

## Data preprocessing

### VCF filtering strategy

- GI: Genome Interval
- CI: Caller Information
- PA: Keep or exclude non-PASS tag
- AV: Artifact Variant filter (FFPE)

### Reject and Accept

- Variants in reject list will be removed
- Variants in accept list will pass.

### MAF filtering strategy

- GI: Genome Interval
- CI: Caller Information
- TE: Tissue Expression
- PAC: Population Allele Count
- HY: Hypermutator



### VCF combination

- Union of VCF files
- At least m callers mention data
- At most n callers include non-PASS tag

### VCF to MAF

or



### MAF combination

- Combine all data into one MAF file

## Analysis and visualization

1. Significantly mutated gene detection — *SigMutatedGeneDetection* [OncodriveCLUST]
2. Known cancer gene annotation — *KnownCancerGeneAnnotation*
3. Mutation burden statistics — *TumorMutationBurden*
4. CoMut plot analysis — *CoMutAnalysis*, *CoMutPlot* (CoMut plot)
5. Mutational signature — *MutationalSignature* (Estimation plot, SBS chart, Bar chart, Heatmap, Donut plot)
6. HRD score — *HRDScore* [scarHRD] (Pie chart, Bar chart)
7. Whole-genome doubling (WGD) and Chromosome instability (CIN) — *WGDnCIN* (Pie chart, Bar chart)
8. HRD, CIN and WGD Comparison — *HCWComparison* (Bar chart, Heatmap)
9. Actionable mutation (drug) annotation — *OncoKBAnnotator* [oncokb-annotator] (Pie chart, Bar chart)