

Preparing Job File

You need to provide two JSON files that describes your job:

`${sample-name}.inputs.json`

```
{
  "Velopipe.countMatrix": "s3://dp-lab-data/collaborators/pi/project/sample/..._dense.csv",
  "Velopipe.bam": "s3://dp-lab-data/collaborators/pi/project/sample/..._Aligned.out.sorted.bam",
  "Velopipe.bai": "s3://dp-lab-data/collaborators/pi/project/sample/..._Aligned.out.sorted.bam.bai",
  "Velopipe.gtf": "s3://seqc-public/genomes/mm38_long_polya/annotations.gtf",
  "Velopipe.barcodeWhitelist": "s3://seqc-public/barcodes/ten_x_v3/flat/3M-february-2018.txt",
  "Velopipe.alreadySortedBam": true
}
```

- `Velopipe.countMatrix`: a dense gene expression matrix generated by SEQC
- `Velopipe.bam`: BAM file generated by SEQC
- `Velopipe.bai`: BAM index file
- `Velopipe.gtf`: GTF file
 - Human: `s3://seqc-public/genomes/hg38_long_polya/annotations.gtf`
 - Mouse: `s3://seqc-public/genomes/mm38_long_polya/annotations.gtf`
- `Velopipe.barcodeWhitelist`: a barcode whitelist
 - 10x v2: `s3://seqc-public/barcodes/ten_x_v2/flat/737K-august-2016.txt`
 - 10x v3: `s3://seqc-public/barcodes/ten_x_v3/flat/3M-february-2018.txt`
- `Velopipe.alreadySortedBam`: `true` if the BAM file is already **position** sorted

Note that SEQC produces a read-name sorted BAM file (this is different from position sorted). If you are not providing a position sorted BAM file, remove `Velopipe.bai` completely from the configuration.

`${sample-name}.labels.json`

```
{
  "pipelineType": "Velopipe",
  "project": "Project 193",
  "sample": "1469_TGFB_LCC-TRL_1_P193",
  "owner": "chunj",
  "destination": "s3://dp-lab-data/Siting/TGFB_LCC_TRL1",
  "transfer": "-",
  "comment": "RNA Velocity"
}
```

- `project`: project ID retrieved from SCRI database
- `sample`: sample name
- `destination`: AWS S3 location where the final output files (e.g. loom) should be saved