

# 1 BIC Variant Pipeline (v2.3.7)

## 1.1 Introduction

The output consists of the raw output files from the TEMPO pipeline along with some simple post-processing of the mutation tables to filter for non-silent events.

## 1.2 Directory Structure

```
{outdir}
├── post
│   ├── reports
│   │   ├── {projectNum}_SNV_Report01_v2.xlsx
│   │   ├── Proj_{projectNum}_facets_v3.xlsx
│   │   ├── Proj_{projectNum}_Filtered_facets_hisens.seg
│   │   └── Proj_{projectNum}_Filtered_facets_purity.seg
│   ├── pipeline_info
│   │   ├── version.txt
│   │   └── output.html
│   └── plots
│       └── facets
│           └── {sample}_purity.CNCF.png
└── tempo
    ├── {projectNum}
    │   ├── bams
    │   ├── cohort_level
    │   ├── runlog
    │   └── somatic
```

## 1.3 Tempo Output

The core output is from the Tempo research pipeline from MSKCC CMO Computational Science group. The raw output from Tempo is in the **tempo** subfolder, and a detailed description of these files can be found here: [//outputs.html](https://tempo.mskcc.org/outputs.html). *N.B.*, to save space, the **bam** and **snp-pileup** files are not delivered by default. If you require them, please contact us.

## 1.4 Post folder

The post folder contains a number of useful files that give an overall summary of the pipeline and results.

### 1.4.1 Reports

The reports folder contains multiple files providing comprehensive analysis results:

**1.4.1.1 SNV Report (`{projectNum}_SNV_Report01_v2.xlsx`)** The SNV (Single Nucleotide Variant) mutation summary report in Excel format with the following three sheets:

1. *Sample Data*: Provides summary information on samples including TMB and sample purity.
2. *Gene Stats*: Lists genes mutated in more than two samples, with mutation count, frequency, and affected samples.
3. *Mutations*: Displays a table of all non-silent mutations detected.

**1.4.1.2 FACETS Report (`Proj_{projectNum}_facets_v3.xlsx`)** A comprehensive Excel report focused on copy number analysis results from FACETS, including sample purity estimates, ploidy information, and copy number profiles.

#### 1.4.1.3 Segmentation Files

- `Proj_{projectNum}_Filtered_facets_hisens.seg`: High sensitivity segmentation data from FACETS analysis
- `Proj_{projectNum}_Filtered_facets_purity.seg`: Purity-based filtered segmentation data from FACETS analysis

These tab-delimited files contain genomic coordinates and copy number segment information for downstream analysis.

### 1.4.2 Pipeline Info

The `pipeline_info` folder contains comprehensive information about the pipeline execution:

- `version.txt`: Details about the software version, timestamp, git repository information, and assay type used for analysis
- `output.html`: HTML version of this output documentation

### 1.4.3 Plots Folder

The plots folder contains copy number profile plots generated by FACET for each sample, providing convenient access to visual representations of copy number alterations.