AnchorWave to GATK to vcf files

Baoxing Song

2022-08-05

AnchorWave to GATK to vcf files

The aim is perform variant calling for AnchorWave genome alignment results and generate VCF files. So that the genome results from AnchorWave could be used by other popular pipeline for population or quantitative genetics research.

MAF to GVCF

This step was done using the biokotlin plugin under TASSEL. The details could be found by the documents wroten by Tianyi Li.

Reference genome preparsion and GVCF files index

I am using the GATK GenomicsDBImport->GenotypeGVCFs pipeline to combine multiple GVCF files into a single vcf file.

```
wget https://download.maizegdb.org/Zm-B73-REFERENCE-NAM-5.0/Zm-B73-REFERENCE-NAM-5.0.fa.gz
gunzip Zm-B73-REFERENCE-NAM-5.0.fa.gz
sed -i 's/>chr/>/g' Zm-B73-REFERENCE-NAM-5.0.fa
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -1 | tail -1 > interval1.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -2 | tail -1 > interval2.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -3 | tail -1 > interval3.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -4 | tail -1 > interval4.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -5 | tail -1 > interval5.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -6 | tail -1 > interval6.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -7 | tail -1 > interval7.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -8 | tail -1 > interval8.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -9 | tail -1 > interval9.list
grep ">" Zm-B73-REFERENCE-NAM-5.0.fa | sed 's/>//g' | head -10 | tail -1 > interval10.list


ls *gvcf | awk '{print("bgzip "$1)}' > bgzipCommands
parallel -j 30 < bgzipCommands
ls *gvcf.gz | awk '{print("tabix -p vcf "$1)}' > tabcommands
parallel -j 30 < tabcommands



samtools faidx Zm-B73-REFERENCE-NAM-5.0.fa
wget https://github.com/broadinstitute/picard/releases/download/2.26.10/picard.jar
java -jar picard.jar CreateSequenceDictionary R=Zm-B73-REFERENCE-NAM-5.0.fa O=Zm-B73-REFERENCE-NAM-5.0.d
```

GenomicsDBImport

It is time consuming to generate a single database for the whole genome. And it is hard to debug.
Here we created a database for each chromosome separately.

```
gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport \
        -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML103ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML228ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML247ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML277ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML322ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML333ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML52ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML69ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/HP301ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Il14HToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki11ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki3ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC350ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC358ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh43ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh7BToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/P39ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tx303ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gvcf.gz \
        --batch-size 5 \
    --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk1 \
        --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 1 #done



gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvc


gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gv



gatk --java-options "-Xmx280g -Xms280g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvc

gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/CML322ToB73.

gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gv


gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gv
```

```
gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gv

gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/Il14HToB73.g


        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        --batch-size 5 \
    --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk1 \
    --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 1 #done




gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvc



gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport \
        -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML103ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML228ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML247ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML277ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML322ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML333ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML52ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML69ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/HP301ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Il14HToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki11ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki3ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC350ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC358ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh43ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh7BToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/P39ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tx303ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gvcf.gz \
        --batch-size 5 \
    --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk5 \
    --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 5 #done
```

```
gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport \
        -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML103ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML228ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML247ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML277ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML322ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML333ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML52ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML69ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/HP301ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Il14HToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki11ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki3ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC350ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC358ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh43ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh7BToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/P39ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tx303ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gvcf.gz \
        --batch-size 5 \
    --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk6 \
    --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 6 #done


gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport \
        -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML103ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML228ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML247ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML277ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML322ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML333ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML52ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML69ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/HP301ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Il14HToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki11ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki3ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC350ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC358ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh43ToB73.gvcf.gz \
```

```
            -V /home/xuql/NAM_anchorwave_song/Oh7BToB73.gvcf.gz \
            -V /home/xuql/NAM_anchorwave_song/P39ToB73.gvcf.gz \
            -V /home/xuql/NAM_anchorwave_song/Tx303ToB73.gvcf.gz \
            -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gvcf.gz \
            --batch-size 5 \
        --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk7 \
        --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 7  #done


gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport \
        -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML103ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML228ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML247ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML277ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML322ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML333ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML52ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML69ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/HP301ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Il14HToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki11ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki3ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC350ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC358ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh43ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh7BToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/P39ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tx303ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gvcf.gz \
        --batch-size 5 \
        --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk8 \
        --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 8   #done


gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport \
        -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML103ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML228ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML247ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML277ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML322ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML333ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML52ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML69ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/HP301ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Il14HToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki11ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki3ToB73.gvcf.gz \
```

```
        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC350ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC358ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh43ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh7BToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/P39ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tx303ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gvcf.gz \
        --batch-size 5 \
    --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk9 \
    --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 9 #done


gatk --java-options "-Xmx480g -Xms480g" GenomicsDBImport \
        -V /home/xuql/NAM_anchorwave_song/B97ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML103ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML228ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML247ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML277ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML322ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML333ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML52ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/CML69ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/HP301ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Il14HToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki11ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ki3ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ky21ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M162WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/M37WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Mo18WToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Ms71ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC350ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/NC358ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh43ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Oh7BToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/P39ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tx303ToB73.gvcf.gz \
        -V /home/xuql/NAM_anchorwave_song/Tzi8ToB73.gvcf.gz \
        --batch-size 5 \
    --genomicsdb-workspace-path /home/xuql/NAM_anchorwave_song/NAM_out_gatk10 \
    --genomicsdb-segment-size 1048576000 --genomicsdb-vcf-buffer-size 10000000000 -L 10 #done
```

Generate VCF files from the database

I was having `buffer overflow` problems to generate VCF file for chr9 and chr10. So we have to skip some
intervals at this moment. I reported this problem to the GATK team. https://github.com/broadinstitute/ga
tk/issues/7976 Hopefully, they will provide helps to us.

```
gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy
```

```
gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy

gatk --java-options "-Xmx100g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy


gatk --java-options "-Xmx10g" GenotypeGVCFs -R Zm-B73-REFERENCE-NAM-5.0.fa -stand-call-conf 0 -ploidy 1
```

todo

Merge those vcf files together and perform left alignment.

vcf files merging has not been tested yet. Tianyi tested the GATK left alignment function. ## NOTE The VT and GATK left alignment functions did not work well on the GVCF files generated by biokotlin.