Specific capture of large DNA fragments by coupling CRISPR/Cas9 targeting and matrix-free separation and enrichment in a capillary electrophoresis

Nicolas MILON[1, 2], Céline CHANTRY-DARMON[3], Arnaud BELLEC[3], Audrey Boutonnet[2], Frédéric GINOT[2], Aurélien BANCAUD[1] and Hélène BERGES[3]

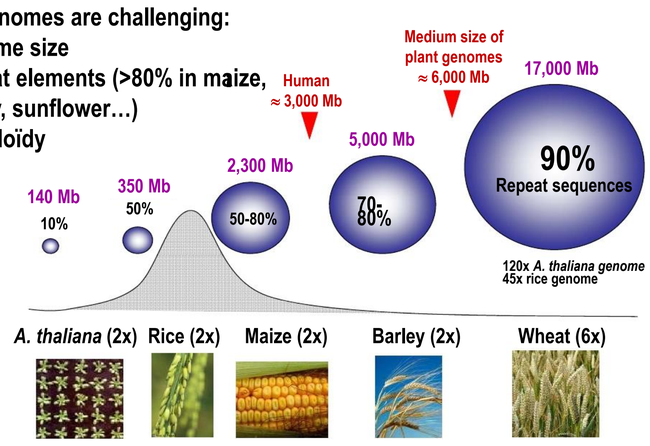
[1]CNRS, LAAS, 7 avenue du colonel Roche, F-31400, Toulouse, France.
[2]Picometrics, 478 Rue de la Découverte, 31670 Labège, France
[3]French Plant Genomic Resource Center, INRA - CNRGV, 24 Chemin de Borde Rouge – Auzeville, CS 52627, 31326 Castanet Tolosan Cedex, France helene.berges@toulouse.inra.fr

Despite the sequencing of long reads with Third Generation Sequencing Technologies, the study of numerous plant genomes remains challenging due to their complexity in terms of size, polyploidy or high percentage of repetitive elements. Moreover, the current state-of-the-art for large DNA separation and preparation is labor-intensive, if not resolution inappropriate for fragments over 20kb. New strategies for targeting large regions of interest are really needed.

We here report a new sequence capture approach for large DNA fragments. We used the first steps of the CATCH method (Cas9-Assisted Targeting of CHromosomal segments as described by Jiang and Zhu, 2016; and Jiang et al., 2015) based on the endonuclease function of the CRISPR/Cas9 system to cut a specific locus of 30kb from a *Medicago truncatula* BAC clone of 100kb.
Then, we adapted the µLAS technology (described in Ranchon et al., 2016) on capillary electrophoresis to perform on-line and automatically the concentration, separation and isolation of the target fragment from the bulk of the BAC DNA. This DNA processing technology does not require a separation matrix, but rather an electrohydrodynamic actuation that allows us sorting DNA molecules according to their sizes. We defined conditions to sort the specific 30kb fragments from 70kb and 100kb molecules. We describe the various steps used and present the collection yield obtained, the amount and purity of the target fraction recovered.

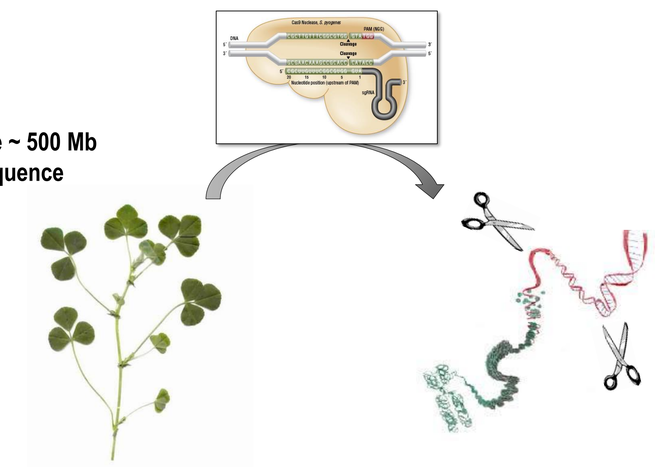
Complexity of the plant genomes

Plant genomes are challenging:
- Genome size
- Repeat elements (>80% in maize, barley, sunflower…)
- Polyploïdy

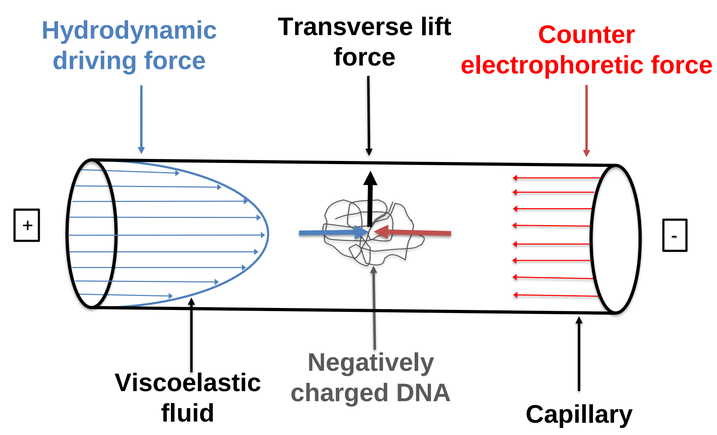


Principe : utilisation of the CRISPR/cas9 system to target sequence of interest in *M. truncatula*

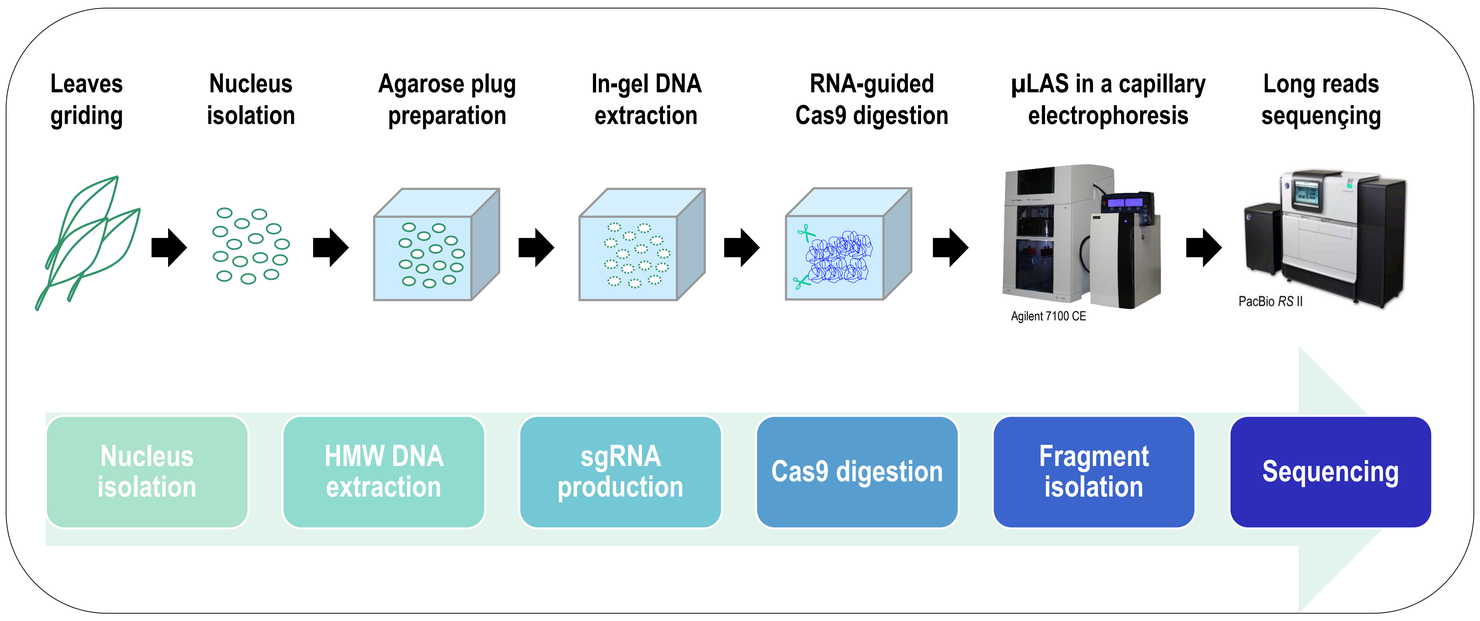
M. truncatula:
- Small genome ~ 500 Mb
- Reference sequence
- BAC libraries



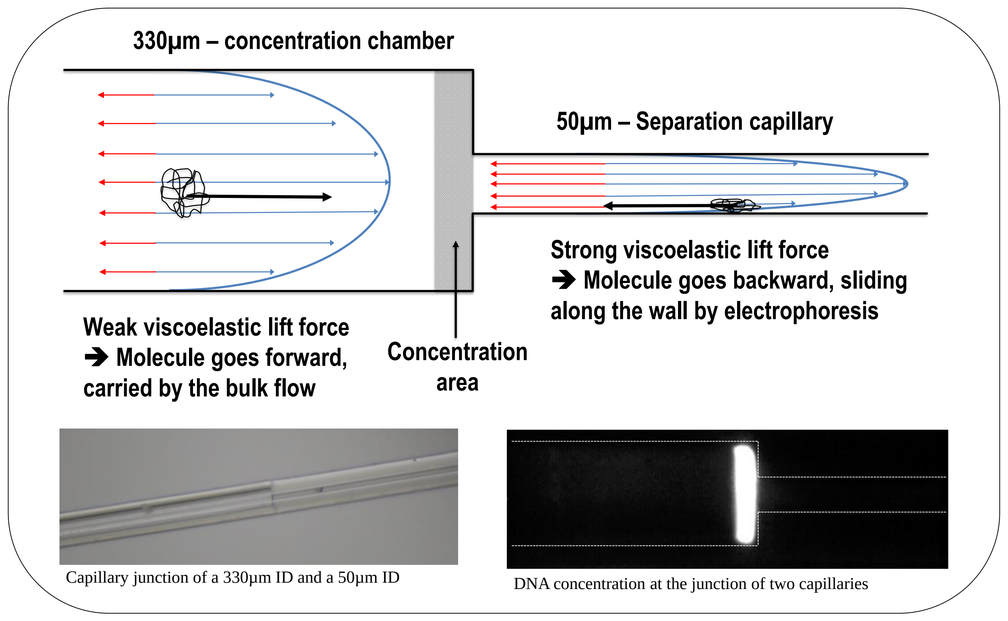
Matrix-free séparation of large DNA fragments



General Workflow



DNA concentration at the junction of 2 capillaries

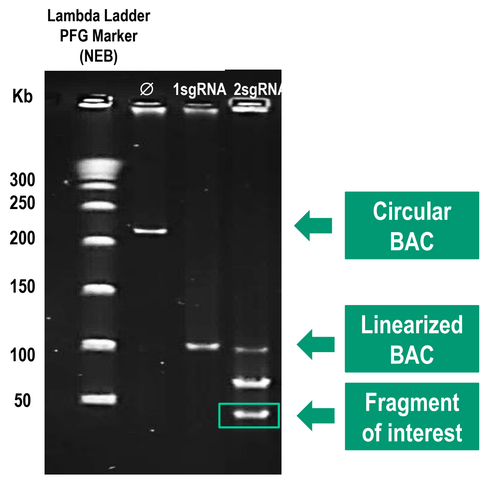


Capillary junction of a 330µm ID and a 50µm ID

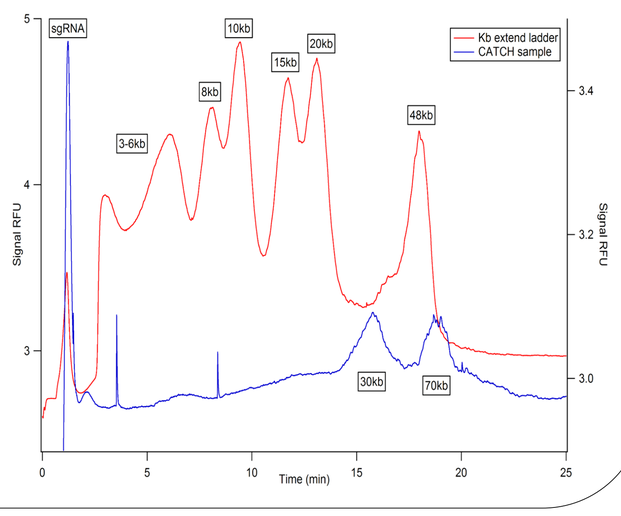
DNA concentration at the junction of two capillaries

Separation of the Cas9 digested sample

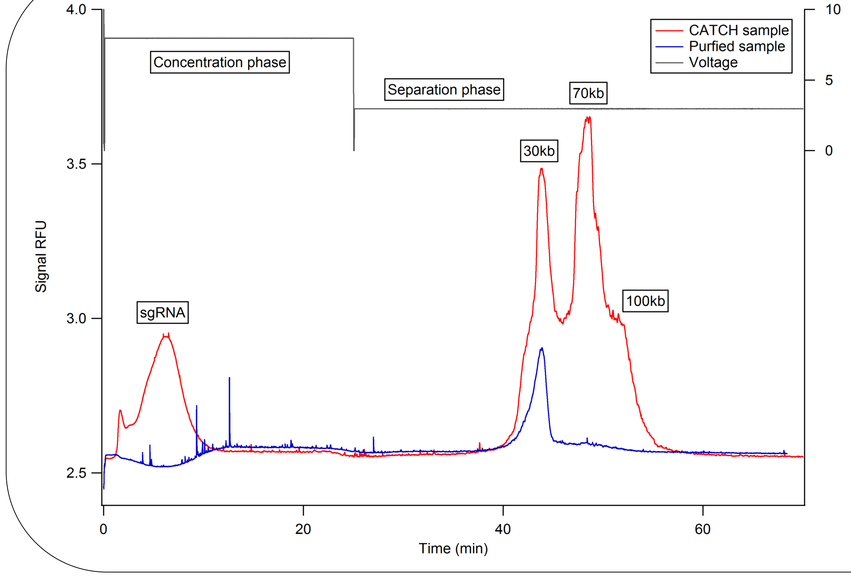
PFGE separation with 100 ng of sample



µLAS separation on CE with 75 pg of sample



Fragment isolation on capillary electrophoresis



- Selection and concentration of DNA fragments > 8kb
- Removal of sgRNA, 70kb and 100 kb fragments
- Isolation of 500pg of 30kb with 65% collection yield

In a first step we used the CATCH method to specifically target a 30kb locus of interest in a *M. truncatula* 100kb BAC clone. Then we developed a new technology for large DNA fragment analysis. We were able to isolate the desired 30kb fragments from the sgRNA, the 70kb and the 100kb fragments. The collection yield is ≈65% and the total amount of 30kb fragment recovered is 500pg. We confirmed that the fragment collected was the fragment of interest using PacBio sequencing .
Altogether we demonstrated that this new technology allows us to capture and sequence large DNA regions of interest. We now aim at using this technology on more complex gDNA.

@CNRGV
@LaasCNRS