**Tutorial PEDCA (Ploidy Estimation by Dynamic Coverage Analysis)**

Pedca is a ploidy estimation algorithm that infers copy number of the contigs submitted as input based on the read coverage that aligns to them. It requires as an input an alignment file in .bam or .sam format of a library or set of libraries aligned to a reference file, containing all the contigs that will be estimated.

Pre-processing the data (5 steps)

We need to align the reads against a reference.

Step 1. Index your reference.

Example using bwa (all command in one single line):

<*path\_to\_bwa\_aligner*>/bwa index -a bwtsw *<path\_to\_reference\_file/your\_reference.fasta>*

Step 2. Align your reads to your reference

Example using bwa and paired end reads (all command in one single line):

*< bwa\_aligner\_path>****/*bwa mem** *<path\_reference\_file/your\_reference.fasta>*  *<readsPath/readsPairEnd1.fasta> < readsPath /readsPairEnd2.fasta>* **>** *<destination\_folder* /*example.sam*>

Step 3. You might want to transform your .sam file into a .bam format

Example using samTools (all command in one single line):

*<samToolsPath>* **/samtools view -Sb** *<destination\_folder* /*example.sam*> **>** *<destination\_folder* /*example.bam*>

Pedca just accepts one input file. If you have several libraries you can put all your bam files in a folder (or create a folder with symbolic links to all files you want to merge) and then:

*<samToolsPath>***/samtools merge** <*bam***\_***destination\_folder*/finalBamFile.bam> **\*.bam**

Step 4. Sort the .bam/.sam file

Example sorting a .bam file using samTools (all command in one single line):

*<samToolsPath>* **/samtools sort -o** *<destination\_folder* /sorted\_*example.bam*>  **-O bam -T** *<temp\_folderPath*/*tempName*> *<destination\_folder* /*example.bam*>

Step 5. Index the sorted .bam/.sam file

Example indexing a sorted bam file using samTools (all command in one single line):

*<samToolsPath>* **/samtools index** *<destination\_folder* /sorted\_*example.bam*>