Fonio – smc++ Analysis

1. Establishing masks files

We will generate masks files as in msmc analysis. To this purpose, we use the individual bam files.

We have to index bam files first. Bam files for all cultivated genotypes were retrieved on a node.

for i in \*;do samtools index $i; done

We also have to calculate their mean sequencing coverage (calculated for chromosome 1 only)

for i in \*.bam;do echo $i; samtools depth -r Dexi\_CM05836\_chr01A $i | awk '{sum += $3} END {print sum / NR}'; done

This results in the following table :

|  |  |
| --- | --- |
| CH1 | 36.9823 |
| CH2 | 38.7963 |
| CH3 | 41.3432 |
| CH4 | 40.219 |
| CM03380 | 42.3589 |
| CM03382 | 39.2431 |
| CM03390 | 34.9306 |
| CM03394 | 39.8188 |
| CM03396 | 32.6504 |
| CM03400 | 37.3045 |
| CM03403 | 38.148 |
| CM03406 | 30.3615 |
| CM03409 | 34.8898 |
| CM03418 | 37.1527 |
| CM03423 | 40.3217 |
| CM03424 | 34.2292 |
| CM03430 | 42.6802 |
| CM03431 | 33.4579 |
| CM03434 | 35.9569 |
| CM03437 | 36.4776 |
| CM03438 | 36.8882 |
| CH1 | 36.9823 |
| CH2 | 38.7963 |
| CH3 | 41.3432 |
| CH4 | 40.219 |
| CM05740 | 38.0802 |
| CM05741 | 39.8058 |
| CM05742 | 38.7123 |
| CM05743 | 33.4318 |
| CM05744 | 37.0117 |
| CM05746 | 32.5717 |
| CM05750 | 36.3727 |
| CM05754 | 32.6783 |
| CM05760 | 48.0135 |
| CM05761 | 39.1567 |
| CM05767 | 32.6268 |
| CM05770 | 28.2194 |
| CM05771 | 30.0599 |
| CM05772 | 31.6816 |
| CM05775 | 30.3252 |
| CM05780 | 31.4678 |
| CM05792 | 31.1121 |
| CM05793 | 33.2765 |
| CM05796 | 31.7906 |
| CM05798 | 34.0828 |
| CM05802 | 29.2789 |
| CM05803 | 28.7477 |
| CM05806 | 32.5678 |
| CM05808 | 35.5561 |
| CM05810 | 31.3953 |
| CM05811 | 33.9755 |
| CM05815 | 31.1212 |
| CM05819 | 33.3287 |
| CM05824 | 35.2749 |
| CM05827 | 32.7049 |
| CM05830 | 35.9794 |
| CM05834 | 32.565 |
| CM05835 | 29.7628 |
| CM05836-C | 40.7102 |
| CM05836-D | 37.7582 |
| CM05839 | 32.8351 |
| CM05840 | 34.2162 |
| CM05841 | 29.4823 |
| CM05842 | 30.0799 |
| CM05843 | 36.0541 |
| CM05844 | 32.8775 |
| CM05847 | 29.5971 |
| CM05849 | 30.4135 |
| CM05853 | 31.4662 |
| CM05854 | 31.4324 |
| CM05855 | 31.6745 |
| CM05856 | 38.7523 |
| CM05857 | 32.0226 |
| CM05858 | 34.6263 |
| CM05863 | 31.0053 |
| CM05865 | 30.6186 |
| CM05869 | 35.4548 |
| CM05870 | 34.3284 |
| CM06496 | 36.6635 |
| CM06501 | 32.7119 |
| CM06505 | 33.3262 |
| CM06510 | 29.88 |
| CM06511 | 31.9704 |
| CM06513 | 32.1095 |
| CM07226 | 35.0701 |
| CM07234 | 34.4326 |
| CM07244 | 34.1424 |
| CM07246 | 33.6162 |
| CM07249 | 40.1842 |
| CH1 | 36.9823 |
| CH2 | 38.7963 |
| CH3 | 41.3432 |
| CH4 | 40.219 |
| CM07268 | 34.7297 |
| CM07270 | 40.3785 |
| CM07280 | 28.3979 |
| CM07281 | 38.0932 |
| CM07289 | 30.5235 |
| CM07292 | 28.2902 |
| CM07295 | 30.5992 |
| CM07310 | 31.7082 |
| CM07317 | 30.278 |
| CM07319 | 30.3319 |
| CM07321 | 33.3567 |
| CM07327 | 31.9674 |
| CM07329 | 31.6509 |
| CM07338 | 31.206 |
| CM07340 | 35.0714 |
| CM07342 | 29.7583 |
| CM07346 | 36.2485 |
| CM07350 | 32.683 |
| CM07354 | 33.6656 |
| CM07360 | 37.5521 |
| CM07885 | 29.6212 |
| CM07888 | 44.0049 |
| CM07890 | 33.1311 |
| CM07891 | 29.7695 |
| CM07892 | 39.964 |
| CM07895 | 30.7162 |
| CM07900 | 29.1774 |
| CM07901 | 34.6936 |
| CM07902 | 32.9944 |
| CM07903 | 28.6976 |
| CM07908 | 30.5691 |
| CM07909 | 37.7569 |
| CM08512 | 32.0162 |
| CM08517 | 30.5986 |
| CM08524 | 29.8363 |
| CM08525 | 31.7827 |
| CM08526 | 27.8712 |
| CM08537 | 31.5073 |
| CM08551 | 29.2926 |
| CM08562 | 33.0245 |
| CM08563 | 33.1462 |
| CM08573 | 33.0975 |
| CM08593 | 36.5354 |
| CM08598 | 31.2429 |
| CM08613 | 38.8658 |
| CM08617 | 37.1236 |
| CM08627 | 29.1192 |
| CM08628 | 35.1288 |
| CM08635 | 31.8591 |
| CM08642 | 35.2638 |
| CM08652 | 40.9889 |
| CM08653 | 32.5138 |
| CM08655 | 35.6605 |
| CM08665 | 34.8455 |
| CM08669 | 32.4168 |
| CM08671 | 34.3295 |
| CM08672 | 36.9261 |
| CM08677 | 36.7081 |
| CM08681 | 43.1989 |
| CM08684 | 30.4744 |
| CM12246 | 40.7621 |

We iterate on the individuals names to apply msmc\_tools to it.

while read NAME DEPTH; do qsub /data2/projects/africrop\_models/fonioWGS/smcpp/run\_bamCaller\_multipleChr\_fonio.bash $NAME $DEPTH ; done < /data2/projects/africrop\_models/fonioWGS/smcpp/fonio\_cult\_depth.txt

This product a mask containing “SNPable” positions according to the msmc\_tools approach. Note that this produce as many mask as chromosomes per individuals, and we want to merge all individual masks for each chromosome in only one mask per chromosome. As we want a mask that indicates positions that are NOT SNPable, we have to substract these masks to the reference sequence in order to produce usable masks. In order to do that, we first create a bed file that contains the different regions with their lengths as exhibited by the fasta reference file.

We then use bedtools subtract option to remove all the “SNPable” positions from the reference file, by iterating over the different masks.

bedtools subtract -a /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/150419\_Digitaria\_Exilis\_v1.0\_pseudomolecules.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr01A.mask.bed.gz > tmp\_mask1A.bed

bedtools subtract -a /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/tmp\_mask1A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr01A.mask.bed.gz > tmp\_mask1Ab.bed

bedtools subtract -a tmp\_mask1Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr01B.mask.bed.gz > tmp\_mask1B.bed

bedtools subtract -a tmp\_mask1B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr01B.mask.bed.gz > tmp\_mask1Bb.bed

bedtools subtract -a tmp\_mask1Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr02A.mask.bed.gz > tmp\_mask2A.bed

bedtools subtract -a tmp\_mask2A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr02A.mask.bed.gz > tmp\_mask2Ab.bed

bedtools subtract -a tmp\_mask2Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr02B.mask.bed.gz > tmp\_mask2B.bed

bedtools subtract -a tmp\_mask2B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr02B.mask.bed.gz > tmp\_mask2Bb.bed

bedtools subtract -a tmp\_mask2Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr03A.mask.bed.gz > tmp\_mask3A.bed

bedtools subtract -a tmp\_mask3A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr03A.mask.bed.gz > tmp\_mask3Ab.bed

bedtools subtract -a tmp\_mask3Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr03B.mask.bed.gz > tmp\_mask3B.bed

bedtools subtract -a tmp\_mask3B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr03B.mask.bed.gz > tmp\_mask3Bb.bed

bedtools subtract -a tmp\_mask3Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr04A.mask.bed.gz > tmp\_mask4A.bed

bedtools subtract -a tmp\_mask4A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr04A.mask.bed.gz > tmp\_mask4Ab.bed

bedtools subtract -a tmp\_mask4Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr04B.mask.bed.gz > tmp\_mask4B.bed

bedtools subtract -a tmp\_mask4B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr04B.mask.bed.gz > tmp\_mask4Bb.bed

bedtools subtract -a tmp\_mask4Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr05A.mask.bed.gz > tmp\_mask5A.bed

bedtools subtract -a tmp\_mask5A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr05A.mask.bed.gz > tmp\_mask5Ab.bed

bedtools subtract -a tmp\_mask5Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr05B.mask.bed.gz > tmp\_mask5B.bed

bedtools subtract -a tmp\_mask5B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr05B.mask.bed.gz > tmp\_mask5Bb.bed

bedtools subtract -a tmp\_mask5Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr06A.mask.bed.gz > tmp\_mask6A.bed

bedtools subtract -a tmp\_mask6A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr06A.mask.bed.gz > tmp\_mask6Ab.bed

bedtools subtract -a tmp\_mask6Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr06B.mask.bed.gz > tmp\_mask6B.bed

bedtools subtract -a tmp\_mask6B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr06B.mask.bed.gz > tmp\_mask6Bb.bed

bedtools subtract -a tmp\_mask6Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr07A.mask.bed.gz > tmp\_mask7A.bed

bedtools subtract -a tmp\_mask7A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr07A.mask.bed.gz > tmp\_mask7Ab.bed

bedtools subtract -a tmp\_mask7Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr07B.mask.bed.gz > tmp\_mask7B.bed

bedtools subtract -a tmp\_mask7B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr07B.mask.bed.gz > tmp\_mask7Bb.bed

bedtools subtract -a tmp\_mask7Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr08A.mask.bed.gz > tmp\_mask8A.bed

bedtools subtract -a tmp\_mask8A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr08A.mask.bed.gz > tmp\_mask8Ab.bed

bedtools subtract -a tmp\_mask8Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr08B.mask.bed.gz > tmp\_mask8B.bed

bedtools subtract -a tmp\_mask8B.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr08B.mask.bed.gz > tmp\_mask8Bb.bed

bedtools subtract -a tmp\_mask8Bb.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr09A.mask.bed.gz > tmp\_mask9A.bed

bedtools subtract -a tmp\_mask9A.bed -b indiv\_masks\_wild/\*Dexi\_CM05836\_chr09A.mask.bed.gz > tmp\_mask9Ab.bed

bedtools subtract -a tmp\_mask9Ab.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr09B.mask.bed.gz > tmp\_fonio\_mask.bed

bedtools subtract -a tmp\_fonio\_mask.bed -b indiv\_masks\_cult/\*Dexi\_CM05836\_chr09B.mask.bed.gz > all\_fonio\_mask\_final.bed

Creation of input files for smcpp, we used a set a 20 randomly chosen genotypes as distinguished lineages

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CH2,CM05815,CM07329,CM03380,CM05830,CM07350,CM05841,CM03406,CM03418,CM07892,CM03430,CM07901,CM05857,CM03438,CM07909,CM08525,CM05740,CM06505,CM05750,CM08627};

do smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/cult\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_cultivated-only.recode.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/TEST/chr$c\.smc.masked.fonio\_cultivatedTEST.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated.csv`;

done

done

Finally, smc++ analysis is made using the “estimate” and “cv” functions.

Analysis was also made for the wild individuals. Here are the results for the depth analysis :

|  |  |
| --- | --- |
| 67996 | 20.2357 |
| IfDlon12 | 3.37375 |
| IfDlon1 | 11.3163 |
| L121 | 18.3399 |
| L17 | 11.0958 |
| L3027 | 8.47104 |
| L4148 | 7.42473 |
| L8249 | 13.3405 |
| L9512 | 9.97875 |
| L953 | 8.43502 |
| M1070 | 5.701 |
| M883 | 7.78894 |
| M905 | 14.854 |
| M949 | 14.3745 |

The analysis was restricted to only the G1 wild group (the ones that are the closest to the cultivated according to chloroplast data):

67996 If-Dlon1 If-Dlon12 L17 L3027 L4148 L8249 L9512 M905 M949 M1070

EDIT 7 november 2019

We ran SMC++ on the 6 different cultivated groups.

Creation of input files for smcpp, we used all genotypes as distinguished lineages within each group

Group1

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM05841,CM05843,CM05856,CM05842,CM06496,CM05857,CM05844,CM05839,CM05840,CM05853,CM05835,CM06501,CM05858,CM05847,CM05849,CM05836-C};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Group1/chr$c\.smc.masked.fonio\_cultivatedG1.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_G1.csv`

done

done

Group2

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CH4,CM03394,CH2,CM07903,CM07908,CM07890,CH1,CM03382,CM07900,CH3,CM03409,CM03396,CM04489,CM07892,CM03390,CM03418,CM03403,CM03406,CM04493,CM03400,CM04487,CM07885};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Group2/chr$c\.smc.masked.fonio\_cultivatedG2.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_G2.csv`

done

done

Group3

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07244,CM07249,CM07327,CM07329,CM08665};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Group3/chr$c\.smc.masked.fonio\_cultivatedG3.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_G3.csv`

done

done

Group4

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07354,CM08628,CM08635,CM08627,CM08613,CM05792,CM07350,CM07340,CM07342,CM08655,CM07346,CM08642,CM07360,CM05796,CM07246,CM08652,CM08653,CM07234,CM07338,CM05811,CM05819,CM05810,CM05772,CM08669,CM05808,CM07268,CM08537};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Group4/chr$c\.smc.masked.fonio\_cultivatedG4.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_G4.csv`

done

done

Group5

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07244,CM07249,CM07327,CM07329,CM08665};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Group5/chr$c\.smc.masked.fonio\_cultivatedG5.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_G5.csv`

done

done

Group6

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07244,CM07249,CM07327,CM07329,CM08665};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Group6/chr$c\.smc.masked.fonio\_cultivatedG6.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_G6.csv`

done

done

UPDATE 19 november 2019 : Analysis made also on several samples of 77 individuals. 20 dist individuals were selected from the sampled accessions

Sample1

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07908,CM07354,CM05808,CM05819,CM03439,CM03430,CM05857,CM07270,CM07350,CM07258,CM08671,CM08681,CM08652,CM05854,CM08537,CM08512,CM05870,CM03390,CM03424,CM05865};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample1/chr$c\.smc.masked.fonio\_cultivatedSample1.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample1.csv`

done

done

Sample2

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM05858,CM05834,CM07317,CM07901,CM03438,CM07266,CM08525,CM03423,CM05754,CM07226,CM05806,CM05824,CM08598,CM05853,CM03418,CM05760,CM08512,CM07270,CM07895,CM05802};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample2/chr$c\.smc.masked.fonio\_cultivatedSample2.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample2.csv`

done

done

Sample3

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07891,CM05810,CM03380,CM08635,CM05830,CM05754,CM08642,CM08551,CH3,CM08653,CM07903,CM08677,CM07292,CM05839,CM05819,CM07321,CM08672,CM08524,CM05806,CM07342};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample3/chr$c\.smc.masked.fonio\_cultivatedSample3.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample3.csv`

done

done

Sample4

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM05857,CM05798,CM03437,CM06510,CM05849,CM05853,CM05810,CM06496,CM07280,CM08598,CM07268,CM05870,CM07295,CM05840,CM08593,CM05744,CM07895,CM03403,CM07289,CM07350};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample4/chr$c\.smc.masked.fonio\_cultivatedSample4.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample4.csv`

done

done

Sample5

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM08613,CH3,CM05815,CM08669,CM07280,CM08681,CM07246,CM05855,CM12246,CM06496,CM03424,CM07354,CM08665,CM06510,CM05853,CM07321,CM07903,CM07327,CM05857,CM08635};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample5/chr$c\.smc.masked.fonio\_cultivatedSample5.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample5.csv`

done

done

Sample6

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM05853,CM05743,CM07360,CM03406,CM07895,CM07340,CM07292,CM07295,CM07270,CM03431,CM08524,CM05819,CM05798,CM05839,CM03424,CM05750,CM05741,CM08562,CM07310,CM05843};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample6/chr$c\.smc.masked.fonio\_cultivatedSample6.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample6.csv`

done

done

Sample7

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07900,CM08524,CM08652,CM07902,CM07888,CM05808,CM05793,CM07350,CM06511,CM07234,CM07903,CM08512,CM05780,CM05767,CM07354,CM07226,CM08551,CM08613,CM05836-C,CM05734};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample7/chr$c\.smc.masked.fonio\_cultivatedSample7.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample7.csv`

done

done

Sample8

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM05834,CM08627,CM03438,CM08653,CM06505,CM05869,CM07310,CM07292,CM05855,CM07895,CM07234,CM05853,CM07249,CM07340,CM05761,CM03400,CM07360,CM07350,CM05839,CM08613};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample8/chr$c\.smc.masked.fonio\_cultivatedSample8.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample8.csv`

done

done

Sample9

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM05836-D,CM05863,CM05737,CM03382,CM05847,CM05842,CM08563,CM05741,CM06511,CM07902,CM08512,CM07289,CM05865,CM05771,CM07266,CM05870,CM03423,CM05796,CM03396,CM05811};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample9/chr$c\.smc.masked.fonio\_cultivatedSample9.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample9.csv`

done

done

Sample10

for c in {01A,01B,02A,02B,03A,03B,04A,04B,05A,05B,06A,06B,07A,07B,08A,08B,09A,09B}

do

for i in {CM07891,CM08635,CM05806,CM07350,CM07268,CM06496,CM07280,CM05836-C,CM05754,CM08627,CM05808,CM05830,CM05865,CM05827,CM07310,CM07908,CM07338,CM03418,CM05856,CM08524};

do /home/cubry/miniconda3/bin/smc++ vcf2smc -m /data2/projects/africrop\_models/fonioWGS/SNPable\_masks/all\_fonio\_mask\_final.bed.gz -d $i $i /data2/projects/africrop\_models/fonioWGS/Fonio\_filtered-final.vcf.gz /data2/projects/africrop\_models/fonioWGS/smcpp/Sample10/chr$c\.smc.masked.fonio\_cultivatedSample10.$i\.gz Dexi\_CM05836\_chr$c `cat /data2/projects/africrop\_models/fonioWGS/smcpp/list\_fonio\_cultivated\_Sample10.csv`

done

done