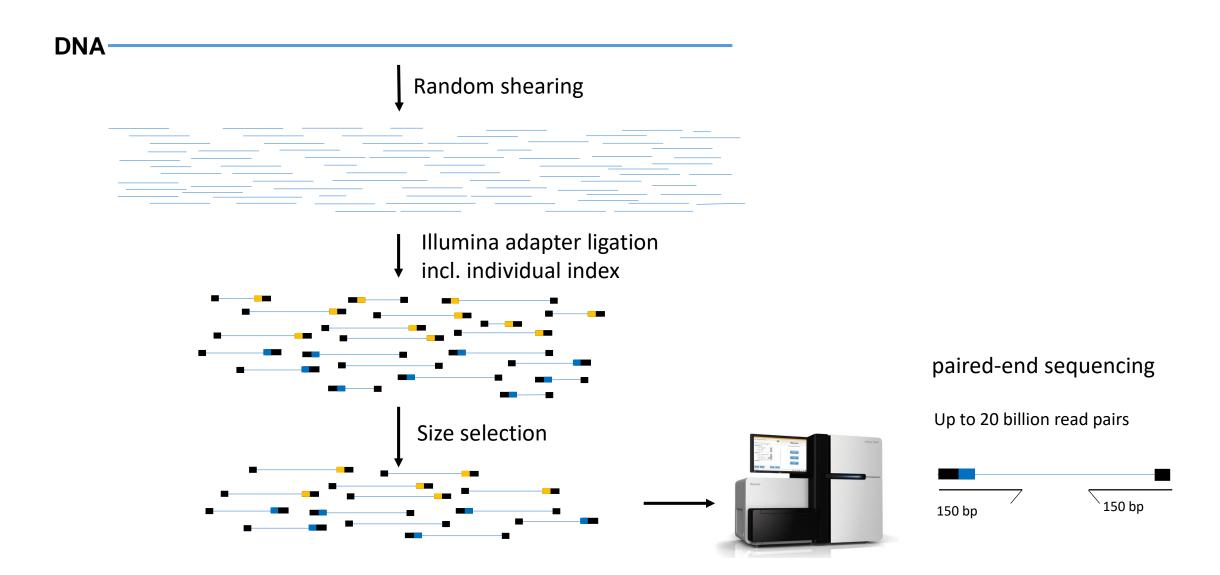
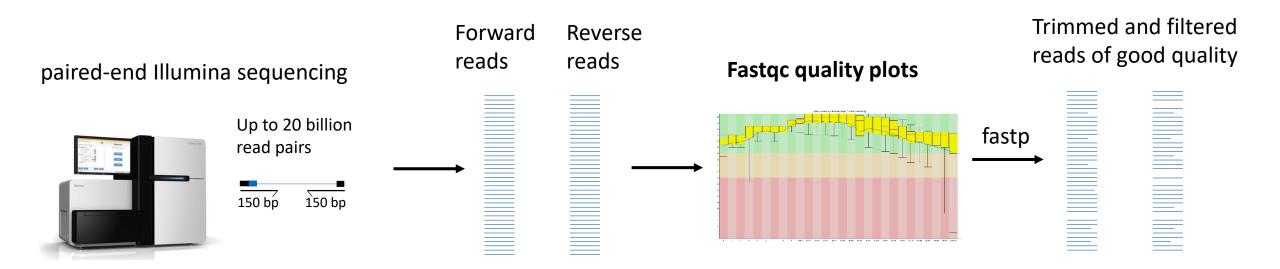
Whole-genome short-read sequencing

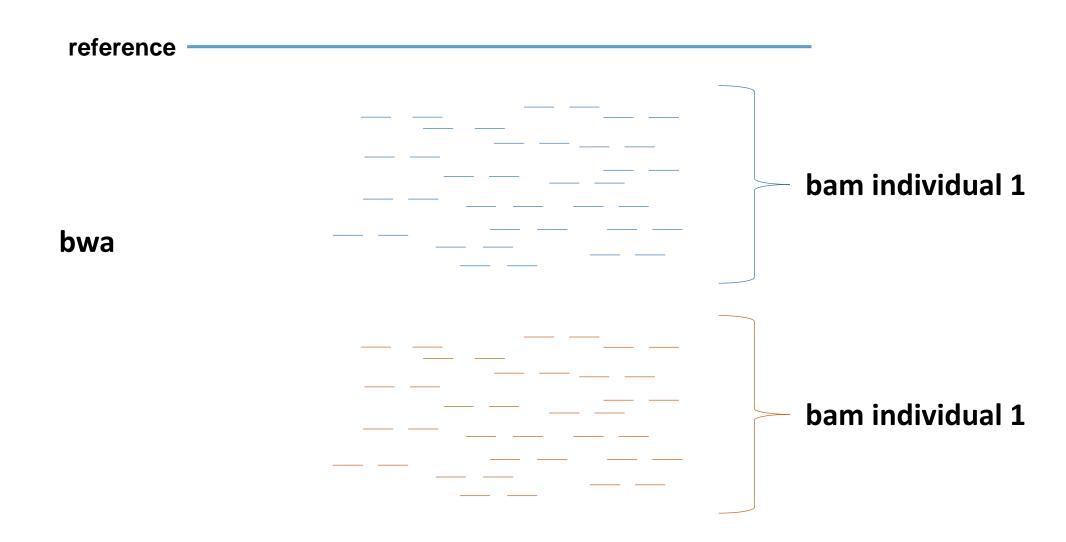


1. Quality check and trimming raw reads

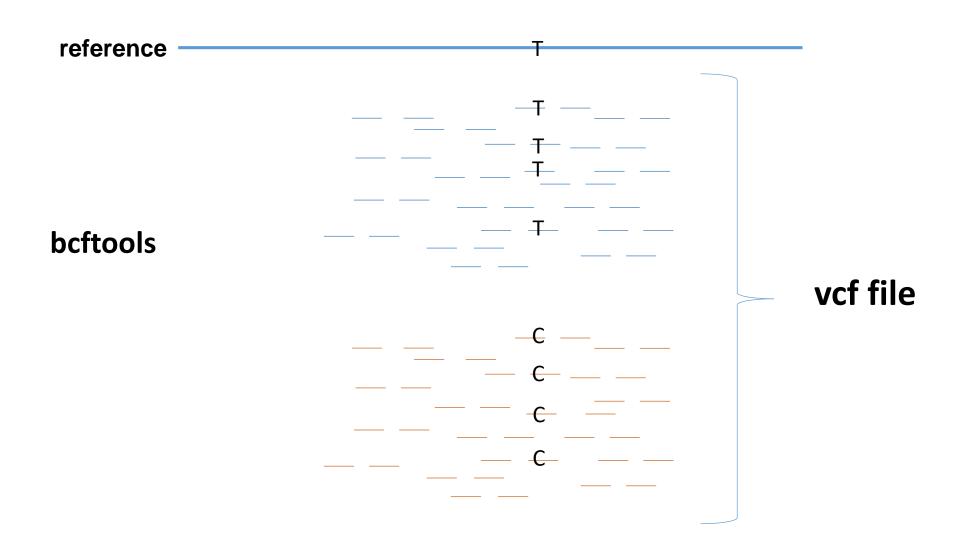


Fastq file of reads

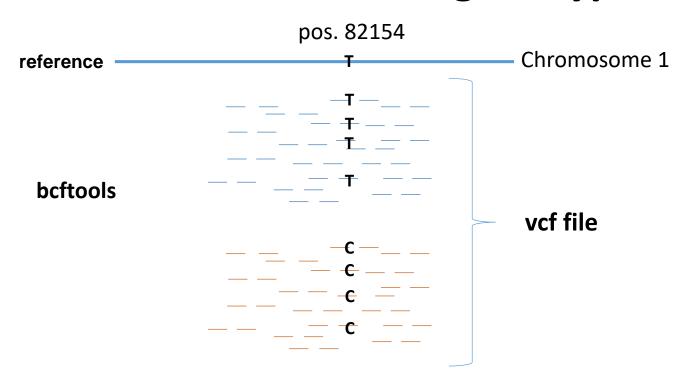
2. Alignment to the reference genome with bwa



3. Variant and genotype calling with bcftools

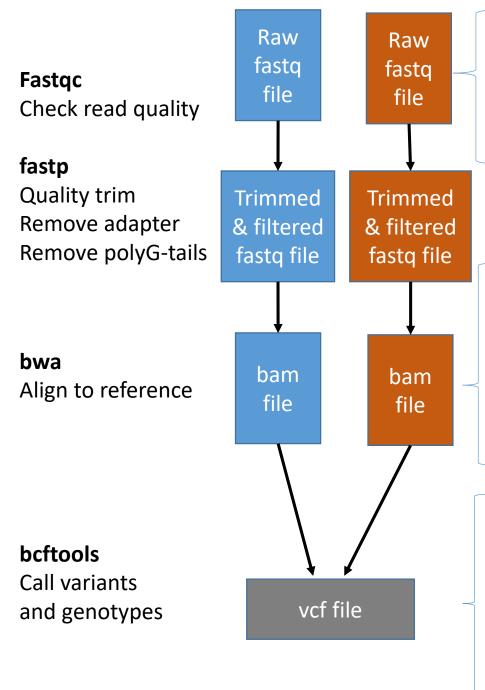


3. Variant and genotype calling with bcftools



vcf file: Genotypes for each individual at genomic sites

##FORMAT= <id=gt, <="" description="Genotype" number="1," th="" type="String,"></id=gt,>												
#CHROM	POS ID	REF	ALT	QUAL		FILTE	RI	NF0	F	ORMA [°]	T G	EN
chr1	82154	•	Т	С		GT	0/0	1/1	0/0	1/1	0/0	1/1
chr1	752566		Т	•		GT	0/0	0/0	0/0	0/0	0/0	0/0
chr1	752721		Т	C		GT	1/1	1/1	1/1	1/1	1/1	1/1
chr1	752721	•	Α			GT	./.	0/0	0/0	0/0	0/0	0/0



Fastq: raw reads with sequencing quality information

sam file: reads mapped to the reference genome -> binary version: bam file

```
HWI-ST1145:74:C101DACXX:7:1114:2759:41961
                                                       193953 50
                                                                     100M
      TGCTGGATCATCTGGTTAGTGGCTTCTGACTCAGAGGACCTTCGTCCCCTGGGGCAGTGGACCTTCCAGTGATTCCCCTGACATAAGGGGCATGGACGA
    DCDDDDEDDDDDDDDDDDDDCCCDDDCDDDDDEEC>DFFFEJJJJJIGJJJJJIHGBHHGJIJJJJJJGJJJJIJJJJJJHHJHHHFFFFFCCC
                 XM:i:3 X0:i:0 XG:i:0
                                     MD:Z:60G16T18T3 NM:i:3 NH:i:1
HWI-ST1145:74:C101DACXX:7:1204:14760:4030
                                                       270877 50
                                                                     100M
      DDDDDDDDDDDDDDDDDDDDDDDDDEEEEEEEFFFEFFEGHHHHFGDJJHJJJJJJJIIIIIGGFJJJHIIIJJJJJJJJGHHFAHGFHJHFGGHFFFDD@BB
   AS:i:-11
                 XM:i:2 X0:i:0 XG:i:0
                                     MD:Z:0A85G13
                                                    NM:i:2 NH:i:1
HWI-ST1145:74:C101DACXX:7:1210:11167:8699
                                                       271218 50
                                                                     50M4700N50M
             GTGGCTCTTCCACAGGAATGTTGAGGATGACATCCATGTCTGGGGTGCACTTGGGTCTCCGAAGCAGAACATCCTCAAATATGACCTCTCG
```

vcf file: Genotypes for each individual at genomic sites

```
##FORMAT=<ID=GT, Number=1, Type=String, Description="Genotype"
         POS ID
                 REF ALT QUAL
                                    FILTER
#CHROM
                                             INF<sub>0</sub>
                                                      FORMAT
        82154
chr1
         752566
chr1
                                      GT
chr1
         752721
                                     GT
chr1
         752721
                                               0/0 0/0 0/0 0/0 0/0
```