THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Biologie fonctionnelle

École doctorale GAIA
Portée par

L'unité de recherche AGAP

Vers le développement d'une nouvelle approche pour étudier les mécanismes moléculaires impliqués dans la formation des aérenchymes racinaires chez le riz

Présentée par Hans Thibault Mounier
Le 12/12/2019

Sous la direction de Christophe Périn

Devant le jury composé de

Fernanda De-Carvalho-Niebel, HDR, CNRS	Rapporteur
Fabien Nogue, HDR, INRA	Rapporteur
Jean-Luc Verdeil, HDR, CIRAD	Examinateur
Olivier Panaud, Professeur, UPVD	Président du jury
Pascal Gantet, Professeur Université Montpellier	Examinateur

UNIVERSITÉ DE MONTPELLIER

Montpellier SupAgro

Remerciements

Tout d'abord, un grand merci à ma **mère et à mes petits frères**, parce qu'idéalement, ça commence par-là ! Vous êtes pour moi une source d'énergie incroyable, même à des milliers de kilomètres. Ces trois années ont vu des hauts et des bas, vous êtes resté un socle solide qui m'aide à garder l'équilibre, en toute circonstance. MI AIM A ZOT FOR ! Qu'est-ce que j'ai hâte de vous voir !

A Christophe Périn, A.K.A Grand Manitou, mon encadrant de thèse. Tu as fait partie des personnes qui m'ont donné envie de continuer mes « études ». Ta vision de la recherche, tes connaissances et ton humanité m'ont donné envie de me lancer dans cette aventure qu'est la thèse. C'est un beau cadeau que tu m'as fait là. Pour toutes ces discussions, pour ton soutien, scientifique ou pas, merci !

A **Sergi Navarro-Sanz**, A.K.A el Puto Amo, l'homme inébranlable. On ne peut pas rêver d'un meilleur collègue de travail, un sang-froid à faire frissonner un bonhomme de neige ! On est parti à l'assaut des protocoles ensemble, moi fébrilement, et toi avec l'assurance qu'on y arrivera, toujours. Et on y est arrivé, de justesse à chaque fois, mais on y est arrivé ! Ça aurait été très « compliqué » (quel euphémisme) sans toi, là, ça a été un plaisir!

Un grand merci à l'équipe DAR : **Mumu** mon petit soleil et **Aurore** de l'underground team, vous transformez des cals de riz comme personne. **Anne-Cécile**, A.K.A CRISPR master, **Léo, Ian, Martine, Donaldo,** on en aura passé du temps ensemble à pipeter, vortexer, PCRiser jusqu'à la crampe de pouce! A **Eve, Remi et Christian**, A.K.A the greenhouse team, c'est un plaisir d'avoir travaillé avec vous ! A **Fabrice**, un enseignant plus qu'impliqué dans sa mission. A **Anne** qui doit régulièrement nous redescendre sur terre. Je ne saurai même pas par où commencer pour décrire la synergie qui anime cette équipe, il faut le voir pour y croire. Merci pour ce deuxième foyer que vous m'avez offert !

A la start-up IAGE, Franz, Antoine et Maxime pour cette émulation dans mon bureau. Travailler à vos côtés est un plaisir ! On va en sortir quelque chose de solide !

A **Charlotte Bureau**, avec qui on a commencé le développement du protocole de laser microdissection, les débuts ont été folkloriques ! Franchement, il y'a eu des moments je croyais vraiment qu'on n'y arriverait pas. Et au final, on a mis au point ce protocole dont je suis très fier, une belle collaboration ! Ravi pour ton poste au CNRS !

Au stagiaires avec qui j'ai collaboré durant ma thèse, **Emilie Clastres, Dylan Gallo et Louis Prieur**. Dylan, merci pour les résultats que tu as produit, tu as pu démontrer que mes facteurs de transcription synthétiques marchaient ! Ça faisait plus d'un an que je les savais prêt, au congélateur,

et qu'ils n'attendaient que la rencontre avec un protoplaste… Louis, on aura pu démontrer que mon protocole LM pouvait être facilement adapté à d'autres questions de recherche, ça fait du bien de se dire qu'on ne développe pas quelque chose juste pour soi.

A mon groupe des 4 fantastics, merci pour votre soutien tout le long de la rédaction, et de cette folle année 2019. Jordan tu en aura passer des heures à juste m'offrir ta présence silencieuse dans ma grotte de rédaction, je me serai senti bien seul sinon!

A tous mes copains, les anciens, les nouveaux, les qui s'en vont et qui reviennent, avec qui j'ai échangé des sourires tranches papayes, des sessions de kitesurf, des soirées de fou. Je me sens super bien entouré !

Merci à **Eole**, d'avoir abondamment fourni le Sud de France en vent durant ces 5 dernières années. Tu m'as offert des sessions de kitesurf mémorables qui étaient plus que bienvenues pour relâcher la pression.

MERCI A LA VIE !!

Essai personnel sur ma vision des approches en biologie

Mes 3 premières années d'études supérieures ont eu majoritairement pour sujet les sciences de l'écologie. Un premier stage en agroécologie en deuxième année de DUT m'a vite amené à me diriger vers le monde de la recherche. J'avais pu m'immerger dans ce monde à part, en ressentir la stimulation intellectuelle. Goûter à la grande satisfaction de mettre en place une méthodologie permettant de décrire, voire de caractériser, un phénomène biologique.

J'ai donc continué sur une licence orientée recherche, mais malgré tout l'intérêt que j'ai pu porter aux sciences de l'écologie, quelque chose me manquait. En effet, l'appréhension du fonctionnement d'un écosystème passe par des données descriptives et une compréhension empirique de phénomènes environnementaux. Les méthodologies expérimentales qui sont employées dans ces domaines, rigoureuses et d'une qualité scientifique certaine, ne me satisfaisait pas totalement niveau stimulation intellectuelle. Pour résumer, j'étais en manque de compréhension fine des mécanismes, notamment les mécanismes biologiques sous-jacents aux phénomènes que l'on pouvait décrire dans le domaine de l'écologie.

Je me suis donc dirigé vers un master de biologie fonctionnelle des plantes. La biologie moléculaire me semblait être le meilleur choix si je voulais rester dans la biologie et m'intéresser aux mécanismes régulant le vivant et son interaction avec son environnement. Quelques mois d'adaptation au langage, et aux raisonnements liés à la biologie fonctionnelle ont été nécessaires. J'ai commencé, à travers le master, les stages, les discussions, à accumuler des connaissances sur les mécanismes moléculaires sous-jacents aux processus biologiques et sur la façon de les disséquer au niveau expérimental.

Après mon premier stage en master 1, je me suis dit que j'étais maintenant de « l'autre côté de la biologie »: on s'attaque à l'étude de phénomènes complexes, avec des niveaux multiples de régulations et d'interactions, en décomposant cette complexité en sous-groupes d'éléments pouvant être étudié indépendamment les uns des autres (Von Bertalanffy, 1950). Cette approche mécanistique a permis de grandes avancées et l'acquisition de connaissances et d'outils incroyables : le séquençage, l'épiséquençage, la technologie CRISPR/Cas9, les expérimentations en cellules isolées, la sélection assistée par marqueurs, pour n'en citer que quelques-uns. Cependant, de mon point de vue, la seule accumulation de connaissances

provenant de ces expérimentations ne pourront jamais permettre une compréhension globale d'un phénomène biologique.

Une image me vient en tête pour exprimer cette pensée : comment comprendre une pièce de théâtre dans son intégralité si on en voit que quelques acteurs à la fois ? Si on la joue avec des scènes sans ligne temporelle et des dialogues manquants? De même, pour un exemple en biologie : comment comprendre les mécanismes sous-jacents à la formation des aérenchymes, tissu répondant à tout un panel de stress et de processus développementaux, si on étudie leur formation presque strictement dans le cadre de la tolérance à la submersion. Quid des autres interactions ? Comment peut-on comprendre un processus biologique sans en faire apparaître la « dynamique », c'est-à-dire des notions de temps, d'espace, d'interactions dans nos modèles ?

Je me retrouvai donc, en fin de master 1, en quelque sorte dans une situation inversée par rapport à mes années de licence : là ou avant, en écologie, je voyais une pièce de théâtre dans sa globalité mais sans entendre les dialogues discrets entre acteurs, je me voyais ensuite suivre un acteur ou une scène en oubliant tout le reste de la représentation. Il semble difficile de lier vision globale et décorticage fin des mécanismes d'un phénomène biologique.

J'ai donc commencé mon stage de master 2 avec une envie de me consacrer à des travaux qui s'apparentaient à du développement plutôt qu'à de la recherche fondamentale. Je voulais en quelque sorte m'échapper des réflexions soulevées dans le paragraphe précédent en utilisant mes compétences pour répondre à un objectif simple et arrêté : développer un outil moléculaire permettant d'améliorer l'efficacité du génome *editing* chez le riz. Poursuivre en thèse ne m'intéressait alors pas beaucoup, n'ayant pas trouvé de thème de recherche qui me satisfaisait totalement.

Quelques mois après le début du stage, beaucoup de bibliographie et de discussions avec mon tuteur de stage, Christophe Périn, et on se met à discuter de l'éventualité de monter un projet de thèse ensemble. Deux choses m'ont décidé à dire oui: l'environnement humain dans lequel je baignais me laisser penser que ça allait être une aventure formidable, et bien sûr le sujet de thèse qui s'est développé au cours de mon stage me plaisait beaucoup.

En effet, en ayant à l'époque beaucoup moins de recul sur le projet, je pouvais déjà « sentir » que la stratégie de recherche que l'on voulait mettre au point était une tentative

pour faire un lien entre vision globale et compréhension fine en biologie moléculaire. Notre stratégie s'apparente à ce que l'on appelle les « *data-driven approach* » : intégrer et structurer des données massives de sortes que l'on puisse interroger facilement et avec pertinence cette base pour diriger l'expérimentation (Pouvreau *et al.*, 2018, Farré *et al.*, 2015, Greene and Troyanskaya, 2010). Ce type d'approche en biologie moléculaire a été rendue possible grâce à l'avènement de nouveaux outils :

- Le développement d'outils moléculaires permettant de manipuler finement le génome, l'épigénome, le transcriptome (Shalem *et al.*, 2015, Xu et Qi, 2019, Shi *et al.*, 2016). L'équipe possède une expertise sur les technologies CRISPR/Cas9, permettant la mutation, la manipulation de l'expression génique et de l'épigénome, une belle boite à outils moléculaires !
- Les technologies –omiques à moyen/haut débit qui sont de plus en plus fiables et peu couteuses.
- Les progrès informatiques en « exploration de données » et modélisation (Wang et Huang, 2014, Nicolle *et al.*, 2015). Cette dernière partie fait l'objet d'une collaboration avec une équipe de recherche spécialisées en bio-informatique qui possède déjà les outils nécessaires à une exploration et une structuration pertinente de nos données (équipe I3Bionet).

Un chercheur finalement, c'est une base de données structurée que l'on interroge *via* sa capacité de synthèse et surtout sa mémoire. On a maintenant un moyen d'accélérer l'acquisition et de virtualiser cette connaissance, de la structurer et de la nourrir collectivement (Thiele *et al.*, 2013, Demir *et al.*, 2010). De plus, ce type d'approche permet de sortir d'une finalité strictement tournée vers la publication dans la recherche : toute information tirée d'expérimentations pourra servir à nourrir ces bases de données pour en améliorer les prédictions. Ainsi, même les études n'amenant pas de résultats « publiables » seront valorisées. Créer de la connaissance, pour la connaissance, et l'utiliser pour produire de nouvelles connaissances!

Le but ultime est la simulation informatique de phénomènes biologiques faisant intervenir des régulations transcriptionnelles intra et inter-cellulaire, des voies métaboliques, l'environnement cellulaire ou extra-cellulaire, le temps, l'espace (Karr *et al.*, 2012, Ritchie *et al.*, 2015, Hutchison *et al.*, 2016). Pour certaines de ces données, il est difficile, voire

impossible, de les mesurer expérimentalement, du moins avec les outils actuels. Comment peut-on alors appréhender des phénomènes que l'on ne peut mesurer expérimentalement? Comment disséquer la complexité du fonctionnement du vivant en menant une multitude de recherches indépendantes les unes des autres ? L'approche biologie des systèmes permet de pallier quelques-unes de ces limites puisqu'elle ne demande pas de connaissances à priori sur la mécanistique des processus biologiques étudiés, tout en permettant de générer des hypothèses. La théorie générale des systèmes tel qu'énoncé par Ludwig Von Bertalanffy en 1950 est un domaine logico-mathématique dont l'objet est la formulation et la déduction des principes valables pour les « systèmes » en général. De nombreux principes ont été développés en sciences des systèmes, notamment sur la notion de robustesse des systèmes, qui implique des redondances fonctionnelles des éléments, des boucles de rétro-action, de la modularité et une stabilité structurelle (Emmert-Streib et Dehmer, 2010, Kitano *et al.*, 2002, Kitano *et al.*, 2002 (b)). Principes que l'on retrouve expérimentalement en biologie moléculaire.

On peut considérer cette approche comme étant une approche heuristique, générer de la connaissance en employant une méthodologie logique qui ne garantit pas d'être optimale, ni de refléter totalement la réalité, mais uniquement de pouvoir prédire et diriger les futures expérimentations. Ainsi, un organisme, une cellule, un organite, serait ici considéré comme un système, c'est-à-dire un ensemble d'éléments ayant des propriétés particulières et interagissant les uns avec les autres selon certains principes (Cramer *et al.*, 2011, Kitano *et al.*, 2002). Le fonctionnement global d'un organisme serait donc la résultante des interactions entre les ensembles d'éléments qui le composent. Partant de ce principe, l'intégration d'informations biologiques massives et multi-échelles est nécessaire (Hong *et al.*, 2016, Cramer *et al.*, 2011, Greene and Troyanskaya, 2010). L'obtention de données –omiques, physiologiques, couplées à une description précise de l'échantillonnage également. Il découle de cette considération que la biologie des systèmes n'est pas seulement une approche de modélisation et de valorisation de données, mais est aussi une approche expérimentale nouvelle car elle nécessite l'acquisition de données pertinentes et idéalement de l'expérimentation à haut débit. C'est de ce côté de la biologie des systèmes et des approches « *data-driven* » que je me place maintenant, à l'issue de ma thèse.

Table des matières

Remerciements **3**

Essai personnel sur ma vision des approches en biologie **5**

Liste des abréviations **16**

Introduction générale **21**

1. Le riz, Importance agronomique et sécurité alimentaire 21

1.1 Le riz et l'Homme. 21

1.1.1 Domestication du riz 21

1.1.2 Importance agronomique et sociale 21

1.1.3 Les différents systèmes de culture 23

1.1.4 Le riz et la sécurité alimentaire mondiale 25

1.2 Le riz, une céréale modèle 25

1.2.1 Importance de développer un modèle monocotylédone 25

1.2.2 Le riz, un modèle idéal 26

2. Le système racinaire, parent pauvre des programmes de sélections 29

2.1 Architecture racinaire, idéotypes et génétique directe 31

2.1.1 Architecture fasciculée des racines du riz 31

2.1.2 A la recherche d'idéotypes d'architecture racinaire 32

2.1.3 Etude de l'architecture racinaire et tolérance au stress abiotique : approches QTL 32

2.2 Anatomie racinaire, idéotypes 35

2.2.1 Anatomie racinaire du riz 35

2.2.2 Tissus internes et adaptation aux stress 35

2.2.3 Formation du cortex 37

2.3 Aérenchymes, des enjeux agronomiques et écologiques 41

2.3.1 La submersion, un problème affectant la sécurité alimentaire 41

2.3.2 Les aérenchymes et la tolérance à la submersion 41

2.3.3 Les tissus externes forment une arrière à la diffusion radiale d'oxygène 43

2.3.4 Les aérenchymes racinaires un trait d'intérêt agronomique, au-delà de la tolérance à la submersion 43

3. Le cycle Design-Build-Test-Learn (DBTL) pour l'identification du réseau de gènes intervenant dans la mise en place des aérenchymes 47

3.1 Le cycle DBTL, utiliser des séries de tests itératifs pour affiner des hypothèses..... 47

3.3 Objectifs de la thèse 48

Chapitre 1, article soumis pour publication dans Trends in Plant Science: Root aerenchyma: this is more than just holes 53

ABSTRACT 53

MAIN TEXT 55

REFERENCES 63

GLOSSARY 65

Chapitre 2: Développement d'un protocole de Laser Microdissection (LM) couplé à la technologie RT-dd-PCR..... 69

1.Introduction..... 69

1.1 Soyons plus spécifique : Récolter des informations –omiques pertinentes 69

1.2 Isoler des groupes de tissus 69

1.3 Laser Microdissection, pertinence pour l'étude du développement des aérenchymes 71

1.4 Bref état de l'art de la technologie LM : limites des approches qui ont précédées .. 71

1.5 LM/RT-ddPCR, une méthode fiable et robuste de déterminer des profils d'expressions tissus spécifiques..... 72

2. OBJECTIFS 73

3. Article soumis dans BMC Cellular and Molecular Biology: *A fast, efficient and high throughput procedure for tissue specific expression profiling of rice roots using Laser Microdissection and RT-digital droplet PCR* 74

ABSTRACT 75

BACKGROUND 76

RESULTS 79

DISCUSSION 87

CONCLUSIONS 90

METHODS 91

REFERENCES 96

ANNEXE 98

4. RESULTATS ET DISCUSSION : COMPLEMENTS SUR LE DEVELOPPEMENT DU PROTOCOLE 107

3.1 La température et le temps de traitement ont un effet sur la qualité de l'ARN. 107

3.2 La qualité histologique est dépendante des contraintes mécaniques auxquelles sont soumises l'échantillon 109

3.3 Les limites du LM 109

3.4 Un protocole LM adapté à différents types d'expérimentations 111

3.5 Le couplage LM/ddPCR, vers une quantification fiable et rapide de l'expression génique dans des groupes de tissus spécifiques (voir l'article intégré, Mounier *et al.*, soumis) 113

5. Conclusions du chapitre 2 113

Chapitre 3 : Construction de banques d'ADNc tissu et stade de développement spécifique 116

1. Introduction 116

1.1 Les données moléculaires sur les acteurs de formation des aérenchymes portent exclusivement sur les aérenchymes *induits* 116

1.2 Les données tissues spécifiques publiées jusqu'à présent sur les céréales ne permettent pas l'investigation sur la dynamique développementale aboutissant à la formation des aérenchymes 117

1.3 Du développement du cortex à la formation d'aérenchymes matures : un processus en 4 phases 119

1.4 Laser microdissection (LM), avantages et limites techniques 120

1.5 Coupler la technologie LM et le RNA-seq, contraintes causées par la faible quantité d'ARN extraite 120

1.6 Limiter les biais générés lors de la création de la banque d'ADNc 121

2. Objectifs 122

3. Matériels et méthodes 123

3.1 Microdissection laser des trois tissus 123

3.2 Extraction des ARN, évaluation de la qualité et de la quantité 123

3.3 Construction du kit Ovation® SoLo RNA-Seq Library Preparation, analyses des séquences ARNr chez le riz 123

3.4 Création des banques RNAseq et analyses de qualités 125

4.Résultats 127

4.1 Microdissection des 4 stades de développement du cortex (**Figure 23**) 127

4.2 Création des banques RNAseq, profils de répartition des fragments par taille 127

5. Discussion : un protocole qui permet de construire des banques RNAseq à partir de très petites quantité d'ARN pour le riz : limites 129

5.1 La microdissection de sections racinaires au niveau de la phase de propagation de la mort cellulaire n'a pas permis d'extraire d'ARNs 129

5.2 LM tissus/stades de développement spécifiques et qualité des échantillonnages. 131

5.3 Le kit kit Nugen Ovation® SoLo RNA-Seq Library permet de partir de quantités plus faibles d'ARN et permet une meilleure conservation de la stœchiométrie entre les transcrits 135

5.4 La déplétion des ARNr permet d'optimiser le nombre de lectures lors du séquençage 137

5.5 La production d'une librairie brin sens spécifique permettra potentiellement la détection de nouvelles espèces d'ARN 139

6. Conclusion 139

Chapitre 4, *Première version* d'un futur article: CRISPR based synthetics transcriptions factors to decipher gene network in rice 142

1. Etapes *Design* et *Build* du cycle DBTL, utilisation de facteurs de transcriptions synthétiques basés sur la technologie CRISPR/Cas en cellules isolées 142

2. CRISPR based synthetics transcriptions factors to decipher gene network in rice 143

ABSTRACT 143

Introduction 145

Results 151

Experimental design 151

CRISPR RNA scaffold system effectively activate VENUS YFP expression 151

CRISPR RNA scaffold effectively repress VENUS YFP expression 155

Discussion 155

Conclusion 159

Materials and methods 160

scRNA sequence design 160

Coding protein sequence design 160

Protoplast extraction and transformation 161

Visualization and quantification of the fluorescence 161

3. Facteurs de transcription synthétiques CRISPR/Cas et protoplastes, vers un système expérimental permettant de tester les hypothèses du cycle DBTL 163

DISCUSSION, CONCLUSION, ET PERSPECTIVES DU PROJET 167

1. Transcriptome tissu/stade développemental spécifique de la formation des aérenchymes, vers l'inférence d'un réseau de co-régulation de gènes. 167

2. Les facteurs de transcriptions synthétiques basés sur la technologie CRISPR/Cas9, des premiers résultats prometteurs et quelques pistes pour améliorer le système 168

3. Approche comparative riz-sorgho pour élucider les mécanismes moléculaires de formation des aérenchymes chez les céréales: cas du gène SbNAC.D du Sorgho..... 169

4. Vers une approche DBTL pour l'étude des mécanismes moléculaires de formation des aérenchymes racinaires 171

REFERENCES GENERALES.....176

ANNEXES 203

Annexe 1 : Séquences des ARN ribosomaux du riz utilisé pour le design des sondes de déplétion par Nugen 203

Annexe 2: Cell sorting..... 209

Annexe 3 : spacer and guide RNA sequences 209

Annexe 4: dCas9, VP64 and 3xSRDX sequences..... 210

Annexe 5; MCP-VP64 and COM-3xSRDX sequences 211

Annexe 6: promoter and reporter gene sequences..... 211

Annexe 7, article accepté dans Methods in Molecular Biology: Efficient genome editing in rice protoplasts using CRISPR/CAS9 construct 213

1 Introduction..... 214

2. MATERIALS 217

3. METHODS 221

4 Notes 231

5. Tables..... 233

REFERENCES 235

Liste des abréviations

ACC	1-Aminocyclopropane-1-carboxylic acid
1-MCP	1-Methylcyclopropene
ACO	ACC Oxydase
ACS	ACC Synthase
ADNc	ADN complémentaire
ARNr	ARN ribosomal
CDPK	Calcium-Dependent Protein Kinase
CRISPR	Clustered regularly interspaced short palindromic repeats
CEEI	Common-Endodermis-Epidermis-Initial
CEI	Cortex-Endodermis-Initial
Cas	CRISPR associated
CRISPRi	CRISPR inhibitor
CRISPRa	CRISPRactivator
DRO1	DEEP ROOTING 1
DOCS1	DEFECTIVE IN OUTER CELL LAYER SPECIFICATION 1
DBTL	Design-Built-Test-Learn
ddPCR	Droplet Digital PCR
SRDX	EAR-repression domain
RBOHH	Respiratory Burst Oxidase Homolog protein H
FACS	Fluorescence Activated Cell Sorting
FAO	Food and Agriculture Organization
GRN	Gene Regulatory Network
GFP	Green Fluorescence Protein
INTACT	Isolation of Nuclei TAgged in specific Cell Types
KRAB	Krüppel Associated Box
LLR	Large Lateral Root
LM	Laser Microdissection
MT	MetalloThionéine
MCP	MS2 Coat Protein
NCBI	National Center for Biological Information
NAC	NAM, ATAF1,2, CUC
NBT	New Breeding Tool
NRT1.1	Nitrate Transporter 1.1
NUE	Nitrogen Use Efficiency
NBC	Nitrotetrazolium Blue Chloride
NLS	Nuclear Localisation Sequence
PSTOL1	Phosphate-Starvation TOLerant 1
Pup1	Phosphate-UPtake 1
PCD	Plant Cell Death
PCR	Polymerase Chain Reaction
QTL	Quantitativ Trait Loci
qSOR1	Quantitative Trait Loci for Soil Surface Rooting 1
ROL	Radial Oxygen Lost
ROS	Reactive Oxygen Species

RT-qPCR	Real-time Quantitative PCR
RT	Reverse Transcription
RIN	RNA Integrity Number
RCA	Root Cortex Aerenchyma
scRNA	Scaffolded RNA
SCR	Scarecrow
SHR	Shortroot
sgRNA	Single Guide RNA
SPIA	Single Primer Isothermal Amplification
SLR	Small Lateral Root
Sk1/Sk2	Snorkelin 1/2
SUB1A	Submergence 1
SAM	Synergic Activation Mediator
TALE	Transcription Activator-Like Effector
T-DNA	Transfer-DNA
TARGET	Transient Assay Reporting Genome-wide Effects of Transcription factors
VLCFA	Very Long Chain Fatty Acid
YFP	Yellow Fluorescent Protein

INTRODUCTION GENERALE

Introduction générale

1. Le riz, importance agronomique et sécurité alimentaire

1.1 Le riz et l'Homme.

1.1.1 Domestication du riz

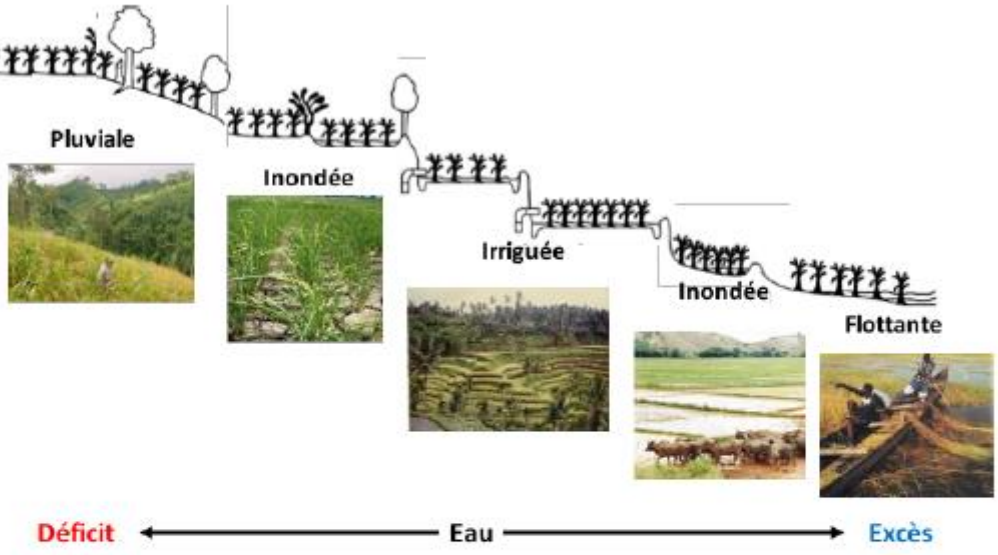
Il existe deux espèces de riz cultivées, *Oryza glaberryma*, également appelé riz africain ou de Casamance, cultivé essentiellement en Afrique de l'ouest, et *Oryza sativa*, à la distribution mondiale (Food and Agriculture Organization, FAO). L'origine du riz africain fait l'objet d'un consensus et correspond à un évènement de domestication unique d'une espèce de riz sauvage, *Oryza barthii,* autour du fleuve Niger il y a 3500 à 5000 ans.

Le riz asiatique, *Oryza sativa*, peut se décomposer en deux sous-espèces, *japonica* (le riz à grains court), *indica* (riz long grain) et comprend également *Oryza sativa aus*, considéré dans le passé comme une sous-espèce (Li *et al.*, 2014). L'origine et la domestication d'*Oryza sativa,* il y a 7000 à 10000 ans, est plus controversée que celle d'*Oryza glaberryma*. Deux grandes hypothèses co-existent encore à l'heure actuelle. La première repose sur de multiples évènements de domestication ayant conduit à la sélection des deux sous-espèces, théorie défendue par Civán *et al.*, en 2016 et Kovach *et al.*, en 2007. La deuxième hypothèse suppose un évènement unique de domestication qui aurait donné naissance à la sous-espèce *japonica* dont certains allèles domestiqués auraient été par la suite introgressés dans des variétés ancestrales de riz proto-indica pour donner les variétés *indica* actuelle (Gross et Zhao, 2014), théorie défendue notamment par Choi et Purugganan en 2018 et Huang *et al.,* en 2012 et qui semble être maintenant la théorie dominante.

1.1.2 Importance agronomique et sociale

"Probably more so than any crop, it drove societies and economies to become densely populated, potentially more urbanized, and it also transformed landscapes" Dorian Fuller, archéobotaniste à l'université de Londres à propos de la place du riz dans les cultures asiatiques.

A



B

	Pluviale	Inondée	Irriguée	Eau profonde	Total (ha)
Afrique (%)	38	34	20	8	9,1
Amérique latine (%)	45		55		6,6
Asie (%)	9	28	57	6	136
USA, UE, Aust. (%)			100		4
Total (%)	10	30	55	5	100
Superficies (Mha)	16	48	86	7	157
Rendement (t/ha)	1 à 5	1 à 5	3 à 10	1 à 3	4,2

Figure 1: Diversité des systèmes rizicoles, quelques chiffres

A. Les 4 grands types de système rizicoles se distinguent en fonction du contrôle de l'irrigation : riziculture pluviale, inondée, irriguée ou flottante (eau profonde). Illustration d'Ahmadi et al., 2017.

B. Tableau présentant les surfaces occupées par chaque type de riziculture et les rendements moyens attendus (FAO 2005).

Près de la moitié des calories consommées par la population humaine provient de trois plantes de la famille des Poacées : le blé, le maïs et le riz. Si les superficies cultivées du maïs et du blé dépassent les 200 millions d'hectares de par le monde, le riz, céréale cultivée sur près de 170 millions d'hectares, est la céréale majeure pour l'alimentation humaine (FAOSTATS 2017). En effet, cette monocotylédone dont la domestication remonte à plus de 7000 ans, est à la base de l'alimentation de plus de la moitié de la population mondiale (Izawa et Shimamoto, 1996). Le continent asiatique consomme et cultive 90% du riz produit mondialement, ce qui représente 35% à 75% des apports caloriques pour plus de 3 milliards d'êtres humains (Zhang et Wing 2013). Les deuxièmes producteurs et consommateurs sont l'Amérique latine et l'Afrique, le riz représentant un tiers de leurs apports caloriques pour plus de 1,5 milliard d'êtres humains. Cette céréale est ainsi cultivée par plus de 144 millions d'agriculteurs dans plus d'une centaine de pays et reste, pour beaucoup, une culture vivrière dont ils dépendent directement (Laborte *et al.*, 2017). **Le maintien de sa production est donc critique pour assurer la sécurité alimentaire et la stabilité politique et sociale de nombreuses populations.**

1.1.3 Les différents systèmes de culture

Le riz est cultivé dans quatre grands types d'agro-écosystèmes rizicoles, qualifiés en fonction de leur hydrologie: irriguée, inondée, pluviale, flottante (Figure 1). A chacun de ces agrosytèmes correspond des types variétaux distincts. Le riz est une plante semi-aquatique, et par conséquent, le rendement est fortement dépendant du contrôle des flux d'eau. La riziculture irriguée est définie par un contrôle des flux d'eau entrant et sortant dans la rizière. Elle représente 55% de la surface cultivée de riz et 75% de la production mondiale, ce qui en fait le système rizicole le plus productif mais aussi le plus polluant à travers les émissions de méthane et d'oxyde nitrique et la consommation d'intrants (Cai *et al.*, 2000, Yan *et al.*, 2005). Ce mode de culture est majoritaire en Asie et en Amérique latine. Environ 30% des rizières sont périodiquement inondées, un mode de culture où les flux d'eau ne sont pas maitrisés mais où il subsiste une lame d'eau tout le long de la culture. La riziculture flottante représente environ 5% des superficies rizicoles dans le lit de fleuves et de rivières. Enfin, la culture pluviale, pour environ 10% de la surface cultivée en riz, est la seule riziculture qui dépend entièrement de la pluviométrie. Ce mode de culture est voué à prendre une place de plus en plus importante car moins gourmand en eau (Nascente *et al.*, 2013).

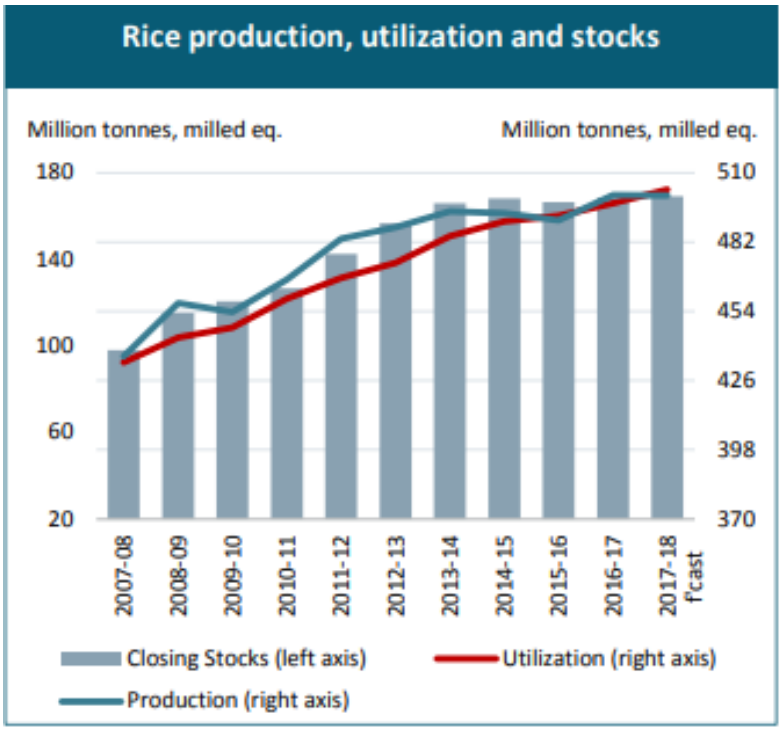


Figure 2: Production, consommation et stock mondial de riz d'après les données de la FAO
La saison 2015-2016 a marqué le début d'une consommation égale à supérieure à la production.

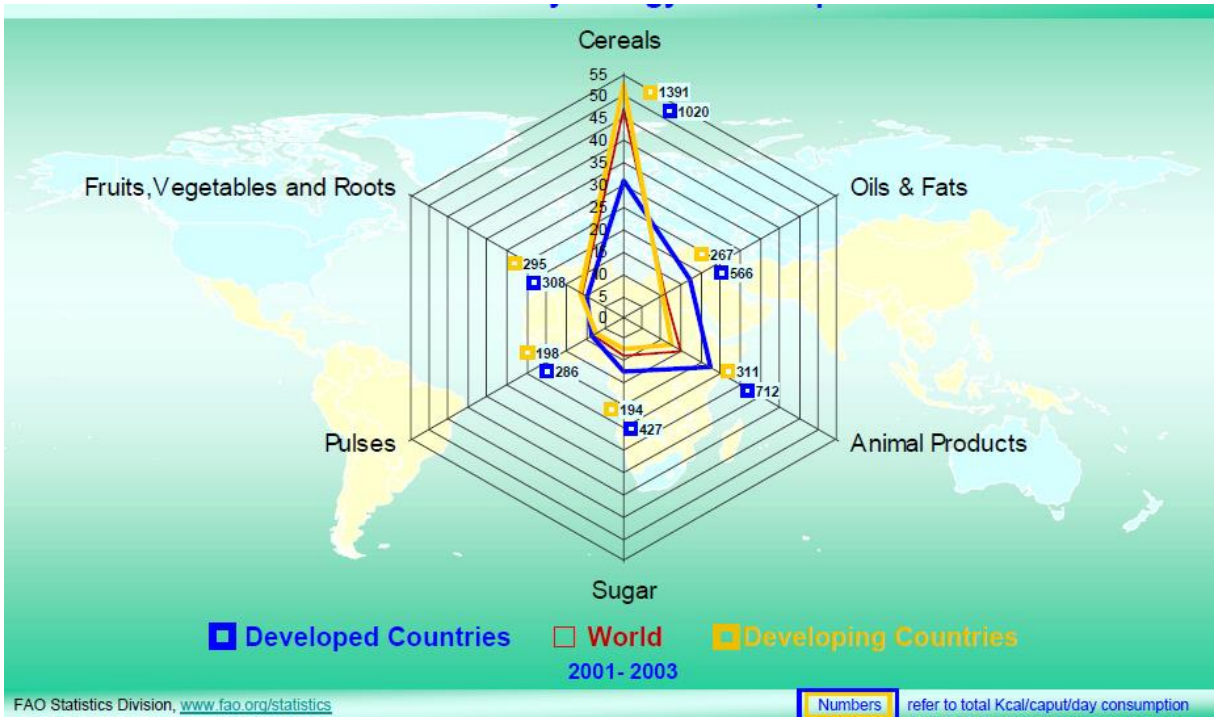


Figure 3: Source des calories consommées par l'humanité d'après les données de la FAO
Presque 50% des calories de l'humanité proviennent de la consommation de céréales, et s'élève à plus de 50% dans les pays en voie de développement.

1.1.4 Le riz et la sécurité alimentaire mondiale

D'après la FAO, l'augmentation de la population mondiale, et le changement climatique, poseront, dans quelques années, un véritable défi pour la riziculture. En effet, la consommation mondiale de riz n'a cessé d'augmenter ces dernières décennies, et la saison 2015-2016 a marqué le début d'une consommation mondiale égale à la production, ce qui pose la question de la pérennité des stocks de riz pour les prochaines décennies (**Figure 2**). Ainsi, le maintien d'une augmentation de production de 1% par an est nécessaire pour soutenir la consommation mondiale (Sasaki 2017). Li *et al.* en 2014 ont estimés qu'une hausse de la production de 25% sera nécessaire d'ici 2030. Contraintes supplémentaires, ces augmentations de rendement devront se faire avec une diminution des intrants et devront être résilients au changement climatique, hausse de température et sécheresse en premier lieu. Enfin, le riz contribue au réchauffement climatique *via* la libération de méthane principalement en riziculture irrigué et inondé (Schimel, 2000). **Il y a donc un besoin croissant de variétés de riz capables de maintenir leurs rendements en conditions de stress abiotiques et biotiques, accompagné d'une réduction de l'impact de la riziculture sur l'environnement.**

1.2 Le riz, une céréale modèle

1.2.1 Importance de développer un modèle monocotylédone

Le riz fait partie de la famille des poacées (ou graminées), qui comprend les céréales et la canne à sucre. Cette famille de plantes est la plus importante sur le plan socio-économique, puisqu'elle fournit presque la moitié des calories de l'humanité au début du 21ème siècle (**Figure 3**). Les poacées font partie des monocotylédones, l'un des deux grands groupes de plante à fleur, ce qui les éloigne donc phylogénétiquement de la plante modèle dicotylédone, *Arabidopsis thaliana.* **Si les travaux en cours sur *A. thaliana* ont permis de faire des progrès considérables, la nécessité de travailler sur une plante modèle complémentaire a rapidement émergé (Izawa et Shimamoto, 1996, Delseny 2003).**

1.2.2 Le riz, un modèle idéal

A ce titre, le riz a commencé à s'imposer comme la plante modèle chez les céréales dès la fin du XXème siècle (Izawa et Shimamoto, 1996). Le riz a un cycle de reproduction court, entre 100 et 210 jours, et plus précisément entre 130 et 150 jours pour la plupart des variétés domestiquées. Dans certaines conditions, il est possible de réaliser 3 voire 4 cycles en une seule année (Vergara, 1991). C'est la première céréale à avoir été génétiquement transformée, et ce, dès 1989, grâce aux travaux de Shimamoto *et al.* et elle est à ce jour la céréale la plus facile à transformer. **Ainsi, l'obtention de multiples générations de plantes après transformation ou sélection facilite l'analyse fonctionnelle de gènes d'intérêt agronomiques.**

Son génome est relativement petit (430 Mb) comparé aux autres céréales cultivées comme le maïs (2,4 Gb) et le blé (17 Gb), ce qui lui a valu d'être choisie pour être la deuxième plante dont le génome a été entièrement séquencé en 2005 (International Rice Genome Sequencing Project, 2005, Delseny 2003). La proximité phylogénétique entre les céréales, génome et synthénie compris, renforce l'intérêt du riz comme espèce modèle pour les études de génomique comparative au sein de cette famille (Gale et Devos., 1998, Akhunov *et al.*, 2013). Son cycle de reproduction court et la facilité de la transformation génétique ont permis de développer des projets de coopération à l'échelle internationale pour faciliter l'analyse fonctionnelle de gènes d'intérêt agronomiques comme la création de collection de lignées d'insertion (Miyao *et al.,* 2003, Kolesnik *et al*., 2004, Sallaud *et al.*, 2004, Hsing *et al.*, 2007). Il existe également de très nombreuses données d'expressions et bases de données les centralisant (plateforme d'expression par RT-qPCR : Caldana *et al.*, 2007 ; base de données puce à ADN : Sato *et al.*, 2013, http://ricexpro.dna.affrc.go.jp/, base de données séquençage : Xia *et al.*, 2017, http://expression.ic4r.org/) au sein de projets collaboratifs internationaux.

Il existe également une grande diversité génétique intra-espèce et intra-groupe variétal chez le riz (Li et Rutger, 2000, Yu *et al.*, 2003). Cette diversité est à l'origine de nombreux travaux de cartographie et de clonage de QTL pour des traits d'intérêt agronomique (Courtois *et al.*, 2009). Il existe ainsi une base de donnée publique

permettant leur recensement et leur annotation (Yonemaru *et al.*, 2010). La grande diversité génétique, moléculaire et phénotypique a également permis de développer de nouvelles procédures d'analyse statistique pour affiner ces QTL en se basant uniquement sur des données publiques (Yang *et al.*, 2019). Enfin, cette diversité génétique et phénotypique est explorée systématiquement à l'échelle du pangénome grâce au re-séquencage massif de nombreuses variétés de riz. Il existe ainsi une base de données publique comprenant plus de 3000 génomes de riz (The 3000 rice genomes project, 2014). De nombreux projets en cours vont conduire à une augmentation exponentielle de la quantité de données dans les prochaines années comme le séquençage et la cartographie de variétés de riz sauvage (Projet OMAP, Wing *et al.*, 2007, Bisht *et al.*, 2018).

Le riz est également une espèce modèle pour le développement des nouvelles technologies d'édition CRISPR/Cas (clustered regularly interspaced short palindromic repeats/ CRSIPR associated protein) chez les plantes. Ces technologies ont révolutionné l'analyse fonctionnelle en facilitant la mutagénèse ciblée de gènes d'intérêt chez les animaux et les végétaux. Tous les développements récents autour de cette technologie (base editor, nucléase Cpf1, CRISPR activator…) ont d'ailleurs tous été réalisés en premier lieu chez le riz (Li *et al.*, 2012, Jiang *et al.*, 2013, Miao *et al.*, 2013, Wang *et al.*, 2017, Hao *et al.*, 2019, Qin *et al.*, 2019). L'efficacité est telle chez le riz, et dans une moindre mesure chez d'autres espèces, que ces technologies sont maintenant testées et utilisées pour leur intégration dans les programmes de sélection sous la dénomination de NBT (New Breeding Tools). **Par exemple, Zhou *et al.* en 2019 ont pu ainsi augmenter le rendement de deux variétés élites en éditant trois gènes, correspondant à trois QTL isolés par de précédentes études (Fan *et al.* 2006; Mao *et al.* 2010, Song *et al.*, 2017, Ashikari *et al.*, 2005). Autre exemple, une technologie dérivée de CRISPR/Cas a été utilisée pour éditer et modifier un SNP dans un transcepteur de nitrate, le gène NRT1.1 impliqué dans la NUE (Nitrogen Use efficiency) (Qin *et al.*, 2019).**

La grande diversité génétique, le cycle de production court, les ressources moléculaires, la facilité de transformation, et le statut de plante d'intérêt agronomique majeur font du riz une plante modèle idéale. Sa proximité avec les autres céréales cultivées lui permet de compléter avantageusement *Arabidopsis*. **Ainsi, le riz présente un intérêt à la fois pour l'identification de gènes d'intérêt agronomique, le décryptage de**



Figure 4: Comparaison du système racinaire fasciculé du riz et pivotant *d'Arabidopsis thaliana*

A. Le système racinaire du riz est fasciculé, il ne s'organise pas autour d'une racine principale mais est composé de multiples racines pré ou post embryonnaire.

B. Le système racinaire d'*Arabidopsis thaliana* est pivotant, il s'organise autour de l'unique racine pré-embryonnaire d'où partent des racines latérales.

mécanismes biologiques fondamentaux, et l'amélioration génétique des céréales.

2. Le système racinaire, parent pauvre des programmes de sélections

Les programmes de sélection chez le riz ont jusque-là essentiellement porté sur l'augmentation du rendement en ciblant l'architecture aérienne de la plante (diminution de la taille, augmentation du nombre de panicules) ou des traits reliés directement à la qualité et au rendement (Fan *et al.* 2006; Mao *et al.* 2010, Song *et al.*, 2017, Ashikari *et al.*, 2005). Avec la volonté de développer de nouveaux idéotypes dans les années 90, le système racinaire est devenu une cible importante en sélection (Cassman, 1993, Shen *et al.*, 2012, Dievart *et al.*, 2013, Ahmadi *et al.*, 2014).

Le terme d'idéotype a été premièrement proposé par Donald en 1968 et décrit l'apparence idéalisée d'une variété de plante qui est censé fonctionner ou se comporter de manière particulière dans un environnement défini (Donald, 1968). L'idée est d'imaginer des modèles nouveaux de plantes adaptés à différentes conditions de culture (communauté de plantes), qui utilisent mieux que les variétés actuelles, les ressources du milieu (lumière, eau, éléments minéraux) afin de maximiser et stabiliser le rendement. Ce besoin est critique chez le riz, espèce cultivée dans une large gamme d'agro-écosystèmes, comparativement à d'autres céréales. Les racines jouent un rôle primordial dans l'ancrage et la captation d'eau et de nutriments à travers une interaction complexe avec le milieu environnant. Maintenir et optimiser ses fonctions est un challenge majeur dans des contextes d'instabilité climatique ou de limitation en intrants.

Comprendre la contribution d'un trait phénotypique racinaire au rendement final dans différents scénarios et introgresser les QTL correspondant dans des variétés élites est à l'heure actuelle une priorité dans les programmes de sélection chez le riz (Ahmadi *et al.*, 2014, Zheng *et al.*, 2003, Zheng *et al.*, 2006, Qu *et al.*, 2008, Dorlodot *et al.*, 2017, Courtois *et al.*, 2009, Uga *et al.*, 2011, Uga *et al.*, 2013, Uga *et al.*, 2012).

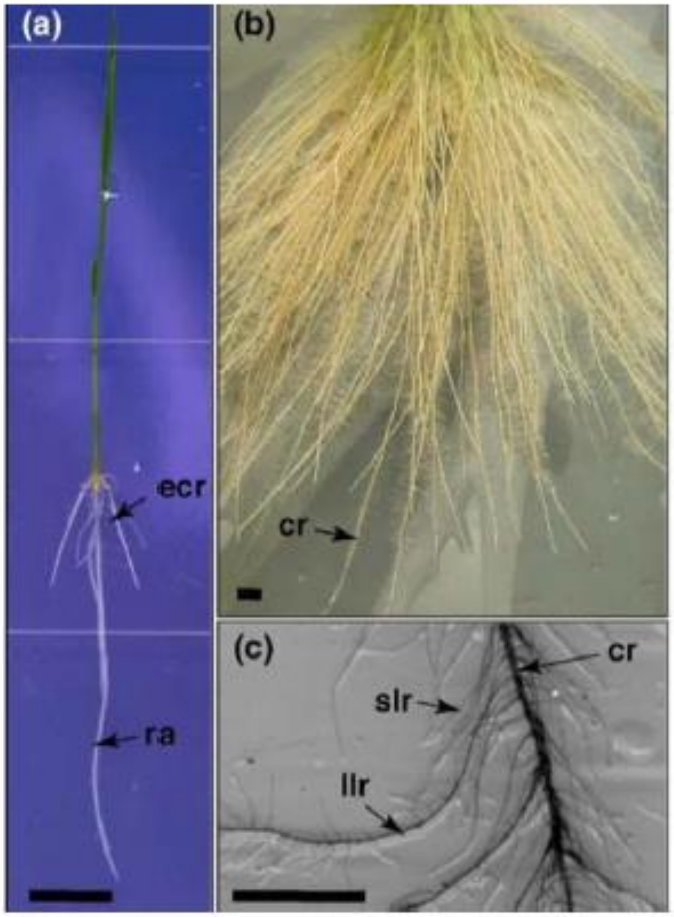


Figure 5: Architecture fasciculée de la racine de riz (Coudert et al., 2010)
A. Architecture racinaire d'une plantule d'*Oryza sativa* cv *Nipponbare* agée de 6 jours
B. Architecture racinaire d'une plantule d'*Oryza sativa* cv *Nipponbare* agée de 40 jours
C. Détail d'une racine coronaire
Barre d'échelle= 1cm. Abbréviation: ra: radicule ; ecr:racines embryonnaires coronaires; cr:racines coronaires ; slr: petites racines latérales ; llr : longue racine latérale.

Cependant, la forte plasticité phénotypique du système racinaire combinée à la difficulté de phénotypage des paramètres architecturaux racinaires, est un challenge pour son intégration dans les programmes de sélection variétale. **A ce jour, il est donc nécessaire d'acquérir des connaissances qui pourraient contribuer à la conception d'idéotypes racinaires adaptés à différentes conditions de stress. Des mécanismes d'adaptation peuvent être observés à différentes échelles, à l'échelle de l'organe à travers l'architecture racinaire, à l'échelle tissulaire à travers une modification de l'anatomie racinaire ou au niveau cellulaire avec par exemple des modifications dans la composition des parois cellulaires (Rebouillat *et al.*, 2009, Geldner *et al.*, 2013).**

2.1 Architecture racinaire, idéotypes et génétique directe

2.1.1 Architecture fasciculée des racines du riz

Un système racinaire fasciculée est la norme chez les monocotylédones. Il est caractérisé par le développement de nombreuses racines adventives, portant des racines latérales, qui a maturité, assurent la nutrition minérale et hydrique (**Figure 4**). Par contraste, le système racinaire d'*Arabidopsis thaliana,* et de la majorité des plantes dicotylédones, est un système racinaire pivotant : une racine principale d'où émergent des racines latérales (Smith et de Smet, 2012, Péret *et al.*, 2009).

Les céréales possèdent différents types de racines embryonnaire et post-embryonnaire. Chez le riz, cinq types de racines peuvent être identifiées (Ahmadi *et al.*, 2014) : la radicule et les 5 racines coronaires embryonnaires, les racines coronaires, les petites racines latérales (SLR) et les grandes racines latérales (LLR) post embryonnaires (**Figure 5**). La radicule est la première racine à émerger suivi des cinq racines coronaires embryonnaires qui émergent du nœud du coléoptile. Viennent ensuite les racines coronaires post-embryonnaires qui émergent au niveau des nœuds de la tige en une ou deux rangées. Les racines latérales comprennent les SLR et les LLR et sont présentes sur toutes les racines primaires. Les SLR ont une croissance déterminée, ne sont pas ramifiées et ne répondent pas à la gravité. Les LLR quant à elle ont une croissance indéterminée, répondent positivement à la gravité et se ramifient par l'émergence d'autres racines latérales d'ordre supérieur.

2.1.2 A la recherche d'idéotypes d'architecture racinaire

L'instabilité climatique, la nécessité d'étendre la riziculture à des terres non submergées (riziculture pluviale) et de réduire les intrants impactent grandement la pérennité des récoltes et le besoin d'augmenter la production. Une façon de minimiser cet impact est de comprendre et de manipuler les traits phénotypiques racinaires ayant un impact sur l'assimilation de l'eau et de nutriments. Il faut ensuite déterminer des idéotypes racinaires qui pourront être adaptés à différents agro-écosystèmes rizicoles. Pour le riz et les autres céréales, les recherches menées sur l'architecture racinaire tendent à développer des systèmes racinaires plus profonds afin d'optimiser la captation de minéraux et d'eau qui sont généralement plus abondants dans les couches profondes (Lynch, 2013, Arai-Sanoh *et al.*, 2015, Uga *et al.*, 2013), une problématique importante pour la riziculture pluviale (Kondo *et al.*, 2000). En revanche, Lynch a suggéré en 2011 qu'un système racinaire plus dense dans les horizons de terre peu profonds favoriserait l'extraction d'autre nutriments comme le phosphore, théorie validée au moins sur deux plantes d'intérêt agronomique que sont le soja (*Glycine max*) et le maïs (Lynch et Brown 2001, Zhu *et al.*, 2004). Il n'existe donc pas un seul idéotype racinaire « idéal » mais une gamme en fonction des agrosystèmes considérés.

2.1.3 Etude de l'architecture racinaire et tolérance au stress abiotique : approches QTL

Dans le cas de la résistance aux stress abiotiques, de nombreuses approches de génétique directe ont été menées pour identifier des QTL portant sur l'architecture racinaire (Zheng *et al.*, 2003, Zheng *et al.*, 2006, Li *et al.*, 2005, Courtois *et al.*, 2009). A ce jour, plus de 675 QTL intervenant dans l'architecture racinaire ont été identifiés chez le riz (Courtois *et al.*, 2009). Parmi les QTL identifiés, quelques-uns ont fait l'objet d'études poussées afin de les cartographier précisément et de caractériser leurs fonctions en isolant les gènes en cause. Par exemple qSOR1 (quantitative trait loci for Soil surface rooting 1) favorise un enracinement superficiel (Uga *et al.*, 2012) et Phosphate-uptake 1 (Pup1) qui promeut un développement racinaire précoce et une tolérance à la carence en phosphate (Wissuwa *et al.*, 1998). Uga *et al.*, en 2011 ont croisé 2 approches, identification de QTL et détermination d'idéotypes racinaires, ici imaginé pour améliorer

la résistance au stress hydrique *via* un enracinement plus profond, ce qui a permis le clonage positionnel de DEEP ROOTING 1 (DRO1) qui joue un rôle dans l'angle racinaire.

Les approches QTL ont donc permis la découverte de nombreux loci impliqués dans la variation de l'architecture racinaire du riz. Néanmoins, l'intervalle de temps entre la cartographie d'un QTL et la caractérisation du polymorphisme moléculaire et du gène en cause, peut être très long. Par exemple, une quinzaine d'années se sont écoulées entre la cartographie du QTL Pup1 en 1998 et l'identification du gène sous-jacent, PSTOL1 en 2012 (Phosphate-starvation tolerant 1, Gamuyao *et al.*, 2012). Plus récemment, Uga *et al.* ont réussi à identifier le gène DRO1 en deux ans seulement (Uga *et al.*, 2013). Pour la majorité des QTL détectés, comme qSOR1, les variations génétiques responsables des variations phénotypiques observées ne sont pas connues. Cela est en partie dû à la difficulté de cartographier précisément les QTL en lien avec l'architecture racinaire, pour plusieurs raisons comme les difficultés de phénotyper le système racinaire mais aussi aux différences dans la structure des génomes entre les groupes variétaux utilisés. Par exemple, PSTOL1 n'existe pas chez Nipponbare, le cultivar de référence pour lequel une séquence du génome de haute qualité est disponible. Enfin, la caractérisation fonctionnelle des QTL, et donc la compréhension des mécanismes moléculaires influençant la variabilité d'un trait racinaire d'intérêt, nécessite des travaux d'investigations par génétique inverse (Gamuyao *et al.*, 2012, Uga *et al.*, 2013). Par exemple, le gène DRO1 n'a aucune similarité avec des gènes de fonction connue (Uga *et al.*, 2013).

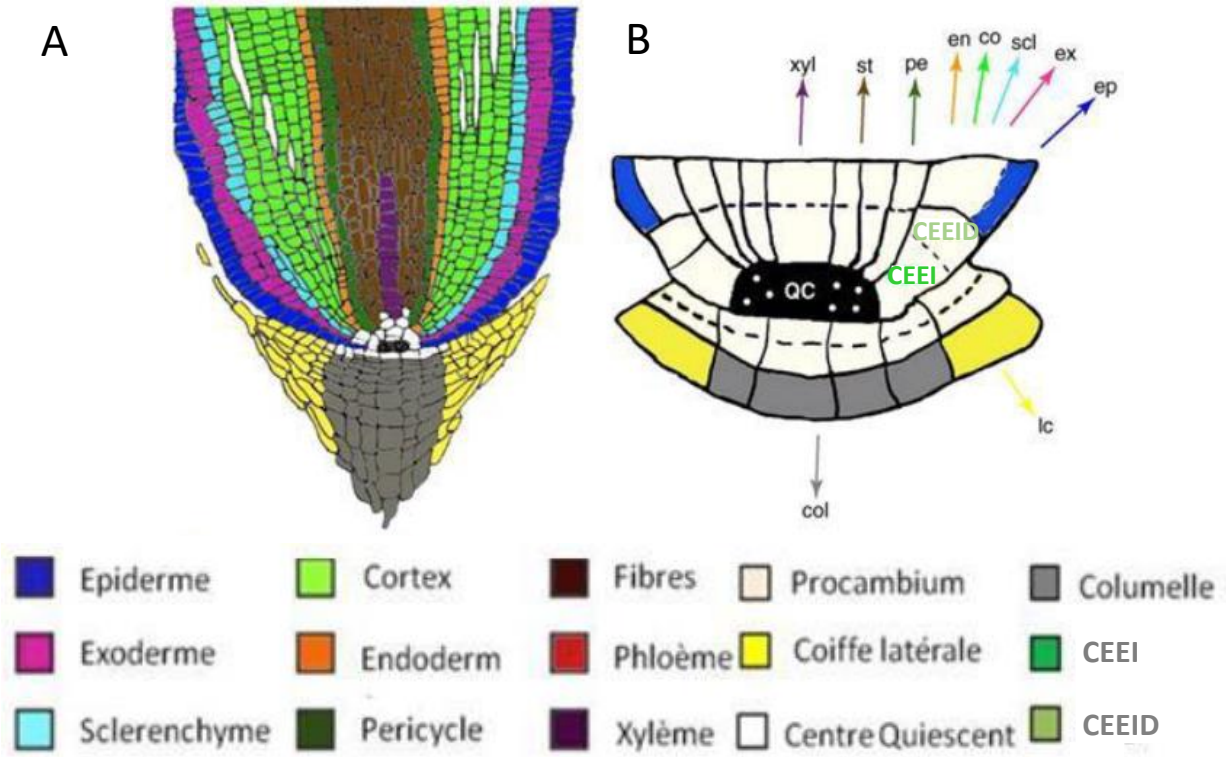


Figure 6 : Mise en place des tissus racinaires chez le riz *(*Rebouillat et al. 2009)
A : Mise en place des tissus internes au niveau du méristème racinaire (RAM) chez le riz.
B : Zoom sur le méristème racinaire du riz. La CEEI (Common Epidermis-Endodermis- initial) va subir une première division anticlinale qui donne la CEEID (CEEI Daughter). La CEEID va ensuite entamer une série de divisions périclinales asymétriques successives, qui forment l'épiderme et les tissus internes.

2.2 Anatomie racinaire, idéotypes

2.2.1 Anatomie racinaire du riz

L'anatomie racinaire radiale chez le riz est typique d'une plante semi-aquatique avec des tissus permettant la croissance racinaire en condition submergée (Rebouillat *et al.*, 2009). De l'extérieur à l'intérieur, ces tissus sont : l'épiderme, puis les tissus internes (exoderme, sclérenchyme, cortex et endoderme) et enfin le cylindre central, composé des tissus vasculaires, phloème et xylème, et du péricycle, à l'origine des racines latérales (**Figure 6A**). A un stade de développement plus avancé, les cellules corticales débutent un processus de mort cellulaire programmée (**cf chapitre 1**), ce qui a pour conséquence la création de cavités appelées aérenchymes lysogènes (**cf page 41, Figure 7A**). Le riz forme ces aérenchymes de manière constitutive comparativement à d'autres plantes cultivées comme le maïs ou le tournesol où les aérenchymes sont uniquement induits par différents stress, dont la submersion. Les différents types racinaires du riz ont une anatomie assez similaire et diffèrent essentiellement pour le nombre de couches de cellules corticales (**Figure 8**).

L'anatomie racinaire du riz et des autres céréales cultivées diffèrent de celle de la plante modèle *Arabidopsis thaliana*, notamment au niveau du nombre de couches de cellules corticales (**Figure 9**, Smith et de Smet., 2012, Hochholdinger et Zimmermann, 2008, Pauluzzi *et al.*, 2012) et de la présence de tissus externes comme l'exoderme et le sclérenchyme qui sont absents chez Arabidopsis.

2.2.2 Tissus internes et adaptation aux stress

Les tissus internes sont impliqués dans la tolérance à de nombreux stress abiotiques en servant d'interface entre l'épiderme, en contact avec l'environnement extérieur, et le cylindre central. L'endoderme et l'exoderme sont connus pour leur rôle de barrières pour les nutriments et l'eau, contrôlant ainsi leur transport vers le système vasculaire, aussi bien que son reflux du système vasculaire vers l'extérieur (Doblas *et al.*, 2017). Pour cela, ces tissus forment deux barrières imperméables composées de polymères, la lignine, qui forme la bande de Caspari, et la subérine, qui forme la lamelle

de subérine. La subérisation est régulée en réponse à des stress environnementaux (Andersen *et al.*, 2015, Kreszies *et al.*, 2018).

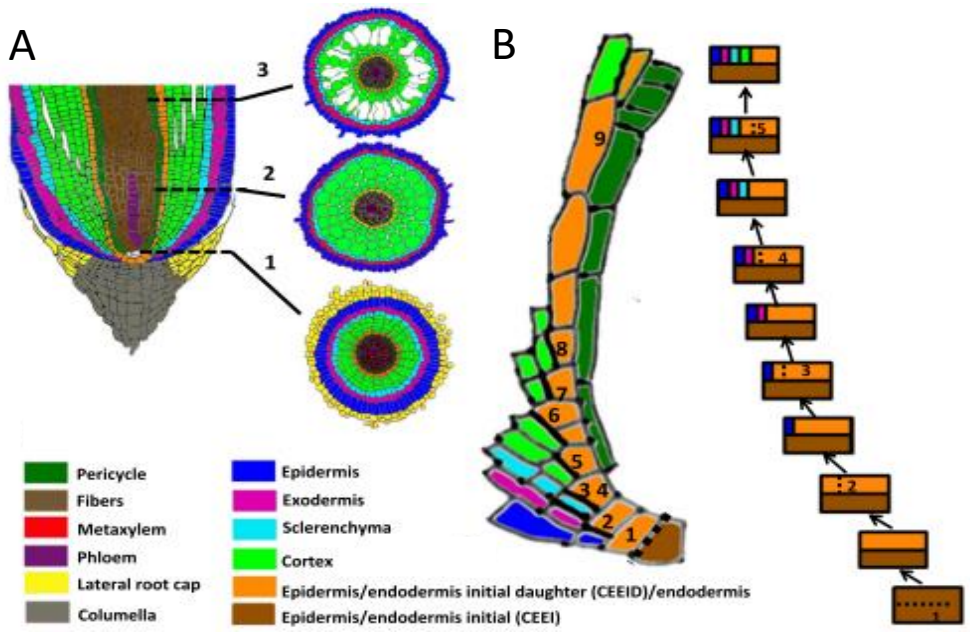


Figure 7: Anatomie racinaire du riz et profil de différentiation cellulaire des tissus internes
A. Il existe 4 zones distinctes de cellules initiales (Coudert et al., 2010, Rebouillat et al., 2009). La zone vasculaire, la zone de la columelle, la coiffe racinaire latérale, et enfin la CEEI qui contient l'épiderme, l'exoderme, le sclerenchyme, le cortex et l'endoderme. A droite 3 coupes radiales à différents stades de développement, on peut observer l'apparition d'aérenchymes dans le cortex.
B. Chez les céréales, les tissus internes (endoderme, cortex, sclérenchyme et exoderme) et l'épiderme dérivent tous d'une même cellule initiale. Celle-ci est appelée Common-Endodermis-Epidermis-Initial chez le riz (CEEI). Cette cellule initiale se divise de façon anticlinale pour donner une cellule initiale dite « fille » (CEID chez *A. thaliana* et CEEID chez le riz). Ces cellules initiales « filles » entament ensuite une succession de divisions périclinales asymétriques et forment ainsi les différents tissus internes racinaires chez le riz

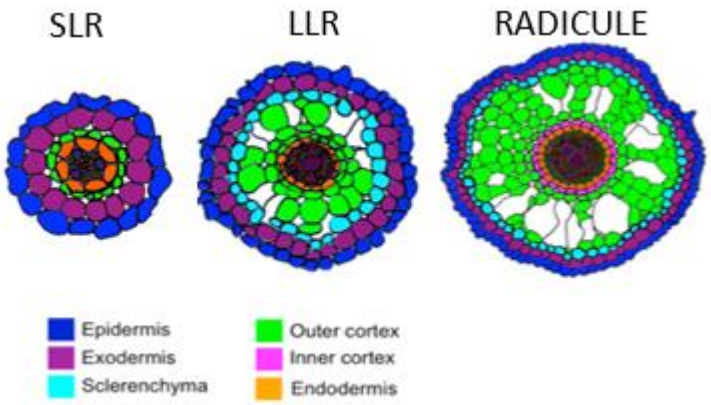


Figure 8: Anatomie comparée de trois types racinaires du riz (adapté de Pauluzzi et al., 2012)
L'anatomie des différents types racinaire du riz est assez similaire et diffèrent essentiellement pour le nombre de couches de cellules corticales. Une seule couche de cellules corticales est présente dans les petites racines latérales (SLR), qui ne forment pas d'aérenchymes.

Certaines plantes halophytes, tolérantes à de fortes concentrations en Na^+, présentent un doublement de la couche d'endoderme (Naz *et al.*, 2013, Inan *et al.*, 2004), les cadres de caspari formant une barrière imperméable aux ions dont le K^+, le Zn^{2+} et le Na^+ (Peng *et al.*, 2004, Ranathunge *et al.*, 2005, Doblas *et al.*, 2017). Un autre tissu interne clef est l'aérenchyme, tissu lacuneux formé à partir du cortex qui jouent un rôle critique dans la tolérance à la submersion.

2.2.3 Formation du cortex

Chez les angiospermes, les tissus internes (endoderme, cortex, sclérenchyme et exoderme) dérivent tous d'une même cellule initiale (Lux *et al.*, 2004, **Figure 6B et Figure 7B**). Celle-ci est appelée « Cortex Endodermis Initial » chez *A.thaliana* (CEI) et Common-Endodermis-Epidermis-Initial chez le riz (CEEI). Cependant, si chez les dicotylédones, dont *A. thaliana*, il existe une origine cellulaire commune entre la coiffe racinaire et de l'épiderme, chez les monocotylédones une seule initiale est à l'origine de l'épiderme et des tissus internes (CEEI) qui est distincte de celle de la coiffe (Clowes, 2000). Cette cellule initiale se divise de façon anticlinale pour donner une cellule initiale dite « fille » (CEID chez *A. thaliana* et CEEID chez le riz). Ces cellules initiales « filles » entament ensuite une succession de divisions périclinales asymétriques qui forment les différents tissus internes racinaires chez le riz et le cortex, multicouche (Rebouillat *et al.*, 2009, Pauluzzi *et al.*, 2012). **Les mécanismes moléculaires à l'origine de la formation et de l'organisation de ces différents tissus restent à ce jour largement méconnus** chez le riz.

Chez *A. thaliana*, un réseau de gènes, impliquant principalement SHR et SCR, deux membres de la famille de facteurs de transcription GRAS, sont impliqués dans la formation du cortex (Benfey *et al.*, 1993, Cui *et al.*, 2007, Wu *et al.*, 2014). Brièvement, le gène *SCR* est transcrit dans l'endoderme et le gène *SHR* dans la stèle. La protéine SHR est capable de migrer dans l'endoderme au niveau de la CEI, ou elle active *SCR*. *SCR* et *SHR* activent une cycline D, conduisant à la division périclinale de la CEI et la séparation en deux couches de l'initiale. SCR se lie à SHR par interaction protéine-protéine séquestrant la protéine SHR dans le noyau de la couche interne empêchant ainsi SHR de migrer dans la couche adjacente. Enfin, un peptide, CIF2 a lui aussi la capacité de migrer dans l'endoderme et conjointement avec SHR, induit la différenciation de la couche interne adjacente à la stèle en endoderme (Li *et al.*, 2018, Drapek *et al.*, 2018).

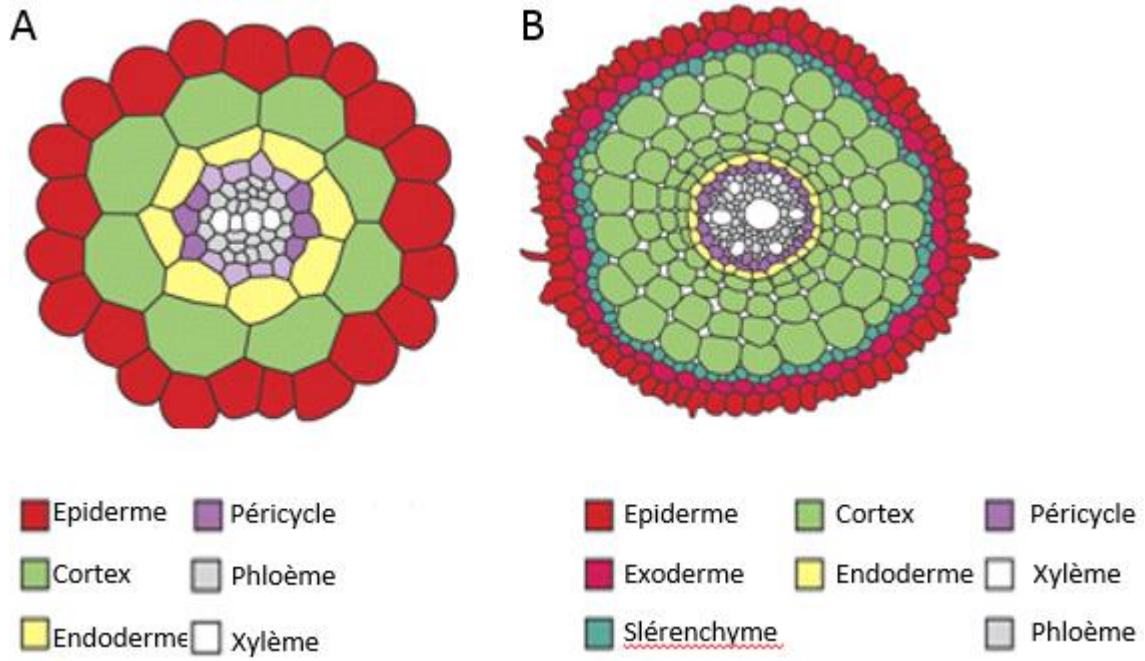


Figure 9: Anatomie racinaire comparée du riz et d'*Arabidopsis thaliana*
A. La racine d'*A. thaliana* possède une seule couche de cortex et n'a pas de sclérenchyme, ni d'exoderme
B. La racine de riz possède un nombre variable de couche de cellules corticales et possède un sclérenchyme et un exoderme.

Contrairement à *A. thaliana*, le riz forme un cortex multicouche qui comprend d'une à plus de 10 couches en fonction du type de racine (**Figures 8 et 9**). Il possède deux orthologues aux gènes *SHR* et *SCR* et leur fonction dans la mise en place du cortex n'a été que partiellement élucidé. La complémentation d'un mutant perte de fonction pour SHR chez *A. thaliana* a néanmoins permis de montrer qu'il existait une conservation partielle de fonction. La complémentation par les gènes OsSHR1 et OsSHR2 du mutant shr restaure un endoderme fonctionnel mais comparativement à AtSHR, induit la formation de couches surnuméraires, jusqu'à six pour OsSHR2 (Wu *et al.*, 2014). Les protéines OsSHR2 et OsSHR1 sont également plus mobiles que AtSHR et pouvaient migrer dans 5 couches de tissu interne pour une seule pour AtSHR, et induire des couches de cortex surnuméraires. La surexpression des gènes *SHR* chez le riz induit également la formation de couches de cortex supplémentaires, jusqu'à plus de 60 couches de cortex chez cette espèce et la localisation de la protéine OsSHR2 n'est pas limité à l'endoderme puisqu'elle est présente dans toutes les couches de tissus externes, cortex, exoderme et sclérenchyme compris (Henry *et al.*, 2017). Enfin, les protéines SHR et SCR chez le riz interagissent entre elles, exactement comme pour les protéines SHR et SCR chez *A. thaliana* (Wu *et al.*, 2014). Les premières analyses de mutant perte de fonction pour OsSHR2 et OsSHR1 ont permis d'établir qu'ils n'avaient pas d'anomalies au niveau de la formation des tissus internes, suggérant une possible redondance fonctionnelle qui reste à déterminer (résultats non publiés).

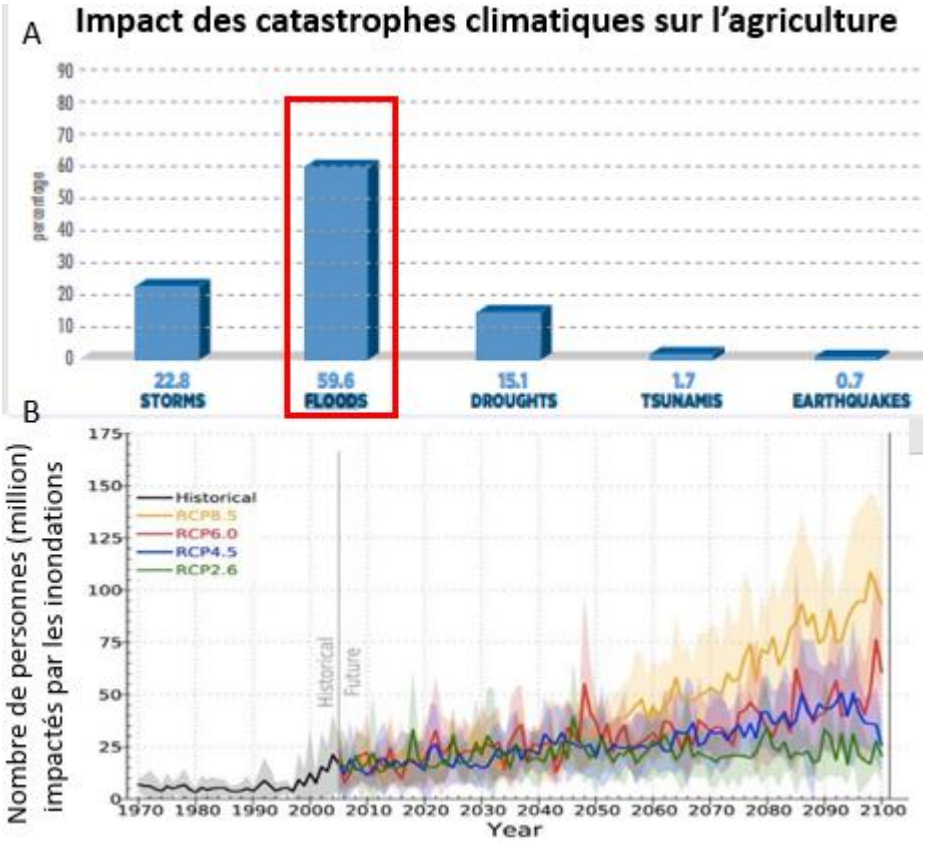


Figure 10: Dommages liés aux catastrophes climatiques dans le monde, zoom sur les inondations. D'après des données de la FAO exploités par Bhushan *et al.* en 2016.
A: Pourcentage des dommages et pertes liés aux catastrophes climatiques. On peut constater que les dommages liés aux inondations représentent 60% des pertes liés aux catastrophes climatiques dans l'agriculture, loin devant la sécheresse.
B: Observations et prédictions du nombre de personnes touchées par les inondations entre 1970 et 2100 d'après différents modèles climatiques

2.3 Aérenchymes, des enjeux agronomiques et écologiques

2.3.1 La submersion, un problème affectant la sécurité alimentaire

L'eau est un milieu dans lequel la diffusion de l'oxygène est 10 000 fois plus faible que dans l'air (Armstrong, 1980, Colmer et Flowers, 2008). Ainsi, les plantes submergées font face à un stress qui impacte directement deux processus biologiques essentiels, la photosynthèse et la respiration, affectant leur croissance, et le rendement (Nishiuchi *et al.*, 2012). D'après une étude menée de 2003 à 2013 sur des données de la FAO, les inondations des cultures sont responsables de près de 60% des pertes de rendement dues aux catastrophes climatiques (**Figure 10A**, Bhushan *et al.*, 2016), loin devant la sécheresse ou les cyclones. De plus, cette tendance va se renforcer dans les prochaines années selon les prévisions climatiques (**Figure 10B**), augmentant la fréquence des stress de submersion temporaires.

2.3.2 Les aérenchymes et la tolérance à la submersion

La tolérance à la submersion est très variable chez les plantes et résulte de nombreux mécanismes impliquant l'anatomie, l'architecture et le métabolisme (Voesenek *et al.*, 2006, Loreti *et al.*, 2016, Pedersen *et al.*, 2017). Parmi ces mécanismes, les aérenchymes se forment chez de nombreuses espèces cultivées en condition de saturation des sols en eau (Mounier et al, soumis, cf chapitre 1). Les aérenchymes sont des tissus lacuneux qui se différencient à partir du cortex (**Figure 11**). Ils connectent de manière longitudinale l'environnement « gazeux » des parties aériennes à celui des parties racinaires. Il en existe trois grands types, lysogène, schizogène et expansigène (Seago *et al.*, 2005, **Figure 11**). Les aérenchymes schizogènes ont pour origine la séparation des parois cellulaires des cellules corticales. Les aérenchymes expansigènes sont formés par division et élongation différentielle des cellules corticales. Et enfin les aérenchymes lysogènes, présents constitutivement chez le riz, sont formés par la mort cellulaire de cellules du cortex.

Parmi les poacées, certaines espèces forment des aérenchymes constitutivement et en condition de submergence (e.g le riz), ou uniquement en conditions de submergence (e.g blé, maïs, sorgho) ou enfin, ne forment jamais d'aérenchymes (e.g *Brachypodium distachyon*, Shiono et Yamada, 2014, Yamauchi *et al.*, 2013).

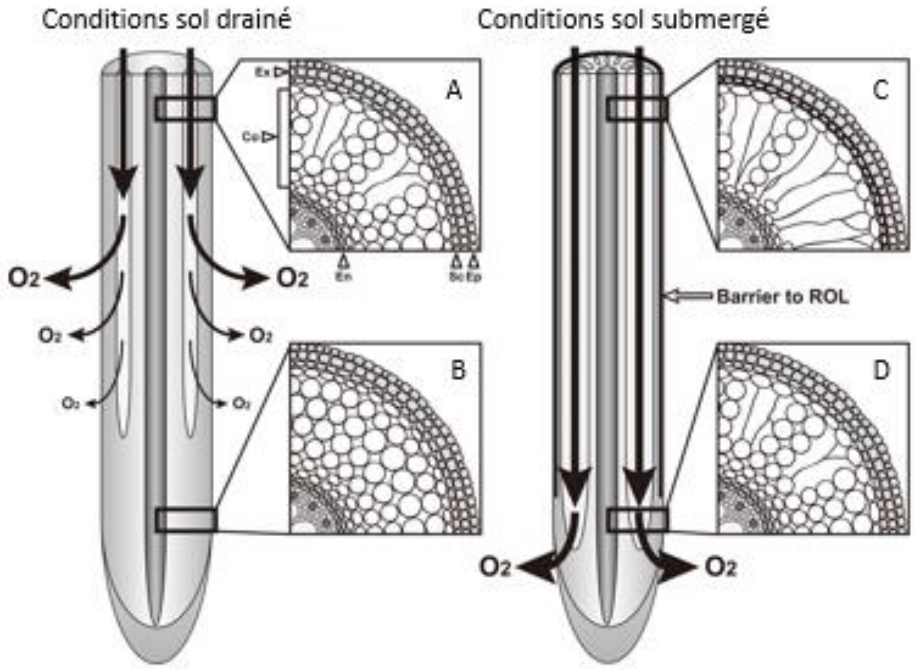


Figure 12: Formation des aérenchymes et de diffusion du dioxygène en condition de submersion ou pluviale (drainé), Yamauchi *et al*., 2013

En condition pluviale, les aérenchymes sont formés de manière constitutive chez le riz, mais il n'y a pas formation de barrières contre la perte radiale d'oxygène par diffusion (ROL, Yamauchi et al., 2018). En condition saturée en eau, la formation d'aérenchymes est stimulée. Une barrière contre le ROL se met également en place, permettant l'approvisionnement en oxygène de l'apex racinaire.

A. En condition drainée, les aérenchymes se forment à distance de l'apex racinaire

B. En condition de submersion, la formation des aérenchymes est plus précoce.

L'épaisseur des flèches reflète la quantité d'oxygène diffusant. Ep, epiderme; Ex, endoderme; Sc, sclérenchyme; Co, cortex; En, endodermis.

Classiquement, les aérenchymes sont présentés comme un mécanisme d'adaptation à la privation d'oxygène dans un sol saturé en eau ou lors d'un stress de submersion (Nishiuchi *et al.*, 2012, Mounier *et al.*, soumis).

2.3.3 Les tissus externes forment une arrière à la diffusion radiale d'oxygène

Dans les aérenchymes, l'oxygène diffuse verticalement vers les méristèmes racinaires, mais également radialement (Radial Oxygen Lost, ROL) à travers les tissus adjacents, conduisant à une diminution rapide de la quantité d'oxygène disponible pour le méristème racinaire (**Figure 12**). Les tissus adjacents au cortex, et plus particulièrement l'exoderme accumule plus précocement de la subérine dans ses parois cellulaires en condition submergée formant une barrière au ROL ou ROB (*Radial Oxygen Barrier*, Nishiuchi *et al.*, 2012, Kulichikhin *et al.*, 2014, Yamauchi *et al.*, 2018, **Figure 13**). Enfin, la stimulation de la lignification des parois cellulaires en condition submergée, principalement celle du sclérenchyme, permet de maintenir la rigidité des racines malgré la formation de cavités aérenchymateuses et contribuer à une meilleure tolérance aux contraintes mécaniques.

Ces deux adaptations anatomiques, sont fréquemment retrouvées chez de nombreuses plantes possédant des aérenchymes (Striker et al., 2007). Si la formation des aérenchymes est induite chez d'autres poacées cultivées comme le maïs, le blé, ou le sorgho on observe en revanche aucune formation de barrières aux ROL chez ces espèces (Yamauchi *et al.*, 2013, Watanabe et al., 2017). **Ces trois mécanismes, formation d'aérenchymes, formation d'une barrière à la diffusion radiale d'oxygène et ceinture rigide de tolérance à la compression mécanique latérale, représentent une adaptation anatomique complète du système racinaire à la submersion chez le riz.**

2.3.4 Les aérenchymes racinaires un trait d'intérêt agronomique, au-delà de la tolérance à la submersion

Si le rôle clef des aérenchymes dans la tolérance à la submersion est connu, leur présence a également un impact sur la tolérance à d'autres stress (Mounier *et al.*, soumis, cf chapitre 1).

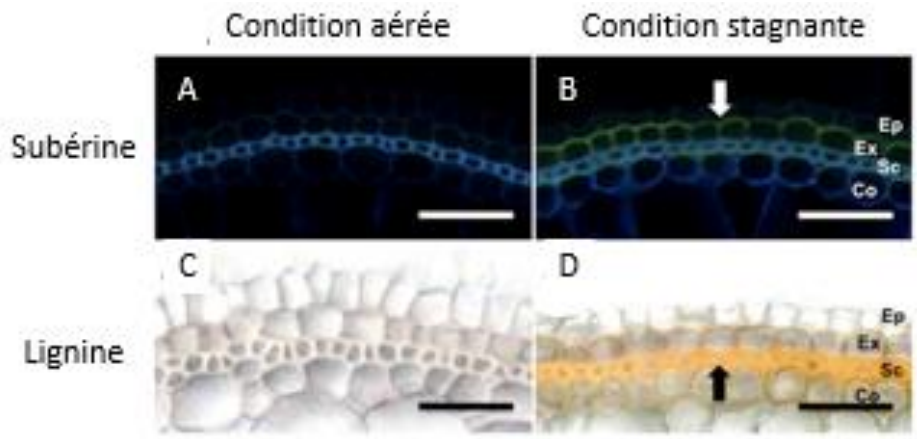


Figure 13: Coloration de la lignine et de la subérine dans les cellules externes de racines de riz âgée de 14 jours, de plantes en condition stagnantes ou aérées (Nishiushi et al., 2012).

A et B Coloration de la subérine au jaune de Fluorol et observation au microscope UV. La subérine n'est visible qu'en condition stagnante au niveau de l'exoderme (flèche blanche).

C et D Coloration de la lignine contenant des groupements cinnamyles aldèhydes au phluroglucinol hydrochloride. La lignine n'est visible dans le sclérenchyme qu'en condition stagnante (flèche noire).

Barre d'échelle: 50µm. Légende : Ex: Exoderme; Sc:Sclérenchyme; Ex: Exoderme; Ep: Epderme.

Ainsi, la présence d'aérenchymes réduit le coût métabolique de l'exploration racinaire, *via* une réduction du nombre de cellules vivantes pour un même diamètre, une même longueur racinaire et surtout un même volume de sol exploré (Lynch *et al.*, 2015). Il existe une corrélation positive entre le rendement et le taux de RCA (Root Cortex Aerenchyma) dans des conditions de stress hydrique chez le maïs (Zhu *et al.*, 2010), et entre rendement et RCA en conditions de carence en azote ou en phosphate chez le riz (Pujol et Wissuwa, 2018) et le maïs (Saengwilai *et al.*, 2014, Galindo-Castaneda *et al.*, 2018). La formation d'aérenchymes est par ailleurs induite par des carences nutritionnelles chez les céréales (Abiko et Obara., 2014, Yamauchi *et al.*, 2013). A contrario, en condition favorable, la présence d'aérenchymes impacte négativement la diffusion radiale de nutriments chez le maïs (Hu *et al.*, 2014) et la croissance. La plus faible proportion de cellules vivantes limite le transfert radial d'eau et de minéraux vers les tissus vasculaires impactant le rendement en conditions non limitantes. **Il semble donc il y ait une corrélation forte entre nutrition, rendement, et taux de RCA et la présence d'aérenchymes est un caractère favorable en conditions environnementales défavorables.**

Par ailleurs, il a été démontré que 90% du méthane produit dans les rizières irrigués ou inondés circulait *via* les aérenchymes (Wang *et al.*, 2017). La présence d'aérenchymes a néanmoins un effet favorable sur la production de méthane grâce à la mise en place de microbiomes aérobies spécifiques *via* une plus grande diffusion radiale et longitudinale de l'oxygène au niveau racinaire (Li *et al.*, 2008). Ces microbiomes aérobies interviennent dans des processus de conversion de NH_4^+ en NO_3^- par des bactéries nitrifiantes, ce qui conduit à la baisse de la production de méthane et augmente l'efficacité d'utilisation de l'azote (NUE) par la plante (Li *et al.*, 2008, Kludze *et al.*, 1993). **La formation des aérenchymes et/ou d'une barrière à la diffusion radiale de l'oxygène influence donc la production rizicole de méthane en condition de submersion** (Cai *et al.*, 2000, Yan *et al.*, 2005).

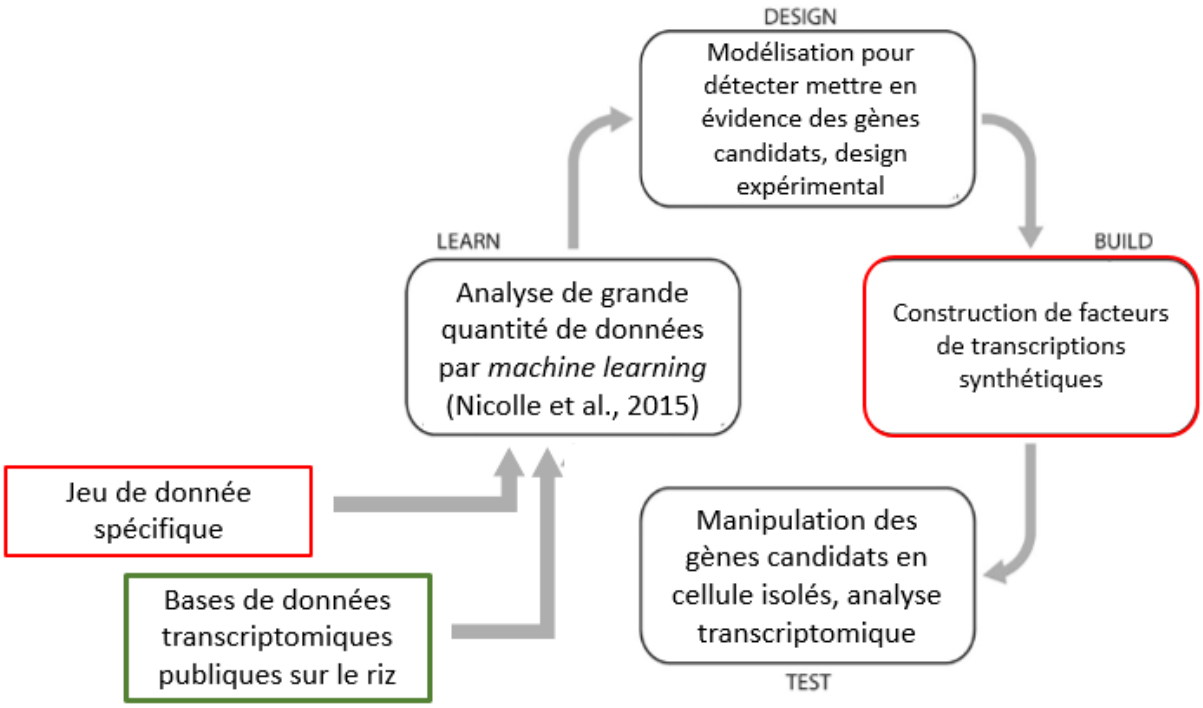


Figure 14: Représentation schématique de l'approche DBTL

Le principe du cycle DBTL est d'affiner des connaissances sur un processus biologique de manière itérative. En quelques mots, le cycle peut se résumer en 4 étapes 1) « Design », choisir une action à mener sur des gènes candidats en se basant sur une modélisation *in silico* d'un processus biologique à partir de bases de données -omiques, 2) « Build », construire des outils moléculaires permettant de manipuler ces gènes, 3) « Test », mener une expérimentation en système biologique et en extraire des informations –omiques 4) « Learn », utiliser des algorithmes de machine learning afin d'analyser les données et de ré-affiner la partie « Design ».

Néanmoins, pour commencer le cycle d'apprentissage itératif, il est nécessaire de disposer d'un grand nombre de jeux de données transcriptomiques (encadré vert) et d'un jeu de données spécifique à la question biologique étudiée (encadré rouge, Nicolle et al., 2015). L'un de mes objectifs est de créer ce jeu de donnée spécifique à l'étude de la formation des aérenchymes. Mon deuxième objectif est le développement d'un système expérimental en cellule isolée avec des facteurs de transcriptions synthétiques basés sur la technologie CRISPR/Cas. Les autres étapes du cycle feront l'objet d'une future collaboration avec l'équipe I3Bionet et ne seront abordés que succinctement ici.

3. Le cycle Design-Build-Test-Learn (DBTL) pour l'identification du réseau de gènes intervenant dans la mise en place des aérenchymes

3.1 Le cycle DBTL, utiliser des séries de tests itératifs pour affiner des hypothèses

Ces dernières années, une nouvelle approche basée sur un *workflow* appelé cycle DBTL (Design-Build-Test-Learn) a été progressivment adoptée par des biologistes pour le décryptage de mécanismes moléculaires (Figure 14, Durán *et al.*, 2016, Liu *et al.*, 2015). Mon travail de thèse est une première contribution au développement d'une stratégie DBTL chez le riz. La stratégie proposée se décompose en 3 phases. Une première phase consiste à l'identification des gènes candidats pour la formation *via* une modélisation des réseaux de gènes impliqués dans des processus biologiques spécifiques («Design») ; dans une deuxième phase il s'agit de construire des outils moléculaires adaptés afin de valider les candidats identifiés (« Build-Test ») ; enfin, les informations –omiques issues de ces expérimentations seront utilisées pour affiner les réseaux (« Learn »), grâce aux approches *machine learning*, dans une troisième phase. Ainsi, ce *worflow* forme un cycle d'apprentissage itératif : tester des hypothèses qui serviront à affiner les hypothèses suivantes.

Dans notre cas, la première phase consistait à établir un transcriptome spécifique du cortex et plus particulièrement de la différenciation du cortex en aérenchymes. Nous avons donc choisi de prélever des phases critiques du développement du cortex (Kawai *et al.*, 1998) afin de réaliser un transcriptome tissu et stade spécifique. La deuxième phase a consisté à développer un système de validation de gènes candidats en cellules isolées (protoplastes). La reconstruction de ces réseaux fait l'objet d'une collaboration avec I3Bionet, une équipe de recherche en bio-informatique qui dispose déjà d'outils dédiés à l'approche DBTL (Nicolle *et al.*, 2015, Chebil *et al.*, 2014, Elati *et al.*, 2007). Ces outils ont déjà été exploités expérimentalement avec succès sur *Yarrowia lipolytica* pour modifier le métabolisme lipidique de manière dirigée (Trébulle *et al.*, 2017). Ces travaux sont en cours, les données de transcriptome n'ont été obtenus qu'en fin de thèse, et ne seront pas présentés dans ce manuscrit.

3.3 Objectifs de la thèse

L'objectif de ce travail de thèse consistait à contribuer à l'élucidation des mécanismes moléculaires impliqués dans la formation des aérenchymes. Dans un premier temps, j'ai réalisé une synthèse bibliographique de l'état de l'art sur la formation des aérenchymes chez le riz et les céréales et l'impact agronomique au sens large de la présence d'aérenchymes. Cet article, « *Root aerenchyma, this is more than just holes »* a été soumis à Trends in Plant Science (*cf* Chapitre 1).

Dans un deuxième temps, j'ai développé les outils pour permettre de réaliser un transcriptome temporel de formation des aérenchymes (chapitres 2 et 3). Le chapitre 2, décrit le développement d'un protocole de laser microdissection (LM) adapté à la réalisation de transcriptomes de tissus racinaires chez le riz (cf chapitre 2). L'objectif était d'obtenir une qualité et une quantité d'ARN suffisante pour réaliser des banques RNAseq à partir de quantités extrêmement faibles d'ARN. J'ai combiné ce protocole avec une nouvelle technologie, la digital droplet PCR ou ddPCR qui m'a permis de réaliser des études d'expression tissu-spécifique et de vérifier l'absence de contamination inter-couches lors des dissections laser. L'ensemble de ce travail a fait l'objet d'un article soumis à BMC Cellular and Molecular Biology. J'ai enfin démontré que nous avions atteint les limites de l'appareil de microdissection laser grâce à ce protocole en tentant de découper des tissus monocouches.

Dans le chapitre 3, le protocole de LM a été utilisé pour isoler trois tissus et trois stades développementaux, avec 4 répétitions biologiques et développer un transcriptome différentiel de formation des aérenchymes. Un kit spécifiquement dédié à la création de banques d'ADN complémentaires (ADNc) à partir de quantités extrêmement faibles d'ARN de riz a été développé en collaboration avec Nugen. Ces banques sont compatibles avec les plateformes de séquençage illumina et ont été optimisées lors de leur construction pour limiter les biais de séquençages et le nombre de lectures nécessaire à un séquençage profond. Je décrirais le développement du kit, des banques et les tests effectués. Les banques, pour un total de 36, ont été envoyées pour être séquencées mais je n'ai pu réaliser leur analyse dans le cadre de ce manuscrit de thèse.

Le chapitre 4 va être axé sur le développement d'outils moléculaires permettant la mise en place de l'étape « Test » de la stratégie DBTL sur le riz. Ici notre système expérimental sera basé sur la modulation de l'expression de gènes candidats en protoplastes. J'ai développé un système de manipulation d'expression orthogonale dérivé de la technologie CRISPR/Cas permettant en même temps de sur-exprimer et de sous-exprimer plusieurs gènes simultanément. J'ai démontré que ce système était fonctionnel en utilisant un gène rapporteur en protoplastes chez le riz.

Chapitre 1

Article soumis pour publication dans Trends in Plant Science: *Root aerenchyma: this is more than just holes*

Chapitre 1, article soumis pour publication dans Trends in Plant Science: Root aerenchyma: this is more than just holes

Root aerenchyma: this is more than just holes

AUTHORS NAMES

Thibault Mounier, Anne Dievart, Fabrice Varoquaux, Emmanuel Guiderdoni, Christophe Périn

Affiliations:

1 CIRAD, UMR-AGAP, Université de Montpellier, Avenue Agropolis, F-34398 Montpellier Cedex 5,

France

2 Université de Montpellier, Cirad-Inra-Montpellier SupAgro, 34000 Montpellier, France

Corresponding author: Christophe Périn: perin@cirad.fr

Aerenchyma, Cortex, Programmed-Cell death, Waterlogging, Rice

ABSTRACT

Root aerenchymas are air-containing cavities allowing the diffusion of oxygen from the aerial to the root parts, ensuring the survival of many aquatic plants in submerged conditions. Recent developments have partially elucidated the physiological and molecular mechanisms underlying the formation of aerenchymas in rice and maize which results from an interplay between ethylene, radical oxygen species (ROS) and PCD. Varieties with a high proportion of root aerenchyma have a higher yield under conditions of nutritional deficiency or water stress. In this article, we propose to outline recent developments and suggest ways to improve knowledge on the formation of aerenchymas that can be used to develop cereals that are better tolerant to submergence, nutritional deficiencies and water stress.

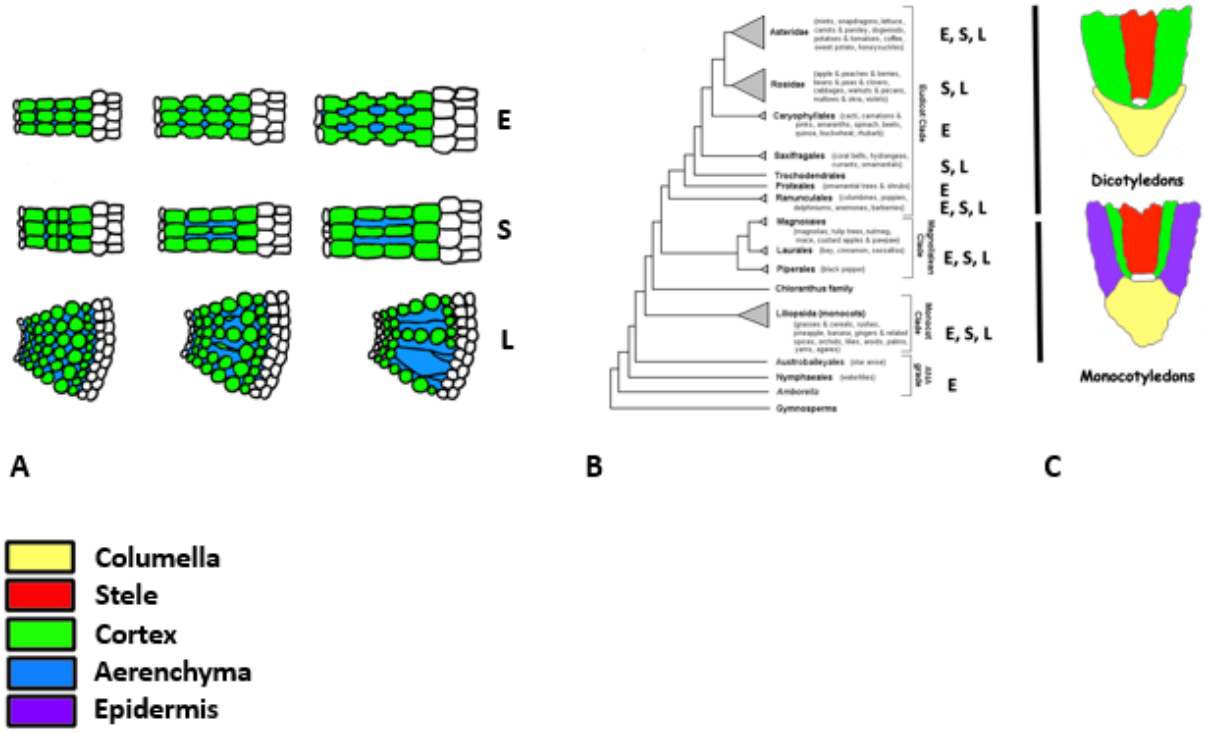


Figure 1: Types of aerenchymas and their distribution in the different groups of flowering plants. Figure 1A. For simplicity, only the three main types of aerenchyma are shown here: expansigenous (E), where aerenchymas are formed by differential cell division and elongation; schizogenous (S), where aerenchymas are formed by the separation of cell walls; and lysogenous (L), where aerenchymas result from programmed cell death (PCD). Figure 1B. Distribution of the different types of aerenchymas in the flowering plant groups. Figure 1C. Formation of the cortex, the tissue differentiating into aerenchymas, in monocotyledons and dicotyledons. A common initial epidermis-endodermis-cortex is responsible for cortex formation in monocotyledons, while an independent initial endodermis-cortex is at the origin of the cortex in dicotyledons. The epidermis is formed from a common initial ring of cells with a lateral root cap in dicotyledons. The colors are labeled in the figure. Adapted from (Seago *et al.* 2005).

MAIN TEXT

Aerenchyma is one adaptation mechanism to flooding in plants

Floods have become more frequent and more severe in recent years in many parts of the world, a trend that is linked to ongoing climate change [1]. By definition, a submerged or flooded plant is a plant that is partially or entirely covered by water [2]. Waterlogging is defined as a condition in which the soil contains an excessive amount of water and is also called soil flooding [2]. Beyond their obvious impact on the affected human populations, flood periods are responsible for significant crop losses, with most cereals being sensitive or even very sensitive to flooding; 60% of the total world crop was lost due to flooding in 2003–2013 [1]. Water is a medium through which gases diffuse very slowly. Submersion very quickly leads to the blocking of two essential biological processes, photosynthesis and respiration, which in turn blocks growth and ultimately biomass production and affects yields [2, 3]. **Survival to flooding is highly variable in the plant kingdom and includes various survival mechanisms involving anatomical, architectural and metabolic adaptations** [4] [5]. In addition to these specific mechanisms, a broader mechanism of flooding tolerance, important for waterlogging and relying on anatomical adaptations termed aerenchyma that are typical of aquatic plants, is present in rice [2].

Aerenchyma is a broad anatomical adaptation to waterlogging in flowering plants

Aerenchymas are air-filled cavities that connect the aerial parts of plants and roots, allowing plants to maintain their respiration and growth under submerged conditions [6]. To simplify, there are three main types of aerenchymas: schizogeny, lysogeny and expansigeny. Schizogenous aerenchyma are formed in the cortex by separation of cell walls. Lysogenous aerenchymas are formed by cell death. A third class of aerenchyma is called expansigenous and results from a combination of cell division and cell elongation [6] (**Figure 1A**). Aerenchymas are found in most flowering plant groups as well as basal angiosperms and are invariably associated with wetland species, confirming the tight link between aerenchyma presence and flooding tolerance [6] (**Figure 1B**). Aerenchymas allow the transport of oxygen to the root-growing areas, the root meristems (**Figure 2A**). Aside this beneficial feature, the presence of aerenchymas poses two constraints.

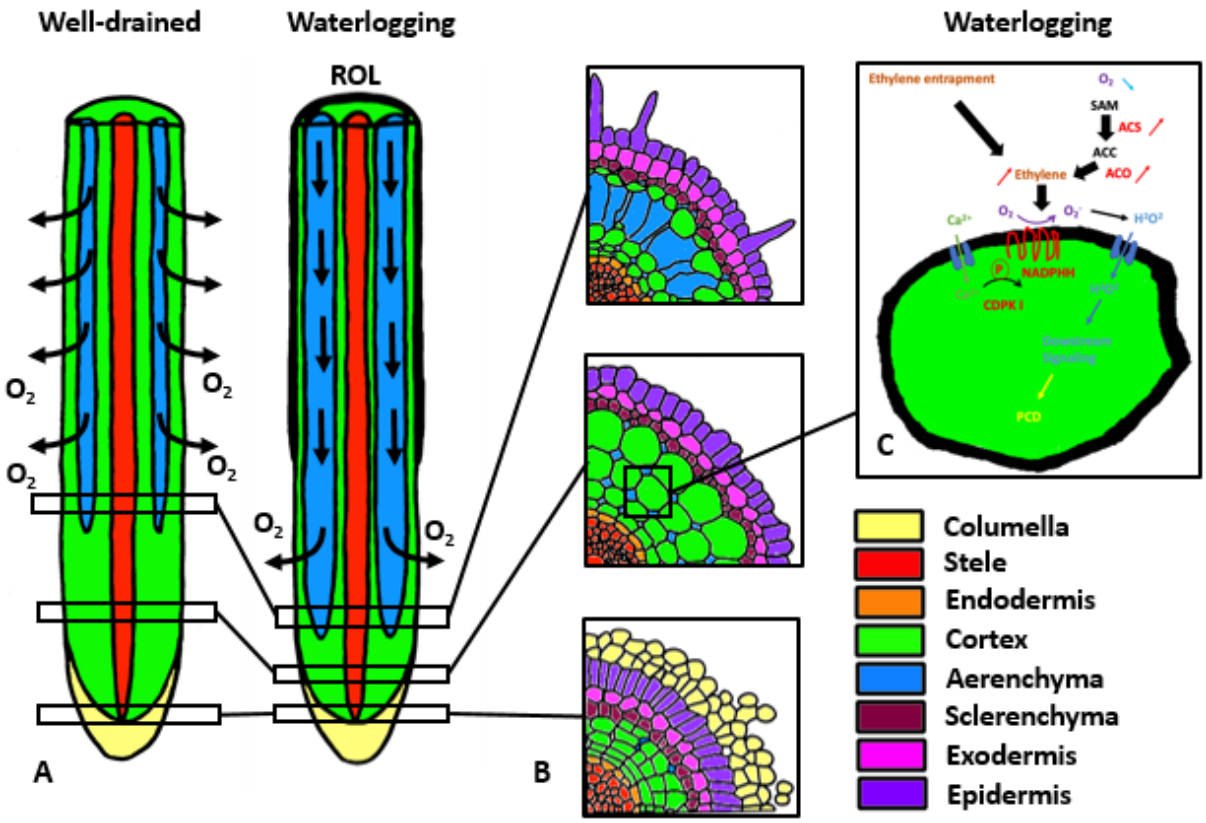


Figure 2: Schematic view of the presence of aerenchymas in rice and the molecular model of aerenchyma formation under waterlogging. Figure 2A. Left, schematic longitudinal section of a rice root under aerated conditions; right, the same but under water logging conditions. Note the earlier appearance, closer to the meristematic tip, of the blue aerenchymas and the formation of a root oxygen loss barrier under waterlogging conditions. Figure 2B. Radial sections showing three stages of aerenchyma formation. From bottom to top, area of cortex division and formation, area of cortex elongation and priming of mid-cortex cells, and area where aerenchymas are formed and mature. Figure 2C. Schematic molecular model of programmed cell death (PCD) of a cortex cell under waterlogging conditions. The red upward arrows represent an increase in transcription (ACS, ACO, CDPK) or accumulation (ethylene, H_2O_2); the blue downward arrows represent a decrease in concentration (O_2). The colors are labeled in the figure. Under waterlogging conditions, the O_2 concentration decreases and induces an increase in ethylene synthesis by stimulating the expression of the ethylene biosynthesis genes ACS and ACO. The ethylene concentration also increases by ethylene entrapment during waterlogging. Ethylene stimulates the expression of the NADPHH gene. CDPK I and Ca2+ entry induce phosphorylation of and activation of NADPHH. Active NADPHH produces ROS (O_2-) and H_2O_2. H2O2 cortex cell entry induces downstream signaling by an unknown mechanism and stimulates PCD and ultimately aerenchyma formation.

. First, Oxygen from the aerial parts diffuses vertically towards the meristems as well as radially, which leads to a decrease in oxygen concentration along the root during root growth. Therefore, many aquatic plants have developed barriers to lateral oxygen diffusion. In rice, external tissues such as the exodermis, which has a structure similar to that of the endodermis, form a barrier to the radial diffusion of oxygen (radial oxygen loss barrier or ROL barrier), which is further enhanced by the early accumulation of suberin under flooding conditions [2, 7] (**Figure 2A**). Second, the formation of aerenchyma also leads to a decrease in mechanical strength during root soil penetration. In most submergence-tolerant species that form aerenchymas, an association with a compensation mechanism to maintain mechanical strength by the formation of a ring of cell layers outside the cortex exists [8, 9]. These mechanisms that affect the outer tissues are also reinforced under submergence conditions in rice [7] and, with the formation of aerenchyma, represent a complete anatomical adaptation mechanism to flooding.

Root cortex aerenchymas reduce the metabolic cost for soil exploration

A promising way to create varieties with enhancer water, nitrogen and phosphate use efficiencies is to reduce respiration and metabolic cost for soil exploration of the root system [10]. The presence of root aerenchymas, reduces the cost of soil exploration by reducing the number of living cells per length unit of root. For instance, under water stress conditions, maize varieties with high levels of root cortex aerenchyma (RCA) have a deeper root system and an 8 times higher yield than varieties with low RCA [11]. Under low-N conditions, varieties with high RCA have a yield 58% higher than controls [12], and low N increases the root aerenchyma area [13]. In rice, P deficiency induces the earlier formation of aerenchymas, and varieties that induce the formation of aerenchymas under P deficiency conditions are more tolerant to that constraint than varieties that do not [14]. Maize cultivars with low RCA also have a better grain yield under low phosphorus availability than do other corn cultivars [15, 16], confirming a possible role of aerenchymas in adapting to nutritional deficiencies. However, the presence of aerenchymas also has a negative impact on some other agronomical traits. Under favorable nutritional conditions, the presence of aerenchymas reduces the radial transport of nutrients and water, which also reduces yield [16, 17].

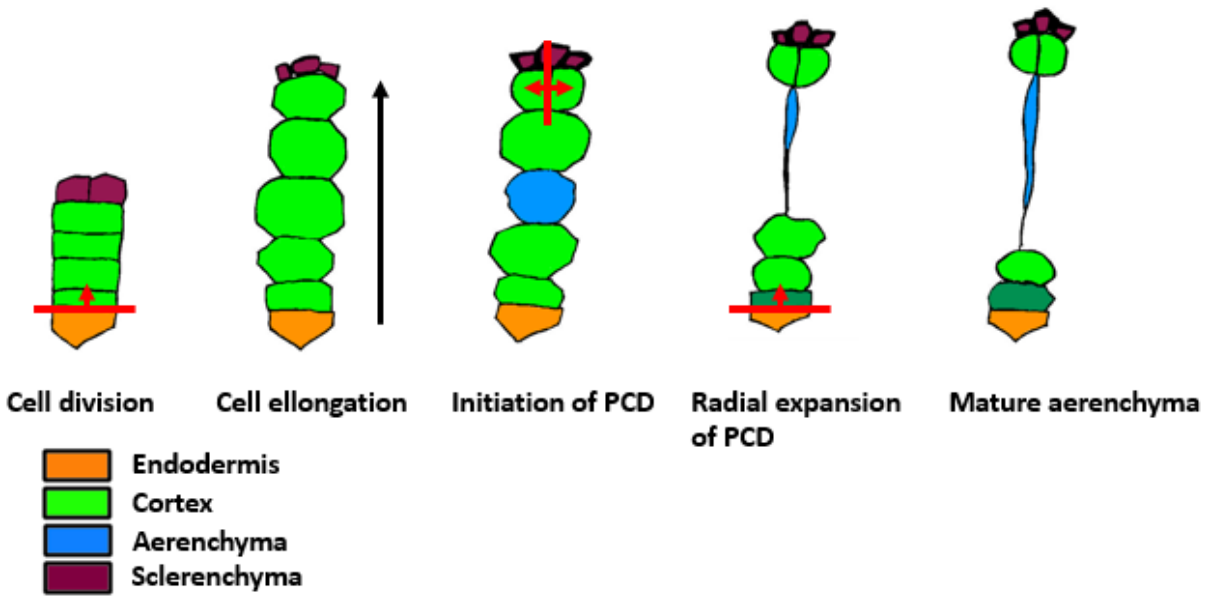


Figure 3: Main stages of aerenchyma formation during root development in rice (Adapted from Kawai et al. 1998). The red arrows indicate cellular divisions, and the black arrows indicate cellular elongation. From left to right, cell division and cortex formation phase, mid-cortex cell elongation and priming phase, PCD initiation phase, PCD radial expansion phase and mature aerenchyma.

The presence or absence of root aerenchymas seems to have a considerable impact on many agronomic traits, and this impact, positive or negative, depends on the agrosystem in question. The development of varieties with contrasting RCA levels is therefore a clear breeding goal for creating cereals adapted to submergence conditions, low nutrient inputs or water stress conditions.

Rhizosphere and its impact on aerenchyma formation

Many so-called endophytic bacteria have the ability to enter roots (for a review, see [19]) and bacteria are able to produce hormones (reviewed in [20]) or affect the production or signal transduction of hormones potentially involved in the formation of aerenchyma. Some bacteria have ACC deaminase activity; removing ACC from the plant leads to a decrease in the plant's ability to produce ethylene [21]. Inoculation of rice with *Phomopsis liquidambari* induces an increase in the amount of ethylene and auxin in the roots as well as the formation of more aerenchyma [22]. This work suggests that it is possible to modify the aerenchyma of a plant *via* inoculation with a microorganism. Taxa and diversity of microorganisms found inside the roots varies according to the plant species. Thus, there are more *Azoarcus spp.* (diazotrophic endophytic) bacteria in wild rice species and land races than in modern cultivars [23]. This finding also opens the possibility of improving the plants in such a way that recruits more endophytes to modify the amount of aerenchyma.

Origin and formation of lysogenic aerenchymas in rice

Aerenchymas are formed mostly from the root cortex. T**he formation of aerenchyma can be subdivided into four phases in rice (Figure 3)** [24] and can be classified as a radial lysogenous aerenchyma [6]. In the first phase inside the root meristem, called cortex formation, the endodermis-cortex stem cell divides several times to form the outer cortex and a layer of endodermis (**Figure 3**). The elongation phase follows the initial phase when the median cortical cells elongate radially. In addition to cell elongation, the acidification of these cells is also one of the first early elements distinguishing these cortex cells from their neighbors [24]. In the third so-called cell death initiation phase, the cells of the median cortex layers begin to undergo programmed cell death (PCD) where the first aerenchyma begins to form [24]. The last phase, the cell death expansion phase, starts when cell death extends radially to other cortical cells (**Figure 3**). When aerenchyma formation is completed, some bridges of living cortical cells as well as the outermost layers of the

cortex remain alive and are called "forks." Clearly, the cortex is a heterogeneous tissue considering the fate of the most distal cortex cell layer, which undergoes radial cell division, and that of the median cortex cell layer, where cell death starts and expands radially. These findings suggest tight developmental control of cortex differentiation and cell death expansion. Rice has constitutive aerenchymas that are formed regardless of the growing conditions (well drained or submerged) but are induced in a more distal position inside root meristems in submerged conditions [25] (**Figure 2A**). Even upland rice cultivars form root aerenchymas despite cultivation under non-submerged conditions.

Molecular mechanisms of lysogenous aerenchyma formation

Programmed cell death is triggered by and under the control of ethylene in rice as well as in maize and other cereals. If the biosynthesis or perception of ethylene is blocked, there is no formation of aerenchymas, confirming the key role of ethylene [26] in the formation of both inducible and constitutive aerenchyma in rice. Most of the data on the molecular mechanisms of aerenchyma formation come from experiments on inducible aerenchymas. The formation mechanisms of constitutive aerenchymas are not known but probably overlap in part with those of inducible aerenchymas. The formation of inducible aerenchymas is blocked in the presence of ethylene biosynthesis inhibitors in corn [27]. The use of cell death inhibitors slows down or blocks the formation of aerenchymas, demonstrating that this process is necessary for their formation [28]. Moreover, only cortical tissues and not all layers undergo cell death, indicating that there is strict developmental control of the cells that undergo PCD. If we summarize the different steps, first there is perception of hypoxia and subsequent synthesis of ethylene, followed by perception of ethylene by the mid-cortex cells. The next steps are the initiation of cell death, followed by the production of hydrolytic parietal enzymes, chromatin condensation, cell death and the reabsorption of water and cell contents by adjacent cells [29].

A molecular model for inducible aerenchyma formation was recently established in rice (Figure 2C). In response to flooding, ethylene accumulates and the oxygen concentration decreases in root tips. The decrease in oxygen concentration stimulates the transcription of the ACC synthase (ACS) and ACC oxidase (ACO) ethylene biosynthesis genes and increases the local concentration of ethylene; ethylene also passively accumulates due to restricted gas exchange during waterlogging. In parallel, the

decrease in oxygen concentration stimulates the production of Very long chain fatty acid (VLCFA) in outer cell layers, which, through an unknown mechanism, increases the transcription of the ACS gene [30]. The increase in ethylene concentration will then stimulate the biosynthesis of reactive oxygen species (ROS) by activating respiratory burst oxidase homolog protein H (RBOHH), which will increase H_2O_2 production [31]. The induction of RBOHH requires the entry of Ca_{2+} into the cortical cells in parallel with ruthenium inhibition of ROS production. This Ca_{2+} will allow the activation of RBOHH following its phosphorylation by CDPK5 and/or CDPK3 [31]. The H_2O_2 produced will enter or be transported into the cell, activating the downstream pathway to cell death [32]. Under well-drained conditions, metallothionein (MT) genes acting as ROS detoxifiers are expressed in all tissues but repressed in cortical cells during waterlogging, suggesting that they play an important role in tissue-specific PCD in both species. Maintaining the expression of MT in stele cells and external tissues inhibits cell death in these cells despite the production of H_2O_2 by RBOHH in the cortex under submerged conditions in rice and maize [32, 33]. As lysogenous aerenchyma is initiated in the mid-cortex, these cells are more sensitive than other cells to ethylene. The radial expansion of cell death in rice may thus be due to H_2O_2 diffusion originating from these mid-cortex cells to lateral cortex cells. This hypothesis is partially confirmed by visualization of the radial expansion of O_{2-} in cortical cells using the Nitrotetrazolium Blue chloride (NBT) stain [31]. However, the control of mid-cortex priming seems independent of ethylene. When ethylene biosynthesis is completely blocked using $AgNO_{3+}$, mid-cortex radial expansion still occurs, while aerenchyma formation is completely blocked, suggesting that aerenchyma formation is controlled developmentally and/or hormonally in rice by independent mechanisms and signals [26].

Rice forms constitutive aerenchymas, but aerenchyma formation is stimulated under submerged conditions, while other cereals such as maize form aerenchymas only under waterlogging conditions (Figure 2A). If ethylene is involved in aerenchyma formation in these two species, recent results suggest that differences in the response between these two cereals may be due to differences in the regulatory mechanisms of ethylene biosynthesis [34]. In rice, the expression level of the ACO gene (OsACO5) is high under well-drained conditions, and this level would be sufficient to ensure the formation of aerenchymas even under aerobic conditions. The oxygen carried by these constitutive

aerenchymas is sufficient level for the activity of the ACO gene, as ACO activity depends on O2 availability. In maize, the low level of ACO (ZmACO15) expression would not allow aerobic aerenchyma formation, but its expression would increase under anaerobic conditions, stimulating aerenchyma production. The low oxygen level present in the root meristems would then be insufficient to maintain a high level of ethylene biosynthesis and could explain the adaptive advantage of rice in producing constitutive aerenchymas for submergence tolerance [34].

Future questions

To date, no key genes for aerenchyma formation, such as cortical cell death regulators or priming genes for the initiation phase of cell death have been identified. Moreover, **due to the close phylogenetic relationship of cereals and the existence of a partially conserved mechanism of inducible aerenchyma formation between rice and other cereals, it should be possible to transfer rice flooding tolerance mechanisms existing in rice conferred by aerenchymas to other cereals. This will also allow** to explore the link between RCA rate and several agronomical traits, including nutritional deficiencies and drought stress. Genetic diversity exists for aerenchyma formation. Teosinte is tolerant to flooding, and recently, several major loci for aerenchyma and flooding tolerance loci were identified and introduced in maize, suggesting that there is great potential for developing tolerant cereals using wild species as a source of genetic variability [35, 36]. In sorghum, genetic variation seems also to exist for aerenchyma in specific cultivars [37]. Adaptive genes for flooding were for instance identified in rice and found specific to some rice sub-groups [4, 5]. The *SUB1A* and *SK1*/*SK2* genes involved in quiescence and escape flooding strategies, respectively [4, 5] both part of Group VII of the a family [38] are absent from the Nipponbare japonica reference sequence. These are transcription factors that play a role in coordinating the response of plants to oxygen deficiency. Altogether, these results suggest that large genetic variability for this trait exists and remains underemployed in the identification of key genes involved in aerenchyma formation and harnessing in breeding.

REFERENCES

1. FAO (2015) The impact of natural hazards and disasters on agriculture and food security and nutrition. http://www.fao.org/3/a-i4434e.pdf, (accessed).

2. Nishiuchi, S. *et al.* (2012) Mechanisms for coping with submergence and waterlogging in rice. Rice (N Y) 5 (1), 2.

3. Sasidharan, R. and Voesenek, L.A. (2015) Ethylene-Mediated Acclimations to Flooding Stress. Plant Physiol 169 (1), 3-12.

4. Xu, K. *et al.* (2006) Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442 (7103), 705-8.

5. Hattori, Y. *et al.* (2009) The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. Nature 460 (7258), 1026-30.

6. Seago, J.L., Jr. *et al.* (2005) A re-examination of the root cortex in wetland flowering plants with respect to aerenchyma. Ann Bot 96 (4), 565-79.

7. Kulichikhin, K. *et al.* (2014) Biochemical and molecular characterization of rice (Oryza sativa L.) roots forming a barrier to radial oxygen loss. Plant Cell Environ 37 (10), 2406-20.

8. Striker, G.G. *et al.* (2007) Trade-off between root porosity and mechanical strength in species with different types of aerenchyma. Plant Cell Environ 30 (5), 580-9.

9. Rebouillat, J. *et al.* (2009) Molecular Genetics of Rice Root Development. Rice 2 (1), 15-34.

10. Lynch, J.P. (2015) Root phenes that reduce the metabolic costs of soil exploration: opportunities for 21st century agriculture. Plant Cell Environ 38 (9), 1775-84.

11. Zhu, J. *et al.* (2010) Root cortical aerenchyma improves the drought tolerance of maize (Zea mays L.). Plant Cell Environ 33 (5), 740-9.

12. Saengwilai, P. *et al.* (2014) Root cortical aerenchyma enhances nitrogen acquisition from low-nitrogen soils in maize. Plant Physiol 166 (2), 726-35.

13. Gao, K. *et al.* (2015) A comprehensive analysis of root morphological changes and nitrogen allocation in maize in response to low nitrogen stress. Plant Cell Environ 38 (4), 740-50.

14. Pujol, V. and Wissuwa, M. (2018) Contrasting development of lysigenous aerenchyma in two rice genotypes under phosphorus deficiency. BMC Res Notes 11 (1), 60.

15. Galindo-Castaneda, T. *et al.* (2018) Reduced root cortical burden improves growth and grain yield under low phosphorus availability in maize. Plant Cell Environ 41 (7), 1579-1592.

16. Hu, B. *et al.* (2014) Root cortical aerenchyma inhibits radial nutrient transport in maize (Zea mays). Ann Bot 113 (1), 181-9.

17. Yang, X. *et al.* (2012) Drought-induced root aerenchyma formation restricts water uptake in rice seedlings supplied with nitrate. Plant Cell Physiol 53 (3), 495-504.

18. Bhattacharyya, P. *et al.* (2019) Mechanism of plant mediated methane emission in tropical lowland rice. Sci Total Environ 651 (Pt 1), 84-92.

19. Mano, H. and Morisaki, H. (2008) Endophytic bacteria in the rice plant. Microbes Environ 23 (2), 109-17.

20. Dodd, I.C. *et al.* (2010) Rhizobacterial mediation of plant hormone status. Annals of Applied Biology 157 (3), 361-379.

21. Glick, B.R. (2005) Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. FEMS Microbiol Lett 251 (1), 1-7.

22. Hu, L.Y. *et al.* (2018) Mutualistic fungus Phomopsis liquidambari increases root aerenchyma formation through auxin-mediated ethylene accumulation in rice (Oryza sativa L.). Plant Physiol Biochem 130, 367-376.

23. Engelhard, M. *et al.* (2000) Preferential occurrence of diazotrophic endophytes, Azoarcus spp., in wild rice species and land races of Oryza sativa in comparison with modern races. Environ Microbiol 2 (2), 131-41.

24. Kawai, M. *et al.* (1998) Cellular dissection of the degradation pattern of cortical cell death during aerenchyma formation of rice roots. Planta 204 (3), 277-287.

25. Yamauchi, T. *et al.* (2018) Regulation of Root Traits for Internal Aeration and Tolerance to Soil Waterlogging-Flooding Stress. Plant Physiol 176 (2), 1118-1130.

26. Justin, S.H.F.W. and Armstrong, W. (1991) Evidence for the involvement of ethene in aerenchyma formation in adventitious roots of rice (Oryza sativa L.). New Phytologist 118 (1), 49-62.

27. Drew, M.C. *et al.* (1981) Inhibition by silver ions of gas space (aerenchyma) formation in adventitious roots of Zea mays L. subjected to exogenous ethylene or to oxygen deficiency. Planta 153 (3), 217-24.

28. Drew, M.C. *et al.* (2000) Programmed cell death and aerenchyma formation in roots. Trends Plant Sci 5 (3), 123-7.

29. Evans, D.E. (2004) Aerenchyma formation. New Phytologist 161 (1), 35-49.

30. Yamauchi, T. *et al.* (2015) Ethylene Biosynthesis Is Promoted by Very-Long-Chain Fatty Acids during Lysigenous Aerenchyma Formation in Rice Roots. Plant Physiol 169 (1), 180-93.

31. Yamauchi, T. *et al.* (2017) An NADPH Oxidase RBOH Functions in Rice Roots during Lysigenous Aerenchyma Formation under Oxygen-Deficient Conditions. Plant Cell 29 (4), 775-790.

32. Yamauchi, T. *et al.* (2017) METALLOTHIONEIN genes encoding ROS scavenging enzymes are down-regulated in the root cortex during inducible aerenchyma formation in rice. Plant Signal Behav 12 (11), e1388976.

33. Yamauchi, T. *et al.* (2011) Lysigenous aerenchyma formation in maize root is confined to cortical cells by regulation of genes related to generation and scavenging of reactive oxygen species. Plant Signal Behav 6 (5), 759-61.

34. Yamauchi, T. *et al.* (2016) Ethylene-dependent aerenchyma formation in adventitious roots is regulated differently in rice and maize. Plant Cell Environ 39 (10), 2145-57.

35. Mano, Y. and Omori, F. (2013) Flooding tolerance in interspecific introgression lines containing chromosome segments from teosinte (Zea nicaraguensis) in maize (Zea mays subsp. mays). Ann Bot 112 (6), 1125-39.

36. Mano, Y. and Omori, F. (2008) Verification of QTL controlling root aerenchyma formation in a maize teosinte *Zea nicaraguensis* advanced backcross populatio. Breeding Science 58 (3), 217-223.

37. Jaffuel, S. *et al.* (2016) Anatomical features of an African sorghum landrace adapted to flooded conditions. Australian Journal of Crop Science 10 (10), 1489-1495.

38. Giuntoli, B. and Perata, P. (2018) Group VII Ethylene Response Factors in Arabidopsis: Regulation and Physiological Roles. Plant Physiol 176 (2), 1143-1155.

GLOSSARY

Aerenchyma : Tissue that contain cavities for gas exchange between stems and roots; by extension aerenchyma are gas-containing cavities

Expansigeny: Aerenchymas formed by cell division and elongation

Lysogeny: Aerenchymas formed by programmed cell death

PCD: Programmed Cell Death, cell death induced by internal plant signals

Schizogeny: Aerenchymas formed by the separation of cell walls

RCA : Root Cortex Aerenchyma, part of the cortex that differentiate as aerenchymas

ROS : Reactive Oxygen Species, Oxygen-containing reactive chemical compounds

ROL: Loss of radial oxygen to the external environment during its diffusion to the root tip.

ROL Barrier: Tissue differentiation that limits ROL and channels oxygen diffusion to the root tip

Waterlogging: Complete submergence of the root parts of a plant, to be opposed to submergence or the entire plant is submerged.

Chapitre 2

Développement d'un protocole de Laser Microdissection (LM) couplé à la technologie RT-dd-PCR



Figure 15: Système de microdissection laser (LM) Leica LMD7000
Le système de microdissection Leica LMD 7000 est l'unique système utilisant la récupération des coupes microdisséquées par gravité, ce qui limite les contaminations. Le laser passe par l'objectif du microscope et peut être dévié pour découper n'importe quel forme d'échantillon. Ainsi, avec ce système, le support de lame est immobile à la découpe.

Chapitre 2: Développement d'un protocole de Laser Microdissection (LM) couplé à la technologie RT-dd-PCR

1.Introduction

1.1 Soyons plus spécifique : Récolter des informations –omiques pertinentes

La biologie d'une plante résulte en partie de la dynamique d'un environnement cellulaire complexe. En effet, chaque organe est structuré en tissus, puis en types cellulaires distincts ayant chacun leurs fonctions. Ainsi, la compréhension précise d'un processus biologique qui a lieu dans un groupe de cellules ou au sein d'un tissu nécessite souvent une méthode qui permet d'en extraire des informations –omique. Trois technologies permettent actuellement d'isoler des types cellulaires distincts et d'en isoler l'ARN: le FACS (Fluorescence assisted cell sorting, Birnbraum *et al.*, 2003), l'INTACT (Isolation of nuclei tagged in specific cell types, Deal and Henikoff, 2010; Deal and Henikoff, 2011) et la LM (Laser microdissection, **Figure 15**, Day *et al.*, 2005).

1.2 Isoler des groupes de tissus

Le FACS et l'INTACT utilisent des lignées transgéniques exprimant une protéine fluorescente ou une protéine nucléaire particulière de manière tissu-spécifique. Ces technologies ont donc été essentiellement déployées sur *A. thaliana,* plante modèle pour laquelle on possède le plus de ressources moléculaires, et dans une moindre mesure sur *Oryza sativa* (Evrard *et al.*, 2012; Reynoso *et al.*, 2018). Ces technologies nécessitent la digestion des parois cellulaires afin d'en extraire des protoplastes (FACS) ou des noyaux (INTACT), ce qui peut générer des biais et des contaminations croisées (Dinneny *et al.*, 2008). Elles ont néanmoins été optimisées pour des expérimentations à grandes échelles et ont démontré leur intérêt.

La LM est une technologie complémentaire qui permet de découper au laser des sections fixées ou cryofixées afin d'en extraire les ARN et d'établir un transcriptome type cellulaire spécifique (Day *et al.*, 2005, Sakai *et al.*, 2018; Takahashi *et al.*, 2010).

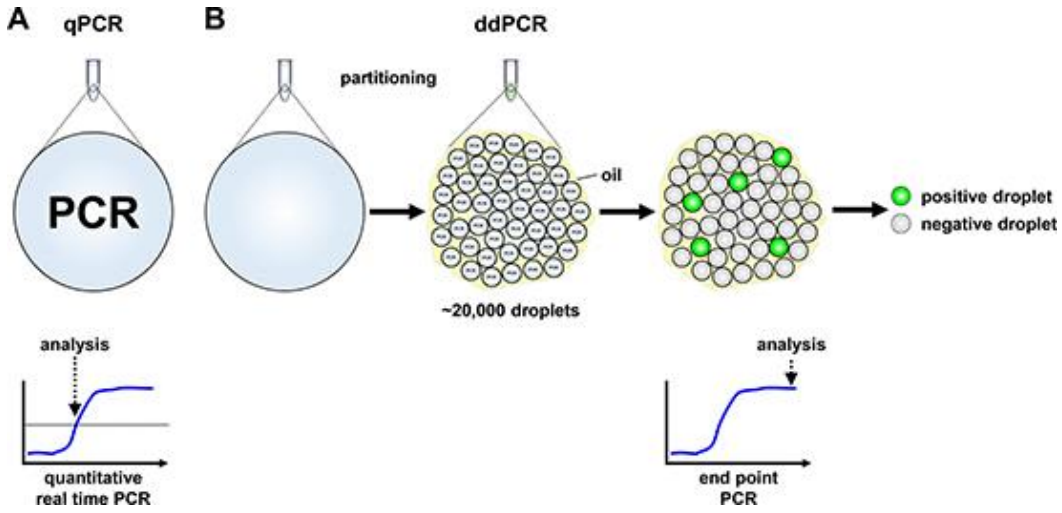


Figure 16: Représentation schématique de la PCR quantitative en temps réel (qPCR) et de la digital droplet PCR (ddPCR), d'après Lodrini et al., 2017.
A. Avec une qPCR en temps réel, l'analyse se fait pendant la phase exponentielle, en temps réel, de la réaction PCR par détection de fluorescence. Ainsi, la qualité de l'analyse est en grande partie lié à la qualité de l'amplification.
B La réaction ddPCR est partitionné en environ 20000 gouttelettes dont la fluorescence sera évaluée en point final, chaque gouttelette sera alors considérée comme positive, ou négative (digitale). La qualité de la PCR est moins influencée par des contaminants, les limites de détection sont plus basses, et d'une manière générale le résultat final n'est que peu influencé par la qualité de la réaction PCR (Lodrini et al., 2017) et son efficacité.

Comme elle ne demande aucun autre prérequis qu'une fixation des échantillons ou un prétraitement (cryofixation), elle a été déployée sur un nombre plus important de plantes comme le maïs (Zhan *et al.*, 2015), *Medicago truncatula* (Roux *et al.*, 2014), *Lotus japonicus* (Takahashi *et al.*, 2010) pour ne citer que trois exemples. Cette technologie a déjà été utilisée sur des racines de riz pour établir un atlas transcriptomique en 2010 et 2012 (Sato *et al.*, 2010, Takehisa *et al.*, 2012). Cependant, les difficultés techniques intrinsèques à la LM a limité sa diffusion.

1.3 Laser Microdissection, pertinence pour l'étude du développement des aérenchymes

Il n'existe aucun marqueur spécifique du cortex chez le riz, ce qui excluait d'entrée la possibilité d'utiliser les technologies INTACT ou FACS dans notre cas. La LM est un hybride entre plusieurs champs d'étude du domaine de l'optique, de la microscopie et la science des lasers. Ainsi, c'est la seule technologie permettant d'isoler des tissus que l'on peut visualiser directement. Au-delà de pouvoir contrôler les contaminations croisées, cette visualisation permet également de distinguer des tissus en fonction de leurs stades de développement. Cette technologie est donc particulièrement adaptée à notre objet d'étude qu'est la formation des aérenchymes. En effet, c'est un processus développemental et tissu-spécifique. Le LM doit nous permettre d'isoler des cellules du cortex afin d'en extraire les ARN et ainsi réaliser à terme un transcriptome lié aux principales phases de développement du cortex (**cf chapitre 1 et 3**).

1.4 Bref état de l'art de la technologie LM : limites des approches qui ont précédées

Deux atlas des transcriptomes de différents compartiments racinaires (cylindre vasculaire/endoderme, cortex, sclérenchyme/exoderme/épiderme) en fonction de leurs distances à la pointe racinaire ont été construits pour les racine chez le riz (http://bioinformatics.med.yale.edu/riceatlas/overview.jspx, Jiao *et al.*, 2009, **http://ricexpro.dna.affrc.go.jp**, Takehisa *et al.*, 2012,). Ces études ne donnent aucune information concernant la qualité des ARN extraits or la qualité des analyses transcriptomiques est fortement impactée lorsque l'ARN est dégradé, notamment en terme de quantification des niveaux d'expressions relatifs (Galego Romero *et al.*, 2014, Vermeulen *et al.*, 2011). De plus, ces approches se basaient sur des puces à ADN, elles-

mêmes développées à partir d'annotations du génome du riz anciennes et incomplètes. Un autre problème, qui n'a pas été évalué dans ces études, est celui du taux de contamination inter-tissus. Dernier problème, le plus important, c'est que dans le cadre de l'étude du développement des aérenchymes, il est critique de microdisséquer le tissu cortical en regard des 4 phases de développement de l'aérenchyme afin de capturer la dynamique de l'expression des gènes intervenant dans la formation de ce tissu (Mounier *et al.*, soumis, Kawai *et al.*, 1998). Aucune de ces études n'a été conçue initialement pour répondre à cet objectif.

1.5 LM/RT-ddPCR, une méthode fiable et robuste de déterminer des profils d'expressions tissus spécifiques

Ces dernières années ont vu l'apparition d'une nouvelle technologie PCR, la droplet digital PCR qui est complémentaire de la PCR quantitative (qPCR) et commercialement disponible depuis 2011 (ddPCR, Kiss *et al.*, 2008). Comme pour une qPCR classique, la ddPCR utilise une polymérase Taq pour amplifier un fragment d'ADN. Cependant, il existe deux différences importantes pour la ddPCR comparativement à la qPCR: le fractionnement de la réaction PCR dans des milliers de gouttelettes avant amplification et l'acquisition des données en « point final » (**Figure 16**, Lodrini *et al.*, 2017, Taylor *et al.*, 2017). Ces deux aspects offrent l'avantage d'une quantification directe, absolue et indépendante de l'échantillon sans nécessiter de courbes d'étalonnage (Gutiérrez-Aguirre *et al.*, 2015, Taylor *et al.*, 2017). En comparaison avec la RT-qPCR, la RT-ddPCR est moins sensibles aux contaminants pouvant perturber la réaction de polymérase ou l'hybridation des amorces, un problème classique en qPCR (Racki *et al.*, 2014, Taylor *et al.*, 2017). De plus, l'analyse en point final permet de s'émanciper des biais induits par les différences d'efficacité des réactions PCR en fonction des échantillons et des couples d'amorces. En effet, avec la RT-ddPCR, chaque gouttelette contenant une réaction PCR est lue indépendamment en un signal pouvant s'apparenter à un signal positif ou négatif (d'où le terme de digital PCR) (Hughesman *et al.*, 2016). Enfin, le fractionnement permet la détection de cibles extrêmement peu abondantes, ou pour de faibles concentrations d'ARN, et ce même en présence de contaminants (Racki *et al.*, 2014, Vogelstein and Kinzler, 1999). Ce dernier point est particulièrement intéressant pour l'analyse d'échantillons issu de LM dont les quantités d'ARN sont faibles, et nécessitent donc en

pratique la collecte de plusieurs centaines de coupes (Takahashi *et al.*, 2010, , Takehisa *et al.*, 2012) ou une optimisation fastidieuse des conditions de qPCR.

2. OBJECTIFS

La technologie LM est aujourd'hui devenue un outil utilisé dans de nombreux domaines de recherche en biologie végétale (pour une revue : Gautam et Sarkar, 2015). Les protocoles publiés nécessitent plus de deux semaines de préparations des échantillons et la microdissection de plusieurs dizaines voire centaines de coupes pour obtenir une quantité d'ARN suffisante pour les analyses postérieures (Takahashi *et al.*, 2010). A ce stade, cette technologie ne peut être utilisée en routine, ce qui rend également laborieux la collecte de tissus à des stades de développement précis (cf partie 3).

En réponse à ces limitations, nous avons développé un protocole de LM permettant une récupération d'une quantité analysable d'ARN de haute qualité à partir d'une dizaine de coupes histologiques fixées. Ce protocole est robuste et applicable en moins de 24 heures, et les échantillons fixés dans la paraffine peuvent se conserver dans le temps sans dégradation de la qualité de l'ARN (Mounier *et al.*, soumis).

Afin d'être à même de pouvoir quantifier des profils d'expression à partir de picogrammes d'ARN issu de la LM, et également tester la contamination inter-tissus nous avons utilisé la technologie RT-ddPCR et un marqueur spécifique des tissus vasculaires. Enfin nous avons démontré la pertinence du couplage LM/ddPCR pour tester l'expression tissu spécifique de gènes candidats (Mounier *et al.*, soumis, fin de la partie 2).

3. Article soumis dans BMC Cellular and Molecular Biology: *A fast, efficient and high throughput procedure for tissue specific expression profiling of rice roots using Laser Microdissection and RT-digital droplet PCR*

Title: A fast, efficient and high throughput procedure for tissue specific expression profiling of rice roots using Laser Microdissection and RT-digital droplet PCR

Running title: Laser Capture Microdissection and RT digital roplet PCR in rice root tissues

Authors: Thibault Mounier*[1], Sergi Navarro Sanz*[1], Charlotte Bureau*[1], Lefeuvre Antoine[2], Fabrice Varoquaux[1], Franz Durandet[2], Christophe Périn[1]

Affiliations:

[1] CIRAD, UMR-AGAP, Université de Montpellier, Avenue Agropolis, F-34398 Montpellier Cedex 5, France

[2] IAGE company, Avenue Agropolis, F-34398 Montpellier Cedex 5, France

*these authors contribute equally to this work

Contact: Christophe Périn: perin@cirad.fr

ABSTRACT

Background: In rice, the cortex and the outer tissues, play a key role in submergence tolerance. Cortex differentiate aerenchyma, air-containing cavities that allow oxygen to flow from shoot to root while suberification of exodermis and lignification of sclerenchyma limit oxygen loss from mature parts of the root forming a root oxygen loss barrier. The genes and gene networks involved in the cellular identity and differentiation of these tissues are still poorly known. Identification and characterization of key regulators of aerenchyma and ROL (Root Oxygen Loss) barrier formation requires obtaining the specific expression profiles of these tissues.

Results: We have optimized an approach combining Laser Microdissection and ddRT-PCR (digital Droplet RT-PCR) to identify specific tissue expression profiles at high throughput. The protocol developed makes it possible to extract high quality RNA in three days for root tissues with a low contamination rate. We have also demonstrated that it is possible to extract RNAs from paraffin blocks stored at 4°C without loss of quality. We included a detailed troubleshooting guide to allow future users to adapt this protocol for other tissues and/or species. We demonstrated our protocol can be used as an alternative to in situ hybridization to characterize gene expression at tissue specific level by combining LCM with ddRT-PCR even with low RNA concentration input. We illustrated the efficiency of this approach by validating three out four potential tissue specific candidate genes issued from RiceXpro Database.

Conclusion: This detailed protocol as well as the critical steps to optimize it for use from other species will democratize the tissue specific transcriptome approaches with LCM in combination with ddRT-PCR in plants

Keywords: Rice, root meristem, Laser Microdissection (LM), cortex, ddRT-PCR, digital droplet PCR.

BACKGROUND

There are mainly three technologies developed to isolate RNA from specific tissues that can be grouped into two groups: on the one hand, the technology using a LM (Laser Microdissection) (Takehisa *et al.*, 2012) and on the other hand, the FACS (Birnbaum *et al.*, 2003) and INTACT (Deal and Henikoff, 2010; Deal and Henikoff, 2011) technologies.

The FACS and INTACT technologies use transgenic lines and specific tissue promoters that express respectively fluorophores or a nuclei-tagged protein (Birnbaum *et al.*, 2003; Deal and Henikoff, 2010; Deal and Henikoff, 2011). Seedlings are then digested to produce protoplasts for FACS and nuclei for INTACT. These protoplasts and nuclei are separated in two populations, GFP-marked and unmarked using a cell sorter (FACS) or using a column of affinity (INTACT). RNAseq or microarray and then used to get an expression profile of the marked cell populations. These technologies have mainly been used in A. thaliana, see for instance (Birnbaum *et al.*, 2003), and to a lesser extent in rice, see for instance (Evrard *et al.*, 2012; Reynoso *et al.*, 2018). In A. thaliana the FACS has made it possible to produce a map of expression profiles of most root cell types (Birnbaum *et al.*, 2003) but also to analyze tissue specific responses to salt stress (Dinneny *et al.*, 2008).

The INTACT and FACS approaches require transgenic plants and tissue or cell-specific promoters. Moreover, for FACS the protoplasting isolation phase generates biases and often cross-contamination that are sometimes difficult to control and requires set of control experiments stress (Dinneny *et al.*, 2008). On the other hand, INTACT requires only frozen tissues to isolate nuclei par affinity-purification (Reynoso *et al.*, 2018). These technologies cannot be used to isolate few cells within a specific tissue unless of course a specific promoter is available. They also require large quantities of biological material, proportional to the number of labelled cells and are well suited for large scale transcriptomics experiments.

LM technology is complementary to FACS and INTACT technologies. It consists of laser cutting paraffin-embedded or frozen tissue sections in order to extract specific RNAs to determine expression profiles using DNA chips or RNAseq (see for instance (Sakai *et al.*, 2018; Takahashi *et al.*, 2010). This technology has been used on a larger number of

species because it does neither require tissue-specific promoters nor the generation of transgenic plants. Theoretically it can be used to do gene specific expression profiling in small scale experiments. In particular, this technology has been used in rice to isolate root tissues (Shiono *et al.*, 2014; Takehisa *et al.*, 2012), but it requires the optimization of many parameters like fixation, dehydration, inclusion in paraffin, laser parameters (Takahashi *et al.*, 2010) and is therefore only mastered by a few laboratories.

In our first experiments to use an LM approach in rice to isolate RNA from the root cortex we started from published protocols (Takahashi *et al.*, 2010). We faced many difficulties to isolate good quality RNA. This led us to take up each step again, and to determine the key parameters and rationally optimize each step by seeking solutions to identify the main sources of variation in the quality and quantity of RNA.

RNAs extracted from LM can be used to perform transcriptomic analyses by RNAseq, microarray or directly to determine the tissue specific expression profiles of candidate genes. These expression analyses are most often done by qPCR, or by RT-PCR but face several difficulties. The quantities of RNA extracted are extremely small, which requires collecting a large number of tissues and/or using amplification systems that potentially introduce bias. qPCR is sensitive to potential contaminants, PCR inhibitors and its reliability requires an almost perfect PCR. Unfortunately, the most interesting samples are those containing small quantities of targets with sometimes small or very small expression differences and potentially contaminated by PCR contaminants present in PPFE samples.

Digital droplet PCR is a recent technology. Like RT-qPCR, the PCR droplet uses Taq polymerase in a PCR reaction to amplify targets but has two important advantages comparatively with RT-qPCR. The PCR reaction is distributed in 20000 independent droplets and the expression data are collected at the end of the PCR reaction. These two differences allow direct quantification without a standard curve with more accurate and repeatable results. The fluorescence measurement at the end of the reaction in each droplet (yes/no, hence the term digital PCR), enables quantification of expression independently of the PCR efficiency. The RT-ddPCR can therefore also be used to measure gene expression for extremely small quantities of targets containing PCR contaminants.

To our knowledge, RT-ddPCR has not yet been used to test gene expression in combination with LM despite the potential of this technology.

We have developed a simplified, high throughput protocol for extracting high quality RNA from rice root tissue using LM and gene expression using ddPCR. We identified key steps and simple solutions for any team wishing to use this protocol for other tissues of other species. We also provided evidence for the low level of inter-tissue contamination using tissue specific markers. We demonstrated that it was possible to store paraffin samples without loss of quality for at least 8 months to make it possible to collect samples in the field and store samples for tropical species where it is sometimes impossible to have a laboratory close to the experimental areas. Lastly, we have shown ddRT-PCR can be used to evaluate tissue specificity of candidate genes directly from RNA extracted from LCM samples offering a powerful and unique alternative to in situ hybridization and in situ RT-PCR and illustrated the efficiency of our approach using some candidate extracted from RiceXpro Database.

The complete protocol and the associated troubleshooting list should make it possible to democratize the LM approach in plants combined with ddRT-PCR with numerous applications in plant development.

RESULTS

LM protocol in a nutshell

The full protocol is detailed in Supplemental file 1 (see also methods), including critical steps (notes) and advice for anyone wishing to move from this protocol to another tissue or species. Briefly, the first step is to germinate the rice seeds in an ARALAB (Supplemental **Figure** S1A) in a hydroponics net floating system (**Figure** S1B) and then collect two-centimeter root tips from 7-day-old seedlings (**Figure** S1C) and impregnate them with fixative all night. Bundles of 8 aligned roots are collected, one root of which is eosin-colored (**Figure** S1D) to visualize the bundles in future paraffin blocks. In a second step, the root fagots are positioned in biopsy cassettes and trapped with biopsy foam (**Figure** S1E). After dehydration the cassettes are immersed in a microwave water bath (**Figure** S1F) to carry out the paraffin inclusion. Finally, the fagots are positioned in liquid paraffin on a cold block (**Figure** S1G) and then soaked in the solidifying paraffin (**Figure** S1H). In a third step, the blocks are cut in microtome about 300µm from the root cap using the eosin-colored root to locate the bundle in the block and the cuts are placed on the blades for cutting at LM. Finally, the sections are dewaxed and the tissues are laser cut at 40x or 63x magnification. The tubes are stored at -80°C until extraction. From sample collection to RNA determination, it takes only 3 days to obtain high quality tissue specific RNAs to perform downstream RNAseq or RT-ddPCR experiments.

We have started from the Takashi protocol published in 2010. During our first tests we quickly noticed that the samples were histologically degraded, in particular that the structure of the root cuts was not preserved. We therefore first sought to preserve the structure of the root tissue.

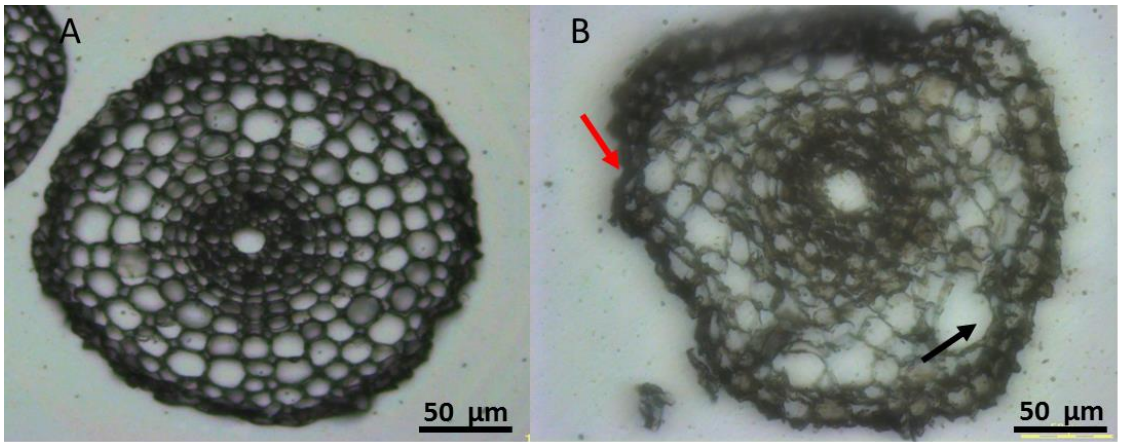


Figure 1: Impact of the use of biopsy foam for dehydration and inclusion steps on root tissue integrity.

A) Root radial section obtained by following the dehydration and inclusion treatments after using biopsy foam. B) Same as A), but without biopsy foam. The root is deformed and no longer have circular (black arrow) and we observed the destruction of the most outer cellular layers that are no longer distinguishable (black arrow) and the bursting of more inner cells such as the cortex (red arrow). Bar=50 µm.

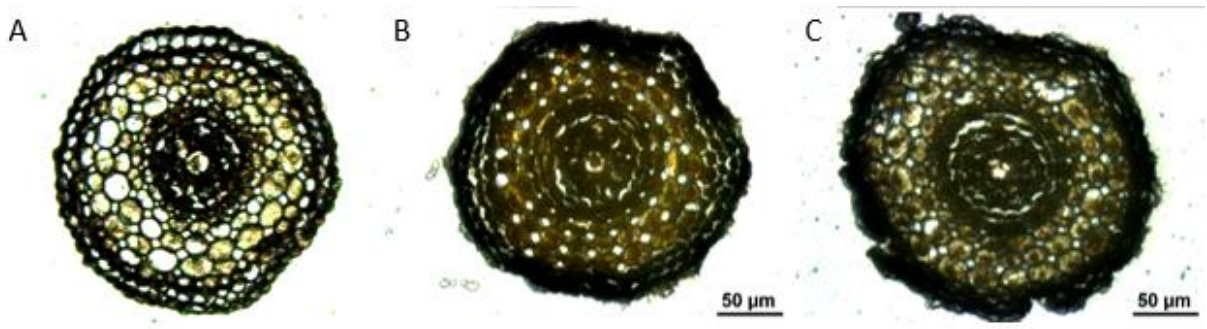


Figure 2: Impact of inclusion time on root tissue integrity. Root cross section obtained after 10 minutes of paraffin inclusion (A), 20 minutes (B) and 30 minutes (C). The first root structure deformation are visible at 20 minutes (B) (arrows) and become very important at 30 minutes (C) (arrows) with a loss of the structure of the external tissues and the impossibility to distinguish the external tissues from each other (epidermis, exodermis and sclerenchyma). Compare with an inclusion time of 10 minutes in paraffin (A) where all tissues are perfectly identifiable and where no deformation is visible.

The use of biopsy foam to immobilize samples preserves root tissues

We immobilized the root tips collected using biopsy foam when positioning the samples in the inclusion cassettes (Supplemental **Figure** S1). Comparing sections obtained without foam (**Figure 1A**) and with foam (**Figure 1B**), it can be seen that biopsy foam probably prevents samples from being tossed during the dehydration and inclusion steps and thus from undergoing repeated shocks with the edges of the cassette or between samples in order to preserve the external tissues and the integrity of the internal tissues.

Reducing the time of paraffin inclusion also preserves the integrity of root tissues

The use of biopsy foam limited root tip degradation but external tissues were still damaged (data not shown). We therefore sought to reduce the inclusion time and measure its effect, assuming that the heat contact time gradually degraded the external tissues. In **Figure 2** we can observe the effect of inclusion time on external tissues. All structures are preserved after 10 minutes of inclusion (**Figure 2A**), the external tissues partially (**Figure 2B**) or completely collapsed (**Figure 2C**) after 20 and 30 minutes of inclusion respectively.

The inclusion time does not affect the quality of the extracted RNAs but their quantity

All root tissues were cut at the LCM after 10, 20 and 35 minutes of paraffin inclusion and the quality (RIN) and quantity (pg/µm2) of extracted RNA were assessed. First of all, 10 minutes of paraffin inclusion allows to extract very good quality RNA (RIN between 8 and 9 **Figure 3A**) as well as 15 and 35 minutes (RIN around 8, not significantly different). On the other hand, increasing the inclusion time decreases the amount of RNA that can be extracted per unit area, from 0.010 pg/µm2 to 10 minutes at 0.0038 pg/µm2 to 35 minutes ($p<0.01$) or less than half as much (**Figure 3B**). We therefore set the paraffin inclusion time at 10 minutes in the following experiments.

The quality of the extracted RNA is correlated with the quantity of tissue collected

We tried to determine if collecting more tissue would increase the quality of the extracted RNAs in order to determine a minimum surface area to be used for future LM experiments.

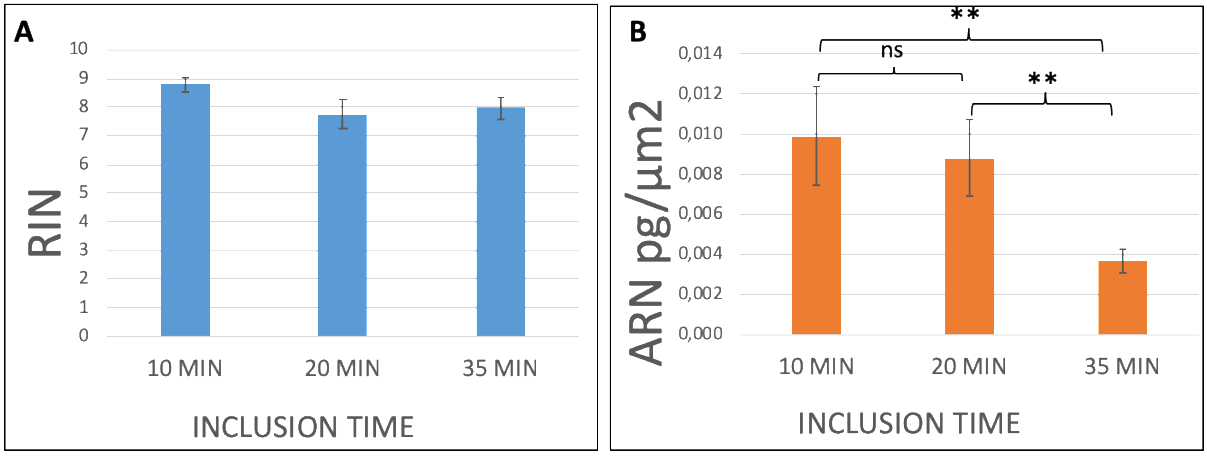


Figure 3: Effect of inclusion time on the quality and quantity of extracted RNA from LM root cross sections. A) Effect of inclusion time on the quality of extracted RNAs (RIN). B) Effect of inclusion time on the amount of RNA extracted (pg/µm2). **: significantly different at p<0.01. ns: not significantly different. Mean and standard errors are calculated from 6 biologicals repetitions.

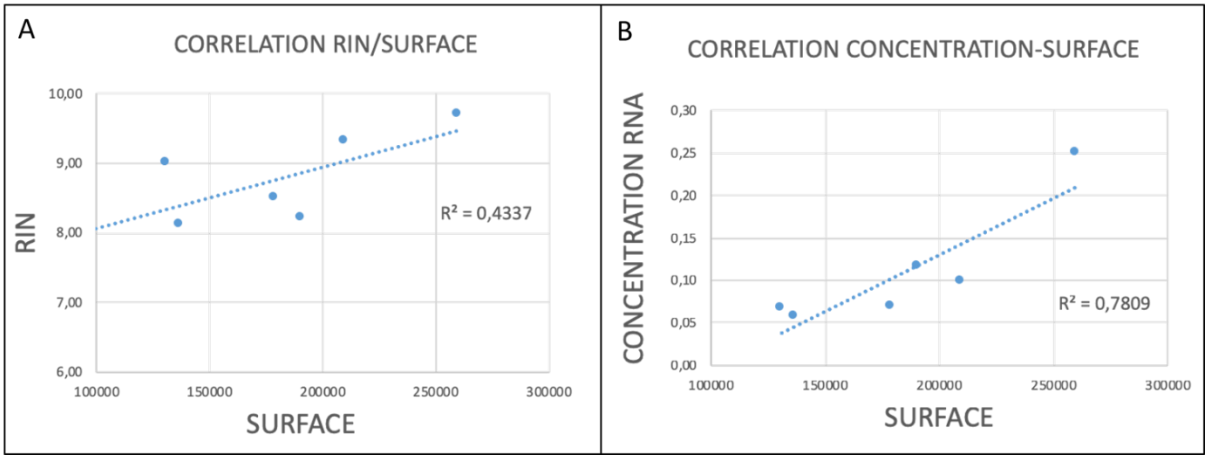


Figure 4: Correlations between the surface of microdissected tissue and the quantity and quality of extracted RNA (n=6). A) Quality of RNA (RIN) extracted as a function of the surface area of microdissected tissue (µm2). B) Quantity of RNA (pg) extracted as a function of the surface area of microdissected tissue (µm2).

. There is indeed a correlation between the quality of the extracted RNAs and the cutting tissue surface (R2=0.4337 p<0.05 **Figure 4A**) but even for the smallest quantities of tissue (**Figure 4A**) we obtained a RIN>8, a RNA quality largely sufficient for RNAseq or RT-ddPCR applications. Less surprisingly, we obtain a strong correlation (R2=0.78 **Figure 4B**) between the cut tissue surface and the amount of RNA extracted (**Figure 4B**).

It is possible to extract RNA without loss of quality from paraffin blocks longer than 8 months.

Our institute works on many tropical species and there is often a gap between the time of collecting samples in the field and their analysis. In addition, it is desirable to be able to decouple the paraffin inclusion steps from the LM cutting in order to store the samples and only make the LM cuts when this device, generally accessible through shared platforms, is available. We therefore tested whether the storage of paraffin blocks at 4°C altered the quality and quantity of the extracted RNA. The integrity of the extracted decrease significantly between 1 day and 15 days of conservation, from a RIN of 8 to a RIN of 6. After that point, no significant variation can be observed concerning RNA integrity even after 8 month of conservation (RIN of 6,7). The RNA extracted after 15 days or 40 days of storage of the paraffin blocks at 4°C have qualities substantially equivalent to those obtained from 1-day blocks (**Figure 5A**). The amounts of extracted RNA are also very close at 1, 15 and 40 days after inclusion in paraffin, with values of 0.01 pg/µm^2, 0.01 pg/µm^2 and 0.014 pg/µm^2 respectively (**Figure 5B**).

Our optimized LCM protocol offer a possible alternative to in situ hybridization and/or RT-PCR in situ experiments using ddPCR

Using this protocol, it is now possible to collect root tissues RNA from 3 biological repetitions and 30 rice roots each in one week by one person and thus offer an alternative to in situ-RT-PCR and in-situ hybridization to evaluate expression profiling in tissue using RT-ddPCR. This technology is highly sensitive, specific and can deal with degraded RNA and very low RNA input. We used the RT-PCR ddPCR kit from BioRad (Table 2) to assay feasibility of gene profiling of low RNA quantity issued from LM experiment.

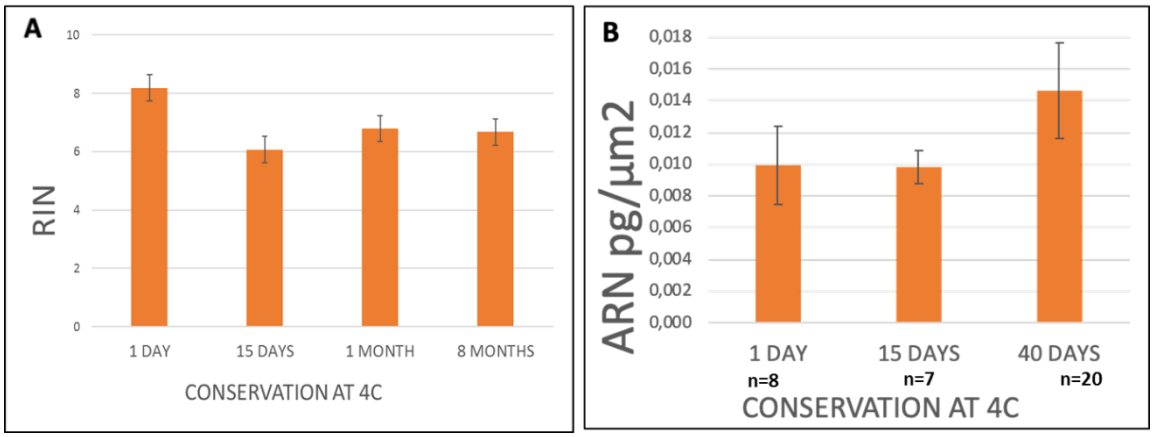


Figure 5: Effect of storage time of paraffin block at 4°C on the quality and quantity of extracted RNA from LM root cross sections. A) Effect of inclusion time on the quality of extracted RNAs (RIN). B) Effect of inclusion time on the amount of RNA extracted (pg/µm2). Mean and standard errors calculated from 3 biological repetitions.

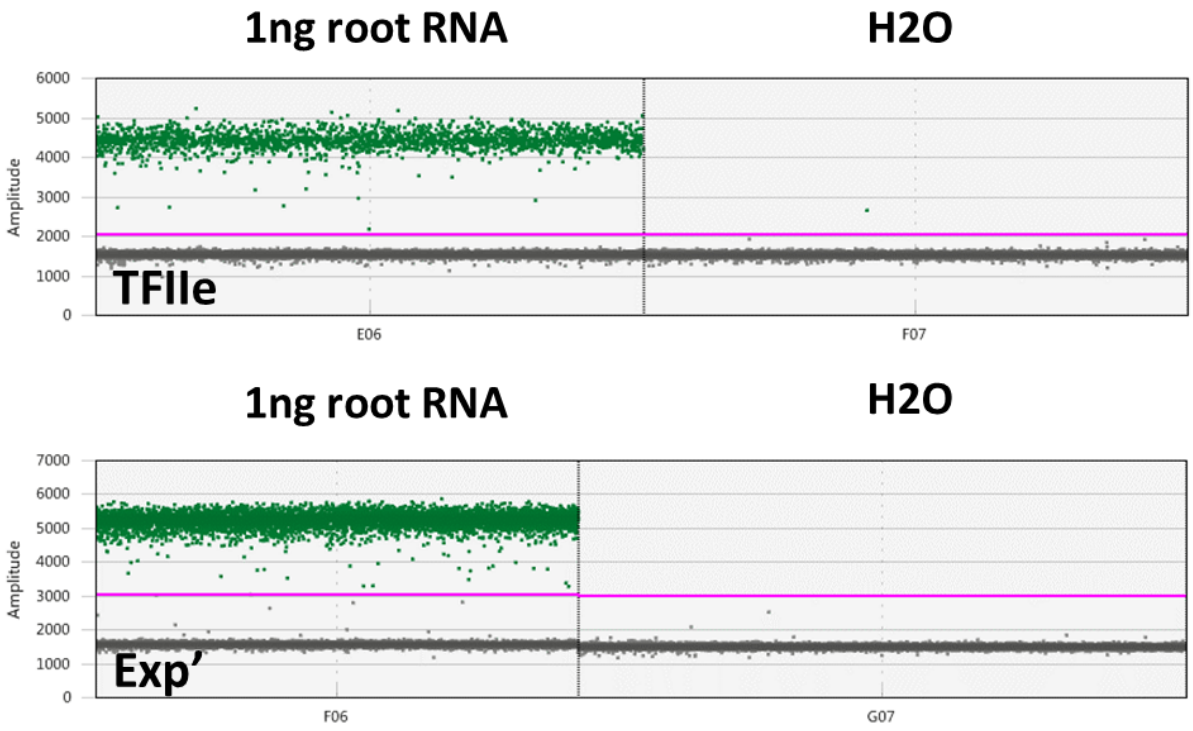


Figure 6: Validation of constitutives control for ddRT-PCR. 1D diagrams of ddRT-PCR for TFIIE (up) and Exp' (down). The red bar is the threshold for detection of positive droplet. 1ng of root RNA is used as input on right. On left, negative control without RNA.

Two genes were first tested as constitutive controls of expression for RT-ddPCR (**Figure 6**). TFIIE, a class IIE transcription factor assumed to be constitutively expressed in all transcriptionally active cells, and EXP' a gene of unknow function previously identified as a gene uniformly expressed based on a microarray expression data set (Caldana *et al.*, 2007) (Table 3). Both genes generate only one or two positive droplets in the negative control. From one ng of total root RNA, TFIIE and EXP' generate 2215 and 5814 positive droplets from a total of 13023 and 13473 droplets respectively. They are therefore sufficiently expressed to be used as standardization controls for small amounts of RNA such as those obtained using LM, EXP' being the more expressed of the two.

In order to validate our specific tissue data, we first used OsSHR1 as a specific tissue control (**Figure 7**). Based on in situ hybridization experiments (Cui *et al.*, 2007), OsSHR1 is expressed specifically in the stele. We confirm that OsSHR1 is almost exclusively expressed in the stele based on our results and is also much less expressed but still easily detectable using ddRT-PCR in cortex and outer cell layers (**Figure 7**). We then choose 4 genes with potentially distinguishable specific tissue profiles, 5NG4, stele specific, a palmytoyltransferase (SP) serine expressed in all tissues, a pollen Ole1 (PO) expressed in outer tissues (epidermis, exodermis, sclerenchyma), and a disease protein (DP) expressed in the cortex based on RiceXpro data (Supplemental **Figure** S1). We confirm the strong stele specific expression of the 5NG4 gene while PD, SP and PO are weakly expressed (**Figure 7**). As expected, PO appears to be epiderme/exoderme specific and SP is expressed at the same low level in all tissues (**Figure 7**) while DP seems to be non-expressed. The 4 genes were initially chosen for their similar level of expression, around 1000-2000 estimated by the signal obtained in microarray and accessible in RiceXpro (**Figure 7 and Supplemental Table 1**) but there is poor correlation between RiceXpro and their real level of expression.

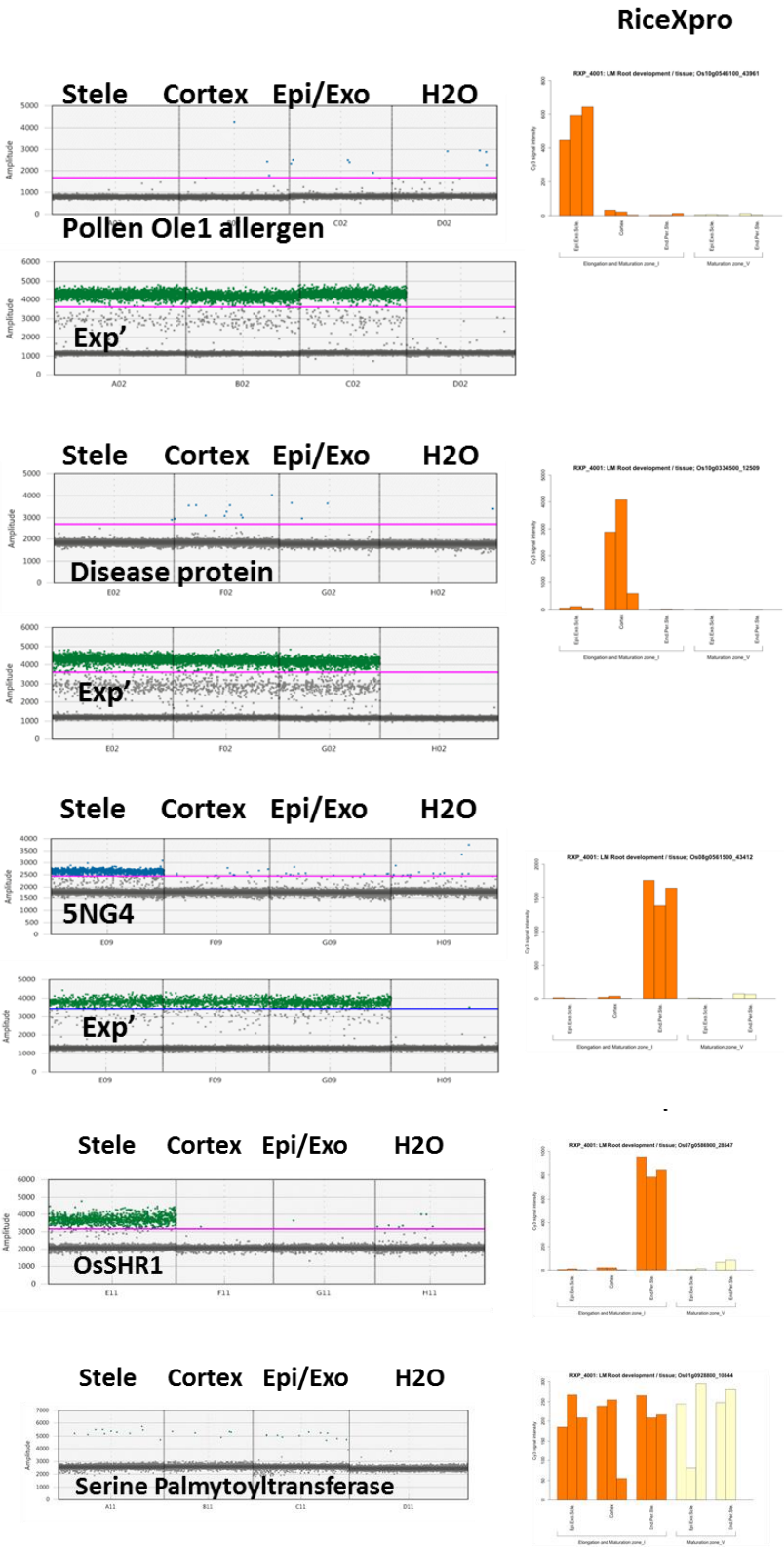


Figure 7: Expression profiling in root tissues of candidates genes from RiceXpro using RT-ddPCR. 1D diagrams of ddRT-PCR, from top and left to right, for 5NG4, Serine Palmytoyltransferase, Pollen Oe1, disease protein and OsSHR1 for three tissues (stele, cortex, epi/exo) and a negative control (H2O). In parallel, expression profiling extracted from RiceXpro Database. OsSHR1 is used as a positive control of stele specific expression.

We performed two RT-ddPCR experiments to estimate the relative level of expression between the stele, the cortex and external tissues for the 5NG4 gene (FAM probe), with the TFIIe and EXP' genes (HEX probe) (**Figure 8**) as normalization controls. 5NG4 is preferentially expressed in the stele but is also expressed at a detectable level in other tissues. By normalizing using EXP', 5NG4 is expressed 15- and 18-fold less in the cortex and in the outer tissues compared to the stele. By normalizing using TFIIe, 5NG4 is expressed 26 and 25-fold less in the cortex and external tissues compared to the stele.

In order to test the sensitivity of RT-ddPCR, we performed serial dilutions of total root RNA from 1ng RNA to 1fg (**Figure 9**). For Exp' we detect its expression up to 100 fg where only 1 positive droplet is visible. There is also a perfect linear relationship between the number of positive droplets and the amount of RNA or copy number/µL up to 10 pg (**Figure 9**). For 5NG4, we also detected gene expression up to 10 pg with once again a perfect linear relationship between the number of positive droplets and the amount of RNA or the number of copies/µL. For OsSHR1, 100 pg of RNA were necessary to detect its expression.

DISCUSSION

A simple and fast protocol to prepare paraffin block and do LM on root tissues

In three days, using our protocol, it is possible to get RNA of high quality from rice root tissue suitable for downstream applications like RT-ddPCR or RNAseq. Compare to the reference protocols (Takahashi *et al.*, 2010) we used thicker tissue sections, less sections, and got higher RNA quality. Moreover, we provide a full detailed protocol with key points to optimize for laboratory that want to adapt the protocol to other tissue/species. The main critical points are embedding time that need to be reduced to the shortest time as possible to ensure paraffin embedding in one hand, but to preserve tissue anatomy in the other hand. Besides these specific points, cold dehydration and using methanol instead water for mounting the slides are also important parameters to consider. Using paraffin instead of frozen samples allow to store embedded tissue to delay LM dissection if one want to collect samples in one site and do LM in a distant laboratory. Moreover, paraffin usually preserve tissue structure better than frozen samples and paraffin blocks can be used also for other complementary applications like in situ hybridization and immunochemistry if required.

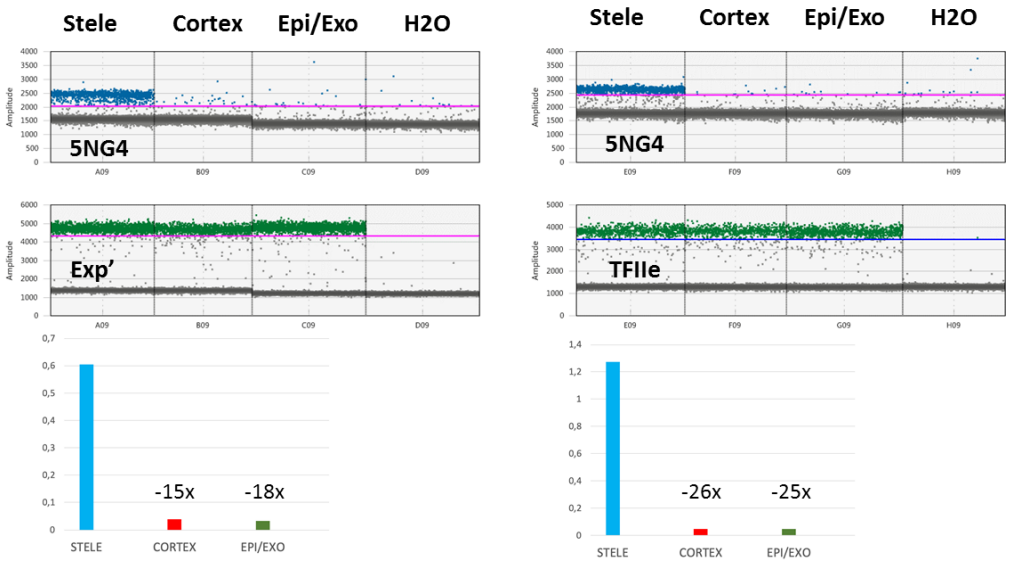


Figure 8: Relative expression profiling of 5NG4 in stele, cortex and epi/exo using TFIIE and EXP' as normalization control of RNA quantity. Top, 1D ddRT-PCR diagram for 5NG4 with EXP' (left) and TFIIe (right) controls for three tissues (stele, cortex and epi/exo) and a negative control (water). Blue, FAM probes, green HEX probes. Bottom, relative expression level in the three normalized with EXP' (left) or TFIIE (right). Abscises represent arbitrary values of expression.

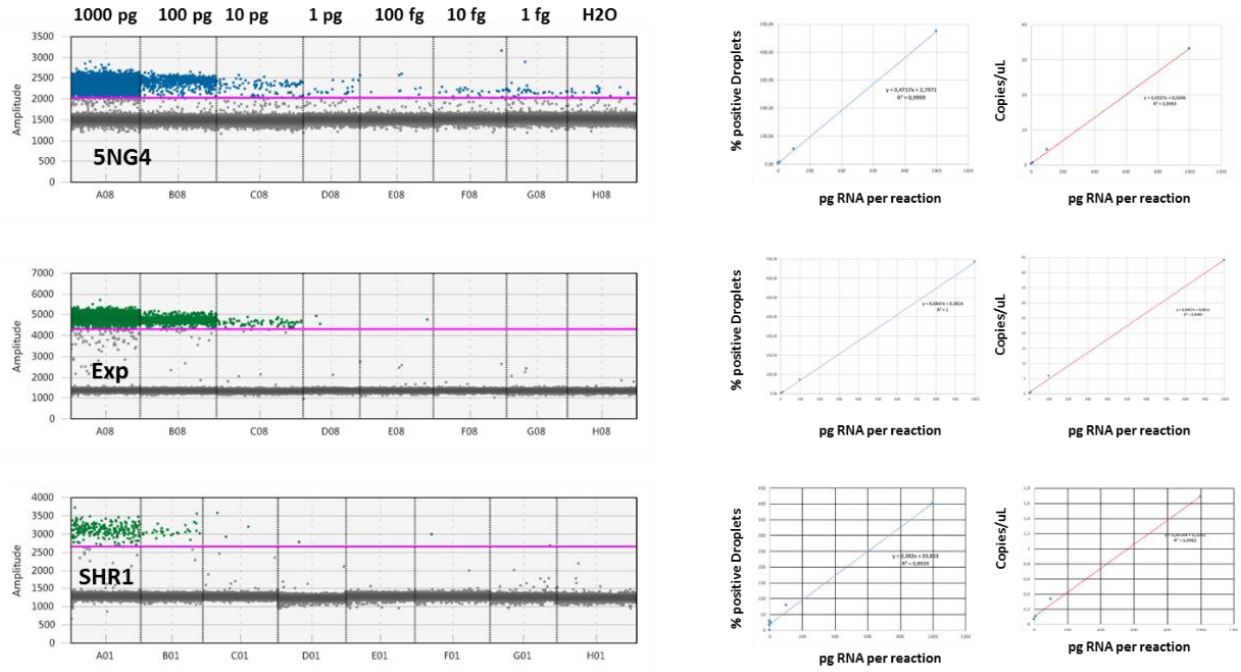


Figure 9: ddRT-PCR on serial dilutions of total root RNA for 5NG4 and EXP'. Right, 1D ddRT-PCR diagrams with serial dilutions of total root RNA from 1 ng to 1 fg plus a negative control (H20) for 5NG4 and EXP'. Right, correlation diagrams for % positives droplet or copies/µL and pg RNA per reaction. Red bar, threshold for droplet detection.

Archiving paraffin blocks was for instance described for tumor samples (Kokkat *et al.*, 2013) in biomedical research. The protocol is a high throughput one so it is possible to collect and extract RNA from hundred roots sections in one or two weeks. This protocol can be used to compare tissue specific expression profiling between several cultivars and/or mutant lines or for several environmental conditions. Compare to FACS or INTACT, it can be used in any species and our protocol should be used in a broader list of plant species with minimal modifications and/or optimizations.

A simple and efficient alternative or complementary protocol to in situ hybridization and/or in situ RT-PCR

One of the key experiment to characterize gene or gene network function is to clarify tissue expression of the candidate genes. Usually this is done by using promoter fusion and/or in situ hybridization and in situ RT-PCR. In the former, this is possible only in species where genetic transformation is available, and in the latter this is laborious and also probe dependent. A much less popular approach used in situ RT-PCR. There is only few papers describing the use of this technology probably because of numerous artifacts that affect the reliability of this technology. Tissue specific RNA extraction offered an attractive alternative but was not considered until now as a true alternative because reproductivity and technicity hampered a day to day use routinely especially in combination with qPCR. Here, we provided a much simpler and reproductible protocol that should help any laboratory wanted to use to do expression profiling between plant tissues, by ddRT-PCR but also using other downstream applications like RNAseq. For instance, we validated three genes out four candidates starting from RiceXpro Database (Jin *et al.*, 2019). Moreover, we obtained quantitative expression for OsSHR1 demonstrating this gene is also expressed outside stele tissue albeit at a much lower level. This protocol will clearly help to democratize the technology for plant applications and help to better understand tissue-and cell specific response during plant development or during interactions with changing environmental conditions including pathogen/biostimulants interactions.

CONCLUSIONS

The protocol developed and the detailed trouble shooting provides should allow any laboratory to develop LM tissue specific approaches in plant and will help to democratize the technology together with RT-ddPCR. This will offer an alternative way to identify and characterize cell and tissue specific response. As the starting material is embedded in paraffine they can be stored for long time for additional experiment to confirm or to give more precise insight using complementary technologies like in situ if needed. We demonstrated using rice root tissues as an example; we can use this protocol to characterize tissue specific response of transcription factor (OsSHR1) but also to get tissue expression profiling of four candidates genes in less than two weeks coupling LM and RT-ddPCR.

METHODS

Plant materials and plant growth conditions

Dehusked seeds of O. sativa cv Nipponbare were surface-sterilized in 70% ethanol for 2 minutes, then rinsed with sterile distilled water and disinfected by dipping in a 40% bleach solution diluted in distilled water with 0.4% Tween 80 (Sigma-Aldrich P4780-500 mL) for 30 minutes under gentle agitation at room temperature. The seeds were then rinsed seven times with distilled water. Seedling were grown hydroponically (ARALAB, Fitoclima 600) in osmotic water for seven days (see Table 1 for light/hygrometric condition setup).

Sample collection and fixation

All commercial reagents and products references can be found in Table 2. Roots tips of 1.5 cm length were hand-dissected in cold RNA secure reagent treated water (AM7005, Thermofischer scientific, USA) and placed in an cold ethanol/acetic acid (EAA, 3:1) solution. At the end of the sample harvest, EAA solution was replaced by a fresh one and vacuum infiltrated for 5 min (0.6 psi). The EAA solution was replaced again by a fresh one and samples were incubated at 4°C overnight. Some roots samples were stained with 2% eosin solution to serve as visual control and for orientation while cutting roots sections in the paraffin block .

Tissue dehydratation, inclusion and embedding

The whole dehydratation process should be done on ice. Eight roots have been aligned by the tips between two biopsy foam pads (M476-1, Simport, Canada) in histology cassettes (12677796 Thermo Scientific, USA) to form a sort of roots fagot with one root stained with eosine per fagot. Then, the samples were dipped 5 minutes in baths using ethanolic solutions with increase ethanol concentrations successively, 75%, 80%, 85%, 90%, 95% and 100%, followed by two baths of 10 min in an ethanol/butanol (1:1) solution and an absolute butanol successively. Samples were transferred in a water-bath at 54°C and transferred to an histology microwave oven (Histos 5 Rapid Tissue Processor, Milestone, Italy). The next bath is a butanol/paraffin (1:1) bath, 5 minutes at 54°C and 300W, followed by two baths of paraffin for 5 minutes at 54°C and 250W. Before the

Time	Temperature	Humidity	Light intensity
720 min	23°C	60%	0%
1min	27°C	60%	10%
59min	27°C	60%	10% to 90%
600min	27°C	60%	90%
60min	27°C	60%	90% to 10%

Table 1: ARALAB conditions for rice seedling growth

Reagent	Chemical formula	Source	Identifier
Absolute ethanol	C_2H_5OH	Honeywell, USA	603-002-00-5
Acetic acid	$C_2H_4O_2$	VWR, USA	0714-2.5L
Butanol	CH_3-CH_2-CH_2-CH_2OH	Sigma-Aldrich, USA	B7906-500ml
Leica-Paraplast EXTRA		Leica, Germany	39603002
Xylene	C6H4(CH3)2	Sigma-Aldrich, USA	214736
Nuclease-free water	H_2O	Ambion, USA	M9932
Eosin	$C_{20}H_8Br_4O_5$	RAL diagnostics, France	312710
RNA secure reagent		Thermofisher scientific, USA	AM7005
RNaseZAP		Sigma-Aldrich, USA	R2020-250ml
Biopsy foam pads		Simport, Canada	M476-1
Biopsy cassette		Square Mesh Cassette, Orange	70072-O
Histology cassette		Thermo Scientific, USA	12677796
PEN membrane glass slide		Leica, Germany	11505190
0.6 ml Microcentrifuge tubes		Molecular Bio-products, USA	3454
ddPCR 96-Well PCR plates		Biorad, USA	12001925
Pierceable Foil Heat Seal		Biorad, USA	1814040
DG8 Cartridges		Biorad, USA	1804008
DG8 Gasket for ddPCR		Biorad, USA	1863009
Droplet Generation Oil for probes		Biorad, USA	1863005
ddPCR Droplet Reader Oil		Biorad, USA	1863004
ddRT-PCR kit Biorad		Biorad, USA	186-4021

Table 2: Reagents and material

embedding step, the roots fagots were quickly removed from the cassettes while paraffin is still liquid and relocated to a cold RNase-free surface. Then the fagots were vertically transferred in a moulding tray (E70182, EMS, USA) upside down. The paraffin blocks were kept at 4°C.

Microtomy and Laser microdissection

Transversal sections of 10µm thick were cut on an RNase-free microtome (RM2255, Leica, Germany). Sections of meristematic and differentiated root tissues were visually identified in the first 500 µm after the first root cap cells. Then, the paraffin sections were mounted on a pre-warmed at 52°C PEN membrane glass slide (11505190, Leica, Germany) with drops of methanol. The sections were air dried until the methanol evaporated and the slide was dewaxed with two 2.5 minutes baths of cold xylene. Once the xylene was evaporated, the slide was immediately processed for laser microdissection (LMD7000, Leica, USA) using the following laser settings (power 21, aperture : 1, speed : 10, balance : 25, Head current 100%, pulse frenquency : 120, Offset : 180). The cut tissues were collected by gravity in a 0.5ml tube cap filled with 25 µl extraction buffer from the PicoPure® RNA isolation kit (Cat no. KIT0204, ThermoFisher Scientific, USA).

RNA extraction and dosage

The RNA extraction was carried out following the PicoPure® RNA isolation kit instruction manual with DNase treatment on column (RNase-Free DNase Set, Cat no. 79254, Qiagen, Germany). The RNA integrity was evaluated with the Agilent 2100 Bioanalyzer system (Cat no. DE72902360, Agilent, USA) using the Agilent RNA 6000 Pico kit (5067-1513, Agilent, USA).

Identification of tissue specific candidates

We first used the RiceXpro database (http://ricexpro.dna.affrc.go.jp) to identify genes with tissue specific expression (see Supplemental **Figure S3** and supplemental Table 1). We then designed primers and Taqman probes for one gene per tissue plus a reference gene with constant expression across root tissues. Primers were designed using Primer3 and Taqman probes following furnisher recommendation (BIORAD, USA) and were first validated by PCR using genomic DNA with the Diamon Taq (TAQ-I021, Eurogentec, Belgium) in a final volume of 25µL, with 2.5µL 10x Taq Mix, 1.5µL MgCl2 (25mM), 2µL

Tissue	Position	Gene	Primer Name	Oligo sequence	Amplicon Size	Probe Name	Probe Sequence	Tm Probe	Fluorophore
All	LOC_Os01g70380	Serine palmitoyltransferase	Serine_F	TTGCCGTCGATAATCCTGAC	196	pSerine	CCTCGTTCGTTCGTCGCTGACGGC	64,2	HEX
			Serine_R	GAGGAAGAGGTCGTCAATGG					
	TFIIE	Transcription factor 2E	TFIIE_F	TTAGCTGTGTTGGTCATGGG	161	pTFIIE	CGGAAGAGCTGCTTCAGGTCATCGTCG	63	HEX
			TFIIE_R	TCCCAGGAGGACATTGTGTA					
	EXP'	Expressed	Exp_F	ATGGGCAGAAGTCGAAGATG	155	pExp	AGCCAGCTTGAGGCCAACAAGAAGGCC	64,9	HEX
			Exp_R	TTTGCACTTGGTCTCAGAGG					
Xtele/endoderm	LOC_Os08g44750	Nodulin-like protein	5NG4_F	GCAGATATGGTGCATCGACA	170	p5NG4	GCCTCCCTCACCCTCGGCGGAGAGC	66,4	FAM
			5NG4_R	CCCAGAGGAGCGAGGTAGAG					
	OsSHR1	SHR1	SHR1_F	CAAGCCGCCTCCG	79	pSHR1	CGTCCTACAACTCGAGG	70	FAM
			SHR1_R	TGGACCCGCTCGAC					
Cortex	LOC_Os10g18820	Plant disease response prot.	Dis_F	AAGGGATCCACACTTCAGGT	152	pDis	GCTGCAAGCAGTGGTGAGTGGTCTGTT	63,2	FAM
			Dis_R	AGTTCTCGAACAGCATCCTC					
Outer	LOC_Os10g39890	Pollen Ole 1 allergen	Ole_F	TTCTACTTCACCCTGTCCCA	179	pOle	GGACGGTGCCACCTACTGATCGACCGT	65,2	FAM
			Ole_R	ACAAAGGCCAAACAACACAC					

Table 3: Primers and probes for Digital RT-PCR

dNTP (10mM), 1µL of forward and reverse primers (10µM) and 0.6µL of Diamond Taq. PCR conditions were: 95°C 10 mn, followed by 40 cycles (94°C 30 s, 60°C 60 s, 72°C 60 s) and the final elongation step 72°C 1 mn in a Thermocycler (Eppendorf™ 6331000041). PCR products were analyzed and validated on a 1% Agarose gel (see Table 3 for primers validated).

Digital droplet RT-PCR for quantification of expression

The RT-ddPCR was performed in a solution containing 2 µL of RNA. A reaction volume of 20µL was used For the droplet generation using the RT-DDPCR reaction kit (Biorad, USA) consisted of 5µL of RT-ddPCR supermix, 2µL of Reverse Transcriptase, 1µL of 300mM DTT, 1µL of primer, 1µL of probe, 2µL of RNA QSP 20µL of RNAse free water. Samples were transferred to 8 channel disposable droplet generation cartridge and a volume of 70 µL of droplet generation oil was added. Cartridge was loaded into the QX200 droplet generator (BIORAD, USA). After Droplet genetration, 40µL of samples were immediately transferred to 96 well PCR plate (ddPCR 96-Well PCR plates, BioRad, USA) and sealed with the PX1 Plate Sealer (BIORAD, USA). RT-ddPCR conditions on a MasterCycler Nexus Gradient (6331000017, Eppendorf, USA), were 60mn 50°C, 10mn 95°C then 40 cycles of two steps 30 s 95°C, 1mn 57°C and a final denaturation step 10 mn 98°C. The QX200 droplet reader (BIORAD, USA) was used to analyze droplet fluorescence. Each droplet was analyzed individually using a two-color detection system (FAM, HEX). PCR-positive and PCR-negative droplets are counted to get absolute quantification of target RNA molecules using the QuantaSoft 1.6 Pro (BIORAD, USA) software. Results are presented as 1D plots.

DECLARATIONS

- Ethics approval and consent to participate : Not applicable
- Consent for publication : Not applicable
- Availability of data and material : All the datasets, including video and picture are included in the article and also available from the corresponding author on reasonable request. All the transgenic lines are also available thought MTA from the corresponding author.

- Competing interests : The authors declare that they have no competing interests in this section

- Funding : Not applicable

- Authors' contributions : CP drafted the manuscript

- Acknowledgements : All LM and histology observations were performed at the Montpellier RIO Imaging facilities at the PHIV site. We thank the MRI and PHIV platforms for their technical support

REFERENCES

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., and Benfey, P.N. (2003). A gene expression map of the Arabidopsis root. Science 302:1956-1960.

Caldana, C., Scheible, W.R., Mueller-Roeber, B., and Ruzicic, S. (2007). A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. Plant Methods 3:7.

Cui, H., Levesque, M.P., Vernoux, T., Jung, J.W., Paquette, A.J., Gallagher, K.L., Wang, J.Y., Blilou, I., Scheres, B., and Benfey, P.N. (2007). An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. Science 316:421-425.

Deal, R.B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. Dev Cell 18:1030-1040.

Deal, R.B., and Henikoff, S. (2011). The INTACT method for cell type-specific gene expression and chromatin profiling in Arabidopsis thaliana. Nat Protoc 6:56-68.

Dinneny, J.R., Long, T.A., Wang, J.Y., Jung, J.W., Mace, D., Pointer, S., Barron, C., Brady, S.M., Schiefelbein, J., and Benfey, P.N. (2008). Cell identity mediates the response of Arabidopsis roots to abiotic stress. Science 320:942-945.

Evrard, A., Bargmann, B.O., Birnbaum, K.D., Tester, M., Baumann, U., and Johnson, A.A. (2012). Fluorescence-activated cell sorting for analysis of cell type-specific responses to salinity stress in Arabidopsis and rice. Methods Mol Biol 913:265-276.

Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.L., Zhang, F., *et al.* (2019). Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. Science 364:292-295.

Kokkat, T.J., Patel, M.S., McGarvey, D., LiVolsi, V.A., and Baloch, Z.W. (2013). Archived formalin-fixed paraffin-embedded (FFPE) blocks: A valuable underexploited resource for extraction of DNA, RNA, and protein. Biopreserv Biobank 11:101-106.

Reynoso, M.A., Pauluzzi, G.C., Kajala, K., Cabanlit, S., Velasco, J., Bazin, J., Deal, R., Sinha, N.R., Brady, S.M., and Bailey-Serres, J. (2018). Nuclear Transcriptomes at High Resolution Using Retooled INTACT. Plant Physiol 176:270-281.

Sakai, K., Taconnat, L., Borrega, N., Yansouni, J., Brunaud, V., Paysant-Le Roux, C., Delannoy, E., Martin Magniette, M.L., Lepiniec, L., Faure, J.D., *et al.* (2018). Combining laser-assisted microdissection (LAM) and RNA-seq allows to perform a comprehensive transcriptomic analysis of epidermal cells of Arabidopsis embryo. Plant Methods 14:10.

Shiono, K., Yamauchi, T., Yamazaki, S., Mohanty, B., Malik, A.I., Nagamura, Y., Nishizawa, N.K., Tsutsumi, N., Colmer, T.D., and Nakazono, M. (2014). Microarray analysis of laser-microdissected tissues indicates the biosynthesis of suberin in the outer part of roots during formation of a barrier to radial oxygen loss in rice (Oryza sativa). J Exp Bot 65:4795-4806.

Takahashi, H., Kamakura, H., Sato, Y., Shiono, K., Abiko, T., Tsutsumi, N., Nagamura, Y., Nishizawa, N.K., and Nakazono, M. (2010). A method for obtaining high quality RNA from paraffin sections of plant tissues by laser microdissection. J Plant Res 123:807-813.

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.

ANNEXE

Supplemental File 1 : Methods

1. Seedling growth

1.1 Seeds disinfection (note 1)

- Select 200 well formed seeds from 200-300 dehusked rice seeds (note 1)
- Desinfect 100-200 seeds in a volume of 50ml of 70% ethanol for 2 minutes
- Rinse once with 50 mL of milliQ water
- Desinfect the seeds shaking during 30 min in 50 mL of 40% bleach (9.6°) + 2 drops of tween 20
- Rinse 7 times with 40 mL of milliQ water
- Prepare four petri dishes (90x14mm) with pre-wetted Whatman paper with 8mL of milliQ
- Distribute 50 seeds per petri dish and seal with Parafilm
- Incubation O-N at 28°C in growth chamber (12h light/12h dark)

1.2 Hydroponics system setup and transfer of imbibited seeds, (**supplemental Figure S2**), (note 2)

-Disinfect four 6L buckets and floating sieves O/N with H2O2 up to 12% at RT
- Rinse them all generously with water and fill the buckets with osmosis water, place the floating sieves in them.
- At 3pm 30 hours after incubation, fill the sieves with 15-20 grains/sector (4 sectors/screen)

1.3 Seedling growth in ARALAB

- Start the program below. Ideally, the light cycle should start at 10 am (5 hours of beating between switching on the lamps and collecting samples (collection at 3pm))

- It is therefore necessary to adjust the program accordingly,

- Here is the detail of the program (60% humidity, 28° during the day and 24° at night):

Segment 0

Segment 1: 1 min ramp to increase to 10% brightness

Segment 2: 59 min of ramp to increase to 90% brightness

Segment 3: plateau from 10h to 90%

<-- if sowing at 15h, the cycle must start here, at segment 3 at 240min -->

Segment 4: 1H of ramp to lower to 10% brightness

Segment 5: 12h at 0% brightness

Segment 6: 0, back to segment 1

2. Root tip harvesting, fixation, de-hydratation and paraffine embedding (note 3)

All steps must be done RNase Free. Aluminium sheet on the work surface, gloves, containers washed with RNaseZAP, then ethanol, RNaseZAP then ethanol on LM microscope, 3 small Histos beakers with their covers, 1 rack for Histos5 cassette, 3 magnetic stirrers. Prepare a large histo beaker that will serve as a water bath (see "materials" below). Put everything in the oven at 54°C.

2.1 Prepare solutions

- Prepare the EAA solution (>200ml), keep it cold

- Distribute in 4 tubes of 30ml plus two tubes of 40ml

- Add 1ml of Eosin 2% in two of the 30ml fixing solution tubes and in one of the 40ml fixing solution tube

- Prepare 120ml of dehydration solutions with absolute Ethanol : 75%, 80%, 85% 90%, 95%, 100%, Ethanol:Butanol 1:1(v/v) and Butanol 100%

- Prepare 150 mL of Butanol:Paraffin 1:1 (v/v), add 75ml of Butanol at 56°C and add 75ml of melted paraffin the next day (see below)

- Store these solutions at 4°C overnight, except for Butanol:Paraffin to be stored at 54°C

- Fill the large histo beaker that will serve as a water bath and store it at 45°C

2.2 Root tips harvesting and fixation (under cold conditions)

- In an ice cold petri dish with 10 ml RNA secure reagent treated water, cut root sections of at least 1.5 cm

- Place several roots immediately in the fixing solution with Eosin. These roots will serve as landmarks when cutting with the microtome

-Place the rest of the roots in another fixing solution without Eosin

- Change respectively their fixing solutions, with or withour Eosin, by pouring carefully.

- Apply a cold vacuum for 5 minutes

- Replace with 40ml their fixing solutions

- Store overnight 4°C

2.3 Root tip transfer to cassette

- Clean and switch on the paraffin embedding station and Histos5 in advance.

- Put 150ml paraffin in 2 of the beakers reserved for 54°C.

- In the third one put the butanol:paraffin solution and keep it at 54°C.

- Place the biopsy cassettes in a glass petri dish filled with cold 75% ethanol

- Put a first biopsy foam bread in it

- Position very carefully the roots on the first foam so that all the root tips are aligned without stacking among them

-Form up to 3 bundles of roots per cassette, each bundle containing 7 roots.

-Add one Eosin-stained root per bundle

- Gently place a 2nd biopsy foam on the roots and close the cassette

2.4 Samples dehydration

Each step, 5min on ice

- First bath 5 min on 75% ethanol, second bath 5 min on 80% ethanol

- Third bath 5 min on 85% ethanol, Fourth bath 5 min on 90% ethanol

- Fifth bath 5 min on 80%ethanol, six bath 5 min on 85% ethanol

- Seven bath 5 min on 90%Ethanol, eight bath 5 min on 95% ethanol

- Nine bath 5 min on 95% ethanol and last bath 5 min 100% ethanol

Infiltration: each step 10min on ice

- Replace last absolute ethanol with a bath of 50%ethanol/50% n-butanol

- Replace by a bath of 100% n-butanol

2.4 Paraffine Embedding (note 4)

The microwave cannot heat directly the paraffin. It is then necessary to use a water bath system directly in Histos 5.

- Replace Butanol 100% solution with paraffin:Butanol 1:1 (v/v)

- Bath for 5min, 54°C, 300W

- Transfer to 100% paraffin bath

- Bath for 5 min, 54°C, 250W

- Transfer to the second paraffin bath

- Bath for 5 min, 56°C, 250W

- The following steps are carried out on the paraffin embedding station.

- Transfer the samples into a petri dish with 100% paraffin in the heating compartment

- Gently grasp the root bundles by the ends opposite the tips

- Gently pinch the root tips together while the paraffin is liquid

- Place them on an RNase-free surface

- Remove all root bundles before inclusion in RNAse-free conditions

- Include them in paraffin by keeping them straight, pointing towards the bottom of the pre-warmed molding tray

- Store the paraffin block at 4°C protected from the light

3. Microtome sectioning and LM

3.1 Microtome sectioning

-Treat for 30 minutes under UV light the LM collector tubes and PEN membrane slide.

- Microtome 10 µm paraffin sections. If the roots are correctly aligned, eosin-stained root indicates you the position of all the root tips.

-Start sectioning as close as possible of the root tips and cut up to 1500 µm

- While cutting sections, warm on a hot plate at 52°C both the PEN membrane slide and methanol

-Mount the sections on the PEN slide by adding few drops of methanol (note 6)

- Wait until the methanol has evaporated completely

3.2 LM (note 7)

- Remove the paraffin by washing in cold Xylene for 2,5 minutes twice.

-Air dry the slide until complete xylene evaporation.

- Microdissect stele, endoderm and cortex.

- Collect the samples in specially designed tubes containing 25 µl of extraction buffer (notes 8,9,10)

- Follow the instructions from Arcturus PicoPure kit

-Store at -80°C until RNA extraction

3.3 RNA isolation, quantification and quality assay

- RNA extraction with the Arcturus PicoPure kit (+ DNase1 treatment)

- Assess the RNA quality and concentration with Agilent RNA 6000 Pico kit

Notes and troubleshooting

Note 1: In order to achieve homogeneous germination, select the best seeds possible, not green, not broken without any fungi and/or possible bacteria contamination.

Note 2: Keep the seeds separated during the whole growing process, it reduces cross-contamination.

Note 3: Be gentle with your samples to preserve tissue integrity.

Note 4: It is critical to minimize paraffin inclusion time to reduce tissue degradation and insure good RNA quality. For new tissue/species, try several fixation, inclusion, embedding times and choose the shortest one that allows good fixation and paraffin embedding.

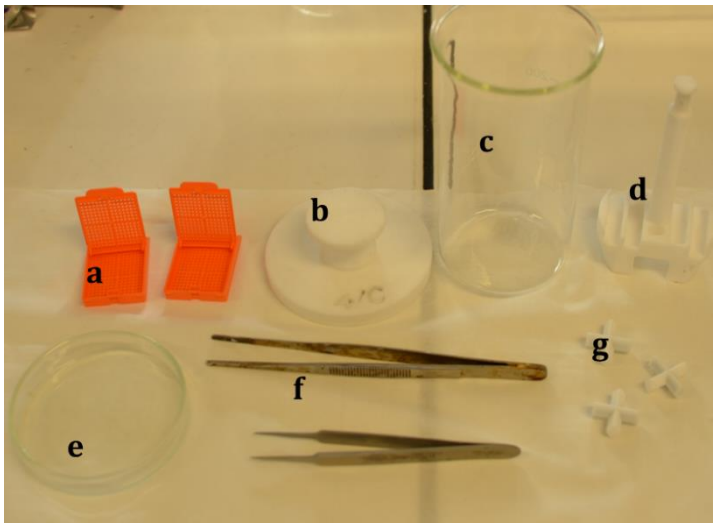
Note 5: Reduce electrostatic deleterious effects. Static electricity might cause LM dissected samples to move out cap. Try to keep static electricity as low as possible by maintain hygrometry above 60% in LM room and treat all LM material by exposing it under UV light for 30 minutes. UV should prevent root section to drop from the slide. A correct laser setting is also critical too to avoid static electricity accumulation.

Note 6: Use as less methanol as you can when mounting the slide. Usually we used a pipette and less than 200 µl of methanol per slide. The drying time is shorten and it prevent RNA degradation and root section to drop from the slide. Proceed to the xylene step when slide is completely dry.

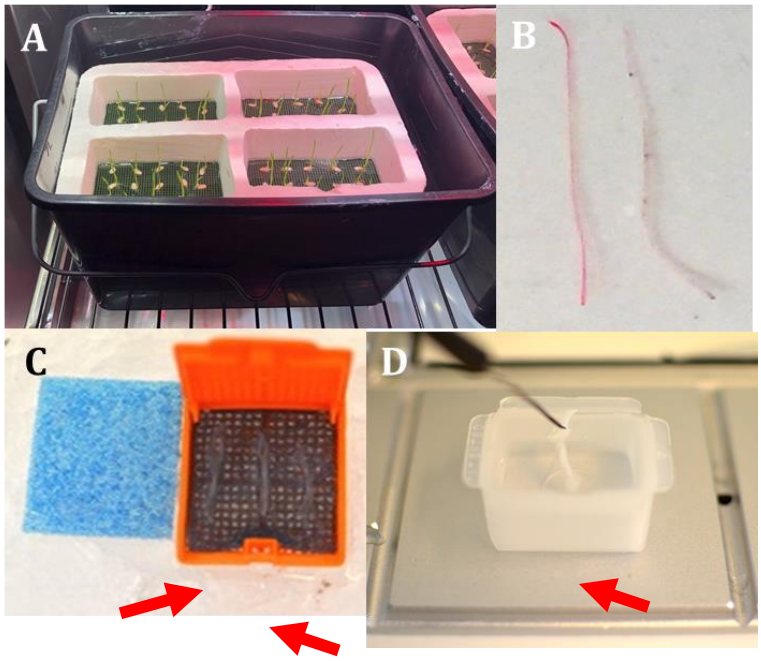
Note 7: Adjusting the laser settings is key to microdissect properly. Refer to the microscope manual and find the best set up for the samples.

Note 8: We never go beyond the two-hour limit to collect samples in a single collecting tube to preserve RNA quality. It is nevertheless possible to pool several tubes filled with a low volume of extraction buffer in order to increase RNA concentration. Adjust the extraction buffer volume in regard of the amount of tissue if you want to increase RNA yield.

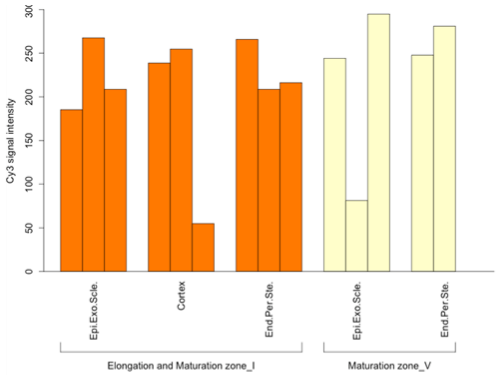
Note 9: We never go beyond four-hour of microdissection per a slide to prevent RNA degradation. It is better again to pool samples than to increase microdissection time of a single slide.



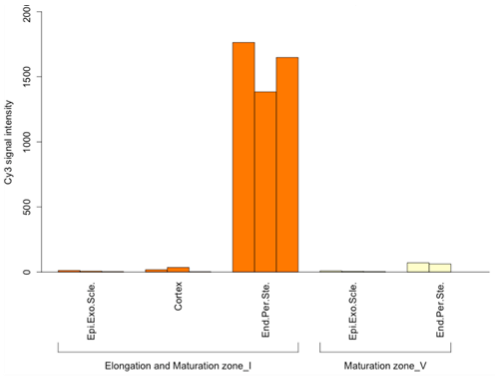
Supplemental Figure S1: RNAse Free material to be prepared before dehydration and inclusion of samples. a) histological cassettes, b), c) and d) elements of the water bath for the microwave. b) lid, c) beaker, d) fixing system for histology cassettes. e) glass petri dish. f) tongs and g) stirrers.



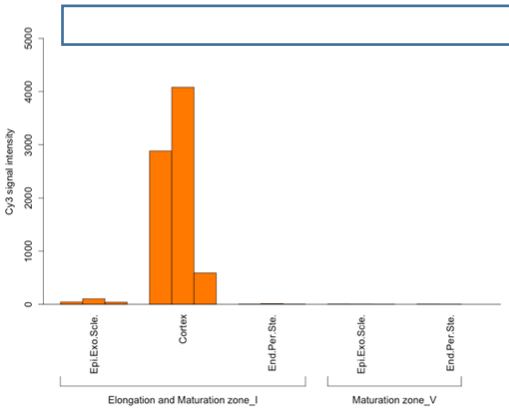
Supplemental Figure S2: Preparation of root tip samples. A) Hydroponic culture system using a floating net. B) Harvesting of 2 cm length root tips, staining of 1/5 root tip with eosin for positioning root fagots (red arrox). C) Inclusion cassette (right) of root tips that are covered with biopsy foam (left). Inclusion cassettes containing the roots after dehydration are placed in the water bath (**see also Supplemental Figure 1**). D) Beginning of soaking of the root tips in the paraffin and embedding of the whole root fagot in the paraffin



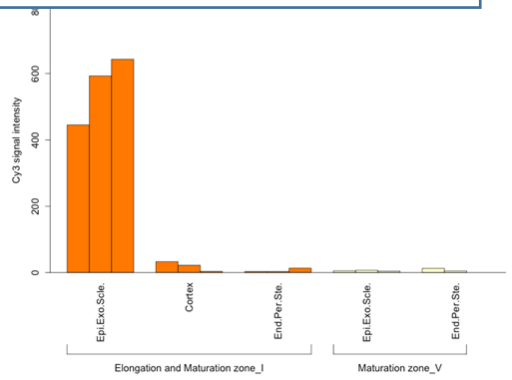
LOC_Os01g70380 Serine palmitoyltransferase



LOC_Os08g44750 Nodulin-like protein



LOC_Os10g18820 Plant disease response prot.



LOC_Os10g39890 Pollen Ole 1 allergen

Supplemental Figure S3: Expression profiling of the putative tissue specific genes extracted from RiceXpro

RAPD ID	MSU ID	Gene function	Tissu Specificity	Signal
Os10g0334500	LOC_Os10g18820	Plant disease resistance response protein family protein.	Cortex	3000
Os12g0114800	LOC_Os12g02300	Similar to Lipid transfer protein.	Cortex	150
Os10g0544200	LOC_Os10g39750	Basic helix-loop-helix dimerisation region bHLH domain containing protein.	Cortex	300
Os07g0676600	LOC_Os07g47960	Helix-loop-helix DNA-binding domain containing protein.	Cortex	150
Os10g0335000	LOC_Os10g18870	Conserved hypothetical protein.	Cortex	7000
Os04g0515900	LOC_Os04g43560	Similar to NAM / CUC2-like protein.	Cortex	600
Os10g0333700	LOC_Os10g18760	Plant disease resistance response protein family protein.	Cortex	5000
Os01g0216000	LOC_Os01g11730	Lipolytic enzyme, G-D-S-L family protein.	Outer cell	2000
Os01g0216000	LOC_Os01g11730	Lipolytic enzyme, G-D-S-L family protein.	Outer cell	2000
Os04g0674800	LOC_Os04g57860	Similar to CEL1=CELLULASE 1 (Fragment).	Outer cell	600
Os03g0416200	LOC_Os03g30250	BRITTLE CULM1.	Outer cell	120
Os10g0546100	LOC_Os10g39890	Pollen Ole e 1 allergen and extensin domain containing protein.	Outer cell	500
Os01g0839900	LOC_Os01g62260	Thaumatin, pathogenesis-related family protein.	Stele+endodermis	10000
Os07g0638500	LOC_Os07g44450	Similar to Dirigent-like protein (Fragment).	Stele+endodermis	2500
Os11g0508600	LOC_Os11g31190	Similar to MtN3 protein precursor.	Stele+endodermis	4000
Os03g0165400	LOC_Os03g06940	Similar to Relative to SR12 protein (Fragment).	Stele+endodermis	1500
Os08g0114200	LOC_Os08g02220	Similar to CEL5=CELLULASE 5 (Fragment).	Stele+endodermis	3000
Os08g0561500	LOC_Os08g44750	Similar to Nodulin-like protein 5NG4.	Stele+endodermis	4000
Os11g0149400	LOC_Os11g05190	Phytosulfokines 2 precursor	Stele+endodermis	2500

Supplemental table 1: Genes with tissue specific expression extracted from RiceXpro.

4. RESULTATS ET DISCUSSION : COMPLEMENTS SUR LE DEVELOPPEMENT DU PROTOCOLE

Le protocole développé ici est basé initialement sur le protocole décrit par Takahashi *et al.* en 2010. Protocole que j'ai optimisé pas à pas afin d'atteindre une qualité et une quantité d'ARN suffisante. La qualité et la quantité d'ARN ont été évaluées grâce à un Agilent 2100 Bioanalyzer system en utilisant les puces Agilent RNA 6000 Pico Kit. **Le protocole est détaillé dans l'article au-dessus (Mounier *et al.*, soumis dans BMC).**

3.1 La température et le temps de traitement ont un effet sur la qualité de l'ARN.

Le protocole détaillé dans Mounier *et al.*, (ci-dessus) permet d'obtenir des échantillons de racines inclus en blocs de paraffine moins de 24 heures après récolte des échantillons. Les protocoles de LM diffusés jusqu'à présent en biologie végétale prennent entre 4 et 7 jours avant inclusion des échantillons en blocs de paraffine (Sakai *et al.*, 2018, Takahashi *et al.*, 2010, Gautam *et al.*, 2016). Ces modifications nous ont permis d'atteindre un RIN autour de 8 ce qui correspond à un échantillon de haute qualité (**Figure 17**, Shroeder *et al.*, 2006, Endrullat *et al.*, 2016, Kashofer *et al.*, 2013, Derveaux *et al.*, 2010). Parmi ces modifications dans les étapes de fixation et d'inclusion en parafffine, cinq ont été critiques :

- Les échantillons (pointes racinaires) ont été prélevés dans du RNA-later®(AM7020) avant fixation pour limiter la dégradation des ARN;
- Le temps de fixation au vide des échantillons a été réduit à 1x5 minutes (3x5minutes sur l'ancien protocole) pour maintenir au maximum l'intégrité des tissus racinaires;
- Le nombre d'étapes de déshydratation a été augmenté, nous faisons maintenant des pas de 5%, de 75% à 100%. Elles sont faites sur glace et non au micro-onde.
- Les temps d'enrobage des échantillons au micro-onde ont été diminués (plus de 5 heures initialement, nous sommes passé à 3x5minutes) également pour maintenir l'intégrité des tissus et la qualité des ARNs.

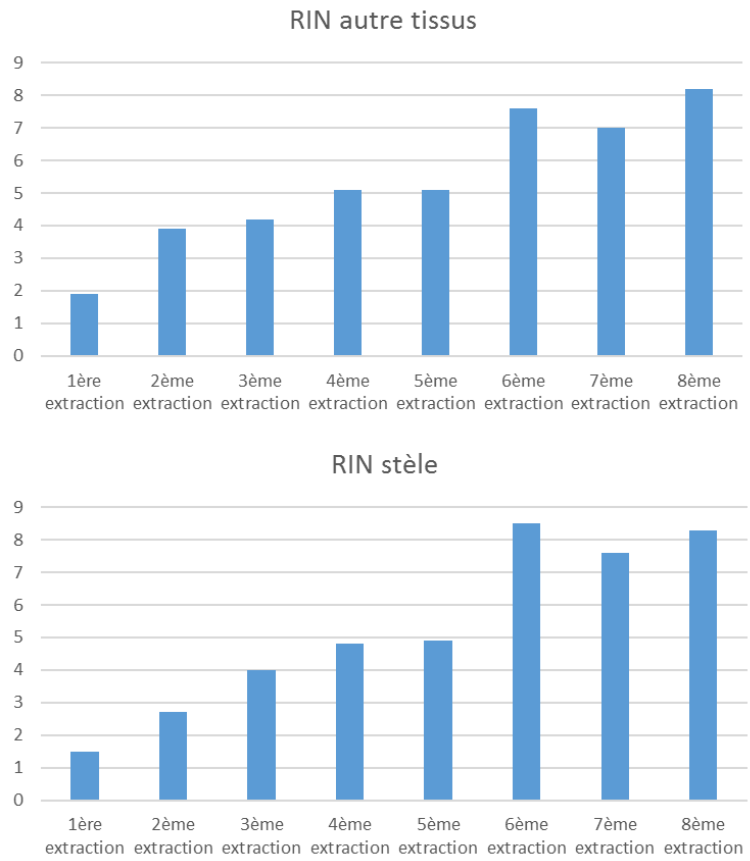


Figure 17: Evolution du RIN des ARN extraits après LM tout au long de l'optimisation du protocole

A partir de la 3ème extraction, les échantillons ont été découpés dans de l'eau RNase free traitée au RNAlater®.

A partir de la 4ème extraction, les temps de différentes étapes de notre protocole ont été diminués progressivement. Particulièrement les étapes de vides et celles au micro-onde.

A partir de la 6ème extraction, les étapes de déshydratation se font sur glace et non au micro-onde.

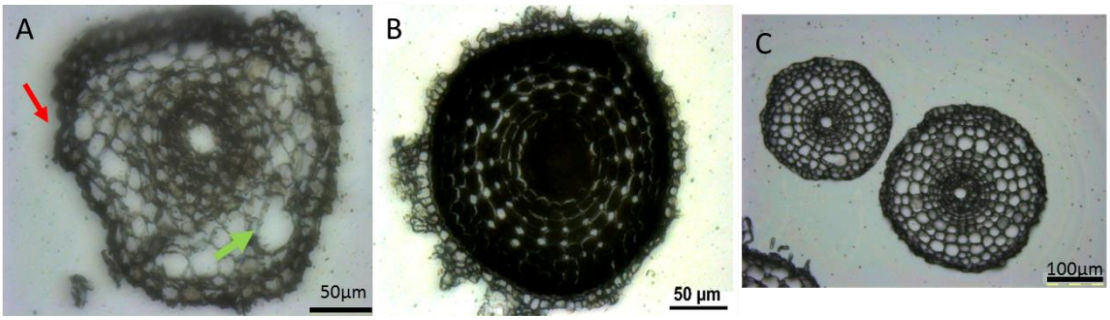


Figure 18: Effet du traitement mécanique sur la qualité histologique des échantillons.
A Coupe racinaire d'échantillons ayant « flottés librement » dans les bains successifs de déshydratation et d'imprégnation. Les 3 derniers tissus sont collapsés (flèche rouge) et on peut observer des zones où les tissus sont arrachés (flèche verte). B et C Coupe racinaire d'échantillons maintenus entre 2 pains de mousse à biopsie dans une cassette d'inclusion. Les tissus sont intacts.

3.2 La qualité histologique est dépendante des contraintes mécaniques auxquelles sont soumises l'échantillon

Malgré l'amélioration de la qualité des ARN (RIN à 8), la qualité histologique des coupes était initialement insuffisante pour une microdissection fiable des différents tissus (**Figure 18A**). Nous avons fait alors l'hypothèse que les contraintes mécaniques auxquelles sont soumises les racines lors des traitements des échantillons altérait la qualité histologique des coupes (cf Mounier *et al.*, soumis, fin de la partie 2), en particulier lors de l'agitation des échantillons. Nous avons alors comparé des coupes de racines laissées libres dans les bains de déshydratation et d'inclusion, à des coupes racines traitées et immobilisées entre deux pains de mousse à biopsie (**Figure 18B**). Cette dernière permet l'immobilisation et le fagotage des racines avant d'effectuer les différents traitements. Moins de 15% des sections de racines traitées sans mousse à biopsie ont leurs tissus externes intacts, contre plus 85% pour les sections de racines traités entre deux pains de mousse (**Figure 19A**). Moins de 5% des sections de racines traitées sans mousse à biopsie sont rondes, similaires aux échantillons avant traitement, pour environ 80% pour les sections de racines traitées entre deux pains de mousse (**Figure 19B**).

3.3 Les limites du LM

Dans le but d'explorer les limites techniques du LM, nous avons tenté de microdisséquer des tissus monocouches, ici l'endoderme (**Figure 20**). Nous avons réalisé 3 essais avec un nombre croissant de découpes et donc différentes surfaces de microdissection (114469 µm², 72978 µm², 79290 µm²). Nous nous sommes assurés que les échantillons microdissequés tombaient dans le capuchon des tubes de collecte. En effet, leur faible surface, les rend particulièrement sensibles à l'électricité statique (communication personnelle d'Anne Roscoff, spécialiste microdissection chez Leica). Les surfaces collectées sont équivalentes aux surfaces collectés pour la stèle ou le cortex et pour lesquelles nous avions réussi à extraire et détecter de l'ARN. En contrôle, nous avons par exemple disséqué 125259 µm² de stèle, qui est un tissu complexe multicouche, et nous avons obtenus 24,3 ng d'ARN avec un RIN de 4,7.

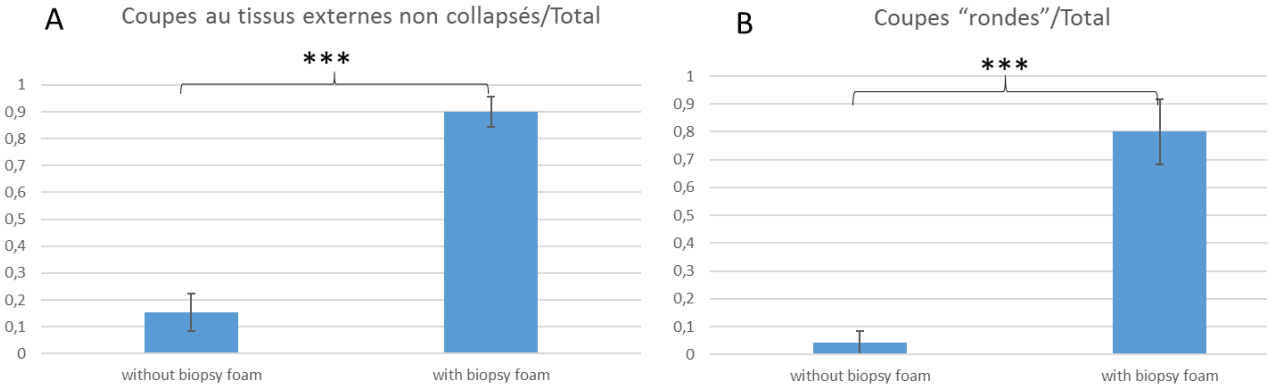


Figure 19: Effet du traitement avec ou sans mousse à biopsie sur la qualité histologique des coupes

A: Effet du traitement mécanique sur l'intégrité des tissus externes de la racine de riz: ratio du nombre de coupes racinaires dont les tissus externes sont collapsés sur le nombre total de coupes.

B: Effet du traitement mécanique sur la déformation de la racine de riz: ratio du nombre de coupes racinaires intactes sur le nombre total de coupes.

La moyenne et l'erreur standard sont calculés à partir de 4 répétitions biologiques sur lesquels 100 coupes racinaires ont été observés t-test, $p<0,001$***).

Lorsque nous avons découpé 60601 µm² de surface de stèle+endoderme, nous avons obtenus 9,4 ng d'ARN avec un RIN de 6,10. Malgré les surfaces équivalentes découpées en moncouche (114469µm², 72978µm², 79290 µm²), nous n'avons pu détecter aucun ARN dans ces 3 échantillons indépendants. L'hypothèse la plus vraisemblable, c'est qu'il est impossible d'obtenir de l'ARN à partir de tissu monocouche car le laser détruit/interfère avec la récupération d'ARN. Cet inconvénient, est aussi un avantage car il garantit l'absence de contamination intercouches lors des découpes au LM. Il serait intéressant de tester d'autres conditions, voir d'autres laser pour déterminer s'il est possible d'aller au-delà et d'aboutir à la découpe monocouche pour séparer par exemple l'endoderme de la stèle et du cortex.

Il semble donc que nous ayons atteint une limite technique de la technologie LM. A ma connaissance, aucun protocole de LM permettant l'isolation d'un tissu monocouche n'a été reporté chez les plantes.

3.4 Un protocole LM adapté à différents types d'expérimentations

La méthode d'enrobage en paraffine permet généralement une meilleure conservation histologique des tissus que la cryoconservation et permet également une conservation des échantillons sur plusieurs mois à 4°C. Ainsi, nous avons pu parfaitement conserver la qualité des ARNs extraits à partir de coupes histologique de racines de riz incluses en paraffine et stockées pendant 8 mois à 4°C (Mounier et al, soumis, cf ci-dessous). La rapidité d'exécution du protocole, la qualité des ARN et la possibilité de conservation des échantillons à long terme permet d'imaginer des campagnes d'échantillonnages, même au champ, avec une possibilité d'exploiter les échantillons sur plusieurs mois. De plus, comparé aux FACS ou à l'INTACT, ce protocole de LM peut être utilisé sur n'importe quelle variété (Birnbraum *et al.*, 2003, Deal and Henikoff, 2010; Deal and Henikoff, 2011). Il a d'ores et déjà été ré-utilisé avec succès dans notre équipe sur le cultivar Azucena cultivé en semi-hydroponie (terre+eau) pour comparer les profils d'expression au niveau du cortex après inoculation avec un biostimulant (données non montrées).

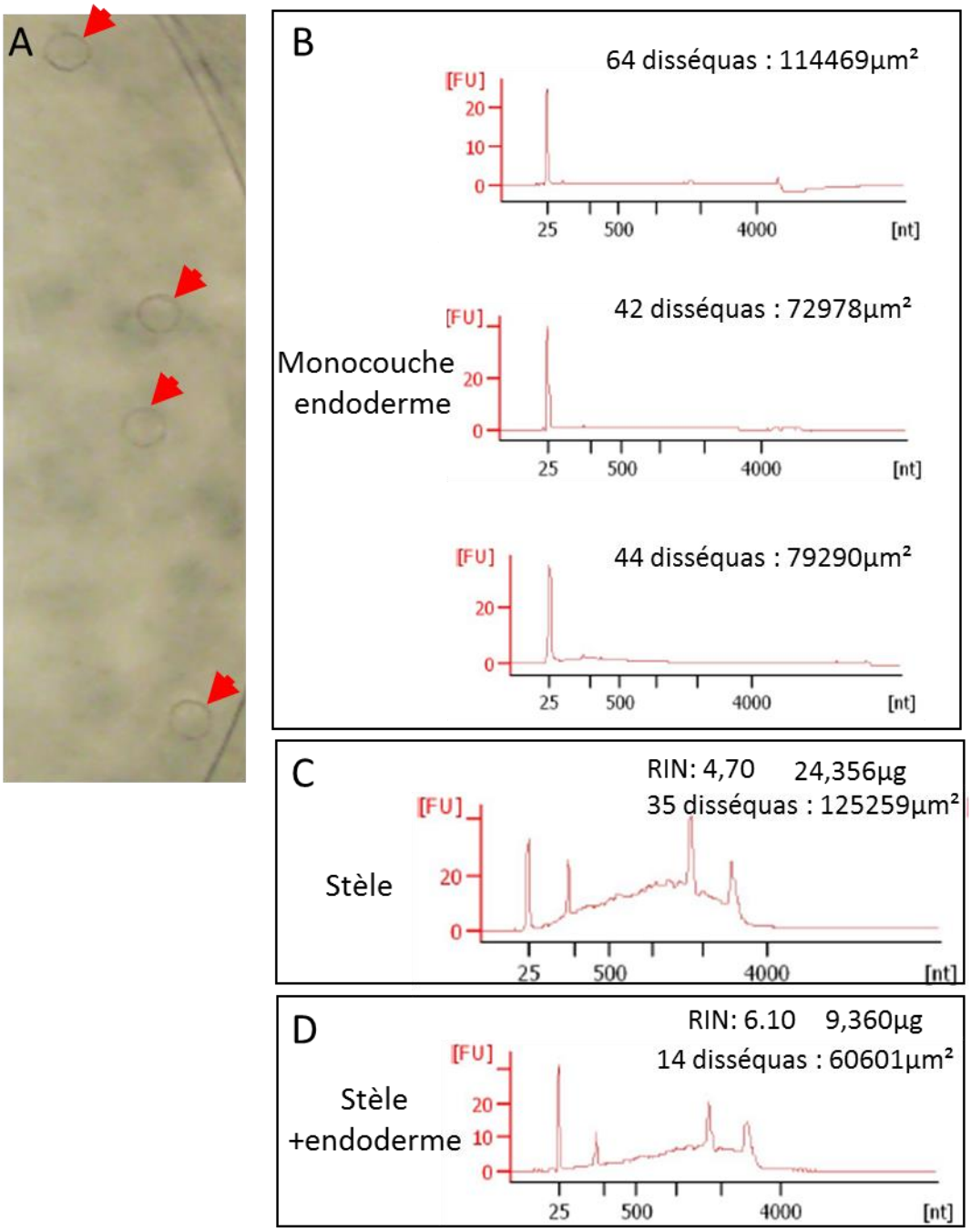


Figure 20: Test de dissection de tissus Monocouches

A. Visualisation des microdisséquas (flèche rouge) de monocouches d'endoderme dans le tube de collection après micro-dissection.
B. Analyse des quantités d'ARN au bioanalyzer après extraction sur microdisséquas de monocouches d'endodermes. Aucune présence d'ARN n'est observée. 3 échantillons biologiques indépendants ont été découpés.
C. Analyse des quantités et de la qualité d'ARN après extraction sur microdisséquas de stèle sur le même échantillon.
D. Analyse des quantités et qualité d'ARN après extraction sur microdisséquas de stèle+endoderme.

Les surfaces découpées sont données en µm², la qualité d'ARN est exprimée en RIN.

3.5 Le couplage LM/ddPCR, vers une quantification fiable et rapide de l'expression génique dans des groupes de tissus spécifiques (voir l'article intégré, Mounier *et al.*, soumis)

Les quantités d'ARN provenant des échantillons disséqués par LM sont souvent faibles ou alors exigent la collection de centaines de sections ou bien de groupes de tissus d'une grande taille (Takahashi *et al.*, 2010, Gautam *et al.*, 2016, Sakai *et al.*, 2018), ce qui complique l'utilisation en routine de cette technologie. Nous avons montré que la technologie RT-ddPCR est capable de détecter l'expression d'un gène de référence, Exp', à partir de 100 fg d'ARN. Nous avons utilisé cette technologie pour quantifier l'expression de 3 gènes, Exp', 5NG4 et SHR1, dans des échantillons dilués en cascade de 1 fg à 1000 pg. Il existe une très bonne **linéarité entre la quantité d'ARN et le nombre de copies détectées par microlitre par RT-ddPCR. Les seuils de détection dépendent du niveau d'expression de chaque gène, mais permettent de travailler avec des quantités supérieures ou égales à 100 fg d'ARN.**

Habituellement, la LM est couplée à la RT-qPCR afin de comparer l'expression d'un ou plusieurs gènes entre deux ensembles de cellules (Anjam *et al.*, 2016, Rahji *et al.*, 2010, Ludwig et Hochholdinger, 2013). Cette technologie est moins sensible que la RT-ddPCR qui en théorie, permettrait la détection d'une seule molécule d'ARN dans un échantillon (Mounier *et al.*, soumis, ci-dessous, Dreo *et al.*, 2014, Rački *et al.*, 2013). Nous avons utilisé cette technologie avec succès pour valider l'expression tissu-spécifiques de 3 gènes issus de la base de données Rice-Xpro et de OsSHR1 (Takehisa *et al.*, 2012, Mounier *et al.*, soumis, ci-dessous, Cui *et al.*, 2007). **Ainsi la combinaison de la technologie LM à la ddPCR devient une alternative intéressante à l'utilisation de technologies telle que l'hybridation *in situ* et la RT-PCR *in situ*.**

5. Conclusions du chapitre 2

Nous disposons maintenant d'un protocole rapide et fiable permettant l'inclusion en paraffine d'échantillon de racines de riz pour la microdissection laser. Si ce protocole a été initialement développé pour la LM, il peut être également utilisé pour des technologies alternatives ou complémentaires comme l'hybridation *in situ.* L'enrobage en

paraffine permet une conservation des échantillons sur plusieurs mois à 4°C, ce qui permet d'envisager de collecter un grand nombre d'échantillons puis de les stocker pour leur exploitation ultérieure. Le protocole publié est très détaillé, comprend la liste complète des produits et appareils utilisés ainsi qu'un trouble shooting lui aussi détaillé afin de faciliter la réutilisation de cette technologie chez le riz mais également chez d'autres espèces de plantes.

Chapitre 3

Construction de banques d'ADNc tissu et stade de développement spécifique

Chapitre 3 : Construction de banques d'ADNc tissu et stade de développement spécifique

1. Introduction

Un état de l'art rapide montre que quelques études ont permis de réaliser des transcriptomes racinaires tissus spécifiques (**http://ricexpro.dna.affrc.go.jp**, Takehisa *et al.*, 2012, http://bioinformatics.med.yale.edu/riceatlas/overview.jspx, Jiao *et al.*, 2009) et que des données moléculaires (pour une revue : Yamauchi *et al.*, 2018, Yamauchi *et al.*, 2015, Yamauchi *et al.*, 2017 (a), Yamauchi *et al.*, 2017 (b)) existent aussi chez le riz sur des acteurs de la formation des aérenchymes (ROS, éthylène, acides gras à très longues chaines…). Dès lors, pourquoi se lancer dans l'obtention d'un transcriptome cortex spécifique *de novo* pour étudier la formation des aérenchymes racinaires chez le riz ?

1.1 Les données moléculaires sur les acteurs de formation des aérenchymes portent exclusivement sur les aérenchymes *induits*

La formation des aérenchymes est un processus à la fois inductible et constitutif chez le riz. Les aérenchymes racinaires se forment en l'absence de submersion et les riz pluviaux ont également des aérenchymes. De nombreuses études ont démontré que les carences nutritionnelles (Abiko et Obara, 2014, Pujol et Wissuwa, 2018), le stress hydrique (Yang *et al.*, 2012) et bien sûr la submersion accéléraient la formation des aérenchymes (Suralta et Yamauchi, 2008), qui se forment alors beaucoup plus tôt et plus proche des méristèmes racinaires. Malgré la présence d'aérenchyme constitutif chez le riz, la plupart des études moléculaires récentes portent sur la formation des aérenchymes induits par un stress abiotique ou un traitement chimique (Shiono *et al.*, 2014, Yamauchi *et al.*, 2O17(a), Steffens *et al.*, 2011, Yamauchi *et al.*, 2018, Yamauchi *et al.*, 2016, Colmer *et al.*, 2006). Deux problèmes se posent alors :

- Nous ignorons si les mécanismes de formation sont les mêmes entre aérenchymes inductibles et constitutifs. Il est possible de faire l'hypothèse, raisonnable, qu'un certain nombre d'acteurs moléculaires sont communs mais rien ne permet d'affirmer que le même processus soit en cause. En effet, chez d'autres céréales, comme le sorgho et le maïs, il existe également des aérenchymes mais qui sont exclusivement

inductible en condition de stress. Il reste donc possible que ce soit deux processus génétiquement séparables.

- Les approches de différentiel d'expression se basant sur la formation d'aérenchymes inductibles ont essentiellement été menées sur des échantillons de racines entières, ce qui ne permet pas d'apprécier la réponse transcriptomique spécifique du cortex au stress imposé à la plante, qui doit par nature être différente de la réponse apportée par les autres tissus ne formant pas d'aérenchymes. Isoler les cellules corticales des autres tissus est donc un premier pas pour comprendre les mécanismes moléculaires localisés spécifiquement dans ce tissu lors de la formation des aérenchymes.

1.2 Les données tissues spécifiques publiées jusqu'à présent sur les céréales ne permettent pas l'investigation sur la dynamique développementale aboutissant à la formation des aérenchymes

Deux bases de données tissu-spécifique pour le riz sont d'ores et déjà disponible (**http://ricexpro.dna.affrc.go.jp**, Takehisa *et al.*, 2012, http://bioinformatics.med.yale.edu/riceatlas/overview.jspx, Jiao *et al.*, 2009). La base de données Rice-Xpro propose même un transcriptome du cortex sur différentes zones définies par leurs distances à la pointe racinaire. Cependant, ces études ne portaient pas spécifiquement sur les stades de développement du cortex et des aérenchymes.

Une première étude sur la réponse transcriptomique de la racine à la submergence a également été publiée, mais elle se concentre essentiellement sur la formation de la barrière à la diffusion radiale de l'oxygène en condition anaérobie et il ne contient aucune description histologique des stades développementaux utilisés (Shiono *et al.*, 2014). Yamauchi *et al.*, en 2017, ont réalisés des RT-qPCR sur de l'ARN extrait de cellules corticales de racines de riz. Ici aussi, ils comparent des plantes après une croissance en condition stagnante (induisant la formation des aérenchymes) ou une croissance en condition aérée. La publication se concentre essentiellement sur l'expression différentielle de 15 gènes codant pour des Métallothionéines (MT) par RT-PCR, elle ne donne donc qu'un aperçu limité du paysage transcriptomique dans les cellules corticales et les tissus adjacents au cours de la formation des aérenchymes.

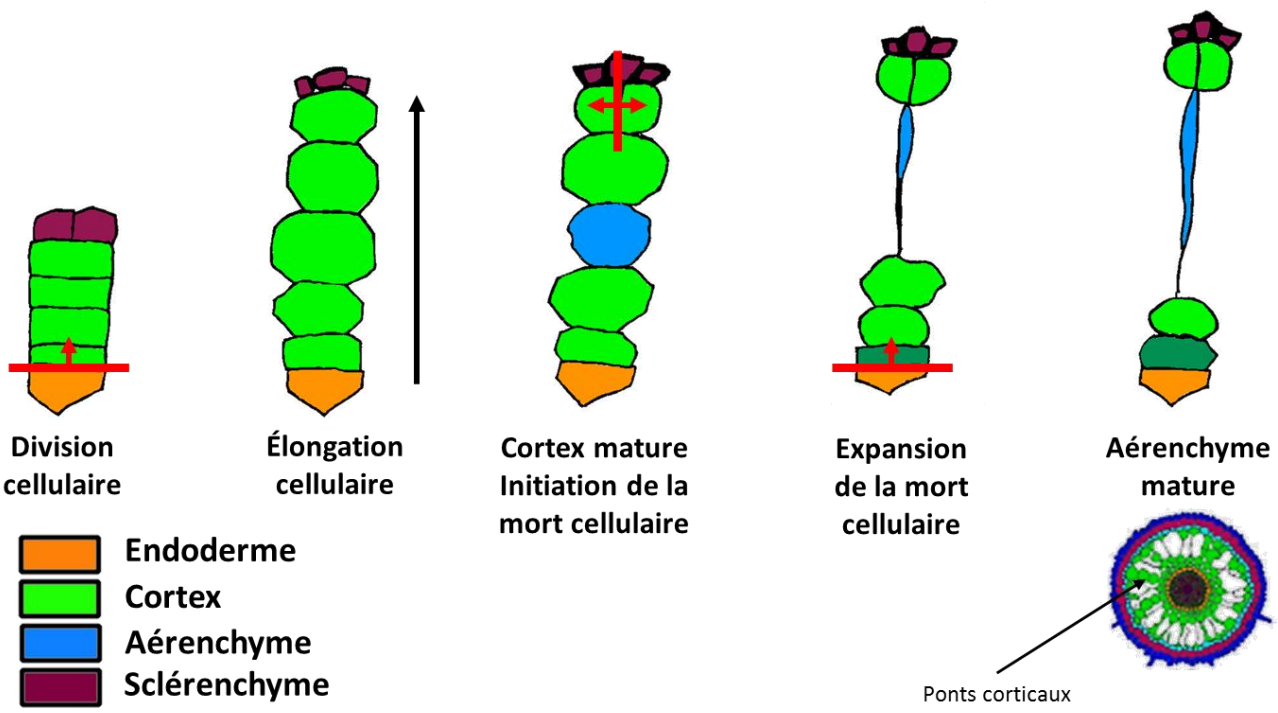


Figure 21: Développement d'un aérenchyme mature, quatre phases de développement du cortex (Kawai et al., 1998)

La formation des aérenchymes à partir des cellules corticales racinaires peut être décomposé en 4 phases. Les flèches rouges représentent le plan dans lesquels les cellules se divisent. La flèche noire représente l'extension radiale des cellules corticales

Dans le cas de l'étude des aérenchymes, une première étude transcriptomique est parue sur l'isolation de cellules corticales provenant de plant de maïs soumis à différents traitements (Rajhi *et al.*, 2011), submergence et/ou traitement avec l'inhibiteur de perception de l'éthylène (1-MCP). Cependant, là encore nous n'avons pas d'informations précises quant au stade de développement des tissus microdisséqués. De plus, les tissus corticaux ne sont pas comparés entre eux directement, le différentiel d'expression se faisant entre cortex et stèle dans un premier temps, puis les gènes présentant une expression différentielle entre ces tissus, sont ensuite comparés pour les différentes conditions de traitement.

Ces études nous révèlent donc le profil d'expression des gènes en fonction des tissus, et potentiellement la réponse de ces tissus à un traitement, mais ne permettront pas de déterminer quels sont les éléments clefs de la différenciation du cortex en aérenchymes. En effet, le fait qu'il n'y ait pas d'aérenchymes dans le cortex proche des zones méristématiques suggère que les cellules corticales acquièrent la capacité de former des aérenchymes au cours de leur développement. En conclusion, seul un transcriptome tissu-spécifique réalisé à partir de différents stades de différenciation précoces du cortex en aérenchymes est à même de permettre d'identifier les déterminants moléculaires de la formation des aérenchymes constitutifs chez le riz.

1.3 Du développement du cortex à la formation d'aérenchymes matures : un processus en 4 phases

La formation des aérenchymes chez le riz passe par quatre stades de développement distinguables histologiquement (Kawai *et al.*, 1998, **Figure 21**, cf chapitre 1 et introduction). Dans une première phase, appelée phase de formation des cellules corticales, la cellules souche CEEID se divise plusieurs fois périclinalement afin de former le cortex, tissu multicouche, et l'endoderme (**Figure 7B, Figure 21**). Vient ensuite la phase d'élongation cellulaire radiale pendant laquelle les cellules du cortex médian s'allongent radialement et leur contenu cellulaire s'acidifie, comparativement aux cellules corticales adjacentes. La troisième phase, la phase où le cortex est mature, est celle de l'initiation de la mort cellulaire au niveau des cellules corticales médianes. La dernière phase correspond à une extension radiale de la mort cellulaire aux cellules voisines. A la fin du processus, il ne reste que quelque couches cellulaires vivantes dans le cortex, deux

couronnes de cellules sur les faces intérieurs et extérieurs des aérenchymes, reliés entre elles par des « ponts « « corticaux » de cellules vivantes (**Figure 21**) qui permettent d'assurer la cohésion des tissus racinaires et le transfert radial d'eau et de nutriments.

1.4 Laser microdissection (LM), avantages et limites techniques

La microdissection laser est l'unique technologie permettant une récupération d'un échantilon de cellule après observation de l'échantillon. Ainsi, au-delà de la dissection tissu-spécifique, elle permet de distinguer les stades développementaux afin de pouvoir établir des transcriptomes différentiels avec une dynamique développementale (Nelson *et al.*, 2006, Nichterwitz *et al.*, 2016).

Une première limite technique à la récupération de stades développementaux précis et tissus spécifique par LM est le nombre limité de sections qui peuvent être récoltées, ce qui a pour conséquence une faible quantité d'ARN disponible après extraction. De plus, les précédents transcriptome tissu spécifique chez le riz ont été établis grâce à des puces à ADN qui nécessitent de grandes quantités d'ARN, particulièrement lorsqu'il s'agit de détecter l'expression de transcrits peu abondants (Wang *et al.*, 2014, Hoen *et al.*, 2012, Nygaard et Hovig 2006) et/ou nécessitent des étapes de préamplification des ARNs qui induisent potentiellement des biais. Ainsi, en routine, un minimum de 10ng d'ARN est nécessaire pour l'utilisation des puces à ADN, ce qui en pratique requiert plusieurs dizaines à centaine de sections pour certains tissus comme la stèle (Kulichikin *et al.*, 2014, Takehisa *et al.*, 2012, Takahashi *et al.*, 2010, Rajhi *et al.*, 2011). Une autre raison qui nous a amené à vouloir développer un transcriptome *de novo* : les précédentes études ont été réalisées à une époque où les annotations du génome du riz, sur lesquelles se sont basées les puces à ADN affymetrix utilisées, étaient partielles, incomplètes et ne représentaient donc qu'un paysage incomplet des gènes exprimés dans ces tissus.

1.5 Coupler la technologie LM et le RNA-seq, contraintes causées par la faible quantité d'ARN extraite

Une alternative aux puces à ADN est le RNA-seq qui requiert une quantité plus faible d'ARN. Néanmoins, cette technologie adoptée par de nombreuses plateformes peut être source de biais d'analyse, notamment lors de la construction de la banque d'ADNc à partir

d'une faible quantité d'ARN (pour des revues Hrdlickova *et al.*, 2016, Dijk *et al.*, 2014). Ces biais peuvent être introduits à différentes étapes, et ont essentiellement deux causes, 1) la perte d'information quant au brin réellement transcrit dans la cellule (Levin *et al.*, 2010), et 2) des étapes PCR avec une amplification non linéaire des transcrits (Fu *et al.*, 2014, Casbon *et al.*, 2011, Shiroguchi *et al.*, 2012).

Une analyse approfondie des technologies de création de banque d'ADN complémentaire (ADNc) proposées nous a rapidement conduit à conclure qu'aucune n'était réellement adaptée à nos conditions (déplétion des ARN ribosomaux, librairie brin spécifique, indexage des transcrits, amplification linéaire de la banque d'ADNc, pour des revues Hrdlickova *et al.*, 2016, Dijk *et al.*, 2014), en particulier en partant de quantités d'ARN de l'ordre de quelques centaines de pico-grammes. Néanmoins, la société Nugen propose un kit nommé Ovation® SoLo RNA-Seq Library Preparation qui propose tous ces traitements et annonce une quantité minimale de 10pg d'ARN total en entrée. Disponible uniquement chez *A. thaliana* pour les plantes, nous avons décidé de l'adapter pour le riz.

1.6 Limiter les biais générés lors de la création de la banque d'ADNc

Une seule technologie disponible sur le marché permet, à partir de quelques centaines de picogrammes d'ARN, de générer des banques d'ADNc brins spécifiques avec déplétion des ARN ribosomaux (ARNr) : le kit Ovation® SoLo RNA-Seq Library Preparation (Dueck *et al.*, 2016, Adiconis *et al.*, 2013, Shanker *et al.*, 2015, Masters *et al.*, 2018, sur échantillons LM : Faris *et al.*, 2017). Ces deux derniers points permettent respectivement, 1) une meilleure analyse des isoformes d'ARN puisque l'on saura précisément quel brin est exprimé, sens ou anti-sens, et 2) une optimisation de la profondeur de séquençage, la majorité des séquences lues sont alors des ARN messagers (ARNm) la plupart des ARNr est alors éliminées par l'étape de déplétion ; les ARNr représentent environ 90% des espèces d'ARN présentent dans une cellule (O'Neil *et al.*, 2013, Mills *et al.*, 2013, Levin *et al.*, 2010). La technologie de Nugen propose aussi deux étapes supplémentaires pour éviter les biais générés lors de l'amplification des banques d'ADNc avant envoi au séquençage, une indexation moléculaire des transcrits (Fu et al., 2014) et une amplification linéaire des ADNc (Kurn et al., 2005).

2. Objectifs

L'objectif ici est le développement d'un protocole permettant de réaliser des banques RNAseq à partir de quantités d'ARN faibles et adapté à l'étude du développement des aérenchymes du riz. Le protocole LM développé dans la partie 2 (Mounier *et al.*, soumis) a été utilisé pour disséquer le cortex sur les 4 stades développementaux aboutissant à la formation des aérenchymes racinaires (Kawai *et al.*, 1998). En parallèle, deux autres groupes de tissus, le cylindre central (stèle+endoderme) et les trois tissus externes (sclérenchymes/exoderme/épiderme), ont été également microdisséqués en portant une attention particulière au stade de développement de ces tissus. Afin de pouvoir effectuer un séquençage d'ADNc à partir des petites quantités obtenues, le kit Ovation® SoLo RNA-Seq Library Preparation a été optimisé afin de pouvoir être utilisé sur le riz. En effet, la déplétion des ARNr avec ce kit nécessite un développement pour chaque espèce et n'est disponible commercialement que pour *A. thaliana*. Enfin, à l'aide de ce kit customisé, nous avons construit un total de 36 banques (3 tissus x 3 stades et 3 répétitions) qui ont été envoyés au séquençage pour réaliser un transcriptome différentiel (séquençage et analyses en cours au moment de la remise du manuscrit). Les technologies proposées par ce kit afin de limiter l'introduction de biais lors de la création de la banque d'ADNc seront discutées.

3. Matériels et méthodes

3.1 Microdissection laser des trois tissus

Le protocole de LM utilisé ici est celui développé dans Mounier *et al.*, soumis (cf partie 2). Les plants de riz utilisés ont été cultivés en hydroponie non-aérée pendant 7 jours (cv Nipponbare). Les radicules de plants de riz de 2 cms sont alors incluses en paraffine et coupées au microtome en section de 12µm à partir de la pointe de la racine jusqu'à environ 2mm de l'extrémité de la coiffe. Les stades de développements du cortex sont inspectés visuellement grâce au dispositif Leica LMD7000® (LMD7000, Leica, USA). Environ une dizaine d'échantillons par tube collecteur sont microdisséqués à l'aide du LM à l'aide du grossissement 63x (puissance 22, aperture : 1, vitesse : 8, Balance : 20, Head current 80%, pulse frenquency : 228, Offset ; 210) ou 40x (puissance 21, aperture : 1, vitesse : 10, Balance : 25, Head current 100%, pulse frenquency : 120, Offset : 180).

3.2 Extraction des ARN, évaluation de la qualité et de la quantité

L'extraction des ARN a été effectuée en suivant le protocole du kit PicoPure® RNA ; extraction avec un traitement DNase sur colonne (RNase-Free DNase Set, Cat no. 79254, Qiagen, Germany). Une légère modification y a été apporté afin d'optimiser l'extraction sur de très petites quantité d'ARN : des sections racinaires provenant de deux répétitions techniques ont été microdisséquées et récupérées dans 25µl de tampon d'extraction (voir le protocole du kit kit PicoPure® RNA isolation). Ces échantillons ont ensuite été assemblés avant extraction pour un volume final de 50µl. Le volume d'élution choisi est le minimum recommandé par Arcturus, 11µl. La quantité et l'intégrité de l'ARN a été évalué avec un bioanalyzer (Bioanalyzer system, Agilent 2100, no. DE72902360) en utilisant le kit RNA Pico 6000 (5067-1513).

3.3 Construction du kit Ovation® SoLo RNA-Seq Library Preparation, analyses des séquences ARNr chez le riz

Les séquences des ARNr nucléaires ont été extraits de la banque de données Rice Annotation Program développé par le NCBI, le numéro d'accession est le PRJNA122 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA122). Les séquences d'ARNr utilisées sont : XR_003238829.1 (ARNr5.8S), XR_003238828.1 (ARNr28S), XR_003238822.1 (ARNr18s), XR_003239795.1 (ARNr5S).

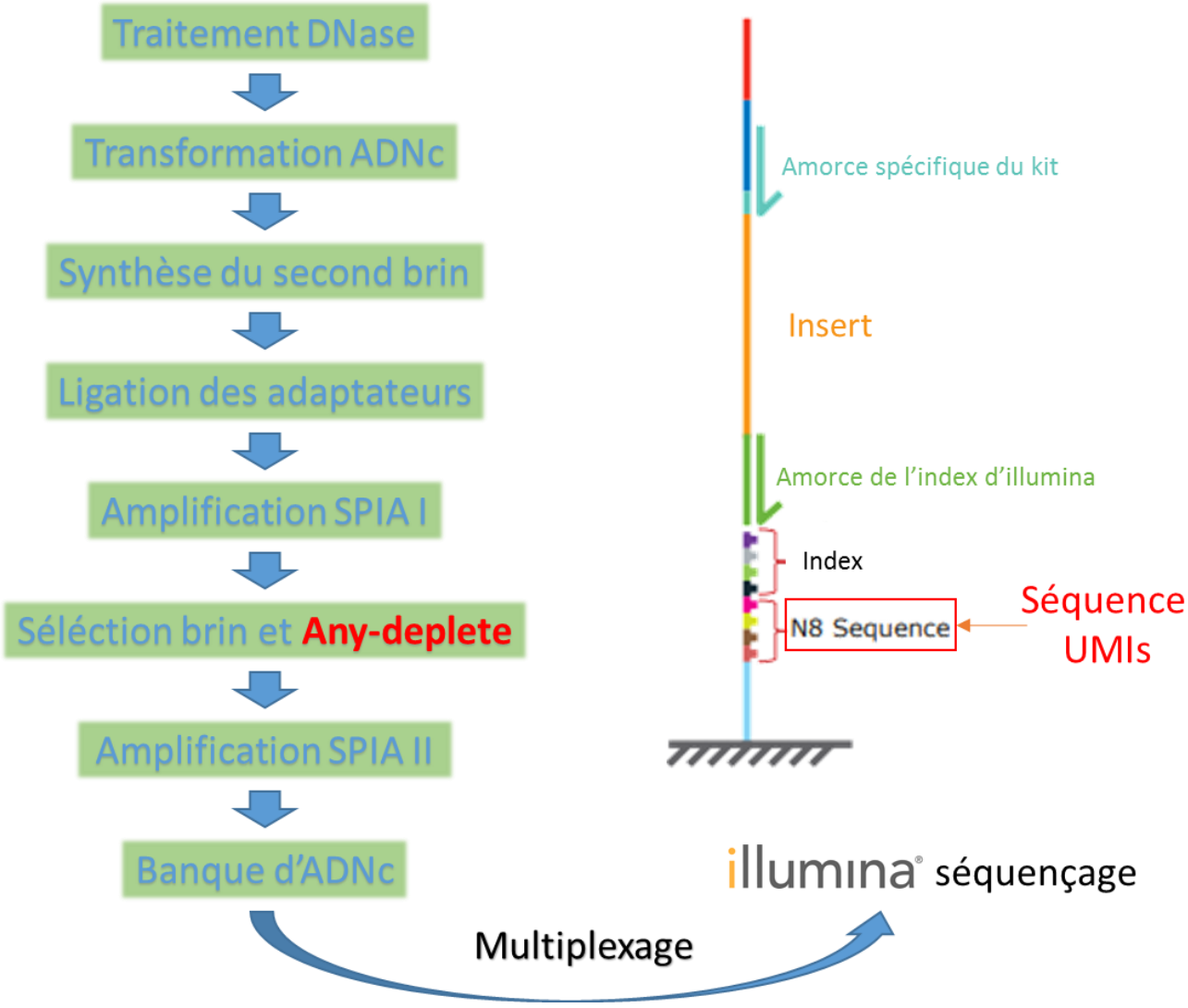


Figure 22: Représentation schématique du protocole du kit Ovation® SoLo RNA-Seq Library Preparation

Les technologies de sélection du brin et Any deplete sont décrites dans la figure 30 et 31. L'amplification SPIA permet une meilleure conservation de la stœchiométrie entre les transcrits. L'insert correspond à l'ADN complémentaire du brin sens de l'ARN transcrit dans nos échantillons. L'indexage des banque d'ADNc permet de les multiplexer avant séquençage. La séquence UMIs est unique et spécifique à chaque transcrit, elle permet de distinguer les doublons générés par les phases d'amplification.

Les séquences des ARNr chloroplastiques ont été extraites du site du NCBI, le numéro d'accession de la banque est MG252500.1. Les accessions d'ARNr utilisées sont: MG252500.1:91253-92743 (ARNr16S), MG252500.1:95166-98053 (ARNr23S), MG252500.1:98149-98243 (ARNr4.5S), MG252500.1:98471-98591 (ARN 5S).

Les séquences des ARNr mitochondriaux ont été extraites du site du NCBI, le numéro d'accession de la banque est KY486275.1. Les accessions d'ARNr utilisés sont les suivantes : KY486275.1:c490935-490814 (ARNr5S), KY486275.1:c492743-491049 (ARNr18S), KY486275.1:457475-461016 (ARNr26S).

Toutes les séquences sont fournies dans l'annexe 1.

3.4 Création des banques RNAseq et analyses de qualités

Le design des amorces de déplétion spécifiques au riz a été effectué par Nugen à partir des séquences d'ARNr que nous leur avons fourni, 836 amorces ont ainsi été générées (Adomas *et al.*, 2010, Hrdlickova *et al.*, 2016). Les banques RNAseq ont été générées en suivant le protocole du kit Ovation® SoLo RNA-Seq Library Preparation de Nugen (résumé du protocole dans la **Figure 22**) avec 1ng d'ARN total en entrée. Les deux contrôles conseillés dans le kit ont été utilisés: un contrôle sans ARN (contrôle négatif pour s'assurer qu'il n'y pas de contamination) et un contrôle avec 1ng d'ARN sans traitement Any deplete (contrôle de la qualité de l'appauvrissement en ARNr). Le nombre de cycles d'amplification nécessaires est évalué par qPCR comme conseillé par le protocole fournit. La qualité et la quantité des ADNc des banques générées ont été évaluées avec un bioanalyzer (Bioanalyzer system, Agilent 2100, no. DE72902360) en utilisant le kit RNA Pico 6000. Nous devions visuellement observer une répartition normale de la taille des transcrits un peu au-dessus de 300pb, d'après le protocole fournit avec le kit. La molarité des échantillons a été évaluée avec un test Takara Library Quantification Kit (638324)

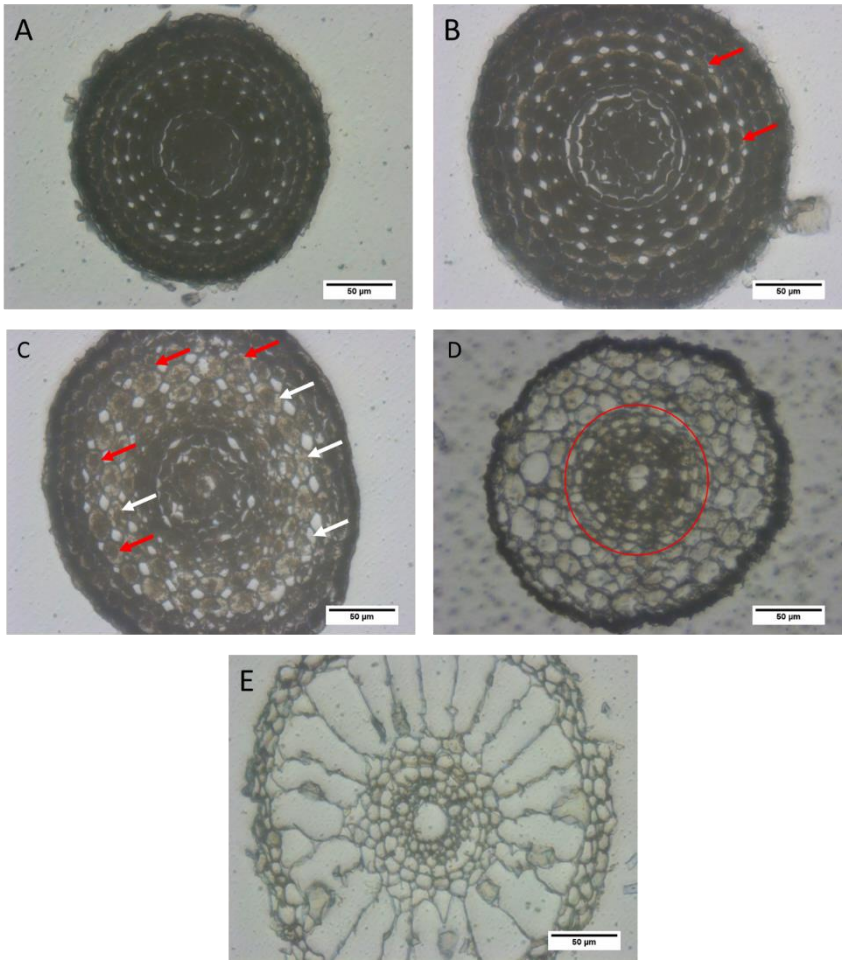


Figure 23: Description histologique des 4 zones de développement du cortex aboutissant à la formation des aérenchymes

A) Phase de formation du cortex: les cellules souches cortex/endoderme sont encore en divisions car le nombre maximal de couche de cortex n'est pas atteint. Le cortex est formé de lignées cellulaires disposée en rayon autour du cylindre central. Le contenu cellulaire est très dense. Les trois tissus externes sont déjà bien différenciés.

B. Phase d'élongation: Le nombre maximal de couches de cellules corticales est atteint, les cellules s'allonge radialement. Le contenu cellulaire est moins dense. Et des irrégularités dans la linéarité des lignées de cellules corticales au niveau de la couche la plus extérieure peuvent être observés (flèche rouge). Les pôles du xylème sont observables dans le cylindre central.

C. Phase mature du cortex: Les cellules corticales ont atteint un volume proche du maximum, le contenu cellulaire est peu dense. Certaines cellules de l'avant dernière couche de cellule corticale sont moins rondes, leurs contenus cellulaires est altéré (flèche blanche). Le phloème est visible dans le cylindre central.

D. Phase de propagation radiale de la mort cellulaire: Le contenu des cellules corticales est vidé. La morphologie des cellules du cortex devient irrégulière, à l'exception des deux ou trois couches de cortex entourant le cylindre central (cercle rouge).

E. Aérenchymes matures: Les parois des cellules corticales mortes forment des bandes disposées en rayon entre les tissus externes et 1 à 2 couches de cellules corticales survivantes entourent le cylindre central (flèche noire).

Les coupes racinaires sont observées au grossissement 40 en lumière blanche, la barre d'échelle représente 50µm.

4.Résultats

4.1 Microdissection des 4 stades de développement du cortex (**Figure 23**)

Les 4 phases de développement du cortex ont été microdisséquées par LM. Initialement nous avions aussi prévu de microdisséquer et d'extraire les ARNs dans une cinquième zone ou les aérenchymes sont matures (toutes les phases de développement prélevées sont décrites avec précision dans la **Figure 23**). Parmi tous les échantillons analysés, nous avons extrait de l'ARN en quantité et en qualité suffisante pour trois phase : la phase de formation du cortex, la phase d'élongation cellulaire et la phase d'initiation de la PCD (**Figure 25**). Les échantillons ont un RIN moyen de 7, avec un minimum à 5,5 et un maximum à 8,3. Les échantillons ont été élués dans 11µl ; par conséquent la quantité minimale d'ARN récolté est d'environ 1600 pg, et la quantité maximale d'environ 46000 pg. En pratique, néanmoins, environ 1µl d'échantillon sont perdus lors de l'élution. Après passage de l'échantillon au bioanalyzer, nous disposions d'environ 8µl d'échantillons à exploiter pour la création de la banque cDNA. Très peu, voir pas d'ARN ou de l'ARN dégradés ont été extraits dans la zone de propagation de la mort cellulaire, et ce dans les trois groupes de tissus microdisséqués (**Figure 24**).

4.2 Création des banques RNAseq, profils de répartition des fragments par taille

Le contrôle sans ARN, (contrôle négatif de contamination) ne présente aucune trace d'ADNc au bioanalyzer (**Figure 26C**). Les échantillons d'ADNc tissus et stades développement spécifiques présentent tous une répartition attendue de taille des fragments d'ADNc avec une moyenne un peu au-dessus de 300 pb (un électrophorégramme type est montré, **Figure 26A**). La concentration des banques d'ADNc générée est comprise entre 0,3 et 7 ng/µl avec une moyenne de 3,67 ng/µl (**Figure 27**). Le contrôle qui n'a pas subi l'étape d'appauvrissement en ARNr (Any-deplete) présente la même répartition normale autour de 300pb, cependant la concentration de la banque est de 25,5 ng/µl (**Figure 26B**). Les molarités des banques sont déterminées grâce à un test de qPCR par le kit Takara Library Quantification et les résultats sont visibles dans la **Figure 28**. Nous avons préparé un pool contenant nos 36 banques à 5nM en utilisant l'application pooling calculator d'Illumina (données non montrées), et ce pool a été envoyé au séquençage sur un HiSeq/NovaSeq.

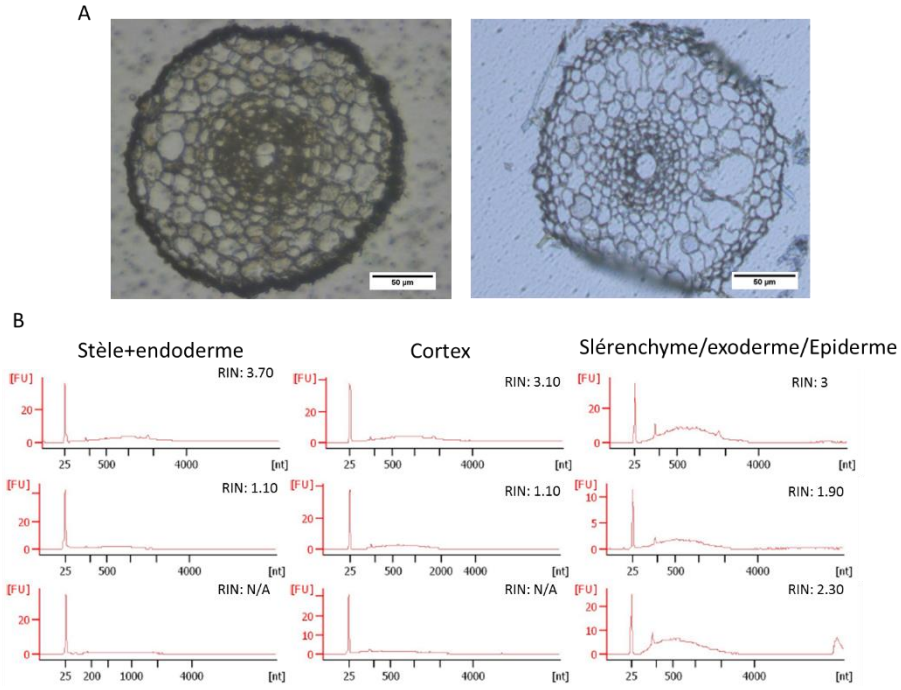


Figure 24: Microdissection laser des trois groupes de tissus dans les zones de propagation de la mort cellulaire des cellules du cortex.

A. Section type de la zone de développement « initiation de formation des aérenchymes (flèche rouge) découpée par laser.

B. Analyse de l'échantillon après extraction d'ARN sur les microdisséquas de compartiment cellulaire au stade d'initiation de formation des aérenchymes.

	Cylindre central pg/µl (RIN)	Cortex pg/µl (RIN)	Tissus externes pg/µl (RIN)
Phase de différenciation cellulaire	146 (6,9)	1277 (7,5)	2753 (7,4)
	303 (6,3)	780 (6,1)	2673 (7,6)
	335 (7,6)	816 (6,2)	1513 (7,2)
	196 (7,3)	412 (6,3)	2001 (6,8)
Phase d'élongation cellulaire	1123 (7,3)	984 (7)	4400 (7,3)
	675 (7,6)	1415 (7,5)	2754 (7,5)
	1245 (8,3)	2678 (6,3)	4209 (6,2)
	1353 (6,8)	1739 (5,5)	4186 (6,8)
Phase mature du cortex	156 (7,6)	768 (7,4)	1989 (7,7)
	183 (7,8)	737 (7,9)	1336 (8,1)
	647 (6,4)	682 (6,8)	2656 (8,3)
	1057 (6,9)	1480 (6,8)	1539 (6,7)
Phase de propagation de la mort cellulaire	0	0	0
	0	0	0
	0	0	0

Figure 25: Concentration et RIN (RNA integrity number) des échantillons d'ARN récoltés en fonction du compartiment cellulaire et du stade de développement des cellules corticales

Les concentrations sont en pg/µl, le RIN est indiqué entre parenthèses. Le RIN moyen est de 7, avec un minimum à 5,5 et un maximum à 8,3. Des traces d'ARN ont pu être observées dans les échantillons pour la phase de propagation de la mort cellulaire, mais trop peu et/ou trop dégradé. Environ 8 µl d'échantillons sont disponibles après analyse pour construire les banques cDNA. Les quatre répétitions biologiques ont été analysées.

5. Discussion : un protocole qui permet de construire des banques RNAseq à partir de très petites quantité d'ARN pour le riz : limites

5.1 La microdissection de sections racinaires au niveau de la phase de propagation de la mort cellulaire n'a pas permis d'extraire d'ARNs

Nous avons utilisé la technologie LM pour microdisséquer trois groupes de tissus racinaires dans la zone développementale qui correspond à l'initiation de la propagation de la mort cellulaire dans le cortex (**Figure 24**) correspondant à une étape importante de la formation des aérenchymes matures chez le riz (Kawai *et al.*, 1998, Yamauchi *et al.*, 2013, Joshi et Kumar 2012, Takahashi *et al.*, 2014). Dans les compartiments stèle et cortex, des ARN dégradés sont visibles. Dans les compartiments sclérenchyme/exoderme/épiderme, nous observons des ARN dégradés plus concentrés avec un RIN autour de 3. La quantité d'ARN extraite dans cette zone est faible et les ARNs sont tous fortement dégradés. Pour le cortex, cela pourrait s'expliquer assez simplement : la formation d'aérenchyme conduit à la mort cellulaire programmée et à la dégradation des ARNs. Mais pour les autres tissus, nous n'avons pas d'hypothèse biologique simple pour expliquer la mauvaise qualité et la faible quantité des ARN extraits. Il reste toutefois possible que les conditions de fixation et d'inclusion soient inadaptés pour les tissus les plus différenciés et qu'il soit nécessaire dans le futur de les ajuster en augmentant le temps de fixation et/ou le temps d'inclusion en paraffine. Le protocole que nous avons mis au point est simple, rapide et permettra facilement d'ajuster différents paramètres et tester leur impact sur la qualité et la quantité des ARN extraits. Nous n'avons pu conduire cette série d'optimisations par manque de temps, mais elles pourront être reprisent assez facilement à la suite de ce travail.

Cette optimisation pourrait permettre d'effectuer des prélèvements dans les tissus entourant le cortex plus haut dans la racine, au stade où les aérenchymes sont complétements matures. En effet, en parallèle à la formation des aérenchymes dans le cortex en condition de submersion, les tissus externes, plus particulièrement l'exoderme, forment une barrière au ROL *via* une subérisation de leurs parois cellulaires (Shiono *et al.*, 2014, Shiono *et al.*, 2010, Kulichikhin *et al.*, 2014). De plus, une couronne d'une ou deux couche de cellules corticales subsistent au stade ou les aérenchymes matures.

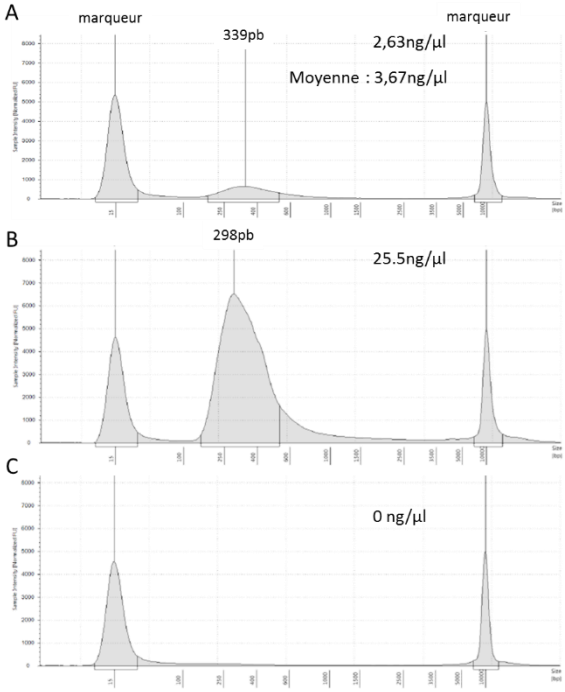


Figure 26: Profils des banques d'ADNc tissu et phase de développement spécifique
A. Profil type d'une banque d'ADNc obtenu avec le kit Ovation® SoLo RNA-Seq Library Preparation de Nugen. La taille des fragments cDNA présentent une répartition normale autour de 300pb, ce qui est la répartition et la taille optimale recommandée.
B. Profil du contrôle sans appauvrissement des ARNr (Any deplete). La répartition normale autour de 300 pb est respectée, mais la quantité d'ADNc est beaucoup plus importante.
C. Contrôle sans matrice, aucune trace d'ARN n'est détectée.

Phase développementale du cortex	Cylindre central ng/µl	Cortex ng/µl	Tissus externes ng/µl
Phase de différenciation cellulaire	0,357	2,23	2,42
	2,63	3,65	3,87
	1,45	8,99	2,45
	1,97	7,95	2,43
Phase d'élongation cellulaire	0,187	6,23	3,5
	0,672	3,75	1,98
	2,37	2,36	5,34
	5,85	5,17	7,04
Phase mature du cortex	6,35	6,05	1,89
	5,54	6,23	1,69
	5,07	3,75	1,54
	2,89	6,23	0,149

Figure 27: Concentration des banques d'ADNc en ng/µl
La concentration des échantillons a été estimée au bioanalyzer avec le kit Bioanalyser RNA Pico 6000. La concentration est exprimée ici en ng/µl.

Extraire l'ARN de ces tissus et réaliser des analyses RNAseq nous permettrait potentiellement d'obtenir des informations sur la raison de la non propagation de la mort cellulaire à ces tissus corticaux. Des résultats récents ont en effet suggéré qu'il existait un différentiel d'expression de gènes impliqués dans la captation des espèces réactives de l'oxygène (ROS) entre les cellules corticales qui entraient dans un processus de mort cellulaire et les cellules des tissus adjacents qui restaient vivantes (Rajhi *et al.*, 2011, Yamauchi *et al.*, 2017(b)). Vérifier si ce différentiel d'expression existe également entre cellules corticales médianes formant des aérenchymes et cellules corticales survivantes avant la phase « aérenchymes matures » serait donc informatif. Ce serait d'autant plus intéressant que des différences morphologiques et de composés pariétaux ont déjà été mis en évidence par notre équipe de recherche entre ces deux types cellulaires corticaux chez le riz (Henry *et al.*, 2016). Par ailleurs, les études de Rajhi *et al.*, 2011 et Yamauchi *et al.*, 2017(b) soulignent également l'efficacité des approches LM pour l'investigation sur les acteurs moléculaires impliqués dans la formation des aérenchymes corticaux. Il restera aussi à optimiser le protocole du LM, en collaboration avec Leica dans l'objectif de microdisséquer plus finement les couches de cortex, par exemple en tentant de séparer couches externes de cortex (qui ne subissent pas la PCD) des couches médianes (qui initient la PCD) puisque, probablement en raison des limites techniques de l'appareil, nous avons échoué à extraire de l'ARN d'une seule couche cellulaire.

5.2 LM tissus/stades de développement spécifiques et qualité des échantillonnages

Les études utilisant la technologie LM pour microdisséquer les tissus racinaires des céréales publiées jusqu'à présent ne fournissent peu voire aucun détail sur les stades de développement du cortex ou sont prélevés les tissus (Takahashi *et al.*, 2010, Takehisa *et al.*, 2012, Rajhi *et al.*, 2011, Shiono *et al.*, 2014, Yamauchi *et al.*, 2017(b)). Le protocole utilisé est celui développé par Takahashi *et al.*, en 2010 et celui-ci nécessite un échantillonnage d'un très grand nombre de sections pour chaque expérimentation de LM, pouvant aller jusqu'à plus d'une centaine de sections pour des tissus comme la stèle, ce qui rend très laborieux l'analyse manuelle de chaque section et du stade prélevé avant microdissection.

Phase développementale du cortex	Cylindre central nM	Cortex nM	Tissus externes nM
Phase de différenciation cellulaire	7.2	8.0	16.6
	5.9	10.7	23.1
	4.3	14.3	20.5
	7.8	20.2	6.8
Phase d'élongation cellulaire	1.9	17.1	10.5
	3.5	8	9.3
	11.9	9.9	11.5
	13.1	7.7	8.9
Phase mature du cortex	17.4	16.1	8.5
	12.7	14.1	16.6
	16.5	13.3	22.7
	12.9	17.4	0.4

Figure 28: Molarité des banque d'ADNc
La molarité des échantillons a été estimé avec un test Takara Library Quantification Kit. La molarité est exprimée ici en nanomoles (nM).

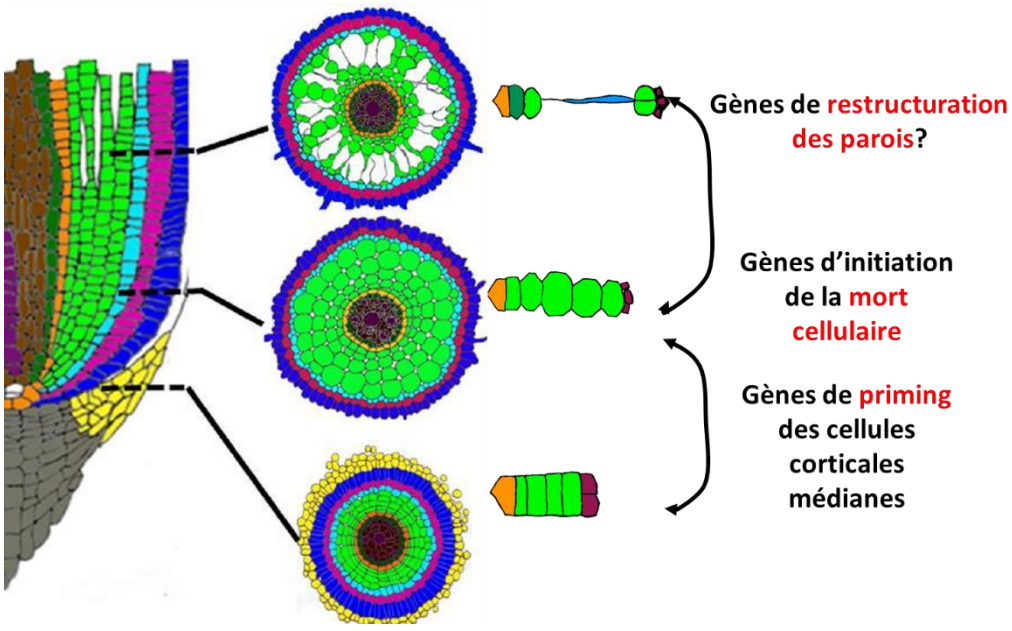


Figure 29: Représentation schématique de nos hypothèses concernant les gènes exprimés lors des transitions entres les phases de développements des aérenchymes décrites par Kawai et al., 1998)

Dans notre étude, les différents stades de développements prélevés sont espacés sur une distance totale de l'ordre de quelques centaines de micromètres le long d'une même racine séminale (données non montrées). Grâce au protocole que nous avons développé, nous n'avons eu besoin de ne microdisséquer qu'une dizaine à quinzaine d'échantillons par stade de développement. La qualité des échantillons, avec un RIN moyen de 7, nous permet d'espérer une analyse transcriptomique optimale (Shroeder *et al.*, 2006, Endrullat *et al.*, 2016, Kashofer *et al.*, 2013, Derveaux *et al.*, 2010).

Si l'on considère la formation des aérenchymes comme un processus développemental tissu spécifique, il est possible de faire l'hypothèse de la présence de gènes potentiellement impliqués dans le priming (amorçage) des cellules corticales pour initier la mort cellulaire dans ce tissu juste avant pendant la phase de maturation du cortex (**Figure 29**). Cette hypothèse est appuyée par les deux études transcriptomiques tissus spécifiques publiées par Rajhi *et al.*, en 2011 et Yamauchi *et al.*, en 2017, respectivement. En effet, les auteurs émettent l'hypothèse de l'implication des MT dans la limitation de la mort cellulaire au cellules corticales. Viendrait ensuite l'expression de gènes impliqués dans les processus ayant attrait à la mort cellulaire, puis potentiellement des gènes de restructuration des parois qui permettent d'obtenir la structure finale des aérenchymes matures chez le riz.

L'utilisation du protocole décrit dans la partie 2 nous a permis d'isoler des quantités d'ARN analysables par ddRT-PCR (cf partie 2) et compatibles avec la préparation de banques RNAseq, et ce, à partir d'une dizaine à vingtaine de sections racinaires microdisséqués. Le protocole de microdissection couplé à la ddRT-PCR et au kit Nugen développé durant cette thèse forment donc un protocole complet qui permettra de mener des analyses transcriptomiques (ddRT-qPCR, cf chapitre 2, et séquençage du transcriptome) à partir des picogrammes d'ARN chez le riz et pourra être repris par d'autres équipes dans le futur.

5.3 Le kit kit Nugen Ovation® SoLo RNA-Seq Library permet de partir de quantités plus faibles d'ARN et permet une meilleure conservation de la stœchiométrie entre les transcrits

Une source majeure de biais est la nécessité de l'étape d'amplification des banques d'ADNc avant envoi au séquençage lorsque l'on travaille avec de petites quantité d'ARN en entrée, la plupart des séquenceurs ayant une capacité de détection limitée (Hrdlickova *et al.*, 2016, Dijk *et al.*, 2014). En effet, l'amplification des transcripts par PCR classique va changer la stœchiométrie entre les transcrits, et sur-représenter les plus abondants, ce qui va conduire à sous-estimer les transcrits les plus faiblement exprimés (Coenen *et al.*, 2015, Puskàs *et al.*, 2002). Par ailleurs, les transcripts les plus abondants sont les ARNr qui seront eux aussi sur-représentés. Ce qui conduit à la nécessité de séquencer avec une plus grande profondeur les banques RNAseq et à une évaluation biaisée du niveau d'expression relatives des transcrits en raison des biais d'amplification introduits. Ce biais dégrade la qualité des informations nécessaires à une inférence correcte d'un réseau de gènes (Fu *et al.*, 2014). Le kit Nugen Ovation® SoLo RNA-Seq Library Preparation propose deux technologies pour pallier à ce biais.

La première technologie repose sur l'indexation moléculaire des transcrits (Fu *et al.*, 2014). Lors de l'étape de ligation des adaptateurs, avant amplification, un UMIs (Unique Molecular identifier) qui est une séquence de 8pb aléatoire, est incorporée aux ADNc. Cette UMI est donc unique et spécifique à chaque transcrit, ce qui permet de distinguer les doublons issus de l'amplification par PCR (Casbon *et al.*, 2011, Shiroguchi *et al.*, 2012). Cette technologie est particulièrement utile sur de très petites quantité d'ARN pour lesquels un grand nombre de cycles d'amplifications sont nécessaires pour limiter les biais d'amplification.

La technologie SPIA (Single Primer Isothermal Amplification) permet une amplification linéaire des ADNc. Cette technologie parue en 2005 (Kurn *et al.*, 2005) utilise un primer unique chimérique pour amplifier les cDNA. Elle a été testée sur de nombreux organismes et l'amplification montre une bonne linéarité de l'amplification entre des transcrits ayant 6 ordres de magnitude de différence (Kurn *et al.*, 2005). L'utilisation de cette technologie sur de l'ARN de rat suffit à stabiliser la stœchiométrie de 93% des gènes exprimés (Faherty *et al.*, 2015).

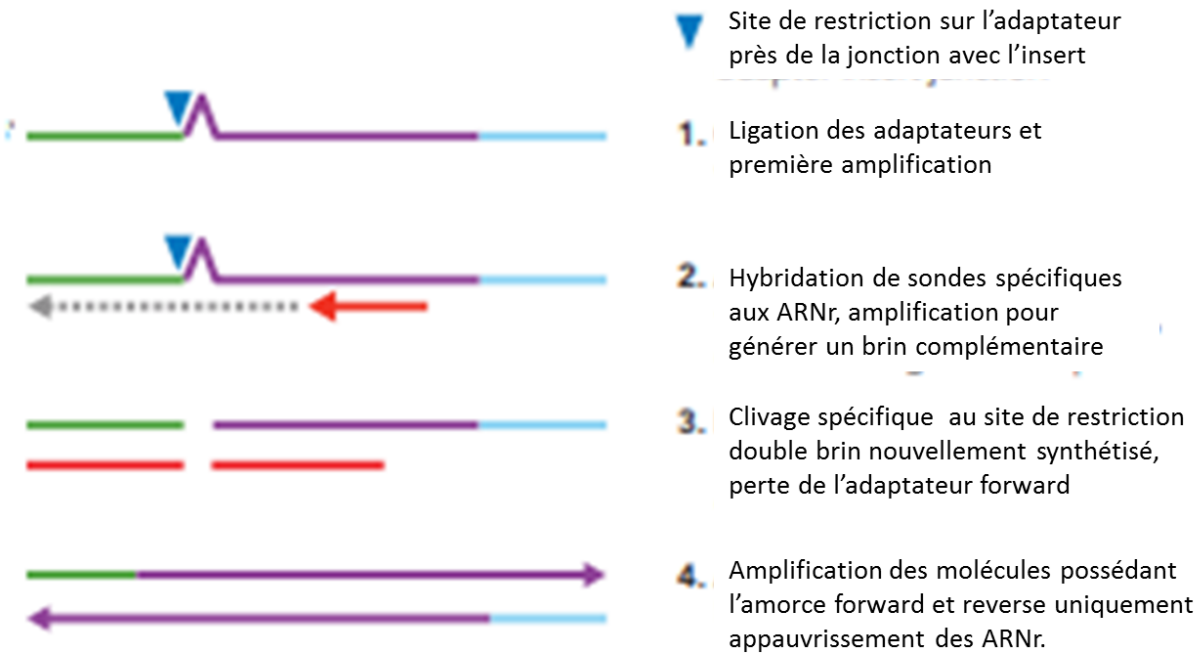


Figure 30: Description de la technologie Any-deplete du kit Ovation® SoLo RNA-Seq Library Preparation de Nugen

5.4 La déplétion des ARNr permet d'optimiser le nombre de lectures lors du séquençage

Chez les eucaryotes, les ARNr représentent habituellement 90% du pool total d'ARN (Conesa *et al.*, 2016, O'Neil *et al.*, 2013, Mills *et al.*, 2013). Ces transcrits sont donc surreprésentés dans les données de séquençages, ce qui réduit grandement le ratio nombre de lectures/profondeur de séquençage alors même que pour des études d'expression, ils n'ont que peu d'intérêt. Classiquement, les ARNm polyA sont extraits, afin d'éliminer les ARNr, pour ensuite être rétrotranscripts afin d'éliminer la majorité des ARNr. Malheureusement, cette technique nécessite une quantité significative d'ARN pour être utilisable et élimine les transcrits qui n'ont pas de queue polyA. Enfin, une sélection sur la taille est réalisée en général, ce qui élimine encore un certain nombre d'ARN de la procédure. La technologie de déplétion des ARNr proposée par Nugen (Any-deplete, anciennement InDA-C, **Figure 30**) est la seule technologie présente sur le marché qui permet un traitement adapté d'échantillons d'ARN de l'ordre du picogramme. La technologie Any-deplete utilise des amorces NSR (Not So Random) spécifiques qui s'hybrident aux séquences nucléiques spécifiques des ARNr. Une étape supplémentaire d'amplification ajoute un site de clivage au niveau de l'adaptateur en reverse, et une étape de digestion va éliminer la partie 3' du cDNA ce qui prévient son amplification lors des étapes d'amplification suivantes. De plus, contrairement aux autres kits proposant cette technologie, la phase d'appauvrissement se fait sur les cDNA après une première phase d'amplification linéaire, ce qui permet de travailler avec des quantités d'ARN initiales extrêmement faibles, notamment après LM comme montré par Farris *et al.*, en 2017. Nos premiers résultats semblent indiquer que la customisation du kit pour le riz a été efficace. On constate en effet une diminution de la quantité d'ARN très importante dans nos échantillons de banque d'ADNc final comparativement au contrôle sans l'étape de déplétion. La banque non appauvrie en ARNr sera également envoyée au séquençage et comparée à la banque déplétée afin de s'assurer que la majeure partie des ARNr ont été éliminés et seulement les ARNr.

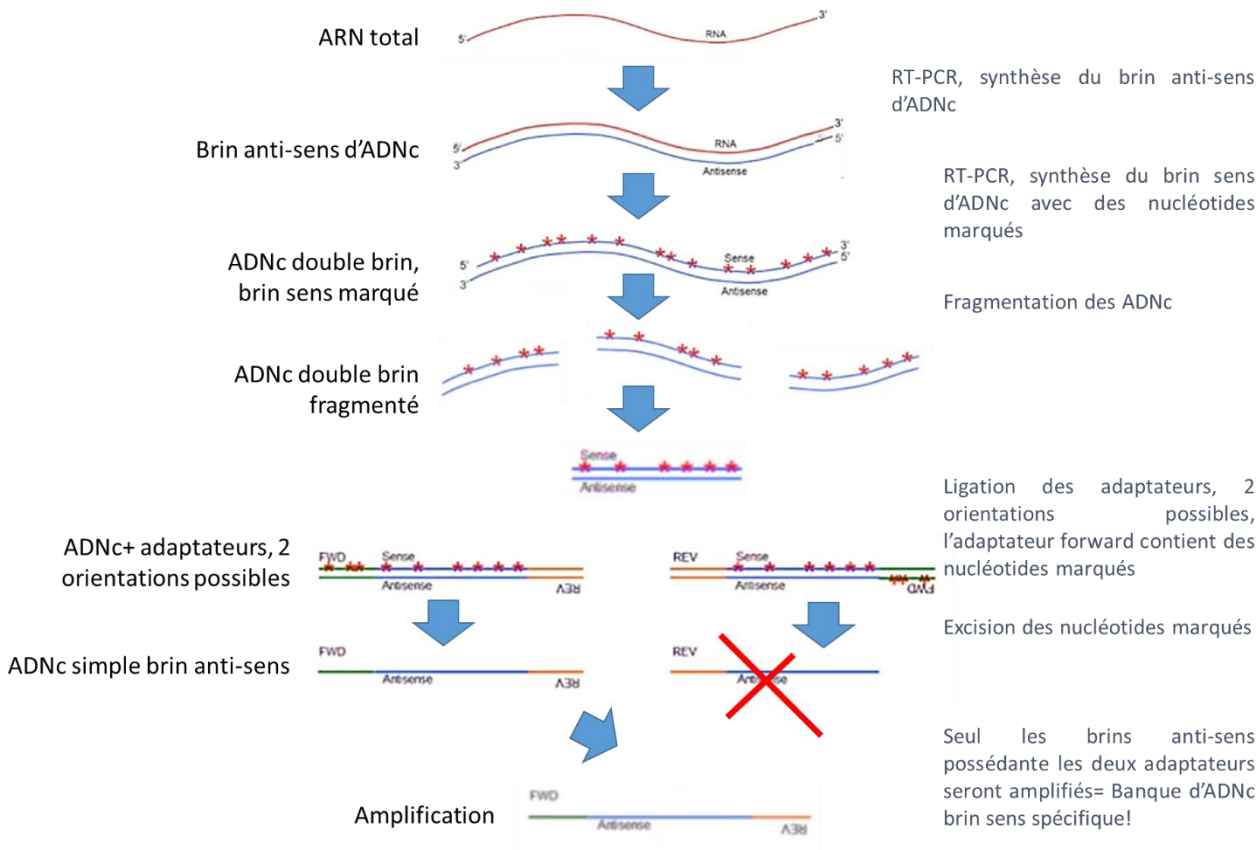


Figure 31: Représentation schématique de la technologie de sélection brin transcrit spécifique proposé par Nugen dans le kit Nugen Ovation® SoLo RNA-Seq Library Preparation
https://www.youtube.com/watch?v=2Ph7V_lTFjg

5.5 La production d'une librairie brin sens spécifique permettra potentiellement la détection de nouvelles espèces d'ARN

Les anciens standards lors de la préparation des banques d'ADNc sont basés sur la ligation d'adaptateurs similaires aux deux extrémités d'un brin d'ADNc issus d'une transcription inverse (RT) avant amplification. Bien que simple et adaptée à de petites quantités d'ARN, cette méthode ne donne pas d'informations quant à la spécificité du brin exprimé, ce qui complique l'identification de transcrits anti-sens et la découverte de transcripts alternatifs, et donc de nouvelles espèces d'ARN dont la transcription n'avait pas encore été reportée (Hrdlickova *et al.*, 2016, Levin *et al.*, 2010). Le kit Nugen Ovation® SoLo RNA-Seq Library Preparation permet une incorporation de nucléotides modifiés durant la ligation du primer forward et la première amplification (**Figure 31**). Ainsi, le brin marqué pourra être dégradé grâce à une enzyme qui clive spécifiquement ces acides nucléiques modifiés, ce qui permet d'obtenir des libraires sens spécifiques analysables.

Le kit permet théoriquement de partir de 10pg d'ARN (Dueck *et al.*, 2016, Adiconis *et al.*, 2013, Shanker *et al.*, 2015, Masters *et al.*, 2018, sur échantillons LM : Faris *et al.*, 2017). Il a été comparé à deux autres kits en terme de performance, le kit Illumina® TruSeq Stranded Total RNA, et le kit Takara® SMARTer universal ultra low input RNA. Ces deux autres kits nécessitent entre 100ng et 10ng d'ARN en entrée, respectivement, pour pouvoir effectuer une étape de déplétion des ARNr. De telles exigences ne sont pas compatibles avec les quantités obtenues pour la construction de nos banques RNAseq.

6. Conclusion

Le protocole LM développé dans la partie 2 a donc permis la collecte d'ARNs de bonne qualité (RIN moyen à 7) provenant de différents groupes de tissus, et ce pour des stades de développement très précis de différenciation du cortex. Notre but était d'obtenir un transcriptome par RNA-seq tissu et stade de développement spécifique. Il nous a tout d'abord fallu adapter chez le riz, le seul kit disponible sur le marché qui permettent un traitement optimal (Hrdlickova *et al.*, 2016, Dijk *et al.*, 2014) de quantités aussi faibles d'ARN pour obtenir les banques d'ADNc les moins biaisées possibles. Ces technologies sont discutées ici et les résultats préliminaires (Any-deplete, molarité et

répartition normale des fragments autour de 300pb) sont très encourageants. Le séquençage est en cours et nous donnera de plus amples précisions sur l'efficacité du kit.

Après séquençage des banques d'ADNc, notre collaboration avec l'équipe I3Bionet nous permettra d'obtenir des réseaux de gènes reflétant la dynamique de l'expression au cours du développement des aérenchymes, et ce, dans le cortex, mais aussi dans les tissus adjacents. Nous espérons ainsi mettre en évidence des régulateurs moléculaires clés de ce processus biologique. Ce protocole, comprenant LM, construction de banques RNAseq à partir de microquantités d'ARN, validation de contamination inter-tissus par RT-ddPCR, pourra en principe être appliqué à un grand nombre d'espèces pour de nombreuses questions biologiques, notamment autour de la biologie du développement.

Chapitre 4

Première version d'un futur article: *CRISPR based synthetics transcriptions factors to decipher gene network in rice*

Chapitre 4, *Première version* d'un futur article: CRISPR based synthetics transcriptions factors to decipher gene network in rice

1. Etapes *Design* et *Build* du cycle DBTL, utilisation de facteurs de transcriptions synthétiques basés sur la technologie CRISPR/Cas en cellules isolées

Les chapitres 2 et 3 de ce manuscrit de thèse sont axées sur l'obtention de jeux de données transcriptomiques qui nous serviront à émettre des hypothèses quant aux régulateurs moléculaires impliqués dans la formation des aérenchymes (étape Learn du cycle DBTL). **Le chapitre 4 sera axé sur le développement d'outils moléculaires et d'un système biologique qui permettront d'éprouver ces hypothèses expérimentalement (étape Design-Build).** La stratégie DBTL repose en grande partie sur la biologie synthétique pour accélérer le cycle. L'année 2012 a vu l'apparition d'une nouvelle technologie révolutionnant le domaine de la biologie synthétique et de l'ingénierie du génome, la technologie CRISPR/Cas (Jinek *et al*., 2012, Jakočiūnas *et al.,* 2017, Kumlehn *et al.,* 2018, Puchta *et al.,* 2017). Brièvement, cette technologie permet de cibler un locus génomique spécifique et d'y recruter des domaines enzymatiques ayant différentes activités, comme une activité de régulation transcriptionnelle. Ainsi, une des applications possible est la construction de facteurs de transcriptions synthétiques (TFs). Dans ce futur article, nous développerons trois TFs qui seront testés pour effectuer une preuve de concept : l'utilisation de facteurs de transcriptions synthétiques CRISPR/Cas pour manipuler l'activité transcriptomique d'un gène rapporteur en protoplastes de riz.

2. CRISPR based synthetics transcriptions factors to decipher gene network in rice

ABSTRACT

Background: While there are currently efficient methods for rebuilding GRN and an ever-increasing amount of accessible expression data, one of the bottleneck remains the fast validation of predicted regulatory relationships. A way to quickly validate predicted regulatory relationships is to test them by modulating the expression of the regulator(s) and to check if the targets are enabled**.** The third generation of CRISPR/CAS9 synthetic transcription factor were developed to allow the simultaneous activation and repression of several genes and is currently the most appropriate tool for the validation of gene networks.

Results: We developed the first steps to perform gene expression modulation in rice using the scRNA technology using a prototyping system in protoplasts. A fluorescent reporter gene and FACS (fluorescent activated cell sorting) were combined to compare the efficiency of first and third generation of constructs of CRISPR/CAS9 synthetic transcription factors. scRNA approach can be used to overexpress or downregulate expression of a reporter gene. Moreover, scRNA strategy is more efficient for repression than direct fusion of a dCAS9 with a 3xSRDX repressor domain.

Conclusion: Our prototyping system combining synthetic promoter teto1x/teto7x for activation and modified p35 driving VENUS expression can be used to evaluate efficiency of constructs to activate or repress gene expression in the future.

Introduction

Biological systems are being further investigated at an unprecedented scale thanks to the large amount of genomic information now available. New methodologies are thus needed for efficient data mining of massive transcriptomics experiments (Elati *et al.*, 2007, Nicolle *et al.*, 2015, Chebil *et al.*, 2014). Gene Regulatory Network (GRN) reconstruction consists of building a network where the nodes represent genes and the edges biological relationships such as protein interactions and/or regulatory links between regulators and their effectors. Most GRN inference approaches consider expression correlation between transcripts to identify possible regulatory links between genes using model assumptions (Hecker *et al.*, 2009, Banf and Rhee, 2017). Edges represent then predicted direct as well as indirect regulation between genes. At this stage, these links represent only correlations that, if they guide biological experiments later on, need to be validated to move from correlation to regulation status. Some GRN reconstruction methods use machine learning algorithms that will gradually improve the reconstructed network by using new data, including validation of regulatory networks, protein-protein interaction data or expression data (Nicolle *et al.*, 2015, Trébulle *et al.*,2017). While there are currently efficient methods for rebuilding GRN and an ever-increasing amount of accessible expression data, one of the bottleneck remains the validation of predicted regulatory relationships. (Konermann *et al.*, 2015, Didovyk *et al.*, 2016, Najm *et al.*, 2018, Boettcher *et al.*, 2018). **The most direct method is then to modify the expression of one or more regulators and determine if the expression of their predicted regulators is modified in return.**

A technology, called Transient Assay Reporting Genome-wide Effects of Transcription factors or TARGET, was developed by Bargmann et al in 2013. This technology consists of merging a transcription factor into a GR domain of the rat steroid receptor, under the control of a strong constitutive promoter, and transformation of the construct into protoplasts. After adding dexamethasone, the chimeric protein is translocated in the nucleus where it will activate its targets. TARGET has made it possible not only to validate predicted targets, but also to discover new targets of candidates transcription factors (Medici *et al.*, 2015, Para *et al.*, 2014). The addition of cycloheximide makes it possible, by blocking translation, to differentiate between primary targets,

(direct regulation), and secondary targets, (indirect regulation). However, there are many limitations to TARGET. Merging into a GR domain can change the function of the transcription factor. Ideally, the functionality of the construction should be checked by complementing a mutant loss of function, a verification very rarely done in practice (Levesque *et al.*, 2006). Then, this system can only be used for FT, not for other regulators such as kinases, peptides or even other classes of regulators. Finally, in practice, this system only allows to test the effect of overactivation of an FT, not its downregulation.

The discovery of TALE, proteins derived from Xantomonas oryzae, a rice pathogen bacterium, has led to the construction of the first artificial transcription factors that target the promoter of an FT or any candidate gene. In 2010, Morbitzer *et al.* showed that activation of a selected genome locus can be monitored by transcription activator like effector (TALE)-type transcription factors. The part of the TALE protein that specifically targets a specific genomic region requires the repetition assembly of 16 amino acids, one repetition being specific to a nucleotide. To target a single region of 20 nucleotides, it is therefore necessary to build a domain composed of 320 amino acids and like the TALE dimer operation, it is necessary to build two domains of 320 amino acids. Soon after the demonstration of TALE activity to control gene expression in plants, come **the first demonstration that a new technology called CRISPR-Cas9, can be used to modify any specific sequence in eukaryotes cells, paving the way for the CAS9 revolution (Jinek *et al.*, 2012).**

CRISRP/CAS9 is now a very popular technology which has spread very rapidly to become the tool of choice for all species where functional analysis is possible (Xu *et al.*, 2019). This is a two components system, an endonuclease Cas (CRISPR Associated protein), complex with an RNA, called single guide RNA (sgRNA). The sgRNA can be programmed to target any locus using twenty base pairs complementary to its target. The technology is inexpensive and accessible to any molecular biology laboratory. Firstly, this technology was used to inactivate a specific locus by introducing frameshift mutations (Jinek *et al.* 2012). **Quickly, the versatility of the CRISPR/CAS9 system, due to the fact the recognition activity and catalytic activity are provided separately by the guide RNA and CAS9 nuclease, has encouraged the development of a large number of derived technologies by using the targeting function of the system (Pickar-Oliver and Gersbach,**

2019). Comparatively, in the TALEN technology, the protein carries both target recognition and catalytic activity.

The first generation of CRISPR/Cas based synthetic transcription factors was developed by Bickard *et al.*, and Cheng *et al.*, in 2013 using a catalytically 'dead' version of Cas9 (dCas9). This dCas9 can be directly transcriptionally fused to an activation or repressive domain, (CRISPRa and CRISPRi systems, respectively, **Figure 32**). These systems has been tested on a wide range of organisms and biological systems (Piatek *et al.*, 2014; Chavez *et al.*, 2016). However, the reported rather low efficiency of the system, has limited the applications and use of this first generation of constructions (Heckl et Charpentier., 2015, Chavez *et al.*, 2016).

The second generation of CRISPR/Cas based synthetic transcription factor were based on the recruitment of multiple gene expression effectors at the targeted locus (Figure 32). A first experiment to obtain a more efficient gene expression manipulation was successfully done by Gilbert *et al.* and Tanenbaum *et al.* in 2014 with the so called Suntag activation system. This system is based on the fusion of the dCas9 to a tandem array of peptides, called a SunTag array, which recruits many copies of the VP64 activation domain fused to an antibody recognizing the array of peptides. In 2015, Chavez *et al.* too overcome the low efficiency of the first generation of synthetic transcriptional activator by designing the VPR activator system. This system use multiple activators domains (VP64, p65, Rta) fused to the dCas9 to synergistically overexpress the target locus. **However, despite their efficiency, these systems have a major disadvantage: they can only activate or repress expression a set of genes.**

The third generation of CRISPR/CAS9 synthetic transcription factor were developed to overcome this limitation using RNA Aptatmers. RNA aptamers, recognized by small viral proteins, were introduced in the sgRNA (Mali et al., 2013, Zalatan *et al.*, 2015, Lowder LG *et al.*, 2018, Konermann *et al.*, 2015, **Figure 32**). These sgRNA are called scaffold RNA (scRNA). Viral protein are then fused to activator or repressor domains and will be recruited by the scRNA. Elegantly, the scRNA itself encodes both target specificity and regulatory function. Using two differents pairs of aptamer and viral protein, it is now possible to activate and repress expression of two distinct sets of genes simultaneously by designing specific scRNA. Thus, Zalatan *et al.*, in 2015 have harnessed this technology

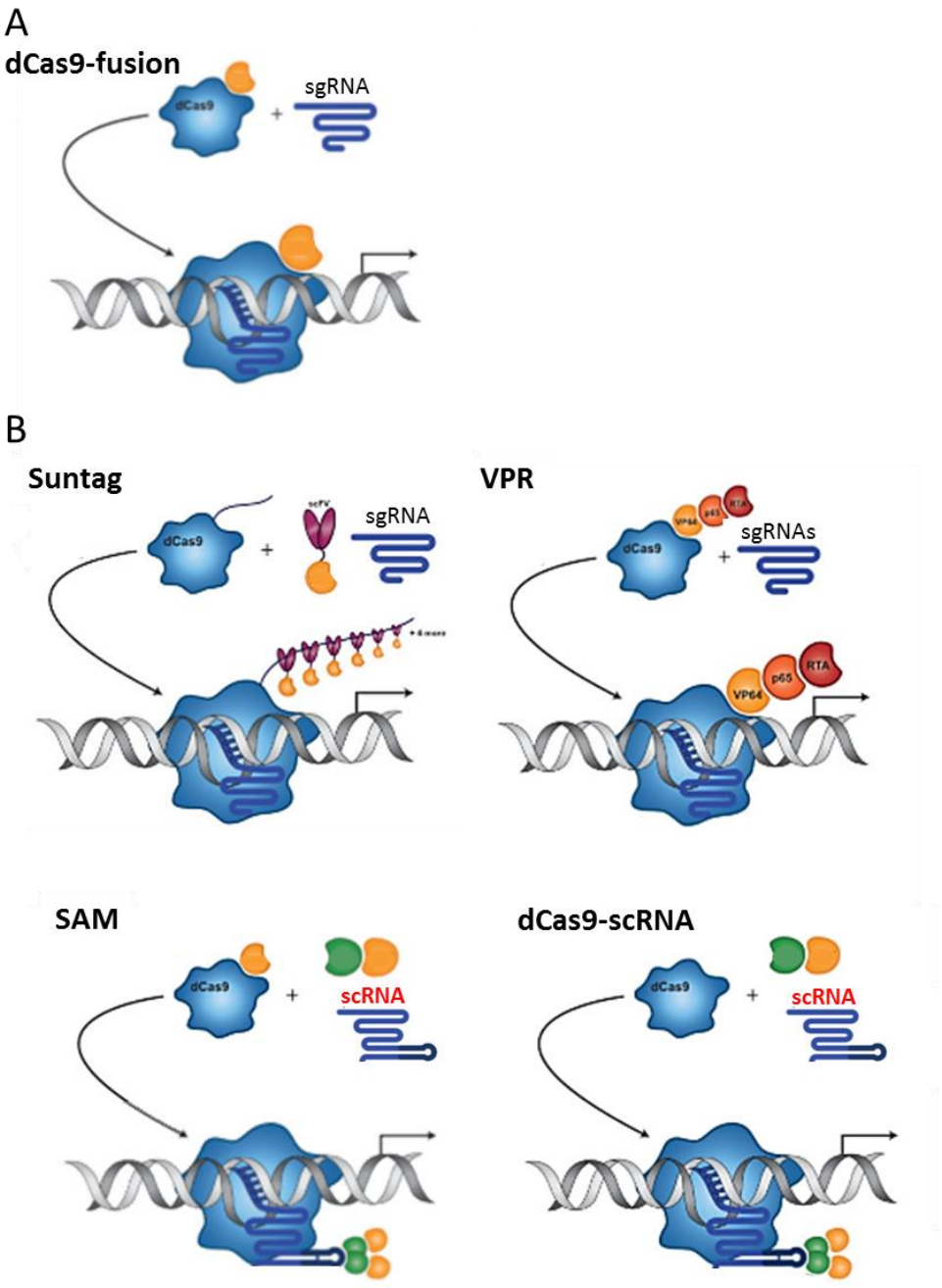


Figure 32: CRISPR synthetic transcription factors, first and second generation

A) First generation of CRISPR/Cas based synthetic transcription factors. dCas9-fusion to one effector trigger gene expression manipulation

B) Second generation of CRISPR/Cas based synthetic transcription factors, recruitment of several effectors at the targeting locus. Suntag uses single chain antibodies to recruit multiple copies of effectors VP64 to the peptide tail. VPR work via three heterogenic effector domains fused to the C-terminus of the dCas9. SAM uses dCas9-VP64, but recruits more activation domains to the gRNA.

C) The third generation of S CRISPR/Cas based synthetic transcription factors. The dCas9-scRNA recruits multiple effectors fuses to a ligands which recognize the sgNA extended with an aptamer sequence.

to flexibly redirect the violacein metabolic pathway in yeast in a tunable manner. **Nowadays, the scRNA strategies are the only simple and generic way to activate and repress simultaneously a set of genes within a cell or an organism.**

We developed the first steps to perform gene expression modulation in rice using the scRNA technology. We designed a prototyping system in protoplasts using fluorescent reporter gene and FACS (fluorescent activated cell sorting) **to compare the efficiency of first and third generation of constructs of CRISPR/CAS9 synthetic transcription factor.** We successfully demonstrated scRNA approach can be used to overexpress or downregulate expression of a reporter gene. We also demonstrated the scRNA strategy is more efficient than direct fusion of a dCAS9 with a repressor domain 3xSRDX, while less efficient for overexpression compare to direct fusion of a dCAS9 with a VP64 activator domain. This work represent a first important step to develop a generic tool to quickly and efficiently manipulate and validate gene networks in rice and beyond in plants.

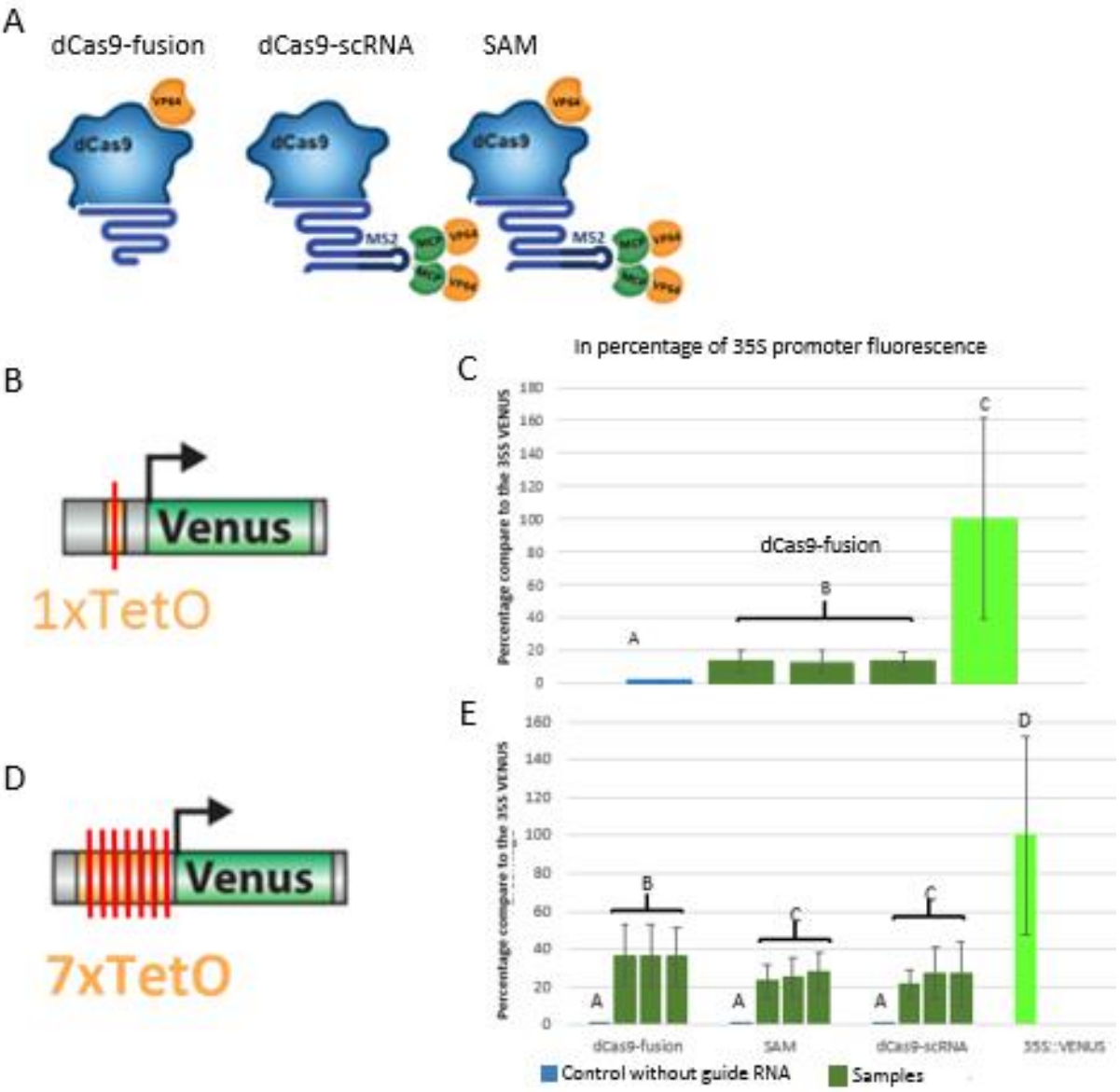


Figure 33: CRISPR-mediated gene expression activation on a teto1x and teto7X VENUS promoter

A. Representation of the CRISPRa two-component dCas9-fusion (dCas9-VP64) and the three-component dCas9-scRNA and SAM strategy. All the constructs were subcloned in individual vectors (see annexes 2,3,4) and co-transformed in protoplasts (5ng of each). B. 1xtetO reporter gene, one CRISPR target is located on the TRE motif of the tetO sequence by adding a 'NGG' PAM sequence (red bar). C. dCas9-fusion mediated VENUS YFP fluorescence in protoplasts population compared to the positive control population transfected with a p35S:VENUS construct. Percentage of the protoplasts population identified as positives. D 7xtetO reporter gene, 7x CRISPR target is located on the 7x TRE motif of the tetO sequence by adding a 'NGG' PAM sequence (red bar). E Comparison of the three CRISPR based transcription factor mediated VENUS YFP fluorescence in protoplasts population compared to the positive control population transfected with a p35S:VENUS construct. A Kruskal-Wallis test was used to calculate the statistical significance of fluorescence difference between protoplasts populations.

Results

Experimental design

Two reporter constructs were first designed to test activation and repression. For the activation experiments, a VENUS YFP fused to a NLS was driven by teto promoter, with two versions, one with a single target, teto1X, a second one with seven repeats of teto1x, the teto7x (**Figure 33**). For the repression experiment, the same VENUS YFP NLS construct is driven by a p35S promoter modified to introduce a 'NGG' PAM (Protospacer Adjacent Motif, Jinek *et al.*, 2012) 120 pb before the TSS (Transcription Start site, **Figure 35, Table 1**).

For activation, a direct dCas9-fusion system with a tetramer of the VP16 domain from the herpes virus, VP64 (4xVP16) was developed (**Figure 32 and Figure 33, Table 1**). This construct correspond to the first generation of CRISPR/Cas based synthetic transcription factors. For the dCas9-scRNA strategy, we added a MS2 (scRNA-MS2) (Lowary and Uhlenbeck, 1987), aptamer to the sgRNA targeting the teto sequence. The MS2 recruit a MCP protein, that we have 3' merged to the VP64 (MCP-VP64) activator domain, (Lowder *et al.*, 2015, Beerli *et al.*, 1998). To evaluate a possible synergy between scaffold and first generation construct, we combineb both through the SAM (Synergistic Activation mediator) strategy that is a hybrid between the dCas9-fusion and the scRNA strategy by co-transformation of both constructs in protoplasts.

For repression, a direct dCas9-fusion system with a trimer of SRDX domain (Lowder *et al.*, 2015) was developed as a control from the first generation of CRISPR/Cas based synthetic transcription factors (**Figure 35**). For the scaffold, we added a com (scRNA-com) aptamer sequence to the sgRNA targeting the p35S promoter 120 pb before the TSS (Hattman, 1999, Zalatan *et al.*, 2015). The com aptamer recruit a COM protein merged to the 3xSRDX (COM-3xSRDX) repressor domain. SAM strategy was also evaluated by co-transformation of both constructs in protoplasts (**Figure 32, Figure 35, Table 1**).

CRISPR RNA scaffold system effectively activate VENUS YFP expression

We first evaluate qualitatively activation of VENUS YFP in protoplasts (**Figure 34**).

Backbones	Constructions
pC5300	pUbi:NLS-**dCas9**:tNos
pC5300	pUbi:**NLS-dCas9-VP64**:tNos
pC5300	pUbi:**NLS-dCas9-3xSRDX**:tNos
pC5300	1xtetO:**NLS-VENUS**:tNos
pC5300	1xtetO:**NLS-VENUS**:tNos pUbi:**NLS-MCP-VP64**:tNos
pC5300	7xtetO:**NLS-VENUS**:tNos
pC5300	7xtetO:**NLS-VENUS**:tNos pUbi:**MCP-VP64**:tNos
pC5300	p35S:**NLS-VENUS**:tNos
pC5300	p35S:**NLS-VENUS**:tNos pUbi:**NLS-COM-3xSRDX**:tNos
pOs-Entry	pOsU3:**scRNA-MS2**:tNos
pOs-Entry	pOsU3:**scRNA-com**:tNos

Table 1: List of the genetic constructs used for protoplast transformation

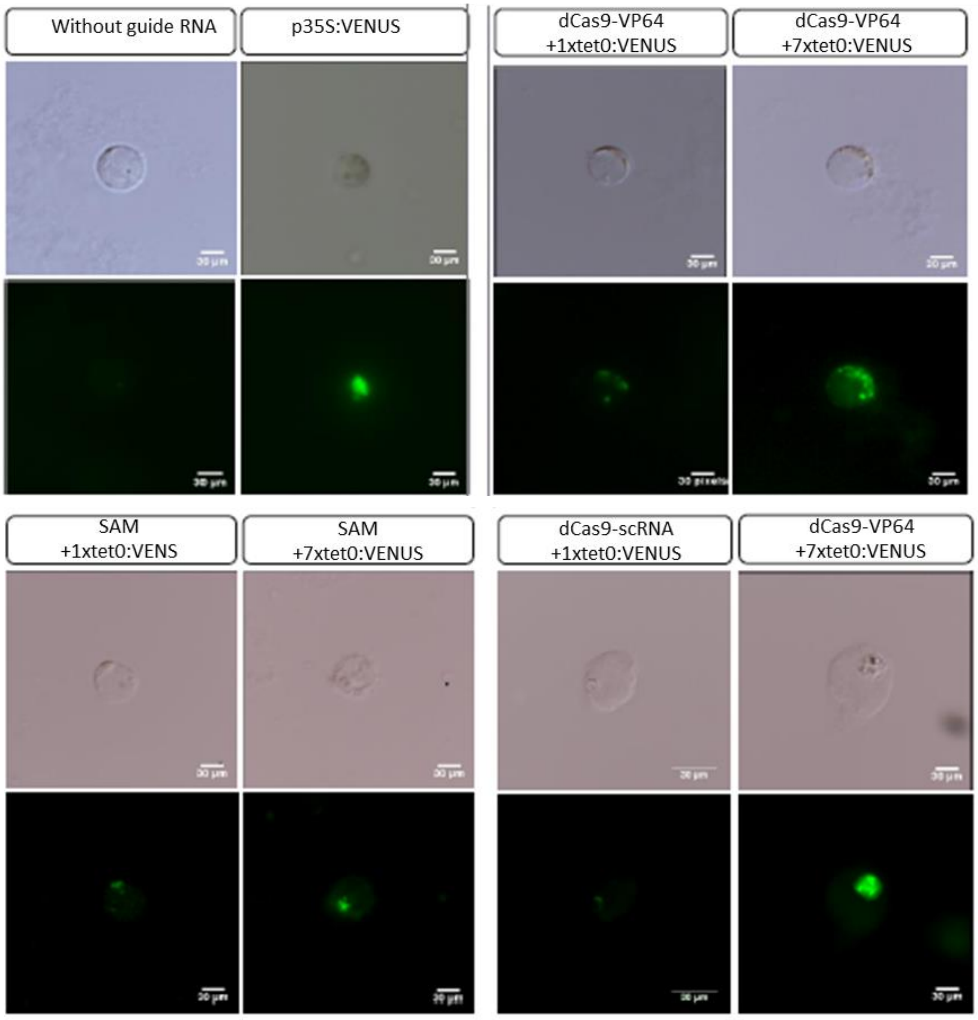


Figure 34: CRISPR-mediated VENUS YFP expression activation, fluorescence microscopy
dCas9-VP64 transfected protoplasts populations without scRNA was used as negative control. P35S driven protoplasts population was used as a positive control (GFP filter, exposition time 300ms, scale bar 20µm)

All systems successfully induce YFP expression and fluorescence in protoplasts. We do not observe any fluorescence for the teto1x and teto7x constructions alone, confirming that these promoters alone do not have cryptic transcriptional activity in rice protoplasts.

FACS was then used to do cell sorting of positive fluorescence protoplasts versus negative and untransformed protoplasts. Mean level of fluorescence of the positive population was then used a proxy to evaluate transcriptional activity of each construct. To calibrate FACS parameters, we used the p35S::VENUS::NLS construct and efficiently separate fluorescent vs non fluorescent protoplasts populations (**Figure 33**, **Annexe 2**). For negative control, we co-transform protoplasts without the RNA scaffold**.** We did no recover positive protoplasts in the negative control population. When we used the FACS to cell sorting teto1X co-transformed protoplasts, we only identified positive protoplasts with the dCas9-fusion system, around 13 percent of the fluorescence intensity observed with a 35S driven VENUS positive control, (**Figure 33**) in a small percentage of the total protoplasts population (9,54%). Compared to qualitive comparison, where we clearly visualized YFP expression for all constructs using teto1X, albeit at different level of fluorescence (**Figure 34**), these results suggest FACS sensibility is too low when the level of activation is moderate. This may be due to lack of sensibility of FACS device or also to the fact we used a YFP with nuclei addressing. Initially, we designed this construct to facilitate difference of “true” fluorescence in protoplasts versus possible autofluorescence as we never observed autofluorescence in nuclei. The protoplast populations used were also heterogenous in term of cell types and cell size, ie coming from stem and leaf, a feature visible during FACS cell sorting (**Annexe 2**). Fluorescence quantification is then somewhat difficult and suggest that using a YFP without any addressing sequence will be more efficient.

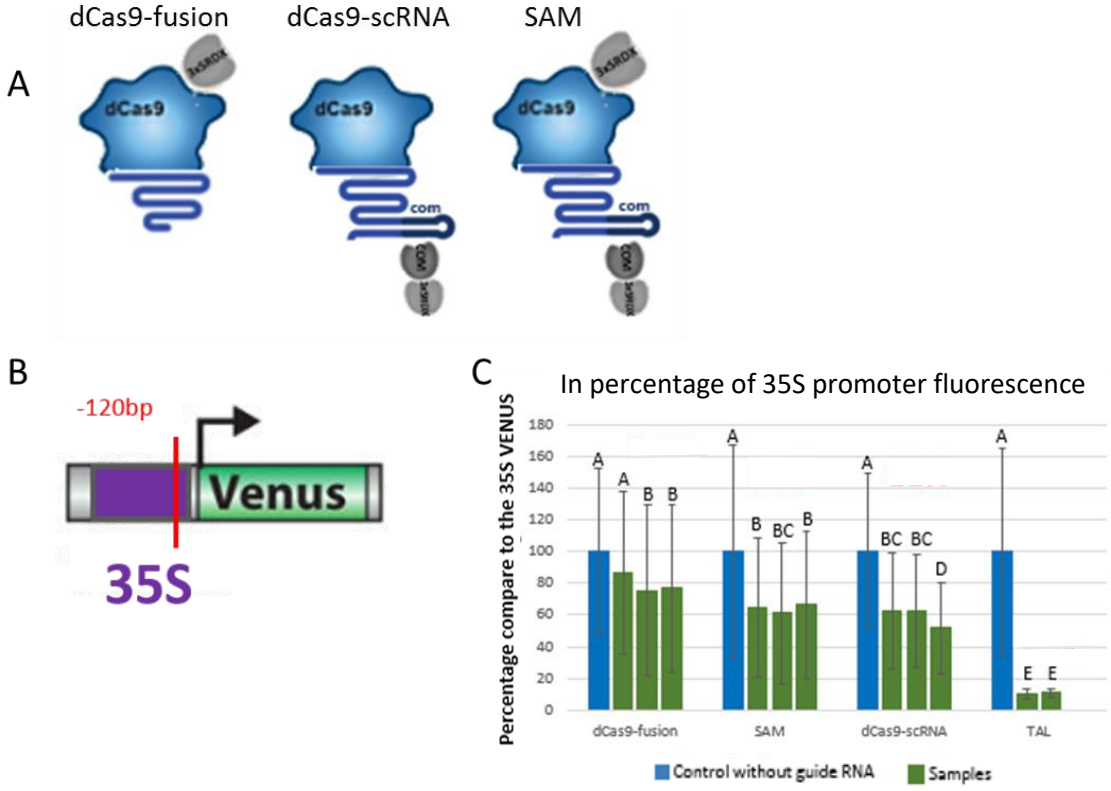


Figure 35: CRISPR-mediated gene expression inhibition on a 35S VENUS promoter
A. Representation of the CRISPRi two-component dCas9-fusion (dCas9-3xSRDX) and the three-component dCas9-scRNA and SAM strategy. All the construct were subcloned in individual vectors (**see annexe 2, 3, 4**) and co-transformed in protoplasts (5ng of each). B. 35S:VENUS reporter gene, one CRISPR target is located 120bp before the TSS in the sens strand by adding a 'NGG' PAM sequence (red bar). C. Three CRISPRi and one TAL systems efficiency to repress p35S driven VENUS fluorescence in transfected protoplasts populations. A Kruskal-Wallis test was used to calculate the statistical significance of fluorescence difference between protoplasts populations.

We shift to the 7xteto VENUS construct to test activation of VENUS YFP for scaffold, direct fusion of VP64 and SAM (Figure 33, Figure 34). In all cases we identified a positive population following FACS cell sorting. dCAS9-VP64 fusion had the highest level of fluorescence, followed by scaffold, corresponding to a 10 and 20x fold induction of YFP fluorescence. Intriguingly, the SAM construction display a fluorescent level similar to dCAS9-VP64 fusion suggesting there is neither synergic nor negative effect of using both system, scaffold and fusion, together. With the dCas9 fusion and the scaffold systems, we achieved a level of fluorescence equivalent to a 36% and 25% of the fluorescence observed the positive control p35S::VENUS::NLS, respectively and around 35% for the SAM.

CRISPR RNA scaffold effectively repress VENUS YFP expression .

We evaluated and compare repression of dCAS9-3XSRDX (fusion), scaffold (COM-3XSRDX) and SAM by targeting the strong 35S promoter driving the expression of the VENUS reporter gene, 120bp in 5' of the transcription start site (TSS, Figure 35). We used a TALEN targeting the coding sequence of the YFP that was successfully used to inactivate GFP expression and take advantage the target is perfectly similar between YFP and GFP as a positive repressive control. After FACS cell sorting, we found a strong reduction of YFP fluorescence with the TALEN (below 10% of a p35S) but also with all three constructs (**Figure 35**). The strongest fluorescence repression observed, beside TALEN, was achieved with the dCas9-scRNA system, with a fluorescence level between 48% and 38% of the p35:VENUS construct. The SAM system repress fluorescence between 34% and 39% and the dCas9 fusion between 14% and 25% of the p35:VENUS positive control fluorescence.

Discussion

Our prototyping system combining synthetic promoter teto1x/teto7x for activation and modified p35 driving VENUS expression can be used to evaluate efficiency of constructs to activate or repress gene expression. First, we easily check qualitatively the functionality of the system using a epifluorescent microscope and quantitatively with a FACS. Our system is similar to the system developed by Zalatan et al 2015 in yeast and human cells albeit using different repressor domains and fluorescence reporter constructs. Nevertheless, our system was developed in rice protoplasts so we

expect to get results more readily useful to develop and applied this technology in plant. For instance, there is no real consensus about which CRISPR synthetic transcription factor is the more efficient across different biological context (Chavez *et al.*, 2016, Mitsunobu *et al.*, 2017). However, the SAM system was reported as one of the most efficient system in several publications and organisms (Chavez *et al.*, 2016, Zhang *et al.*, 2015, Konermann *et al.*, 2015), certainly due to the recruitment of multiple gene expression activator or repressor at the targeted promoter. Our results are in disagreement, suggesting that evaluation in the species where the system will be used is safer to get a better idea of its effectiveness. A limitation of the prototyping system is the low sensitivity of FACS for cell sorting of YFP positive protoplasts. This may be due to the use of a YFP::NLS construct that interfere with FACS detection and also heterogeneity of protoplast population. A simple solution will be to develop an YFP construct without NLS and use of callus instead of stem/leaves protoplasts that are much more homogeneous in size and identity. Nevertheless, due to its simplicity and its modularity, our prototyping system will be used in the future to test other activation/repression domain, scaffold and even other inactivated nuclease like dCpf1 (Tak *et al.*, 2017). Last, we planned to evaluate gene expression of YFP in all protoplast fluorescent populations, using RT-ddPCR. We expect to validate the use of fluorescence level as a proxy of transcription activation or repression. We choice initially the VENUS protein, with an estimated half-life stability of 20 minutes, to ensure the level of fluorescence observed in our experiments was not due to protein accumulation but is direct a readout of YFP transcription level.

Scaffold constructs were used successfully to activate or repress synthetic promoters in rice protoplasts. Using scaffold and MS2-VP64, we achieved a level of expression around 25% of a strong p35S promoter. Moreover, we reduce expression of a strong p35S promoter below 50% of its initial level using com-3XSRDX. These results are a first proof of concept the scaffold strategy can be developed to modulate gene expression in plants. Our results are similar of Zalatan et al 2015, even if we only proved its functionality for synthetic promoters co-transformed with activation/repression constructs. A critical step will be to target rice genes, and demonstrate we can activate and repress their expression first separately, then together. Several other CRISPR based synthetic transcription factors systems were developed recently (For a review Li *et al.*,

2019) and used distinct activation domain. At least, the TV or the VPR domain (Li *et al.*, 2017, Chavez *et al.*, 2015) was recently proved to be superior to VP64. Our prototyping system is a very simple way to test and benchmark novel activation combination. MS2-3XTV could be tested and compared to VP64 in our prototyping system to quantify the activation level gain of this new activator domain. It would be a simple way to improve activation level using scaffold to be compared, again to direct fusion with dCAS9.

Multimerisation of aptamers and optimization of scaffold stability: towards the second generation of scaffolds. A recent publication report the development of an improved scaffold strategy. Konermann *et al.*, 2015 first demonstrated the instability of sgRNA when merging aptamer in the 3' end, but higher stability and expression when aptamers were inserted in the sgRNA tetraloop, resulting in greater efficiency of the system in term of control of gene expression (Chavez *et al.*, 2016). Two years later, Ma *et al.*, 2018 demonstrated the stability of a scaffold made with a 8xMS2 aptamer inserted in tetraloop (called CRISPR-Sirius, **Figure 36**) of the native sgRNA was also compared to a scaffold built with a 12xMS2 aptamer merged in 3' of the sgRNA. In the presence of dCAS9, 8xMS2 aptamer scaffold accumulation in cells increase up 60 fold compared to 3' 12x aptamer scaffold where expression increase only 4 fold compared to a control without dCAS9. These experiments suggest that a critical point is the stability of scaffold. One way to confirm the difference observed between dCAS9-fusion and scaffold in our system is due to scRNA stability, will be to evaluate the level of expression of scaffold and single guide and thus their stability, by digital droplet PCR. We also develop 8x aptamer versions of our scaffolds constructs by inserting the aptamer in the tetraloop, for activation (8xMS2) and repression (8xcom) of YFP. We first evaluate their secondary structure and optimize their sequence to get thermodynamic stable scaffold of second generation (**Figure 36**). We also develop 1x aptamer tetraloop versions of com and MS2 constructs to compare the effect of scaffold stability for activation and repression. All these constructs will be used and compared to identify the best combination for gene activation and repression. Moreover, we expect to get a wider range of level of activation and repression that should be useful in the future if we want to fine tune gene expression modulation.

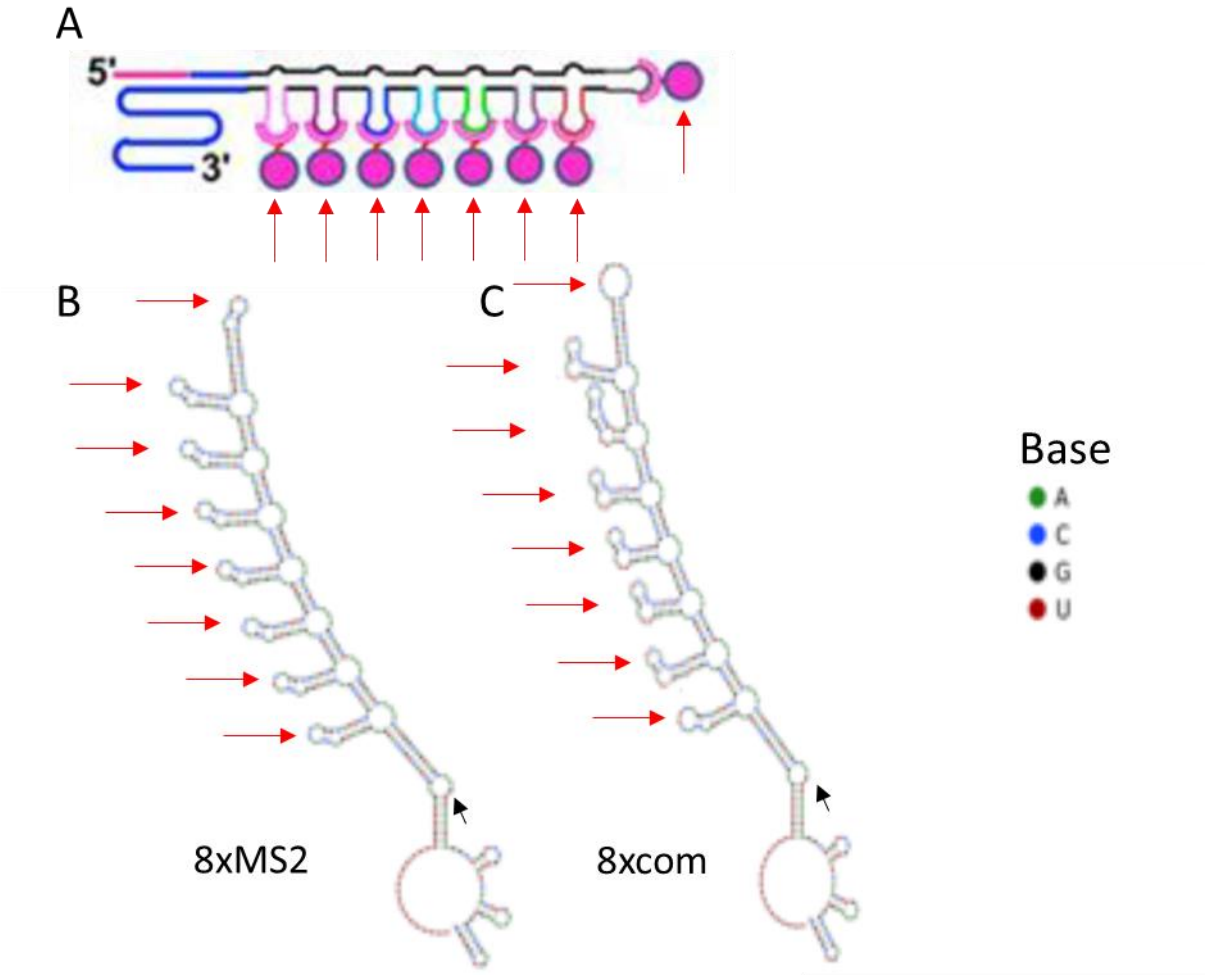


Figure 36: In silico two-dimensional scaffolding modelization of second generation of our scRNA for activation and repression
A. Schematic representation of the SIRIUS scRNA strategy from Ma et al., 2018). B. In silico two-dimensional scaffol representation of our new designed Sirius scRNA with 8xMS2 aptamers sequences C. In silico two-dimensional scaffold representation of our new designed Sirius scRNA with 8xcom aptamers sequences.
Red arrow: Aptamers sequences, Black Arrow: RNA bulge needed for Cas9 recognition of sg and scRNA (for a review, see Nowak et al., 2016).

Scaffold constructs are more efficient for repression than for activation of YFP compared to dCAS9-domain fusions. Our results differ to precedent studies that report a rather lower efficiency of the CRISPR based transcription factor of the first generation (dCas9-fusion) compared to the third generation (SAM, dCas9-scRNA). The differences between the CRIPSRa and CRISPRi assays may be due to independent and not mutually exclusives reasons, including: difference between the targeted promoters, sgRNA efficiency (see below), the use of different domain for repression (KRABS for Zalatan et al 2015, 3xSRDX in our case) for the most probable ones. A first result of Zalatan *et al.*, in 2015 was the demonstration the com aptamer has lower efficiency compared to MS2 aptamer for gene activation thank swapping aptamer experiments. They hypothesized that the difference between the two aptamers was that MS2 recruit MCP, the viral protein recognizing MS2 as a dimer (Chao *et al.*, 2008), bringing two effectors domains at the targeted locus, when com aptamer certainly recruit COM as a monomer (Hattman, 1999). In our experiment, the scRNA strategy was more efficient with the com aptamer, for repression, than the MS2 for activation. This may be also due to the fact we used a different domain of repression. A way to understand differences between our results and Zalatan results will be to do swapping experiments between aptamers in a similar way as described in Zalatan et al 2015.

In conclusion, two modifications seems to lead to an better efficiency of the dCas9-scNA constructs, increasing the stability of the scRNA (Shao *et al.*, 2016, Mu *et al.*, 2019, Konermann *et al.*, 2015) and the insertion of a multiple aptameric sequence into the sgRNA tetraloop in order to recruit multiple effectors at the same locus (Ma *et al.*, 2018, Zalatan *et al.*, 2015, Konnermann *et al.*, 2015).

Conclusion

Our prototyping system can be used to evaluate efficiency of future constructs to activate or repress gene expression. Moreover, we demonstrated scaffold constructs can be used successfully to activate or repress synthetic promoter in rice protoplasts. We also develop news construct to increase efficiency of the dCAS9-scRNA construct by multimerisation of aptamers and optimization of scaffold stability. It remains to prove, using rice genes, the system is also able to modulate rice gene expression. In conclusion, this work is the first critical step to develop validation tool to quickly validate and explore gene networks in rice and beyond plants.

Materials and methods

scRNA sequence design

Aptamers sequences for MS2 (C5 variant) (Lowary and Uhlenbeck, 1987), and com (Hattman, 1999) were fused to the 3' extremity of the sgRNA sequences (Miao *et al.*, 2013) and cloned upstream a rice U3 snoRNA promoter (Pendrack and Roberts., 2011). scRNA sequence secondary structure and stability were evaluated using NUPACK (Zadeh *et al.*, 2011) to check their compatibility with CRISPR CAS9 sgRNA folding (annexe). Spacers for tetO and p35S sequences were designed using CRISPOR-TEFOR software (Haeussler et al., 2016, http://crispor.tefor.net/). The two scRNA were subcloned in a pOs-ENTRY vector (**Table 1**, Miao *et al.*, 2013). All scaffold, minimal sgRNA and spacer sequences can be found in annexe 3.

Coding protein sequence design

The rice codon-optimized dCas9 was developed starting with Cas9 sequence of (Miao *et al.*, 2013), by introduction of two mutations that inactivate the catalytic site of Cas9; the first one in the RuvC domain (D10A), the second one in the HNH domain (H840A) (Cheng *et al.* in 2013) (annexe 4). The dCas9 protein was then merged with one C-terminal VP64 (Beerli *et al.*, 1998) or 3xSRDX (Lowder *et al.*, 2015) domains to dCAS9 thanks with a 5xGS linker (Zalatan *et al.*, 2015, annexe 4). MS2 aptamer ligand, MCP (DFG/V29I mutant) (Lim and Peabody, 1994), and com aptamer ligand COM (Hattman, 1999) were fused to a C-terminal VP64 or 3xSRDX domain using a 5xGS linker plus a N-terminal SV40 NLS using a 1xGS linker (annexe 5). All these constructs were put under the control of the strong constitutive promoter pZmUbi (Muhitch et Shatters, 1998). The VENUS:NLS gene was cloned downstream the 1xTetO, the 7xTetO, and the modified 35S promoter. These three promoter were modified to introduce NGG PAM sequences (Promoter Adjacent Motif (Jinek *et al.*, 2012). Sequences of the constructs are described in **annexe 5** and **annexe 6**, and resume in the **Table 1**. All of these constructs were subcloned in standard pCAMBIA5300 binary vectors. All the constructs were codon optimized for rice. All the vectors were finally checked by sequencing.

Protoplast extraction and transformation

The protocol we used for protoplast extraction, vector cloning and transformation is detailed in Bes *et al.*, accepted in Methods in Molecular Biology (**Annexe 7**). Briefly, leaves of 9-days-old plants growth in a dark room were digested with a enzymatic solution in order to remove cell walls. After several purification steps to separate protoplasts from remaining wall debris, protoplasts were transfected or co-transfected using a PEG 40% solution with a total of 5 µg of each vector at a concentration of 1µg/µl . All the constructs were amplified in DH5α *E. coli* Competent Cells, Thermo Fisher (18265017) with the QIAGEN Plasmid Maxi Kit (Cat No./ID: 12163).

Visualization and quantification of the fluorescence

We used a Nikon ECLIPSE® epi-fluorescence microscope and a BD FACSAria™ III with a GFP filter. The FACS data were analyzed with Flowing Software as shown in **Annexe 2**.

3. Facteurs de transcription synthétiques CRISPR/Cas et protoplastes, vers un système expérimental permettant de tester les hypothèses du cycle DBTL

Tous les TFs que nous avons développés permettent une manipulation de l'expression d'un gène rapporteur en protoplaste de riz avec différents niveaux d'efficacité. Après un test des nouveaux TFs développés (CRISPR-Sirius) sur gène rapporteur et une preuve de concept sur gènes endogènes, l'objectif à terme est le ciblage des régulateurs potentiellement impliqués dans la formation des aérenchymes que l'on aura identifiés (étape *Build* du cycle DBTL, future collaboration avec l'équipe I3Bionet, *cf* chapitre 3). Les conséquences de la perturbation de l'expression de ces gènes candidats sur les transciptomes des populations de protoplastes transformées seront analysés par RT-ddPCR ou par séquençage (étape *Test* du cycle DBTL). Ces informations transcriptomiques seront réinjectées dans nos réseaux de gènes pour les affiner et émettre de nouvelles hypothèses (étape *Learn*).

Le système expérimental combinant TFs basé sur la technologie CRISPR/Cas et protoplastes de riz nous permet d'obtenir les premiers résultats dans un intervalle de quelques semaines entre la conception d'un scRNA ciblant de nouveaux gènes candidats et, les analyses. Ce système expérimental a déjà été utilisé avec succès pour effectuer une cartographie *de novo* de réseaux de régulation de gènes *via* une perturbation du transcriptome suivi d'une analyse de ces perturbations (Rubin et al., 2018, Shen et al., 2017, Najm et al., 2018). Néanmoins, la preuve de concept de l'efficacité de ce système expérimentale dans le cadre de l'étude de la biologie du développement est la redirection du destin cellulaire *in vitro via* la technologie CRISPR/Cas en cellules animales (Truong *et al.,* 2019, Hsu *et al.,* 2019).

DISCUSSION, CONCLUSION, ET PERSPECTIVES DU PROJET

DISCUSSION, CONCLUSION, ET PERSPECTIVES DU PROJET

1. Transcriptome tissu/stade développemental spécifique de la formation des aérenchymes, vers l'inférence d'un réseau de co-régulation de gènes.

Le protocole de LM développé ici permet d'isoler spécifiquement des groupes de tissus. Après extraction, l'ARN obtenu a un RIN moyen supérieur à 7, ce qui correspond à des échantillons de suffisamment bonne qualité pour limiter les biais d'analyses (Shroeder *et al.*, 2006, Endrullat *et al.*, 2016, Kashofer *et al.*, 2013, Derveaux *et al.*, 2010). Le couplage de la technologie LM à la ddPCR nous a permis d'effectuer un profilage différentiel tissu spécifique de l'expression de 6 gènes du riz, validant ainsi des profils d'expressions tissu-spécifiques obtenus lors d'études précédentes. Nous avons pu également constater la grande sensibilité de la technologie ddPCR qui est capable de quantifier l'expression de transcrits pour parfois des quantités d'ARN de moins de 1 pg.

Enfin, nous avons contribué à développer le seul kit présent actuellement sur le marché permettant de construire une banque d'ADNc à partir de 1ng d'ARN de riz (le kit descend à 10pg). Ce protocole intègre par ailleurs l'ensemble des technologies permettant de limiter les biais lors de la construction de banques RNAseq (pour une synthèse, cf Hrdlickova *et al.*, 2016, Dijk *et al.*, 2014). En effet, les autres kits ne permettent pas de réaliser une déplétion des ARN ribosomaux pour des quantités aussi faibles d'ARN. Ce kit a été utilisé sur nos échantillons d'ARN provenant de sections de 3 groupes de tissus dans la racine, dont le cortex, et ce sur les trois premières phases développementales aboutissant à la formation des aérenchymes à partir des cellules corticales (Kawai *et al.*, 1998), pour un total de 4 répétitions biologiques. Les banques d'ADNc sont terminées et ont été envoyées pour être séquencées au moment du dépôt du manuscrit. Les données de séquençage seront ensuite analysées par l'équipe I3Bionet avec qui une collaboration est en cours, afin d'inférer un réseau de co-régulation de gènes. Ce dernier point sera brièvement abordé plus bas.

2. Les facteurs de transcriptions synthétiques basés sur la technologie CRISPR/Cas9, des premiers résultats prometteurs et quelques pistes pour améliorer le système

La technologie CRISPR/Cas9, depuis sa découverte en 2012 par Jinek *et al.*, s'est largement répandue et de nombreuses technologies dérivées sont apparues, notamment pour réaliser des facteurs de transcriptions synthétiques (pour des revues récentes Mitsunobu *et al.*, 2017, Xu *et al.*, 2019). Nous avons testé ici trois types différents de ces facteurs de transcriptions artificiels, et avons pu démontrer avec succès leur fonctionnalité en utilisant un gènes rapporteur. Il reste à démontrer que ces constructions sont capables d'activer des gènes endogènes du riz. Néanmoins, un seul des systèmes testé, le dCas9-scRNA, permet un contrôle orthogonal (activation ET inhibition) de l'expression d'un jeu de gènes. Cette technologie est basée sur l'ajout d'une séquence appelée aptamère à l'ARN guide, aptamère dont le ligand protéique le reconnaissant est fusionné à un domaine activateur ou répresseur de la transcription (Mali *et al.*, 2013). Deux facteurs semblent entrer en jeu dans l'efficacité des différents systèmes utilisés : 1) le nombre de domaines protéiques activateurs ou répresseurs recrutés au niveau du promoteur du gène ciblé, 2) la stabilité de l'ARN guide (Konnermann *et al.*, 2015, Nowak *et al.*, 2016, Chavez *et al.*, 2016). Confirmant ces hypothèses, Ma *et al.* en 2018, ont créé la technologie CRISPR-Sirius, qui permet le recrutement de huit fluorochromes au même locus tout en démontrant la plus grande stabilité de l'ARN guide, améliorant de manière significative les précédents systèmes qui avaient été testés en imagerie *in vivo* utilisant l'approche dCAS9 (Ma *et al.*, 2016). Pour cela, huit séquences d'aptamères ont été ajoutées les uns à la suite des autres dans le prolongement du tetraloop de l'ARN guide (Nishimasu *et al.*, 2014). Nous avons adapté ces modifications à notre ARN guide en utilisant deux aptamères différents (MS2 et com, Lowary and Uhlenbeck, 1987, Hattman, 1999). Ces deux constructions seront prochainement testées pour leur efficacité à manipuler l'expression de notre gène rapporteur à l'aide de la construction 7xtetO:VENUS:NLS.

3. Approche comparative riz-sorgho pour élucider les mécanismes moléculaires de formation des aérenchymes chez les céréales: cas du gène SbNAC.D du Sorgho.

De récentes publications décrivent le rôle d'un gène, SbNAC.D, dans la formation des aérenchymes de la tige chez le Sorgho. A l'origine de cette découverte, des études portant sur la recherche de QTL (*Quantitative trait loci*) qui influencent l'accumulation de sucre dans la tige de différentes populations de Sorgho ont identifié un QTL majeur, qui a été nommé D et qui fait partie de la famille des FTs NAC (NAM, ATAF1,2, CUC-like protein Burks *et al.*, 2015, Srinivas *et al*., 2009, Xu *et al*., 2000). L'accumulation de sucre dans la tige du sorgho est négativement corrélée avec la formation d'aérenchymes. En 2018, une investigation plus poussée menée sur ce QTL a permis d'identifier un gène dont la version tronquée est responsable de la perte des aérenchymes constitutifs dans la tige chez les variétés de sorgho sucrés (Casto et al., 2018). Cependant, les variétés de sorgho sucrées ont conservé leur aptitude à développer des aérenchymes racinaires inductibles en condition de submergence (Promkhambut *et al.,* 2011, Tari et al., 2013). Cela suggère qu'il existe des mécanismes de régulation distincts entre le développement des aérenchymes de la tige et de la racine. La présence de cette protéine exclusivement sous sa forme tronquée chez les variétés de blé et de canne à sucre, qui ne développent pas d'aérenchyme constitutif vient renforcer cette hypothèse (Fujimoto et al., 2018). Il a également été démonté que l'expression de sa version non tronquée, ainsi que celle de ses homologues chez le riz (OsD) et Arabidopsis (ANAC074, **Figure 37**) était suffisante pour déclencher la mort cellulaire en cellule isolée d'Arabidopsis (Fujimoto et al., 2018), confirmant que le gène D et ses paralogues sont probablement des candidats pertinents pour la formation des aérenchymes chez les céréales. Dans un premier temps, les données RNAseq en cours de séquençage pourront être utilisées pour obtenir un profil l'expression, au cours de la formation des aérenchymes, de l'orthologue du gène SbNAC.D chez le riz mais également de ses paralogues. La capacité de ces FTs NAC candidats à induire la mort cellulaire pourra alors être testé à l'aide des constructions dCas9-scRNA que nous avons développés, et leur fonction *in planta* et leur rôle potentiel dans la formation des aérenchymes, sera testé à l'aide de la technologie CRISPR/dCas9.

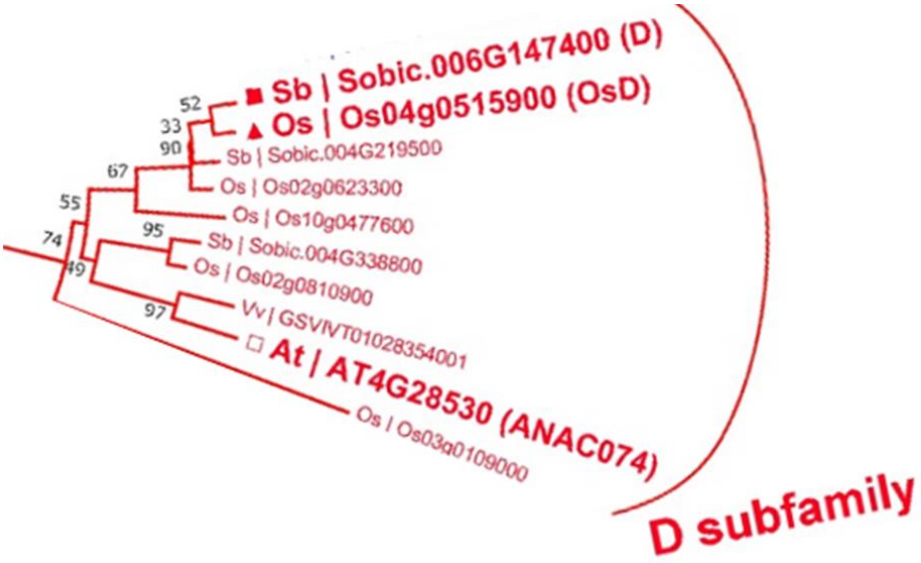


Figure 37: Homologues du gène SbNAC.D du Sorgho et ANAC074 d'Arabidopsis chez différentes espèces, dont le riz (OsD)
D'après Fujimoto *et al*., 2017.

4. Vers une approche DBTL pour l'étude des mécanismes moléculaires de formation des aérenchymes racinaires

L'ère du *big data* et des nouveaux outils moléculaires comme CRISPR/Cas a ouvert la voie à de nouvelles approches en biologie, notamment les approches regroupées sous le terme de *data-driven* (Li et Chen, 2014, Farré *et al.*, 2015, Chang *et al.*, 2011, Wang *et al.*, 2013). Sous ce terme sont regroupé les approches se basant sur 1) l'exploration et l'intégration de données -omiques massives ; 2) des analyses d'ontologie permettant de découvrir et d'agréger des informations moléculaires ; 3) le développement d'outils de visualisation des données permettant d'émette des hypothèses quant aux interactions moléculaires probables dans un contexte biologique spécifique (Greene et Troyanskaya, 2010). Dans cette tendance, mon projet de recherche se veut être une première contribution au développement d'une stratégie DBTL complète chez le riz (**fFgure 38**, pour une revue en plante Pouvreau *et al.*, 2018, pour une première application en plante : Vanhercke et al., 2017, Mortimer, 2019). Cette thèse et la future collaboration avec l'équipe I3Bionet pour la reconstruction et l'analyse de réseaux de gènes, permettra 1) d'identifier des gènes candidats *via* une modélisation des réseaux de gènes qui mettra en évidence les régulateurs majeurs impliqués dans les transition entre les différentes phases développementales du cortex (« Learn-Design»), 2) de construire des outils moléculaires adaptés afin de manipuler finement l'expression des gènes identifiés en cellules isolées (« Build-Learn »), 3) de réinjecter les informations (-omiques) obtenues par expérimentations dans les modèles afin de les affiner (« Learn »), grâce aux approches machine Learning (**Figure 38**), 4) d'orienter enfin les tests *in planta* qui sont long, laborieux, et couteux en limitant leur nombre.

Cette approche est rendue possible en partie grâce aux outils informatiques qui ont été développé au sein de l'équipe I3Bionet, regroupé sous forme de *packages* informatique comme CoRegNet (Nicolle *et al.*, 2015). Ce dernier a déjà été utilisé avec succès pour rediriger le métabolisme lipidique de *Yarrowia lipolytica* (Trébulle *et al.*, 2017).

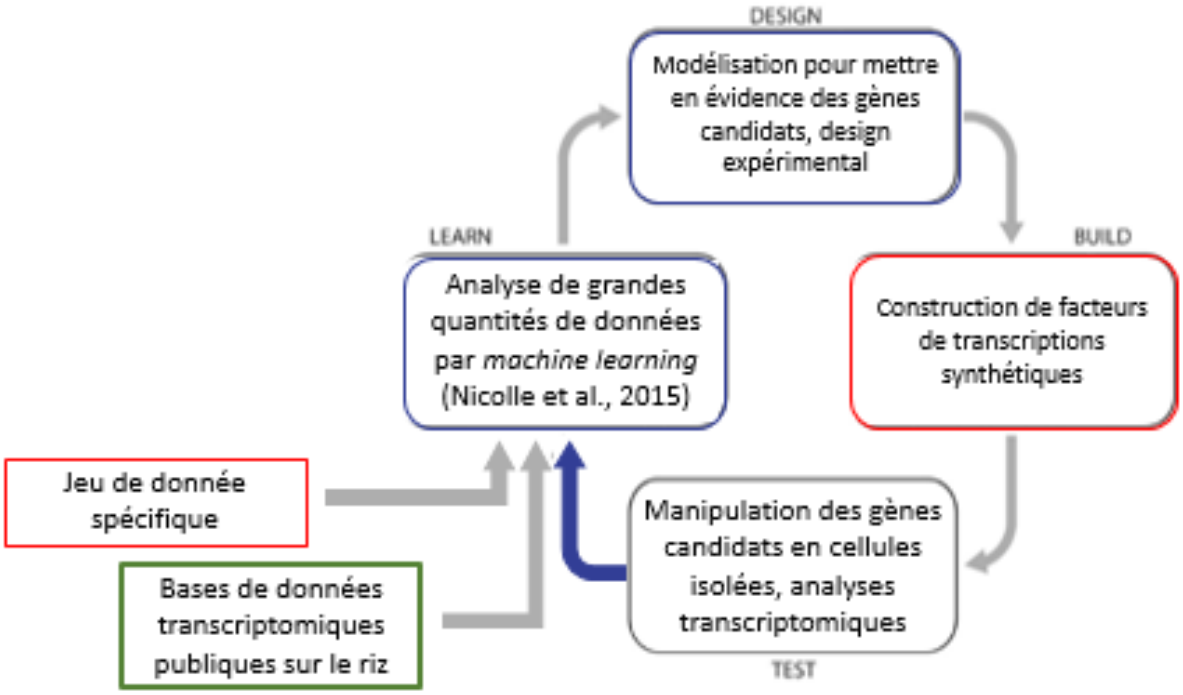


Figure 38: Représentation schématique de l'approche DBTL

Le principe du cycle DBTL est d'affiner des connaissances sur un processus biologique de manière itérative. En quelques mots, le cycle peut se résumer en 4 étapes 1) « Design », choisir une action à mener sur des gènes candidats en se basant sur une modélisation in silico d'un processus biologique à partir de bases de données -omiques, 2) « Build », construire des outils moléculaires permettant de manipuler ces gènes, 3) « Test », mener une expérimentation en système biologique et en extraire des informations –omiques 4) « Learn », utiliser des algorithmes de machine learning afin d'analyser les données et de ré-affiner la partie « Design ».

Néanmoins, pour commencer le cycle d'apprentissage itératif, il est nécessaire de disposer d'un grand nombre de jeux de données de transcriptome (encadré vert) et d'un jeu de données spécifique à la question biologique étudiée (encadré rouge, Nicolle et al., 2015). L'un de mes objectifs est de créer ce jeu de donnée spécifique à l'étude de la formation des aérenchymes (encadré rouge). Mon deuxième objectif est le développement d'un système expérimental en cellule isolée avec des facteurs de transcriptions synthétiques basés sur la technologie CRISPR/Cas. Les autres étapes du cycle entourées en bleu feront l'objet d'une future collaboration avec l'équipe I3Bionet

A partir de données transcriptomiques massives, et donc d'une multitude de jeux de données, ce *package* permet de construire un réseau de co-régulation de gènes *via* un algorithme de *machine learning* appelé h-LICORN (hybrid Learning cooperative regulation networks from gene expression data, Elati *et al.*, 2007, Chebil *et al.*, 2014). Des données massives sont maintenant disponibles chez le riz et chez de nombreuses autres plantes. Ainsi, dans notre cas, cet algorithme pourra être entrainé sur des bases de données issues de séquençages comme TENOR (https://tenor.dna.affrc.go.jp, Kawahara *et al.*, 2019), mais également sur des données d'expression provenant de puces à ADN. Ainsi, des bases de données contenant déjà des informations tissus-spécifiques comme Rice-Xpro (http://ricexpro.dna.affrc.go.jp/) pourront être utilisés. Cet algorithme nécessite également la distinction entre les régulateurs (facteurs de transcriptions, kinase et autres) et les effecteurs. Là encore, des bases de données comme PlantTFDB (http://planttfdb.gao-lab.org/link.php, Jin *et al.*, 2017) pourront être utilisées.

Les données d'expression tissus/phase développementale spécifique provenant de ma thèse serviront à la construction de ce réseau, et ce jeu de données nous servira de base de travail. En effet, CoRegNet dispose aussi d'un outil de visualisation interactif permettant de mettre en évidence les combinaisons de régulateurs actifs au sein d'un jeu de donnée spécifique. Ce qui, par conséquent, permet également la comparaison de ces différentes combinaisons dans les différents jeux de données ayant servi à construire le réseau de gènes. **En d'autres termes, dans notre cas, cet outil de visualisation nous permettra de suivre l'évolution des combinaisons de régulateurs les plus influents dans les différentes phases de développement du cortex que l'on a isolé, ou de comparer les combinaisons de régulateurs actifs entre les différents groupes de tissus pour une même phase de développement.** Nous espérons ainsi mettre en évidence quelques gènes candidats, par exemple les gènes permettant d'initier la mort cellulaire dans le cortex. Leur validation se fera en protoplastes (dCAS9-scRNA), puis pour les candidats les plus pertinents, *in planta* par CRISPR/Cas9. **Le *package* CoRegNet permet enfin d'améliorer progressivement la robustesse du réseau de co-régulation de gènes en intégrant des données expérimentales pour compléter le cycle d'apprentissage itératif.**

J'espère, à travers cette thèse, avoir contribuer au développement des approches DBTL chez le riz en développant des outils et des ressources moléculaires adaptés.

L'ensemble des ressources moléculaires, bases de données et outils expérimentaux, dont on dispose maintenant sur une espèces d'intérêt agronomiques et modèle comme le riz, ouvrent en la voie à de nouvelles approches dans le domaine de la biologie fonctionnelle et de la biologie des systèmes (Pouvreau *et al.*, 2018, Farré *et al.*, 2015 pour des revues sur les plantes, Li et Chen, 2014, Chang *et al.*, 2011, Carbonell *et al.*, 2018, Liu *et al.*, 2015 pour des revues sur des micro-organismes).

REFERENCES

REFERENCES GENERALES

Abiko, T., & Obara, M. (2014). Enhancement of porosity and aerenchyma formation in nitrogen-deficient rice roots. *Plant Science*, *215–216*, 76–83. https://doi.org/10.1016/J.PLANTSCI.2013.10.016

Abiri, R., Shaharuddin, N. A., Maziah, M., Yusof, Z. N. B., Atabaki, N., Sahebi, M., … Hanafi, M. M. (2017). Role of ethylene and the APETALA 2/ethylene response factor superfamily in rice under various abiotic and biotic stress conditions. *Environmental and Experimental Botany*, *134*, 33–44. https://doi.org/10.1016/J.ENVEXPBOT.2016.10.015

Adiconis, X., Borges-Rivera, D., Satija, R., DeLuca, D. S., Busby, M. A., Berlin, A. M., … Levin, J. Z. (2013). Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nature Methods*, *10*(7), 623–629. https://doi.org/10.1038/nmeth.2483

Adomas, A. B., Lopez-Giraldez, F., Clark, T. A., Wang, Z., & Townsend, J. P. (2010). Multi-targeted priming for genome-wide gene expression assays. *BMC Genomics*, *11*(1), 477. https://doi.org/10.1186/1471-2164-11-477

Ahmadi, N., Audebert, A., Bennett, M. J., Bishopp, A., de Oliveira, A. C., Courtois, B., … Xu, J. (2014). The roots of future rice harvests. *Rice (New York, N.Y.)*, *7*(1), 29. https://doi.org/10.1186/s12284-014-0029-y

Akhunov, E. D., Sehgal, S., Liang, H., Wang, S., Akhunova, A. R., Kaur, G., … Gill, B. S. (2013). Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. *Plant Physiology*, *161*(1), 252–265. https://doi.org/10.1104/pp.112.205161

Andersen, T. G., Barberon, M., & Geldner, N. (2015). Suberization — the second life of an endodermal cell. *Current Opinion in Plant Biology*, *28*, 9–15. https://doi.org/10.1016/J.PBI.2015.08.004

Arai-Sanoh, Y., Takai, T., Yoshinaga, S., Nakano, H., Kojima, M., Sakakibara, H., … Uga, Y. (2015). Deep rooting conferred by DEEPER ROOTING 1 enhances rice yield in paddy fields. *Scientific Reports*, *4*(1), 5563. https://doi.org/10.1038/srep05563

Armstrong, W. (1980). Aeration in Higher Plants. *Advances in Botanical Research*, *7*, 225–332. https://doi.org/10.1016/S0065-2296(08)60089-0

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., … Matsuoka, M. (2005). Cytokinin Oxidase Regulates Rice Grain Production. *Science*, *309*(5735), 741–745. https://doi.org/10.1126/SCIENCE.1113373

Astashkina, A., Mann, B., & Grainger, D. W. (2012). A critical evaluation of in vitro cell culture models for high-throughput drug screening and toxicity. *Pharmacology & Therapeutics*, *134*(1), 82–106. https://doi.org/10.1016/j.pharmthera.2012.01.001

Baltes, N. J., & Voytas, D. F. (2015). Enabling plant synthetic biology through genome engineering. *Trends in Biotechnology*, *33*(2), 120–131. https://doi.org/10.1016/J.TIBTECH.2014.11.008

Banf, M., & Rhee, S. Y. (2017). Computational inference of gene regulatory networks: Approaches, limitations and opportunities. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1860*(1), 41–52. https://doi.org/10.1016/J.BBAGRM.2016.09.003

Bargmann, B. O. R., Marshall-Colon, A., Efroni, I., Ruffel, S., Birnbaum, K. D., Coruzzi, G. M., & Krouk, G. (2013). TARGET: a transient transformation system for genome-wide transcription factor target discovery. *Molecular Plant*, *6*(3), 978–980. https://doi.org/10.1093/mp/sst010

Beerli, R. R., Segal, D. J., Dreier, B., & Barbas, C. F. (1998). Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(25), 14628–14633. https://doi.org/10.1073/pnas.95.25.14628

Benfey, P. N., Linstead, P. J., Roberts, K., Schiefelbein, J. W., Hauser, M. T., & Aeschbacher, R. A. (1993). Root development in Arabidopsis: Four mutants with dramatically altered root morphogenesis. *Development*, *119*(1), 57–70. https://dev.biologists.org/content/119/1/57.short

Bhushan C., Singh G., Rattani V. and Kumar V. (2016). Insuring agriculture in times of climate change, Centre for Science and Environment, New Delhi Retrieved from http://www.indiaenvironmentportal.org.in/files/file/Report on Agricultural Insurance.pdf

Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., & Benfey, P. N. (2003). A Gene Expression Map of the Arabidopsis Root. *Science*, *302*(5652), 1956–1960. https://doi.org/10.1126/science.1090022

Bisht, D. S., Solanke, A. U., & Mondal, T. K. (2018). Informatics of Wild Relatives of Rice (pp. 27–40). Springer, Cham. https://doi.org/10.1007/978-3-319-71997-9_2

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., … Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science (New York, N.Y.)*, *326*(5959), 1509–1512. https://doi.org/10.1126/science.1178811

Boettcher, M., Tian, R., Blau, J. A., Markegard, E., Wagner, R. T., Wu, D., … McManus, M. T. (2018). Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nature Biotechnology*, *36*(2), 170–178. https://doi.org/10.1038/nbt.4062

Bortesi, L., & Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, *33*, 41–52. https://doi.org/10.1016/j.biotechadv.2014.12.006

Burks, P. S., Kaiser, C. M., Hawkins, E. M., & Brown, P. J. (2015). Genomewide Association for Sugar Yield in Sweet Sorghum. *Crop Science*, *55*(5), 2138. https://doi.org/10.2135/cropsci2015.01.0057

Butt, H., Eid, A., Momin, A. A., Bazin, J., Crespi, M., Arold, S. T., & Mahfouz, M. M. (2019). CRISPR directed evolution of the spliceosome for resistance to splicing inhibitors. *Genome Biology*, *20*(1), 73. https://doi.org/10.1186/s13059-019-1680-9

Cai, Z. C., Tsuruta, H., & Minami, K. (2000). Methane emission from rice fields in China: Measurements and influencing factors. *Journal of Geophysical Research: Atmospheres*, *105*(D13), 17231–17242. https://doi.org/10.1029/2000JD900014

Caldana, C., Scheible, W.-R., Mueller-Roeber, B., & Ruzicic, S. (2007). A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. *Plant Methods*, *3*(1), 7. https://doi.org/10.1186/1746-4811-3-7

Callaway, E. (2014). Domestication: The birth of rice. *Nature*, *514*(7524), S58–S59. https://doi.org/10.1038/514S58a

Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., … Scrutton, N. S. (2018). An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Communications Biology*, *1*(1), 66. https://doi.org/10.1038/s42003-018-0076-9

Casbon, J. A., Osborne, R. J., Brenner, S., & Lichtenstein, C. P. (2011). A method for counting PCR template molecules with application to next-generation sequencing. *Nucleic Acids Research*, *39*(12), e81–e81. https://doi.org/10.1093/nar/gkr217

Cassman, K. G., Kropff, M. J., Gaunt, J., & Peng, S. (1993). Nitrogen use efficiency of rice reconsidered: What are the key constraints? *Plant and Soil*, *155–156*(1), 359–362. https://doi.org/10.1007/BF00025057

Casto, A. L., McKinley, B. A., Yu, K. M. J., Rooney, W. L., & Mullet, J. E. (2018). Sorghum stem aerenchyma formation is regulated by SbNAC_D during internode development. *Plant Direct*, *2*(11), e00085. https://doi.org/10.1002/pld3.85

Chang R., Shoemaker, R., & Wei Wang. (2011). A Novel Knowledge-Driven Systems Biology Approach for Phenotype Prediction upon Genetic Intervention. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *8*(5), 1170–1182. https://doi.org/10.1109/TCBB.2011.18

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., P R Iyer, E., … Church, G. M. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nature Methods*, *12*(4), 326–328. https://doi.org/10.1038/nmeth.3312

Chavez, A., Tuttle, M., Pruitt, B. W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., … Church, G. (2016). Comparison of Cas9 activators in multiple species. *Nature Methods*, *13*(7), 563–567. https://doi.org/10.1038/nmeth.3871

Chebil, I., Nicolle, R., Santini, G., Rouveirol, C., & Elati, M. (2014). Hybrid Method Inference for the Construction of Cooperative Regulatory Network in Human. *IEEE Transactions on NanoBioscience*, *13*(2), 97–103. https://doi.org/10.1109/TNB.2014.2316920

Cheng, A. W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T. W., … Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Research*, *23122*(23). https://doi.org/10.1038/cr.2013.122

Choi, J. Y., & Purugganan, M. D. (2018). Multiple origin but single domestication led to Oryza sativa. *G3: Genes, Genomes, Genetics*, *8*(3), 797–803. https://doi.org/10.1534/g3.117.300334

Choudhury, S., Panda, P., Sahoo, L., & Panda, S. K. (2013). Reactive oxygen species signaling in plants under abiotic stress. *Plant Signaling & Behavior*, *8*(4), e23681. https://doi.org/10.4161/psb.23681

Civán, P., Craig, H., Cox, C. J., & Brown, T. A. (2016). Multiple domestications of Asian rice. *Nature Plants*, *2*(4), 16037. https://doi.org/10.1038/nplants.2016.37

Clowes, F. A. L. (2000). Pattern in root meristem development in angiosperms. *New Phytologist*, *146*(1), 83–94. https://doi.org/10.1046/j.1469-8137.2000.00614.x

Coenen, C., Liedtke, S., & Kogler, G. (2015). RNA Amplification Protocol Leads to Biased Polymerase Chain Reaction Results Especially for Low-Copy Transcripts of Human Bone Marrow-Derived Stromal Cells. *PLOS ONE*, *10*(10), e0141070. https://doi.org/10.1371/journal.pone.0141070

Colmer, T. D., Cox, M. C. H., & Voesenek, L. A. C. J. (2006). Root aeration in rice (Oryza sativa): Evaluation of oxygen, carbon dioxide, and ethylene as possible regulators of root acclimatizations. *New Phytologist*, *170*(4), 767–778. https://doi.org/10.1111/j.1469-8137.2006.01725.x

Colmer, T. D., & Flowers, T. J. (2008). Flooding tolerance in halophytes. *New Phytologist*, *179*(4), 964–974. https://doi.org/10.1111/j.1469-8137.2008.02483.x

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., … Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, *17*(1), 13. https://doi.org/10.1186/s13059-016-0881-8

Coons, A. H., & Kaplan, M. H. (1949). Localization of antigen in tissue cells: ii. improvements in a method for the detection of antigen by means of fluorescent antibody. *Journal of Experimental Medicine*, *91*(1), 1–13. https://doi.org/10.1084/jem.91.1.1

Courtois, B., Ahmadi, N., Khowaja, F., Price, A. H., Rami, J.-F., Frouin, J., … Ruiz, M. (2009). Rice Root Genetic Architecture: Meta-analysis from a Drought QTL Database. *Rice*, *2*(2–3), 115–128. https://doi.org/10.1007/s12284-009-9028-9

Cramer, G. R., Urano, K., Delrot, S., Pezzotti, M., & Shinozaki, K. (2011). Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biology*, *11*(1), 163. https://doi.org/10.1186/1471-2229-11-163

Cui, H., Levesque, M. P., Vernoux, T., Jung, J. W., Paquette, A. J., Gallagher, K. L., … Benfey, P. N. (2007). An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. *Science (New York, N.Y.)*, *316*(5823), 421–425. https://doi.org/10.1126/science.1139531

Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., … Bock, C. (2017). Pooled CRISPR screening with single-cell transcriptome readout. *Nature Methods*, *14*(3), 297–301. https://doi.org/10.1038/nmeth.4177

Day, R. C., & Macknight, R. C. (2005). Be more specific! Laser-assisted microdissection of plant cells Comparative transcriptomics of sexual and apomictic Boechera species View project Assessing differential DNA methylation with Jensen-Shannon divergence View project. https://doi.org/10.1016/j.tplants.2005.06.006

Deal, R. B., & Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Developmental Cell*, *18*(6), 1030–1040. https://doi.org/10.1016/j.devcel.2010.05.013

Deal, R. B., & Henikoff, S. (2011). The INTACT method for cell typeg-specific gene expression and chromatin profiling in Arabidopsis thaliana. *Nature Protocols*, *6*(1), 56–68. https://doi.org/10.1038/nprot.2010.175

de Dorlodot, S., Forster, B., Pagès, L., Price, A., Tuberosa, R., & Draye, X. (2007). Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends in Plant Science*, *12*(10), 474–481. https://doi.org/10.1016/J.TPLANTS.2007.08.012

Delseny, M. (2003). Science, médiatisation et politique : le séquençage du génome du riz. M/S : médecine sciences 19(4). https://id.erudit.org/iderudit/006507ar

Demir, E., Cary, M. P., Paley, S., Fukuda, K., Lemer, C., Vastrik, I., … Bader, G. D. (2010). The BioPAX community standard for pathway data sharing. *Nature Biotechnology*, *28*(9), 935–942. https://doi.org/10.1038/nbt.1666

Derveaux, S., Vandesompele, J., & Hellemans, J. (2010). How to do successful gene expression analysis using real-time PCR. *Methods*, *50*(4), 227–230. https://doi.org/10.1016/J.YMETH.2009.11.001

Díaz-Triviño, S., Long, Y., Scheres, B., & Blilou, I. (2017). Analysis of a Plant Transcriptional Regulatory Network Using Transient Expression Systems (pp. 83–103). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-7125-1_7

Didovyk, A., Borek, B., Tsimring, L., & Hasty, J. (2016). Transcriptional regulation with CRISPR-Cas9: principles, advances, and applications. *Current Opinion in Biotechnology*, *40*, 177–184. https://doi.org/10.1016/J.COPBIO.2016.06.003

Dievart, A., Coudert, Y., Gantet, P., Pauluzzi, G., Puig, J., Fanchon, D., … Périn, C. (2013). Dissection des bases biologiques de caractéres d'intérêtchez le riz: Architecture et développement du systéme racinaire. *Cahiers Agricultures*, *22*(5), 475–483. https://doi.org/10.1684/agr.2013.0651

Dinneny, J. R., Long, T. A., Wang, J. Y., Jung, J. W., Mace, D., Pointer, S., … Benfey, P. N. (2008). Cell identity mediates the response of Arabidopsis roots to abiotic stress. *Science*, *320*(5878), 942–945. https://doi.org/10.1126/science.1153795

Doblas, V. G., Geldner, N., & Barberon, M. (2017). The endodermis, a tightly controlled barrier for nutrients. *Current Opinion in Plant Biology*, *39*, 136–143. https://doi.org/10.1016/J.PBI.2017.06.010

Dominguez, A. A., Lim, W. A., & Qi, L. S. (2016). Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. *Nature Reviews Molecular Cell Biology*, *17*(1), 5–15. https://doi.org/10.1038/nrm.2015.2

Donald, C. M. (1968). The breeding of crop ideotypes. *Euphytica*, *17*(3), 385–403. https://doi.org/10.1007/BF00056241

Drapek, C., Sparks, E. E., Marhavy, P., Taylor, I., Andersen, T. G., Hennacy, J. H., … Benfey, P. N. (2018). Minimum requirements for changing and maintaining endodermis cell identity in the Arabidopsis root. *Nature Plants*, *4*(8), 586–595. https://doi.org/10.1038/s41477-018-0213-y

Dueck, H. R., Ai, R., Camarena, A., Ding, B., Dominguez, R., Evgrafov, O. V., … Kim, J. (2016). Assessing characteristics of RNA amplification methods for single cell RNA sequencing. *BMC Genomics*, *17*(1), 966. https://doi.org/10.1186/s12864-016-3300-3

Durán, N., Justo, G. Z., Durán, M., Brocchi, M., Cordi, L., Tasic, L., … Nakazato, G. (2016). Advances in Chromobacterium violaceum and properties of violacein-Its main secondary metabolite: A review. *Biotechnology Advances*, *34*(5), 1030–1045. https://doi.org/10.1016/j.biotechadv.2016.06.003b

Elati, M., Neuvial, P., Bolotin-Fukuhara, M., Barillot, E., Radvanyi, F., & Rouveirol, C. (2007). LICORN: learning cooperative regulation networks from gene expression data. *Bioinformatics*, *23*(18), 2407–2414. https://doi.org/10.1093/bioinformatics/btm352

EL-MAAROUF-BOUTEAU, H., SAJJAD, Y., BAZIN, J., LANGLADE, N., CRISTESCU, S. M., BALZERGUE, S., … BAILLY, C. (2015). Reactive oxygen species, abscisic acid and ethylene interact to regulate sunflower seed germination. *Plant, Cell & Environment*, *38*(2), 364–374. https://doi.org/10.1111/pce.12371

Emmert-Streib, F., & Dehmer, M. (2011). Networks for systems biology: Conceptual connection of data and function. *IET Systems Biology*, *5*(3), 185–207. https://doi.org/10.1049/iet-syb.2010.0025

Endrullat, C., Glökler, J., Franke, P., & Frohme, M. (2016). Standardization and quality management in next-generation sequencing. *Applied & Translational Genomics*, *10*, 2–9. https://doi.org/10.1016/J.ATG.2016.06.001

Evrard, A., Bargmann, B. O. R., Birnbaum, K. D., Tester, M., Baumann, U., & Johnson, A. A. T. (2012). Fluorescence-activated cell sorting for analysis of cell type-specific responses to salinity stress in arabidopsis and rice. *Methods in Molecular Biology*, *913*, 265–276. https://doi.org/10.1007/978-1-61779-986-0_18

Faherty, S. L., Campbell, C. R., Larsen, P. A., & Yoder, A. D. (2015). Evaluating whole transcriptome amplification for gene profiling experiments using RNA-Seq. *BMC Biotechnology*, *15*(1), 65. https://doi.org/10.1186/s12896-015-0155-7

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., … Zhang, Q. (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and Applied Genetics*, *112*(6), 1164–1171. https://doi.org/10.1007/s00122-006-0218-1

FAO (2015) The impact of natural hazards and disasters on agriculture and food security and nutrition. http://www.fao.org/3/a-i4434e.pdf, (accessed).

Farré, G., Twyman, R. M., Christou, P., Capell, T., & Zhu, C. (2015). Knowledge-driven approaches for engineering complex metabolic pathways in plants. *Current Opinion in Biotechnology*, *32*, 54–60. https://doi.org/10.1016/j.copbio.2014.11.004

Farris, S., Wang, Y., Ward, J. M., & Dudek, S. M. (2017). Optimized Method for Robust Transcriptome Profiling of Minute Tissues Using Laser Capture Microdissection and Low-Input RNA-Seq. *Frontiers in Molecular Neuroscience*, *10*, 185. https://doi.org/10.3389/fnmol.2017.00185

Fu, G. K., Xu, W., Wilhelmy, J., Mindrinos, M. N., Davis, R. W., Xiao, W., & Fodor, S. P. A. (2014). Molecular indexing enables quantitative targeted RNA sequencing and reveals poor efficiencies in standard library preparations. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(5), 1891–1896. https://doi.org/10.1073/pnas.1323732111

Fujimoto, M., Sazuka, T., Oda, Y., Kawahigashi, H., Wu, J., Takanashi, H., … Tsutsumi, N. (2018). Transcriptional switch for programmed cell death in pith parenchyma of sorghum stems. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(37), E8783–E8792. https://doi.org/10.1073/pnas.1807501115

Fujita, M., Fujita, Y., Noutoshi, Y., Takahashi, F., Narusaka, Y., Yamaguchi-Shinozaki, K., & Shinozaki, K. (2006). Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current Opinion in Plant Biology*, *9*(4), 436–442. https://doi.org/10.1016/J.PBI.2006.05.014

Gale, M. D., & Devos, K. M. (1998). Comparative genetics in the grasses. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(5), 1971–1974. https://doi.org/10.1073/pnas.95.5.1971

Gallego Romero, I., Pai, A. A., Tung, J., & Gilad, Y. (2014). RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biology*, *12*(1), 42. https://doi.org/10.1186/1741-7007-12-42

Galindo-Castañeda, T., Brown, K. M., & Lynch, J. P. (2018). Reduced root cortical burden improves growth and grain yield under low phosphorus availability in maize. *Plant Cell and Environment*, *41*(7), 1579–1592. https://doi.org/10.1111/pce.13197

Galindo-Castañeda, T., Brown, K. M., & Lynch, J. P. (2018). Reduced root cortical burden improves growth and grain yield under low phosphorus availability in maize. *Plant Cell and Environment*, *41*(7), 1579–1592. https://doi.org/10.1111/pce.13197

Gall, J. G., & Pardue, M. L. (1969). Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proceedings of the National Academy of Sciences of the United States of America*, *63*(2), 378–383. https://doi.org/10.1073/pnas.63.2.378

Gamuyao, R., Chin, J. H., Pariasca-Tanaka, J., Pesaresi, P., Catausan, S., Dalid, C., … Heuer, S. (2012). The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature*, *488*(7412), 535–539. https://doi.org/10.1038/nature11346

Gao, J.-P., Chao, D.-Y., & Lin, H.-X. (2007). Understanding Abiotic Stress Tolerance Mechanisms: Recent Studies on Stress Response in Rice. *Journal of Integrative Plant Biology*, *49*(6), 742–750. https://doi.org/10.1111/j.1744-7909.2007.00495.x

Gautam, V., & Sarkar, A. K. (2015, April 18). Laser Assisted Microdissection, an Efficient Technique to Understand Tissue Specific Gene Expression Patterns and Functional Genomics in Plants. *Molecular Biotechnology*. Springer US. https://doi.org/10.1007/s12033-014-9824-3

Gautam, V., Singh, A., Singh, S., & Sarkar, A. K. (2016). An Efficient LCM-Based Method for Tissue Specific Expression Analysis of Genes and miRNAs. https://doi.org/10.1038/srep21577

Geldner, N. (2013). The Endodermis. *Annual Review of Plant Biology*, *64*(1), 531–558. https://doi.org/10.1146/annurev-arplant-050312-120050

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., … Qi, L. S. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, *154*(2), 442–451. https://doi.org/10.1016/j.cell.2013.06.044

Greene, C. S., & Troyanskaya, O. G. (2010). Integrative Systems Biology for Data-Driven Knowledge Discovery. *Seminars in Nephrology*, *30*(5), 443–454. https://doi.org/10.1016/J.SEMNEPHROL.2010.07.002

Gross, B. L., & Zhao, Z. (2014). Archaeological and genetic insights into the origins of domesticated rice. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1308942110

Gutiérrez-Aguirre, I., Rački, N., Dreo, T., & Ravnikar, M. (2015). Droplet Digital PCR for Absolute Quantification of Pathogens (pp. 331–347). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-2620-6_24

Hao, L., Ruiying, Q., Xiaoshuang, L., Shengxiang, L., Rongfang, X., Jianbo, Y., & Pengcheng, W. (2019). CRISPR/Cas9-Mediated Adenine Base Editing in Rice Genome. *Rice Science*, *26*(2), 125–128. https://doi.org/10.1016/J.RSCI.2018.07.002

Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., … Concordet, J.-P. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biology*, *17*(1), 148. https://doi.org/10.1186/s13059-016-1012-2

Hassan, H., Scheres, B., & Blilou, I. (2010). JACKDAW controls epidermal patterning in the Arabidopsis root meristem through a non-cell-autonomous mechanism. *Development (Cambridge, England)*, *137*(9), 1523–1529. https://doi.org/10.1242/dev.048777

Hattman, S. (1999). Unusual transcriptional and translational regulation of the bacteriophage Mu mom operon. *Pharmacology & Therapeutics*, *84*(3), 367–388. https://doi.org/10.1016/S0163-7258(99)00042-X

Hecker, M., Lambeck, S., Toepfer, S., van Someren, E., & Guthke, R. (2009). Gene regulatory network inference: Data integration in dynamic models—A review. *Biosystems*, *96*(1), 86–103. https://doi.org/10.1016/J.BIOSYSTEMS.2008.12.004

Heckl, D., & Charpentier, E. (2015). Toward Whole-Transcriptome Editing with CRISPR-Cas9. *Molecular Cell*, *58*(4), 560–562. https://doi.org/10.1016/J.MOLCEL.2015.05.016

Henry, S., Divol, F., Bettembourg, M., Bureau, C., Guiderdoni, E., Périn, C., & Diévart, A. (2016). Immunoprofiling of Rice Root Cortex Reveals Two Cortical Subdomains. *Frontiers in Plant Science*, *6*, 1139. https://doi.org/10.3389/fpls.2015.01139

Henry, S., Dievart, A., Divol, F., Pauluzzi, G., Meynard, D., Swarup, R., … Périn, C. (2017). SHR overexpression induces the formation of supernumerary cell layers with cortex cell identity in rice. *Developmental Biology*, *425*(1), 1–7. https://doi.org/10.1016/J.YDBIO.2017.03.001

Hirakawa, Y., Shinohara, H., Kondo, Y., Inoue, A., Nakanomyo, I., Ogawa, M., … Fukuda, H. (2008). Non-cell-autonomous control of vascular stem cell fate by a CLE peptide/receptor system. *Proceedings*

of the National Academy of Sciences of the United States of America, *105*(39), 15208–15213. https://doi.org/10.1073/pnas.0808444105

Hoen, P. A. C., Ariyurek, Y., Thygesen, H. H., Vreugdenhil, E., Vossen, R. H. A. M., de Menezes, R. X., … den Dunnen, J. T. (2008). Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Research*, *36*(21), e141–e141. https://doi.org/10.1093/nar/gkn705

Hong, J., Yang, L., Zhang, D., & Shi, J. (2016). Plant Metabolomics: An Indispensable System Biology Tool for Plant Science. *International Journal of Molecular Sciences*, *17*(6), 767. https://doi.org/10.3390/ijms17060767

Hrdlickova, R., Toloue, M., & Tian, B. (2017, January 1). RNA-Seq methods for transcriptome analysis. *Wiley Interdisciplinary Reviews: RNA*. John Wiley & Sons, Ltd. https://doi.org/10.1002/wrna.1364

Hu, B., Henry, A., Brown, K. M., & Lynch, J. P. (n.d.). Root cortical aerenchyma inhibits radial nutrient transport in maize (Zea mays). https://doi.org/10.1093/aob/mct259

Huang, C. F., Yamaji, N., Ono, K., & Ma, J. F. (2012). A leucine-rich repeat receptor-like kinase gene is involved in the specification of outer cell layers in rice roots. *Plant Journal*, *69*(4), 565–576. https://doi.org/10.1111/j.1365-313X.2011.04824.x

Huang, X., Kurata, N., Wei, X., Wang, Z.-X., Wang, A., Zhao, Q., … Han, B. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature*, *490*(7421), 497–501. https://doi.org/10.1038/nature11532

Hutchison, C. A., Chuang, R.-Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., … Venter, J. C. (2016). Design and synthesis of a minimal bacterial genome. *Science*, *351*(6280), aad6253–aad6253. https://doi.org/10.1126/science.aad6253

Hughesman, C. B., Lu, X. J. D., Liu, K. Y. P., Zhu, Y., Poh, C. F., & Haynes, C. (2016). A Robust Protocol for Using Multiplexed Droplet Digital PCR to Quantify Somatic Copy Number Alterations in Clinical Tissue Specimens. *PLOS ONE*, *11*(8), e0161274. https://doi.org/10.1371/journal.pone.0161274

Hsing, Y.-I., Chern, C.-G., Fan, M.-J., Lu, P.-C., Chen, K.-T., Lo, S.-F., … Yu, S.-M. (2007). A rice gene activation/knockout mutant resource for high throughput functional genomics. *Plant Molecular Biology*, *63*(3), 351–364. https://doi.org/10.1007/s11103-006-9093-z

Hsu, M.-N., Chang, Y.-H., Truong, V. A., Lai, P.-L., Nguyen, T. K. N., & Hu, Y.-C. (2019). CRISPR technologies for stem cell engineering and regenerative medicine. *Biotechnology Advances*, 107447. https://doi.org/10.1016/J.BIOTECHADV.2019.107447

Inan, G., Zhang, Q., Li, P., Wang, Z., Cao, Z., Zhang, H., … Zhu, J. K. (2004). Salt cress. A halophyte and cryophyte arabidopsis relative model system and its applicability to molecular genetic analyses of growth and development of extremophiles. *Plant Physiology*, *135*(3), 1718–1737. https://doi.org/10.1104/pp.104.041723

Izawa, T., & Shimamoto, K. (1996). Becoming a model plant: The importance of rice to plant science. *Trends in Plant Science*, *1*(3), 95–99. https://doi.org/10.1016/S1360-1385(96)80041-0

Jaitin, D. A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., … Amit, I. (2016). Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell*, *167*(7), 1883-1896.e15. https://doi.org/10.1016/J.CELL.2016.11.039

Jakočiūnas, T., Jensen, M. K., & Keasling, J. D. (2017). System-level perturbations of cell metabolism using CRISPR/Cas9. *Current Opinion in Biotechnology*. https://doi.org/10.1016/j.copbio.2017.03.014

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., & Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research*, *41*(20), e188–e188. https://doi.org/10.1093/nar/gkt780

Jiao, Y., Lori Tausta, S., Gandotra, N., Sun, N., Liu, T., Clay, N. K., … Nelson, T. (2009). A transcriptome atlas of rice cell types uncovers cellular, functional and developmental hierarchies. *Nature Genetics*, *41*(2), 258–263. https://doi.org/10.1038/ng.282

Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., & Gao, G. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*, *45*(D1), D1040–D1045. https://doi.org/10.1093/nar/gkw982

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Joshi, R., & Kumar, P. (2012, January 8). Lysigenous aerenchyma formation involves non-apoptotic programmed cell death in rice (Oryza sativa L.) roots. *Physiology and Molecular Biology of Plants*. Springer-Verlag. https://doi.org/10.1007/s12298-011-0093-3

Joung, J., Konermann, S., Gootenberg, J. S., Abudayyeh, O. O., Platt, R. J., Brigham, M. D., … Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature Protocols*, *12*(4), 828–863. https://doi.org/10.1038/nprot.2017.016

Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., … Covert, M. W. (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, *150*(2), 389–401. https://doi.org/10.1016/j.cell.2012.05.044

Kashofer, K., Viertler, C., Pichler, M., & Zatloukal, K. (2013). Quality Control of RNA Preservation and Extraction from Paraffin-Embedded Tissue: Implications for RT-PCR and Microarray Analysis. *PLoS ONE*, *8*(7), e70714. https://doi.org/10.1371/journal.pone.0070714

Kawai, M., Samarajeewa, P. K., Barrero, R. A., Nishiguchi, M., & Uchimiya, H. (1998). Cellular dissection of the degradation pattern of cortical cell death during aerenchyma formation of rice roots. *Planta*, *204*(3), 277–287. https://doi.org/10.1007/s004250050257

Kawahara, Y., Oono, Y., Wakimoto, H., Ogata, J., Kanamori, H., Sasaki, H., … Itoh, T. (2016). TENOR: Database for Comprehensive mRNA-Seq Experiments in Rice. *Plant and Cell Physiology*, *57*(1), e7–e7. https://doi.org/10.1093/pcp/pcv179

Kiss, M. M., Ortoleva-Donnelly, L., Beer, N. R., Warner, J., Bailey, C. G., Colston, B. W., … Leamon, J. H. (2008). High-Throughput Quantitative Polymerase Chain Reaction in Picoliter Droplets. *Analytical Chemistry*, *80*(23), 8975–8981. https://doi.org/10.1021/ac801276c

Kitano, H. (2002). Computational systems biology. *Nature*. Nature Publishing Group. https://doi.org/10.1038/nature01254

Kitano, H. (2002)b. Systems biology: A brief overview. *Science*. American Association for the Advancement of Science. https://doi.org/10.1126/science.1069492

Kludze, H. K., DeLaune, R. D., & Patrick, W. H. (1993). Aerenchyma Formation and Methane and Oxygen Exchange in Rice. *Soil Science Society of America Journal*, *57*(2), 386. https://doi.org/10.2136/sssaj1993.03615995005700020017x

Kohl, P., Crampin, E. J., Quinn, T. A., & Noble, D. (2010). Systems Biology: An Approach. *Clinical Pharmacology & Therapeutics*, *88*(1), 25–33. https://doi.org/10.1038/clpt.2010.92

Kolesnik, T., Szeverenyi, I., Bachmann, D., Kumar, C. S., Jiang, S., Ramamoorthy, R., … Ramachandran, S. (2004). Establishing an efficient Ac/Ds tagging system in rice: large-scale analysis of Ds flanking sequences. *The Plant Journal*, *37*(2), 301–314. https://doi.org/10.1046/j.1365-313X.2003.01948.x

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., … Zhang, F. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, *517*(7536), 583–588. https://doi.org/10.1038/nature14136

Kondo, M., Murty, M. V. R., & Aragones, D. V. (2000). Characteristics of root growth and water uptake from soil in upland rice and maize under water stress. *Soil Science and Plant Nutrition*, *46*(3), 721–732. https://doi.org/10.1080/00380768.2000.10409137

Kovach, M. J., Sweeney, M. T., & McCouch, S. R. (2007). New insights into the history of rice domestication. *Trends in Genetics*, *23*(11), 578–587. https://doi.org/10.1016/J.TIG.2007.08.012

pengies, T. (2018). Suberized transport barriers in Arabidopsis, barley and rice roots: From the model plant to crop species. *Journal of Plant Physiology*, *227*, 75–83. https://doi.org/10.1016/J.JPLPH.2018.02.002

Kulichikhin, K., Yamauchi, T., Watanabe, K., & Nakazono, M. (2014). Biochemical and molecular characterization of rice (*O ryza sativa* L.) roots forming a barrier to radial oxygen loss. *Plant, Cell & Environment*, *37*(10), n/a-n/a. https://doi.org/10.1111/pce.12294

Kumlehn, J., Pietralla, J., Hensel, G., Pacher, M., & Puchta, H. (2018). The CRISPR/Cas revolution continues: From efficient gene editing for crop breeding to plant synthetic biology. *Journal of Integrative Plant Biology*, *60*(12), 1127–1153. https://doi.org/10.1111/jipb.12734

Kurn, N., Chen, P., Heath, J. D., Kopf-Sill, A., Stephens, K. M., & Wang, S. (2005). Novel isothermal, linear nucleic acid amplification systems for highly multiplexed applications. *Clinical Chemistry*, *51*(10), 1973–1981. https://doi.org/10.1373/clinchem.2005.053694

Laborte, A. G., Gutierrez, M. A., Balanza, J. G., Saito, K., Zwart, S. J., Boschetti, M., … Nelson, A. (2017). RiceAtlas, a spatial database of global rice calendars and production. *Scientific Data*, *4*(1), 170074. https://doi.org/10.1038/sdata.2017.74

le Sage, C., Lawo, S., Panicker, P., Scales, T. M. E., Rahman, S. A., Little, A. S., … Cross, B. C. S. (2017). Dual direction CRISPR transcriptional regulation screening uncovers gene networks driving drug resistance. *Scientific Reports*, *7*(1), 17693. https://doi.org/10.1038/s41598-017-18172-6

Levesque, M. P., Vernoux, T., Busch, W., Cui, H., Wang, J. Y., Blilou, I., … Benfey, P. N. (2006). Whole-Genome Analysis of the SHORT-ROOT Developmental Pathway in Arabidopsis. *PLoS Biology*, *4*(5), e143. https://doi.org/10.1371/journal.pbio.0040143

Levin, J. Z., Yassour, M., Adiconis, X., Nusbaum, C., Thompson, D. A., Friedman, N., … Regev, A. (2010). Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nature Methods*, *7*(9), 709–715. https://doi.org/10.1038/nmeth.1491

Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R. J., Franklin, L. D., … Stacey, G. (2010). An integrated transcriptome atlas of the crop model Glycine max, and its use in comparative analyses in plants. *The Plant Journal*, *63*(1), no-no. https://doi.org/10.1111/j.1365-313X.2010.04222.x

Li, J.-Y., Wang, J., & Zeigler, R. S. (2014). The 3,000 rice genomes project: new opportunities and challenges for future rice research. *GigaScience*, *3*(1), 8. https://doi.org/10.1186/2047-217X-3-8

Li, P., Yu, Q., Gu, X., Xu, C., Qi, S., Wang, H., … Wu, S. (2018). Construction of a Functional Casparian Strip in Non-endodermal Lineages Is Orchestrated by Two Parallel Signaling Systems in Arabidopsis thaliana. *Current Biology*, 1–10. https://doi.org/10.1016/j.cub.2018.07.028

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., & Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature Biotechnology*, *30*(5), 390–392. https://doi.org/10.1038/nbt.2199

LI, Y. L., FAN, X. R., & SHEN, Q. R. (2008). The relationship between rhizosphere nitrification and nitrogen-use efficiency in rice plants. *Plant, Cell & Environment*, *0*(0), 071018035446001-??? https://doi.org/10.1111/j.1365-3040.2007.01737.x

Li, Y., & Chen, L. (2014). Big Biological Data: Challenges and Opportunities. *Genomics, Proteomics & Bioinformatics*, *12*(5), 187. https://doi.org/10.1016/J.GPB.2014.10.001

Li, Z., Mu, P., Li, C., Zhang, H., Li, Z., Gao, Y., & Wang, X. (2005). QTL mapping of root traits in a doubled haploid population from a cross between upland and lowland japonica rice in three environments. *Theoretical and Applied Genetics*, *110*(7), 1244–1252. https://doi.org/10.1007/s00122-005-1958-z

Li, Z., & Rutger, J. N. (2000). Geographic distribution and multilocus organization of isozyme variation of rice (Oryza sativa L.). *Theoretical and Applied Genetics*, *101*(3), 379–387. https://doi.org/10.1007/s001220051494

Li, Z., Xiong, X., & Li, J.-F. (2019). The working dead: repurposing inactive CRISPR-associated nucleases as programmable transcriptional regulators in plants. *ABIOTECH*, 1–9. https://doi.org/10.1007/s42994-019-00003-z

Liu, R., Bassalo, M. C., Zeitoun, R. I., & Gill, R. T. (2015). Genome scale engineering techniques for metabolic engineering. *Metabolic Engineering*, *32*, 143–154. https://doi.org/10.1016/j.ymben.2015.09.013

Liu, W., & Stewart, C. N. (2015). Plant synthetic biology. *Trends in Plant Science*, *20*(5), 309–317. https://doi.org/10.1016/J.TPLANTS.2015.02.004

Lodrini, M., Sprüssel, A., Astrahantseff, K., Tiburtius, D., Konschak, R., Lode, H. N., … Deubzer, H. E. (2017). Using droplet digital PCR to analyze MYCN and ALK copy number in plasma from patients with neuroblastoma. Oncotarget, 8(49), 85234–85251. https://doi.org/10.18632/oncotarget.19076 Loreti, E., van Veen, H., & Perata, P. (2016). Plant responses to flooding stress. *Current Opinion in Plant Biology*, *33*, 64–71. https://doi.org/10.1016/J.PBI.2016.06.005

Lowary, P. T., & Uhlenbeck, O. C. (1987). An RNA mutation that increases the affinity of an RNA–protein interaction. *Nucleic Acids Research*, *15*(24), 10483–10493. https://doi.org/10.1093/nar/15.24.10483

Lowder, L. G., Zhang, D., Baltes, N. J., Paul, J. W., Tang, X., Zheng, X., … Qi, Y. (2015). A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiology*, *169*(2), 971–985. https://doi.org/10.1104/pp.15.00636

Lu, T., Zhu, C., Lu, G., Guo, Y., Zhou, Y., Zhang, Z., … Han, B. (2012). Strand-specific RNA-seq reveals widespread occurrence of novel cis-natural antisense transcripts in rice. *BMC Genomics*, *13*(1), 721. https://doi.org/10.1186/1471-2164-13-721

Ludwig, Y., & Hochholdinger, F. (2014). Laser Microdissection of Plant Cells (pp. 249–258). Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-62703-643-6_21

Lux, A., Luxova, M., Abe, J., & Morita, S. (n.d.). Root cortex: structural and functional variability and responses to environmental stress. Retrieved from https://www.jstage.jst.go.jp/article/rootres1992/13/3/13_3_117/_pdf

Lynch, J. P., & Brown, K. M. (2001). Topsoil foraging – an architectural adaptation of plants to low phosphorus availability. *Plant and Soil*, *237*(2), 225–237. https://doi.org/10.1023/A:1013324727040

Lynch, J. P. (2013). Steep, cheap and deep: An ideotype to optimize water and N acquisition by maize root systems. *Annals of Botany*. https://doi.org/10.1093/aob/mcs293

Lynch, J. P. (2015). Root phenes that reduce the metabolic costs of soil exploration: opportunities for 21st century agriculture. *Plant, Cell & Environment*, *38*(9), 1775–1784. https://doi.org/10.1111/pce.12451

Ma, H., Tu, L.-C., Naseri, A., Chung, Y.-C., Grunwald, D., Zhang, S., & Pederson, T. (2018). CRISPR-Sirius: RNA scaffolds for signal amplification in genome imaging. *Nature Methods*, *15*(11), 928–931. https://doi.org/10.1038/s41592-018-0174-0

Ma, H., Tu, L. C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., & Pederson, T. (2016). Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nature Biotechnology*, *34*(5), 528–530. https://doi.org/10.1038/nbt.3526

Mai, C. D., Phung, N. T., To, H. T., Gonin, M., Hoang, G. T., Nguyen, K. L., … Gantet, P. (2014). Genes controlling root development in rice. *Rice*, *7*(1), 30. https://doi.org/10.1186/s12284-014-0030-5

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., … Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*, *31*(9), 833–838. https://doi.org/10.1038/nbt.2675

Mao, H., Sun, S., Yao, J., Wang, C., Yu, S., Xu, C., … Zhang, Q. (2010). Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proceedings of the National Academy of Sciences*, *107*(45), 19579–19584. https://doi.org/10.1073/PNAS.1014419107

Masters, T. L., Hilker, C. A., Jeraldo, P. R., Bhagwate, A. V., Greenwood-Quaintance, K. E., Eckloff, B. W., … Patel, R. (2018). Comparative evaluation of cDNA library construction approaches for RNA-Seq analysis from low RNA-content human specimens. *Journal of Microbiological Methods*, *154*, 55–62. https://doi.org/10.1016/J.MIMET.2018.10.008

Medici, A., Marshall-Colon, A., Ronzier, E., Szponarski, W., Wang, R., Gojon, A., … Krouk, G. (2015). AtNIGT1/HRS1 integrates nitrate and phosphate signals at the Arabidopsis root tip. *Nature Communications*, *6*(1), 6274. https://doi.org/10.1038/ncomms7274

Meng, X., Yu, H., Zhang, Y., Zhuang, F., Song, X., Gao, S., … Li, J. (2017). Construction of a Genome-Wide Mutant Library in Rice Using CRISPR/Cas9. *Molecular Plant*, *10*(9), 1238–1241. https://doi.org/10.1016/j.molp.2017.06.006

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., … Qu, L.-J. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Research*, *23*(10), 1233–1236. https://doi.org/10.1038/cr.2013.123

Miller, G., Suzuki, N., Ciftci-Yilmaz, S., & Mittler, R. (2010). Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant, Cell and Environment*, *33*(4), 453–467. https://doi.org/10.1111/j.1365-3040.2009.02041.x

Mills D., J., Kawahara, Y., & Janitz, M. (2013). Strand-Specific RNA-Seq Provides Greater Resolution of Transcriptome Profiling. *Current Genomics*, *14*(3), 173–181. https://doi.org/10.2174/1389202911314030003

Mitsunobu, H., Teramoto, J., Nishida, K., & Kondo, A. (2017). Beyond Native Cas9: Manipulating Genomic Information and Function. *Trends in Biotechnology*, *35*(10), 983–996. https://doi.org/10.1016/J.TIBTECH.2017.06.004

Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K., … Hirochika, H. (2003). Target Site Specificity of the Tos17 Retrotransposon Shows a Preference for Insertion within Genes and against Insertion in Retrotransposon-Rich Regions of the Genome. *The Plant Cell*, *15*, 1771–1780. https://doi.org/10.1105/tpc.012559

Miyashima, S., Koi, S., Hashimoto, T., & Nakajima, K. (2011). Non-cell-autonomous microRNA 165 acts in a dose-dependent manner to regulate multiple differentiation status in the Arabidopsis root. *Development*, *138*(11), 2303–2313. https://doi.org/10.1242/dev.060491

Möller, B. K., Ten Hove, C. A., Xiang, D., Williams, N., González López, L., Yoshida, S., … Weijers, D. (n.d.). Auxin response cell-autonomously controls ground tissue initiation in the early Arabidopsis embryo. https://doi.org/10.1073/pnas.1616493114

Morbitzer, R., Römer, P., Boch, J., & Lahaye, T. (2010). Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21617–21622. https://doi.org/10.1073/pnas.1013133107

Mortimer, J. C. (2019). Plant synthetic biology could drive a revolution in biofuels and medicine. *Experimental Biology and Medicine (Maywood, N.J.)*, *244*(4), 323–331. https://doi.org/10.1177/1535370218793890

Muhitch, M. J., & Shatters, R. G. (1998). Regulation of the maize ubiquitin (Ubi-1) promoter in developing maize (Zea mays L.) seeds examined using transient gene expression in kernels grown in vitro. *Plant Cell Reports*, *17*(6–7), 476–481. https://doi.org/10.1007/s002990050428

Mu, W., Zhang, Y., Xue, X., Liu, L., Wei, X., & Wang, H. (2019). 5′ capped and 3′ polyA-tailed sgRNAs enhance the efficiency of CRISPR-Cas9 system. *Protein & Cell*, *10*(3), 223–228. https://doi.org/10.1007/s13238-018-0552-5

Naz, N., Hameed, M., Nawaz, T., Batool, R., Ashraf, M., Ahmad, F., & Ruby, T. (2013). Structural adaptations in the desert halophyte Aeluropus lagopoides (Linn.) Trin. ex Thw. under high salinity. *Journal of Biological Research (Greece)*, *19*, 150–164.

Najm, F. J., Strand, C., Donovan, K. F., Hegde, M., Sanson, K. R., Vaimberg, E. W., … Doench, J. G. (2018). Orthologous CRISPR–Cas9 enzymes for combinatorial genetic screens. *Nature Biotechnology*, *36*(2), 179–189. https://doi.org/10.1038/nbt.4048

Nascente, A. S., Crusciol, C. A. C., & Cobucci, T. (2013). The no-tillage system and cover crops—Alternatives to increase upland rice yields. *European Journal of Agronomy*, *45*, 124–131. https://doi.org/10.1016/J.EJA.2012.09.004

Nelson, T., Tausta, S. L., Gandotra, N., & Liu, T. (2006). LASER MICRODISSECTION OF PLANT TISSUE: What You See Is What You Get. *Annual Review of Plant Biology*, *57*(1), 181–201. https://doi.org/10.1146/annurev.arplant.56.032604.144138

Nichterwitz, S., Chen, G., Aguila Benitez, J., Yilmaz, M., Storvall, H., Cao, M., … Hedlund, E. (2016). Laser capture microscopy coupled with Smart-seq2 for precise spatial transcriptomic profiling. *Nature Communications*, *7*(1), 12139. https://doi.org/10.1038/ncomms12139

Nicolle, R., Radvanyi, F., & Elati, M. (2015). CoRegNet: reconstruction and integrated analysis of co-regulatory networks. *Bioinformatics (Oxford, England)*, *31*(18), 3066–3068. https://doi.org/10.1093/bioinformatics/btv305

Nielsen, J., & Keasling, J. D. (2016). Engineering Cellular Metabolism. *Cell*, *164*(6), 1185–1197. https://doi.org/10.1016/J.CELL.2016.02.004

Nishiuchi, S., Yamauchi, T., Takahashi, H., Kotula, L., & Nakazono, M. (2012). Mechanisms for coping with submergence and waterlogging in rice. *Rice*, *5*(1), 2. https://doi.org/10.1186/1939-8433-5-2

Nowak, C. M., Lawson, S., Zerez, M., & Bleris, L. (2016). Guide RNA engineering for versatile Cas9 functionality. *Nucleic Acids Research*, *44*(20), 9555–9564. https://doi.org/10.1093/nar/gkw908

Nygaard, V., & Hovig, E. (2006). Options available for profiling small samples: a review of sample amplification technology when combined with microarray profiling. *Nucleic Acids Research*, *34*(3), 996–1014. https://doi.org/10.1093/nar/gkj499

O'Neil, D., Glowatz, H., & Schlumpberger, M. (2013). Ribosomal RNA Depletion for Efficient Use of RNA-Seq Capacity. In *Current Protocols in Molecular Biology* (Vol. 103, pp. 4.19.1-4.19.8). Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/0471142727.mb0419s103

Ozsolak, F., & Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics*, *12*(2), 87–98. https://doi.org/10.1038/nrg2934

Para, A., Li, Y., Marshall-Colón, A., Varala, K., Francoeur, N. J., Moran, T. M., … Coruzzi, G. M. (2014). Hit-and-run transcriptional control by bZIP1 mediates rapid nutrient signaling in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(28), 10371–10376. https://doi.org/10.1073/pnas.1404657111

Pauluzzi, G., Divol, F., Puig, J., Guiderdoni, E., Dievart, A., & Périn, C. (2012). Surfing along the root ground tissue gene network. *Developmental Biology*, *365*(1), 14–22. https://doi.org/10.1016/J.YDBIO.2012.02.007

Pedersen, O., Perata, P., & Voesenek, L. A. C. J. (2017). Flooding and low oxygen responses in plants. *Functional Plant Biology*, *44*(9), iii. https://doi.org/10.1071/FPv44n9_FO

Pendrak, M. L., & Roberts, D. D. (2011). Ribosomal RNA processing in Candida albicans. *RNA (New York, N.Y.)*, *17*(12), 2235–2248. https://doi.org/10.1261/rna.028050.111

Peng, Y.-H., Zhu, Y., Mao, Y., Wang, S., Su, W., & Tang, Z. (2004). Alkali grass resists salt stress through high [K+] and an endodermis barrier to Na+. *Journal of Experimental Botany*, *55*(398), 939–949. https://doi.org/10.1093/jxb/erh071

Péret, B., Larrieu, A., & Bennett, M. J. (2009). Lateral root emergence: A difficult birth. *Journal of Experimental Botany*. https://doi.org/10.1093/jxb/erp232

Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., … Mahfouz, M. M. (2015). RNA-guided transcriptional regulation *in planta* via synthetic dCas9-based transcription factors. *Plant Biotechnology Journal*, *13*(4), 578–589. https://doi.org/10.1111/pbi.12284

Polstein, L. R., Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Bledsoe, P., Song, L., … Gersbach, C. A. (2015). Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Research*, *25*(8), 1158–1169. https://doi.org/10.1101/gr.179044.114

Pouvreau, B., Vanhercke, T., & Singh, S. (2018). From plant metabolic engineering to plant synthetic biology: The evolution of the design/build/test/learn cycle. *Plant Science*, *273*, 3–12. https://doi.org/10.1016/j.plantsci.2018.03.035

Prakadan, S. M., Shalek, A. K., & Weitz, D. A. (2017). Scaling by shrinking: empowering single-cell "omics" with microfluidic devices. *Nature Reviews Genetics*, *18*(6), 345–361. https://doi.org/10.1038/nrg.2017.15

Promkhambut, A., Polthanee, A., Akkasaeng, C., & Younger, A. (2011). *Growth, yield and aerenchyma formation of sweet and multipurpose sorghum (Sorghum bicolor L. Moench) as affected by flooding at different growth stages*. *AJCS* (Vol. 5). Retrieved from http://www.cropj.com/prumkkhambut_5_8_2011_954_965.pdf

Pucciariello, C., & Perata, P. (2017). New insights into reactive oxygen species and nitric oxide signalling under low oxygen in plants. *Plant, Cell & Environment*, *40*(4), 473–482. https://doi.org/10.1111/pce.12715

Puchta, H. (2017). Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Current Opinion in Plant Biology*, *36*, 1–8. https://doi.org/10.1016/J.PBI.2016.11.011

Pujol, V., & Wissuwa, M. (2018). Contrasting development of lysigenous aerenchyma in two rice genotypes under phosphorus deficiency. *BMC Research Notes*, *11*(1), 60. https://doi.org/10.1186/s13104-018-3179-y

Puskás, L. G., Zvara, Á., Hackler, L., & Van Hummelen, P. (2002). RNA amplification results in reproducible microarray data with slight ratio bias. *BioTechniques*, *32*(6), 1330–1340. https://doi.org/10.2144/02326mt04

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, *152*(5), 1173–1183. https://doi.org/10.1016/J.CELL.2013.02.022

Qin, R., Liao, S., Li, J., Li, H., Liu, X., Yang, J., & Wei, P. (2019). Increasing fidelity and efficiency by modifying cytidine base-editing systems in rice. *The Crop Journal*. https://doi.org/10.1016/J.CJ.2019.04.007

Rebouillat, J., Dievart, A., Verdeil, J. L., Escoute, J., Giese, G., Breitler, J. C., … Périn, C. (2009). Molecular Genetics of Rice Root Development. *Rice*, *2*(1), 15–34. https://doi.org/10.1007/s12284-008-9016-5

Rački, N., Dreo, T., Gutierrez-Aguirre, I., Blejec, A., & Ravnikar, M. (2014). Reverse transcriptase droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water samples. *Plant Methods*, *10*(1), 42. https://doi.org/10.1186/s13007-014-0042-6

Rački, N., Morisset, D., Gutierrez-Aguirre, I., & Ravnikar, M. (2014). One-step RT-droplet digital PCR: a breakthrough in the quantification of waterborne RNA viruses. *Analytical and Bioanalytical Chemistry*, *406*(3), 661–667. https://doi.org/10.1007/s00216-013-7476-y

Rajhi, I., Yamauchi, T., Takahashi, H., Nishiuchi, S., Shiono, K., Watanabe, R., … Nakazono, M. (2011). Identification of genes expressed in maize root cortical cells during lysigenous aerenchyma formation using laser microdissection and microarray analyses. *New Phytologist*, *190*(2), 351–368. https://doi.org/10.1111/j.1469-8137.2010.03535.x

Ranathunge, K., Steudle, E., & Lafitte, R. (2005). A new precipitation technique provides evidence for the permeability of Casparian bands to ions in young roots of corn (Zea mays L.) and rice (Oryza sativa L.). *Plant, Cell and Environment*, *28*(11), 1450–1462. https://doi.org/10.1111/j.1365-3040.2005.01391.x

Rebouillat, J., Dievart, A., Verdeil, J. L., Escoute, J., Giese, G., Breitler, J. C., … Périn, C. (2009). Molecular Genetics of Rice Root Development. *Rice*, *2*(1), 15–34. https://doi.org/10.1007/s12284-008-9016-5

Reubinoff B., Pera M., Fong C.Y, Trounson A (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nature America*. Retrieved from https://pdfs.semanticscholar.org/5186/44b2db79fd080d436f165e0b031a5b03bdc3.pdf

Reynoso, M. A., Pauluzzi, G. C., Kajala, K., Cabanlit, S., Velasco, J., Bazin, J., … Bailey-Serres, J. (2018). Nuclear transcriptomes at high resolution using retooled INTACT. *Plant Physiology*, *176*(1), 270–281. https://doi.org/10.1104/pp.17.00688

Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A., & Kim, D. (2015). Methods of integrating data to uncover genotype–phenotype interactions. *Nature Reviews Genetics*, *16*(2), 85–97. https://doi.org/10.1038/nrg3868

Roux, B., Rodde, N., Jardinaud, M.-F., Timmers, T., Sauviac, L., Cottret, L., … Gamas, P. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *The Plant Journal*, *77*(6), 817–837. https://doi.org/10.1111/tpj.12442

Rubin, A. J., Parker, K. R., Satpathy, A. T., Qi, Y., Wu, B., Ong, A. J., … Khavari, P. A. (2019). Coupled Single-Cell CRISPR Screening and Epigenomic Profiling Reveals Causal Gene Regulatory Networks. *Cell*, *176*(1–2), 361-376.e17. https://doi.org/10.1016/J.CELL.2018.11.022

Ryu, K. H., Huang, L., Kang, H. M., & Schiefelbein, J. (2019). Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiology*, pp.01482.2018. https://doi.org/10.1104/pp.18.01482

Saengwilai, P., Nord, E. A., Chimungu, J. G., Brown, K. M., & Lynch, J. P. (2014). Root cortical aerenchyma enhances nitrogen acquisition from low-nitrogen soils in maize. *Plant Physiology*, *166*(2), 726–735. https://doi.org/10.1104/pp.114.241711

Sakai, K., Taconnat, L., Borrega, N., Yansouni, J., Brunaud, V., Paysant-Le Roux, C., … Dubreucq, B. (2018). Combining laser-assisted microdissection (LAM) and RNA-seq allows to perform a comprehensive transcriptomic analysis of epidermal cells of Arabidopsis embryo. *Plant Methods*, *14*(1), 10. https://doi.org/10.1186/s13007-018-0275-x

Sallaud, C., Gay, C., Larmande, P., Bès, M., Piffanelli, P., Piégu, B., … Guiderdoni, E. (2004). High throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. *The Plant Journal*, *39*(3), 450–464. https://doi.org/10.1111/j.1365-313X.2004.02145.x

Sakai, K., Taconnat, L., Borrega, N., Yansouni, J., Brunaud, V., Paysant-Le Roux, C., … Dubreucq, B. (2018). Combining laser-assisted microdissection (LAM) and RNA-seq allows to perform a comprehensive transcriptomic analysis of epidermal cells of Arabidopsis embryo. Plant Methods, 14(1), 10. https://doi.org/10.1186/s13007-018-0275-x

Sasaki, T. (2017). Achieving sustainable cultivation of rice. Volume 1: breeding for higher yield and quality. *Achieving Sustainable Cultivation of Rice. Volume 1: Breeding for Higher Yield and Quality.* Retrieved from https://www.cabdirect.org/cabdirect/abstract/20173221667

Sato, Y., Takehisa, H., Kamatsuki, K., Minami, H., Namiki, N., Ikawa, H., … Nagamura, Y. (2013). RiceXPro version 3.0: expanding the informatics resource for rice transcriptome. *Nucleic Acids Research*, *41*(Database issue), D1206-13. https://doi.org/10.1093/nar/gks1125

Sato, Y., Antonio, B. A., Namiki, N., Takehisa, H., Minami, H., Kamatsuki, K., … Nagamura, Y. (2011). RiceXPro: a platform for monitoring gene expression in japonica rice grown under natural field conditions. *Nucleic Acids Research*, *39*(Database issue), D1141-8. https://doi.org/10.1093/nar/gkq1085

Schimel, J. (2000). Rice, microbes and methane. *Nature*, *403*(6768), 375–377. https://doi.org/10.1038/35000325

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., … Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology*, *7*(1), 3. https://doi.org/10.1186/1471-2199-7-3

Seago, J. L., Marsh, L. C., Stevens, K. J., Soukup, A., Votrubová, O., Enstone, D. E., & ENSTONE, D. E. (2005). A re-examination of the root cortex in wetland flowering plants with respect to aerenchyma. *Annals of Botany*, *96*(4), 565–579. https://doi.org/10.1093/aob/mci211

Shalem, O., Sanjana, N. E., & Zhang, F. (2015). High-throughput functional genomics using CRISPR–Cas9. *Nature Reviews Genetics*, *16*(5), 299–311. https://doi.org/10.1038/nrg3899

Shanker, S., Paulson, A., Edenberg, H. J., Peak, A., Perera, A., Alekseyev, Y. O., … Nicolet, C. M. (2015). Evaluation of commercially available RNA amplification kits for RNA sequencing using very low input amounts of total RNA. *Journal of Biomolecular Techniques : JBT*, *26*(1), 4–18. https://doi.org/10.7171/jbt.15-2601-001

Shao, S., Zhang, W., Hu, H., Xue, B., Qin, J., Sun, C., … Sun, Y. (2016). Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system. *Nucleic Acids Research*, *44*(9), e86–e86. https://doi.org/10.1093/nar/gkw066

Shen, J., Li, C., Mi, G., Li, L., Yuan, L., Jiang, R., & Zhang, F. (2013). Maximizing root/rhizosphere efficiency to improve crop productivity and nutrient use efficiency in intensive agriculture of China. *Journal of Experimental Botany*, *64*(5), 1181–1192. https://doi.org/10.1093/jxb/ers342.

Shen, J. P., Zhao, D., Sasik, R., Luebeck, J., Birmingham, A., Bojorquez-Gomez, A., … Mali, P. (2017). Combinatorial CRISPR–Cas9 screens for de novo mapping of genetic interactions. *Nature Methods*, *14*(6), 573–576. https://doi.org/10.1038/nmeth.4225

Shih, P. M., Vuu, K., Mansoori, N., Ayad, L., Louie, K. B., Bowen, B. P., … Loqué, D. (2016). A robust gene-stacking method utilizing yeast assembly for plant synthetic biology. *Nature Communications*, *7*(May), 13215. https://doi.org/10.1038/ncomms13215

Shimamoto, K., Terada, R., Izawa, T., & Fujimoto, H. (1989). Fertile transgenic rice plants regenerated from transformed protoplasts. *Nature*, *338*(6212), 274–276. https://doi.org/10.1038/338274a0

Shiono, K., Yamauchi, T., Yamazaki, S., Mohanty, B., Malik, A. I., Nagamura, Y., … Nakazono, M. (2014). Microarray analysis of laser-microdissected tissues indicates the biosynthesis of suberin in the outer part of roots during formation of a barrier to radial oxygen loss in rice (Oryza sativa). *Journal of Experimental Botany*, *65*(17), 4795–4806. https://doi.org/10.1093/jxb/eru235

Shiono, K., & Yamada, S. (2014). Waterlogging tolerance and capacity for oxygen transport in Brachypodium distachyon (Bd21). *Plant Root*, *8*, 5–12. https://doi.org/10.3117/plantroot.8.5

Shiono, K., Ogawa, S., Yamazaki, S., Isoda, H., Fujimura, T., Nakazono, M., & Colmer, T. D. (2011). Contrasting dynamics of radial O2-loss barrier induction and aerenchyma formation in rice roots of two lengths. *Annals of Botany*, *107*(1), 89–99. https://doi.org/10.1093/aob/mcq221

Shiroguchi, K., Jia, T. Z., Sims, P. A., & Xie, X. S. (2012). Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(4), 1347–1352. https://doi.org/10.1073/pnas.1118018109

Smith, S., & de Smet, I. (2012). Root system architecture: Insights from Arabidopsis and cereal crops. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1595), 1441–1452. https://doi.org/10.1098/rstb.2011.0234

Song, X.-J., Huang, W., Shi, M., Zhu, M.-Z., & Lin, H.-X. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genetics*, *39*(5), 623–630. https://doi.org/10.1038/ng2014

Srinivas, G., Satish, K., Madhusudhana, R., Nagaraja Reddy, R., Murali Mohan, S., & Seetharama, N. (2009). Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellite markers in sorghum. *Theoretical and Applied Genetics*, *118*(8), 1439–1454. https://doi.org/10.1007/s00122-009-0993-6

Steffens, B., Geske, T., & Sauter, M. (2011). Aerenchyma formation in the rice stem and its promotion by H 2 O 2. *New Phytologist*, *190*, 369–378. https://doi.org/10.1111/j.1469-8137.2010.03496.x

Striker, G. G., Insausti, P., Grimoldi, A. A., & Vega, A. S. (2007). Trade-off between root porosity and mechanical strength in species with different types of aerenchyma. *Plant, Cell and Environment*, *30*(5), 580–589. https://doi.org/10.1111/j.1365-3040.2007.01639.x

Suralta, R. R., & Yamauchi, A. (2008). Root growth, aerenchyma development, and oxygen transport in rice genotypes subjected to drought and waterlogging. *Environmental and Experimental Botany*, *64*(1), 75–82. https://doi.org/10.1016/j.envexpbot.2008.01.004

Tabata, T., Takei, Y., Willemsen, V., Linstead, P., Poethig, S., Roberts, K., & Scheres, B. (2004). Morphogens, their identification and regulation. *Development*, *131*(4), 703–712. https://doi.org/10.1242/dev.01043

Tak, Y. E., Kleinstiver, B. P., Nuñez, J. K., Hsu, J. Y., Horng, J. E., Gong, J., … Joung, J. K. (2017). Inducible and multiplex gene regulation using CRISPR–Cpf1-based transcription factors. *Nature Methods*, *14*(12), 1163–1166. https://doi.org/10.1038/nmeth.4483

Takahashi, H., Kamakura, H., Sato, Y., Shiono, K., Abiko, T., Tsutsumi, N., Nakazono, M. (2010). A method for obtaining high quality RNA from paraffin sections of plant tissues by laser microdissection. *Journal of Plant Research*, *123*(6), 807–813. https://doi.org/10.1007/s10265-010-0319-4

Takahashi, H., Yamauchi, T., Colmer, T. D., & Nakazono, M. (2014). Aerenchyma Formation in Plants (pp. 247–265). Springer, Vienna. https://doi.org/10.1007/978-3-7091-1254-0_13

Takehisa, H., Sato, Y., Igarashi, M., Abiko, T., Antonio, B. A., Kamatsuki, K., … Nagamura, Y. (2012). Genome-wide transcriptome dissection of the rice root system: implications for developmental and physiological functions. *The Plant Journal*, *69*(1), 126–140. https://doi.org/10.1111/j.1365-313X.2011.04777.x

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., & Vale, R. D. (2014). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*, *159*(3), 635–646. https://doi.org/10.1016/J.CELL.2014.09.039

Tari, I., Laskay, G., Takács, Z., & Poór, P. (2013). Response of Sorghum to Abiotic Stresses: A Review. *Journal of Agronomy and Crop Science*, *199*(4), 264–274. https://doi.org/10.1111/jac.12017

Taylor, S. C., Laperriere, G., & Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Scientific Reports*, *7*(1), 2409. https://doi.org/10.1038/s41598-017-02217-x

The 3,000 rice genomes project. (2014). *GigaScience*, *3*(1), 7. https://doi.org/10.1186/2047-217X-3-7

Thiele, I., Swainston, N., Fleming, R. M. T., Hoppe, A., Sahoo, S., Aurich, M. K., … Palsson, B. Ø. (2013). A community-driven global reconstruction of human metabolism. *Nature Biotechnology*, *31*(5), 419–425. https://doi.org/10.1038/nbt.2488

Trébulle, P., Nicaud, J.-M., Leplat, C., & Elati, M. (2017). Inference and interrogation of a coregulatory network in the context of lipid accumulation in Yarrowia lipolytica. *Npj Systems Biology and Applications*, *3*(1), 21. https://doi.org/10.1038/s41540-017-0024-1

Truong, V. A., Hsu, M.-N., Kieu Nguyen, N. T., Lin, M.-W., Shen, C.-C., Lin, C.-Y., & Hu, Y.-C. (2019). CRISPRai for simultaneous gene activation and inhibition to promote stem cell chondrogenesis and calvarial bone regeneration. *Nucleic Acids Research*, *47*(13), e74–e74. https://doi.org/10.1093/nar/gkz267

Uga, Y., Hanzawa, E., Nagai, S., Sasaki, K., Yano, M., & Sato, T. (2012). Identification of qSOR1, a major rice QTL involved in soil-surface rooting in paddy fields. *Theoretical and Applied Genetics*, *124*(1), 75–86. https://doi.org/10.1007/s00122-011-1688-3

Uga, Y., Okuno, K., & Yano, M. (2011). Dro1, a major QTL involved in deep rooting of rice under upland field conditions. *Journal of Experimental Botany*, *62*(8), 2485–2494. https://doi.org/10.1093/jxb/erq429

Uga, Y., Sugimoto, K., Ogawa, S., Rane, J., Ishitani, M., Hara, N., … Yano, M. (2013). Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. *Nature Genetics*, *45*(9), 1097–1102. https://doi.org/10.1038/ng.2725

Vanhercke, T., Divi, U. K., El Tahchy, A., Liu, Q., Mitchell, M., Taylor, M. C., … Petrie, J. R. (2017). Step changes in leaf oil accumulation via iterative metabolic engineering. *Metabolic Engineering*, *39*, 237–246. https://doi.org/10.1016/J.YMBEN.2016.12.007

Vergara, B. S. (1991). Rice Plant Growth and Development. In *Rice* (pp. 13–22). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4899-3754-4_2

Vermeulen, J., De Preter, K., Lefever, S., Nuytens, J., De Vloed, F., Derveaux, S., … Vandesompele, J. (2011). Measurable impact of RNA quality on gene expression results from quantitative PCR. *Nucleic Acids Research*, *39*(9), e63–e63. https://doi.org/10.1093/nar/gkr065

Voesenek, L. A. C. J., Colmer, T. D., Pierik, R., Millenaar, F. F., & Peeters, A. J. M. (2006). How plants cope with complete submergence. *New Phytologist*, *170*(2), 213–226. https://doi.org/10.1111/j.1469-8137.2006.01692.x

Voesenek, L. A. C. J., & Sasidharan, R. (2013). Ethylene - and oxygen signalling - drive plant survival during flooding. *Plant Biology*, *15*(3), 426–435. https://doi.org/10.1111/plb.12014

Vogelstein, B., & Kinzler, K. W. (1999). Digital PCR. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(16), 9236–9241. https://doi.org/10.1073/pnas.96.16.9236

van Dijk, E. L., Jaszczyszyn, Y., & Thermes, C. (2014). Library preparation methods for next-generation sequencing: Tone down the bias. *Experimental Cell Research*, *322*(1), 12–20. https://doi.org/10.1016/J.YEXCR.2014.01.008

LUDWIG VON BERTALANFFY. (1950). *AN OUTLINE OF GENERAL SYSTEM THEORY i Parallel Evolution in Science*. Retrieved from https://academic.oup.com/bjps/article-abstract/I/2/134/1514614

Wang, C., Gong, B., Bushel, P. R., Thierry-Mieg, J., Thierry-Mieg, D., Xu, J., … Tong, W. (2014). The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nature Biotechnology*, *32*(9), 926–932. https://doi.org/10.1038/nbt.3001

Wang, C., Lai, D. Y. F., Sardans, J., Wang, W., Zeng, C., & Peñuelas, J. (2017). Factors Related with CH4 and N2O Emissions from a Paddy Field: Clues for Management implications. *PloS One*, *12*(1), e0169254. https://doi.org/10.1371/journal.pone.0169254

Wang, Y. X. X. R., & Huang, H. (2014). Review on statistical methods for gene network reconstruction using expression data. *Journal of Theoretical Biology*, *362*, 53–61. https://doi.org/10.1016/j.jtbi.2014.03.040

Wang, Y., Loake, G. J., & Chu, C. (2013). Cross-talk of nitric oxide and reactive oxygen species in plant programed cell death. *Frontiers in Plant Science*, *4*, 314. https://doi.org/10.3389/fpls.2013.00314

Watanabe, K., Takahashi, H., Sato, S., Nishiuchi, S., Omori, F., Malik, A. I., … Nakazono, M. (2017). A major locus involved in the formation of the radial oxygen loss barrier in adventitious roots of teosinte *Zea nicaraguensis* is located on the short-arm of chromosome 3. *Plant, Cell & Environment*, *40*(2), 304–316. https://doi.org/10.1111/pce.12849

Wing, R. A., Kim, H., Goicoechea, J. L., Yu, Y., Kudrna, D., Zuccolo, A., … Jackson, S. (2007). The Oryza Map Alignment Project (OMAP): A New Resource for Comparative Genome Studies within Oryza. In *Rice Functional Genomics* (pp. 395–409). New York, NY: Springer New York. https://doi.org/10.1007/0-387-48914-2_15

Wolter, F., & Puchta, H. (2018). Application of CRISPR/Cas to Understand Cis- and Trans-Regulatory Elements in Plants (pp. 23–40). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-8657-6_2

Wu, S., Lee, C.-M., Hayashi, T., Price, S., Divol, F., Henry, S., … Gallagher, K. L. (2014). A plausible mechanism, based upon Short-Root movement, for regulating the number of cortex cell layers in roots. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(45), 16184–16189. https://doi.org/10.1073/pnas.1407371111

Xia, L., Zou, D., Sang, J., Xu, X., Yin, H., Li, M., … Zhang, Z. (2017). Rice Expression Database (RED): An integrated RNA-Seq-derived gene expression database for rice. *Journal of Genetics and Genomics*, *44*(5), 235–241. https://doi.org/10.1016/J.JGG.2017.05.003

Xu, X., & Qi, L. S. (2019). A CRISPR–dCas Toolbox for Genetic Engineering and Synthetic Biology. *Journal of Molecular Biology*. Academic Press. https://doi.org/10.1016/j.jmb.2018.06.037

Xu, W., Subudhi, P. K., Crasta, O. R., Rosenow, D. T., Mullet, J. E., & Nguyen, H. T. (2000). Molecular mapping of QTLs conferring stay-green in grain sorghum (Sorghum bicolor L. Moench). *Genome*, *43*(3), 461–469. https://doi.org/10.1139/g00-003

Yamaguchi, M., Goué, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J. C., … Demura, T. (2010). VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiology*, *153*(3), 906–914. https://doi.org/10.1104/pp.110.154013

Yamauchi, T., Colmer, T. D., Pedersen, O., & Nakazono, M. (2018). Regulation of Root Traits for Internal Aeration and Tolerance to Soil Waterlogging-Flooding Stress. *Plant Physiology*, *176*(2), 1118–1130. https://doi.org/10.1104/pp.17.01157

Yamauchi, T., Fukazawa, A., & Nakazono, M. (2017(b)). METALLOTHIONEIN genes encoding ROS scavenging enzymes are down-regulated in the root cortex during inducible aerenchyma formation in rice. *Plant Signaling & Behavior*, *12*(11), e1388976. https://doi.org/10.1080/15592324.2017.1388976

Yamauchi, T., Shimamura, S., Nakazono, M., & Mochizuki, T. (2013). Aerenchyma formation in crop species: A review. *Field Crops Research*, *152*, 8–16. https://doi.org/10.1016/j.fcr.2012.12.008

Yamauchi, T., Shiono, K., Nagano, M., Fukazawa, A., Ando, M., Takamure, I., … Nakazono, M. (2015). Ethylene Biosynthesis Is Promoted by Very-Long-Chain Fatty Acids during Lysigenous Aerenchyma Formation in Rice Roots. *Plant Physiology*, *169*(1), 180–193. https://doi.org/10.1104/PP.15.00106

Yamauchi, T., Yoshioka, M., Fukazawa, A., Mori, H., Nishizawa, N. K., Tsutsumi, N., … Nakazono, M. (2017(a)). An NADPH Oxidase RBOH Functions in Rice Roots during Lysigenous Aerenchyma Formation under Oxygen-Deficient Conditions. *The Plant Cell*, *29*(4), 775–790. https://doi.org/10.1105/tpc.16.00976

Yang, M.-H., Wu, D.-H., & Kao, C.-H. (2019). A Statistical Procedure for Genome-Wide Detection of QTL Hotspots Using Public Databases with Application to Rice. *G3 (Bethesda, Md.)*, *9*(2), 439–452. https://doi.org/10.1534/g3.118.200922

Yang, X., Li, Y., Ren, B., Ding, L., Gao, C., Shen, Q., & Guo, S. (2012). Drought-Induced Root Aerenchyma Formation Restricts Water Uptake in Rice Seedlings Supplied with Nitrate. *Plant and Cell Physiology*, *53*(3), 495–504. https://doi.org/10.1093/pcp/pcs003

Yan, X., Yagi, K., Akiyama, H., & Akimoto, H. (2005). Statistical analysis of the major variables controlling methane emission from rice fields. *Global Change Biology*, *11*(7), 1131–1141. https://doi.org/10.1111/j.1365-2486.2005.00976.x

Yonemaru, J., Yamamoto, T., Fukuoka, S., Uga, Y., Hori, K., & Yano, M. (2010). Q-TARO: QTL Annotation Rice Online Database. *Rice*, *3*(2–3), 194–203. https://doi.org/10.1007/s12284-010-9041-z

Yu, S. B., Xu, W. J., Vijayakumar, C. H. M., Ali, J., Fu, B. Y., Xu, J. L., … Li, Z. K. (2003). Molecular diversity and multilocus organization of the parental lines used in the International Rice Molecular Breeding Program. *Theoretical and Applied Genetics*, *108*(1), 131–140. https://doi.org/10.1007/s00122-003-1400-3

Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., … Pierce, N. A. (2011). NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, *32*(1), 170–173. https://doi.org/10.1002/jcc.21596

Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., Russa, M. La, … Lim, W. A. (2015). Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell January*, *1511*(1600). https://doi.org/10.1016/j.cell.2014.11.052

Zhang, L., Spratt, S. K., Liu, Q., Johnstone, B., Qi, H., Raschke, E. E., … Case, C. C. (2000). Synthetic zinc finger transcription factor action at an endogenous chromosomal site. Activation of the human erythropoietin gene. *The Journal of Biological Chemistry*, *275*(43), 33850–33860. https://doi.org/10.1074/jbc.M005341200

Zhang, Q., & Wing, R. A. (2013). *Genetics and genomics of rice*. (Q. Zhang & R. A. Wing, Eds.), *Genetics and Genomics of Rice*. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-7903-1

Zhang, Y., Yin, C., Zhang, T., Li, F., Yang, W., Kaminski, R., … Hu, W. (2015). CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs. *Scientific Reports*, *5*(1), 16277. https://doi.org/10.1038/srep16277

Zheng, B. S., Yang, L., Zhang, W. P., Mao, C. Z., Wu, Y. R., Yi, K. K., … Wu, P. (2003). Mapping QTLs and candidate genes for rice root traits under different water-supply conditions and comparative analysis across three populations. *Theoretical and Applied Genetics*, *107*(8), 1505–1515. https://doi.org/10.1007/s00122-003-1390-1

Zheng, B. S., Yang, L., Mao, C. Z., Zhang, W. P., & Wu, P. (2006). QTLs and candidate genes for rice root growth under flooding and upland conditions. *Acta Genetica Sinica*, *33*(2), 141–151. https://doi.org/10.1016/S0379-4172(06)60033-4

Zhan, J., Thakare, D., Ma, C., Lloyd, A., Nixon, N. M., Arakaki, A. M., … Yadegari, R. (2015). RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. *The Plant Cell*, *27*(3), 513–531. https://doi.org/10.1105/tpc.114.135657

Zhou, J., Xin, X., He, Y., Chen, H., Li, Q., Tang, X., … Zhang, Y. (2019). Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Reports*, *38*(4), 475–485. https://doi.org/10.1007/s00299-018-2340-3

Zhu, J., Kaeppler, S. M., & Lynch, J. P. (2005). Topsoil foraging and phosphorus acquisition efficiency in maize (Zea mays). *Functional Plant Biology*, *32*(8), 749. https://doi.org/10.1071/FP05005

Zhu, J., Brown, K. M., & Lynch, J. P. (2010). Root cortical aerenchyma improves the drought tolerance of maize (*Zea mays* L.). *Plant, Cell & Environment*, *33*(5), 740–749. https://doi.org/10.1111/j.1365-3040.2009.02099.x

ANNEXES

ANNEXES

Annexe 1 : Séquences des ARN ribosomaux du riz utilisé pour le design des sondes de déplétion par Nugen

ARN ribosomaux du riz

>XR_003238829.1_Oryza_5.8SrRNA_(LOC112936523)

GACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTAGCGAAATGCGATACCTGGTGTGAATTGCAGAATCCCGTGAAC
CATCGAGTCTTTGAACGCAAGTTGCGCCCGAGGCCATCCGGCCGAGGGCACGCCTGCCTGGGCGTCA

>XR_003238828.1_Oryza_28SrRNA_(LOC112936522)

CGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATATAAATAAGCGGAGGAGAAGAAACTTACGAGGATTCCCCTAGTAACG
GCGAGCGAACCGGGAGATGCCCAGCTTGAGAATCGGGCGGCCGCGCCGTCCGAATTGTAGTCTGGAGAGGCGTCCTCAGCGACGGACC
GGGCCCAAGTCCCCTGGAAAGGGGCGCCTGGGAGGGTGAGAGCCCCGTCCGGCCCGGACCCTGTCGCCCCACGAGGCGCCGTCAACG
AGTCGGGTTGTTTGGGAATGCAGCCCAAATCGGGCGGTAAACTCCGTCCAAGGCTAAATACAGGCGAGAGACCGATAGCGAACAAGTA
CCGCGAGGGAAAGATGAAAAGGACTTTGAAAAGAGAGTCAAAGAGTGCTTGAAATTGCCGGGAGGGAAGCGGATGGGGGCCGGCGA
TGCGCCCCGGCCGTATGCGGAACGGCTTCGGCTGGTCCGCCGATCGGCTCGGGGCGTGGACTGTTGTCGGCCGCGCCGGCGGCCAAAG
CCCGGGGGCTCCGCGCCCCCGGCAGCCGTCGTCGGCGCAGCCGGTCACCGCGCGCCTCTGGCGCGCCCCTCGGGGCGCTGCGCCGCAA
CGGCCTGCGGGCTCCCCATCCGACCCGTCTTGAAACACGGACCAAGGAGTCTGACATGCGTGCGAGTCGACGGGTTCTGAAACCTGGGA
TGCGCAAGGAAGCTGACGAGCGGGAGGCCCTCACGGGCCGCACCGCTGGCCGACCCTGATCTTCTGTGAAGGGTTCGAGTTGGAGCAC
GCCTGTCGGGACCCGAAAGATGGTGAACTATGCCTGAGCGGGGCGAAGCCAGAGGAAACTCTGGTGGAGGCTCGAAGCGATACTGAC
GTGCAAATCGTTCGTCTGACTTGGGTATAGGGGCGAAAGACTAATCGAACCATCTAGTAGCTGGTTCCCTCCGAAGTTTCCCTCAGGATA
GCTGGAGCCCATTACGAGTTCTATCGGGTAAAGCCAATGATTAGAGGCATCGGGGGCGCAACGCCCTCGACCTATTCTCAAACTTTAAAT
AGGTAGGACGGCGCGGCTGCTCCGGTGAGCCGCGCCACGGAATCGGGAGCTCCAAGTGGGCCATTTTTGGTAAGCAGAACTGGCGATG
CGGGATGAACCGGAAGCCTGGTTACGGTGCCGAACTGCGCGCTAACCTAGAACCCACAAAGGGTGTTGGTCGATTAAGACAGCAGGAC
GGTGGTCATGGAAGTCGAAATCCGCTAAGGAGTGTGTAACAACTCACCTGCCGAATCAACTAGCCCCGAAAATGGATGGCGCTGAAGC
GCGCGACCCACACCAGGCCATCTGGGCGAGCGCCATGCCCCGATGAGTAGGAGGGCGCGGCGGCCGCCGCAAAACCCGGGGCGCGAG
CCCGGGCGGAGCGGCCGTCGGTGCAGATCTTGGTGGTAGTAGCAAATATTCAAATGAGAACTTTGAAGGCCGAAGAGGAGAAAGGTTC
CATGTGAACGGCACTTGCACATGGGTAAGCCGATCCTAAGGGACGGGGTAACCCCGGCAGAGAGCGCGACCACGCGCGTGCCCCGAAA
GGGAATCGGGTTAAGATTTCCCGAGCCGGGACGTGGCGGTTGACGGCGACGTTAGGAAGTCCGGAGACGCCGGCGGGGGCCTCGGGA
AGAGTTATCTTTTCTGCTTAACGGCCCGCCAACCCTGGAAACGGTTCAGCCGGAGGTAGGGTCCAGCGGCCGGAAGAGCACCGCACGTC
GCGCGGTGTCCGGTGCGCCCCCGGCGGCCCTTGAAAATCCGGAGGACCGAGTACCGTCCACGCCCGGTCGTACTCATAACCGCATCAGG
TCTCCAAGGTGAACAGCCTCTGGCCAATGGAACAATGTAGGCAAGGGAAGTCGGCAAAACGGATCCGTAACTTCGGGAAAAGGATTGG
CTCTGAGGGCTGGGCTCGGGGGTCCCGGCCCCGAACCCGTCGGCTGCCGGCGGACTGCTCGAGCTGCTCGCGCGGCGAGAGCGGGCC
GCCGCGTGCCGGCCGGGGGACGGACCGGGAACGGCCCCCTCGGGGGCCTTCCCCGGGCGTCGAACAGCCGACTCAGAACTGGTACGG
ACAAGGGGAATCCGACTGTTTAATTAAAACAAAGCATTGCGATGGTCCTCGCGGATGCTGACGCAATGTGATTTCTGCCCAGTGCTCTGA
ATGTCAAAGTGAAGAAATTCAACCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAA
TTAGTGACGCGCATGAATGGATTAACGAGATTCCCACTGTCCCTGTCTACTATCCAGCGAAACCACAGCCAAGGGAACGGGCTTGGCGG
AATCAGCGGGGAAAGAAGACCCTGTTGAGCTTGACTCTAGTCCGACTTTGTGAAATGACTTGAGAGGTGTAGGATAAGTGGGAGCCCTC

GGGCGCAAGTGAAATACCACTACTTTTAACGTTATTTTACTTATTCCGTGAGTCGGAAGCGGGGCCTGGCCCCTCCTTTTGGCTCTAAGGC
CCGAGTCCCTCGGGCCGATCCGGGCGGAAGACATTGTCAGGTGGGGAGTTTGGCTGGGGCGGCACATCTGTTAAAAGATAACGCAGGT
GTCCTAAGATGAGCTCAACGAGAACAGAAATCTCGTGTGGAACAAAAGGGTAAAAGCTCGTTTGATTCTGATTTCCAGTACGAATACGA
ACCGTGAAAGCGTGGCCTATCGATCCTTTAGACCTTCGGAGTTTGAAGCTAGAGGTGTCAGAAAAGTTACCACAGGGATAACTGGCTTG
TGGCAGCCAAGCGTTCATAGCGACGTTGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTGAAGCAGAATTCACCAAGTGTTGGA
TTGTTCACCCACCAATAGGGAACGTGAGCTGGGTTTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGACCGTGCCGCGATAGT
AATTCAACCTAGTACGAGAGGAACCGTTGATTCACACAATTGGTCATCGCGCTTGGTTGAAAAGCCAGTGGCGCGAAGCTACCGTGTGC
CGGATTATGACTGAACGCCTCTAAGTCAGAATCCAAGCTAGCAAGCGGCGCCTGCGCCCGCCGCCCGCCCCGACCCACGTTAGGGGCGC
AAGCCCCCAAGGGCCCGTGCCACCGGCCAAGCCGGCCCGGCCGACGCGCCGCGGCCGGCCGCCTCGAAGCTCCCTTCCCAACGGGCGG
CGGGCTGAATCCTTTGCAGACGACTTAAATACGCGACGGGGCATTGTAAGTGGCAGAGTGGCCTTGCTGCCACGATCCACTGAGATCCA
GCCCCGCGTCGCACGGATTCGT

>XR_003238822.1_Oryza_18SrRNA_(LOC112936517)

TACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTCTCAAAGATTAAGCCATGCATGTGCAAGTATGAACTAATTCGAACTGTGAAACTGC
GAATGGCTCATTAAATCAGTTATAGTTTGTTTGATGGTACGTGCTACTCGGATAACCGTAGTAATTCTAGAGCTAATACGTGCAACAAACC
CCGACTTCCGGGAGGGGCGCATTTATTAGATAAAAGGCTGACGCGGGCTCCGCCCGCTGATCCGATGATTCATGATAACTCGACGGATC
GCACGGCCCTCGTGCCGGCGACGCATCATTCAAATTTCTGCCCTATCAACTTTCGATGGTAGGATAGGGGCCTACCATGGTGGTGACGG
GTGACGGAGAATTAGGGTTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGCAAATTACCC
AATCCTGACACGGGGAGGTAGTGACAATAAATAACAATACCGGGCGCTTTAGTGTCTGGTAATTGGAATGAGTACAATCTAAATCCCTTA
ACGAGGATCCATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTTAAGTTGTTGCAGTTAAAAA
GCTCGTAGTTGGACCTTGGGCCGGGCCGGCCGGTCCGCCTCACGGCGAGCACCGACCTGCTCGACCCTTCTGCCGGCGATGCGCTCCTG
GCCTTAACTGGCCGGGTCGTGCCTCCGGCGCCGTTACTTTGAAGAAATTAGAGTGCTCAAAGCAAGCCATCGCTCTGGATACATTAGCAT
GGGATAACATCATAGGATTCCGGTCCTATTGTGTTGGCCTTCGGGATCGGAGTAATGATTAATAGGGACAGTCGGGGGCATTCGTATTTC
ATAGTCAGAGGTGAAATTCTTGGATTTATGAAAGACGAACAACTGCGAAAGCATTTGCCAAGGATGTTTTCATTAATCAAGAACGAAAGT
TGGGGGCTCGAAGACGATCAGATACCGTCCTAGTCTCAACCATAAACGATGCCGACCAGGGATCGGCGGATGTTGCTTATAGGACTCCG
CCGGCACCTTATGAGAAATCAAAGTCTTTGGGTTCCGGGGGGAGTATGGTCGCAAGGCTGAAACTTAAAGGAATTGACGGAAGGGCAC
CACCAGGCGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTTACCAGGTCCAGACATAGCAAGGATTGACAGACTGAGAGCT
CTTTCTTGATTCTATGGGTGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGCGATTTGTCTGGTTAATTCCGTTAACGAACGAGACCTCAGC
CTGCTAACTAGCTATGCGGAGCCATCCCTCCGCAGCTAGCTTCTTAGAGGGACTATGGCCGTTTAGGCCACGGAAGTTTGAGGCAATAAC
AGGTCTGTGATGCCCTTAGATGTTCTGGGCCGCACGCGCGCTACACTGATGTATCCAACGAGTATATAGCCTTGGCCGACAGGCCCGGG
TAATCTTGGGAAATTTCATCGTGATGGGGATAGATCATTGCAATTGTTGGTCTTCAACGAGGAATGCCTAGTAAGCGCGAGTCATCAGCT
CGCGTTGACTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTCCTACCGATTGAATGGTCCGGTGAAGTGTTCGGATCGCGGCGACGG
GGGCGGTTCGCCGCCCCCGACGTCGCGAGAAGTCCATTGAACCTTATCATTTAGAGGAAGGAGAAGTCGTAACAAGGTTTCCGTAGGTG
AACCTGCGGAAGGATCATTG

>XR_003239795.1_Oryza_5SrRNA_(LOC112937280)

GGATGCGATCATACCAGCACTAAAGCACCGGATCCCATCAGAACTCCGAAGTTAAGCGTGCTTGGGCGAGAGTAGTACTAGGATGGGT
GACCTCCTGGGAAGTCCTCGTGTTGCATCCC

ARN ribosomaux chloroplastique du riz

>MG252500.1:91253-92743_Oryza_chloroplast_16S

TCTCATGGAGAGTTCGATCCTGGCTCAGGATGAACGCTGGCGGCATGCTTAACACATGCAAGTCGAACGGGAAGTGGTGTTTCCAGTGG
CGAACGGGTGAGTAACGCGTAAGAACCTGCCCTTGGGAGGGGAACAACAACTGGAAACGGTTGCTAATACCCCGTAGGCTGAGGAGCA
AAAGGAGAAATCCGCCCAAGGAGGGGCTCGCGTCTGATTAGCTAGTTGGTGAGGCAATAGCTTACCAAGGCGATGATCAGTAGCTGGT
CCGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTCCGCAATGGGCGAA
AGCCTGACGGAGCAATGCCGCGTGGAGGTGGAAGGCCCACGGGTCGTCAACTTCTTTTCTCGGAGAAGAAACAATGACGGTATCTGAG
GAATAAGCATCGGCTAACTCTGTGCCAGCAGCCGCGGTAAGACAGAGGATGCAAGCGTTATCCGGAATGATTGGGCGTAAAGCGTCTG
TAGGTGGCTTTTCAAGTCCGCCGTCAAATCCCAGGGCTCAACCCTGGACAGGCGGTGGAAACTACCAAGCTGGAGTACGGTAGGGGCA
GAGGGAATTTCCGGTGGAGCGGTGAAATGCATTGAGATCGGAAAGAACACCAACGGCGAAAGCACTCTGCTGGGCCGACACTGACACT
GAGAGACGAAAGCTAGGGGAGCAAATGGGATTAGAGACCCCAGTAGTCCTAGCCGTAAACGATGGATACTAGGTGCTGTGCGACTCGA
CCCGTGCAGTGCTGTAGCTAACGCGTTAAGTATCCCGCCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGGCCC
GCACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAAAGCGAAGAACCTTACCAGGGCTTGACATGCCGCGAATCCTCTTGAAAGAGAG
GGGTGCCCTCGGGAACGCGGACACAGGTGGTGCATGGCTGTCGTCAGCTCGTGCCGTAAGGTGTTGGGTTAAGTCTCGCAACGAGCGC
AACCCTCGTGTTTAGTTGCCACTATGAGTTTGGAACCCTGAACAGACCGCCGGTGTTAAGCCGGAGGAAGGAGAGGATGAGGCCAAGT
CATCATGCCCCTTATGCCCTGGGCGACACACGTGCTACAATGGGCGGGACAAAGGGTCGCGATCTCGCGAGGGTGAGCTAACTCCAAAA
ACCCGTCCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGCAGGAATCGCTAGTAATCGCCGGTCAGCCATACGGCGGTGAAT
CCGTTCCCGGGCCTTGTACACACCGCCCGTCACACTATAGGAGCTGGCCATGTTTGAAGTCATTACCCTTAACCGTAAGGAGGGGGATGC
CTAAGGCTAGGCTTGCGACTGGAGTGAAGTCGTAACAAGGTAGCCGTACTGGAAGGTGCGGCTGGATCACCTCCTTT

>MG252500.1:95166-98053_Oryza_chloroplast_23S

TTCAAAAGAGGAAAGGCTTGCGGTGGATACCTAGGTACCCAGAGACGAGGAAGGGCGTAGCAAGCGACGAAATGCTTCG
GGGAGTTGAAAATAAGCATAGATCCGGAGATTCCCAAATAGGTCAACCTTTTGAACTGCCTGCTGAATCCATGAGCAGGC
AAGAGACAACCTGGCGAACTGAAACATCTTAGTAGCCAGAGGAAAAGAAAGCAAAAGCGATTCCCGTAGTAGCGGCGAG
CGAAATGGGAGCAGCCTAAACCGTGAAAACGGGGTTGTGGGAGAGCAATACAAGCGTTGTGCTGCTAGGCGAAGCGGTT
GAGTGCCGCACCCTAGATGGCTAAAGTCCAGTAGCCGAAAGCATCACTAGCTTACGCTCTGACCCGAGTAGCATGGGGCA
CGTGGAATCCCGTGTGAATCAGCAAGGACCACCTTGCAAGGCTAAATACTCCTGGGTGACCGATAGCGAAGTAGTACCGT
GAGGGAAAGGTGAAAAGAACCCCCAGTGGGTAGTGAAATAGAACGTGAAACCGTGCTGAGCTCCCAAGCAGTGGGAGG
GGAAAGTGATCTCTGACCGCGTGCCTGTTGAAGAATGAGCCGGCGACTCATAGGCAGTGGCTTGGTTAAGGGAACGGAA
CCCACCGGAGCCGTAGCGAAAGCGAGTCTTCATAGGGCGATTGTCACTGCTTATGGACCCGAACCTGGGTGATCTATCCAT
GACCAGGATGAAGCTTGGATGAAACTAAGCAGAGGTCCGAACCGACTGATGTTGAAGAATCAGCGGATGAGTTGTGGTT
AGGGGTGAAATGCCACTCGAACCCAGAGCTAGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAGTTGACTGGACATCT
AGGGGTAAAGCACTGTTTCGGTGCGGGCTGCGCGAGCGGTACCAAATCGAGGCAAACTCTGAATACTAGATATGACCCAA
AAATAACAGGGGTCAAGGTCGGCCAGTGAGACGATGGGGGATAAGCTTCATCGTCGAGAGGGAAACAGCCCGGATCACC
AGCTAAGGCCCCTAAATGACCGCTCAGTGATAAAGGAGGTGGGGGTGCAAAGACAGCCAGGAGGTTTGCCTAGAAGCAG
CCACCCTTTAAAGAGTGCGTAATAGCTCACTGATCGAGCGCCCTTGCGCTGAAGATGAACGGGGCTAAGCGATCTGCCGA
AGCTGTGGGATGTCAAAATGCATCGGTAGGGGAGCGTTCCGCCTTAGAGGGAAGCAACCGCGAAAGCGGGGGTCGACG
AAGCGGAAGCGAGAATGTCGGCTTGAGTAACGAAAACATTGGTGAGAATCCAATGCCCCGAAAACCCAAGGTTTCCTCCG
CAAGGTTCGTCCACGGAGGGTGAGTCAGGGCCTAAGATCAGGCCGAAAGGCGTAGTCGATGGACAACAGGTCAATATTC
CTGTACTACCCCTTGTTGGTACGGAGGGACGGAGGAGGCTAGGTTAGCCGAAAGATGGTTATAGGTTTAAGGACACAAG
GTGACCCTGCTTTTTCAGGGTAAGAAGGGGTAGAGAAAATGCCTCGAGCCGAGGTCCGAGTACCAAGCGCTGCAGCGCT
GAAGTATGAGCCCCGTGGACTAGCCATTGCTTCTCCACGAGGCTCATACCAGGCGCTACGGCGCTGAAGTATGTAACCCAT

GCCATACTCCCAGGAAAAGCTCGAACGACCTTCAACAAAGGGGTACCTGTACCCGAAACCGACACAGGTGGGTAGGTAG
AGAATACCTAGGGGCGCGAGACAACTCTCTCTAAGGAACTCGGCAAAATAGCCCCGTAACTTCGGGAGAAGGGGTGCCC
CCTCGCAAAAGGGGGTCGCAGTGACCAGGCCCGGGCGACTGTTTACCAAAAACACAGGTCTCCGAAAGTCGTAAGACCAT
GTATGGGGGCTGACGCCTGCCCAGTGCCGGAAGGTCAAGGAAGTTGGTGAACTGATGACAGGGAAGCCGGCGACCGAA
GCCCCGGTGAACGGCGGCCGTAACTATAACGGTCCTAAGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCCGCACGAA
AGGCGTAACGATCTGGGCACTGTCTCGGAGAGAGACTCGGTGAAATAGACATGTCTGTGAAGATGCGGACTACCTGCACC
TGGACAGAAAGACCCTATGAAGCTTTACTGTTCCCTGGGATTGGCTTTGGGCCTTTCCTGCGCAGCTTAGGTGGAAGGCGA
AGAAGGCCCCCTTCCGGGGGGGCCCGAGCCATCAGTGAGATACCACTCTGGAAGAGCTCGGATTCTAACCTTGTGTCAGA
CCCGCGGGCCAAGGGACAGTCTCAGGTAGACAGTTTCTATGGGGCGTAGGCCTCCCAAAAGGTAACGGAGGCGTGCAAA
GGTTTCCTCGGGCCAGACGGACATTGGTCCTCGAGTGCAAAGGCAGAAGGGAGCTTGACTGCAAGACTCACCCGTCGAG
CAGAGACGAAAGTCGGCCTTAGTGATCCGACGGTGCCGAGTGGAAGGGCCGTCGCTCAACGGATAAAAGTTACTCTAGG
GATAACAGGCTGATCTTCCCCAAGAGTCCACATCGACGGGAAGGTTTGGCACCTCGATGTCGGCTCTTCGCCACCTGGAGC
TGTAGGTGGTTCCAAGGGTTGGGCTGTTCGCCCATTAATGCGGTACGTGAGCTGGGTTCAGAACGTCGTGAGACAGTTCG
GTCCATATCCGGTGTGGGCGTTAGAGCATTGAGAGGACCTTTCCCTAGTACGAGAGGACCGGGAAGGACGCACCTCTGGT
GTACCAGTTATCGTGCCTACGGTAAACGCTGGGTAGCCAAGTGCGGAGAGGATAACTGCTGAAAGCATATAAGTAGTAAG
CCCACCCCAAGATGAGTGCTCTCT

>MG252500.1:98149-98243_Oryza_chloroplast_4.5S

TAAGGTAGCGGCGAGACGAGCCGTTTAAATAGGTGTCAAGTGGAAGTGCAGTGATGTATGCAGCTGAGGCATCCTAACGAACGAACGA
TTTGAAC

>MG252500.1:98471-98591_Oryza_chloroplast_5S

TATTCTGGTGTCCTAGGCGTAGAGGAACCACACCAATCCATCCCGAATTTGGTGGTTAAACTCTACTGCGGTGACGATACTGTAGGGGAG
GTCCTGCGGCAAAATAGCTCGATGCCAGAAT

ARN ribosomaux chloroplastique du riz

>KY486275.1:c490935-490814_DPA mitochondrion_5S

AAACCGGGCACTACGGTGAGACGTGAAAACACCCGATCCCATTCCGACCTCGATATATATGTGGAATCGTCTTGCGCCATATGTACTGAA
ATTGTTCGGGAGACATGGTCAAAGCCCGGGGT

>KY486275.1:c492743-491049 Oryza sativa Indica Group cultivar DPA mitochondrion_18S rRNA, complete genome

ATCATAGTCAAAATCTGAGTTTGATCCTGGCTCAGAAGGAACGCTAGCTATATGCTTAACACATGCAAGTCGAACGTTGTTTTCGGGGAG
CTGGGCAGAAGGAAAAGAGGCTCCTAGCTAAAGTAGTCTCGCCCTGCTTCAAAACTACAGGGCGCGCGCTACGGCTTTGACCTAACGGC
CTCCGTTTGCTGGAATCGGAATAGTTGAGAACAAAGTGGCGAACGGGTGCGTAACGCGTGGGAATCTGCCGAACAGTTCGGGCCAAAT
CCTGAAGAAAGCTCAAAAGCGCTGTTTGATGAGCCTGCGTAGTATTAGGTAGTTGGTCGGGTAAAGGCTGACCAAGCCAATGATGCTTA
GCTGGTCTTTTCGGATGATCAGCCACACTGGGACTGAGACACGGCCCGGACTCCCACGGGGGGCAGCAGTGGGGAATCTTGGACAATG
GGCGAAAGCCCGATCCAGCAATATCGCGTGAGTGAAGAAGGGCAATGCCGCTTGTAAAGCTCTTTCGTCGAGTGCGCGATCATGACAG

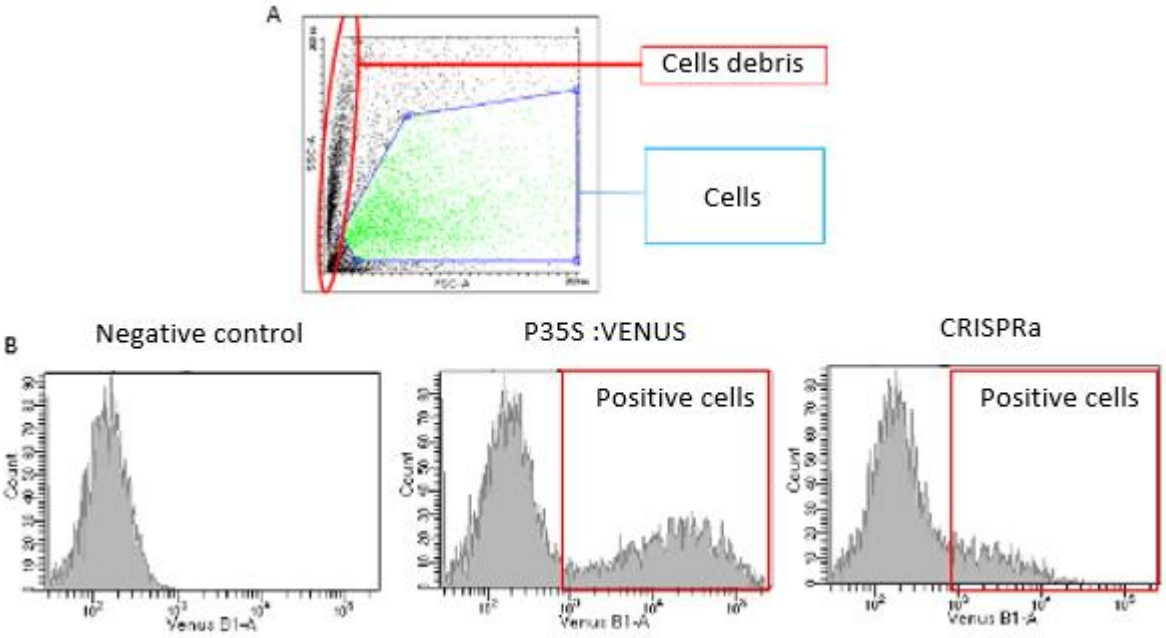
GACTCGAGGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAAGACGGGGGGGGCAAGTGTTCTTCGGAATGACTAGGCGTA
AAGGGCACGTAGGCGGTGAATCGGGTTGAAAGTGAAAGTCGCCAAAAAGTGGCGGAATGCTCTCGAAACCAATTCACTTGAGTGAGAC
AGAGGAGAGTGGAATTTCGTGTGTAGGGGTGAAATCCGTAGATCTACGAAGGAACGCCAAAAGCGAAGGCAGCTCTCTGGGTCCCTAC
CGACGCTGGGGTGCGAAAGCATGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGATGAGTGTTCGCCCTTGGTC
TACGCGGATCAGGGGCCCAGCTAACGCGTGAAACACTCCGCCTGGGGAGTACGGTCGCAAGACCGAAACTCAAAGGAATTGACGGGG
GCCTGCACAAGCGGTGGAGCATGTGGTTTAATTCGATACAACGCGCAAAACCTTACCAGCCCTTGACATATGAACAACAAAACCTGTCCT
TAACGGGATGGTACTGACTTTCATACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTTGGTCAAGTCCTATAACGAGCGA
AACCCTCGTTTTGTGTTGCTGAGACATGCGCCTAAGGAGAAATTGTGGTACGTAGTGGTAATAGTACGCGCCCCGCTCCGAAACAAAGA
AAAAGGTGCGTGCCGCACTCACGAGGGACTGCCAGTGAGATACTGGAGGAAGGTGGGGATGACGTCAAGTCCGCATGGCCCTTATGG
GCTGGGCCACACACGTGCTACAATGGCAATGACAATGGGAAGCAAGGCTGTAAGGCGGAGCGAATCCGGAAAGATTGCCTCAGTTCGG
ATTGTTCTCTGCAACTCGGGAACATGAAGTTGAAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATATGTACCCGGGCCCTGTAC
ACACCGCCCGTCACACCCTGGGAATTGGTTTCGCCCGAAGCATCGGACCAATGATCACCCATGACTTCTGTGTACCACTAGTGCCACAAA
GGCCTTTGGTGGTCTTATTGGCGCATACCACGGTGGGGTCTTCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGTGG
CTGGATTGAATCC

>KY486275.1:457475-461016_Oryza_mitochondrion_26S

ATGTTCGACGCTATCGAAAATCATGCATTGGATGGATGCCCGGGCATTGAGAAGGAAGGACGCTTTCAGAGGCGAAAGGCCATGGGGA
GAGAGAGCGTCTGTGATCCATGGATCTCCGATCGGGAAACCGTATCCAAGCTCCGTGGCTAGTCTGCGCTCTTTGGACTTTTCAAACTTA
GCGAACTGAAACATCTGAGTAGCTAAAGGAAGGAAAATCAACCGAGACCCCGTTAGTAGCGGCGAGCGAGAGCGGAACAAGGGTTTTC
ATCAAAAGAAATCCGAAGCGGTTTCATTCGATTTGTTGTGGATTGGATGATGGAAAAACCAGCAAGCAGCAAGCGTAGGCTGTGTTGCC
GTAGCGCGCCCTACGGAGTTGTACAAAGTCAGCAACCGTTTTGGGGCGCAAGCATAGGCAACACAGGACTGATGACGGGGTGGGGCGC
GCAGTGAACTGGTTTTCTAAAAAGATTGGAAGATCCGGCCAAAGAAGGTGATAGCCCTGTTCATTCGTTCCCATGGTTCGATCCTTCCCA
GTAAAACGCGGCGTGTTCGAATGATGATCGCTTTTACGCGAGAAAGGGGGACCACCCTCTAAGCCTAAGTATTCCTCAATGACCGATAG
CGTACAAGTACCGTGAGGGAAAGGTGAAAAGAACCCCAATCGGGGAGTGCAATAGAGAACCTGAGATCCGATGCGAACAATCAGTCGA
AGGAGCGGAGCGCGGGCGCACTCACTCTAACGGCGTACCTTTCGCATGATGGGTCAGCGAGGAAATGGGAACAGCGGCTTAAGCCATT
AGGTGTAGGCGCTTTCCAGAGGTGGAAGATTTACGTTCTTCCTATTTGACCCGAAACCGATCGATCTAGCCATGAGCAGGTTGAAGAGA
GCTCTAACAGGCCTTGGAGGACCGAACCCACGTATGTGGCAAAATACGGGGATGACTTGTGGCTAGGGGTGAAAGGCCAACCAAGATC
GGATATAGCTGGTTTTCCGCGAAATCTATTTCAGTAGAGCGTATGATGTCGATGGCCCGAGGTAGAGCACTCAATGGGCTAGGGTGGCC
CCCATTTCGCTTTACCAACCCCAGGGAAACTCCGAATACAGGCCGTCATCCTTAGTACAGACAGACTGATTGGGTGCTAAGATCCAAAGT
CGAGAGGGAAACAGCCCAGATCGTACGCTAAGGTCCCTAAGCAATCACTTAGTGGAAAAGGAAGTGATCGAGCGATGACAACCAGGAG
GTGGGCTTGGAAGCAGCCATCCTTTGAAGAAAGCGTAATAGCTCACTGGTCTAGCTCCATGGCACCGAAAATGTCTCAGGGCTCAAGTG
ATTCACCGAAGCGACGAGACCTTGAAAGCTGCTTTTTCAAGTGTCAGTAGCGGGACGTTCTGTCAATCGGGGAAGGTTTTTGGTGACAA
GACCTGGAGATATCAGAAGTGAGAATGCTGACATGAGTAACGAGAAATCCTGTGAAAAACACGATCGCCTGCCAGTGGAAGGCTTTCTG
CGTTCAGTCAATCTACGCAGAGTGAATCGGTCCCTAAGGAACCCCCGAAAGGGCTGCCGTCCGATGGGTACACGAAAGTGACGAAGTTG
CTTTGACTACTGAACCATGCCTGGATCGTCGGAGCGAATTGGATGATCGGGCCGAGGGCTGCCCCCTCTTCCCCTCGCTCTCCTTTCCTTA
ATATTCACCGTCGAGTCATCAAAGCCGTCAACTAATTCAGAGGGGCTCGGCTGGCCCGGTCGCCCTACGCTACTGGCGCTTCCAAAGGCG
AAGCTCTCGTCTTTGGCGACCAGCAAACTCTGGGCTAAAACCTGTAGTTTTGAAGCAAGGGATGAGCGCGAAAGCCGTTGCGCTTCCGT
ACACGCGCATACGTTTTCTTGCTCCGTCGCGCGTTAGTACAATAACGCGCAGCCGTTTTCTTGCTGGGGCGGACACGGATTTCTCTCTAAA
GGCGAAGACTCAGAAGTGGGCGAACACTCGGCCCAGCGGGCGAACATGCCGGCTCCGGCTCCGCGCACCTGCTGACACCTTTCGAAGC
ACTTTCACGTGTGAACCGAAGTCGTCTTGCCGAACTCCTTCCTTTCTCATTTCAGTCCTCGCCTTTTTCATCTTCCGTAGGGGCCTTTAGTCT
TTTGATTAGAGTAGAGGTCGCGAGAGAGCAGAGCGTACCGCCCTGCCATAGTCGCGAGGATCTTAATAGTCGGTCGCGACTGTTGTCAT
AGTCAACGACGGTTGAAACTTCCAGGAAAAAACTTCGAATTGGGAGGGCGATCCTCCCGGTGAGCTGACCGTACCCCAAACCGACACAG

GTGAACAAGTAGAGTATACTAGGGCGCGTCGAGAGAACCATGTCGAAGGAACTCGGCAAAATGACCCCGTAACTTCGGGAGAAGGGGT
GCTCTCCTATCTTTTGATTAGGAAAGCGGCACATACCAGGGGGTAGCGACTGTTTATTAAAAACACAGGACTCTGCTAAGTGGTAACACG
ATGTATAGAGTCTGACACCTGCCCGGTGCTGGAAGGTCGGAAGGAGAAGTGTGATAAGCTTTGAATGGAAGCCCCGGTAAACGGCGGC
AGTAACTCTAACTGTCCTAAGGTAGCGAAATTCCTTGTCGCATAAGTAGCGACCTGCACGAATGGTGTAACGACTGCCCCGCTGTCTCCG
ACATGGACCCGGTGAAATTGAATTCTCCGTGAAGATGCGGAGTACCAACGGCTAGACGGTAAGACCCCGTGCACCTTGACTATAGCTTC
GCAGTGACAACCTTGATCGAATGTGTAGGATAGGTGGGAGGTGGTGACACACAACGACCAATCCTGAAAGACCACTCTTTCGTCTAAGG
ATGCCTAACCGCCGCACCGATCATTCGGGGGGGGCGGGACACTGCGAGGTGGGTAGTTTATCTGGGGCGGATGCCTCCTAAAGAGTAA
CGGAGGTGTGCGAAGGTAGGCTCAAGCGTTGATTCTGCTCGTGAGCGTAATGGTATAAGCCTGCCTGACTGTGAGACCGACTGGTCGAA
CAGAGACGAAAGTCGGCCATAGTGATCCGGGAGTCCCGTGTGGAAGGGCTCTCGCTCAACGGATCAAAGGTACGCCGGGGATAACAGG
CTGATGACTCCCAAGAGCTCTTATCGACGGAGTCGTTTGGCACCTCGATGTCGACTCATCACATCCTGGGGTTGAAGAAGGTCCCAAGGG
TTCGGTTGTTCGCCGATGAAAGTGGTACGTGAGTTGGGTTTAGAACGTCGTGAGACAGTTCGGTTCCTATCTACCGTTGGTGTTAAAGGG
AGAACTGCGAGGAGCCAACCCTAGTACGAGAGGACTGGGTTGGGTCA
ACCTATGGTGTACCGGTTGTTATGCCAATAGCAGCGCCGGGCAGCTAAGTTGGTATGGAAGAACTGCTGCGCCGCGGGAAATCCTTCTC
TATACAAGTTCTCGTACGAGGTTTTTGAACAGAACTTCGATAGGCGAGAGGTGTAAGCACCGCGAGGTGTGAAGCGATCTCGTACTAAA
CGAATGGAACTTTCCAACCT

Annexe 2: Cell sorting



Annexe 2: Cell sorting
A. Classification of sorted evenements, B. Cell number according to their fluorescence

Annexe 3 : spacer and guide RNA sequences

>minimal_guide_RNA_sequence (without the 5' spacer)

TTTTAGAGCTATGCTGAAAAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACGAGTCGGTGC

>scRNA-MS2_sequence (**tetO+_spacer-minimal guide-MS2 sequence**)

CTCCCTATCAGTGATAGAGAGTTTTAGAGCTATGCTGAAAAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGCACGAGTCGGTGC**GCGCACATGAGGATCACCCATGTGC**

>scRNA-com_sequence (**-120pb+_35S_Spacer-minimal guide-com**)

ATAAGGAAGTTCATTTCATTGTTTTAGAGCTATGCTGAAAAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGCACGAGTCGGTGC**CTGAATGCCTGCGAGCATC**

Annexe 4: dCas9, VP64 and 3xSRDX sequences

>NLS-dCas9_sequence (in red , V5 tag- in green NLS sequence- in yellow RuvC domain (D10A) and HNH domain (H840A)mutation)

ATGGGAAAACCTATTCCTAATCCTCTGCTGGGCCTGGATTCTACCGCCCCAAAGAAGAAGCGCAAGGTCGGTATTCATGGAGTTCCTGCT
GCGGACAAGAAGTACTCCATCGGCCTCGCCATCGGCACCAATTCTGTTGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCGTCCAA
GAAGTTCAAGGTCCTCGGCAACACCGACCGCCACTCCATCAAGAAGAATCTCATCGGCGCCCTGCTGTTCGACTCTGGCGAGACAGCCG
AGGCTACAAGGCTCAAGAGGACCGCTAGACGCAGGTACACCAGGCGCAAGAACCGCATCTGCTACCTCCAAGAGATCTTCTCCAACGAG
ATGGCCAAGGTGGACGACAGCTTCTTCCACAGGCTCGAGGAGAGCTTCCTCGTCGAGGAGGACAAGAAGCACGAGCGCCATCCGATCT
TCGGCAACATCGTGGATGAGGTGGCCTACCACGAGAAGTACCCGACCATCTACCACCTCCGCAAGAAGCTCGTCGACTCCACCGATAAG
GCCGACCTCAGGCTCATCTACCTCGCCCTCGCCCACATGATCAAGTTCAGGGGCCACTTCCTCATCGAGGGCGACCTCAACCCGGACAAC
TCCGATGTGGACAAGCTGTTCATCCAGCTCGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCGATCAACGCCTCTGGCGTTGACGCC
AAGGCTATTCTCTCTGCCAGGCTCTCTAAGTCCCGCAGGCTCGAGAATCTGATCGCCCAACTTCCGGGCGAGAAGAAGAATGGCCTCTTC
GGCAACCTGATCGCCCTCTCTCTTGGCCTCACCCCGAACTTCAAGTCCAACTTCGACCTCGCCGAGGACGCCAAGCTCCAGCTTTCCAAGG
ACACCTACGACGACGACCTCGACAATCTCCTCGCCCAGATTGGCGATCAGTACGCCGATCTGTTCCTCGCCGCCAAGAATCTCTCCGACG
CCATCCTCCTCAGCGACATCCTCAGGGTGAACACCGAGATCACCAAGGCCCCACTCTCCGCCTCCATGATCAAGAGGTACGACGAGCACC
ACCAGGACCTCACACTCCTCAAGGCCCTCGTGAGACAGCAGCTCCCAGAGAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAACGGC
TACGCCGGCTACATCGATGGCGGCGCTTCTCAAGAGGAGTTCTACAAGTTCATCAAGCCGATCCTCGAGAAGATGGACGGCACCGAGGA
GCTGCTCGTGAAGCTCAATAGAGAGGACCTCCTCCGCAAGCAGCGCACCTTCGATAATGGCTCCATCCCGCACCAGATCCACCTCGGCGA
GCTTCATGCTATCCTCCGCAGGCAAGAGGACTTCTACCCGTTCCTCAAGGACAACCGCGAGAAGATTGAGAAGATCCTCACCTTCCGCAT
CCCGTACTACGTGGGCCCGCTCGCCAGGGGCAACTCCAGGTTCGCCTGGATGACCAGAAAGTCCGAGGAGACAATCACCCCCTGGAACT
TCGAGGAGGTGGTGGATAAGGGCGCCTCTGCCCAGTCTTTCATCGAGCGCATGACCAACTTCGACAAGAACCTCCCGAACGAGAAGGTG
CTCCCGAAGCACTCACTCCTCTACGAGTACTTCACCGTGTACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGGATGAGGAAGCC
AGCTTTCCTTAGCGGCGAGCAAAAGAAGGCCATCGTCGACCTGCTGTTCAAGACCAACCGCAAGGTGACCGTGAAGCAGCTCAAGGAG
GACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTCGAGATCTCCGGCGTCGAGGATAGGTTCAATGCCTCCCTCGGGACCTACCACGAC
CTCCTCAAGATTATCAAGGACAAGGACTTCCTCGACAACGAGGAGAACGAGGACATCCTCGAGGACATCGTGCTCACCCTCACCCTCTTC
GAGGACCGCGAGATGATCGAGGAGCGCCTCAAGACATACGCCCACCTCTTCGACGACAAGGTGATGAAGCAGCTGAAGCGCAGGCGCT
ATACCGGCTGGGGCAGGCTCTCTAGGAAGCTCATCAACGGCATCCGCGACAAGCAGTCCGGCAAGACGATCCTCGACTTCCTCAAGTCC
GACGGCTTCGCCAACCGCAACTTCATGCAGCTCATCCACGACGACTCCCTCACCTTCAAGGAGGACATCCAAAAGGCCCAGGTGTCCGGC
CAAGGCGATTCCCTCCATGAGCATATCGCCAATCTCGCCGGCTCCCCGGCTATCAAGAAGGGCATTCTCCAGACCGTGAAGGTGGTGGA
CGAGCTGGTGAAGGTGATGGGCAGGCACAAGCCAGAGAACATCGTGATCGAGATGGCCCGCGAGAACCAGACCACACAGAAGGGCCA
AAAGAACTCCCGCGAGCGCATGAAGAGGATCGAGGAGGGCATTAAGGAGCTGGGCTCCCAGATCCTCAAGGAGCACCCAGTCGAGAA
CACCCAGCTCCAGAACGAGAAGCTCTACCTCTACTACCTCCAGAACGGCCGCGACATGTACGTGGACCAAGAGCTGGACATCAACCGCCT
CTCCGACTACGACGTGGACGCTATTGTGCCGCAGTCCTTCCTGAAGGACGACTCCATCGACAACAAGGTGCTCACCCGCTCCGACAAGAA
CAGGGGCAAGTCCGATAACGTGCCGTCCGAAGAGGTCGTCAAGAAGATGAAGAACTACTGGCGCCAGCTCCTCAACGCCAAGCTCATCA
CCCAGAGGAAGTTCGACAACCTCACCAAGGCCGAGAGAGGCGGCCTTTCCGAGCTTGATAAGGCCGGCTTCATCAAGCGCCAGCTCGTC
GAGACACGCCAGATCACAAAGCACGTGGCCCAGATCCTCGACTCCCGCATGAACACCAAGTACGACGAGAACGACAAGCTCATCCGCGA
GGTGAAGGTCATCACCCTCAAGTCCAAGCTCGTGTCCGACTTCCGCAAGGACTTCCAGTTCTACAAGGTGCGCGAGATCAACAACTACCA
CCACGCCCACGACGCCTACCTCAATGCCGTGGTGGGCACAGCCCTCATCAAGAAGTACCCAAAGCTCGAGTCCGAGTTCGTGTACGGCG
ACTACAAGGTGTACGACGTGCGCAAGATGATCGCCAAGTCCGAGCAAGAGATCGGCAAGGCGACCGCCAAGTACTTCTTCTACTCCAAC
ATCATGAATTTCTTCAAGACCGAGATCACGCTCGCCAACGGCGAGATTAGGAAGAGGCCGCTCATCGAGACAAACGGCGAGACAGGCG
AGATCGTGTGGGACAAGGGCAGGGATTTCGCCACAGTGCGCAAGGTGCTCTCCATGCCGCAAGTGAACATCGTGAAGAAGACCGAGGT
TCAGACCGGCGGCTTCTCCAAGGAGTCCATCCTCCCAAAGCGCAACTCCGACAAGCTGATCGCCCGCAAGAAGGACTGGGACCCGAAGA
AGTATGGCGGCTTCGATTCTCCGACCGTGGCCTACTCTGTGCTCGTGGTTGCCAAGGTCGAGAAGGGCAAGAGCAAGAAGCTCAAGTCC
GTCAAGGAGCTGCTGGGCATCACGATCATGGAGCGCAGCAGCTTCGAGAAGAACCCAATCGACTTCCTCGAGGCCAAGGGCTACAAGG
AGGTGAAGAAGGACCTCATCATCAAGCTCCCGAAGTACAGCCTCTTCGAGCTTGAGAACGGCCGCAAGAGAATGCTCGCCTCTGCTGGC
GAGCTTCAGAAGGGCAACGAGCTTGCTCTCCCGTCCAAGTACGTGAACTTCCTCTACCTCGCCTCCCACTACGAGAAGCTCAAGGGCTCC
CCAGAGGACAACGAGCAAAAGCAGCTGTTCGTCGAGCAGCACAAGCACTACCTCGACGAGATCATCGAGCAGATCTCCGAGTTCTCCAA
GCGCGTGATCCTCGCCGATGCCAACCTCGATAAGGTGCTCAGCGCCTACAACAAGCACCGCGATAAGCCAATTCGCGAGCAGGCCGAGA
ACATCATCCACCTCTTCACCCTCACCAACCTCGGCGCTCCAGCCGCCTTCAAGTACTTCGACACCACCATCGACCGCAAGCGCTACACCTCT
ACCAAGGAGGTTCTCGACGCCACCCTCATCCACCAGTCTATCACAGGCCTCTACGAGACACGCATCGACCTCTCACAACTCGGCGGCGAT
TGA

>VP64_sequence (5xGS_LINKER-VP64)

AAGGGAGATGGCAGTGGGTCCGGGAGCGGGTCTGGTTCAGTGGACGACGCACTTGATGACTTTGATCTCGTATGTTGGATGCC
TTGGACGATTTCGACCTTGACATGCTTGATGCTCTTGATGATTTTGACCTTGACATGCTGGATGCACTGGACGATTTTGATTT
GGATATGCTCTAG

>3xSRDX_sequence (5xGS_LINKER-3xSRDX)

AAGGGAGATGGCAGTGGGTCCGGGAGCGGGTCTGGTTCAGTGGACCTTGACCTGGATTTGGAGTTGAGACTGGCTTCGCTCTT
GATCTGGACCTCGAACTGCGCTTGGGTTTCGCCCTTGATTTGGATCTCGAATTGAGATTGGATTCGCGTAG

Annexe 5; MCP-VP64 and COM-3xSRDX sequences

>NLS-1xGS_Linker-MCP-1xGS_Linker-VP64

ATGCCTAAAAAGAAGAGGAAGGTGGGCTCCATGGCAAGTAACTTTACACAATTTGTTCTGGTTGATAACGGGGCACGGGGGAT
GTGACCGTGGCACCAAGTAACTTCGCAAACGGGATTGCGGAGTGGATTAGTTCTAACTCTAGGAGCCAAGCGTACAAAGTTAC
GTGTTCGGTGCGGCAGAGCAGCGCCCAGAATCGGAAATACACCATTAAGGTTGAGGTGCCTAAAGGTGCCTGGCGCAGTTACT
TGAACATGGAACTGACCATTCCTATCTTCGCTACAAACAGTGATTGCGAGTTGATTGTCAAGGCAATGCAAGGATTGCTCAAA
GATGGAAATCCAATCCCGTCTGCAATCGCAGCCAATTCTGGAATCTACGGTAGTGACGCACTTGATGACTTTGATCTCGATAT
GTTGGATGCCTTGGACGATTTCGACCTTGACATGCTTGATGCTCTTGATGATTTTGACCTTGACATGCTGGATGCACTGGACG
ATTTTGATTTGGATATGCTCTAG

>NLS-1xGS_Linker-COM-1xGS_Linker-3xSRDX

ATGCCTAAAAAGAAGAGGAAGGTGGGCTCCATGAAGAGCATCAGGTGCAAGAATTGTAACAAACTCCTCTTAAGGCTGACTCA
TTCGATCATATCGAAATCAGATGTCCCCGCTGTAAGAGACACATAATAATGCTTAACGCCTGTGAGCATCCGACTGAAAAACA
TTGCGGAAAGAGAGAGAAAATCACACACTCAGATGAGACAGTCAGGTACGGTAGTCTTGACCTGGATTTGGAGTTGAGACTTG
GCTTCGCTCTTGATCTGGACCTCGAACTGCGCTTGGGTTCGCCCTTGATTTGGATCTCGAATTGAGATTGGGATTCGCGTAG

Annexe 6: promoter and reporter gene sequences

>1xtetO_minimalCamV_35S (mutation to add a 'NGG' PAM sequence, spacer_sequence)

CTCCCTATCAGTGATAGAGAAGGTATGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGACACGCTG

>7xtetO_minimalCamV_35S (mutation to add a 'NGG' PAM sequences, spacer_sequence)

CTCCCTATCAGTGATAGAGAAGGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAAGGTATGCAGACTTTACTCCCTATCAG
TGATAGAGAAGGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAAGGTATGACCAGTTTACTCCCTATCAGTGATAGAGAAG
GTATCTACAGTTTACTCCCTATCAGTGATAGAGAAGGTATATCCAGTTTACTCCCTATCAGTGATAGAGAAGGTATGCAAGAC
CCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGACACGCTG

>35S_promoter (spacer_sequence)

CCGACACTCTCGTCTACTCCAAGAATATCAAAGATACGGTCTCAGAAGACCAAAGGGCTATTGAGACTTTTCAACAAAGGGTA
ATATCGGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTCATCAAAAGGACAGTAGAAAAGGAAGGTGGCACCTA
CAAATGCCATCATTGCGATAAAGGAAAGGCTATCGTTCAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCA
CGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGG
GATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGACACGCTGAAA
TCACCAGTCTCTCTCTACAAATCTATCTCT

Annexe 7, article accepté dans Methods in Molecular Biology: Efficient genome editing in rice protoplasts using CRISPR/CAS9 construct

Martine Bes[1,2§], Leo Herbert[1,2§], Thibault Mounier[1,2§], Anne-Cécile Meunier[1,2], Franz Durandet[3], Emmanuel Guiderdoni[1,2], Christophe Périn[1,2]

[1] CIRAD, UMR AGAP, 34398 Montpellier cedex 5, France

[2] Université de Montpellier, CIRAD , INRA , Montpellier SupAgro, 34000 Montpellier, France

[3] Present Adress: IAGE company, Avenue Agropolis, F-34398 Montpellier Cedex 5,

§ All authors contributed equally to this work

Abstract

Genome editing technologies, mainly CRISPR/CAS9, are revolutionizing plant biology and breeding. Since the demonstration of its effectiveness in eukaryotic cells, a very large number of derived technologies has emerged. Demonstrating and comparing the effectiveness of all these new technologies in entire plants is a long, tedious and labor-intensive process that generally involves the production of transgenic plants and their analysis. Protoplasts, plant cells free of their walls, offer a simple, high throughput system to test the efficiency of these editing technologies in a few weeks timespan. We have developed a routine protocol using protoplasts to test editing technologies in rice. Our protocol, allows to test more than 30 constructs in protoplasts prepared from leaf tissues of 100, 11-days-old seedlings. CRISPR/CAS9 constructs effectiveness can be clearly established within less than a week. We provide here a full protocol, from designing sgRNA to mutation analysis.

Keywords Rice, CRISPR/CAS9, protoplast, genome editing

1 Introduction

The CRISPR/CAS9 technology became the dominant genome editing technology soon after the first demonstration of its eukaryotic cell efficacy in 2012 (Jinek et al., 2012). Since then, its use has increased exponentially, demonstrating its functionality in most species where it has been tested (Hess et al., 2017), including plants (Yin et al., 2017). Beyond its obvious interest in biology, recent results obtained in rice (Li et al., 2017a; Mundt et al., 1993; Shen et al., 2016; Xu et al., 2016) and tomato (Rodriguez-Leal et al., 2017) have demonstrated its interest also as a plant breeding tool. Many CRISPR/CAS9 derived technologies have emerged and continue to appear in a regular basis. Among the most emblematic examples include the discovery of other nucleases like CPF1 (Moreno-Mateos et al., 2017; Zetsche et al., 2015); the development of CAS9 variants capable of targeting a modified Protospacer Adjacent Motif (Hu et al., 2016; Kleinstiver et al., 2015); the use of CAS9 variants where the catalytic activity of the nuclease was inactivated and fused with various domains for transcriptional regulation of target (Gilbert et al., 2013; Li et al., 2017b) or single nucleotide changes (Gaudelli et al., 2017; Komor et al., 2016; Shimatani et al., 2017; Zong et al., 2017); optimization of DNA of repair sequence to induce targeted modifications or replacement of targets (Gil-Humanes et al., 2017; Paix et al., 2017).

These new systems have often been tested initially in isolated cells, with human cells being one of the most popular prototyping system (see for instance (Cong et al., 2013; Rahdar et al., 2015; Zetsche et al., 2015)). It is possible to have a similar approach in plants by using protoplasts, which are isolated cells with their walls removed. This approach has been used by some research teams to demonstrate the effectiveness of CRISPR/CAS9 constructions (see for instance (Collonnier et al., 2017; Lin et al., 2017; Malnoy et al., 2016), but protoplasts are used much less systematically than human cells. One of the reasons for their underutilization is the supposed difficulty of producing and using protoplasts, which requires fine-tuning protocols. In response to these difficulties, we have therefore decided to develop a simple, detailed and robust protocol for the production of protoplasts and its routine use to evaluate the effectiveness of CRISPR/CAS9 constructs in rice.

In this article, we describe in detail the production, processing, DNA extraction of protoplasts and molecular analyses for testing CRISPR/CAS9 technology in rice. This protocol makes it possible to carry out the transformation and evaluation of 20-30

constructions from one hundred 11-days-old rice seedlings in less than a week. In several places, we indicate the very critical points but also demonstrate the multipurpose of this protocol as quality points are also indicated at key steps to ensure the reproducibility of the protocol and its transferability to other laboratories. We believe this detailed protocol will be a useful resource for the rice and plant community and will popularize the use of protoplasts as a prototyping system for CRISPR technology evaluation.

2. MATERIALS

REAGENT OR RESSOURCE	SOURCE	IDENTIFIER
Tube, 13 ml, 95 x 16.8 mm, round base, PS	Starsted	55468001
Cytometric tubes with a round bottom	Falcon	352054
Nylon Mesh 70 µm BD Falcon pour tube 50 ml	Falcon	352350
Nylon mesh 30µm	Sefar AG	03-37/24
Petri dish	Greiner	1608043
COREX 25ml Centrifuge Round Bottom Tubes	Corex	8446
Cut ended 200µL pipette tips	StarLab	S1111-1706
2 Funnels	Nalgene 80mm	4252-0080
3 erlenmeyers 250ml	Schot Duran	1 216 36
Biological samples		
DNA and leaf tissues from cv Nipponbare	This work	N/A
Chemicals		
PolyEthylene-Glycol	Fluka	81242
Macerozyme R10 Onozuka	Duchefa	651803
Cellulase RS Onozuka	Duchefa	578704
Percoll	Sigma	P1644
Fluorescein Di Acetate	Sigma	F7378
Commercial Assays		
Endonuclease I T7	New England Biolabs	M0302L
PureLink PCR Purification	Fisher SCIENTIFIC	K31000
DNAeasy plant extraction kit	Qiagen	69104
RQ1 Rnase Dnase Free	Promega	M6101
KOD Hot start Poymerase	VWR	71086
HiSpeed Plasmid Maxi Kit	Qiagen	12663
Experimental models and organisms		
Rice *Oryza sativa* cv Nipponbare	See Methods	
Subcloning Efficiency DH5a Chemically Competent Cells	Fisher SCIENTIFIC	11583117
Oligonucleotides		
SE-CAO1-F: CGGGACGCTGATGACGTGC	This work	N/A
SE-CAO1-R: GCGTGCAGCTTTCCCCCAGT	This work	N/A
CR-CAO1-F: GGCAACCGGACCACCTCAAGAGCT	This work	N/A
CR-CAO1-R: AAACAGCTCTTGAGGTGGTCCGGT	This work	N/A
UBI-RAC1: TGCAATGCTCATTATCTCTAGAG	This work	N/A
M13-RP: CAGGAAACAGCTATGACC	This work	N/A

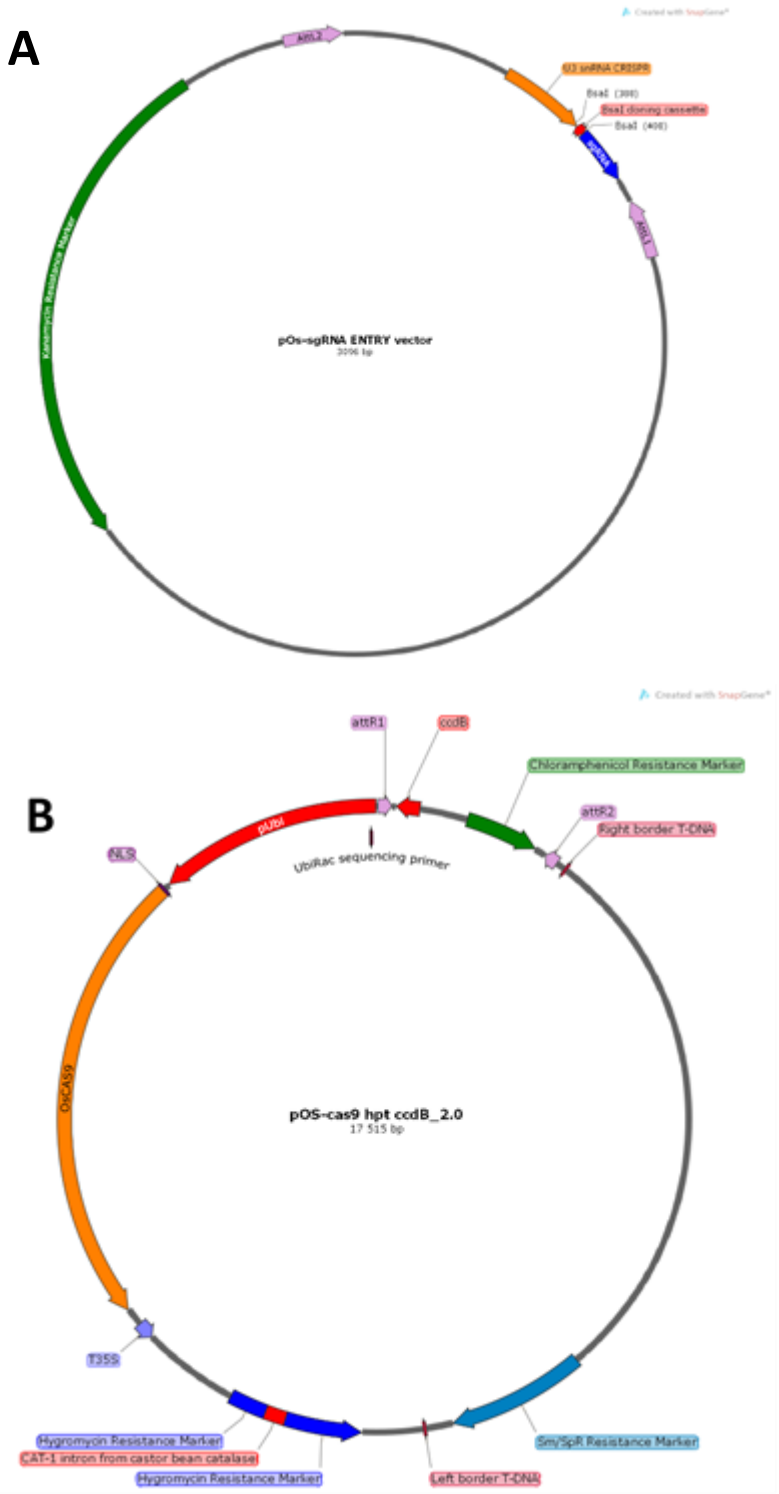


Figure 1: Plasmid maps of pOs-sgRNA ENTRY (A) and pOs-CAS9 hpt ccd_2.0 (B) BINARY vectors

Recombinant DNA

See also **Figure** 1A & 1B		
pOs-sgRNA ENTRY	(Miao et al., 2013)	N/A
pOs-CAS9 hpt ccdB-2.0	(Miao et al., 2013)/This work	N/A
PHKN29	(Kumagai and Kouchi, 2003)	N/A

Software and Algorithms

TEFOR	(Haeussler et al., 2016)
RGEN	(Bae et al., 2014)
ORYGENESDB	(Droc et al., 2006, 2009)

Other

Microscope (X 10 or X 20 objectives)	Leica	DM RXA2
Bioanalyseur Agilent 2100	AGILENT technologies	G2939A
DNA 1000 ScreenTape x7	AGILENT	5067-5582
Fuchs and Rosenthal cell counting cell	VWR	631-1125
Stove 26°C (with a stirrer in it)	Memmert	N/A
Centrifuge	Heraus Multifuge 3SR	N/A
Swing buckets	Heraus	75006441H

Table 1: Materials

2.1 Mix for digestion/ligation and CRISPR vector construction

1. The pENTRY vector is digested using BsaI enzyme. Add sequentially 2 µg of pOs-sgRNA ENTRY Miniprep vector 2µL of Bsa*I*-HF restriction enzyme plus 5 µL of 10x Bsa*I*-HF buffer QSP 50 µL H2Odd.

2. The ligation mix is prepared by mixing together 6 µL of dsDNA primers at 30ng/µL, 2.5 µL of pENTRY Bsa*I* digested vector (see below), 1 µL of T4 Ligase (1U/ µL) QSP 18 µL H_2Odd

2.2 Solutions for protoplasts digestion and purification

1. A volume of 200mL of Mannitol solution is prepared with 21.86g of Mannitol with 76mg of MES. The pH is adjusted to 5.7 and the solution is sterilized by filtration.

2. The first digestion solution is sterilize by filtration. A volume of 200mL is prepared by adding 222mg of $CaCl_2$, $2H_2O$, 21.86g Mannitol, 390mg MES and pH adjusted to 5.7

3. The second digestion solution is also sterilized by filtration. A volume of 20mL is prepared by combining 150mg of Macerozyme R10, 300mg of Cellulase RS, 200 µL of à 100x stock solution of BSA 10%, QSP 20mL completed with first digestion solution.

4. 200mL of W5 solution sterilized by filtration is prepared with 8.92g of NaCl, 18.36g of $CaCl_2$ $2H_2O$, 372mg of KCl, 388mg of MES buffer and pH adjusted to 5.7

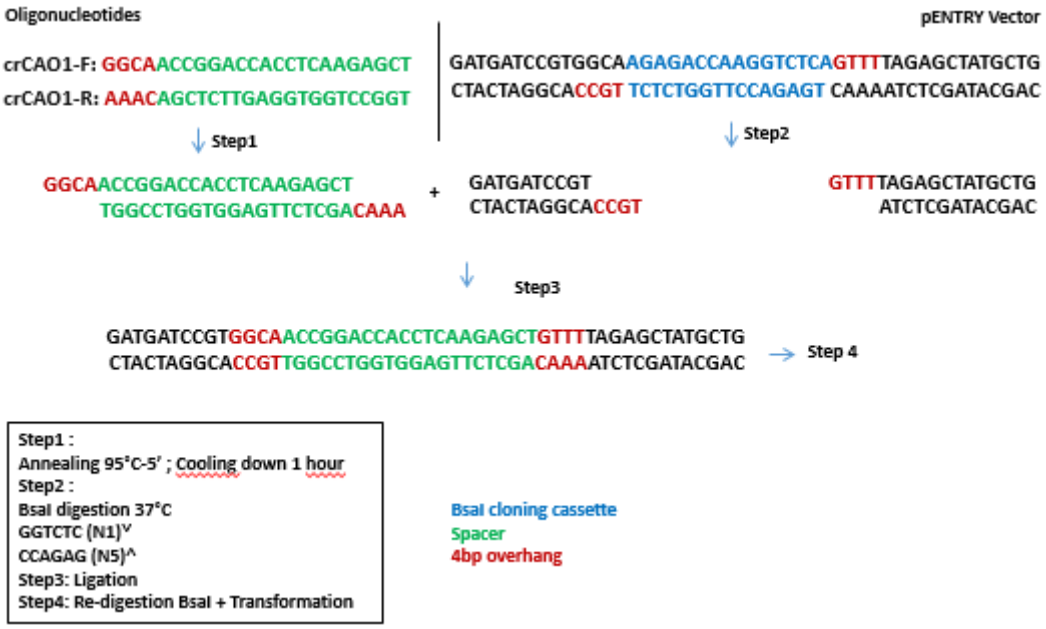


Figure 2: Workflow of protospacer subcloning in pOs-sgRNA ENTRY using CR-CAO1-F and CR-CAO1-R oligonucleotides. Step 1, annealing of oligonucleotides (see 3.2.1), step 2 digesting pOs-sgRNA ENTRY vector with BsaI (see 3.2.2) and step 3 ligation of oligonucleotides and pOs-sgRNA ENTRY BsaI digested plasmid (see 3.2.2)

2.4 Solutions for protoplasts transformation

1. Prepare 10mL of PEG solution, sterilize by filtration, aliquot in 2mL and store to -20°C until use. Add 364mg of Mannitol, 1mL of CaCl2 2H2O of stock solution 1M, 4g of PEG4000 Fluka®, pH adjusted to 5.7

---

3. METHODS

3.1 sgRNA design (*see* Note 1)

1 - Copy your targeted DNA sequence into the online TEFOR software (*see* **Note 2**) (Haeussler et al., 2016).

2 - Select the corresponding species, *Oryza sativa*

3 - Warning! The genome sequence in the database is from the reference cultivar Nipponbare.

4 - The software outputs a potential protospacer list and their position on your target genomic sequence.

5 - Choose two different protospacers, e. g. close to the start codon to generate knock-outs.

6 - Choose protospacers that have minimal potential off-target (Bae et al., 2014) (*see* **Note 3**)

7 – Complementary oligonucleotides corresponding to the variable part of the sgRNA have to be designed for cloning in the right way

8 – 4bp overhangs GGCA are added to the 5’ end of the sequence followed directly by NGG (Protospacer Adjacent Motif or PAM) on the targeted strand

9- 4bp overhangs AAAC are added to the 5’ end of the reverse complementary strand following CCN on the complementary reverse strand

10 - Be careful not to add the PAM (NGG) in your oligonucleotide sequence

11 - For example, for CAO1, the protospacer is ACCGGACCACCTCAAGAGCT*NGG*; therefore the CR-CAO1-F oligonucleotide is GGCAACCGGACCACCTCAAGAGCT and the reverse CAO1 oligonucleotide is CR-CAO1-R AAACAGCTCTTGAGGTGGTCCGGT

12 - Order oligonucleotides, with no primer purification and modification step

3.2 Cloning in pDest vector

3.2.1 Annealing oligonucleotides

1 - The oligonucleotides are diluted to 100µM with distilled water.

2 - 2µg of each oligonucleotide are added in a final volume of 50µL of distilled water.

3 - DNA is denatured at 95°C for 5 min in a dry block

4 - The block is switched off and the tube containing oligonucleotides is left in the dry bath for 1 hour

5 - Determine oligonucleotide concentration and dilute dsDNA to 100 ng/µL

3.2.2 Digestion/ligation in the pOs-sgRNA ENTRY vector (Figure 2)

1 - Prepare a digestion mix with 2µg of the pENTRY vector (**Table 2 & Table 3**) and do a negative control without dsDNA primers

2 - Incubate at 37°C OVN

3 - Check on a 1.2% agarose gel the complete digestion of the vector by loading 3µL of the mix

4 - Prepare the ligation mix (**Table 3**) (*see* **Note 4**)

5 - Incubate at 26°C for 2h-4h or overnight at 16°C

7 - Add 1uL of additional BsaI-HF and incubate one additional hour

3.2.3 Transformation in DH5alpha

1 - Follow your usual genetic transformation protocol (electroporation, heat shock) and select on Kanamycin

2 - Isolate 15 colonies

3 - Make the minipreps and check each plasmid using BsaI digestion; add a negative control (empty vector pOs-sgRNA ENTRY)

4 - Positive clones, with integrated spacer have lost their BsaI sites

5 - Sequence the spacer using M13 RP primer (**Table 1**)

3.2.4 Subcloning in vector pOs-CAS9 hpt ccdB-2.0

1 - After verification by sequencing, the construct containing pU3-sgRNA is subcloned by LR recombinaison following supplier's recommendations by adding 100ng of pOs-sgRNA ENTRY plasmid and 150 ng of pOs-CAS9 hpt ccdB-2.0 plasmid.

2 - Follow the classical bacterial genetic transformation protocol (electroporation, heat shock) and select on Spectinomycin (50µg/mL)

3 - Isolate fifteen colonies, if negative control is really negative (no colonies with the empty vector)

4 - Make the minipreps and check each plasmid with a SpeI/StuI double digestion; add a positive control (empty vector, pOs-CAS9 hpt ccdB-2.0)

5 - Positive vectors containing the insert should yield two bands at 9920 and 6527 bp

6 - Sequence two independent clones using UbiRAC primer (**Table 1**) to check correct insertion of the pU3-sgRNA sequence

3.3 Plant growth for protoplast preparation (*see* Note 5)

1 - Sterile 20 x20 Petri dishes containing 250 mL of MS/2 medium with 4.5g/L phytagel stored at 4°C are left at room temperature for 40 mn

2 - Meanwhile, 100 seeds are dehusked and then surface disinfected by immersing in 70°C ethanol in a 50 mL Falcon tube for 1mn

3 - After rinsing with sterile distilled water, seeds are disinfected in 50 mL of a 9.6° sodium hypochlorite solution with two drops of tween 20

4 - After 30 minutes of incubation under slow agitation (40 rpm), the seeds are rinsed three times in sterile distilled water

5 - The seeds are spread in a sterile petri dish under the hood and then inoculated in the 20x20 square Petri dish containing medium

6 - 25 seeds are inoculated per plate at equal distance from each other with an angle of 45°C with respect to the medium surface, embryo upwards

7 - Four plates (100 seeds) are incubated in the growth room in the dark at 26°C for 9 days.

3.4 Protoplast extraction (*see* Note 6)

1 - Prepare the Mannitol solution (**Table 4**) and the enzyme solution (**Table 5 & 6**) for protoplast isolation

2 - Collect 5-7 cm-long shoot segments located just above the root from the seedlings, discard the rest of the shoot and root tissues.



Figure 3: finely chopped 10-day-old rice seedling shoots.

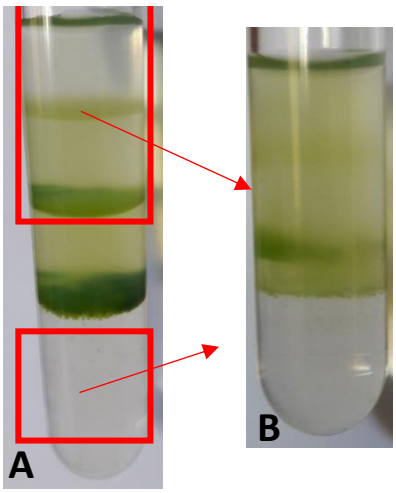


Figure 4: Process to remove plant cell debris from percoll centrifugated tubes. After centrifugation (A), purified protoplasts, up and down plant cell debris, were pooled together in a new tube (B) (red squares)

3 - Using a scalpel, chop the shoot tissues into the smallest possible pieces (1-2 mm) in a Petri dish (**Figure 3**); change the scalpel blade at least once during the process.

4 – Chop the tissues in rotating the cutting angle by 60° twice to favor the subsequent release of protoplasts

5 - Add 20mL of Mannitol solution and incubate for 10 minutes in the dark at room temperature (25°C)

6 - Remove the Mannitol solution using a pipet across a 70 µM sieve

7 - Add 20 mL of the enzyme solution (**Table 6**) into the Petri dish.

8 - Incubate in the dark at 26°C on a giratory shaker under slow motion (15 rpm) for 5 hours.

3.5 Protoplast purification (*see* Note 7)

1 - After 5 hours of incubation, transfer the protoplast and debris suspension to a sterile, 50 mL Erlenmeyer.

2 - Add 20 mL of W5 (**Table 7**) solution and gently shake the Erlenmeyer first clockwise during 10s then counter clockwise for 10 more seconds.

3 - Filter the suspension into a sterile Erlenmeyer through a 30µm nylon sieve mesh

4 - Add twice 15 mL of medium W5 and gently squeeze the filter to wring it out

5 - Filter the suspension again through a new sieve into a new Erlenmeyer

6 - Transfer protoplast suspension to three Corning tubes of 25 mL in glass or sterile polystyrene with a cap

7 - Centrifuge at 250g for 10 mn at room temperature in swing buckets with acceleration and deceleration parameters setup at 2.

8 - Discard the supernatant and leave a 1.5mL volume at the bottom of each tube

9 – Gather the 3 x 1,5ml tubes in a single tube

10 - Fill in the tube to a 25ml final volume with W5 medium

11 - Centrifuge at 250 g for 10 minutes with acceleration and deceleration at 2 at room temperature in swing buckets rotors

12 - While centrifugation is operating, prepare 5mL of the 60%, 40%, 20% and 10% Percoll solutions with W5 solution.

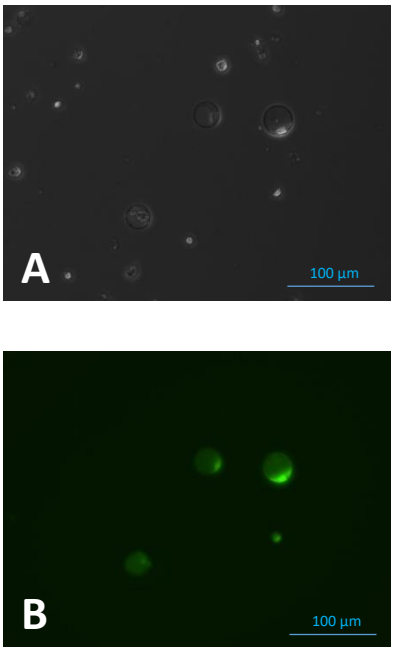


Figure 5: FDA test of rice protoplasts, view on bright field (A) and on UV plus GFP filter (B). FDA positive and thus living protoplasts were fluorescent on UV plus GFP filter (B)

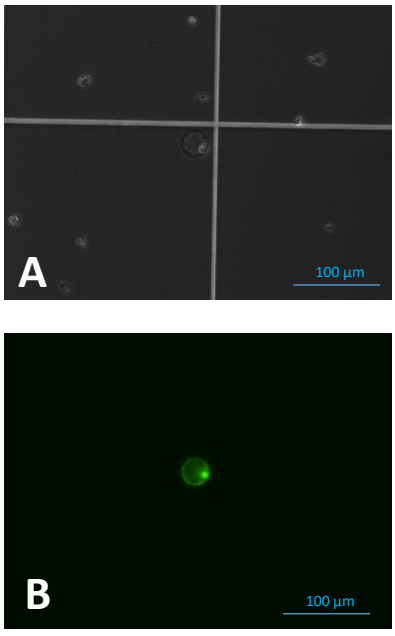


Figure 6: Transformed protoplast with the pHKN29 GFP plasmid, view on bright field (A) and on UV plus GFP filter (B). GFP positive transformed protoplasts were fluorescent on UV plus GFP filter (B).

13 - Prepare the Percoll gradient by successively adding 2 mL of each solution to 30 mL tubes starting with the 60% solution by slowly pouring each solution along the tube, tilting the pipette

14 - Eliminate the supernatant after centrifugation

15 - Leave 2mL of liquid containing the protoplast suspension in the bottom of the tube

16 - Pipet these 2mL and place them at the top of the Percoll gradient slowly, by gravity by tilting the pipet without disturbing the gradient

17 - Centrifuge at 250g for 10 minutes with acceleration and deceleration to 2 at room temperature in swinging buckets

18 - Remove the cellular debris by gently pipetting the debris layers (**Figure 4)**

19 - Gently mix the remaining layers by inversion

20 - Dispense the mix to 3x13 mL polystyrene tubes with caps

21 - Complete each tube to 13 ml with W5 medium

22 - Centrifuge at 250g for 10 minutes with acceleration and deceleration at 2 at room temperature in swinging buckets

23 - Eliminate the supernatant as much as possible; be careful the pellets are often invisible

24 – Pellets are pooled in a single 13mL centrifugation tube, filled with W5 solution

25 - Centrifuge at 250g for 10 minutes at room temperature with acceleration and deceleration parameters setup at 2

26 – Eliminate as much as supernatant as possible, be careful, pellet is invisible at this step

27 - If you want to stop here and carry on the transformation of protoplasts the next day, then resuspend the pellet in 1ml of W5 solution and store the tubes inclined in the dark at 26°C until the next day, then start again from point 25

28 - If you want to readily perform the transformation, resuspend the pellets in 500uL MMg solution (**Table 8**).

29 - Evaluate the protoplast viability using the FDA test

30 - Dilute the FDA stock solution by adding 450 uL of acetone to 5uL of 5mg/mL FDA stock solution.

31 - Add 2.5uL of this solution to 50uL of protoplast suspension in a 500uL tube and incubate for 5 minutes.

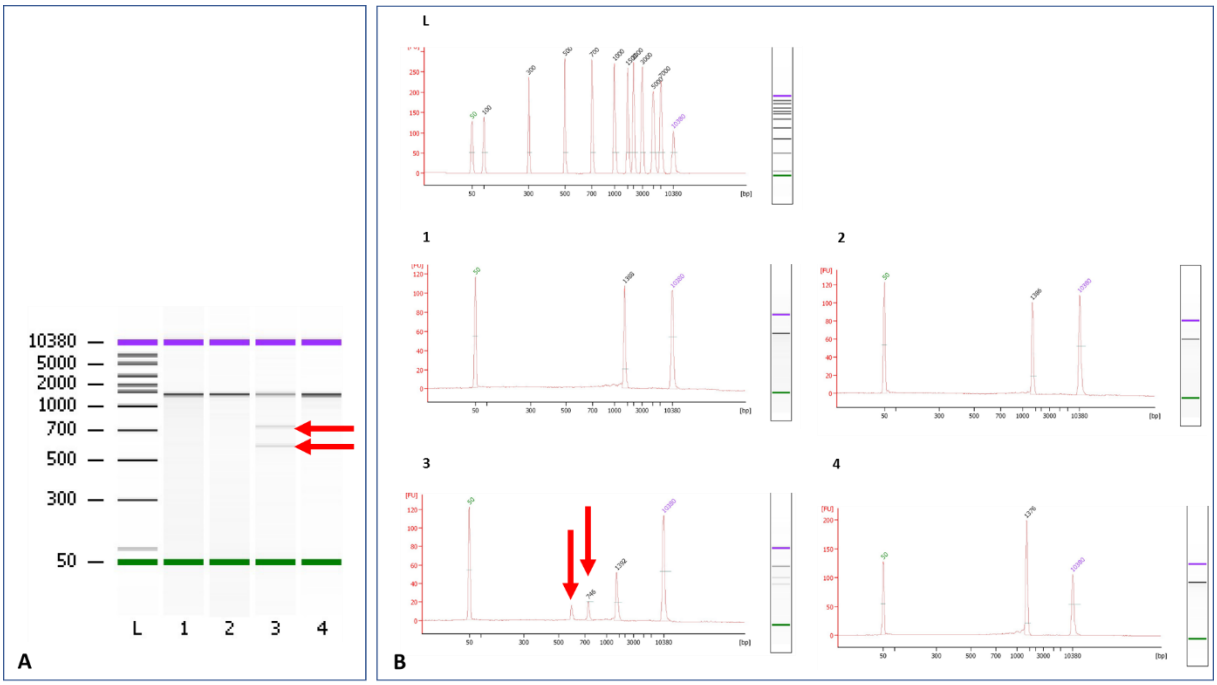


Figure 7: Bionalayser electrophoregram and gel of protoplast transformed with the pOs-CAS9 hpt ccd_2.0 BINARY vectors targeting the CAO1 gene. A) Gel of T7-EI treated PCR products B) Corresponding electrophoregramms. Red arrows indicate positive digested band from T7-EI reaction. L=Ladder, 1=Nipponbare protoplasts, 2=pHK29 GFP plasmid transformed in protoplasts, 3= pOs-CAS9 hpt ccd_2.0 BINARY vectors targeting the CAO1 gene transformed in protoplasts, 4=Nipponbare seedling DNA.

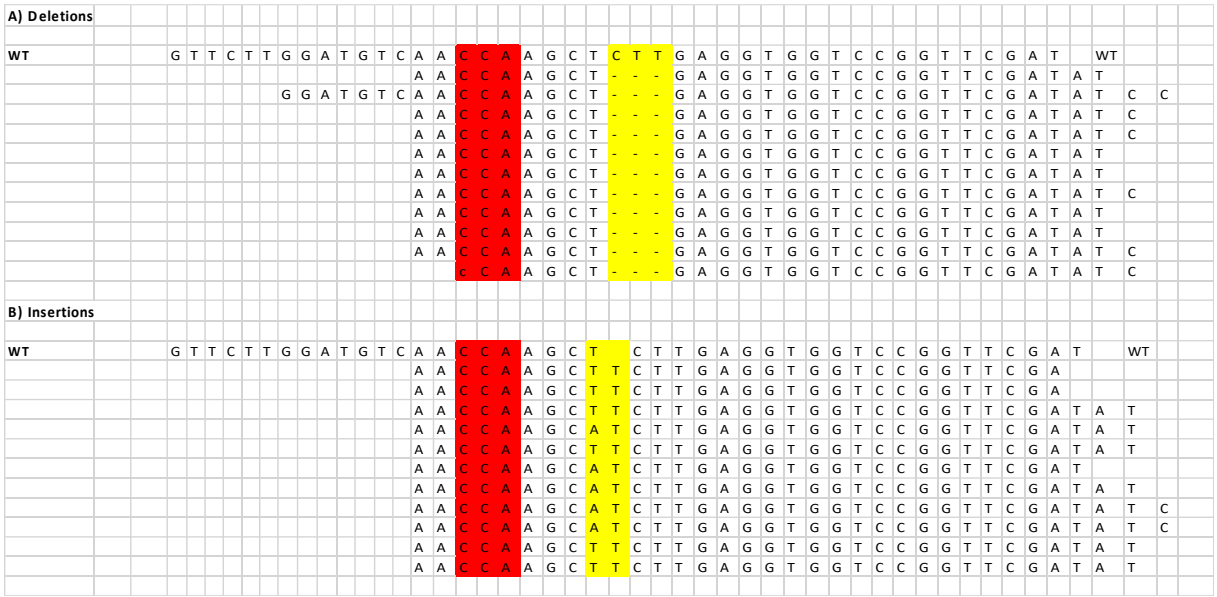


Figure 8: Sequences alignment of CA01-CRISPR-targeting sequences for 23 independent clones. A) Deletions, B) Insertions.

32 - Readily and rapidly (less than 10 minutes because fluorescence is rapidly fainting) estimate the number of fluorescent and non-fluorescent protoplasts per mL using a Mallasey's cell or a Rosenthal cell placed under UV light to visualize fluorescein (absorption at 494nm and emission at 512 nm)

33 - All fluorescent protoplasts are alive (**Figure 5)**. Usually, the viability percentage is close to 100%

34 - The minimum concentration required to perform a transformation is $2.10e^5$ protoplasts/mL and the optimal concentration is $2.10e^6$ protoplasts/mL

35 - Whether the protoplast suspension is too concentrated, adjust the concentration to $2.10e^6$ protoplasts/mL using MMg solution.

36 – Let the protoplasts incubate for 1 hour before transformation

3.6 Protoplast transformation (*see* Note 8)

1 - Prepare a number of tubes corresponding to the number of transformations you wish to perform

2 - Add 10µg of purified plasmid DNA (HiSpeed Plasmid Maxi Kit) in a maximum volume of 20µL in each 5mL cytometric tube.

3 - Add 200 µL of protoplasts in MMg solution, ideally at a density of $2.10e^6$ protoplasts/mL, in each tube.

4 - Then add 200µL of PEG 40% (*see* **Note 9**) (**Table 9**) by slowly pouring the PEG along the tube wall while achieving a long rotation movement inside the tube.

5 - Let incubate for 15 minutes in darkness at room temperature.

6 - To stop the reaction, slowly add 900µL of W5 solution stirring gently by inversion.

7 - Centrifuge at 250g for 10 minutes with acceleration and deceleration to 2 at room temperature in swinging buckets

8 - Remove the supernatant, be careful not removing the pellet, which is invisible

9- Add 500µL of W5 medium and incubate at 26°C in the dark for 24h and up to 48h, tubes left horizontally.

10 - Carry out a FDA viability test at 24h after transformation

11 - Dilute the FDA stock solution by adding to 5µL of 5mg/mL stock solution a volume of 450µL acetone

12 - Add 2.5µL of this solution to 50µL of protoplast solution in a 500µL tube and incubate for 5 mn

13 - Readily and rapidly (less than 10 minutes because fluorescence is rapidly fainting) estimate the number of fluorescent and non-fluorescent protoplasts per mL using a Mallasey's cell or a Rosenthal cell placed under UV light to visualize fluorescein (absorption at 494nm and emission at 512 nm).

3.7 DNA extraction

1 - Add 45µL of DNA RQ1 buffer and 45µL of DNAse to 450µL of protoplast solution

2 - Incubate at 26°C for 30 mn

3 - Add 45µL of STOP solution

4 - Centrifuge at 500g for 3 mn

5 - Eliminate the supernatant

6 - Add 400µL of Dneasy Plant Extraction Kit AP1 buffer and incubate at 65°C for 10 mn.

7 - Extract the DNA by following the conditions of the DNAeasy plant extraction kit.

8 - Elute DNA in a volume of 100µL of RNA and DNAse free water

3.8 Mutation analysis (*see* Note 10)

1 - A PCR is carried out from the DNA isolated from protoplasts, using Nipponbare gDNA as control, exemplified here using the CAO1-fw and CAO1-rv primers and the following PCR conditions (Table 1)

2 - After amplification control on a 1% agarose gel, the PCR products are purified with a PureLink PCR Purification column and dosed.

3 - 200ng of purified PCR product are denatured with a mixture containing 2µL of 10XNEB buffer 2, qsp 19µL of water at 95°C 5min, then hybridized from 95°C to 85°C at 2°C/s, then from 85°C to 25°C at 0.1°C/s in a thermocycler.

4 – 1 µL of A T7E-I is added to the reaction and incubated at 37°C for 15 minutes.

5 - A volume of 1.5µL of EDTA at 0.25M is added to stop the reaction.

6 - 10µL are loaded on an agarose gel and 1µL on the Agilent Nano chip and then loaded to the Agilent 2100 Bioanalyzer. The presence of two bands at expected size after

digestion of the PCR product is a confirmation of the activity of the CAS9 to induce mutations in your target. For instance, see the results obtained with a CRISPR/CAS9 construction targeting CAO1 in **Figure 7**.

7 - The PCR product can also be mass cloned in a PCR vector such as pGEMt following the supplier's recommendations pGEM-t Vector System I Promega A3600 (**Material**).

8 – For instance, 23 independent clones for CAO1 were randomly selected and their inserts sequenced, aligned to get an idea of the diversity of CAS9-induced mutations (**Figure 8**)

4 Notes

1. The design uses the entry and modified binary vectors of (Miao et al., 2013.) respectively "pOs-sgRNA ENTRY" (**Figure 1A**) and "pOs-CAS9 hpt ccdB-2.0" (**Figure 1B**). The vector pOs-CAS9 hpt ccdB-2.0 was obtained by removing the promotor–hygromycin-terminator insert and subcloning the p70s-Hygromycin resistance gene (containing catalase intron)-t35s from pCAMBIA5300 using KpnI/AflII restriction site.

2. Start by defining the gene you want to knock-out and then the area to target, close to the start codon if you want to create loss-of-function mutations. In the following protocol, the CAO1 gene will be taken as an example. The software used to predict spacer is indicative, it is possible to use other softwares like cas-offinder (Bae et al., 2014.) on the rgen website (http://www.rgenome.net).

3. The off-target probability cleavage is lower if mismatches are found in the 12 PAM proximal nucleotides

4. Trick : if you have several dsDNA primers to clone into ENTRY vector, don’t prepare a mix without primers as it increase the probability of vector religation

5. All handling, sowing and disinfection are carried out under a laminar flow hood following the usual instructions. Protoplast production is performed on 9 days-old seedlings.

[6.] All first steps are performed in sterile conditions, if possible in the dark or in low light, by gently manipulating protoplasts. Cut cones of 200mL are used to transfer or mix protoplasts. Swing rotors must be used for all centrifugations

[7.] If you wish to perform a transformation directly after extraction of protoplasts, thaw the PEG in ice 45 minutes before purification of protoplasts

[8]. Do a control transformation using a construct allowing strong constitutive expression of a reporter gene. For instance the pHK29-p35:GFP construct can be used as a positive control for protoplast transformation (**Figure 6**)

[9]. When you defroze you PEG aliquot, it should be transparent

10. The NEB T7-EI kit is used for the detection of CRISPR/CAS9 induced mutations

ACKNOWLEDGEMENTS

This work was partly supported by the French government through the programme "Investissements d'Avenir" and by the French National Research Agency, in the frame of the research project called "GENIUS" (Genome ENgineering Improvement for Useful plants of a Sustainable agriculture), ANR11-BTBR-0001-GENIUS.. Thibault Mounier PhD is partly funded by CRP-Rice and Montpellier SupAgro. Léo Herbert is the recipient of a CIRAD and INRA PhD fellowship.

ABBREVIATIONS

CRISPR	Clustered regularly interspaced short palindromic repeats
gRNA	guide RNA
dsDNA	Double-stranded DNA
PAM	Protospacer Adjacent Motif

5. Tables

Reagent	Final concentration	Vol 1X
pOs-sgRNA ENTRY miniprep	40ng/µL	variable µL
BsaI-HF	1X	2µL
Buffer BsaI-HF	U/µL	5µL
H2O dd	-	QSP 50µL

Table 2: digestion mix for pOs-sgRNA ENTRY

Reagent	Final concentration	Vol 1X
dsDNA primers	30ng/µL	6µL
pOs-sgRNA ENTRY digested (table 2)	25ng/µL	2.5µL
T4 Ligase	U/µL	1µL
H2O dd	-	8.5µL

Table 3: Ligation mix for pOs-sgRNA ENTRY

Reagent	Final concentration	Qte for 200mL
Mannitol	600mM	21.86g
MES Buffer	2mM	76mg
pH 5.7		

Table 4: Mannitol solution

Reagent	Final concentration	Qte for 200mL
$CaCl_2, 2H_2O$	7.5mM	222mg
Mannitol	600mM	21.86g
MES Buffer	10mM	390mg
pH 5.7		

Table 5: Solution n°1 for digestion

Reagent	Final concentration	Qte for 20mL
Macerozyme R10	0.75%	150mg
Cellulase RS	1.5%	300mg
BSA 10%	0.1%	200µL stock solution 100X
Solution n°1 digestion (Table 5)	-	QSP 20mL

Table 6: Solution n°2 for digestion

Reagent	Final concentration	Qte for 200mL
NaCl	154mM	8.92g
$CaCl_2, 2H_2O$	125mM	18.36g
KCl	5mM	372mg
MES Buffer	2mM	388mg
pH 5.7		
Osmolarity 635 mosmol		

Table 7: W5 solution

Reagent	Final concentration	Qte for 200mL
Mannitol	400mM	14.56g
$MgCl_2$, $6H_2O$	15mM	610mg
$CaCl_2, 2H_2O$	5mM	111mg
MES Buffer	4mM	160mg
pH 5.7		
Osmolarity 464 mosmol		

Table 8: MMg solution

Reagent	Final concentration	Qte for 10mL
Mannitol	200mM	364mg
$CaCl_2, 2H_2O$	100mM	1mL Stock solution 1M
PEG4000 Fluka®	40%	4g

Table 9: PEG solution

Prepare 2mL aliquote and store at -20°C until use

REFERENCES

Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30:1473-1475.

Collonnier, C., Epert, A., Mara, K., Maclot, F., Guyon-Debast, A., Charlot, F., White, C., Schaefer, D.G., and Nogue, F. (2017). CRISPR-Cas9-mediated efficient directed mutagenesis and RAD51-dependent and RAD51-independent gene targeting in the moss Physcomitrella patens. Plant biotechnology journal 15:122-131.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science 339:819-823.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551:464-471.

Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C.V., Sanchez-Leon, S., Baltes, N.J., Starker, C., Barro, F., Gao, C., et al. (2017). High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. The Plant journal : for cell and molecular biology 89:1251-1262.

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.

Haeussler, M., Schonig, K., Eckert, H., Eschstruth, A., Mianne, J., Renaud, J.B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome biology 17:148.

Hess, G.T., Tycko, J., Yao, D., and Bassik, M.C. (2017). Methods and Applications of CRISPR-Mediated Base Editing in Eukaryotic Genomes. Molecular cell 68:26-43.

Hu, X., Wang, C., Fu, Y., Liu, Q., Jiao, X., and Wang, K. (2016). Expanding the Range of CRISPR/Cas9 Genome Editing in Rice. Molecular plant 9:943-945.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816-821.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P., Li, Z., Peterson, R.T., Yeh, J.R., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature 523:481-485.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533:420-424.

Kumagai, H., and Kouchi, H. (2003). Gene silencing by expression of hairpin RNA in Lotus japonicus roots and root nodules. Molecular plant-microbe interactions : MPMI 16:663-668.

Li, X., Zhou, W., Ren, Y., Tian, X., Lv, T., Wang, Z., Fang, J., Chu, C., Yang, J., and Bu, Q. (2017a). High-efficiency breeding of early-maturing rice cultivars via CRISPR/Cas9-mediated genome editing. Journal of genetics and genomics = Yi chuan xue bao 44:175-178.

Li, Z., Zhang, D., Xiong, X., Yan, B., Xie, W., Sheen, J., and Li, J.F. (2017b). A potent Cas9-derived gene activator for plant and mammalian cells. Nature plants 3:930-936.

Lin, C.S., Hsu, C.T., Yang, L.H., Lee, L.Y., Fu, J.Y., Cheng, Q.W., Wu, F.H., Hsiao, H.C., Zhang, Y., Zhang, R., et al. (2017). Application of protoplast technology to CRISPR/Cas9 mutagenesis: From single cell mutation detection to mutant plant regeneration. Plant biotechnology journal.

Malnoy, M., Viola, R., Jung, M.H., Koo, O.J., Kim, S., Kim, J.S., Velasco, R., and Nagamangala Kanchiswamy, C. (2016). DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. Frontiers in plant science 7:1904.

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J., Gu, H., and Qu, L.J. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. Cell research 23:1233-1236.

Moreno-Mateos, M.A., Fernandez, J.P., Rouet, R., Vejnar, C.E., Lane, M.A., Mis, E., Khokha, M.K., Doudna, J.A., and Giraldez, A.J. (2017). CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. Nature communications 8:2024.

Mundt, D.J., Kelsey, J.L., Golden, A.L., Panjabi, M.M., Pastides, H., Berg, A.T., Sklar, J., and Hosea, T. (1993). An epidemiologic study of sports and weight lifting as possible risk factors for herniated lumbar and cervical discs. The Northeast Collaborative Group on Low Back Pain. The American journal of sports medicine 21:854-860.

Paix, A., Folkmann, A., Goldman, D.H., Kulaga, H., Grzelak, M.J., Rasoloson, D., Paidemarry, S., Green, R., Reed, R.R., and Seydoux, G. (2017). Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. Proceedings of the National Academy of Sciences of the United States of America 114:E10745-E10754.

Rahdar, M., McMahon, M.A., Prakash, T.P., Swayze, E.E., Bennett, C.F., and Cleveland, D.W. (2015). Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. Proceedings of the National Academy of Sciences of the United States of America 112:E7110-7117.

Rodriguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E., and Lippman, Z.B. (2017). Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. Cell 171:470-480 e478.

Shen, L., Wang, C., Fu, Y., Wang, J., Liu, Q., Zhang, X., Yan, C., Qian, Q., and Wang, K. (2016). QTL editing confers opposing yield performance in different rice varieties. Journal of integrative plant biology.

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., Teramura, H., Yamamoto, T., Komatsu, H., Miura, K., et al. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. Nature biotechnology 35:441-443.

Xu, R., Yang, Y., Qin, R., Li, H., Qiu, C., Li, L., Wei, P., and Yang, J. (2016). Rapid improvement of grain weight via highly efficient CRISPR/Cas9-mediated multiplex genome editing in rice. Journal of genetics and genomics = Yi chuan xue bao 43:529-532.

Yin, K., Gao, C., and Qiu, J.L. (2017). Progress and prospects in plant genome editing. Nature plants 3:17107.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163:759-771.

Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., Qiu, J.L., Wang, D., and Gao, C. (2017). Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nature biotechnology 35:438-440.

Toward the development of a new approach to decipher the molecular mechanisms involved in root aerenchyma formation in rice

Author: Thibault Mounier

Supervisor: Christophe Périn, CIRAD

Resume

Aerenchyma are cavities that develop from the cortex cells of many cultivated plants and play a key role in submergence tolerance. They ensure the diffusion of oxygen from the shoot to the root and are part of the adaptation mechanisms to flooding in rice. Aerenchyma also has a favorable role in many limiting crop conditions, such as nutritional deficiencies or water stress. The genes and molecular mechanisms involved in its formation are mostly unknown. However, their knowledge is an essential step in better understanding their formation but also in developing varieties that are more tolerant to flooding or many environmental stresses.

The formation of aerenchyma takes place in the cortex, a ground tissue, and is divided into four developmental phases**. A microdissection (LM) laser protocol has been developed that has allowed us to construct tissue-specific RNA extraction for all stages of aerenchyma formation from extremely small amounts of biological material.** The LM technology was coupled to the digital droplet PCR and allowed us to analyze the tissue-specific expression of root genes, also from extremely small amounts of RNA. We then worked with Nugen to develop a protocol for the construction of RNAseq banks from low quantities of RNA. **This protocol was used to build RNAseq banks covering all stages of formation and differentiation of aerenchyma and adjacent tissues.** Analyses are ongoing and the regulatory networks will be rebuilt thanks to the collaboration with the I3Bionet team. The function of the most influential transcription factors will then be analyzed, first in protoplasts, then *in planta*.

In order to validate quickly and at high throughput in protoplasts, the genes potentially involved in the formation and differentiation of aerenchyma we have developed a technology derived from CRISPR/Cas9 allowing the construction of synthetic transcript factors (FTS). We were able to activate or repress a reporter gene and report the efficiency of the different designed FTs. The next step will be to demonstrate that we can simultaneously activate and suppress gene expression for rice genes and use this tool to validate at high throughput the genes potentially involved in the formation of aerenchyma.

Beyond their application to elucidating the molecular mechanisms of root aerenchyma formation, the tools developed in this work are sufficiently generic to be used for other species or for the study of other biological mechanisms at the molecular level.

Vers le développement d'une nouvelle approche pour étudier les mécanismes moléculaires impliqués dans la formation des aérenchymes racinaire chez le riz

Auteur: Thibault Mounier

Directeur de thèse: Christophe Périn, CIRAD

Résumé

Les aérenchymes sont des cavités qui se développent à partir des cellules du cortex de nombreuses plantes cultivées et jouent un rôle clef dans la tolérance à la submersion. Elles assurent la diffusion d'oxygène des parties aériennes vers les parties racinaires et font partie des mécanismes d'adaptation à la submersion chez le riz. Les aérenchymes ont également un rôle favorable dans de nombreuses conditions de cultures limitantes, comme les carences nutritionnelles ou les stress hydriques. Les gènes et mécanismes moléculaires qui interviennent dans sa formation sont pour la plupart inconnus. Leur connaissance est pourtant une étape indispensable pour mieux comprendre leur formation mais également pour permettre le développement de variétés plus tolérantes à la submersion ou à de nombreux stress environnementaux.

La formation des aérenchymes se déroule dans le cortex, un tissu racinaire interne et se décompose en quatre phases. **Un protocole de laser microdissection (LM) a été développé qui nous a permis de construire d'extraire de l'ARN tissus-spécifique pour toutes les étapes de formation des aérenchymes à partir de quantités extrêmement faibles de matériel biologique.** La technologie LM a été couplée à la droplet digital PCR et nous a permis d'analyser l'expression tissu-spécifique de gènes racinaires, à partir la aussi de quantités extrêmement faibles d'ARN. Nous avons alors construit avec la société Nugen un protocole pour la construction de banques RNAseq à partir de ces petites quantité d'ARN**. Ce protocole a été utilisé pour construire des banques RNAseq couvrant toutes les étapes de formation et de différenciation des aérenchymes et des tissus adjacents.** Les analyses sont en cours et les réseaux de régulation seront reconstruits grâce à une future collaboration avec l'équipe I3Bionet. Les fonctions des facteurs de transcription les plus influents seront alors analysées, dans un premier temps en protoplastes, puis *in planta*.

Afin de valider rapidement et à haut débit en protoplastes, les gènes potentiellement impliqués dans la formation et la différenciation des aérenchymes nous avons développé une technologie dérivée de CRISPR/Cas9 permettant la construction de facteurs de transcriptions synthétiques (FTS), les systèmes CRIPR activation (CRISPRa) et inhibition (CRISPRi). Nous avons pu démontrer qu'en système protoplaste nous étions capable d'activer ou de réprimer un gène rapporteur efficacement. La prochaine étape sera de démontrer que nous pouvons simultanément activer et réprimer l'expression de gènes chez le riz et utiliser cet outil pour valider à haut débit les gènes potentiellement impliqués dans la formation des aérenchymes.

Au-delà de leur application pour l'élucidation des mécanismes moléculaires de formation des aérenchymes racinaires, les outils développés au cours de ce travail sont suffisamment génériques pour être utilisés pour d'autres espèces ou pour l'étude d'autres mécanismes biologiques à l'échelle moléculaire.