



HAL
open science

TraceAncestor: a Tool to Infer Unphased Phylogenomic Karyotypes of Admixed Plant Genomes

Franck Curk, Aurore Comte, Dalel Ahmed, Alexis Dereeper, Patrick Ollitrault

► To cite this version:

Franck Curk, Aurore Comte, Dalel Ahmed, Alexis Dereeper, Patrick Ollitrault. TraceAncestor: a Tool to Infer Unphased Phylogenomic Karyotypes of Admixed Plant Genomes. 5. Conference on Plant Genome Evolution, Sep 2019, Sitges, Spain. , 1 p., 2019. hal-02789659

HAL Id: hal-02789659

<https://hal.inrae.fr/hal-02789659v1>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

TraceAncestor: a Tool to Infer Unphased Phylogenomic Karyotypes of Admixed Plant Genomes

Franck CURK, Cirad, UMR Agap, F-34398 Montpellier, France. Agap, Univ Montpellier, Cirad, Inra, Montpellier SupAgro, Montpellier, France.
 Aureo COMTE, IRD, UMR IPME, F-34394 Montpellier, France. IPME IRD, Cirad, Université de Montpellier.
 Dalel AHMED, Cirad, UMR Agap, F-20230 San Giuliano, France. Agap Univ Montpellier, Cirad, Inra, Montpellier SupAgro, Montpellier, France.
 Alexis DEREPPER, IRD, UMR IPME, F-34394 Montpellier, France. IPME IRD, Cirad, Université de Montpellier.
 Patrick OLLITRAULT, Cirad, UMR Agap, F-20230 San Giuliano, France. Agap Univ Montpellier, Cirad, Inra, Montpellier SupAgro, Montpellier, France.



For many plants, including citrus (Figure 1), reticulate evolution results in admixed mosaics of large genomic fragments from the ancestral taxa. For those plants, we developed a pipeline called **TraceAncestor** in order to infer phylogenomic karyotypes from GBS or re-sequencing sequence data and diagnostic SNPs (DSNPs) of the ancestral taxa.

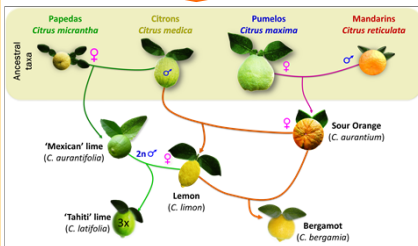


Figure 1: The *Citrus* genus, an example of reticulate evolution coupled with reproductive features, limiting further interspecific recombinations.

TraceAncestor is based in the relative read numbers of ancestral alleles in successive genomic windows (Figure 2). It allows the user to define the number of DSNPs per window, the sequencing error rate and the threshold for LOD values of the maximum likelihood test. **There is no limit to the number of ancestral taxa considered** (which is automatically defined by the reference matrix of DSNPs).

TraceAncestor allowed revealing for the first time the origin and phylogenomic karyotypes of several cultivated *Citrus* (Ahmed *et al.* 2019) such as diploid bergamot *Citrus bergamia*, and triploid 'Tahiti' lime, *C. latifolia* (Figure 3).

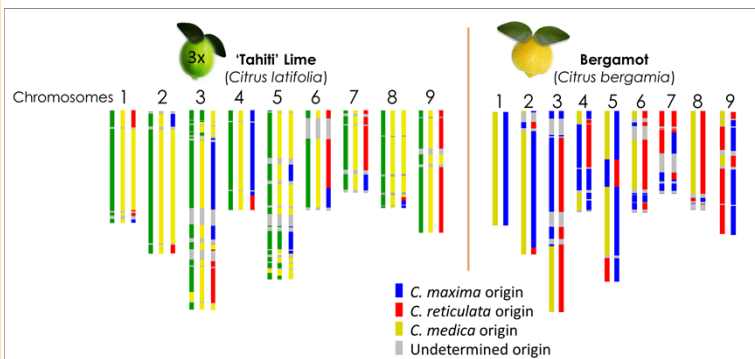


Figure 3: Triploid 'Tahiti' Lime and diploid Bergamot Karyotypes.

Reference : Ahmed D., Comte A., Curk F., Costantino G., Luro F., Dereeper A., Mournet P., Froelicher Y., Ollitault P., - 2019. Genotyping by sequencing can reveal the complex mosaic genomes in gene pools resulting from reticulate evolution: a case study in diploid and polyploid citrus. *Annals of Botany* : 21 p. <https://doi.org/10.1093/aob/mcz029>

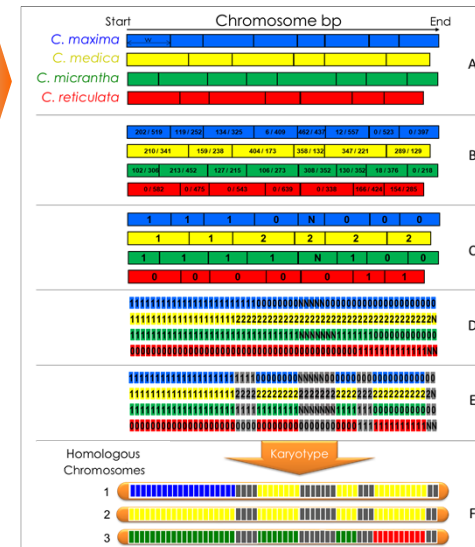


Figure 2: Example of local ancestor allele dose estimation for a triploid accession. (A) Definition of non-overlapping windows of ten DSNPs for each ancestral taxon; w, window of ten DSNPs. (B) Number of reads of the ancestor allele / number of reads of the other allele. (C) Estimation of allelic dosage of each ancestor per window of ten DSNPs. (D) Division of the chromosome into non-overlapping windows of 100 kb; (E) If the sum of allelic dosage of the four classes of DSNPs is different from the expected ploidy (here 3) -> indeterminacy (grey). (F) Unphased karyotype automatic drawings. Blue, *C. maxima*; yellow, *C. medica*; green, *C. micrantha*; red, *C. reticulata*; grey, indeterminacy.

TraceAncestor is efficient for inferring phylogenomic structures of diploid, triploid and tetraploid of any species.

It can also be used for the first generations of interspecific breeding schemes.



This pipeline is available as a Galaxy workflow at: <http://galaxy.southgreen.fr/galaxy/>



and for download at: https://github.com/SouthGreenPlatform/galaxy-wrappers/tree/master/Galaxy_SouthGreen/traceancestor



This work was supported by the FEDER Guadeloupe 'CAVALBIO' project and the Agropolis Fondation 'GenomeHarvest' project (ID 1504-006, 'Investissements d'avenir' programme, Labex Agro:ANR-10-LABX-0001-01)

