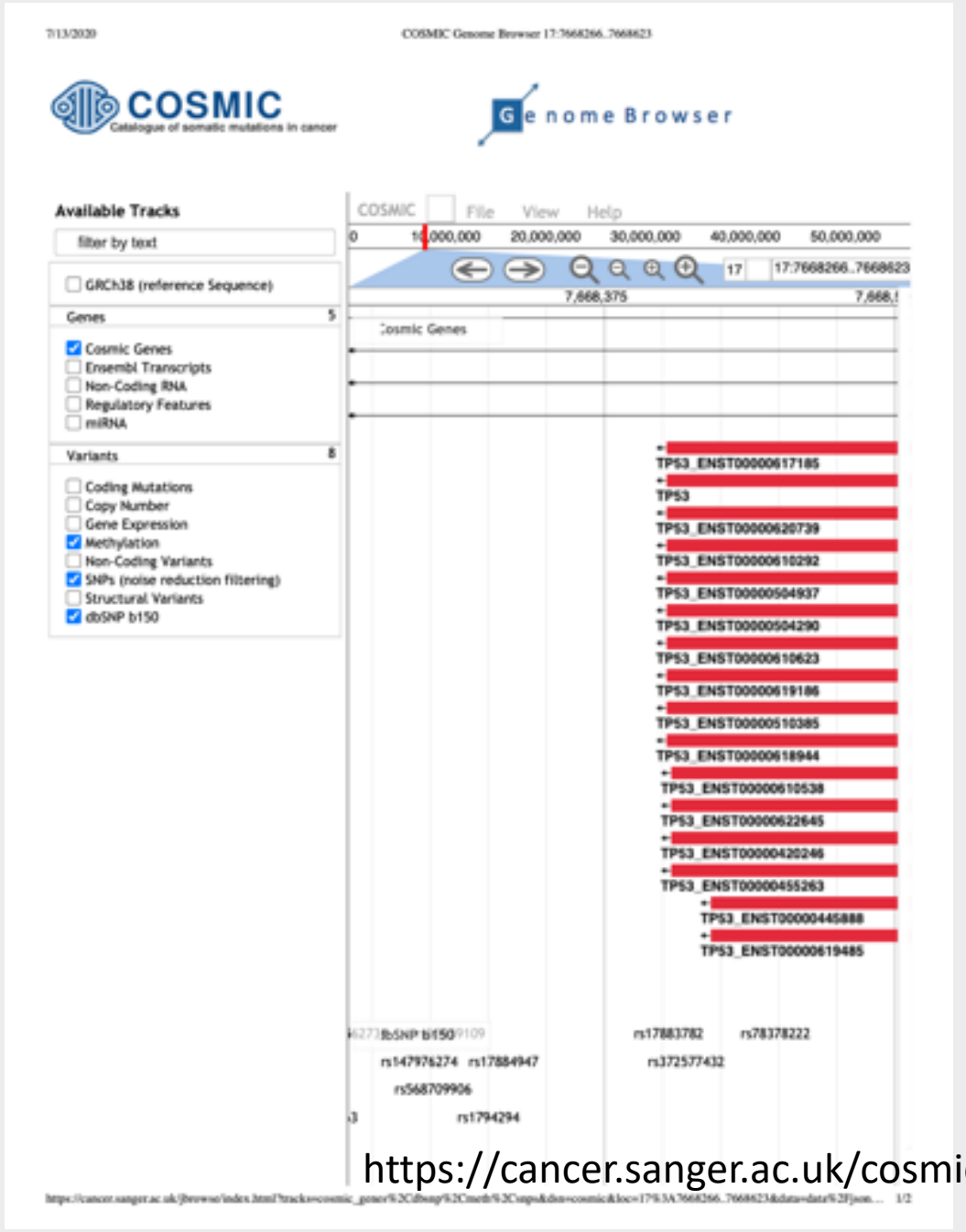


http://epigenomegateway.wustl.edu/support/workshop2019ENCODE.html

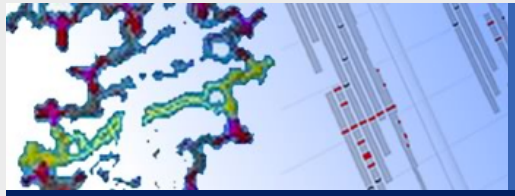


https://www.genome.gov/genetics-glossary/Mutation

MUTATION



https://cancer.sanger.ac.uk/cosmic



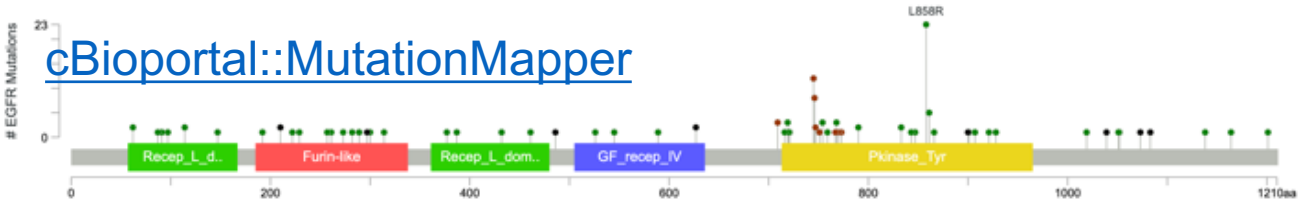
dbSNP

dbSNP contains human single nucleotide variations, microsatellites, and small-scale insertions and deletions along with publication, population frequency, molecular consequence, and genomic and RefSeq mapping information for both common variations and clinical mutations.

Genome browsers and viewers

Lollipop/needle plots

cBioportal::MutationMapper

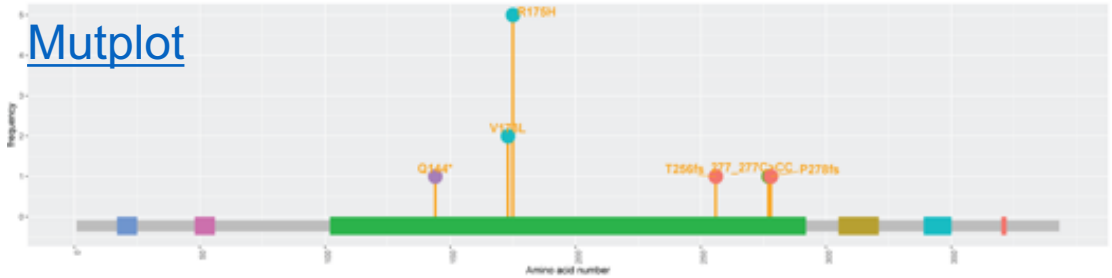


pfam::lollipops

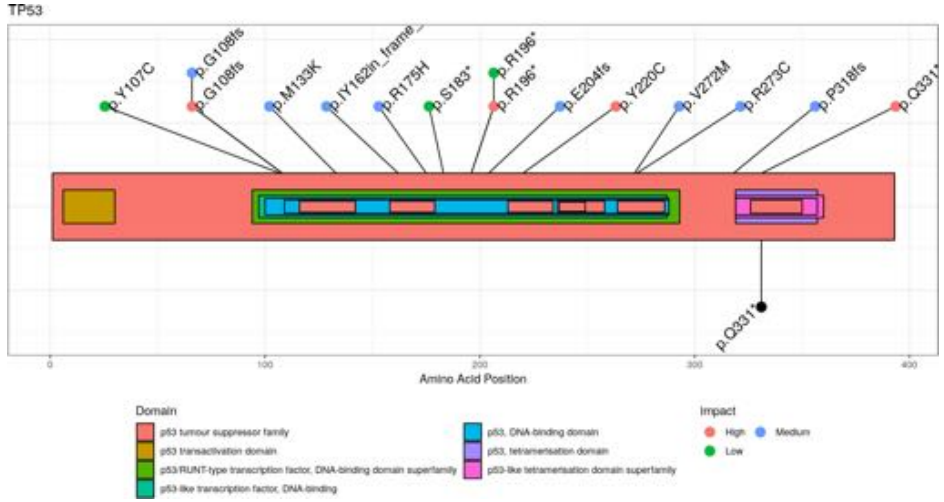
lollipops

Plot Protein

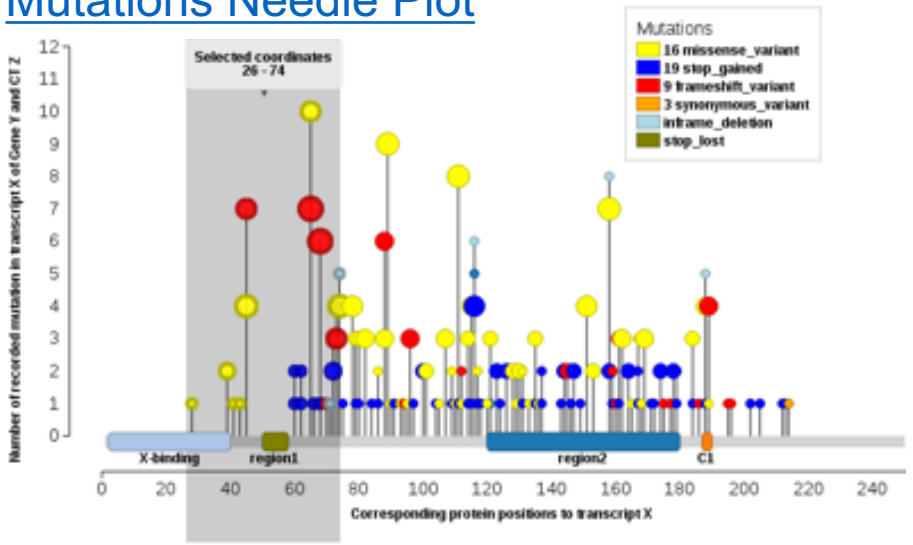
Mutplot



•GenVisR



Mutations Needle Plot



Jay et.al., 2016. doi: 10.1371/journal.pone.0160519

WHY `trackViewer`

pros:

– minimized: simple, clean design for the publication

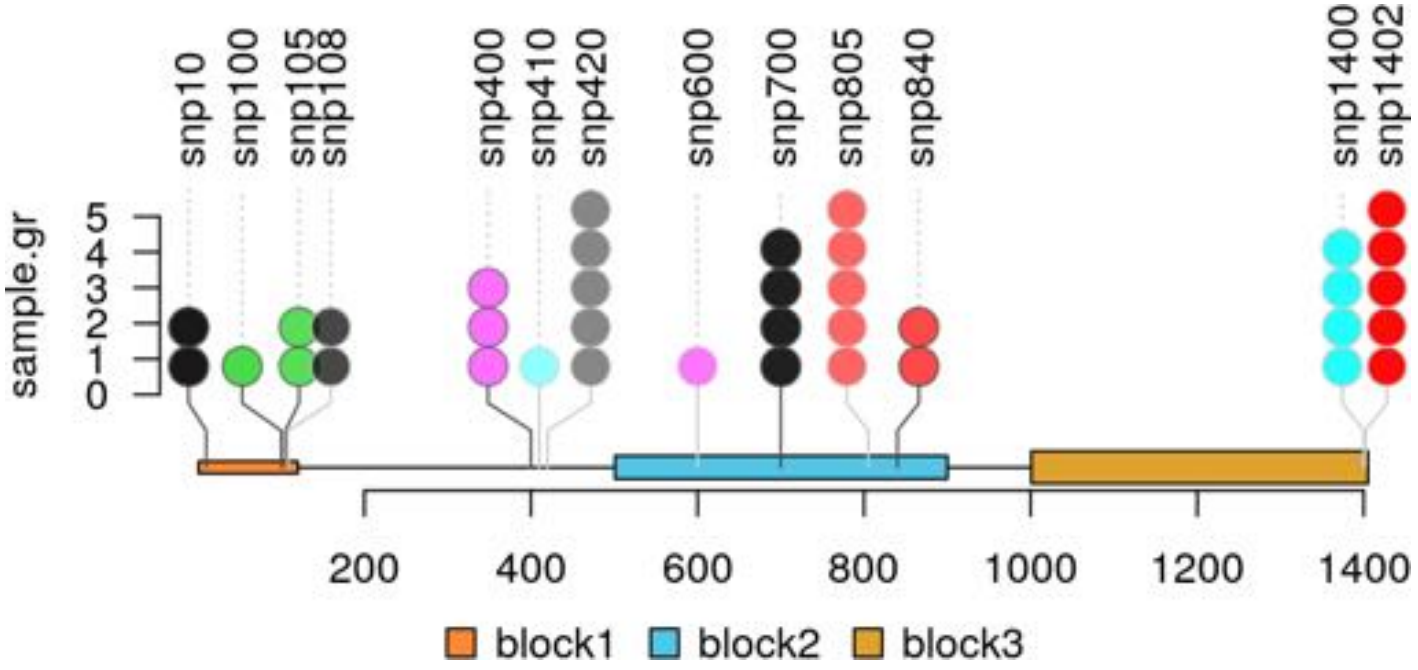
– highly customizable: labels, symbols, colors, opacity and size

– powerful: handle high dense data

cons:

– learning curve

– no embedded database



SNPs/INDELs

SNP: Single Nucleotide Polymorphism
INDEL: INsertion/DELetion

Name	Input formats	Output formats	Called variants
Atlas-SNP2	BAM	VCF	SNPs
GATK (UnifiedGenotyper)	BAM	VCF	SNPs, InDels
SAMtools (samtools mpileup)	BAM	VCF	SNPs, InDels
SNVer	BAM	VCF	SNPs, InDels
SOAPsnp	SOAP	Text format, GLFv2	SNPs
Varscan2	Pileup/mpileup	Text format, VCF	SNPs, InDels, CNA

Mielczarek et.al., 2016. doi: 10.1007/s13353-015-0292-7

VCF Variant Call Format

```
##fileformat=VCFv4.3
##fileDate=20090805
##source=myImputationProgramV3.1
##reference=file:///seq/references/1000GenomesPilot-NCBI36.fasta
##contig=<ID=20,length=62435964,assembly=B36,md5=f126cdf8a6e0c7f379d618ff66beb2da,species="Homo sapiens",taxonomy=x>
##phasing=partial
##INFO=<ID=NS,Number=1,Type=Integer,Description="Number of Samples With Data">
##INFO=<ID=DP,Number=1,Type=Integer,Description="Total Depth">
##INFO=<ID=AF,Number=A,Type=Float,Description="Allele Frequency">
##INFO=<ID=AA,Number=1,Type=String,Description="Ancestral Allele">
##INFO=<ID=DB,Number=0,Type=Flag,Description="dbSNP membership, build 129">
##INFO=<ID=H2,Number=0,Type=Flag,Description="HapMap2 membership">
##FILTER=<ID=q10,Description="Quality below 10">
##FILTER=<ID=s50,Description="Less than 50% of samples have data">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FORMAT=<ID=GQ,Number=1,Type=Integer,Description="Genotype Quality">
##FORMAT=<ID=DP,Number=1,Type=Integer,Description="Read Depth">
##FORMAT=<ID=HQ,Number=2,Type=Integer,Description="Haplotype Quality">
#CHROM POS     ID        REF ALT    QUAL FILTER INFO                              FORMAT      NA00001        NA00002        NA00003
20     14370   rs6054257 G   A      29   PASS   NS=3;DP=14;AF=0.5;DB;H2           GT:GQ:DP:HQ 0|0:48:1:51,51 1|0:48:8:51,51 1/1:43:5:.,.
20     17330   .         T   A      3    q10    NS=3;DP=11;AF=0.017               GT:GQ:DP:HQ 0|0:49:3:58,50 0|1:3:5:65,3   0/0:41:3
20     1230237 .         T   .      47   PASS   NS=3;DP=13;AA=T                   GT:GQ:DP:HQ 0|0:54:7:56,60 0|0:48:4:51,51 0/0:61:2
20     1234567 microsat1 GTC G,GTCT 50   PASS   NS=3;DP=9;AA=G                    GT:GQ:DP    0/1:35:4       0/2:17:2       1/1:40:3
```

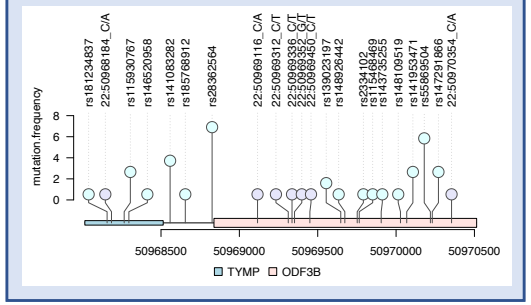
https://en.wikipedia.org/wiki/Variant_Call_Format

Visualization of methylation and mutation data in different styles using `trackViewer`

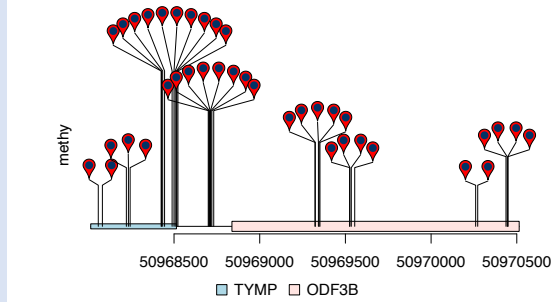
VCF/BED/BedGraph

TxDb annotation package

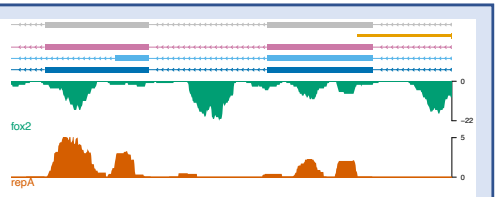
Low dense lollipop plot



High dense dandelion plot

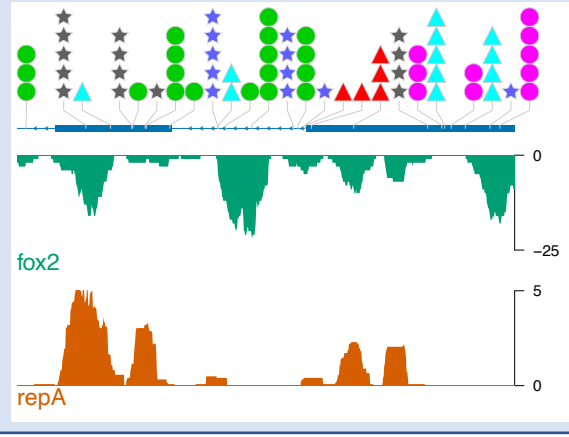


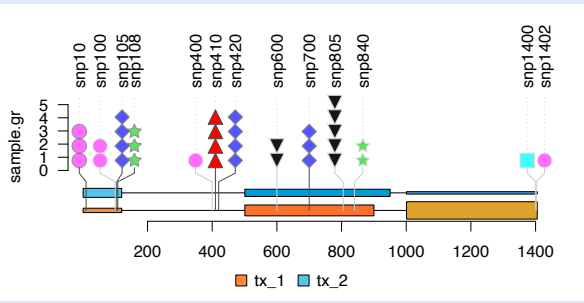
Coverage plot



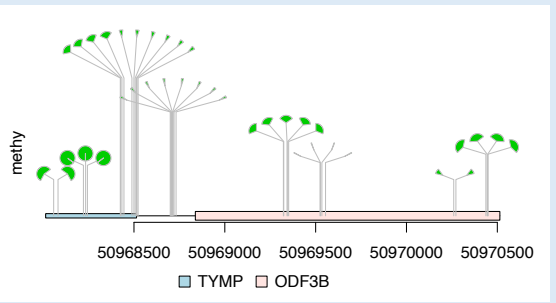
WIG/BigWig/
BED/BedGraph/Bam

Integrated plot

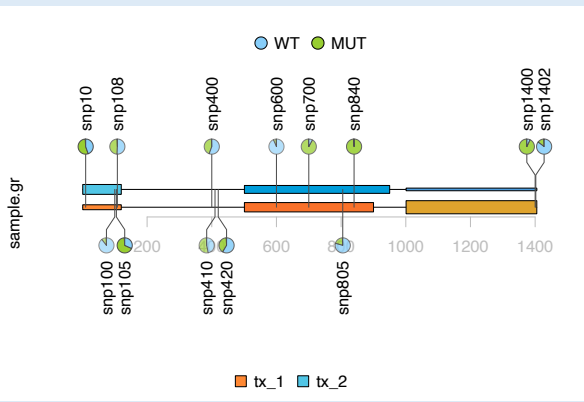




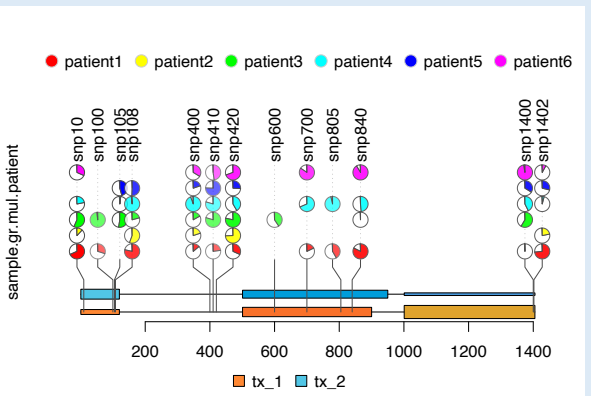
Lollipop plot with different symbols



Dandelion plot for depicting dense methylation data

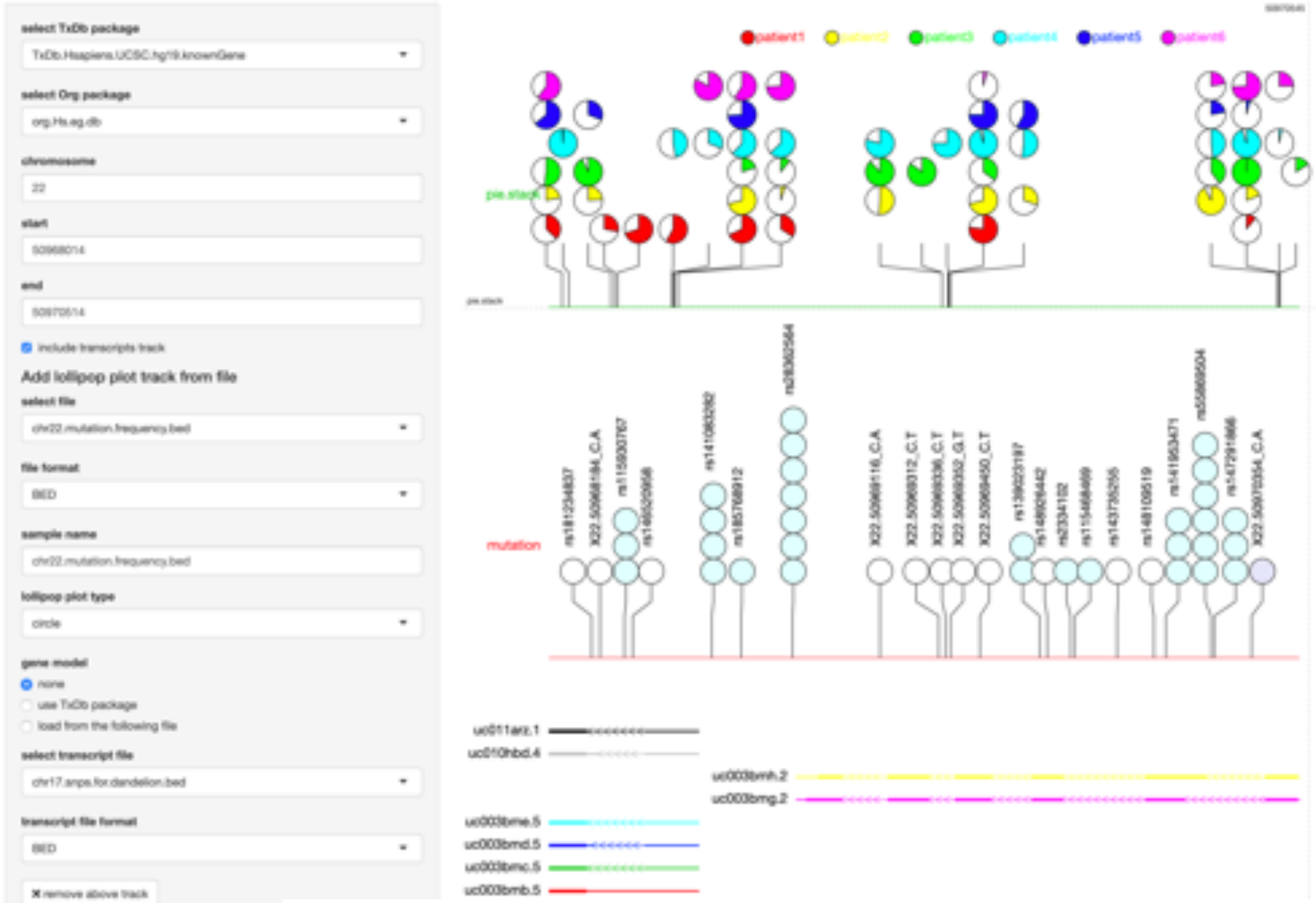


Lollipop plot in caterpillar layout



Lollipop plot in pie.stack layout

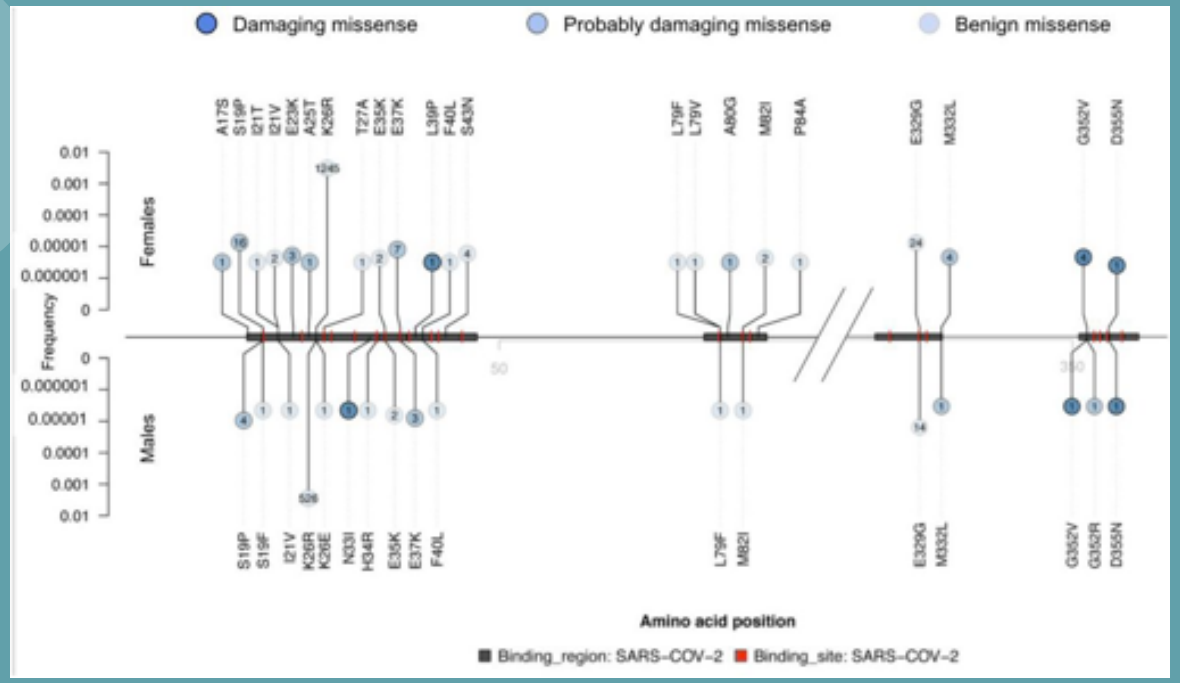
trackViewer



https://github.com/jianhong/trackViewer.documentation/tree/master/trackViewerShinyApp

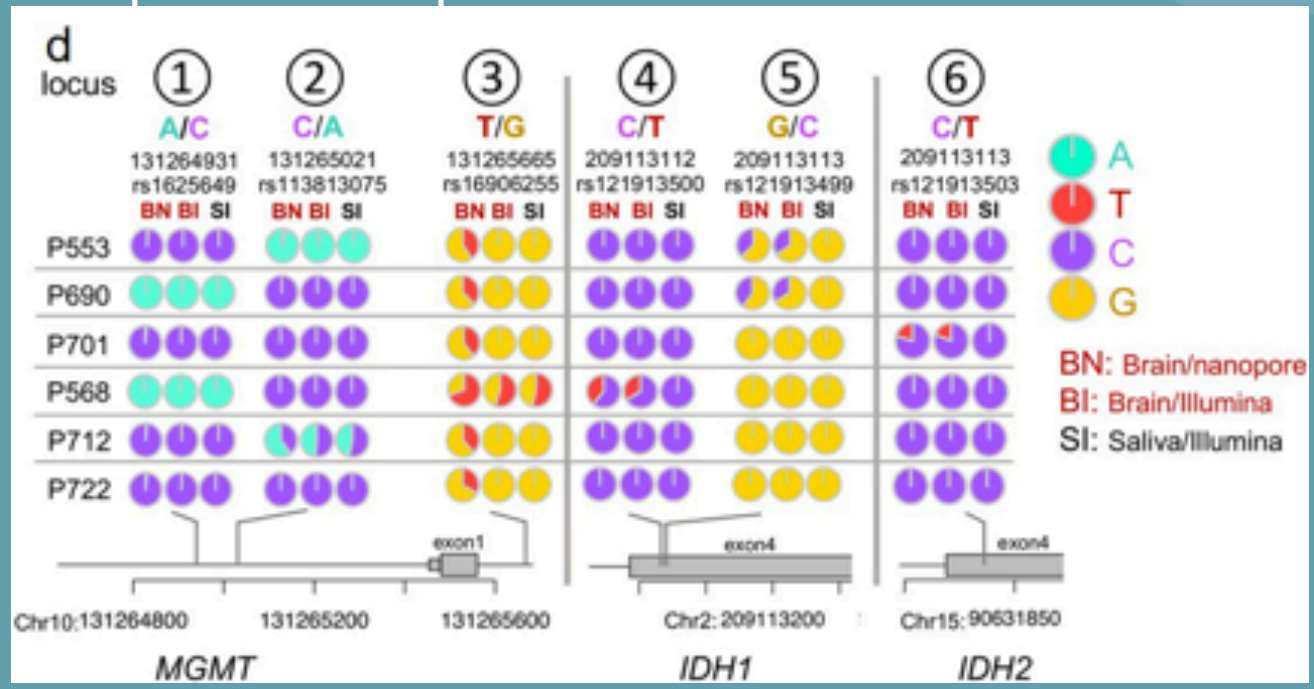
Examples using the `trackViewer` package

Lolliplot of missense variants in the regions of ACE2 that bind to SARS-CoV-2.



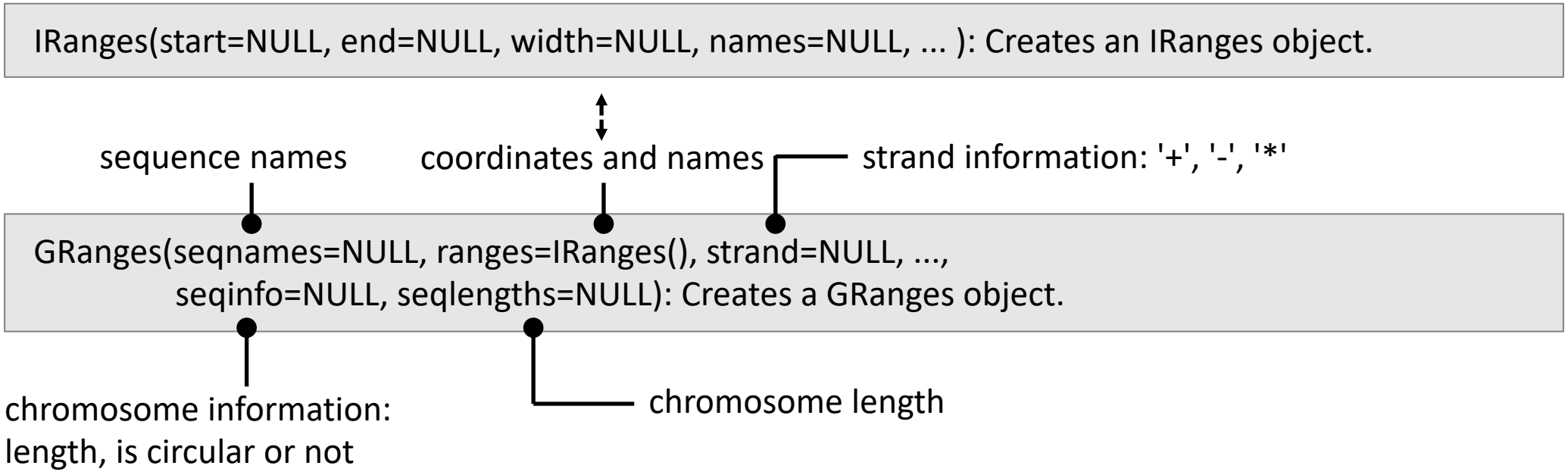
Cirulli et.al., 2020. doi: 10.1101/2020.04.07.030544

SNVs in MGMT and IDH1/2 in tumor and saliva samples from 6 patients



Wongsurawat et.al., 2020. doi: 10.1186/s40478-020-00963-0

GRanges Class



GOTO VIGNETTE

https://jianhong.github.io/trackViewerBiocAsia2020Workshop/

trackViewer CAN ...

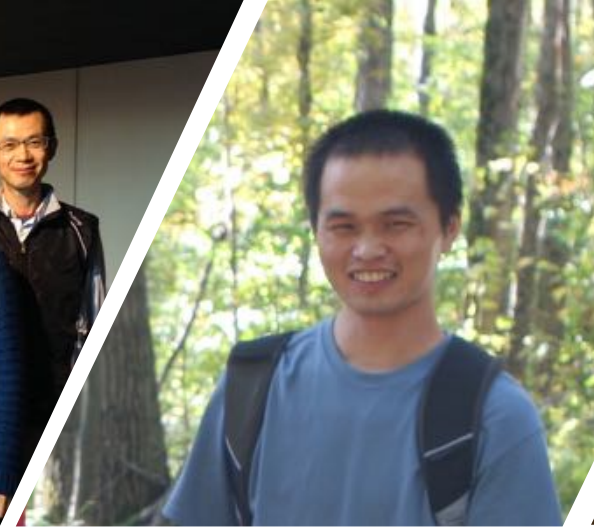
generate lollipop plot to depict the methylation and SNP/mutation status

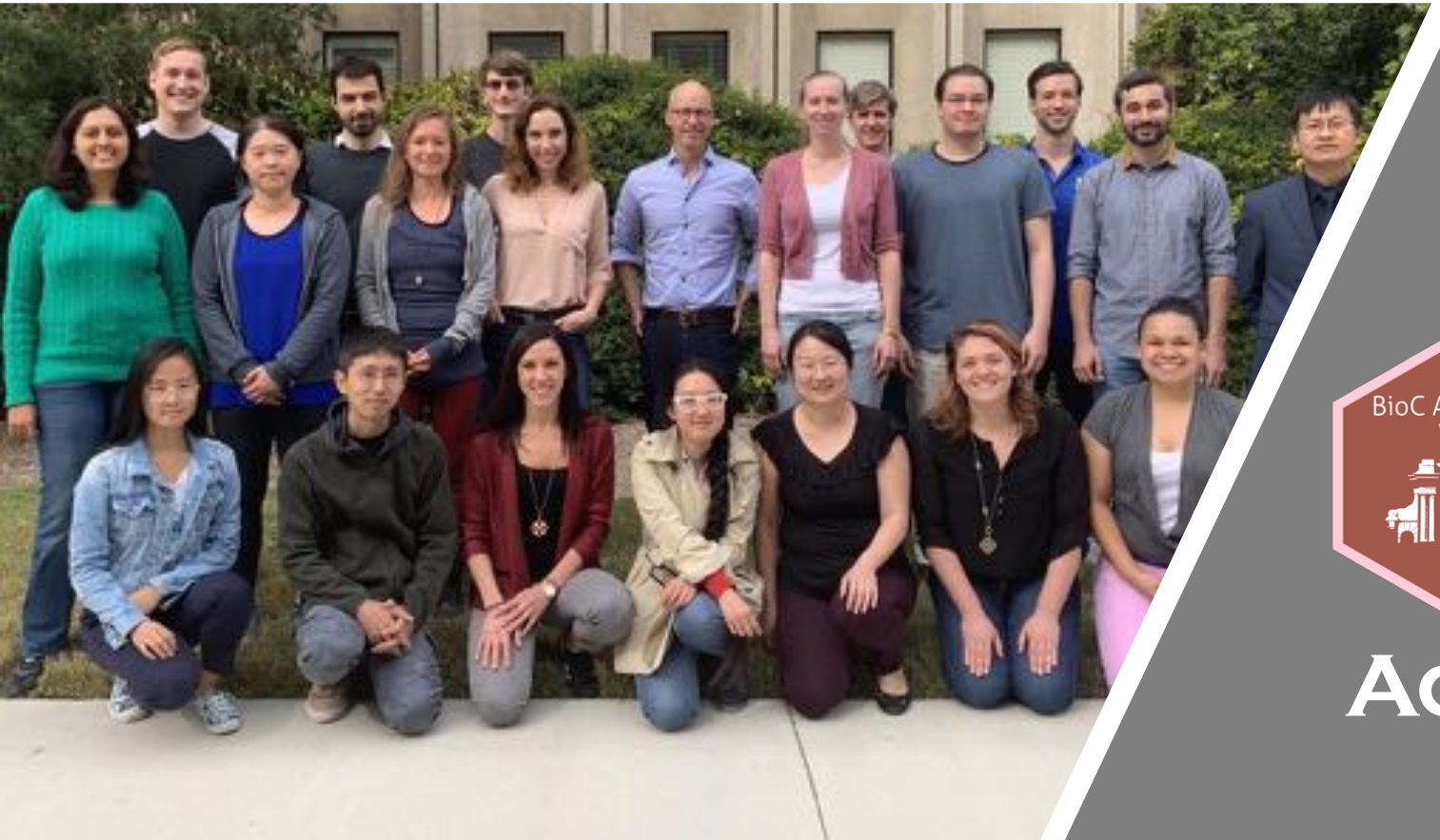
visualize regular read coverage tracks

be easily integrated into standard analysis pipeline for various high-throughput sequencing dataset such as ChIP-seq, RNA-seq, methylation-seq or DNA-seq

produce highly customized figure for publication







BioC Asia 2020
Virtual

regenerationNEXT

Acknowledgement