

# Practical Haplotype Graph (PHG)

Wheat RefSeq\_v1

## Setup

1. get jar files from Buckler Lab (phg.jar and sTASSEL.jar)
2. edit scripts/LoadGenomeIntervals.sh to change Xmx=50G
3. build image
4. create BED file from high confidence genes

## Pipeline

1. **LoadGenomeIntervals.sh**
  - a. TASSEL requires 50G for this assembly, add Xmx=50G to configSQLite.txt
2. **LoadAssemblyAnchors.sh** - do not use, it is being rewritten
3. **CreateHaplotypes.sh**
  - a. change gatk-launch to gatk
  - b. fix arguments
  - c. index documents - <https://gatkforums.broadinstitute.org/gatk/discussion/1601/how-can-i-prepare-a-fasta-file-to-use-as-reference>
    - i. java -jar picard.jar CreateSequenceDictionary R=/data2/sra/161010\_Chinese\_Spring\_v1.0\_pseudomolecules.fasta  
java -jar picard.jar CreateSequenceDictionary R=tmpFileDir/data/reference/161010\_Chinese\_Spring\_v1.0\_pseudomolecules.fasta  
fasta  
samtools faidx 161010\_Chinese\_Spring\_v1.0\_pseudomolecules.fasta
  - d. intervals.bed - change contigs from 1A to chr1A and so on