|  |  |  |
| --- | --- | --- |
| **Job** | **Tools with version** | **Parameter** |
| **Alignment** | BWA-MEM 0.7.17, SAMtools view 1.6 (for conversion from SAM to BAM) | Defaults for bwa index (indexing the reference genome)  Defaults for bwa mem (alignment)  –b (for SAMtools view) |
| **BAM file sorting** | Samtools sort 1.6 | Defaults |
| **Marking Duplicates** | Picard 2.14.0 MarkDuplicates | ASSUME\_SORTED=true VALIDATION\_STRINGENCY=SILENT MAX\_FILE\_HANDLES\_FOR\_READ\_ENDS\_MAP=1024 TMP\_DIR=<temp\_directory> METRICS\_FILE=<metrics\_filename> |
| **Adding read groups** | Picard 2.14.0 AddOrReplaceReadGroups | RGLB=library RGPL=illumina RGPU=barcode RGSM=<sample\_name> CREATE\_INDEX=true |
| **Call variants per sample in GVCF mode** | GATK 4.0.4.0 Haplotypecaller | -ERC GVCF |
| **Consolidate GVCFs** | GATK 4.0.4.0 GenomicsDBImport | --TMP\_DIR <temp\_directory> --sample-name-map <sample\_name\_mapping\_file> --reader-threads 2 --genomicsdb-workspace-path <chromosomes\_and unplaced\_scaffolds\_name> -L <chromosomes\_and unplaced\_scaffolds\_name> |
| **Joint call cohort** | GATK 4.0.4.0 GenotypeGVCF | -new-qual -V gendb:// <chromosomes\_and unplaced\_scaffolds\_name> |
| **Concatenate variants called per chromosome/unplaced\_scaffold into a final multi-sample VCF file** | Picard 2.14.0 GatherVcfs | Defaults |
| **Index the final VCF file** | Tabix | Defaults |
| **Get only biallelic SNVs** | BCFtools view 1.6 | -v snps -m2 -M2 |
| **Hard filter biallelic variants** | BCFtools filter 1.6 | -i 'QD >= 15 & FS <= 60 & SOR <= 2 & MQ >= 50 & MQRankSum >= -2.5 & ReadPosRankSum >= -2.5' |