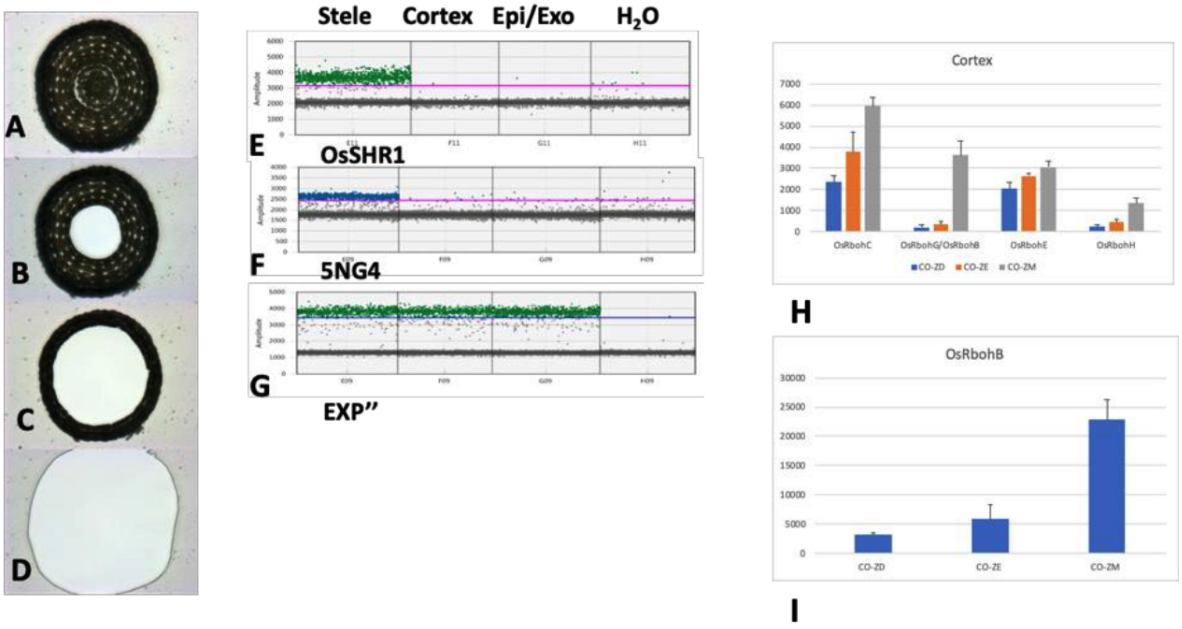
Position of the project as it relates to the state of the art

The mechanisms of formation of aerenchyma are not known, in particular genes involved in its initiation. Their identification would make it possible to understand how this adaptive mechanism to submergence, present in many flowering plant species, is implemented. This would also open up the possibility of developing new submergence-tolerant cereal species. Rice is the perfect model to identify these genes and mechanisms that are naturally present in this species. Functional analysis, by genome editing, is routinely performed; there is a wide genetic diversity and rice is the model species of cereals with a perspective of transferring submergence tolerance to other cereals. The formation of aerenchyma results probably from the cooperation between many regulators (TF, kinases, peptides) and effectors (enzymes, structural proteins) that are specifically expressed in the cortex. This complexity makes it difficult to identify by a simple transcriptomic approach the key regulator(s) in the formation of these aerenchyma as it is the gene network as a whole that is responsible for cortex transition to aerenchyma and not just a single gene. At the reconstruction level, both the context-specific and the combinatorial nature of regulatory elements pose a great challenge for reconstructing Gene Regulatory Networks (GRNs). These networks are inferred from large-scale data samples (transcriptome of coding and non-coding RNAs, ChIP-seq data and proteomic data if available) and evidence of physico-chemical interactions (protein-protein interactions, transcription factor binding sites). The bioinformatics approach proposed in this project is original mainly in terms of the incorporation into the model of the factor of cooperativity between co-regulators (Elati et al., 2007; Nicolle et al., 2015), rendering it closer to the biological reality than most other approaches. Inference techniques cannot yet identify all the components of a network accurately, but they have already identified master regulators for root systems (Hill et al., 2016).

One key breakthrough in the exploration of networks was to consider the influence of the regulators by evaluating the expression of target genes rather than the expression of regulators themselves with the aim of detecting master regulators (Nicolle et al., 2015). Partner 2 has already developed a number of user-friendly and open-source computational network tools e.g., PEPPER - (Winterhalter et al., 2014) and downloaded more than 10,000 times, CoRegNet - (Nicolle et al., 2015) and downloaded more than 7,000 times, GREAT published in 2016 (Bouyioukos et al., 2016). For instance, partner 2 has recently achieved Regulatory network inference of lipid accumulation with Yarrowia lipolytica yeasts using those GRN tools (Trebulle et al., 2017). The most influential TFs on lipid accumulation were identified using tools such as CoRegnet and LICORN and their function has been validated for several of them (Trebulle et al., 2017). This approach has also been used on S. cerevisiae and in human cells with success demonstrating the genericity and validity of the approach. Moreover, CoRegnet was also used to identify 12 microglia specific transcriptional regulators of the human core genes (Galatro et al. 2017) and 10 major regulators driving the transition in nonalcoholic liver disease (Lou et al. 2017) demonstrating its efficiency to extract GRN in complex organisms.

In order to verify whether this approach was feasible, we reconstructed (partner 2 and 1) a network of regulators from RiceXpro data (Sato et al., 2013; Takehisa et al., 2012). We looked for regulators with maximum influence for external tissues, cortex and stele (Sato et al., 2013; Takehisa et al., 2012) for a root mature region from RiceXpro data as a proof of concept. Strikingly, two regulators, a histone deacetylase and an HMGT transcription factor clearly emerged from these analyses as the main regulators for the differentiation of external tissues and cortex. Strikingly, a recent paper published in A. thaliana has just shown tha HDA19, a histone deactylase homolog, as a major role to specify root cortex identity and indirectly for the differentiation of the epidermis (Chen et al., 2019). This is an indirect and unexpected validation and a demonstration of the powerfulness and effectiveness of this approach. The approach of reconstructing and navigating across GRNs for differentiation of aerenchyma is an attractive approach to identify gene network for flooding tolerance but requires several prerequisites to be done in rice.

The formation of aerenchyma can be subdivided into 4 phases in rice (Kawai et al., 1998). A prerequisite is therefore to obtain transcriptomic data in specific tissues but also at specific stages by separating external tissues and cortex. Obtaining RNAseq profiles from specific tissues is technically delicate but not impossible. LCM (Laser Capture Microdissection) consists of using a laser coupled to a microscope to cut fixed tissues that are then transferred into an RNA extraction buffer (Chavan et al., 2018). From there, it is possible to perform expression by RT-PCR or to build RNAseq library. In rice, for example, the technique was used to establish that certain ethylene biosynthesis or ROS production genes were specifically expressed in the cortex (Takehisa et al., 2012). In 2012, by coupling microarray and LCM, expression profiles of the cortex and internal tissues of the mature zone of rice coronary roots were obtained (Takehisa et al., 2012). But to our knowledge, no one has tried to combine tissue-specificity and development stage at the root level at the same time. There are several reasons for this, the main one being the technical difficulty of obtaining RNAs in sufficient amount and quality. Recently, Partner 1 have developed and optimized a protocol that allows us to obtain high quality RNA from LCM sections in sufficient quantity to consider conducting a time course experiment for the differentiation of the cortex and outer tissues with a reduced amount of root tissues (Mounier T. et al 2020b). We also demonstrated specificity of the root tissue laser microdissected using tissue specific markers by ddRTPCR **(Figure 1 A-G)**.



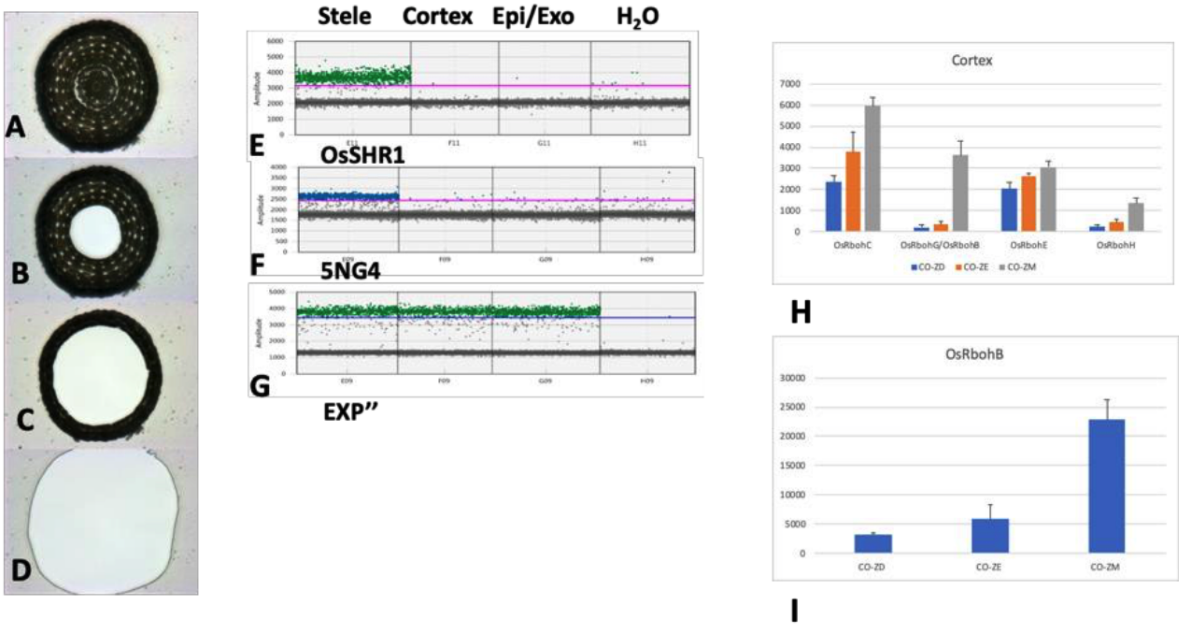


Fig. 1: Microdissection of root tissues. A-D Cutting of the root tissues by microdissection, the tissues are extracted successively from the inside to the outside. A) Paraffin section before cutting. B) After cutting the stele + endodermis. C) After cutting the cortex. D) After cutting the external tissues, epidermis/exodermis and sclerenchyma. E-F ddRTPCR of specific markers. E) Expression of OsSHR1 in the stele, cortex and external tissues. F) Expression of 5NG4 in the stele, cortex and external tissues. G) Expression of a constitutive gene expressed in all three tissues, Exp'. The positive droplets, materializing the presence of transcript are in green, the negative droplets in grey. The red bar represents the detection threshold. These results demonstrate that there is no cross-contamination between the microdissected tissues. H) and I) Expression profiles of RboH family genes during cortex maturation. Y-axis, number of RNAseq readings per bank obtained after normalization, in the division zone (CO-ZD), elongation zone (CO-ZE) and mature zone (CO-ZM). The standard error bar corresponds to four biological replicates.

Using this technological development, we constructed and sequenced 36 RNAseq libraries for 3 tissues, the first 3 developmental steps and four biological repetitions to identify the expression profiles of the different stages of differentiation of aerenchyma. To verify that our RNAseq data capture the main steps of cortex differentiation into aerenchyma, we analyzed the expression of genes of the RboH family **(Figure 1H, I)**. Genes of this family are involved in cell death in rice and their expression increases during submergence (Yamauchi T. et al., 2017). Five genes of this family have a significant increase in expression during cortex maturation **(Figure 1H, I)**, confirming that our data allows us to monitor cortex differentiation into aerenchyma. The identification and classification of possible influential regulators of aerenchyma need to aggregate information from other sources and databases including but not limited to orthologs, gene family classification to explore GRN in a practical way.

The next step after the reconstruction of the GRN is to identify the most relevant regulators in a set comprising several hundreds of co-regulated genes. Network visualization has shifted its focus from simple network models, to more complex ones with many unresolved challenges (von Landesberger et al., 2011), such as dynamic networks (Beck et al., 2017), compound networks, multivariate networks (Kerren et al., 2014) and multilayer networks (Kivelä et al., 2014; McGee et al., 2019). Navigating and comparing reconstructed networks is very difficult, especially when more than two networks are involved. Recent visualization techniques aim to ease such comparisons. For instance, hive plots are used to extract qualitative and quantitative information from these networks and compare them with each other (Krzywinski et al., 2012); their usefulness has been demonstrated with

a use case about disease gene networks, among others (Krzywinski et al., 2012). The hive plot is a network representation which places the nodes belonging to distinct subnetwork on distinct axes, while the connections are depicted as curves between these nodes. Several axes can be represented in a star pattern **(Figure 2)**.

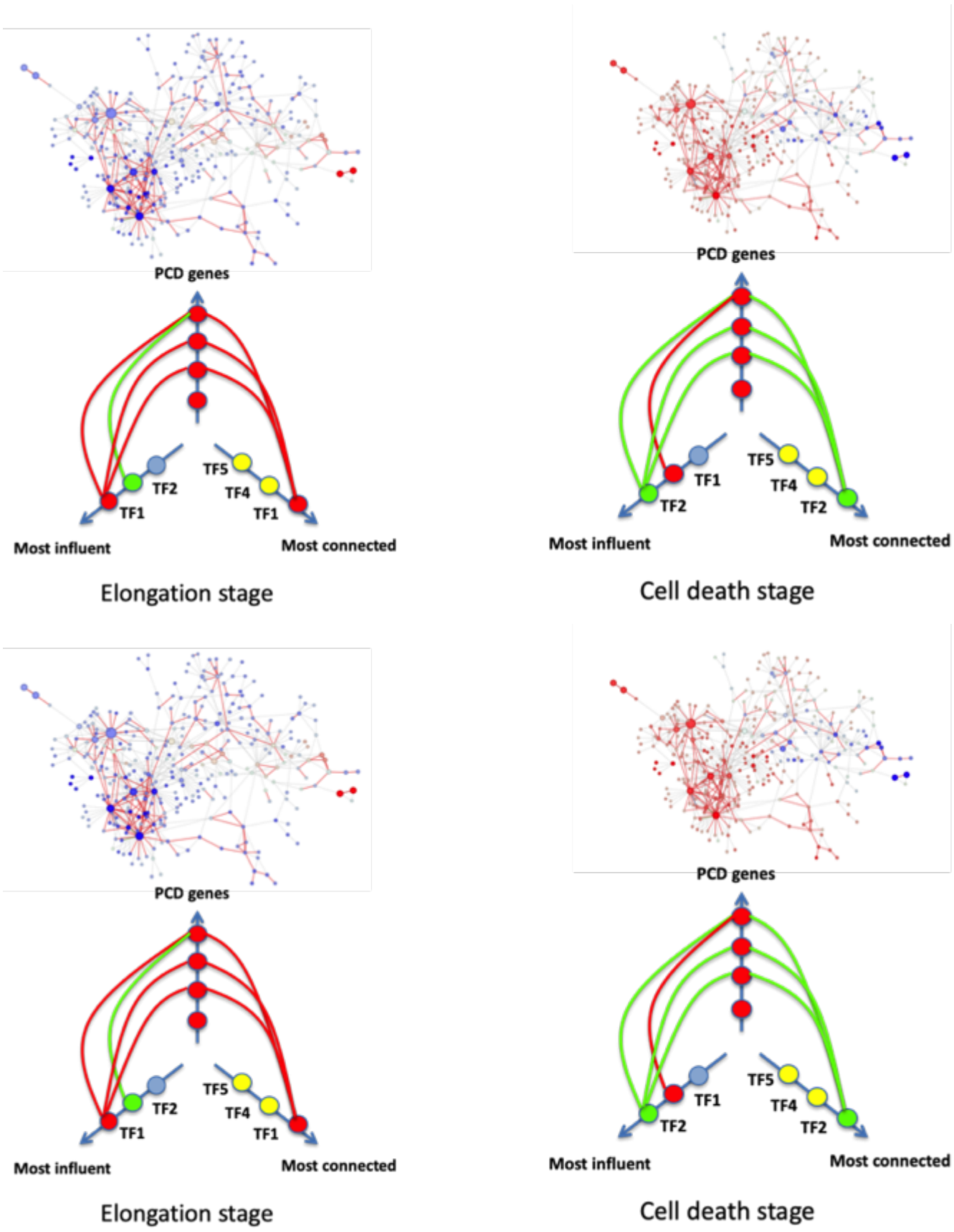


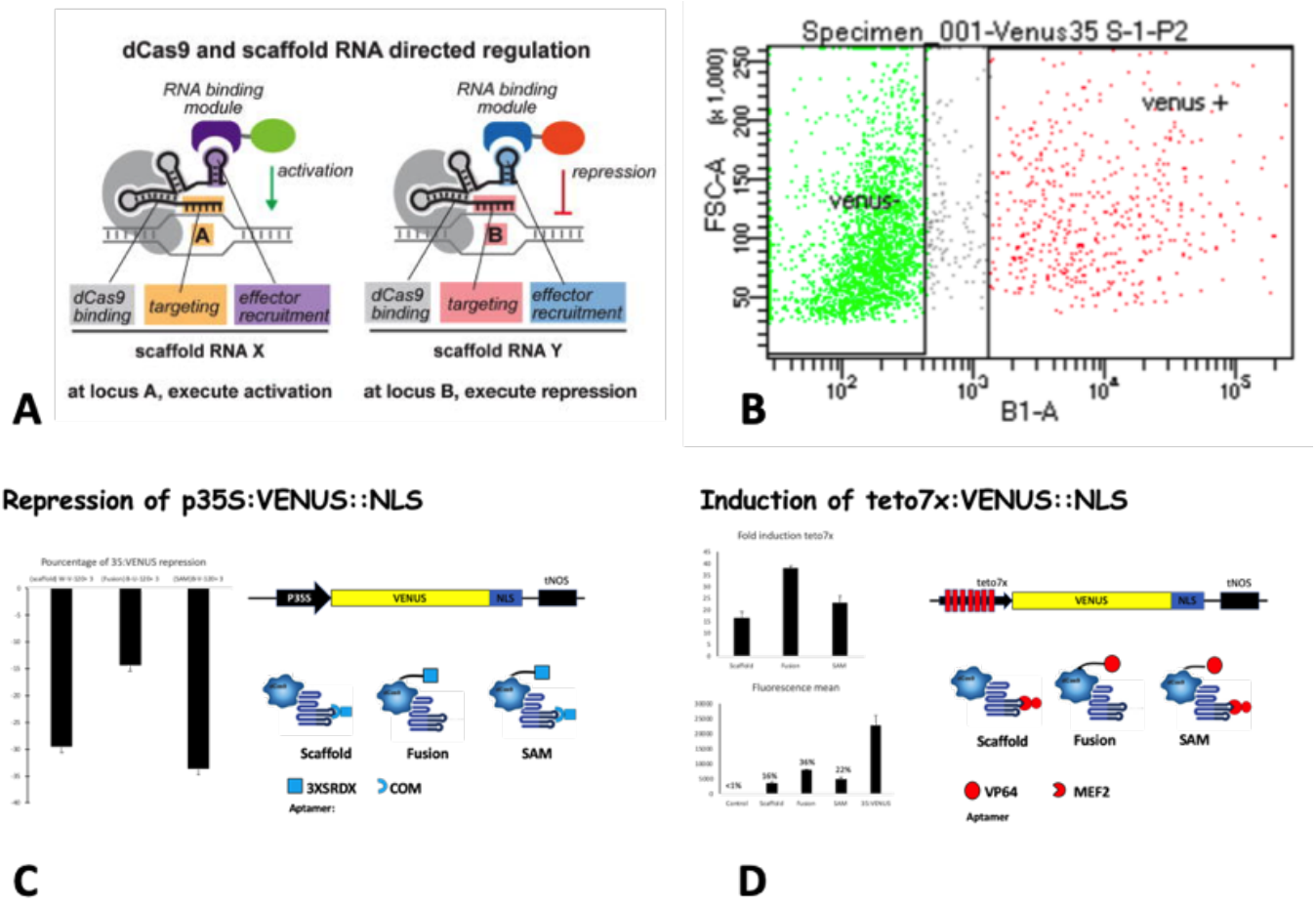
Fig. 2: The hive plots: visualize and compare gene networks. In this hypothetical example, two

networks of transcription factors and their effectors are compared for two stages of cortex differentiation into aerenchyma, A) the elongation and B) initiation of cell death stages. Top are hypothetical, but classical representation of gene networks. Bottom are hive plots generated using gene network data. The 3 axes correspond to: a list of 4 effector genes involved in cell death (PCD genes); the most influential transcriptional factors on these effectors (most influential) on a second axis; the transcription factors most connected to these effectors on the last axis (most connected). The last two axes are ranked. With the help of the hives plots, we can directly compare the two networks and immediately visualize there is a shift between the TF1 (red color) and TF2 (green color) effect during the initiation stage of cell death. This suggest that TF1 is the earliest master regulator of PCD, initiating the cell death process while TF2 become dominant during the cell death initiation step itself. The axes can be arbitrary determined by a user to explore and compare gene networks in almost any way, while direct comparison of entire gene networks seems difficult and most of the time impossible. Curves represent network edges

For example, in the case of searching for the most important regulators from a list of cell death effectors, one axis will represent these effectors, another axis will carry the regulators sorted according to a quantitative criterion e.g. in descending order “from the regulators which are connected the most to effectors to the ones that are connected the least” and a third axis will represent the regulators sorted by another quantitative criterion in descending order “from the most to the least influential”. The interactions between regulators and effectors will be represented as curves connecting them. This network representation allows to identify at a glance the most influential regulators, the ones connected the most to these effectors synthesizing information from several thousand or even ten thousand pieces of information **(Figure 2)**. The ability to order nodes on their axes based on a biological criterion facilitates the exploration of networks and promotes the repeatability of network navigation as opposed to other popular representations. This type of representation also makes it possible to compare networks with each other. By generating this hive plot for several stages of cortex differentiation we can determine if the same regulators are active at these different stages or if there is shift **(Figure 2)**. Finally, another type of application can be to identify new effectors or regulators of unknown function. This time, one can start by mapping regulators known for their function in cell death on one axis and map on another axis the most co-regulated “unknown” effectors on another axis (using all effectors annotated as ‘predicted’ or ‘unknown function’). As the layout of hive plots is much cheaper to compute than popular force-directed layouts, they can be generated on the fly to fulfil various types of analyses. We will rethink in the project, the meaning of Hive plot (HP) axes for the GRN analysis tasks with links with a biological meaning such as the up/down-regulation links between TFs and genes. We propose also to visualize and simulate GRN modification using HP in frame of the GREENER project to understand the effect of local perturbations on the interaction network of genes and TFs.

In practice, the regulators that will be identified act in cooperation to induce the expression of many effectors (enzymes…) necessary for the formation of differentiation of aerenchyma. Validation of their role in combination is difficult and requires the development of a system capable of both inducing and repressing the expression of several genes at a time and then identifying co-regulated genes. Protoplasts -i.e. isolated plant cells without their cell walls - form a powerful model system for deciphering GRN in plants (Diaz-Trivino et al., 2017). A technique, called TARGET was developed in 2013 with the aim to identity TF targets using TF fused to glucocorticoid receptor (GR) (Bargmann et al., 2013). Following dexamethasone treatment of transformed protoplasts, the GR-tagged TF enters into the nucleus where it activates its targets (Bargmann et al., 2013). Despite its interest and successes, this approach has several limitations, the most important one being the impossibility to activate and repress simultaneously several TF. The emergence of the CRISPR/Cas9 technology offer a powerful alternative for the manipulation of the expression of several genes simultaneously. The first generation was developed by merging an activation domain with a CAS9 protein whose double-

stranded nuclease sites have been mutated and inactivated (deadCas9 or dCas9) (Cheng et al., 2013; Gilbert et al., 2013; Lowder et al., 2015), but this system could not be used to activate and repress simultaneously two different set of genes. Recently, a powerful technology using CRISPR RNA Scaffold (Zalatan et al., 2015) was described **(Figure 3A)**. This system is based on the fusion of the sgRNA with an aptamer. The chimeric sgRNA guided was called scRNA or RNA scaffold and each aptamer can be recognized by a different viral protein. Thus, by using different scaffold RNAs, recognizing therefore different binding-effector modules - some fused to activators, some fused to repressors -, it should be possible to activate and repress simultaneously several transcription factors within the same cell **(Figure 3.A)** (Zalatan et al., 2015).



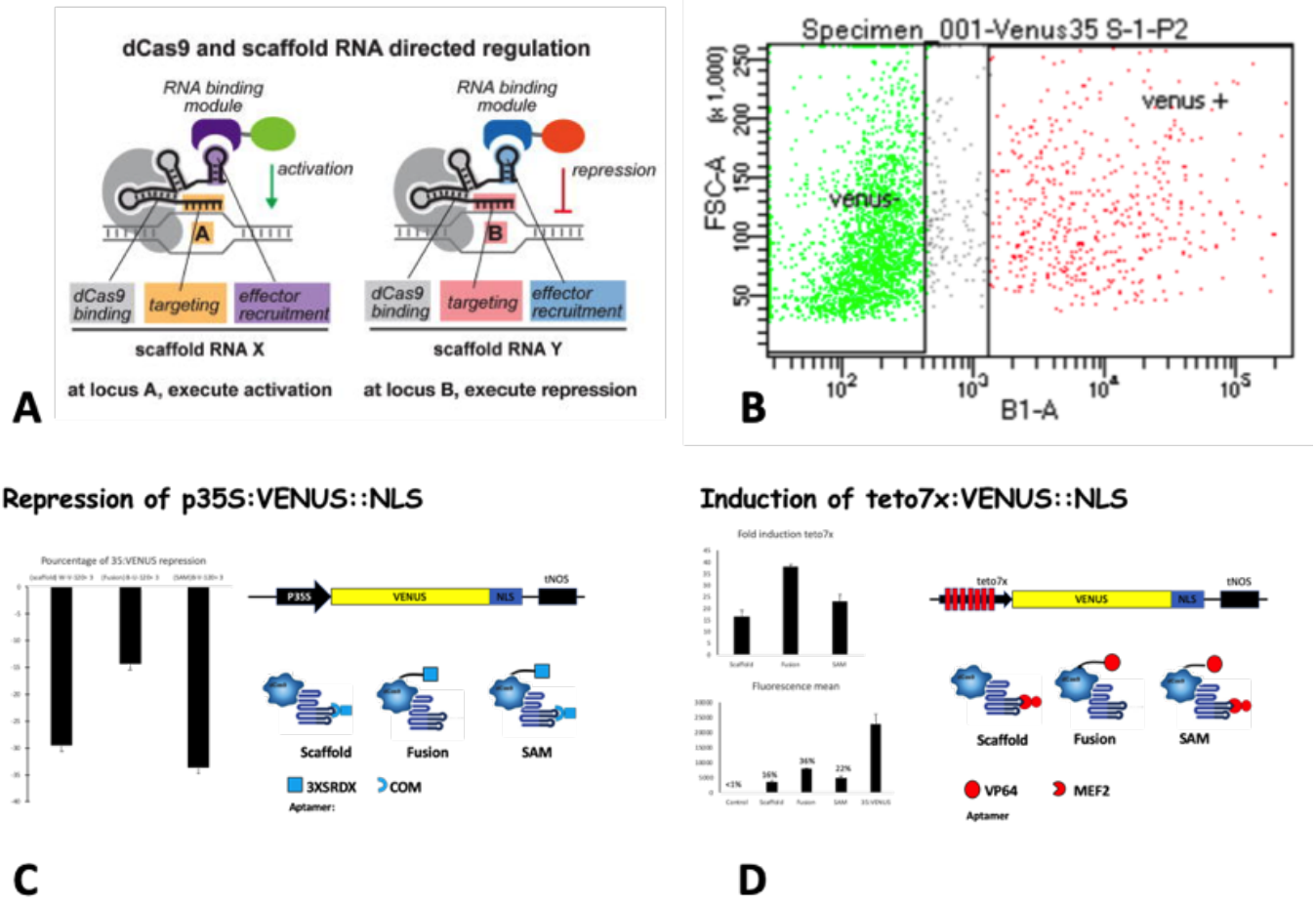


Fig. 3: Activation and repression of a reporter gene, VENUS YFP, by the scaffold system. A) CRISPR RNA scaffold system: principe (from Zalatan 2015). CRISPR RNA scaffold-based allows simultaneous activation or repression of independent gene targets. sgRNA molecules are extended with aptamer recognize by RNA-binding proteins that are fused to functional effectors, either activator or repressor. This approach allows simultaneous activation and repression of several genes in the same time on the same cell. B) Illustration of a sorting of protoplasts transformed by a p35::VENUS:NLS construct. On the left, in green, non-fluorescent untransformed protoplasts, on the left transformed protoplasts. The average fluorescence of the positive protoplasts is used as a proxy for the activation or repression efficiency in C) and D). C) Induction of a VENUS YFP by the scaffold, fusion and SAM systems (combination of scaffold and fusion system). The VENUS promoter comprises 7 teto sequences recognized by the sgRNA used in the three systems. The graph in the upper left corner represents the activation fold of the VENUS by each system. The bottom graph represents the percentage of activation compared to the strong and constitutive p35:VENUS promoter. The best system achieves an induction equivalent to 36% of a 35S promoter. D) Repression of a 35S promoter coupled to a VENUS YFP by all three systems. The graph represents the percentage of repression achieved by each system. The best system achieves nearly 40% repression of a strong promoter

We have recently demonstrated (unpublished results) the feasibility of this approach in rice protoplasts. We used a fluorescent reporter gene (YFP) and a cell sorter **(Figure 3B)** to evaluate the activation of a silent promoter and the repression of a strong promoter. We were able to obtain an activation equivalent to 40% of that of a strong promoter **(Figure 3C)** and a repression of more than 35% of a strong 35S promoter **(Figure 3D)**. We propose to use and develop this technology in rice callus protoplasts - i.e. totipotent cells - in the frame of the GREENER project to validate in combination some of the TF candidates to validate candidate TF and gene network in rice protoplasts. Expression of downstream effectors will be followed by digital droplet PCR and RNAseq.

References

Bargmann, B.O., Marshall-Colon, A., Efroni, I., Ruffel, S., Birnbaum, K.D., Coruzzi, G.M., and Krouk, G. (2013). TARGET: a transient transformation system for genome-wide transcription factor target discovery. Mol Plant 6:978-980.

Beck, F., Burch, M., Diehl, S., and Weiskopf, D. (2017). A Taxonomy and Survey of Dynamic Graph Visualization. Computer Graphics Forum 36:133-159.

Chavan, S., Schnabel, E., Saski, C., and Frugoli, J. (2018). Fixation and Laser Capture Microdissection of Plant Tissue for RNA Extraction and RNASeq Library Preparation. Curr Protoc Plant Biol 3:14-32.

Chen, W.Q., Drapek, C., Li, D.X., Xu, Z.H., Benfey, P.N., and Bai, S.N. (2019). Histone deacetylase HDA19 affects root cortical cell fate by interacting with SCARECROW. Plant Physiol.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res 23:1163-1171.

Diaz-Trivino, S., Long, Y., Scheres, B., and Blilou, I. (2017). Analysis of a Plant Transcriptional Regulatory Network Using Transient Expression Systems. Methods Mol Biol 1629:83-103.

Elati, M., Neuvial, P., Bolotin-Fukuhara, M., Barillot, E., Radvanyi, F., and Rouveirol, C. (2007). LICORN: learning cooperative regulation networks from gene expression data. Bioinformatics 23:2407-2414.

Galatro, TF., Holtman, IR., Lerario, AM., Vainchtein, ID., Brouwer, N., Sola, PR., Veras, MM., Pereira, TF., Leite, REP., Möller, T., Wes, PD., Sogayar, MC., Laman, JD., den Dunnen, W., Pasqualucci, CA., Oba-Shinjo, SM., Boddeke, EWGM., Marie, SKN., Eggen, BJL. (2017) Aug;20(8):1162-1171. doi: 10.1038/nn.4597. Epub 2017 Jul 3. Transcriptomic analysis of purified human cortical microglia reveals age-associated changes.Nat Neurosci.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154:442-451.

Kawai, M., Samarajeewa, P.K., Barrero, R.A., Nishiguchi, M., and Uchimiya, H. (1998). Cellular dissection of the degradation pattern of cortical cell death during aerenchyma formation of rice roots. Planta 204:277-287.

Kerren, A., Purchase, H.C., and Ward, M.O. (2014). Introduction to multivariate network visualization. Lecture Notes in Computer Science 8380:1-9.

==Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., and Porter, M.A. (2014). Multilayer networks. Journal of Complex Networks 2:203-271. Krzywinski, M., Birol, I., Jones, S.J., and Marra, M.A. (2012). Hive plots–rational approach to visualizing networks. Brief Bioinform 13:627-644.==

Lou, Y., Chen, YD., Sun, FR., Shi, JP., Song, Y., Yang, J. (2017) Cell Physiol Biochem. 41(1):239-251. doi: 10.1159/000456061. Epub 2017 Jan 30. Potential Regulators Driving the Transition in Nonalcoholic Fatty Liver Disease: a Stage-Based View.

Lowder, L.G., Zhang, D., Baltes, N.J., Paul, J.W., 3rd, Tang, X., Zheng, X., Voytas, D.F., Hsieh, T.F., Zhang, Y., and Qi, Y. (2015). A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. Plant Physiol 169:971-985

Mounier, T., Navarro Sanz, S., Bureau, C., Lefeuvre, A., Varoquaux, F., Durandet, F., Périn, C. A fast, efficient and high-throughput procedure involving laser microdissection and RT droplet digital PCR for tissue-specific expression profiling of rice roots. (2020 in revision for BMC Cellular and Molecular Biology)

McGee, F., Ghoniem, M., Melançon G., Otjacques B., and Pinaud B. (2019), "The State of the Art in Multilayer Network Visualization," Computer Graphics Forum, vol. 38, no. 6, pp. 125-149.

Nicolle, R., Radvanyi, F., and Elati, M. (2015). CoRegNet: reconstruction and integrated analysis of co-regulatory networks. Bioinformatics 31:3066-3068.

Sato, Y., Takehisa, H., Kamatsuki, K., Minami, H., Namiki, N., Ikawa, H., Ohyanagi, H., Sugimoto, K., Antonio, B.A., and Nagamura, Y. (2013). RiceXPro version 3.0: expanding the informatics resource for rice transcriptome. Nucleic Acids Res 41:D1206-1213.

Takehisa, H., Sato, Y., Igarashi, M., Abiko, T., Antonio, B.A., Kamatsuki, K., Minami, H., Namiki, N., Inukai, Y., Nakazono, M., et al. (2012). Genome-wide transcriptome dissection of the rice root system: implications for developmental and physiological functions. Plant J 69:126-140.

Trebulle, P., Nicaud, J.M., Leplat, C., and Elati, M. (2017). Inference and interrogation of a coregulatory network in the context of lipid accumulation in Yarrowia lipolytica. NPJ Syst Biol Appl 3:21.

Winterhalter, C., Nicolle, R., Louis, A., To, C., Radvanyi, F., and Elati, M. (2014). Pepper: cytoscape app for protein complex expansion using protein-protein interaction networks. Bioinformatics 30:3419-3420.

Yamauchi, T., Yoshioka, M., Fukazawa, A., Mori, H., Nishizawa, N.K., Tsutsumi, N., Yoshioka, H., and Nakazono, M. (2017). An NADPH Oxidase RBOH Functions in Rice Roots during Lysigenous Aerenchyma Formation under Oxygen-Deficient Conditions. Plant Cell 29:775-790.

Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., et al. (2015). Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell 160:339-350.