Single-cell TARGET
Study of the combinatorial effect of Transcription Factors

Amélie Emanuel[1], Jean-Baptiste Carluer[1,2], Angelo Pasquino[3], Samantha Frangos[3], Sandrine Ruffel[1], Wojciech Szponarski[1], Bastiaan Bargmann[4], Gloria Coruzzi[3], Etienne Delannoy[5], Gabriel Krouk[1]

1. Institute for Plant Sciences of Montpellier (IPSIM)
2. Institut Montpelliérain Alexander Grothendieck (IMAG)
3. Center for Genomics and Systems Biology, New York University
4. Virginia Polytechnic Institute and State University, Blacksburg
5. Institute of Plant Sciences Paris-Saclay (IPS2)

Context & Method

Gene expression is largely modulated by transcription factor (TF) activities. Nearly 2000 genes encode TFs in *Arabidopsis thaliana* genome (Riechmann et al. 2000). However, for many TFs, their direct targets remain unknown, which affects the establishment of transcriptional networks orchestrating genome expression. Furthermore it is now known that TF don't act in isolation and that combinatorial logic is at game (Brooks et al. 2019, Rieu et al. 2022).

The TARGET (Transient Assay Reporting Genome-wide Effects of Transcription factors) system is a rapid technique that aims to determine the direct targets of a transcription factor of interest (Bargmann et al. 2013). Here, we propose a digital version of this system, the single-cell-TARGET, and show that it also enables the study of the combinatorial logic of TF activity.

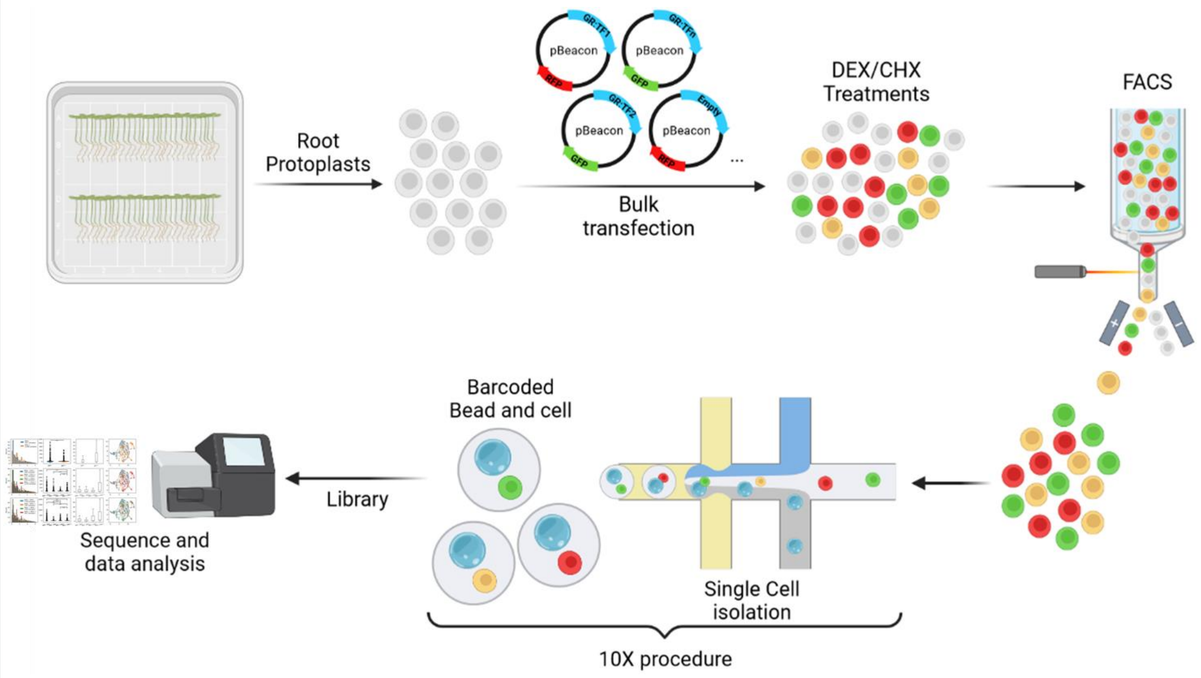

Figure 1 : Single-cell TARGET method

Arabidopsis root protoplasts are bulked transformed with 23 plasmids harboring different TF-GR fusions then treated with dexamethasone (triggers GR–TF fusion entrance in the nucleus) and cycloheximide (translation inhibitor that prevents activation of indirect targets). After FACS selection (based on RFP/GFP signal provided by an independent cassette in the plasmid), transformed protoplasts are subjected to single-cell RNA-sequencing (10X genomics). Data is then analyzed to detect regulatory interactions and combinatorial logic between TF in the control of target gene expressions.

Protoplasts co-transformation

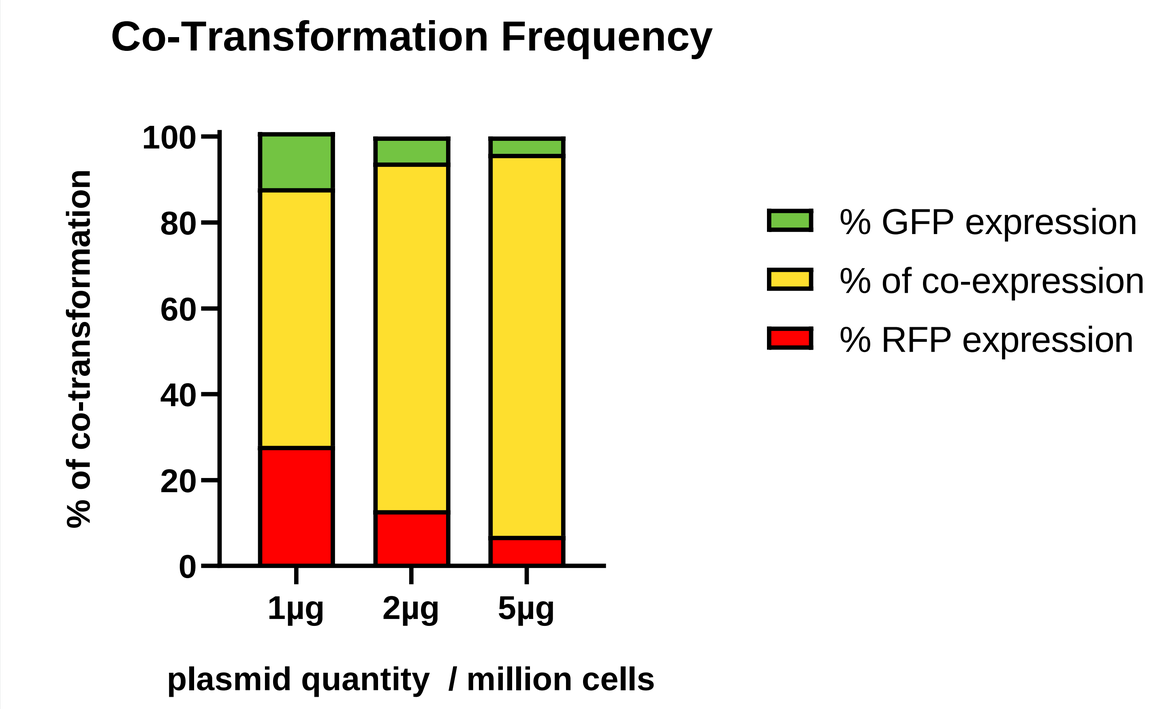

Figure 2 : Protoplasts co-transfection frequency with GFP and RFP plasmids for different DNA concentration (1, 2 or 5 µg each per 1 million cells).

High plasmid concentration (5 µg) leads to a non random transformation, an important proportion of protoplasts are transformed with 2 plasmids. With a lower plasmid concentration (1 µg), co-transfection decreases allowing potential random transformation of protoplasts. This is essential for the following procedure.

TFs expression in single-cell RNA-sequencing

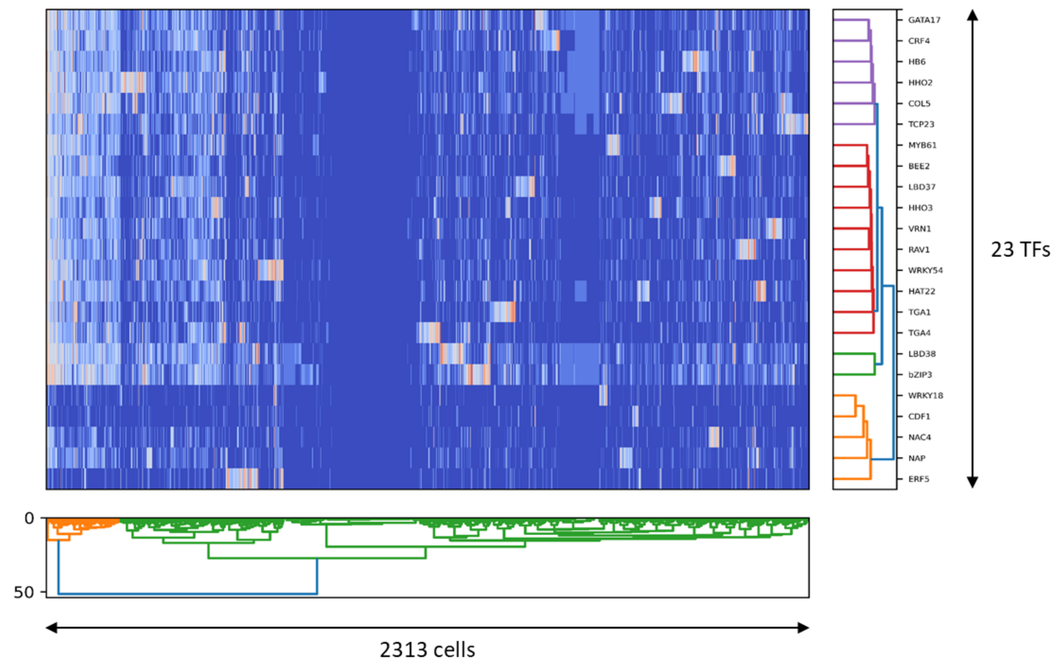

Figure 3 : Expression of the 23 transfected TFs in the 2313 cells retrieve after scRNA-seq.

This cluster represents the heatmap reporting read counts for the transfected TFs-GR fusion. Reads counts are related to mapping on GR-TF tags allowing to be sure that the recorded signal reports the transfection event and not endogenous TF presence. It is worth noting that empty TARGET vectors are also transfected to serve as control. The blue cluster in the middle is likely the report of such transfection events (empty vectors) allowing a good amount of negative controls.

Deconvolution of TF expression combinatory

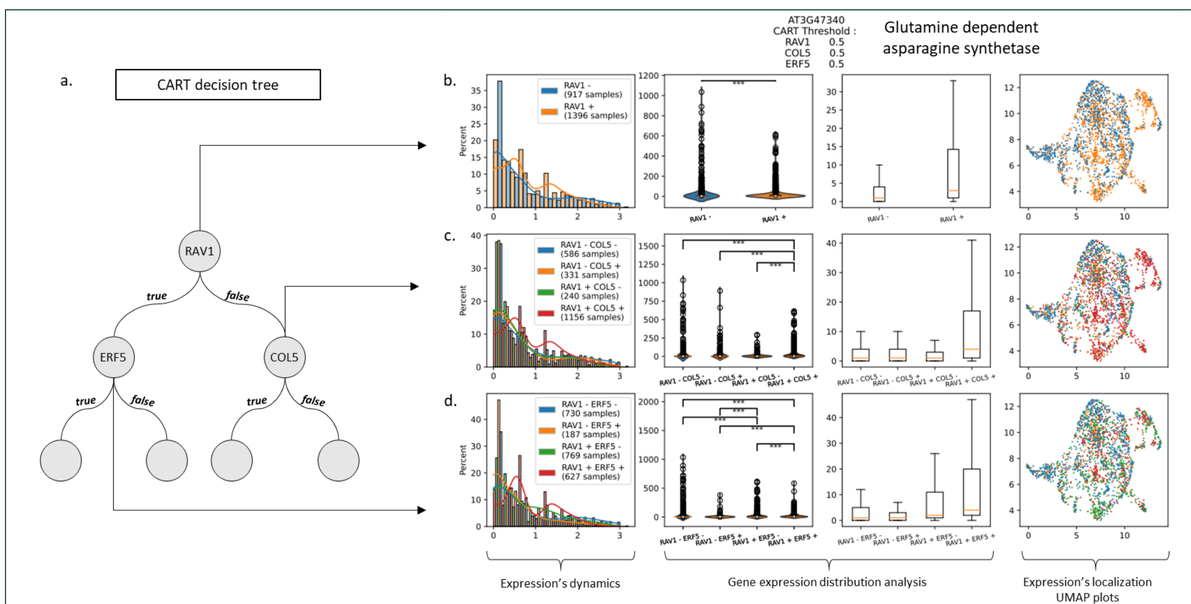

Figure 4 : CART decision tree deconvolutes ASN1 gene expression across the 2313 Cells.

(a) Schematic representation of CART decision tree. (b) Expression of ASN1 according to TF1 (RAV1) absence or presence based on CART threshold (0.5). Expression of ASN1 according to TF1 (RAV1) and (c) TF2 (COL5) or (d) TF3 (ERF5) interaction based on CART threshold (0.5 and 0.5 respectively).

Statistics (Mann-Whitney and ANOVA) are applied to measure the effect of each TF and TF combination of the TARGET (here ASN1) expression. This analysis is repeated throughout the whole genome for each gene. Statistical thresholds are used to build the following (Figure 5) network.

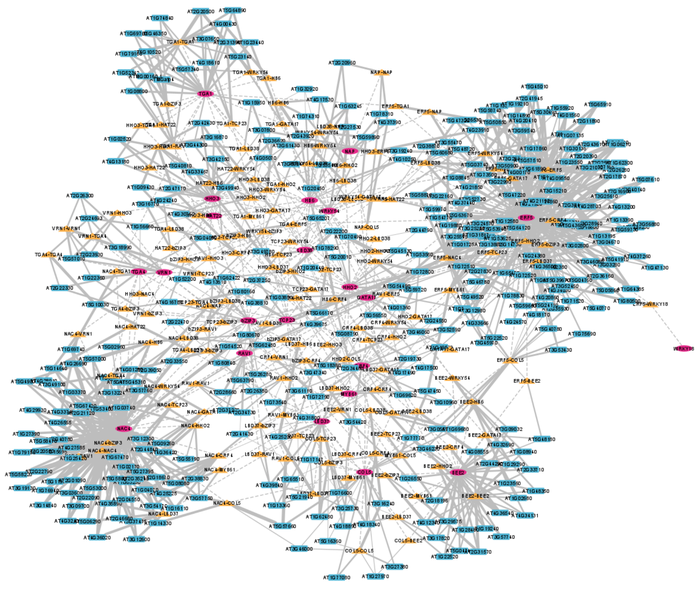

Figure 5 : Gene regulatory cooperative network of 23 TFs.

Transfected TF are drawn as pink circles. Blue rectangles are target genes. Yellow triangles depict TF-TF interactions. The width of the edges is related to the gini score for CART models. ERF5 and NAC4 seem to be master regulators in this network 154 and 93 interactions respectively. ERF5 and CRF4 cooperatively control 49 genes being the most proheminent interactors from this analysis.

Conclusion & Perspectives

This new method derivated from the classic TARGET seems to be a good way to predict the interaction between TF and their combinatorial logic to control target genes.

- We used 23 nitrogen-regulated TFs ➡ try with more TFs.
- ➡ validate the interactions found by looking at the expression of the target gene by qPCR or RNA-seq methods.

Riechmann, J.L., *et al.* (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105-2110.
Brooks MD, *et al.* (2019). Network Walking charts transcriptional dynamics of nitrogen signaling by integrating validated and predicted genome-wide interactions. Nature communications 10: 1569.
Rieu, P., *et al.* (2023). The F-box protein UFO controls flower development by redirecting the master transcription factor LEAFY to new *cis*-elements. Nat. Plants 9, 315–329.
Bargmann, B.O., *et al.* (2013). TARGET: a transient transformation system for genome-wide transcription factor target discovery. Mol Plant 6:978-980.