



**HAL**  
open science

## Discovering domestication mutations in cauliflower

Mathis Pochon, Alexis Mergez, Fabrice Legeai, Jeremy Lucas, Cyril Falentin,  
Mathieu Rousseau-Gueutin, Anne-Marie Chèvre, Romain Blanc-Mathieu,  
François Parcy

► **To cite this version:**

Mathis Pochon, Alexis Mergez, Fabrice Legeai, Jeremy Lucas, Cyril Falentin, et al.. Discovering domestication mutations in cauliflower. Journées scientifiques 2025 Agroécologie et Numérique, Jan 2025, Dijon, France. . cea-04932621

**HAL Id: cea-04932621**

**<https://cea.hal.science/cea-04932621v1>**

Submitted on 6 Feb 2025

**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.

# Discovering domestication mutations in cauliflower

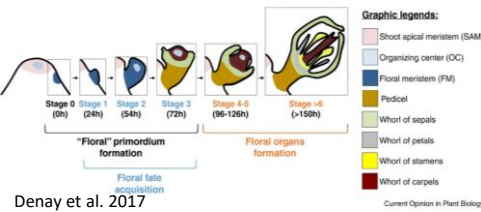
Mathis POCHON<sup>1</sup>, Alexis MERGEZ<sup>2</sup>, Fabrice LEGEAÏ<sup>2</sup>, Jérémy LUCAS<sup>1</sup>, Cyril FALENTIN<sup>2</sup>, Mathieu ROUSSEAU-GUEUTIN<sup>2</sup>, Anne-Marie CHEVRE<sup>2</sup>, François PARCY<sup>2</sup>, Romain BLANC-MATHIEU<sup>1</sup>

1.Laboratoire Physiologie Cellulaire et Végétale, Université Grenoble Alpes, CNRS, CEA, INRAE, IRIG-DBSCI-LPCV, 17 Avenue des Martyrs, F-38054, Grenoble, France | 2.UMR IGEPP, INRA, Agrocampus Ouest, Université de Rennes 1, BP 35327, 35653, Le Rheu Cedex, France

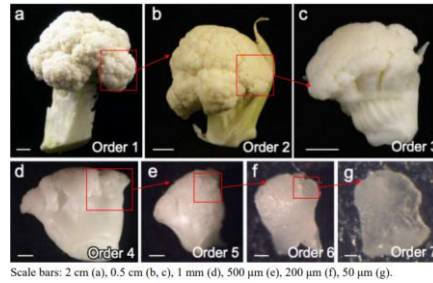
## Abstract

During the development of agriculture over the centuries, humans have selected plants with advantageous agronomic traits. Identifying mutations responsible for these characteristics would enhance our knowledge in agronomy and facilitate the neo-domestication of plants or the improvement of existing varieties. This project aims to identify mutations underlying the domestication of plants in cabbages (*Brassicaceae oleraceae* systems). It will first focus on mutations responsible for deregulation of the genetic network responsible for the aborted flower in cauliflower. A population level pangenome was built to identify genetic variation in 373 *Brassica oleracea* accessions. Genotype-to-morphotypes association analysis and known floral development genes are used to extract candidates variants. These variants, in cis or trans regulatory elements, are further analyzed *in silico* to identify those that most likely disrupt the interaction between TFs and DNA to alter expression of floral developmental genes. Candidates causatives mutations will be validated experimentally. The approach will be apply to other traits in cabbages.

## Cauliflower are successive aborted flowers

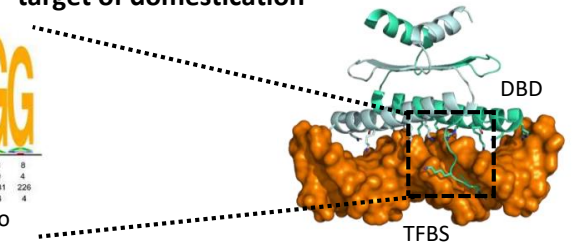
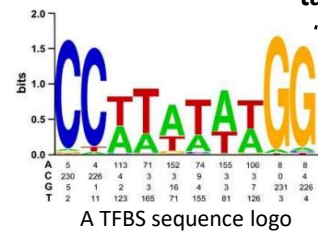


- Flowering: Development of the inflorescence meristem
- Flower development: formation of floral organs.
- Lateral meristems emerge in a spirally arranged way
- Each floral meristem leads to flower formation



- Mutations alter flower development by recursively triggering lateral meristem development
- Heads of cauliflower results from a serie of spirally arranged floral development arrests

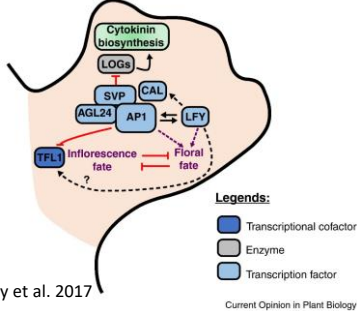
## Transcription factor (TF) and their DNA binding sites (TFBS): target of domestication



- Transcription factors (TFs) are a group of proteins capable of binding short and specific DNA fragments thanks to a DNA binding domain (DBD) to regulate gene expression. Mutations in TF and TFBS have often been associated with domestication traits (Swinnen et al. 2016)
- TFBS can be predicted using position weight matrices (PWM) and represented as sequence logo.

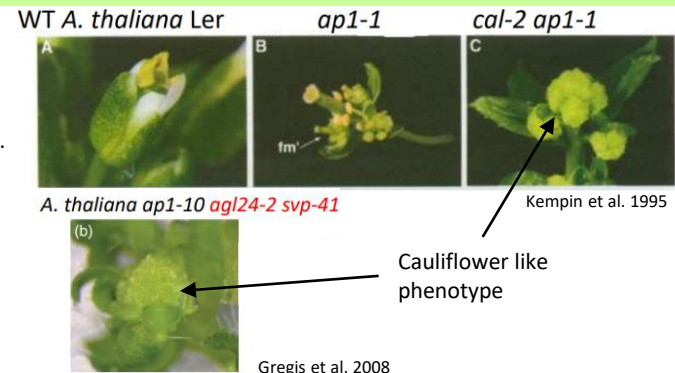
The DNA/TF interaction site is essential for gene regulation, making it a good candidate for identifying causal mutations

## Gene regulatory network (GRN) of floral development

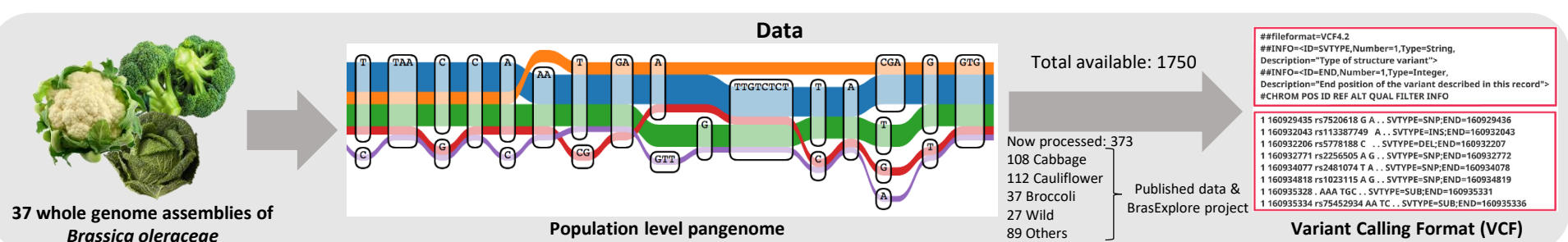


## Causal genes

- In *Arabidopsis thaliana*, a loss of function of transcription factors AP1 and CAL or AGL24 and SVP leads to a cauliflower-like phenotype.
- The CAL ortholog in cauliflower (BoCAL) has a premature stop codon its coding region. Our preliminary analysis indicates this causative mutation is absent in 5% of cauliflower.
- However no mutation that would disrupt BoAP1 function have been identified so far.
- We hypothesize that cauliflowers result from multiple mutations in cis or trans element of the flower development GRN.



## Approach



## Functional annotation of non-coding variants

1. Extract 120pb surrounding a variant
2. Using >10 species phylogenetically close to *Brassica oleraceae* from Plaza database

3. Identify orthologous sequences with an alignment

4. Identify TFBS by scanning PWM and estimating its phylogenetic conservation

5. Find conserved modules (group of TFBS which appear together in a region) and calculate metrics about conservation

6. Variants within conserved TFBS modules are candidates functional cis-element

## Variant classification

Candidates cis-regulatory variant  
+  
Candidates in trans regulatory region (eg. mutation disrupting DNA binding domain)

In vitro validation of candidate variant affecting TF-DNA interaction (EMSA+reporter assay)

perspectives: applying to other *Brassica oleraceae* morphotypes + in planta validation

## Association study

### Genotype-phenotype matrix

variants	1	2	3	4	...	n	morphotype
lines							
1							WT
2							Broccoli
.							Cauliflower
.							Green cabbages
.							etc
n>1000							

- Search if there are regions or genes in both coding and non-coding region can be associated to a morphotype
- Single loci association test (PLINK)
- Multiple loci association tests
  - LAMPLINK (Terada et al 2016)
  - Machine learning tools to detect variant and combination of variant associated with morphotypes (Decision tree, Out of bag random forest)
  - Variants can be weighed according to their predicted effect (ex: disruption of TF-DNA binding,)

⇒ Candidats variant or candidate regions associated with morphotypes