



**HAL**  
open science

# Impact of AP2/ERF transcription factors on rice panicle structure evolution

Ai-My Luong

► **To cite this version:**

Ai-My Luong. Impact of AP2/ERF transcription factors on rice panicle structure evolution. Agricultural sciences. Université Montpellier, 2019. English. NNT : 2019MONTG035 . tel-03455119

**HAL Id: tel-03455119**

**<https://theses.hal.science/tel-03455119>**

Submitted on 29 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biologie du Développement

École doctorale n°584 GAIA : Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Unité de recherche DIADE : Diversité, Adaptation et Développement des Plantes

## IMPACT DES FACTEURS DE TRANSCRIPTION AP2/ERF SUR L'ÉVOLUTION DE LA STRUCTURE DU PANICULE DU RIZ

## IMPACT OF AP2/ERF TRANSCRIPTION FACTORS ON RICE PANICLE STRUCTURE EVOLUTION

Présentée par Ai-My LUONG

Le 25 Novembre 2019

Sous la direction de Dr. Stéphane JOUANNIC

Devant le jury composé de

Mme. Karine ALIX, Professeure à AgroParisTech

M. Bruno FAVERY, Directeur de Recherche INRA, Université Nice Sophia Antipolis

Mme. Veronica GREGIS, Associate Professor, Université de Milano

M. Michel LEBRUN, Professeur Université Montpellier

M. Stéphane JOUANNIC, Chercheur CR1, IRD Montpellier

Rapporteur

Rapporteur

Examineur

Examineur - Président du jury

Directeur de thèse



UNIVERSITÉ  
DE MONTPELLIER

## Abstract

Rice is a staple food for more than half of the world's population. The goal of increasing or sustainably maintaining rice production in a context of climate change and decreasing water and arable land availability requires the establishment of high-yield plants in contrasting environments. Rice yield is a complex trait, governed by genetic and epigenetic factors. It is directly dependent on 3 related traits: the number of panicles per plant, the number of grains per panicle and grain weight. During rice domestications in Asia and Africa, grain number was one of the main traits under selection, resulting in a phenotypic convergence between the crop plants that emerged on the two continents in relation to their yield potential. Panicle development in *Oryza sativa* has been well documented but the evolution of panicle architecture from the wild to the cultivated form remains poorly studied, especially with regard to the underlying molecular regulatory processes. To address this issue, two different transcriptomic resources were developed in the host laboratory. The first involved a comparison between the transcriptomes of different panicle meristem types in *O. sativa*. The second was a study of gene expression during panicle branch initiation in Asian and African cultivated species (*O. sativa* and *Oryza glaberrima*) with respect to their wild relatives (*Oryza rufipogon* and *Oryza barthii*, respectively). The two sets of genes thus identified displayed a significant enrichment in AP2/ERF family genes. Eighty four of the 170 AP2/ERF genes reported in *O. sativa* from the different sub-families (DREB, RAV, AP2 and Soloist) are expressed in the panicle and some are differentially expressed between the different types of meristems or stages of development and/or species. The euANT/PLT group within the AP2 subfamily comprises 4 genes (*OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22*) that are of particular interest in the context of the control of panicle development and architectural diversity between wild and cultivated species. A functional analysis of these 4 genes in the same genetic background was initiated by CRISPR-Cas9 approaches in *O. sativa* cv Kitaake. The results obtained revealed an effect of induced mutations on panicle architecture (number of primary and secondary branches) and/or on the size of the panicle (length of the branches and internodes).

**Keywords:** AP2/ERF, PLETHORA, panicle, meristem, domestication, rice

## Résumé

Le riz est un aliment de base pour plus de la moitié de la population mondiale. Une augmentation ou un maintien durable de la production de riz dans un contexte de changement climatique et de diminution des disponibilités en eau et en surface cultivable nécessite l'établissement de variétés au haut rendement sous différentes conditions environnementales. Le rendement en riz est un caractère complexe, se fondant sur des facteurs génétiques et épigénétiques. Il est directement dépendent de 3 caractères liés : le nombre de panicules par plante, le nombre de grains par panicule et le poids en grains. Durant les deux domestications (en Asie et en Afrique), le nombre de grains a été un des caractères principaux sous sélection, illustrant une convergence phénotypique en relation avec le potentiel de rendement. Le développement de la panicule chez *Oryza sativa* est bien documenté mais l'évolution de l'architecture de la panicule des espèces sauvages vers les espèces cultivées reste peu étudiée et surtout les bases moléculaires associées. Deux transcriptomes ont été développées dans le laboratoire d'accueil. Une 1<sup>ère</sup> correspond à l'étude comparative des transcriptomes des différents types méristématiques de la panicule chez *O. sativa*. Le second correspond à l'étude comparative des étapes de branchement de la panicule chez les espèces cultivées asiatique et africaine (*O. sativa* et *Oryza glaberrima*) avec les espèces sauvages apparentées (respectivement *Oryza rufipogon* et *Oryza barthii*). Parmi ces gènes d'intérêt, un enrichissement en gènes codant des facteurs de transcription de la famille AP2/ERF a pu être mis en évidence. 84 des 170 gènes de la famille AP2/ERF des différentes sous-familles (DREB, RAV, AP2 et Soloist) décrits chez *O. sativa* sont exprimés dans la panicule et certains sont différentiellement exprimés entre les différents types de méristèmes ou stades de développement et/ou d'espèces. Le groupe euANT/PLT de la sous-famille AP2 possède 4 gènes (*OsPLT7*, *OsPLT8*, *OsPLT9* et *AP2/EREBP22*) présentant un grand potentiel dans le contrôle du développement de la panicule et la diversité architecturale entre espèces sauvages et cultivées. Une analyse fonctionnelle de ces 4 gènes dans un même fond génétique a été initiée par des approches de CRISPR-Cas9 chez *O. sativa* cv Kitaake. Les mutations induites ont un impact sur l'architecture (nombre de branches primaires et/secondaires) et/ou la taille de la panicule (longueur des branches et des entre-nœuds) de manière différentielle.

**Mots clés :** AP2/ERF, PLETHORA, panicule, méristème, domestication, riz

## **Acknowledgments**

First and foremost, I would like to show my greatest appreciation to my supervisor, Dr. Stefan Jouannic who has supported me throughout my thesis with his patience and knowledge whilst allowing me work in my own way. I would like to thank Dr. James Tregear for his help and caring. His optimistic cheer me up from the first day I went to IRD. I would like to thank Dr. H  l  ne Adam for her constant encouragement, and expertise. Without their encouragement and guidance, this project would not be possible.

Also, I would like to take this opportunity to thank to the DIADE team (IRD Montpellier) and Joint Laboratory (LMI Rice - Vietnam) for providing me with a good environment and facilities to complete this project. I never forget their kindness and supports during the period of my project work. I want to thank Dr. Francois Sabot for his help in bio-informatic analysis. I want to thank Carole Gauron for her assistance and I wish all the best for her and her baby. I want to thank Sophie Cherron and Harold Chrestin for plant care in greenhouse. I also need to thank to Myriam Colin, a kind-hearted woman who not only give me the technical advices for histological work but also help me get acquainted with IRD laboratory from the beginning; Daniel Moukouanga for his enthusiasm when I was a newbie in the transfo zone. And my old colleague Otho Mantegazza, I have learned a lot from him. Besides, we would like to thank to Thi Nhu Le, Dr. Ngan Giang Khong for their enthusiasm and advice during the time I was in Vietnam. I also wish to express my gratitude to Prof. Pascal Gantet and Prof. Michel Lebrun for their supports and the opportunity they gave me to do my PhD for administrative and others in IRD and LMI. I am indebted to the IRD centre and French Embassy for providing research facilities and funding this work.

The guidance and supports received from all the members who contributed and who are contributing to this project, was important for the success of the project. I am grateful for their constant support and help.

I would like to express my gratitude towards all the friends, many of them I have come to know in Montpellier for less than 3 years, but their friendship and companionship give me a lot of energy. Many thanks to Tuan-Anh, Doaa, Louise, Trixie, C  cile, S  line, Nguyet, Ngan, Hieu, Chi, Ryan, Souhila, Thomas, Elvira, Oanh, Dung, Hoa and other friends. I am very happy for all the smiles and tears (and food during lunch time) we shared together.

Finally, yet importantly, I would like to express my heartfelt thanks to my beloved family for their understanding, endless love and supporting spiritually through of my life.

## ABBREVIATIONS

ANT	AINTEGUMENTA
AIL	AINTEGUMENTA-like
AM	axillary meristem
AP2/ERF	APETALA2/Ethylene Responsive Factor
ARF	AUXIN RESPONSE FACTOR
Aux/IAA	AUXIN/INDOLE-3-ACETIC ACID
CIM	Callus-inducing medium
Dex	dexamethasone
DMSO	Dimethylsulfoxide
ECR	Enoyl-CoA reductase
ePBM	elongated PBM
FM	floral meristem
GFP	GREEN FLUORESCENT PROTEIN
GUS	$\beta$ -glucuronidase
IAA	Indole-3-acetic acid
LR	Lateral root
LRP	Lateral root primordium
MS	Murashige and Skoog
NAA	Naphthalene-1-acetic acid
NPA	1-naphthylphthalamic acid
PBM	primary branch meristem
PLT	PLETHORA
PCR	Polymerase chain reaction
RM	rachis meristem
SM	spikelet meristem

# FRENCH SUMMARY

**Résumé français des travaux de thèse**

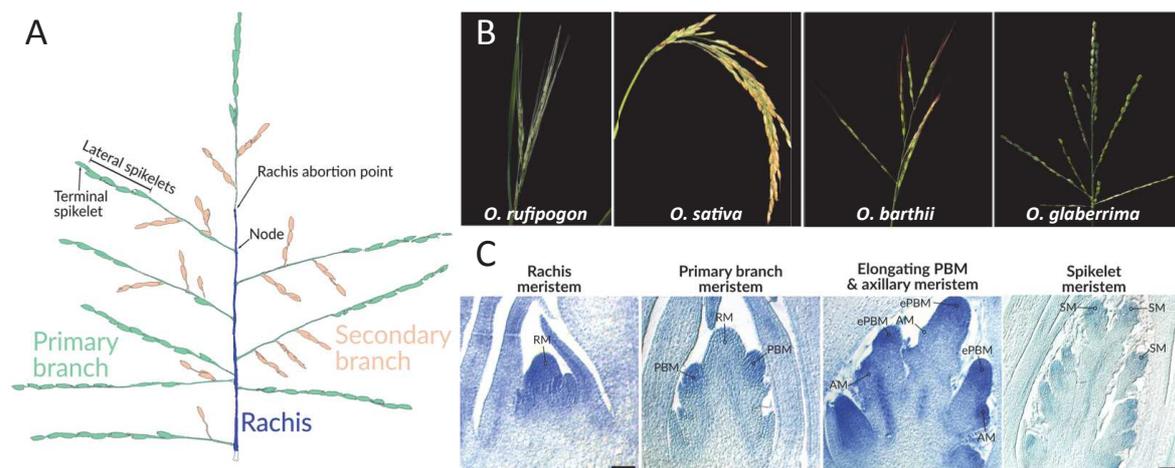
## Contexte général

Le développement agricole est essentiel pour assurer la production et la sécurité alimentaire pour une population croissante, en particulier dans les pays en développement. Le riz est un aliment de base pour plus de la moitié de la population mondiale, incluant un grand nombre de pays en développement. Une augmentation ou un maintien durable de la production de riz dans un contexte de changement climatique et de diminution des disponibilités en eau et en surface cultivable, nécessite la mise place de plantes à haut rendement dans des environnements contrastés. Depuis les années 1960 via la révolution verte, l'amélioration du riz a eu un énorme impact en se concentrant sur des caractères agronomiques qui affectent le potentiel de rendement : réduction de la taille des plantes, augmentation de la capacité de tallage, augmentation du nombre de grain par inflorescence (ou panicule), augmentation de la capacité photosynthétique, la capacité de réponse aux fertilisants (Khush 2005). L'établissement de variétés à haut rendement et capable de maintenir ses rendements dans différentes conditions environnementales reste un enjeu majeur des programmes d'amélioration moderne. Dans ce contexte un nouvel idiotype (New Plant Type, NPT) a été proposé avec une capacité réduite de tallage mais avec une plus grande efficacité de formation de panicules, un grand nombre de grains par panicule avec des plantes d'une taille de l'ordre d'1m avec un système végétatif plus robuste, un système racinaire adapté et un cycle de croissance de l'ordre de 100-130 jours (Peng *et al.* 2008). Ces caractères devraient permettre aux plantes de riz de transformer plus d'énergie dans la production de grains, augmentant le potentiel de rendement notamment chez les riz irrigués. Dans ce contexte d'amélioration, la compréhension des mécanismes moléculaires qui gouvernent le développement et les interactions entre le génome et 'environnement (GxE) qui déterminent les caractéristiques de la panicule comme sa complexité de branchement et sa plasticité reste très importante pour une amélioration durable du potentiel de rendement.

Le rendement en riz est un caractère complexe, se fondant sur des facteurs génétiques et épigénétiques. Il est directement dépendent de 3 caractères liés : le nombre de panicule par plante, le nombre de grain par panicule et le poids en grain (Xing et Zhang, 2010); et indirectement lié à d'autres caractères comme la taille de la plante, le temps de floraison, nombre de talles, le nombre de feuilles, la capacité photosynthétique, etc. (Adriana *et al.*, 2016; Huang *et al.*, 2011; Sakamoto *et al.*, 2006; Takai *et al.*, 2013; Ikeda *et al.*, 2013). Il est progressivement défini au cours du cycle de vie de la plante, d'abord au cours de la phase végétative, où le nombre de talles fertiles est établi, puis durant les phases de reproduction et

de remplissage du grain. Tandis que le poids en grain dépend de la taille des grains (longueur, largeur, épaisseur) et du degré de remplissage des grains, le nombre de panicule est dépendent des capacités de tallage. Par contre le nombre de grain par panicule est directement dépendent de l'architecture de la panicule : la complexité de branchement de la panicule impacte directement le nombre de grains par panicule. L'architecture de la panicule est constituée par une série de différents ordres de branchement : rachis, branches primaires, branches secondaires, potentiellement branches tertiaires et finalement les épillets (latéraux et terminaux) (Fig. 1a). Les épillets correspondent à des branches courtes portant une seule fleur (florete) chez le riz. Par conséquent le nombre d'épillets déterminera le nombre de grain par panicule (Ikeda *et al.*, 2004; Xing et Zhang, 2010).

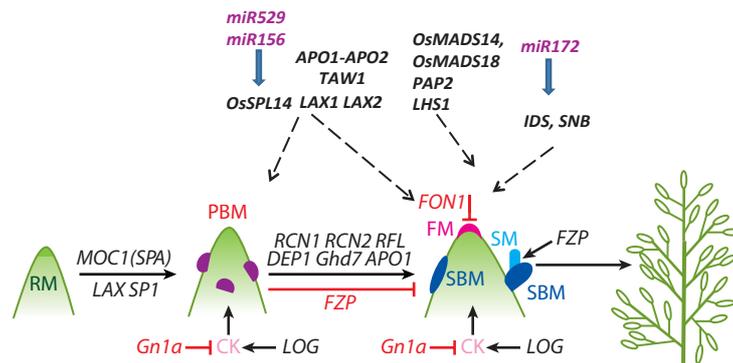
Une grande diversité inter- et intra-spécifique de complexité de l'architecture de la panicule est observée au sein du genre *Oryza* (Fig. 1b). Ce genre comprend deux espèces cultivées, *O. sativa* et *O. glaberrima*, domestiquées de façon indépendantes (respectivement en Asie et en Afrique) respectivement à partir des espèces sauvages *O. rufipogon* et *O. barthii*. Durant les deux domestications, le nombre de grains a été un des caractères principaux sous sélection, avec par conséquent une plus grande complexité de branchement des panicules chez les espèces domestiquées par rapport aux espèces sauvages, illustrant une convergence phénotypique en relation avec le potentiel de rendement. Une diversité intra-spécifique de branchement est également observée chez les espèces sauvages et cultivées des deux continents.



**Figure 1. La panicule de riz.** (a) structure de la panicule mature chez *O. glaberrima* illustrant les différents composants morphologiques de la panicule. (b) comparaison de panicules de riz chez les espèces cultivée asiatique et africaine (*O. sativa* et *O. glaberrima*) et les espèces sauvages apparentées (*O. rufipogon* et *O. barthii*). (c) structure des méristèmes aux stades précoces de développement de la panicule chez *O. sativa*. D'après Ta *et al.* 2017 et Harrop *et al.* 2016.

Cette diversité repose sur l'activité de méristèmes et de leur identité. Au cours du développement de la panicule, les méristèmes vont suivre différentes transitions d'identité (Fig. 1c). Après transition florale, le méristème végétatif se transforme en méristème reproducteur (méristème de rachis, RM) qui va contribuer à la mise en place de méristèmes axillaires (PBM, pour primary branch meristems). Ces derniers vont tous contribuer à la mise en place de branches primaires. Après allongement des PBM, de nouveaux méristèmes axillaires vont être mis en place. Ces derniers vont contribuer soit aux épillets latéraux soit aux branches secondaires qui elles-mêmes porteront des méristèmes axillaires contribuant aux épillets (voire à un ordre de branchement supérieur, i.e. branche tertiaire). Les méristèmes d'épillets se développent en méristèmes floraux qui vont produire les fleurs (ou floret chez les poaceae). Après la différenciation des méristèmes terminaux et axillaires en épillets, la complexité de branchement de la panicule est fixée. Ensuite, le rachis et les branches s'allongent rapidement pour former la panicule mature émergente porteuse de fleurs différenciées.

De nombreux travaux sur l'amélioration du rendement dans le but de créer de nouvelles variétés à haut rendement ont été effectués (Xing et Zhang, 2010; Ikeda *et al.*, 2013). Parmi ces travaux l'identification d'allèles bénéfiques à partir de QTL a eu une contribution importante. La cartographie et la caractérisation de QTL issus de populations biparentales chez *O. sativa* a permis d'identifier un grand nombre de gènes liés aux composants du rendement comme la complexité de branchement et la taille de la panicule (par exemple Gn1a, DEP1, IPA1/WFP), taille ou le poids des grains (comme GS3, GW2, qSW5/GW5) et le remplissage du grain (GIF1), la taille de la feuille paniculaire (qTSN), la capacité photosynthétique (GPS), etc. (Song *et al.*, 2007; Huang *et al.*, 2009; Miura *et al.*, 2010; Jiao *et al.*, 2010; Adriani *et al.*, 2016; Fujita *et al.*, 2013; Takai *et al.*, 2013). Quelques allèles spécifiques de ces gènes furent sélectionnés durant la domestication et/ou lors des programmes d'amélioration plus récents (Ikeda *et al.*, 2013). De plus l'analyse de mutants de développement chez *O. sativa* a permis d'identifier un panel de gènes nécessaires pour l'initiation et le développement de la panicule, tout comme des gènes contrôlant le nombre et la taille des grains et des panicules (Xing et Zhang, 2010 ; Wang et Li, 2011). Certains des gènes liés au développement de la panicule sont impliqués dans l'établissement et/ou le fonctionnement de méristèmes axillaires. D'autres sont impliqués dans le contrôle de la transition d'identité de méristème (méristème de branche vs. méristème d'épillet) (Wang et Li, 2011) (Fig. 2).



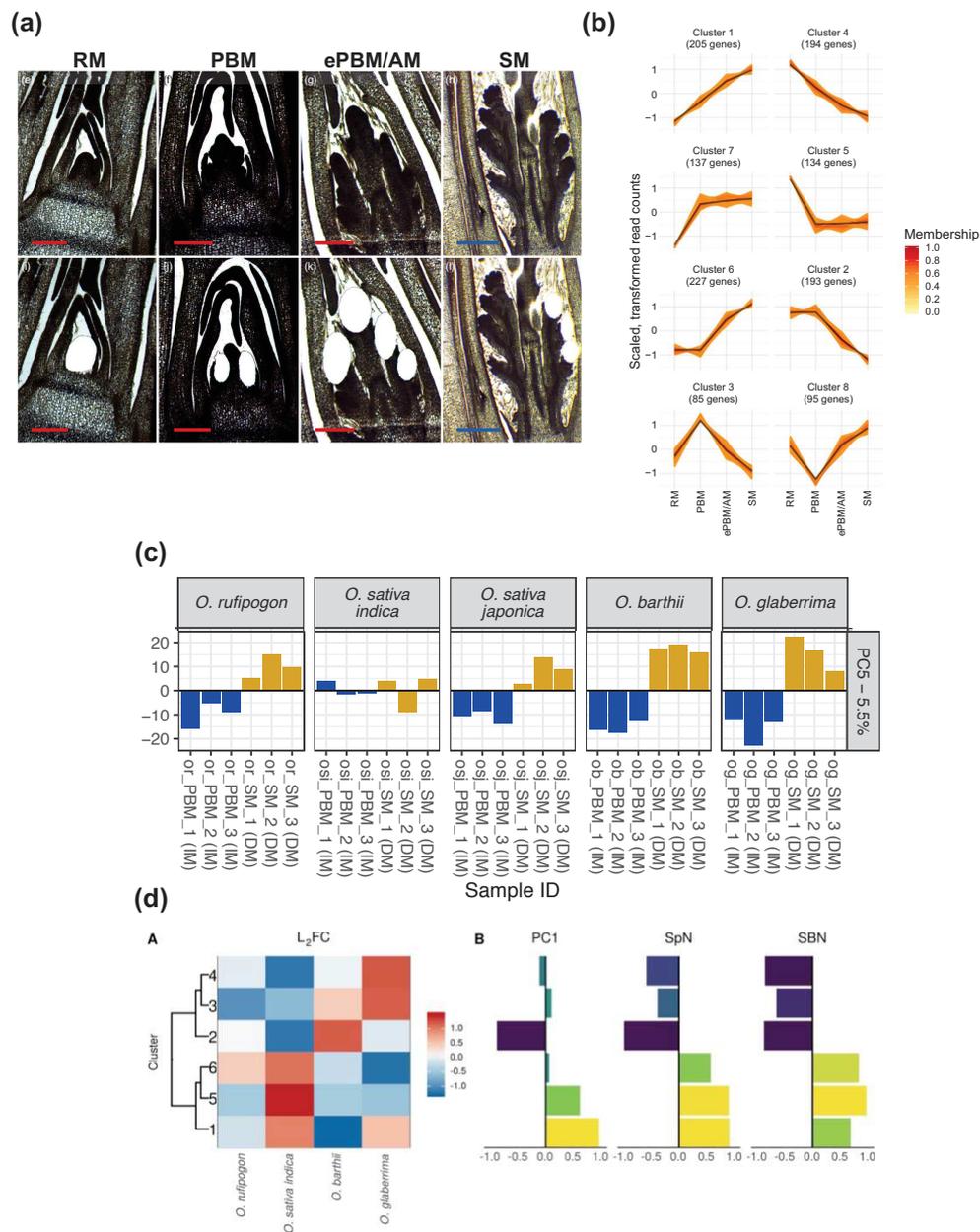
**Figure 2. Exemples de gènes connus pour leur implication dans le développement de la panicule chez *O. sativa*.** D'après Xing et Zhang 2010.

Le développement de la panicule chez *O. sativa* est bien documenté mais l'évolution de l'architecture de la panicule des espèces sauvages vers les espèces cultivées reste peu étudiée et surtout les bases moléculaires associées. Dans ce contexte, le projet de recherche de l'équipe d'accueil est de comprendre les bases cellulaires et moléculaires de la diversité d'architecture observée afin de mieux comprendre les processus évolutifs d'un trait morphologique et également de mieux appréhender les possibilités d'adaptation ou d'amélioration chez le riz en relation avec le potentiel de rendement. Pour cela différentes approches ont été développées sur la base d'une étude comparative des espèces sauvages et cultivées des continents africain et asiatique :

1. Analyse morphologique et cellulaire des panicules à maturité et au cours des phases précoces de développement
2. Approche génétique en association avec les études phénotypiques
3. Analyse de l'expression du génome (gènes candidats, transcriptomique).

Afin de comprendre les mécanismes moléculaires associés à la différenciation des méristèmes paniculaires et leur relation avec la diversité morphologique observée, des études d'expression du génome via des approches de RNA-seq en illumina ont été initiées. Pour cela, deux approches ont été développées. Une 1<sup>ère</sup> correspond à l'étude comparative des transcriptomes des différents types méristématiques de la panicule chez *O. sativa*. La seconde correspond à l'étude comparative des étapes de branchement de la panicule chez les espèces cultivées asiatique et africaine (*O. sativa* et *O. glaberrima*) avec les espèces sauvages apparentées (respectivement *O. rufipogon* et *O. barthii*).

Ces deux études ont permis de mettre en évidence un certains nombre de gènes différentiellement exprimés entre les différents types de méristèmes chez *O. sativa* et entre les espèces sauvages et cultivées aux stades de transition entre les états indéterminés et déterminés des méristèmes de panicules en lien avec la domestication (Fig. 3) (Harrop *et al.*, 2016; Harrop *et al.*, 2019). Parmi ces gènes d'intérêt, un enrichissement en gènes codant des facteurs de transcription de la famille AP2/ERF a pu être mis en évidence.



**Figure 3. Analyse transcriptomique sur méristèmes de panicules isolés par dissection au laser et sur des panicules aux stades de branchement et de différenciation des épillets chez 4 espèces de ri.** (a) coupes histologiques illustrant les différents types d'échantillons obtenus par dissection au laser. (b) analyse de clustering de co-expression de gènes différentiellement exprimés entre les différents types méristématique chez *O. sativa*. RM : méristème de rachis ; PBM : méristème de branche primaire ; ePBM/AM : méristèmes axillaires et de branche primaire ; SM : méristème

d'épillets. D'après Harrop *et al.* 2016. (c) Analyse en composantes principales des nombres de reads pour chaque banque pour l'analyse transcriptomique inter-spécifique. PBM : stade de branchement (IM : méristèmes indéterminés). SM : stade de différenciation des épillets (DM : méristèmes déterminés). Or : *O. rufipogon* ; Osi : *O. sativa indica* ; Osj : *O. sativa japonica* ; Ob : *O. barthii* ; Og : *O. glaberrima*. La composante principale PC5 sépare les 2 stades de développement par espèce et représente 5,4% de la variabilité totale. Les 4 premières composantes principales (non montrées) expliquent 51,7% de la variabilité totale et séparent les échantillons en fonction des espèces. (d) analyse de clustering de co-expression de gènes différentiellement exprimés entre les 2 stades de développement et la corrélation avec la composante principale PC1 (séparation des espèces sauvages et domestiquées), le nombre de branches secondaires (SBN) et le nombre d'épillets (SpN).

## Objectifs généraux

Dans ce contexte, mon projet de thèse a visé à analyser des gènes de la famille AP2/ERF à l'issue des analyses bio-informatique des données RNA-seq avec l'objectif de déterminer la contribution de certains de ces gènes dans le contrôle de l'architecture de la panicule. Le travail a été organisé pour répondre aux questions suivantes :

1. Y a-t-il une relation structure-fonction chez les gènes de la famille AP2/ERF exprimés dans la panicule et les différents types de méristèmes ? (Chapitre 1). Cette partie intègre une publication sur l'analyse bio-informatique de l'étude comparative de la transition indéterminé-déterminé chez les espèces sauvages et cultivées de riz asiatique et africain (Harrop *et al.* 2016)
2. Quelle est la structure et la diversité des gènes d'un clade particulier de la famille AP2/ERF chez les différentes espèces de riz : les gènes *AINTEGUMENTA-LIKE/PLETHORA* (*euANT/PLT*) ? (Chapitre 2)
3. Quelle est la contribution des gènes *euANT/PLT* dans l'établissement de la structure de la panicule chez *O. sativa* ? (Chapitre 3)

## Résultats

Les résultats organisés en 3 parties afin de répondre aux questions précédentes sont principalement rédigés sous forme d'article scientifique.

### **Chapitre 1 : Analyse phylogénétique des gènes de la famille AP2/ERF exprimés dans la panicule.**

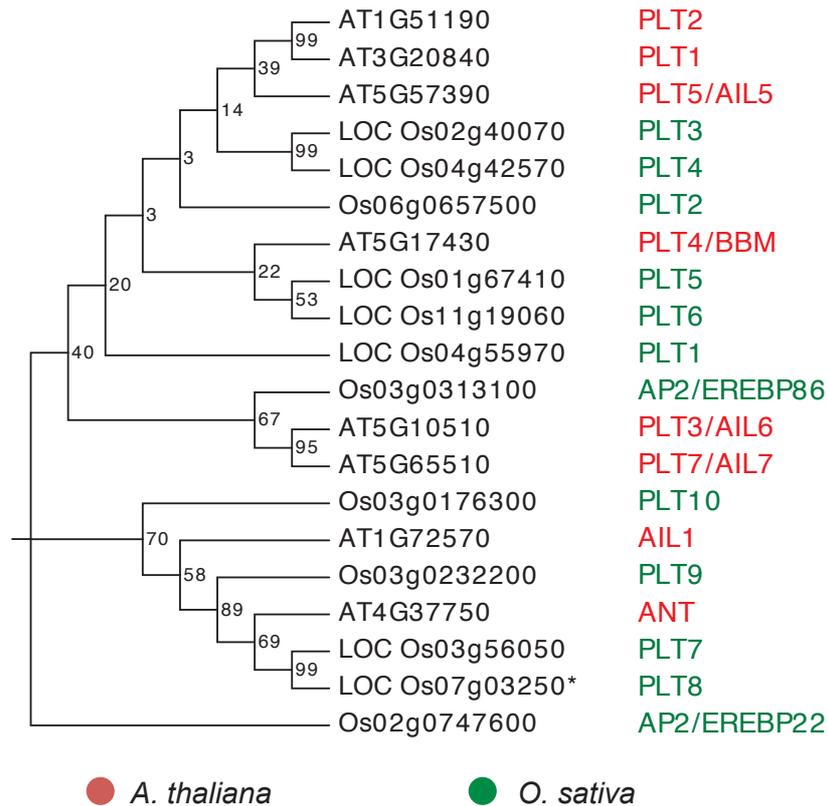
Cette partie intègre les données de RNA-seq obtenues pour l'étude comparative de la transition indéterminé-déterminé chez les espèces sauvages et cultivées de riz asiatique et africain qui ont fait le sujet d'une publication (Harrop *et al.*, 2019). Ces travaux ont permis de

mettre en évidence deux ensembles de gènes sur la base de leurs comportements d'expression : *i.* un ensemble de gènes avec un profil d'expression conservés entre les 4 espèces pouvant constitué un « core-set » de gènes liés à la transition indéterminé-déterminé, *ii.* un ensemble de gènes avec des profils d'expression différents entre les espèces et corrélés avec la domestication. Ces deux ensembles de gènes se caractérisent par un enrichissement significatif en gènes de la famille AP2/ERF.

J'ai complété ces travaux par une analyse phylogénétique exhaustive des gènes de la famille AP2/ERF détectés dans les 2 transcriptomes. Les analyses montrent que 84 des 170 gènes de la famille AP2/ERF chez le riz sont exprimés dans la panicule et que certains sont différentiellement exprimés entre les différents types de méristèmes ou stades de développement et/ou d'espèces. Notamment, des différences d'expression en lien avec la domestication peuvent exister avec un différentiel d'expression entre les espèces sauvages et cultivées. Cette étude a permis de montrer l'absence de relation structure-fonction quant aux transcrits détectés, les gènes détectés au sein de la panicule étant issus des différentes sous-familles (DREB, RAV, AP2 et Solist) indépendamment de leurs profils d'expression. Par ailleurs, un groupe de gènes de la sous-famille AP2 a particulièrement attiré notre attention de par la diversité de profils d'expression entre les différents types de méristèmes paniculaires et selon les espèces de riz : les gènes de la famille *AINTEGUMENTA-LIKE/PLETHORA* (*euANT/PLT*).

## **Chapitre 2 : Analyses *in silico* et d'expression durant le développement de la panicule des gènes *euANT/PLT* du riz.**

Les gènes *AINTEGUMENTA/PLETHORA* (*euANT/PLT*) sont connus notamment chez *A. thaliana* pour leur implication dans de nombreux processus développementaux, dont la maintenance et le fonctionnement des méristèmes (racinaire, végétatif et reproducteur), *via* le contrôle de la prolifération et de la différenciation cellulaire (Scheres et Krizek, 2018; Horstman *et al.*, 2014). Cependant, le rôle des gènes *euANT/PLT* dans le développement de la panicule de riz reste mal connu. Cette partie présente l'identification et l'analyse complète des gènes *euANT/PLT* chez les différentes espèces de riz (structure génique, expression, séquence promotrice), qui pourraient fournir des informations sur l'élucidation de leurs fonctions biologiques (Fig. 4).



**Figure 4. Arbre relationnel des gènes euANT/PLT chez *O. sativa* et *A. thaliana*.** Arbre de distance obtenu sur la base de séquences de 1<sup>er</sup> domaine AP2. Les valeurs de bootstrap sont indiquées aux nœuds correspondants

Ce travail a permis de mettre en évidence 12 gènes du groupe *euANT/PLT* chez *O. sativa* très conservés par rapport aux gènes d'*A. thaliana*, avec deux sous-clades : les gènes proches d'*AIL6* et *BBM* et les gènes proches de *ANT*. Néanmoins les annotations des gènes *euANT/PLT* sont très variables (structure et longueur) entre les deux grandes bases de données génomique RAPDB et MSU. En conséquence la validation expérimentale de la structure de ces gènes est primordiale avant toute analyse fonctionnelle. Seuls les gènes *OsPLT7*, *OsPLT8*, *OsPLT9* et *AP2/EREBP22*, proches du gène *ANT* d'*A. thaliana*, sont détectés dans les deux jeux de données de transcriptomiques et présentent des profils d'expression différents entre les différents types de méristèmes paniculaires chez *O. sativa*. La structure de ces gènes est conservée chez les 4 espèces de riz sauvage et cultivés asiatique et africain. Les profils d'expression de ces gènes au cours du développement de la panicule ont été analysés par qRT-PCR à haut débit (Fluidigm®) chez les espèces sauvages et cultivées asiatiques et africaines. et ont permis de confirmer les profils d'expression de ces gènes. Les gènes *AP2\_EREBP22*, *OsPLT7*, *OsPLT8* gènes se caractérisent par un différentiel d'expression entre les espèces sauvages et cultivées, avec un pic d'expression dans le

méristème de rachis pour *OsPLT7* et *OsPLT8*. Le gène *OsPLT9* se caractérise par une expression plus forte dans les méristèmes axillaires sans différence significative entre les espèces. En conséquence ces 4 gènes possèdent un potentiel fort quant à leur implication dans le contrôle du développement de la panicule et de la diversité d'architecture entre les espèces sauvage et cultivées. Les analyses des séquences promotrices de ces gènes chez les 4 espèces sauvages et cultivées ont montré la présence d'un grand nombre de motifs de réponse aux hormones (auxine, cytokinine, éthylène, gibbérellines) différents d'un paralogue à l'autre chez une même espèce. De plus on observe un polymorphisme (SNP, INDEL) entre les espèces sauvages et cultivées, la divergence étant moins marquée pour les espèces africaines. Ces analyses suggèrent que l'expression différentielle de ces gènes entre les espèces pourrait être liée à des modifications de la régulation en *cis* au niveau des régions promotrices.

### **Chapitre 3 : Impact fonctionnel des gènes *euANT/PLT* sur l'architecture paniculaire chez *O. sativa***

Un mutant de perte de fonction pour le gène *OsPLT8* avait été caractérisé au préalable : le mutant *crown rootless5 (crl5)* qui se caractérise par un défaut de développement des racines coronaires à la base de la tige foliaire (Kitomi *et al.*, 2011). Le contrôle du développement des racines coronaires par le gène *OsPLT8/CRL5* est dépendante de la voie de signalisation de l'auxine en amont et implique une répression de la voie des cytokinines en aval (Kitomi *et al.*, 2011). J'ai caractérisé l'architecture paniculaire chez ce mutant. Ces panicules se caractérisent par une réduction du nombre de branches primaires avec une augmentation du nombre de branches secondaires par branche primaire. Ceci indique que le gène *OsPT8/CRL5* pourrait aussi être impliqué dans le fonctionnement des méristèmes paniculaires et que les 4 gènes *PLT* constituent des candidats intéressants quant au développement de la panicule.

Afin de déterminer l'impact de ces gènes sur l'architecture paniculaire, une analyse fonctionnelle de ces 4 gènes dans un même fond génétique a été initiée par des approches de CRISPR-Cas9 chez *O. sativa* cv *Kitaake*. Les ARN guides ont été élaborés afin d'induire des délétions dans chacun de ces gènes et également pour induire une double mutation les gènes *OsPL7* et *OsPLT8*, deux proches paralogues. Cependant les plantes transgéniques porteuses de la construction ciblant le gène *OsPLT7* et celles ciblant simultanément les gènes *OsPLT7* and *OsPLT8* n'ont pu être analysées suite à des problèmes de génotypage. Des lignées T2 homozygotes sans T-DNA représentant différents allèles édités avec délétions ont pu être obtenues pour les gènes *OsPLT8*, *OsPLT9* et *AP2/EREBP22*. La caractérisation du phénotype

paniculaire de ces différentes lignées indiquent que les mutations induites ont un impact sur l'architecture (nombre de branches primaires et/secondaires) et/ou la taille de la panicule (longueur des branches et des entre-nœuds) de manière différentielle. D'autres impacts ont pu être mesurés affectant le temps de floraison, l'architecture végétative et racinaire. Ceci suggère des fonctions partiellement redondantes entre ces différents gènes dans le contrôle du développement de la panicule. Par ailleurs une analyse d'expression *in situ* a pu être conduite pour le gène *OsPLT9* indiquant une expression dans les méristèmes axillaires et les tissus vasculaires de la panicule ainsi que dans les méristèmes axillaires végétatifs à l'origine du tallage. Ces profils spatiaux d'expression sont en adéquation avec les phénotypes observés pour les mutants édités *plt9*.

## Discussion et perspectives

L'objectif global de cette thèse était d'identifier et de caractériser les gènes AP2/ERF liés au développement de la panicule et à sa diversité dans le genre *Oryza*, puis de déterminer si ces gènes étaient affectés par les deux domestications indépendantes au sein de ce genre. La plupart des travaux portaient sur les facteurs de transcription euANT/PLT, un petit groupe de la superfamille AP2/ERF.

Vingt années de recherche sur les gènes *euANT/PLT* chez *A. thaliana* ont montré que ces facteurs de transcription sont essentiels pour la spécification de la niche des cellules souches, le maintien du méristème, l'initiation des organes et la croissance (Horstman *et al.*, 2014). Néanmoins très peu de données étaient disponibles quant à la fonction des 12 gènes *euANT/PLT* chez le riz et notamment au cours du développement paniculaire, à l'exception du gène *CRL5/PLT8* connu pour son implication dans le développement des racines coronaires (Kitomi *et al.*, 2011). Plus récemment, Khanday *et al.* (2019) ont pu démontrer, dans le cadre de l'étude de mutants embryonnaires, que l'expression de *OsPLT6/BBM1* au sein du génome mâle déclenche le programme embryonnaire dans la cellule œuf fécondée. De plus, les gènes de type BBM favorisent la régénération à partir de la culture de tissus, suggérant qu'ils fonctionnent comme des facteurs de pluripotence (Lowe *et al.*, 2016). Par ailleurs, le profil d'expression des gènes *OsPLT1-OsPLT6* dans la racine de riz indique qu'ils sont tous exprimés dans le primodium de la racine coronaire et principalement dans les cellules initiales adjacentes au contrôle de la qualité des racines primaires, coronaires et latérales (Li et Xue, 2011).

J'ai contribué à montrer que les gènes AP2/ERF étaient des facteurs clés dans la

ramification de l'inflorescence que la domestication du riz serait associée à une altération de l'expression de certains de ces gènes. De plus, j'ai effectué une analyse détaillée de l'expression des gènes *euANT/PLT* du riz chez différentes espèces de riz, pour l'élucidation de leurs fonctions biologiques. Six gènes *euANT/PLT* (*OsPLT4*, *OsPLT7*, *OsPLT8*, *OsPLT9*, *OsPLT10*, *AP2/EREBP22* et *AP2/EREBP86*) sont exprimés dans des méristèmes de panicules de riz. Néanmoins, seuls les gènes *OsPLT7*, *OsPLT8*, *OsPLT9* et *AP2/EREBP22* sont exprimés de manière différenciée à travers les différents stades de développement de la panicule et entre les différentes espèces de riz, ce qui implique que ces gènes pourraient jouer un rôle important dans le fonctionnement des méristèmes paniculaires et revêtent une importance dans le contexte de la domestication. Enfin, j'ai généré des mutants édités par CRISPR-Cas9 pour les gènes *OsPLT7*, *OsPLT8*, *OsPLT9* et *AP2/EREBP22* afin d'étudier leur impact sur l'architecture et le développement de la panicule de riz. Par le phénotypage des caractères paniculaires des mutants, j'ai pu rassembler un grand nombre de données appuyant l'hypothèse d'une implication de ces gènes dans l'établissement de l'architecture paniculaire. Concrètement, les mutants *plt8* produisent moins de branches secondaires alors que les mutants *plt9* produisent plus de branches secondaires et que les mutants *ap22/erebp22* produisent davantage de branches primaires et secondaires.

Cependant, plusieurs questions sur le rôle de ces gènes dans le développement de la panicule restent à traiter. Le premier point est que notre étude doit être complétée par d'autres analyses génétiques et fonctionnelles, afin de déterminer si cette altération de l'expression de ces gènes entraîne une différence d'architecture paniculaire observée à différents stades et entre les différentes espèces de riz. De plus, il sera important de déterminer si d'autres facteurs (hormones, nouveaux gènes, etc.) interagissent avec ces gènes pour leur fonction au cours du développement de la panicule du riz.

Plusieurs perspectives au cours ou long terme peuvent découler de ces travaux. Tout d'abord, il faudra continuer la caractérisation des lignées mutantes *plt7* et *plt7plt8* qui n'ont pu être analysées dans la cadre de cette thèse. Deuxièmement, une analyse de l'expression des gènes *euANT/PLT* dans les fonds génétiques mutants pourront être effectués afin de déterminer si ce sont mutants KO ou non et s'il existe des interactions au niveau transcriptionnel entre ces gènes. De plus une analyse des profils d'expression de gènes marqueurs du développement de la panicule dans ces fonds génétiques mutants pourra donner des indications quant à l'interaction potentielle entre ces différents gènes.

Des analyses plus fines de ces lignées éditées seront nécessaires quant au phénotype paniculaire aux stades précoces de développement durant la phase de ramification (histologie,

taille des méristèmes, etc.), aux profils d'expression *in situ* de *OsPLT8*, *OsAP2/EREBP22*, *OsPLT7* et d'autres gènes marqueurs du développement, ainsi qu'au phénotype racinaire et floral. Ces analyses pourront fournir plus d'informations sur les rôles de ces gènes *euANT/PLT* dans le développement du riz. En parallèle, différentes lignées de sur-expression des gènes *OsPLT8*, *OsPLT9* et *AP2/EREBP22* sont en cours de sélection et leurs phénotypes seront prochainement caractérisés afin de compléter l'analyse fonctionnelle de ces gènes.

A plus long terme, il sera important d'étudier l'interaction de ces gènes avec d'autres facteurs notamment les hormones comme l'auxine et les cytokinines, connues pour leur rôle dans l'architecture de la panicule (Yang *et al.*, 2017 ; Ashikari *et al.*, 2005). Par ailleurs, Kitomi *et al.* (2011) ont montré que le développement des racines coronaires du riz est dépendant de l'activation du gène *OsPLT8/CRL5* par l'auxine qui entraîne une répression de la voie de signalisation des cytokinines en activant un gène de type ARR, *OsRRI*.

L'analyse des profils d'expression *in situ* des gènes *OsPLT7*, *OsPLT8*, *OsPLT9*, *OsAP2/EREBP22* dans les autres espèces de riz étudiés pourra être effectuée afin de déterminer le niveau de conservation des domaines d'expression. Si la transformation génétique d'*O. glaberrima* est optimisée avec succès et efficacité, des mutants similaires pourraient être créés chez cette espèce, ce qui permet de comparer les effets de ces gènes sur le phénotype paniculaire entre les riz domestiqués asiatique et africain.

Dans l'objectif de rechercher des liens possibles entre la diversité intra-spécifique de l'architecture paniculaire et les gènes *euANT/PLT*, une analyse des profils d'expression et de la diversité génétique de ces gènes pourrait être réalisée à l'aide du panel ou d'un sous-panel de variétés vietnamiennes d'*O. sativa* développées par le LMI RICE et AGI (Phung *et al.*, 2014). Ce panel a permis d'identifier par génétique d'association de nouveaux QTL liés à la diversité de l'architecture de la panicule (Ta *et al.*, 2018).

## REFERENCES

- Adriani DE, Dingkuhn M, Dardou A, Adam H, Luquet D, Lafarge T. Rice panicle plasticity in Near Isogenic Lines carrying a QTL for larger panicle is genotype and environment dependent. *Rice* 2016, 9:28.
- Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M. Cytokinin oxidase regulates rice grain production. *Science* 2005, 309:741–745.
- Fujita D, Trijatmiko KR, Tagle AG, Sapasap MV, Koide Y, Sasaki K, Tsakirpaloglou N, Gannaban RB, Nishimura T, Yanagihara S, Fukuta Y, Koshiba T, Slamet-Loedin IH, Ishimaru T, Kobayashi N. *NALI* allele from a rice landrace greatly increases yield in modern indica cultivars. *Proceedings of the National Academy of Sciences USA*. 2013 110:20431-20436
- Harrop TWR, Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, Kater MM. Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *The Plant Journal* 2016, 86:75-88.
- Harrop TWR, Mantegazza O, Luong AM, Béthune K, Lorieux M, Jouannic S, Adam H. A set of AP2-like genes is associated with inflorescence branching and architecture in domesticated rice. *Journal of Experimental*

- Botany 2019, doi:10.1093/jxb/erz340.
- Horstman A, Willemsen V, Boutilier K, Heidstra R. AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends in Plant Science* 2014, 19:146-157
- Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, *et al.* Natural variation at the DEP1 locus enhances grain yield in rice. *Nature Genetics* 2009, 41:494-497
- Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C, Fan D, Lu Y, Weng Q, Liu K, Zhou T, Jing Y, Si L, Dong G, Huang T, Lu T, Feng Q, Qian Q, Li J, Han B. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics* 2011, 44:32-39
- Ikeda K, Sunohara H, Nagato Y. Developmental Course of Inflorescence and Spikelet in Rice. *Breeding Science* 2004, 54:147-156
- Ikeda M, Miura K, Aya K, Kitano H, Matsuoka M. Genes offering the potential for designing yield-related traits in rice. *Current Opinion in Plant Biology* 2013, 16:213-220
- Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, *et al.* Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nature Genetics* 2010, 42:541-544
- Khanday I, Skinner D, Yang B, Mercier R, Sundaresan V. (2019). A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature* 2019, 565:91–95
- Khush GS. What it will take to Feed 5.0 Billion Rice consumers in 2030. *Plant Molecular Biology* 2005, 59:1–6
- Kitomi Y, Ito H, Hobo T, Aya K, Kitano H, Inukai Y. The auxin responsive AP2/ERF transcription factor CROWN ROOTLESS5 is involved in crown root initiation in rice through the induction of *OsRR1*, a type-A response regulator of cytokinin signaling. *The Plant Journal* 2011, 67:472–484
- Miura K, Ikeda M, Matsubara A, Song X-J, Ito M, Asano K, *et al.* OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nature Genetics* 2010, 42:545–549
- Peng S, Khush GS, Virk P, Tang Q, Zou Y. Progress in ideotype breeding to increase rice yield potential. *Field Crops Research* 2008, 108:32–38
- Phung NT, Mai CD, Mournet P, Frouin J, Droc G, Ta NK, Jouannic S, Lê LT, Do VN, Gantet P, Courtois B. Characterization of a panel of Vietnamese rice varieties using DArT and SNP markers for association mapping purposes. *BMC Plant Biology* 2014 19: 371
- Sakamoto T, Morinaka Y, Ohnishi T, Sunohara H, Fujioka S, Ueguchi-Tanaka M, Mizutani M, Sakata K, Takatsuto S, Yoshida S, Tanaka H, Kitano H, Matsuoka M. Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. *Nature Biotechnology* 2006, 24:105-109
- Scheres B, Krizek BA. Coordination of growth in root and shoot apices by AIL/PLT transcription factors. *Current Opinion in Plant Biology* 2018, 41:95–101
- Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genetics* 2007, 39:623-630
- Ta KN, Khong NG, Ha LT, Nguyen DT, Mai DC, Hoang TG, Phung TPN, Bourrie I, Courtois B, Tran TTH, Dinh BY, La TL, Do NV, Lebrun M, Gantet P, Jouannic S. A genome-wide association study using a Vietnamese landrace panel of rice (*Oryza sativa*) reveals new QTLs controlling panicle morphological traits. *BMC Plant Biology* 2018, 18:282.
- Takai T, Adachi S, Taguchi-Shiobara F, Sanoh-Arai Y, Iwasawa N, Yoshinaga S, *et al.* A natural variant of NAL1, selected in high-yield rice breeding programs, pleiotropically increases photosynthesis rate. *Scientific Reports* 2013, 3:731-711
- Wang Y, Li J. Branching in rice. *Current Opinion in Plant Biology* 2011, 14:94-99
- Xing Y, Zhang Q. Genetic and Molecular Bases of Rice Yield. *Annual Review of Plant Biology* 2010, 61:421-442
- Yang J, Yuan Z, Meng Q, Huang G, Périn C, Bureau C, Meunier AC, Ingouff M, Bennett MJ, Liang W, Zhang D. Dynamic Regulation of Auxin Response during Rice Development Revealed by Newly Established Hormone Biosensor Markers. *Frontiers in Plant Science* 2017, 8:9–17

## TABLE OF CONTENTS

ENGLISH ABSTRACT .....	
FRENCH ABSTRACT.....	
FRENCH SUMMARY.....	
<b>GENERAL INTRODUCTION.....</b>	
1. THE EVOLUTION OF PLANT KINGDOM .....	1
1.1. THE FUNCTIONAL DIVERGENCE OF DUPLICATED GENES.....	2
1.2. THE EXPRESSION PATTERN DIVERGENCE OF CONSERVED GENES.....	3
1.3. “ <i>DE NOVO</i> ” FORMATION OF NEW CODING GENES.....	5
1.4. DOMESTICATION.....	7
2. INFLORESCENCES AND MERISTEMS .....	13
2.1. MERISTEM FUNCTIONING.....	13
2.2. INFLORESCENCE ARCHITECTURES .....	18
3. PANICLE DEVELOPMENT IN RICE .....	29
3.1. THE IMPORTANCE OF RICE CROPS.....	29
3.2. RICE DOMESTICATION .....	30
3.3. RICE PANICLE ARCHITECTURE.....	33
3.4. THE DEVELOPMENT OF THE RICE PANICLE.....	35
3.5. GENETIC CONTROL OF TILLERING .....	39
3.6. AXILLARY MERISTEM INITIATION DURING INFLORESCENCE DEVELOPMENT.....	40
3.7. AXILLARY MERISTEM OUTGROWTH DURING INFLORESCENCE DEVELOPMENT.....	42
3.8. THE TIMING OF SPIKELET DIFFERENTIATION .....	45
3.9. FLORAL ORGAN PATTERNING .....	46
4. TRANSCRIPTION FACTORS IN THE PLANT KINGDOM.....	50
4.1. STRUCTURE OF TRANSCRIPTION FACTORS.....	51
4.2. DNA RECOGNITION BY TRANSCRIPTION FACTORS .....	52
4.3. DNA-BINDING DOMAINS .....	52
4.4. REGULATION OF TRANSCRIPTION FACTOR FUNCTION .....	54
4.5. DEFINITION, CLASSIFICATION, STRUCTURE OF AP2/ERF TRANSCRIPTION FACTORS.....	55

4.6. AP2/ERF TRANSCRIPTION FACTORS IN ARABIDOPSIS AND RICE .....	57
5. OBJECTIVES .....	63
<b>CHAPTER I</b> .....	
GENERAL INTRODUCTION .....	65
RESEARCH PAPER .....	66
GENERAL CONCLUSIONS AND PERSPECTIVES .....	88
REFERENCES .....	96
SUPPLEMENTAL FIGURES AND TABLES .....	98
<b>CHAPTER II</b> .....	
INTRODUCTION .....	101
MATERIALS AND METHODS .....	103
Plant materials and growth conditions .....	103
<i>In silico</i> identification of <i>euANT/PLT</i> genes in <i>Arabidopsis</i> and various rice genomes	103
<i>OsPLT</i> gene sequence validation .....	104
RNA extraction and Fluidigm® high throughput qPCR .....	104
RESULTS .....	105
The <i>euANT/PLT</i> group in <i>O. sativa</i> .....	105
Gene and protein structure of <i>euANT/PLT</i> group members in <i>O. sativa</i> .....	107
<i>O. sativa</i> and <i>A. thaliana</i> <i>euANT/PLT</i> genes relationship .....	113
Orthologs of <i>euANT/PLT</i> genes in other rice species .....	115
<i>In silico</i> expression profiling analysis of the <i>euANT/PLT</i> group in <i>O. sativa japonica</i> .	120
Expression patterns in different rice species of <i>euANT/PLT</i> genes preferentially	
expressed in the panicle .....	120
<i>Cis</i> -element analysis of promoter regions of <i>OsPLT</i> genes .....	124
DISCUSSION .....	127
<i>euANT/PLT</i> gene sequences and annotations show divergences between the two common	
rice databases .....	127
<i>OsPLT</i> proteins are highly conserved between <i>Arabidopsis</i> and different rice species .	127
<i>OsPLTs</i> gene acting in hormone signaling pathways .....	130
<i>OsPLT7</i> , <i>OsPLT8</i> , <i>OsPLT9</i> and <i>AP2/EREBP22</i> are differentially expressed during the	
early stages of panicle development in rice .....	131
ACKNOWLEDGEMENTS .....	132

FUNDING .....	132
REFERENCES.....	133
SUPPLEMENTAL FIGURES AND TABLES .....	137
<b>CHAPTER III</b> .....	
INTRODUCTION.....	145
MATERIALS AND METHODS .....	147
Plant materials and growth condition .....	147
Generation and selection of CRISPR-Cas9 edited plants .....	147
Plant phenotyping .....	148
RNA <i>in situ</i> hybridization .....	148
RESULTS .....	149
Panicle phenotyping of AP2/ERF-related mutants.....	149
Generation of CRISPR-Cas9 edited mutants of rice <i>ANT</i> -related genes .....	150
The <i>plt8</i> mutants produce fewer secondary branches.....	155
Larger numbers of secondary branches produced in the <i>plt9</i> mutant lines.....	158
The <i>ap22/erebp22</i> mutants produce more primary and secondary branches.....	158
Comparison between the sister lines and the natural wild-type plants .....	161
Spatial expression profiling of <i>OsPLT</i> genes in early stages of panicle development...	163
DISCUSSION .....	163
Second AP2 domain mutations affect the function of <i>OsPLT</i> candidate genes.....	163
<i>OsPLT</i> candidate genes impact on the rice panicle architecture .....	164
CONCLUSIONS.....	167
REFERENCES.....	169
SUPPLEMENTAL FIGURES AND TABLES .....	171
<b>Chapter IV</b> .....	
<i>PLETHORA</i> genes control meristem development.....	179
<i>OsPLT7</i> , <i>OsPLT8</i> , <i>OsPLT9</i> and <i>AP2/EREBP22</i> may have played a role in panicle architecture evolution .....	181
The short/mid-term perspectives.....	182
The long-term perspectives.....	183
REFERENCES.....	185

<b>CHAPTER V</b> .....	
1. Chemicals and kits .....	187
2. Plant materials and growth condition.....	187
3. <i>In Silico</i> analysis of <i>AP2/ERF</i> family.....	188
3.1. Identification of <i>AP2/ERF</i> genes in rice genome.....	188
3.2. Phylogenetic and MEME motif analysis .....	188
3.3. <i>In silico</i> gene expression analysis .....	188
3.4. <i>In silico</i> identification of <i>AP2</i> sub-family genes in African and Asian wild and domesticated rice and <i>Arabidopsis</i> genomes.....	189
4. Primer Design .....	189
5. <i>AP2/ERF</i> gene expression at inflorescence branching in wild and domesticated rice.....	190
5.1. Tissue collection and RNA sequencing.....	190
5.2. Fluidigm high throughput qPCR.....	191
6. Generating the transgenic lines of <i>PLETHORA</i> genes.....	191
6.1. Generating Knock-Out mutants of <i>PLETHORA</i> genes by using CRISPR/Cas9 system.....	191
6.2. Generating the Over-Expressing mutants of <i>PLETHORA</i> genes.....	197
6.3. <i>Agrobacterium tumefaciens</i> -mediated rice transformation .....	198
7. Genotyping and Phenotyping Knock-Out transgenic lines .....	200
7.1. Genomic DNA isolation .....	200
7.2. Genotyping of genome-edited progeny .....	200
7.3. Plant phenotyping.....	200
8. <i>In situ</i> hybridization.....	201
8.1. Preparation of sense and antisense RNA probes.....	201
8.2. Fixation of tissues .....	202
8.3. Impregnation in paraplast .....	202
8.4. Probe hybridization and labelling .....	203
9. MEDIAS, SOLUTIONS AND BUFFERS.....	204
SUPPLEMENTAL FIGURES AND TABLES .....	206
GENERAL REFERENCES.....	212

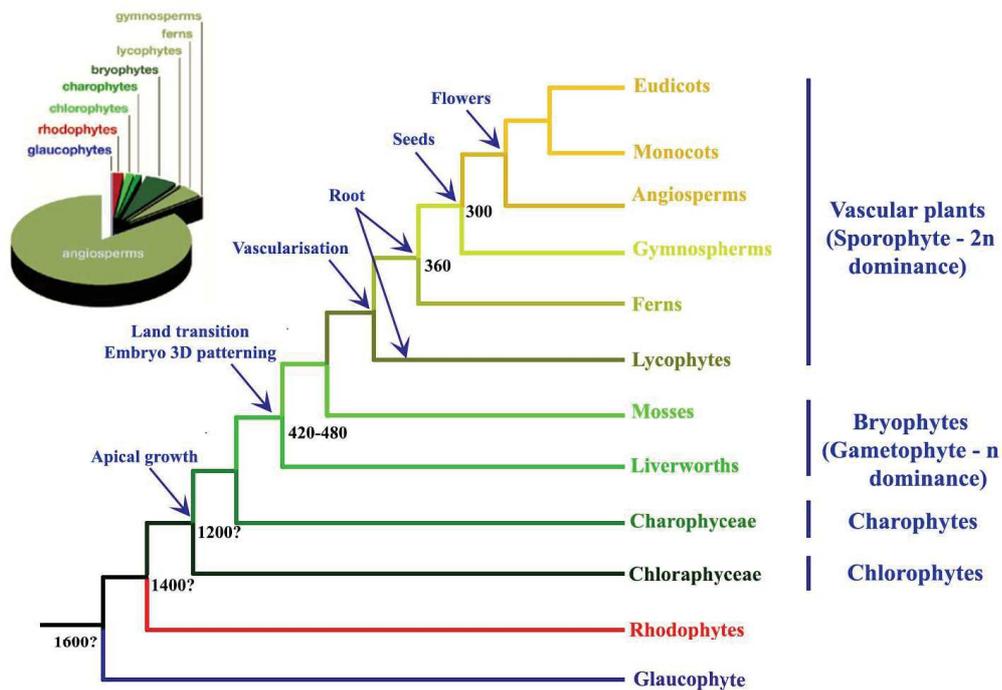
# GENERAL INTRODUCTION

## 1. THE EVOLUTION OF PLANT KINGDOM

During evolution, plants have given rise to a staggering complexity of morphological structures with different shapes, colours, and functions. However, all plants have a common ancestor: a single eukaryotic cell, which acquired a photosynthetic cyanobacterium as an endosymbiont (the ancestral plastid). The plant kingdom may be divided into three main groups: the glaucophytes (little-known freshwater algae), the rhodophytes (red algae), and the green plants (which include green algae and land plants) (**Fig. 1**). The first land plant (liverwort) appeared around 450 million years ago in the Ordovician period. In early Devonian-age rocks, which are approximately 400 million years old, fossils of simple vascular and nonvascular plants can be seen. Ferns, lycopods, horsetails and early gymnosperms became prominent during the Carboniferous period (approximately 300-360 million years ago). Gymnosperms were the dominant flora during the Age of Dinosaurs, the Mesozoic era (250 million years ago). More than 130 million years ago, from the Jurassic period to early in the Cretaceous period, the first angiosperm plants (phylum Anthophyta) arose. Angiosperms are also known as flowering plants because flowers define the most important characteristics of this group compared with other land plants: the presence of flowers, endosperm within the seeds, and the production of fruits containing the seeds. Over the following 40 million years, angiosperms (including eudicot and monocot species) became the world's dominant plants that today occupy almost every habitats on earth with approximately 235 000 species (Bowman and Eshed, 2000; Soltis *et al.*, 2007). This species diversification makes angiosperm evolution one of the most fascinating areas of study in biology.

To gain insights into the morphological diversity of angiosperm, it is essential to understand the evolution of mechanisms underlying the developmental process in a field known as “Evo-Devo” – evolutionary developmental biology. The key question in Evo-Devo is how DNA sequence changes relate to the evolution of morphological diversity. New genomic resources and techniques enable biologists to assess for the first time the evolution of developmental regulatory networks on a global scale. Numerous theories have been proposed to explain diversification and speciation (Slack and Ruvkun 1998; Arthur 2002; Koes 2008; Carroll 2008). Several key molecular mechanisms have been proposed to underlie diversification during evolution, including: (i) the functional divergence of duplicated genes (the neo/sub-functionalization of paralogues), (ii) the expression pattern divergence of conserved genes (through mutations in the *cis* or/and *trans*-regulatory regions, and (iii) “de novo” gene formation (*i.e.* exon shuffling, transposon-based exchanges, alternative splicing).

These three concepts will be detailed in the following sub-sections.



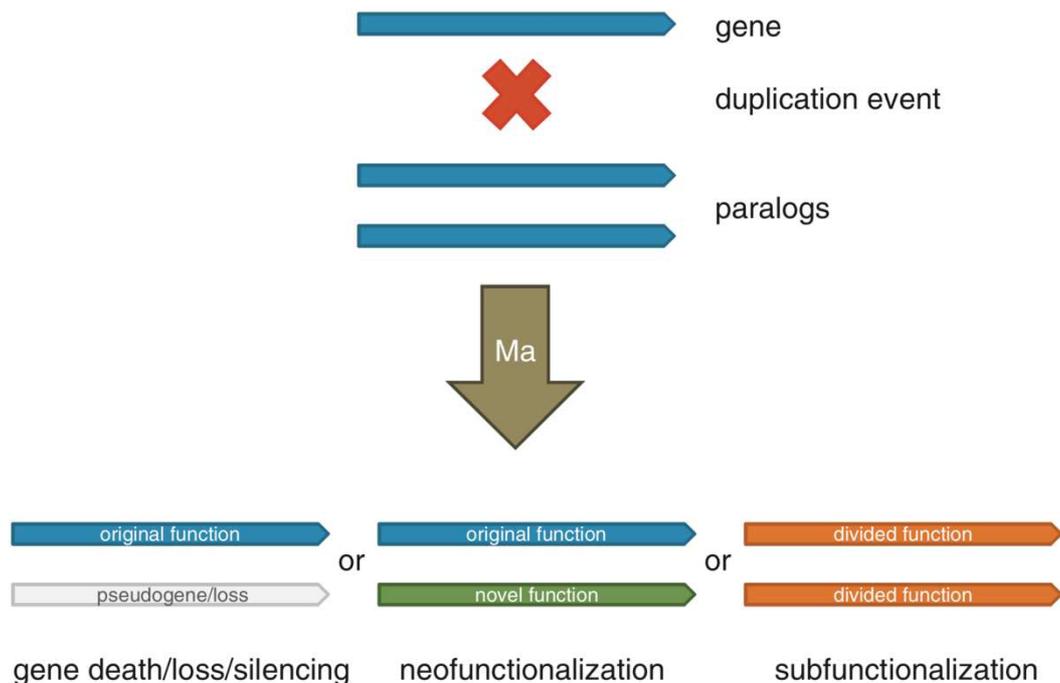
**Figure 1. Phylogenetic and morphological innovations among plants.** Depicted are relationships among the three lineages of plants: glaucophytes (freshwater algae; blue), rhodophytes (red algae; red), and the green plants (chlorophytes, charophytes, and land plants; from green to orange). Estimated dates for some nodes are listed in millions of years before present. Major events in the evolution of land plants are demarcated with arrows. Pie chart shows the relative species richness of the major clades. The vast majority of species within the Plantae consists of angiosperms.

### 1.1. THE FUNCTIONAL DIVERGENCE OF DUPLICATED GENES

Genes can duplicate at single-gene, chromosome, and whole genome level (Freeling, 2009). Many innovations in metabolic networks come from individual duplications of genes encoding enzymes (Caetano-Anollés *et al.*, 2009). On the other hand, a whole genome duplication might create larger-scale change in molecular network than a single-gene duplication might. One example of a whole genome duplication is seen with MADS-box proteins, which illustrate the evolution of a protein–protein interaction network of transcription factors and have been studied in numerous plant species (Veron, Kaufmann and Bornberg-Bauer, 2006). After a duplication event, genes can either be lost or retained in the population of the species. If a new allele is selectively neutral, compared with pre-existing alleles, it only has a small probability of being maintained during evolution (Kimura, 1991). For those that become fixed, the long-term evolutionary fate of duplication will still be determined by the functions of the duplicate genes. The birth and death of genes is a key element in gene family

and genome evolution (Nei, Rogozin and Piontkivska, 2000) with those genes involved in physiological processes that vary greatly among species (*e.g.* immunity, reproduction and sensory systems) probably having higher rates of gene birth and death.

Pseudogenization or non-functionalization is a purely neutral process that ultimately eliminates one of the duplicated copies and is the most common fate. Sub-functionalization, as a neutral process where the two copies partition the ancestral function, has been proposed as an alternative mechanism driving duplicate gene retention in small populations. Neo-functionalization is an adaptive process whereby one mutated copy confers a new function that was not determined by the original gene. Neo-functionalization can include the evolution of a completely new binding capability or modification/improvement of existing binding capabilities under positive selection after removal of pleiotropic constraints (Kramer, Jaramillo and Di Stilio, 2004; Rastogi and Liberles, 2005; Freeling, 2009) (**Fig. 2**).



**Figure 2. Schematic representation biogenesis of sub- and neo-functionalization.** From (Rensing, 2014). Ma, mega-annum, one million years.

## 1.2. THE EXPRESSION PATTERN DIVERGENCE OF CONSERVED GENES

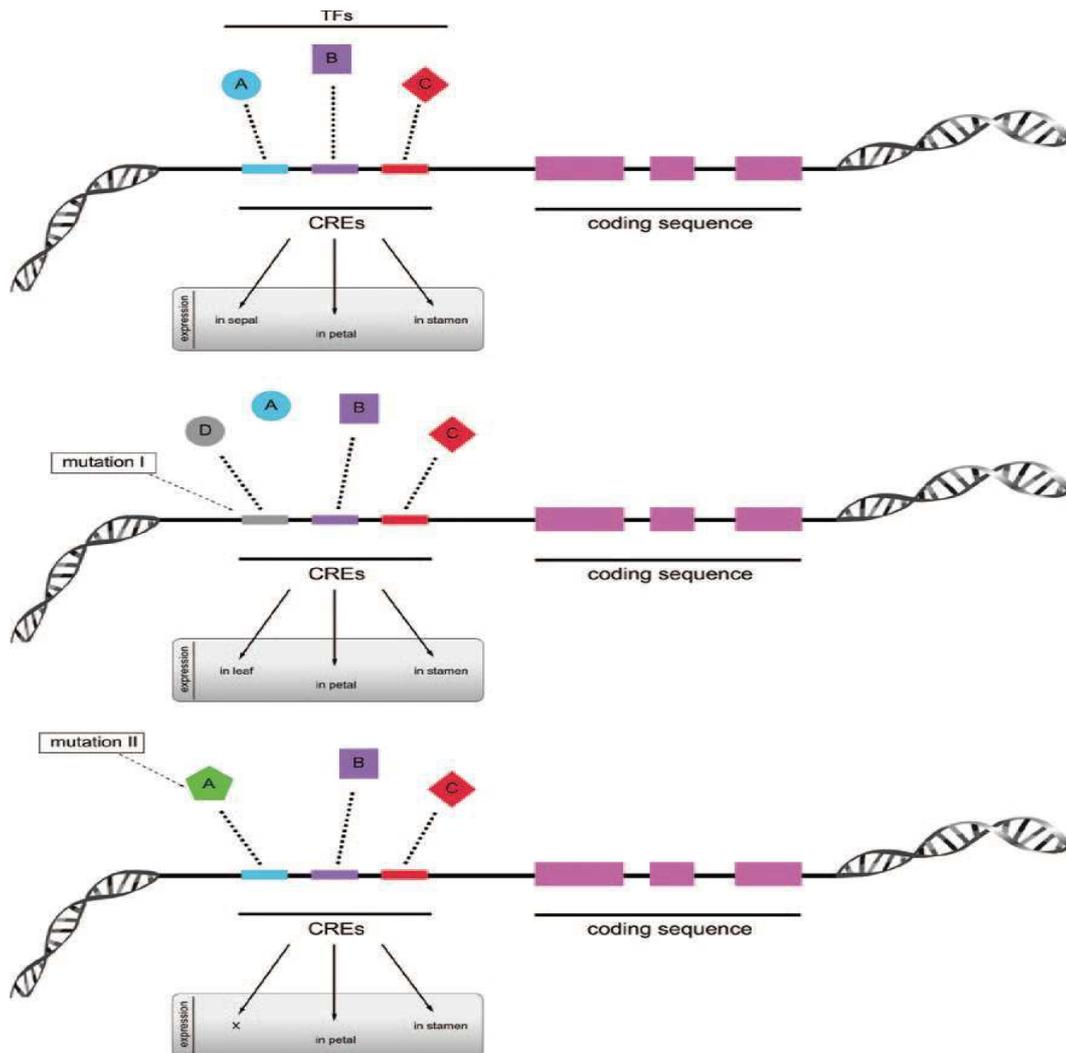
Although it is widely accepted that morphological variation between organisms arose from genetic alterations, the molecular mechanisms underlying these variations remain poorly understood. Nevertheless, studies to date suggest that morphological variation has been shaped more by alterations to the expression patterns of functionally conserved genes rather than

through the emergence of new genes and functions (Wray *et al.*, 2003; Martin, Ellis and Rook, 2010).

As an illustration, mammalian HOX proteins, a conserved homeodomain transcription factor family found in vertebrates, can still functionally replace their *Drosophila* homologues (Mallo, Wellik and Deschamps, 2010). In plants, homologs of B function MADS-box genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) from *A. thaliana* are responsible for the establishment of petal and stamen identities in the second and third whorls of floral meristem. In cases of *Arabidopsis*, *Antirrhinum* and rice, these genes are highly conserved in structure and function (Weigel and Meyerowitz, 1994; Nagasawa *et al.*, 2003; Kanno *et al.*, 2007). However, in some species (e.g. petunia, maize, tulips, lilies, etc.) AP3 and/ or PI homologs show a different spatial expression pattern within the flower, indicating a divergence of expression domains of conserved genes during evolution among angiosperm species, which may be associated with altered floral organ identities (Soltis *et al.*, 2007; Rijpkema *et al.*, 2010). These findings suggest that variations in gene expression is an important source of phenotypic diversity.

Gene expression patterns are governed by complex gene regulatory networks that include *cis*-regulatory and *trans*-regulatory elements. Consistent with the original definitions, *cis*-regulatory DNA elements have an allele-specific effect on gene expression, and map near the target gene whereas *trans* elements affect the expression of distant genes, through indirect regulation. *Trans*-regulatory elements work through intermolecular interactions to regulate the target genes through intermediaries such as transcription factors or inhibitors that regulate transcription initiation or small interfering RNAs that regulates RNA stability. On the other hand, *cis*-regulatory are physically and genetically linked to the gene (or mRNA) that they regulate (*i.e.* in a gene or an adjacent regulatory element near the target genes), examples including promoter regions, enhancers and boundary elements, which regulate transcription initiation, or poly-A signals and siRNA binding sites, which regulate RNA stability (Wray *et al.*, 2003; Gilad, Rifkin and Pritchard, 2008).

Although a number of studies have indicated that variations in *cis*-regulatory elements play important roles in Evo-Devo biology (*e.g.* *TEOSINTE BRANCHED* in maize, *ULTRABITHORAX* and *YELLOW* in fruitfly), we still know little about *trans*-regulator element (Wray, 2007). **Figure 3** represents several cases of potential mutation in *cis*-regulatory elements (CREs) that could affect the transcription process, and as a result, lead to morphological variation.

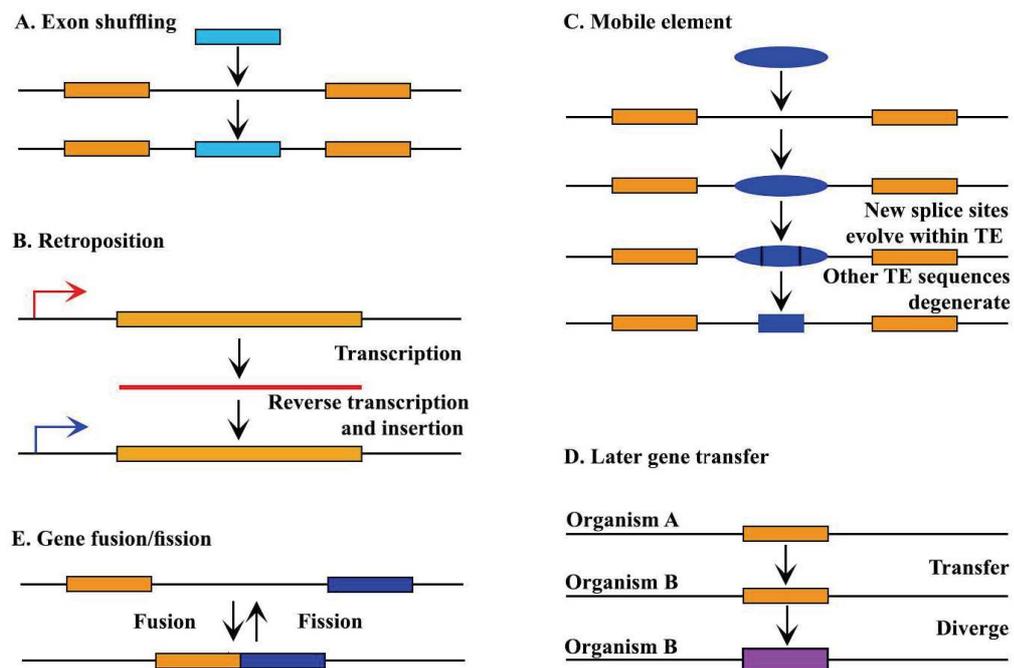


**Figure 3. Schematic representation of a gene with its *cis*-regulatory elements (CREs) and the potential mutations that can affect transcriptional processes.** CREs (A, B, C) together with their respective transcription factors (TFs) allow expression of a gene in a specific organ (or tissue). Middle panel: mutation in one CRE (in this case, the binding site of A became D) leads to loss of expression in sepals but the gene acquires expression in leaves. Bottom panel: mutation in a TF (in this case, A) leads to lack of activation of the gene in a specific organ (in this case, sepals). Adapted from (Della Pina, Souer, & Koes, 2014).

### 1.3. “*DE NOVO*” FORMATION OF NEW CODING GENES

The formation of new genes is an important mechanism generating genetic novelties during the evolution of an organism. *De novo* formation is a process creating new protein-coding genes from non-coding DNA or/and other coding DNA through several mechanisms such as exon shuffling, gene fission/fusion, retroposition, and lateral gene transfer (**Fig. 4**) (Long *et al.*, 2003).

Exon shuffling created around 19% of exons in eukaryotic genes through ectopic recombination of exons and domains from distinct genes (Patthy, 1996, 1999). Morgante et al. (2005) indicated that some genic insertions occurring in maize share the structural hallmarks of Helitron rolling-circle transposons. DNA segments defined by Helitron termini contain multiple gene-derived fragments that are located in multiple genomic locations. Some of the transcripts produced contain segments from different genes, supporting the idea that these transposition events have a role in exon shuffling and in the creation of new proteins (Morgante *et al.*, 2005).



**Figure 4.** *De novo* formation of novel protein-coding genes. (A) exon shuffling, (B) retroposition, (C) Mobile element, (D) later gene transfer, (E) gene fusion/fission (adapted from Long *et al.*, 2003).

Retroposition is a mechanism relating to functional retrogenes when new duplicated genes are created in new genomic positions by reverse transcription or other processes (Betrán and Long, 2002; Wang *et al.*, 2002). New functional retrogenes have been reported in various organisms, especially mammals and *Drosophila melanogaster* (Long *et al.*, 2003; Betran *et al.*, 2004). In plants, beside the few retrogenes that have been identified in the actin gene family of potato (*Solanum tuberosum*), other examples include the alcohol dehydrogenase gene family in *Leavenworthia* and the *Bs1* retrotransposons in maize (Drouin and Dover, 1990; Jin and Bennetzen, 1994; Charlesworth, Liu and Zhang, 1998). H. Wang *et al.* 2005 reported on the abundance of retrogenes in rice, maize and sorghum genomes, suggesting that retroposition shapes the genomes of grass species in general.

The other mechanism, which was reported often in prokaryotes and recently in plants, is lateral gene transfer. This process occurs when a gene is laterally transmitted between organisms (Ochman, 2001; Bergthorsson *et al.*, 2003). The model proposes that two adjacent genes can fuse into a single gene or that a single gene can split into two genes through the deletion, insertion or mutation of a translation stop codon and that alteration of transcription termination signals can allow the creation of new gene functions (Nurminsky *et al.*, 1998) (Fig. 4).

Almost all new gene functions are created from ancient genes which undergo continuous changes in sequence and structure to establish further diverged functions. In contrast, the *de novo* gene origin process, whereby a whole protein-coding gene is created from a fragment of non-coding sequence, is considered to be rare (Long *et al.*, 2003). Nevertheless, Snel *et al.* (2002) suggested that *de novo* evolution not only plays an important role in generating the initial common ancestral protein repertoire but also contributes to the subsequent evolution of an organism. However, it is nearly impossible to identify the non-coding origin of the initial ancestral proteins because of long-term accumulation of mutations. How exactly non-coding regions in a genome create new functions, and the role of non-coding regions in genome, is still an open question.

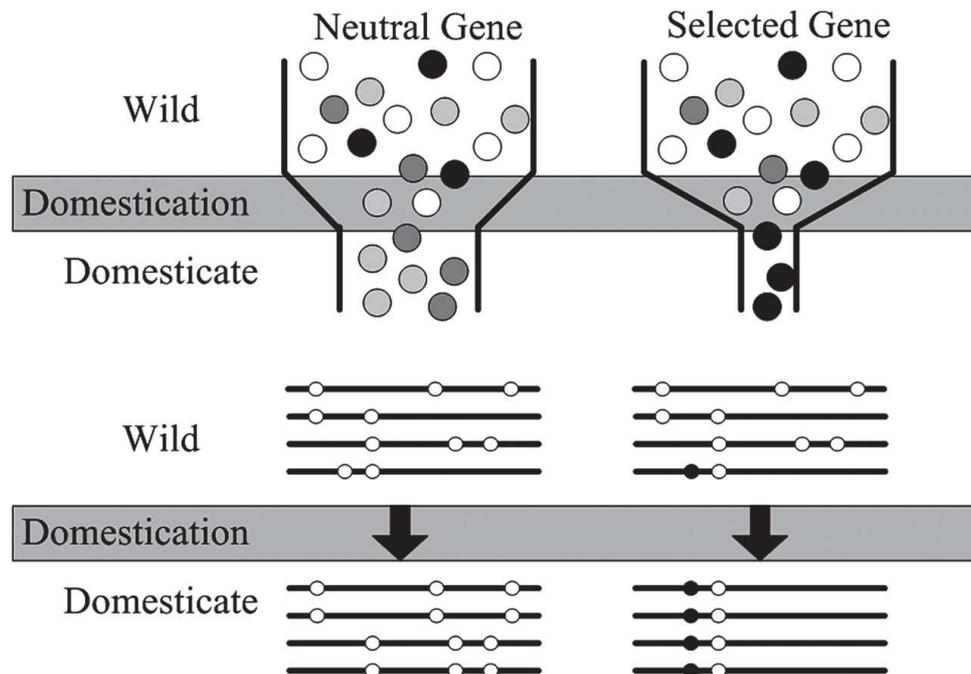
## **1.4. DOMESTICATION**

### **1.4.1. Domestication process**

The process of evolution leads to the increased adaptation of an organism to a changing environment, whereas the domestication process has led to increased adaptation of plants and animals to cultivation or rearing or utilization in general by human beings. Research aimed at understanding domestication has also been a tremendous help in understanding evolution. Domestication originated around 10 000 years ago when agriculture began to encourage the growth of edible wild plants (John F. Doebley, Gaut and Smith, 2006). In contrast to the earlier hunting-gathering period, humans started to select and re-sow grasses (i.e. cereals) from the previous season for the next season. Once the process had been repeated a number of times, the proportion of plants in the field that had desirable traits would be increased (Chen, Gols and Benrey, 2015).

For many crops, such as maize and cauliflower, domestication has rendered the plant completely dependent on humans such that it is no longer capable of propagating itself in

nature. However, other crops, such as hemp, carrot, and lettuce, have been more modestly modified compared to their progenitors, and they can either revert to the wild or become self-propagating weeds. Compare to their ancestor, domesticated crops typically show synchronization of flowering time, enlargement of reproductive organs (*i.e.* bigger fruits, more grain, etc), lost natural seed dispersal (*i.e.* seeds remain attached to the plant for easy harvesting by humans), increased apical dominance, and other features collectively known as the “domestication syndrome” (John F. Doebley, Gaut and Smith, 2006).



**Figure 5. Schematic representation of a population bottleneck and its effect on a neutral gene and a selected gene.** In the upper section, shaded circles represent genetic diversity. The bottleneck reduces diversity in neutral genes, but selection decreases diversity beyond that caused by the bottleneck alone. Lower illustrates sequence haplotypes of these two hypothetical genes. The neutral gene loses several haplotypes through the domestication bottleneck, but the selected gene is left with only one haplotype containing the selected site. From (Ross-Ibarra, Morrell and Gaut, 2007).

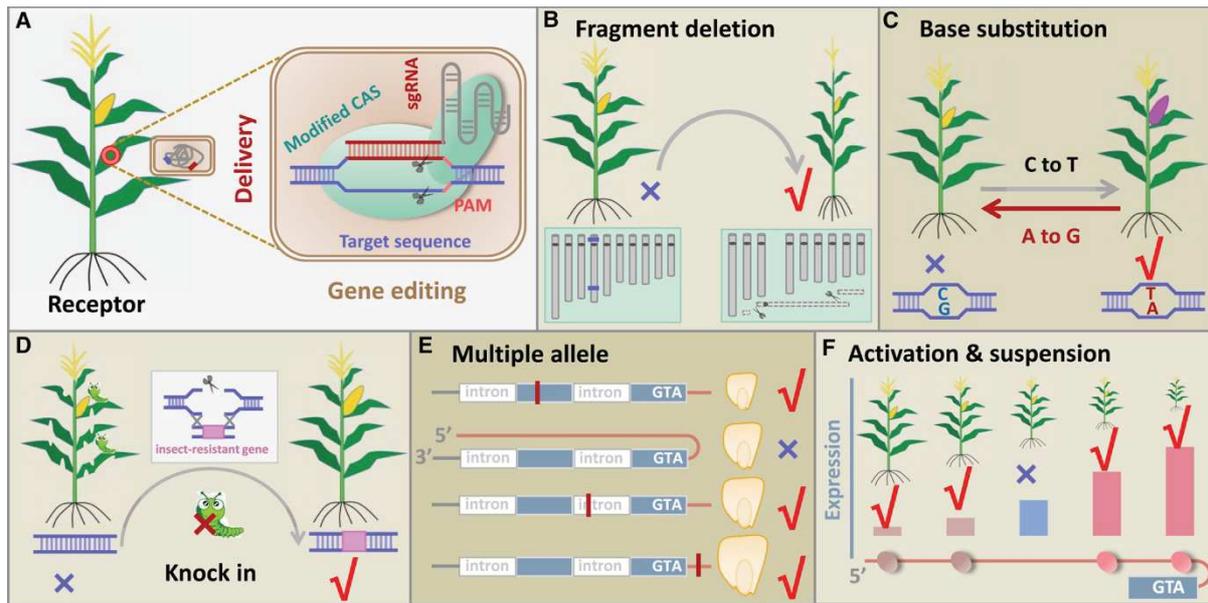
During domestication, these early agricultural practices left their signatures on the patterns of genetic diversity in the genomes of crop plants. Because early farmers used only a limited number of individuals of the progenitor species, much of the genetic diversity in the progenitor was left behind. Moreover, with each generation during the domestication process, only seed from the best plants formed the next generation. This caused a genetic bottleneck, which reduced genetic diversity throughout the genome (**Fig. 5**) (Doebley, 1993). The extent of this loss of diversity depends on the population size during the domestication period and the duration of that period (Eyre-Walker *et al.*, 1998). It should be noted that the loss in diversity is not experienced equally by all genes in the genome. For genes that do not influence favoured

phenotypes (which are called neutral genes), the loss in diversity results only from the strength of the bottleneck in terms of the population size and duration (**Fig. 5**). However, genes that influence desirable phenotypes, i.e. domestication genes, inevitably experience a more drastic loss of diversity, (**Fig. 5**). This is because plants carrying favoured alleles contribute proportionally more to the progeny of each subsequent generation while other alleles may be eliminated from the population (Wright and Gaut, 2004; Wright *et al.*, 2005).

One unknown in the domestication process is the extent to which new mutations *versus* pre-existing genetic variation in the wild species contribute to the evolution of crop phenotypes. For example, in a few cases, crops possess alleles of major genes that disrupt seed shattering (Li, Zhou and Sang, 2006) or the protective casing surrounding the seed (Wang *et al.*, 2005) that are not found in the progenitor species. However, alleles of genes that contribute to increased fruit size in tomato (Nesbitt and Tanksley, 2002) or increased apical dominance in maize (Clark *et al.*, 2004) are also found in their wild relatives, albeit at lower frequencies. Given the large reserve of genetic variation in the progenitor species, it seems reasonable to infer that domestication mostly involves enrichment of the best alleles from pre-existing allelic variation in crop ancestors, although new mutations in key developmental pathways may have been instrumental for some traits.

#### **1.4.2. *De novo* domestication**

Recently, *de novo* domestication was described as a form of domestication that may be achieved by the application of genome-editing. Two parallel approaches have in fact been suggested for the *de novo* domestication of wild plants: traditional breeding approaches (Runck *et al.*, 2014; DeHaan *et al.*, 2016) and gene editing (Altpeter *et al.*, 2016; Kantar *et al.*, 2016). For gene editing, the CRISPR/Cas9 approach has become the method of choice (Altpeter *et al.*, 2016; Pacher and Puchta, 2017; Scheben *et al.*, 2017). This genome-editing tool, which is modified from a prokaryotic immunity-determining system, induces double-stranded DNA breaks by the action of Cas9 nuclease at a genomic location corresponding to a designed guide RNA (Altpeter *et al.*, 2016) (**Fig. 6**). However, one prerequisite for the use of CRISPR/Cas9 in *de novo* domestication is that the genome of the plant to be edited must be sequenced in order to identify known orthologs of the domestication-related genes of interest (Fernie and Yan, 2019).



**Figure 6. A Simplified Schematic Representation of Genome-Editing Techniques in Plants and Their Potential Application.** (A) Gene-editing model, including editing element delivery and modified model with single guide RNA (sgRNA) and different functional CAS protein or protein complex. (B and C) Delete any sequence including large chromosomal fragments or (B) even the entire chromosome via paired sgRNA (C) to achieve any base substitution. (D–F) Add genes that do not exist in the original genome (D), create multiple different alleles of any gene (E), and activate or suspend the function of any gene (F). Ovals represent activator complex (red) and repressor complex (pink). From (Fernie and Yan, 2019).

### 1.4.3. Domestication genes

Several genes that were targeted during domestication or crop improvement have been identified in pathways governing fruit size and shape, seed dispersal, tiller number, seed colour, and many other traits (John F Doebley, Gaut and Smith, 2006; Izawa *et al.*, 2009). Because the traits involved are mostly quantitative in nature, the approach to identify these genes involves the mapping of quantitative trait loci (QTL) in progenitor crop hybrid populations, followed either by positional cloning or cloning using a combination of positional information and candidate gene analysis.

The form and nature of the genetic mutations associated with transitions from wild to domesticated plants is highly variable (Fig.7). Some mutations causing frameshifts or premature termination of the protein product are considered non-functional ‘knock-outs’ in the domesticated species. Meanwhile, some mutations are in the regulatory elements and modify spatiotemporal expression patterns and/or levels (Doebley *et al.*, 2006). In the case of *sh4* (Li, Zhou and Sang, 2006), amino acid substitutions appear to disrupt the interaction of the protein with downstream targets. The *sh4* is a major QTL controlling whether the seed fall off the plant

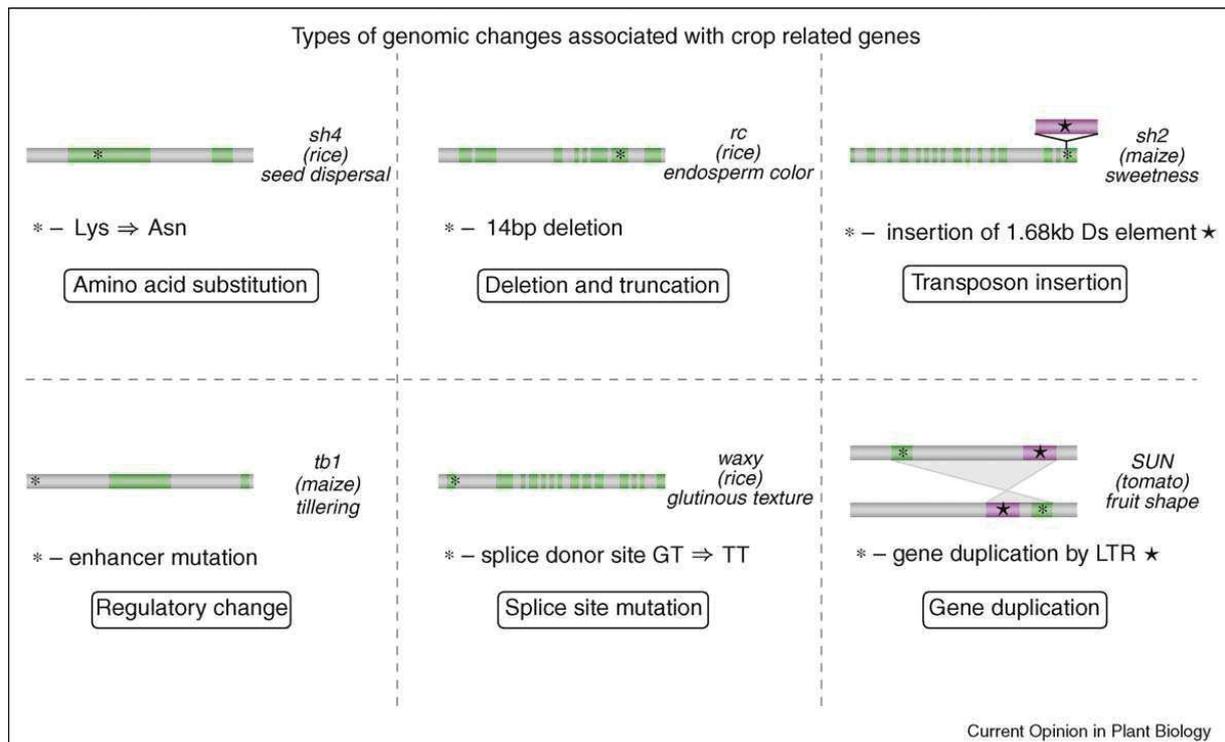
(shatter) as in wild rice or adhere to the plant as in cultivated rice (Li, Zhou and Sang, 2006).

The other QTL controlling shattering in rice, namely *qSH1*, encodes a homeobox containing transcription factor (Konishi *et al.*, 2006). The authors demonstrate that a single nucleotide change in a *cis*-regulatory element of *qSH1* eliminated the expression of the homeobox gene at the provisional abscission layer in the shattering zone, thus preventing shattering (Konishi *et al.*, 2006). It has also been demonstrated that selection for the *qSH1* allele was not as intense and expansive as the selection for the *SH4* allele.

Two examples of domestication genes in rice are the *Rc* and *waxy* genes. *Rc* encodes a bHLH transcription factor involved in the change from red pericarp (in wild rice) to white pericarp (in most cultivated rice cultivars). The gene's function is impaired in the domesticated form by a 14-bp frame-shift deletion that truncates the protein upstream from the bHLH domain, thus producing white pericarp. This mutation is common within all *O. sativa* sub-populations (Sweeney *et al.*, 2006). The *Waxy* gene encodes a granule bound starch synthase (GBSS), whereas in some domesticated lines an altered intron splice donor site in the gene leads to a glutinous ("sticky") grain that lacks amylase (Wang *et al.*, 1995; Olsen *et al.*, 2006).

In maize, *Teosinte branched1* (*tb1*) encodes a transcription factor involved in the regulation of cell cycle genes. It was identified as a major QTL controlling the difference in apical dominance between maize and its progenitor, teosinte (Doebley, Stec and Hubbard, 1997). The maize *tb1* mutation represses the outgrowth of the axillary meristems and branch elongation via its repressive effect on the cell cycle, thus maize plants typically have a single stalk with short branches tipped by ears, whereas teosinte plants are more highly branched (Doebley, Stec and Hubbard, 1997; Wang *et al.*, 1999).

In tomato, *Fruitweight2.2* (*Fw2.2*) and *SUN* are two domestication-related genes that regulate fruit shape. The gene *Fw2.2*, encoding a protein that inhibits cell division in the fruit, was identified via a QTL that controlled 30% of the difference in fruit mass between wild and cultivated tomato (Frary *et al.*, 2000; Cong, Liu and Tanksley, 2002). However, the large- and small-fruited alleles have no differences in protein sequence, supporting the hypothesis that changes in the regulation of *Fw2.2*, rather than its gene sequence, underlie the evolution of tomato fruit size (Nesbitt and Tanksley, 2002). *SUN*-over-expression causes a gene duplication event mediated by the long terminal repeat retro-transposon (Frary *et al.*, 2000; Cong, Liu and Tanksley, 2002).



**Figure 7. Types of changes associated with crop-related genes.** One specific example is given for each type of genomic change: amino acid substitution (*sh4* in rice), deletion and truncation (*rc* in rice), transposon insertion (*sh2* in maize), regulatory change (*tb1* in maize), splice site mutation (*waxy* in rice) and gene duplication (*Sun* in tomato) (From by (Tang, Sezen and Paterson, 2010).

The history of crop domestication parallels the most significant period of human history to date, that occurred over the past 12 000 years. Crop domestication has allowed the development of civilizations based on agriculture, by enabling human beings to transition from a nomadic hunting lifestyle to a self-sufficient modern day life (Fernie and Yan, 2019). Obviously understanding domestication has already been and will continue to be a tremendous help in understanding the mechanisms of evolution. Studies of crop plants are thus vital and the knowledge obtained will also provide a solid foundation for the engineering of new varieties in the future.

## 2. INFLORESCENCES AND MERISTEMS

An inflorescence is a plant reproductive structure comprising a cluster of flowers. This structure is established in different ways leading to different types of organization and complexity levels. This structure is diversified among different species and plays a crucial role in plant reproduction as it strongly impacts on pollination and fruit set (Robert, 1982). Inflorescence development is regulated by different regulatory gene networks. Additionally, the final architecture is initially determined by the activity of apical and axillary meristems that define the branching pattern along with flower positioning (Yamburenko, Kieber and Schaller, 2017). In the following sections, the scientific background of meristem functioning along with current knowledge of related regulatory gene networks will be described. Moreover, the diversity of inflorescence architecture and its modeling to explain its evolution will be clarified.

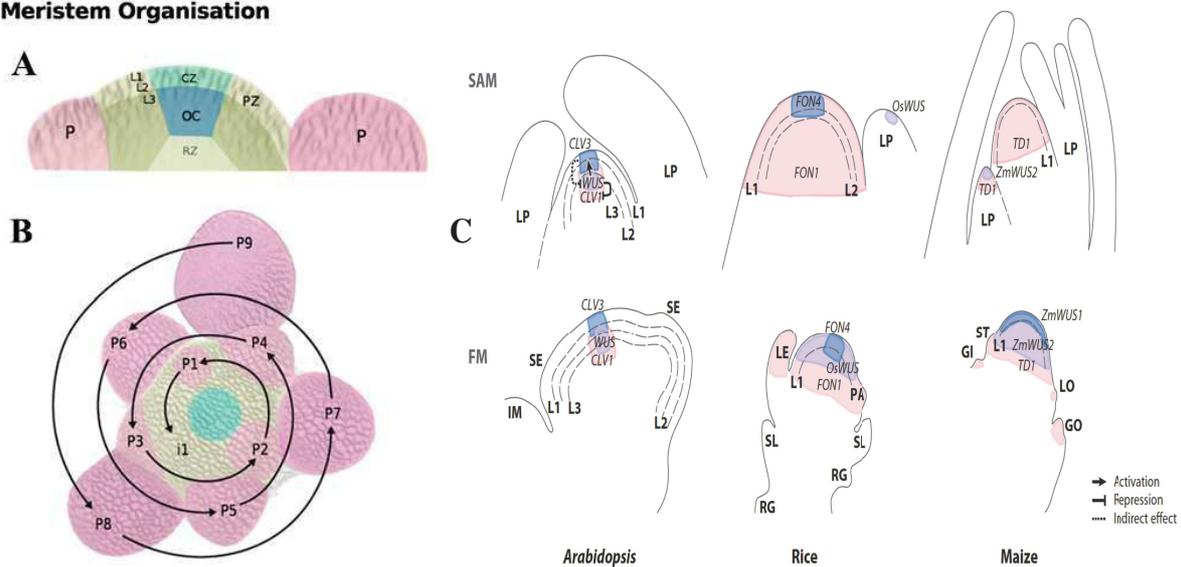
### 2.1. MERISTEM FUNCTIONING

In plants, there are specific zones where a self-renewing population of undifferentiated cells divide and grow. These cells, called meristematic cells, possess the unique property of totipotency, which means the ability to divide and produce all differentiated cells in an organism. In embryogenesis, the apical-basal axis is defined with the root apical meristem (RAM) at one extremity and the shoot apical meristem (SAM) at the other (Schmitz and Theres, 2005). The apical meristems of both RAM and SAM are primarily undifferentiated (indeterminate) meristems giving the main blueprint for the rest of plant development (Brukhin and Morozova, 2011). Moreover, the post-embryogenic growth of plants depends on the persistent function of these meristems which are established during embryogenesis. The SAM is responsible for production of all above-ground plant organ stems including leaves and flowers while the RAM generates all cell types of the root system. Floral meristems (FMs) are products of the reproductive SAM which maintain a transient stem cell reservoir for flower formation. FM activities are regulated by feedback loops shared between the SAM and floral-specific factors (Ha, Jun and Fletcher, 2010).

The structure of SAMs is conserved in different plant species with cell layers and central zones (**Fig. 8**). In *A. thaliana*, cells in the outermost layer (L1) divide in anticlinal orientation and develop the epidermal layer. The L2 layer is internal to the L1 and mainly generates mesophyll tissue. The interior of the meristem is defined as the L3 and includes multiple cell layers which form the internal tissues, mesophyll and vascular tissues. Based on

cytoplasmic densities and cell division rates, the SAM may be subdivided into the peripheral zone (PZ) and the central zone (CZ) comprising the organizing center (OC) and the rib zone (RZ) (**Fig. 8A-B**) (Ha, Jun and Fletcher, 2010; Murray *et al.*, 2012). In addition, the lateral organs are established from cells recruited from the PZ while stem tissue is derived from cells recruited from the RZ. The CZ acts as a reservoir of stem cells, which build up both the peripheral and rib zones, as well as maintaining the integrity of the central zone (Lenhard and Laux, 1999). While both *A. thaliana* and rice have three layers in the SAM, with a tunica consisting of two clonal layers (designated L1 and L2) and the corpus (designated L3), maize has only one obvious tunica layer (L1) along with the inner corpus (1) (**Fig. 8C**).

**Meristem Organisation**



**Figure 8. Structural and functional organization of the SAM.** In *A. thaliana* (**A**) The different zones and layers of the SAM. (**B**) Primordia are spaced according to a regular pattern or phyllotaxis. P9 indicates the oldest primordium and P1 the youngest, (i1) is the next primordium. Adapted from (Murray *et al.*, 2012). (**C**) Conserved and diversified features of *CLAVATA* (*CLV*) signaling in *A. thaliana*, rice and maize grasses. Expression domains are shown in blue for *CLV3* and *FLORAL ORGAN NUMBER 4* (*FON4*); in pink for *CLV1*, *FON1*, and *THICK TASSEL DWARF1* (*TDI*); and in purple for *WUSCHEL* (*WUS*), rice *WUS* (*OsWUS*), maize *WUS1* (*ZmWUS1*), and *ZmWUS2*. FM, floral meristem; GI, inner glume; GO, outer glume; IM, inflorescence meristem; LE, lemma; LO, outer lemma; LP, leaf primordium; PA, palea; RG, rudimentary glume; SAM, shoot apical meristem; SE, sepal; SL, sterile lemma; ST, stamen. Adapted from Zhang and Yuan, (2014).

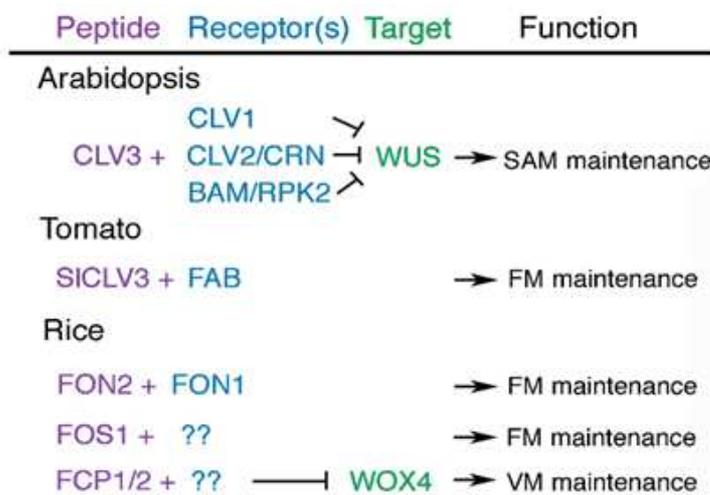
The CLV-WUS signaling pathway, which is one of the best characterized signaling pathways (Zhang and Yuan, 2014b), plays a central role in maintaining shoot and floral stem cell homeostasis in *Arabidopsis*. In *Arabidopsis*, cells in the OC express the transcription factor WUSCHEL (WUS), which promotes the expression of *CLAVATA3* (*CLV3*) gene, encoding a small peptide that moves into the surrounding tissue (Kondo *et al.*, 2006; Müller, Bleckmann

and Simon, 2008a). *CLV3* is a founding member of the CLAVATA3/EMBRYO SURROUNDING REGION (CLE) family of polypeptides (Suzaki, Yoshida and Hirano, 2008), which are present throughout the plant kingdom (Whitewoods *et al.*, 2018). The *WUS* gene is required to maintain stem cell fate during vegetative and reproductive development (Laux *et al.*, 1996). In addition, *WUS* is a bi-functional protein that can both repress and activate gene transcription in the SAM (Ikeda, Mitsuda and Ohme-Takagi, 2009). The *WUS* protein migrates between cells through plasmodesmata into the apical stem cell (Daum *et al.*, 2014) where it induces the expression of the *CLV3* gene in a dosage-dependent fashion (Yadav *et al.*, 2011). Then, *CLV3* is bound by CLAVATA1 (*CLV1*), an extracellular leucine-rich repeat receptor-like kinase (LRR-RLK) produced in cells beneath the stem cell reservoir (Clark, Williams and Meyerowitz, 1997; Ogawa *et al.*, 2008). Additionally, CORYNE (*CRN*), a presumptive pseudo-kinase protein, functions as a CLAVATA2 (*CLV2*) co-receptor that is a leucine-rich-repeat protein with no kinase domain. *CRN* mediates localization of *CLV2/CRN* complex to the plasma membrane, where they can directly interact with *CLV1* heterodimers (Bleckmann *et al.*, 2010; Guo *et al.*, 2010). In contrast to the *CLV1* gene, *CLV2* and *CRN* are expressed throughout the entire SAM, and the *CLV2-CRN* complex functions largely independently of *CLV1* in *CLV3* signal transduction (Müller, Bleckmann and Simon, 2008; Zhu *et al.*, 2010).

Other receptors appear to mediate *CLV3* signaling predominantly on the flanks of the meristem. Three LRR-RLK genes that are closely related to *CLV1*, *BARELY ANY MERISTEM1*, 2 and 3 (*BAM1-3*), act redundantly to maintain stem cell fate on the meristem periphery (DeYoung *et al.*, 2006), and both *BAM1* and *BAM2* directly bind to *CLV3* (Shinohara and Matsubayashi, 2015). The *BAM1* protein physically cooperates with the LRR receptor-like kinase RECEPTOR-LIKE PROTEIN KINASE2 (*RPK2*) (Kinoshita *et al.*, 2010), which itself does not bind *CLV3* and thus is supposed to promote meristem maintenance (Shinohara and Matsubayashi, 2015). *CLV3*-mediated signaling through these receptor complexes limits stem cell accumulation by restricting the *WUS* expression domain to the OC (Brand *et al.*, 2000; Schoof *et al.*, 2000). Thus, the *CLV-WUS* pathway functions as a negative feedback loop to maintain SAM homeostasis.

Even though eudicots and monocots have histological differences in SAM structure, increasing evidence suggests that the *CLV* signaling pathway is partially conserved between them (**Fig. 9**). In tomato, the major quantitative trait loci (QTL) *fasciated (fas)* and *locule number (lc)* contain genes that affect the number of tomato fruit locules, and most cultivated

tomato varieties include mutations in either the *fas* or the *fas* and *lc* genes (Lippman and Tanksley, 2001). The multilocular *fas* phenotype is caused by a mutation in the regulatory region of a *CLV3*-related gene, *SlCLV3* (Xu *et al.*, 2015), likewise the *lc* trait is due to two single nucleotide polymorphisms (SNPs) in a repressor element downstream of a *WUS* gene homolog (Muños *et al.*, 2011). It is suggested that selection at both loci took place during tomato domestication to improve the fruit locule number (Muños *et al.*, 2011; Xu *et al.*, 2015). A forward genetic screen for more complex inflorescence branching and fruit locule number allowed the discovery of new *CLV* pathway members (Xu *et al.*, 2015). Plants with the *fasciated and branched (fab)* multilocular phenotype possess a mis-sense mutation in the closest tomato homolog of *CLV1*, which affects the kinase domain. Interestingly, both the *fasciated inflorescence (fin)* and *fab2* phenotypes are caused by mutations in arabinosyltransferase genes (Ogawa-Ohnishi, Matsushita and Matsubayashi, 2013). Moreover, the addition of arabinosylated *SlCLV3* peptides can rescue the tomato *fin* phenotype (Xu *et al.*, 2015). Available data thus demonstrates that arabinosyltransferase encoding genes are essential components of the *CLV-WUS* stem cell signaling pathway that can be targeted to improve the crop productivity traits.



**Figure 9. Components of CLV-WUS signaling pathways and their functions in model and crop plants.** Proteins with characterized genetic and/or biochemical interactions are listed. Unidentified peptides and receptors are denoted by question marks. Arrows depict positive regulation and bars depict negative regulation. SAM, shoot apical meristem; FM, floral meristem; VM, vegetative meristem; IM, inflorescence meristem. From (Fletcher, 2018).

The role of the *CLV-WUS* pathway in mediating shoot meristem maintenance is partially conserved in agronomically important grass species (Fletcher, 2018). In rice, several distinct pathways regulate stem cell maintenance, depending on the type of meristem.

The *FLORAL ORGAN NUMBER (FON1)* gene encodes the rice ortholog of the CLV1 receptor kinase (Suzaki *et al.*, 2004). *FON1* and *FON2* genes specifically prevent stem cell accumulation in floral meristems, without affecting vegetative or inflorescence meristem activity (Nagasawa *et al.*, 1996). However, *FON1* is expressed within the floral meristems but also within the shoot meristem throughout development, suggesting that related receptor kinase genes share functional redundancy with *FON1* in vegetative and inflorescence tissues (Suzaki *et al.*, 2004). Such genes, however, remain to be described (Fletcher, 2018). The *FON2* gene, also referred to as *FON4*, functions in the same genetic pathway as *FON1* and encodes a CLV3-related protein (Chu *et al.*, 2006; Suzaki *et al.*, 2006). Like *CLV3*, *FON2* is expressed at the apex of both shoot and floral meristems (Chu *et al.*, 2006; Suzaki *et al.*, 2006). Thus, in rice floral meristems, the *FON1-FON2* system corresponds to the CLV1-CLV3 peptide-receptor kinase signaling system in *A. thaliana* (**Fig. 9**). Several other *CLV* genes also play roles in orchestrating rice meristem maintenance. QTL analysis identified the *FON2 SPARE1 (FOS1)* gene in *indica* varieties as a suppressor of the *fon2* floral organ number phenotype in *japonica* varieties, indicating that *FOS1* can substitute for *FON2* activity in rice floral meristems (Suzaki *et al.*, 2009). Two other *CLV* genes, *FON2-LIKE CLE PROTEIN1 (FCP1)* and *FCP2*, encode proteins that differ in the CLE domain by one amino acid and act redundantly to negatively regulate vegetative stem cell activity and promote leaf initiation (Tanaka *et al.*, 2015). *FCP1* represses the expression of rice *WOX4*, an ortholog of *A. thaliana WOX4* (Nardmann and Werr, 2006), which promotes the undifferentiated state of the vegetative SAM (Tanaka *et al.*, 2015). Thus the rice *WOX4* gene functions similarly to the *A. thaliana WUS* gene (Mayer *et al.*, 1998), whereas the *WUS* ortholog in rice, called *TILLERS ABSENT1 (TAB1)*, is required for axillary meristem initiation but not for shoot or floral meristem maintenance (Tanaka *et al.*, 2015). These studies identify additional *CLE* signaling peptide genes besides *CLV3* as potential targets for genome editing to enhance yield traits in crop plants, particularly grasses (Fletcher, 2018).

In rice, during the vegetative phase, the SAM establishes leaf primordia on its periphery, and then generates the secondary shoots or tillers. Once the appropriate environmental and developmental signals have been recognised, plants switch to the reproductive phase. The vegetative meristem converts into an inflorescence meristem (IM) such as a rachis meristem that then produces branch meristems, and floral meristems. During this period, the meristem changes its fate and transforms from an indeterminate meristem in apical and axillary meristems (i.e. self-maintaining activity on) into a determinate meristem in

the case of flowers (i.e. self-maintaining activity off, but organogenesis on).

The process of establishment of apical vs. axillary meristems (number, timing, spatial organization) and the transition from indeterminate to determinate fates are different factors contributing to the diversity of inflorescences observed in nature. Details of inflorescence architecture will be described in the next section.

## 2.2. INFLORESCENCE ARCHITECTURES

In general, inflorescence architecture comprises several units including bracts (the terminal leaves associated with a flower), pedicels (the stalk bearing a flower) and flowers. Based on the determinacy of shoot meristems, inflorescences are basically divided into two categories, depending on whether the primary inflorescence axis terminates into a flower or not. According to this classification, determinate inflorescences are those in which, after floral transition, the SAM promotes the identity of a floral meristem that generates a terminal flower (TFL) (Weberling, 1989). This type of inflorescence sets up from extremely simple architectures, such as that of *Tulipa* sp (i.e. a single flowered inflorescence), to highly complex forms such as some *Solanaceae* species (Zachary B. Lippman *et al.*, 2008) where after formation of the TFL by the primary axis, growth continues from lateral axes that repeat this pattern. In contrast, the SAM is never converted into a floral meristem in indeterminate inflorescences and the inflorescence meristem continues producing floral meristems until senescence, such as in the model plant species *A. thaliana* (**Fig. 10**) (Weberling, 1989; Benlloch *et al.*, 2007). It was proposed that the indeterminate form of inflorescence architecture was derived from an determinate inflorescence structure independently several times during evolution (Stebbins, 1974).

Another similar way to classify inflorescences involves the designation of three main types, namely cyme-type (e.g. tomato, petunia), raceme-type (e.g. *A. thaliana*) and panicle-type (e.g. rice). In cyme inflorescences, the apex also transforms into a terminal flower, but growth of the inflorescence continues through lateral axes produced below the terminal flower (**Fig 10 C-D**). These lateral axes again form terminal flowers and this process is reiterated several times. Thus, multiple terminal flowers are generated on a single inflorescence (Souer *et al.*, 1998). In the raceme-type of *A. thaliana*, the main inflorescence meristem grows indefinitely and generates either floral meristems (FMs) or primary branch meristems (PBMs) that continuously produce FMs (**Fig. 10G**) (Remizowa *et al.*, 2012). The panicle-type inflorescences are largely characteristic of grasses such as rice (*O. sativa*) and oat (*Avena*

*sativa*). The main inflorescence meristems of these plants terminate after producing a series of lateral branch meristems, which eventually terminate in flowers after generating either flowers or higher-order branches (**Fig. 10A**) (Yamaki *et al.*, 2010). In general, structural variation among inflorescences can be attributed to three main factors: (i) the determinacy or indeterminacy of meristems within the shoot system; (ii); extent of growth in each of three dimensions of stem and stem-like structures (i.e. internode length) and (iii) relative positions of lateral shoots and/or flowers (i.e. phyllotaxy). (Prusinkiewicz, Erasmus, Lane, Lawrence D. Harder, *et al.*, 2007). These different points will be illustrated in the following sub-sections.

### **2.2.1. Indeterminate and determinate inflorescence architecture**

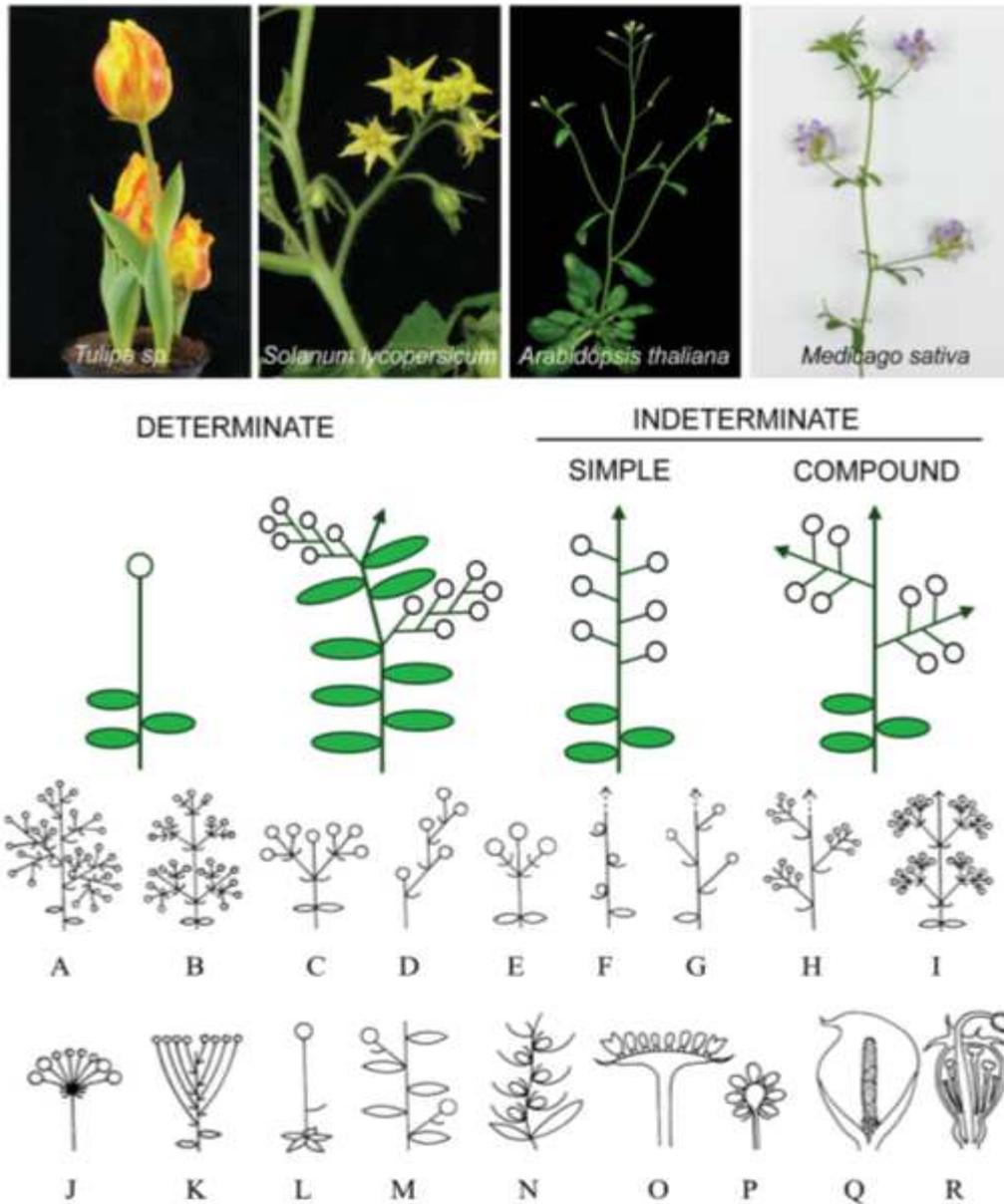
In species that produce indeterminate inflorescences, the apical meristem remains indeterminate and produces lateral meristems that become flowers (**Fig. 11A**). Inflorescences in which flowers are directly formed from the main axis are called simple racemes, such as in *Antirrhinum majus* and *Arabidopsis thaliana*. The other inflorescences where flowers are formed from secondary or higher order branch meristems are called compound racemes such as in the leguminous species pea (*Pisum sativum*), *Medicago truncatula* or *Lotus japonicas*.

In species that produce determinate inflorescences, all shoot meristems in the inflorescence eventually become floral meristems (**Fig. 11B**). In this case, the inflorescence structures are called cymes. Cyme inflorescences lack a main axis: the main shoot terminates in a flower, while growth continues through lateral axes produced below the terminal flower. These lateral axes again form terminal flowers and this process is reiterated several times. Cymose inflorescences display structural variation, from a simple form as seen in *Silene latifolia*, tobacco (*Nicotiana tabacum*) to a sympodium as seen in tomato (*Solanum lycopersicum*) (Benlloch *et al.*, 2007). This type of inflorescence may be terminated by a flower (as in pepper or petunia), by five to six flowers (as in tomato), or by dozens of flowers (as in the Chilean potato vine) (Hake, 2008).

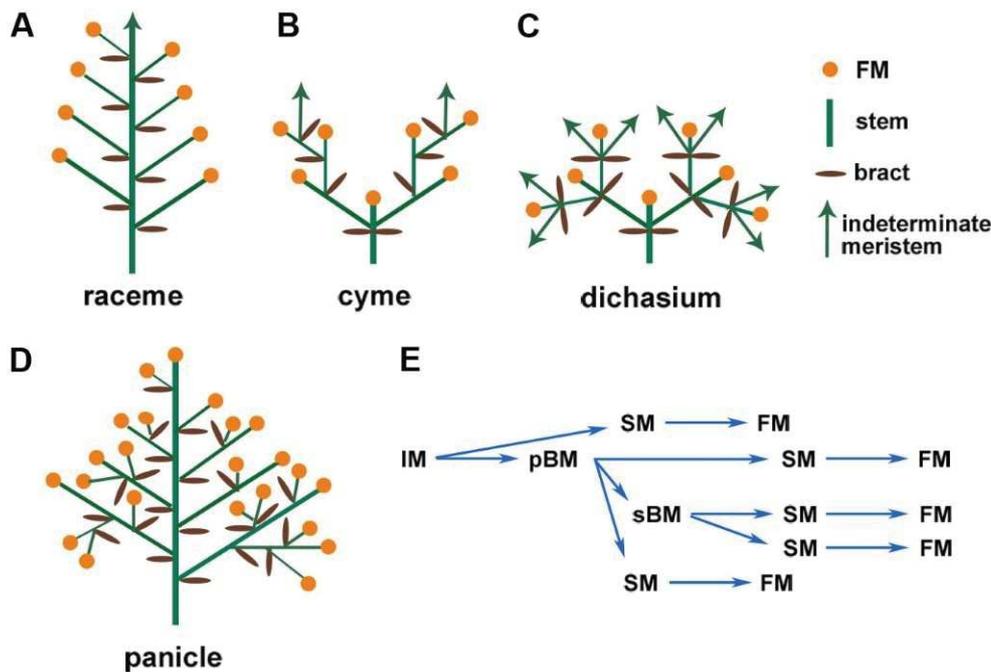
In contrast to this initiation pattern, the floral meristem and branching meristem can also initiate laterally from a terminal flower, either from the axil of a leaf-like organ (such as petals) or they can initiate without subtending lateral organs. These types of inflorescence are termed a dichasium and a pleiochasium (**Fig. 11C**), depending on the number of lateral branches, and can be considered a specialized cyme.

The other determinate inflorescence architecture type is the panicle (**Fig. 11D**). In contrast to the cyme, in this type of inflorescence a clear main shoot axis exists but it is

terminated by a floret meristem (Benlloch *et al.*, 2007; Zachary B Lippman *et al.*, 2008).



**Figure 10. Structures of different types of inflorescences.** Images from plant species representative of main inflorescence types and the corresponding diagrams of the architecture of their inflorescences. Open circles represent flowers and arrows represent indeterminate shoots. Adapted from Benlloch *et al.*, 2015. (A–E) Determinate inflorescences: (A–B) panicle type (A: panicle; B: thyrsoid); (C–E) cyme type (C: dichasium; D: monochasium, E: triad); (F–Q) Indeterminate inflorescences (the raceme type) (F) spike; (G) raceme; (H) panicle-like; (I) thyrsoid; (J) umbel; (K) corymb; (L) solitary on a scape; (M) solitary in axils of leaves; (N) spikelet; (O) Capitulum (P) head with small receptacle; (Q) spadix; (R) cyathium. Adapted from <http://plantnet.rbg.gov.au/>.

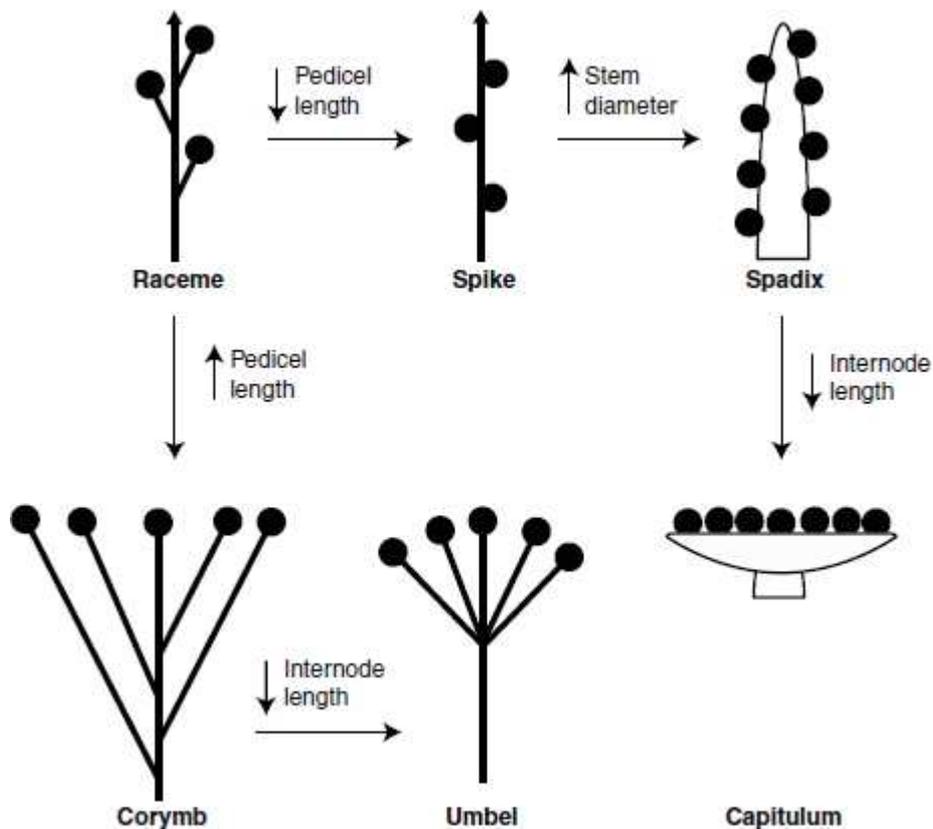


**Figure 11. Schematic representation of common types of inflorescences.** (A) Simple raceme, which is indeterminate and unbranched; (B) cyme; (C) dichasium, which are determinate and branched; (D) panicle, which is determinate and branched; IM, inflorescence meristem; pBM, primary branch meristem; sBM, secondary branch meristem; SM, spikelet meristem; FM, floral meristem. From Han, Yang and Jiao (2014).

### 2.2.2. Internode length effect on inflorescence architecture

The node is the area of a stem from which one or more organs such as leaves, roots, branches or flowers (in the case of the inflorescence) grow; whereas the internode is the distance between two successive nodes on the stem axis. In inflorescence architecture, internode length is contributory factor to diverse inflorescence topologies. In rosette plants such as *A. thaliana*, the transition from vegetative stage to reproductive stage is accompanied by internode elongation. The marked difference in internode length between the vegetative and reproductive stems distinguishes the height of plant and determines whether the flowers are presented to pollinators. Moreover, relative internode and pedicel lengths within the reproductive portion of the plant play important roles to distinguish inflorescence topologies (Fig. 12). A raceme is characterized by lateral flowers, with pedicels forming in sequential axils. They are separated by visibly identifiable internodes. Pedicel length can be completely reduced, leading to sessile flowers and to an inflorescence called a spike. A radial increase of the growth stem in this case leads to the conversion of a spike into a spadix. The spadix type may be converted into a capitulum if the internode lengths reduce. Conversely, elongation of pedicel length interacting with internode length can allow the generation of corymb and umbel

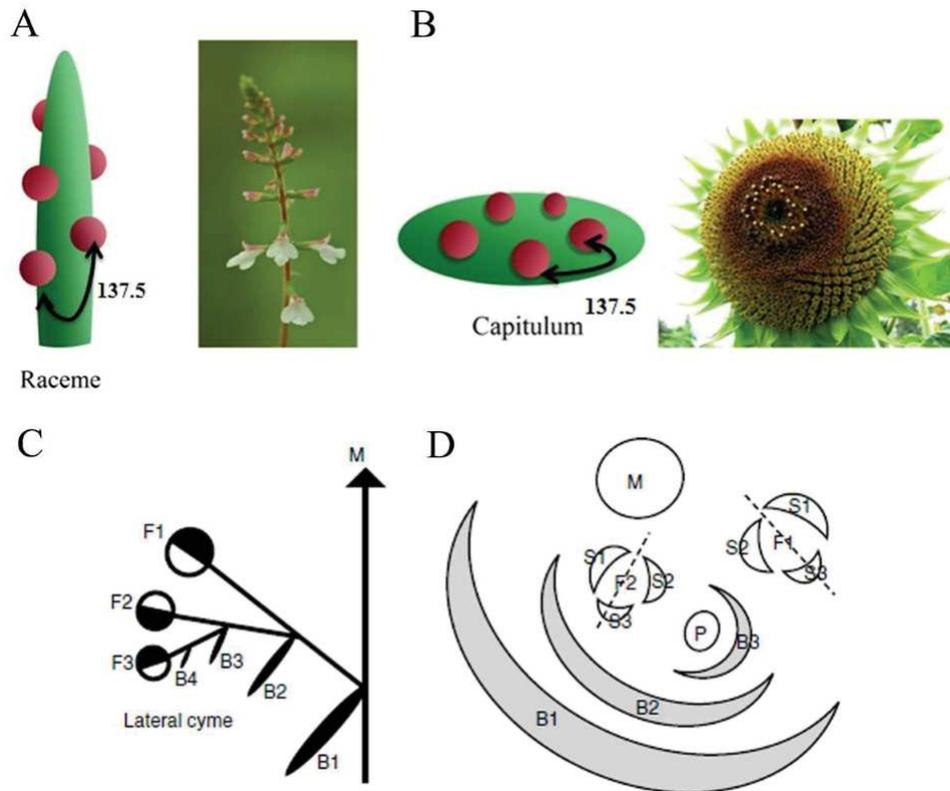
types of inflorescence architecture.



**Figure 12.** The length and diameter of stems and pedicels determine inflorescence architecture. (From Ainsworth, 2006).

### 2.2.3. Phyllotaxy of inflorescence architecture

The term of *phyllotaxis* means “leaf arrangement” in Greek. It means that if we look down from above on the plant and measure the angle formed between a line drawn from the stem to the leaf and a corresponding line for the next leaf, we will find a fixed angle, called the divergence angle. As with the leaf arrangement, phyllotactic changes allow the production of new structures in inflorescence architecture. Floral shoots or flowers that form in axils with alternate, decussate or spiral phyllotaxy contribute to inflorescences with distinctive morphologies. Further variations occur among spiral patterns that correlate with the relative rates of shoot apex growth and primordial initiation, yielding patterns corresponding to different sequential Fibonacci numbers (*i.e.* 137.5 degree divergent angle) (Richards 1951; Jean 1988). Observations of inflorescence architecture revealed that phyllotaxy not only applies to leaf arrangement, but also to organ placement in the inflorescence (Kirchoff, 2003) (Fig. 13).



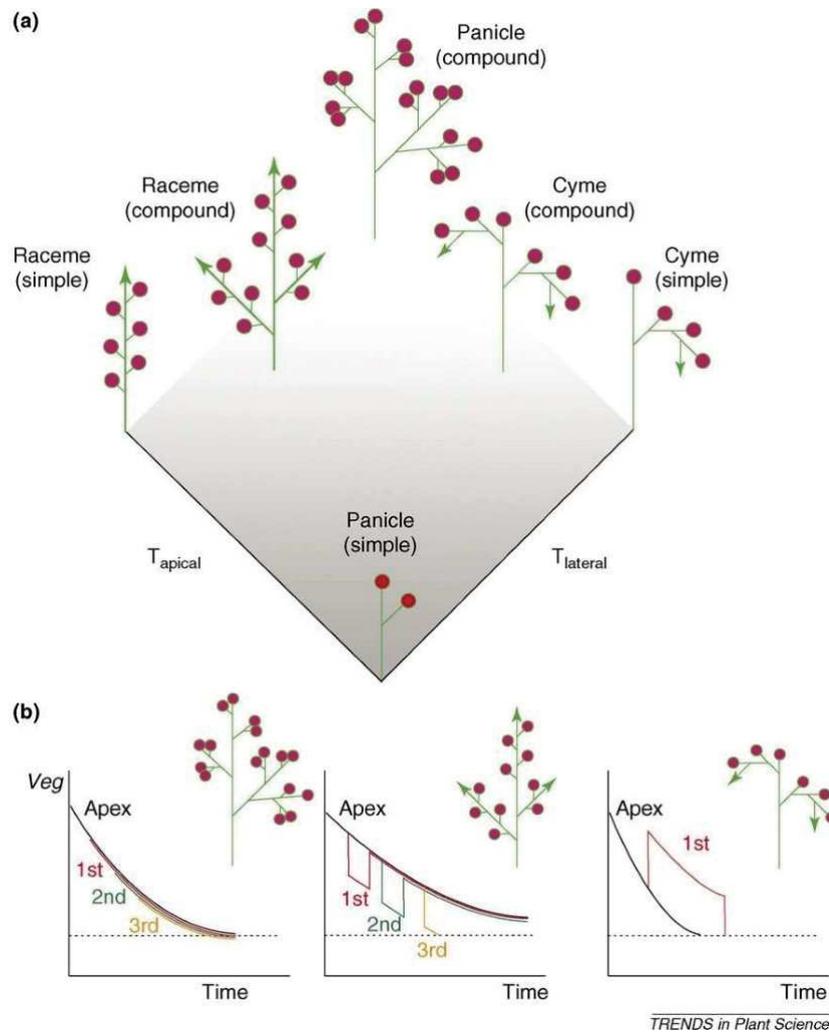
**Figure 13. Inflorescence architecture is affected by phyllotaxy of axillary meristems.** (A-B) Inflorescence architecture of raceme and capitulum are affected by phyllotaxy of axillary meristems initiation, respectively. (C) Schematic of a lateral cyme of *Phenakospermum guyannense* (Strelitziaceae). (D) Cross section of the lateral cyme shown in (C), but earlier in development. The plane of floral symmetry in these bilaterally symmetric flowers is indicated by a dashed line. B: bract; F: flower; M: terminal inflorescence meristem; S: sepal. Organs are numbered based on their order of initiation (Adapted from Ainsworth, 2006).

#### 2.2.4. Modeling of inflorescence architecture evolution

As mentioned previously, inflorescence architecture depends on when and where floral meristem identity is acquired. In Evo-Devo, it is often stated that evolutionary changes were regulated by developmental time or “heterochrony” which have been proposed to explain much of the observed morphological diversity, especially in animals (Slack and Ruvkun, 1998). Similarly, Prusinkiewicz et al., (2007) presented a model of inflorescence architecture evolution based on differences in the time required for apical and lateral meristems to acquire floral fate.

In this model, the state of meristem was defined by the factor “*Vegetativeness*” (*Veg*). If “*Veg*” is high, a meristem will produce a new lateral meristem, but if “*veg*” drops below a certain threshold, the meristem converts to a floral meristem (i.e. a determinate meristem). Depending on the tendency of “*Veg*” timing during inflorescence development, panicle, cymes

or racemes would be specified (Fig. 14). The model was supported and improved by molecular genetic analysis, which identified several genes from different species (e.g. *A. thaliana*, petunia, tomato and rice) as factors related to the “*Veg*” parameter. This will be detailed in the following section.



**Figure 14. Development of distinct virtual inflorescence structures. (A)** Structure of inflorescences and position in morphospace. Flowers are indicated by red circles and meristems by green arrows. The inflorescence types are positioned in a 2D morphospace defined by the time required for apical ( $T_{apical}$ ) and lateral ( $T_{lateral}$ ) meristems to acquire floral fate. **(B)** Expression of *Veg* in compound panicle (left), raceme (middle) and cyme (right). The black line depicts *Veg* levels in the primary apical meristem. The colored lines depict *veg* in the first (red), second (green) and third (orange) lateral meristems formed by the primary apex. Adapted From Koes (2008).

### 2.2.5. Molecular bases of inflorescence architecture

According to Prusinkiewicz’s model (Prusinkiewicz, Erasmus, Lane, Lawrence D Harder, *et al.*, 2007), the main types of inflorescence (ie. raceme, cyme and panicles) could be

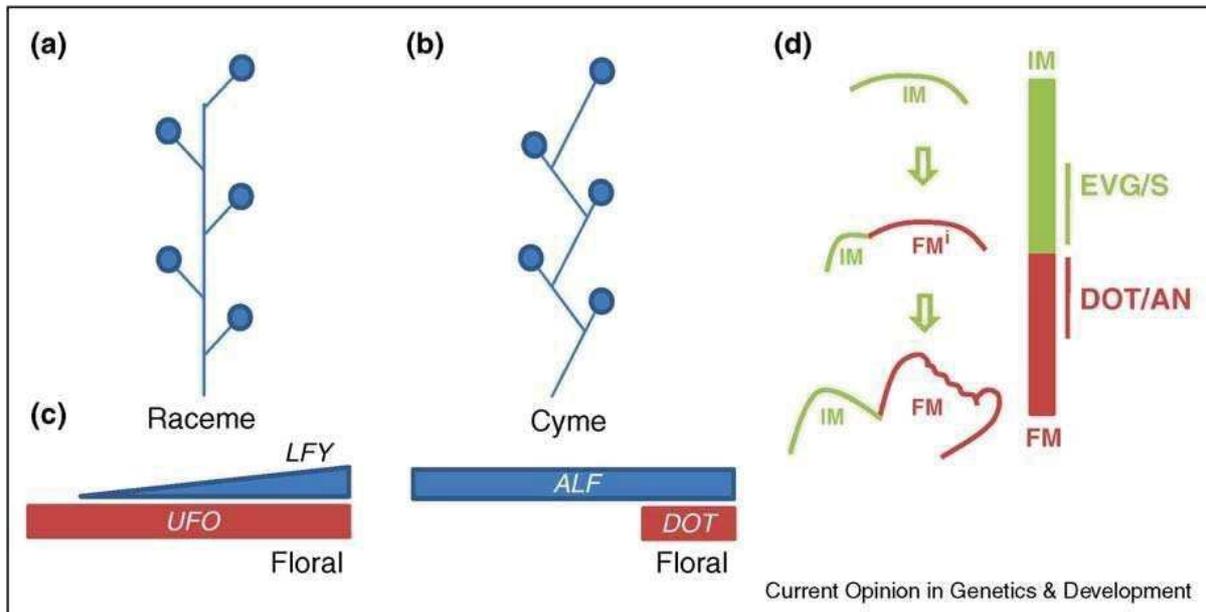
explained by differential expression of a parameter "Veg". Examples of the molecular components that may be comprised by this parameter include the conserved activities of *LEAFY* (*LFY*) and *UNUSUAL FLORAL ORGANS* (*UFO*) genes orthologs, which are regulated in a species-specific manner (Souer *et al.*, 2008; Moyroud *et al.*, 2010).

In *A. thaliana*, *LFY* encodes a transcription factor that promotes the transition from the inflorescence meristem (IM, indeterminate fate) to the floral meristem (FM, determinate fate) (Jack 2004; Irish 2010), whereas *TERMINAL FLOWER1* (*TFL1*) suppresses this process (Shannon and Meeks-Wagner, 1991; Bradley *et al.*, 1997). Considering Prusinkiewicz's model, *LFY* represses "Veg" factor while *TFL1* promotes "Veg". Because *LFY* and *TFL1* down-regulate each other (Jack, 2004), *TFL1* is expressed in the apical meristem (Conti and Bradley, 2007) whereas *LFY* is expressed in lateral floral meristems. These findings, as well as additional genetic data, have been incorporated into the model to compute "Veg", resulting in an *A. thaliana* inflorescence that recreates the wild-type architecture, in addition to single and double mutants containing gain and/or loss-of-function alleles of *TFL1* and *LFY* (Prusinkiewicz, Erasmus, Lane, Lawrence D Harder, *et al.*, 2007). The results suggest that expression of *LFY* and *TFL1* during inflorescence development lead to the raceme-type of *A. thaliana*.

In contrast, the petunia *LFY* ortholog, *ABERRANT LEAF AND FLOWER* (*ALF*) is expressed in both vegetative and reproductive stages (Souer *et al.*, 1998; Molinero-Rosales *et al.*, 1999). Whereas in tobacco, which is closely related to petunia, constitutive expression of *LFY* results in a solitary terminal flower (Ahearn *et al.*, 2001), indicating that the activity of *LFY* plays important role for the formation of a cyme as predicted by the theoretical model (Koes, 2008). In addition, the ortholog of the *A. thaliana* *UFO* gene in petunia, namely *DOUBLE-TOP* (*DOT*), plays an important role to identify FM in this species. *DOT* and *UFO* genes encode F-box proteins that interact with *ALF* and *LFY* in petunia and *A. thaliana* respectively to regulate homeotic gene expression in flowers (Samach *et al.*, 1999; Souer *et al.*, 2008). Thus, *ALF* and *LFY* as well as *DOT* and *UFO* are functionally similar proteins, but they acquired widely divergent expression patterns, which seems to have been a key factor in the evolution of the distinct raceme type in *A. thaliana* and the cyme type of the petunia inflorescence (**Fig. 15**) (Souer *et al.*, 2008).

Interestingly, the orthologs of *LFY* and *UFO* were found also in grasses including rice, but their functions were not similar to those found in eudicot species. Indeed, *LFY* and *UFO* orthologs in grasses suppress the transition from IM to FM to determine inflorescence

morphology (Ikeda *et al.*, 2007; Rao *et al.*, 2008). This finding suggests a conserved mechanism for these genes among grass species (McKim and Hay, 2010).

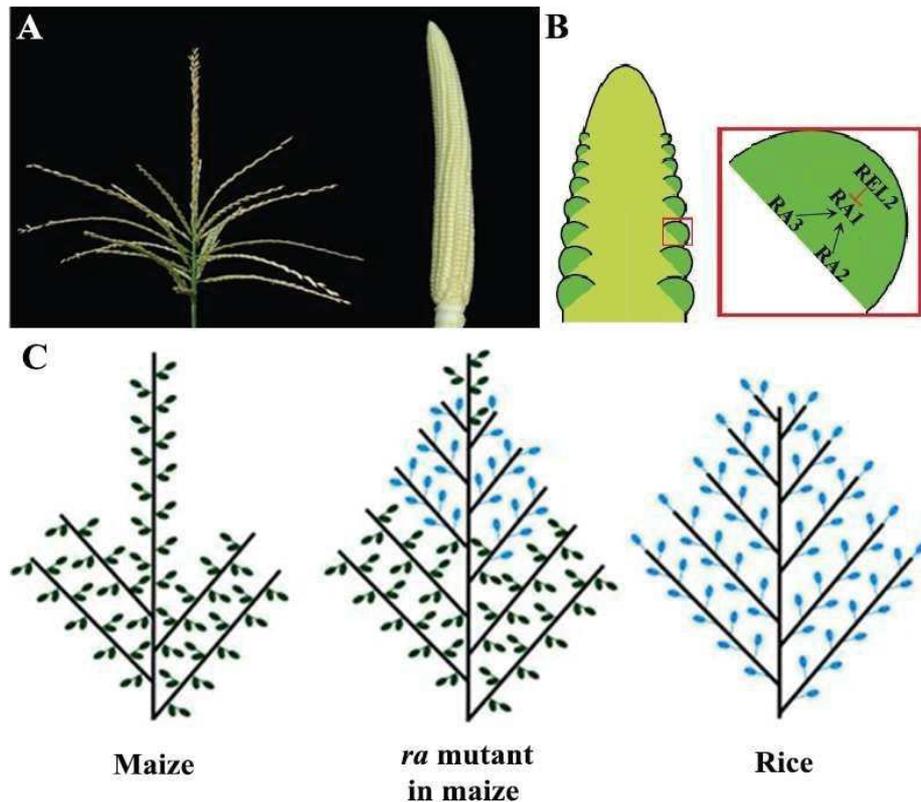


**Figure 15. Modulation of a conserved mechanism regulates diverse flowering architectures.** (A) Raceme and cyme type of inflorescences. Flowers (shown as blue circles) arise laterally from an apical IM in racemes. (B) In cymes (petunia), each flower originates as a lateral IM that transits into a FM after producing a new lateral IM, which repeats this branching pattern to generate a zig-zag pattern. (C) Overlapping expression of *LFY/UFO* and *ALF/DOT* specifies floral meristem identity in both racemes and cymes. This interaction is determined in *A. thaliana* raceme by *LFY* expression, and in petunia cyme by *DOT* expression. (D) Transient expression of *EVG* and *S* in IMs of Petunia and tomato, respectively, is required for lateral IM branching, which promotes expression of *DOT* and *AN* in apical FMs of Petunia and tomato, respectively. IM, inflorescence meristem; FMi, immature and FM, mature floral meristem (From McKim and Hay 2010).

At this point it must be asked: how did the conserved mechanisms diverge during evolution, and to what extent did this divergence contribute to the evolution of distinct inflorescence architectures? A study in petunia identified the *EVERGREEN* (*EVG*) gene, which encodes a WOX (WUSCHEL-related homeobox) transcription factor essential for activation of the *DOT* gene and for the specification of floral identity (Laux *et al.*, 1996; Rebocho *et al.*, 2008). Unlike the *WOX* orthologs in *A. thaliana*, *EVG* is expressed exclusively in the inflorescence meristem and is switched off when *DOT* is up-regulated (Rebocho *et al.*, 2008). However, the temporal expression of *EVG* is conserved in other cymes such as tomato, where *COMPOUND INFLORESCENCE* (*S*, the *EVG* ortholog) is expressed in the inflorescence meristem until *ANANTHA* (*AN*, the *DOT* ortholog) is expressed to identify the floral meristem (Zachary B. Lippman *et al.*, 2008) (Fig. 15D). Because *EVG/S* and *DOT/AN* are sequentially expressed during the gradual phase transition from inflorescence meristem to floral meristem,

the loss of function of either gene delays flower formation, resulting in additional branching and a loss of FM identity in the apex. Lippman et al. (2008) suggested that the transient nature of the expression of these genes provides a flexible mechanism to modulate the duration of the inflorescence phase before commitment to the floral fate, and hence to generate species-specific cyme architectures.

The other mechanism controlling inflorescence architecture which is highly conserved among numerous species is the *RAMOSA* pathway. This pathway was described first in maize plants that display long branches at the base of a main spike with spikelet pairs covering the long branch and main spike, in contrast to rice which is characterised by long branches bearing a single spikelet (**Fig. 15A-C**). In the *ramosa1* (*ra1*), *ramosa2* (*ra2*) and *ramosa3* (*ra3*) mutants of maize, spikelet-pair meristems assume the identity and fate of branch meristems and give rise to highly branched inflorescences (Vollbrecht *et al.*, 2005; Bortiri *et al.*, 2006; Satoh-Nagasawa *et al.*, 2006). In the *ra1* mutant, tassel (male inflorescence) architecture has a transformation of spikelet pairs from short branches into long branches bearing single or paired spikelets (Vollbrecht *et al.*, 2005) (**Fig. 15C**), whereas the ear (female inflorescence) produces more higher-order branches in *ra1* mutants, resulting in reduced fertility (Vollbrecht *et al.*, 2005; McSteen, 2006). Compare with *ra1*, *ra2* is expressed earlier during the branching process in maize. Cross-comparison and genetic analysis suggest that *RA1* functions downstream of *RA2* and *RA3* but in a different pathway (Vollbrecht *et al.*, 2005; Bortiri *et al.*, 2006; Satoh-Nagasawa *et al.*, 2006). Taken together, *RA1*, *RA2*, and *RA3* coordinate to regulate meristem identity and determinancy in the maize inflorescence fate (**Fig. 15B**) (Vollbrecht *et al.*, 2005; Bortiri *et al.*, 2006; Satoh-Nagasawa *et al.*, 2006). In 2010, Gallavotti et al. reported a new regulator of meristem fate in maize: the *RAMOSA1 ENHANCER LOCUS2* (*REL2*) gene that physically interacts with *RA1*, indicating that this complex plays a role in repressing the transcription of target genes.



**Figure 16. Inflorescence architecture of maize and the ramosa pathway.** (A) Maize tassel (left) and ear (right); (B) *RAMOSA* pathway controlling the maize inflorescence architecture; (C) Simplified schematic of inflorescence morphology in maize and rice compared with the phenotype of the *ra* mutants in maize. Thick black lines represent the main spike and the lateral branches, green paired ovals represent paired spikelets, and blue ovals represent single spikelets. The diagram is simplified to illustrate the differences in branching pattern and presence of single versus paired spikelets but does not represent the total number of branches or spikelets. *ra*: *ramosa*, *REL2*: *RAMOSA1 ENHANCER LOCUS2* (Adapted from (McSteen, 2006; Wang and Li, 2008).

In the context of inflorescence architecture evolution, mutants with different levels of *ra1* activity produce long branches and spikelet multimers, resembling architectures of other grasses (Jacobs and Everett 2000). For example, *Miscanthus sinensis* produces a visually simple inflorescence with discrete, long branches similar to the base of the maize tassel and *Sorghum bicolor*, generates a dense, multi-branched head that resembles a *ramosa* mutant. *RA1* orthologs were identified in these species (Vollbrecht et al., 2005). Interestingly, detailed analysis of the early stages of inflorescence development indicated that *RA1* activity regulates long branch architecture similarly in these three species, by imposing spikelet pair identity on the appropriate order of meristem (Vollbrecht et al., 2005). However, in rice and other more distantly related grasses, spikelets are single, and no *RA1* homologue has been identified (Vollbrecht et al., 2005). This led to the hypothesis that the *RAMOSA* pathway and, in particular *RA1*, plays a central role in the evolution of grass inflorescence morphologies but was confined to the *Andropogoneae* tribe (Vollbrecht et al., 2005; McSteen, 2006; Kellogg, 2007).

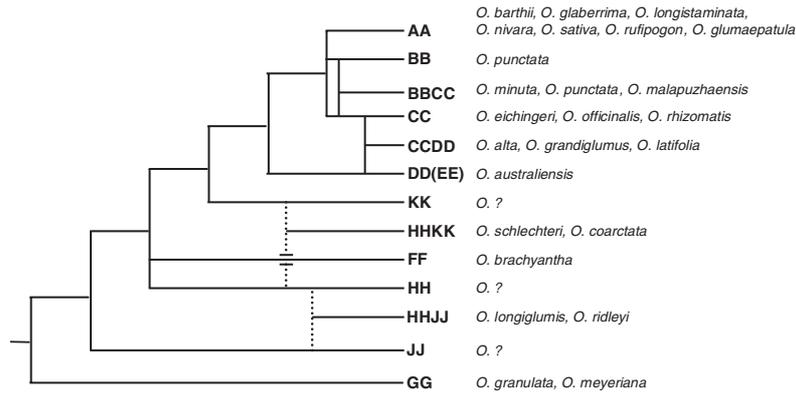
### 3. PANICLE DEVELOPMENT IN RICE

#### 3.1. THE IMPORTANCE OF RICE CROPS

The family Poaceae (grasses) includes over 10 000 species displaying an enormous diversity of morphology. This family includes many essential domesticated species, including: *Oryza sativa* (rice), *Triticum aestivum* (wheat), *Zea mays* (maize) and *Sorghum bicolor* (Sorghum). These crops feed the world through the grains produced by their inflorescence (Barazesh and McSteen, 2008). Within the grass family, rice provides the staple food for over half of the world's population and about one billion depend on rice cultivation for their livelihoods. In 2018, rice accounts for 16.5% of global caloric intake and approximately 19.7% of the world's cereal production (FAO, <http://www.fao.org/home/en/>). Moreover, in Africa, Latin America and Caribbean countries, the demand for rice is increasing day by day (IRRI, <http://irri.org/>).

Rice yields have been increasing since the 1960s, but since the 1990s, rice production has been unable follow the increase in world population. In 2009, nearly one billion people were living in poverty, including 640 million in Asia where rice is the staple food (FAO, <http://www.fao.org>). Indeed, it is anticipated that rice production will need to increase by 30% by 2025 to feed the growing population of rice consumers (IRRI, <http://irri.org/>). However, climate change, especially access to water, soil erosion, desertification, sea level rise, and other problems (pest, urbanization etc.) threaten rice production.

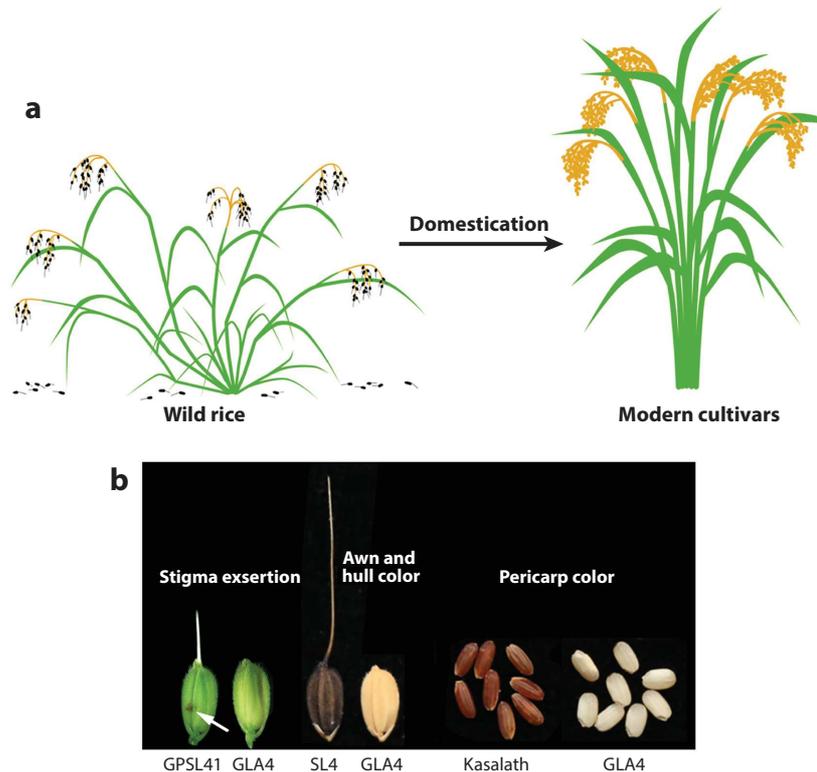
The rice genus (*Oryza*) comprises approximately 27 species and is represented cytogenetically by 11 genome types, 6 of which are diploid ( $n = 12$ : AA, BB, CC, EE, FF and GG) and 5 of which are polyploid ( $n = 24$ : BBCC, CCDD, HHJJ, HHKK and KKLL) (Stein *et al.*, 2018) (**Fig. 17**). The AA genome, also called the *Oryza sativa* complex, is represented by eight diploid species. Among them, the most extensively cultivated species is *O. sativa* L., which consists of two subspecies, *O. sativa* L. ssp. *japonica* and *O. sativa* L. ssp. *Indica* (hereafter referred to as *japonica* and *indica* respectively) which have a worldwide distribution. The other cultivated species is *O. glaberrima* Steud., commonly referred to as African rice, which is localized in West Africa (Wambugu *et al.*, 2015). While Asian rice is currently the world's second largest worldwide food crop, African rice is grown primarily in tropical West Africa.



**Figure 17. Phylogenetic tree of species from *Oryza* genus.** The genome types are indicated as a leaf tree. The corresponding species are indicated on the right. Question marks indicate categories for which no species have been reported. From (Wing *et al.*, 2005).

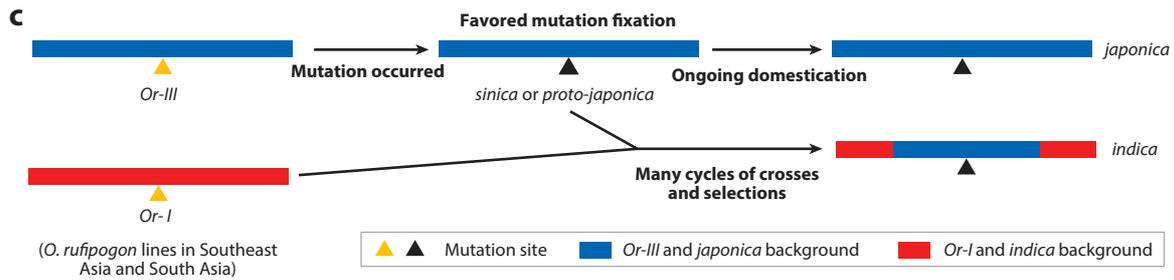
### 3.2. RICE DOMESTICATION

The genus *Oryza* has two independently domesticated species: cultivated Asian rice (*Oryza sativa*) and cultivated African rice (*Oryza glaberrima*). (Second, 1982; Vaughan, Lu and Tomooka, 2008). In both cases, domestication was observed to be associated with morphological modifications to the plants, notably with respect to flower and seed traits (**Fig. 18**). Domestication of *O. sativa* from its wild progenitor, *Oryza rufipogon*, is thought to have been initiated about 10 000 years ago (Khush, 1997; Choi *et al.*, 2017). However, the story of Asian rice domestication is still in debate, even after recent in-depth analyses of domestication sweeps and genome-wide patterns. Although molecular phylogenetic analyses indicated that *indica* and *japonica* originated independently (Londo *et al.*, 2006; He *et al.*, 2011), the well-characterized domestication genes in rice were shown to be common to both subspecies with the same alleles and a genome-wide survey of SNP polymorphisms provided stronger support for a single domestication origin of *O. sativa* (Li, Zhou and Sang, 2006; Zhang *et al.*, 2009; Molina *et al.*, 2011; Huang *et al.*, 2012; Choi *et al.*, 2017).



**Figure 18. The different phenotypes of wild rice and cultivars.** (A) Overview of wild rice and cultivars. (B) Differences between three standing domestication traits in wild rice and cultivars: stigma exertion (*white arrow*) in GPSL41 (a chromosome segment substitution line (CSSL) from progeny of *indica* variety Guangluai4 (GLA4)) and wild rice W1943, and *Oryza sativa* GLA4 ; awn and hull color in SL4 (a CSSL from progeny of *indica* GLA4 and wild rice *Oryza rufipogon* W1943) and GLA4 ; pericarp color in Kasalath and GLA4. (Chen *et al.*, 2019)

The close ancestors (*O. rufipogon*) of Asian cultivated rice are divided into three main types, *Or-I*, *Or-II*, and *Or-III*. Useful mutations may have occurred randomly in some *Or-III* populations of wild rice species that were then selected to generate the *sinica* or *proto-japonica* varieties. The *indica* varieties were subsequently developed due to acquisition of favored mutations through crosses between the *sinica* or *proto-japonica* varieties and the *O. rufipogon* *Or-I* varieties (in Southeast Asia and South Asia) after many cycles of crosses and selections. The modern *japonica* varieties were domesticated through ongoing selection. Introgression between *indica* and *japonica* of diverse natural gene variants was then followed by wide distribution to obtain adaption to local environments. (Chen *et al.* 2019) (**Fig. 19**).



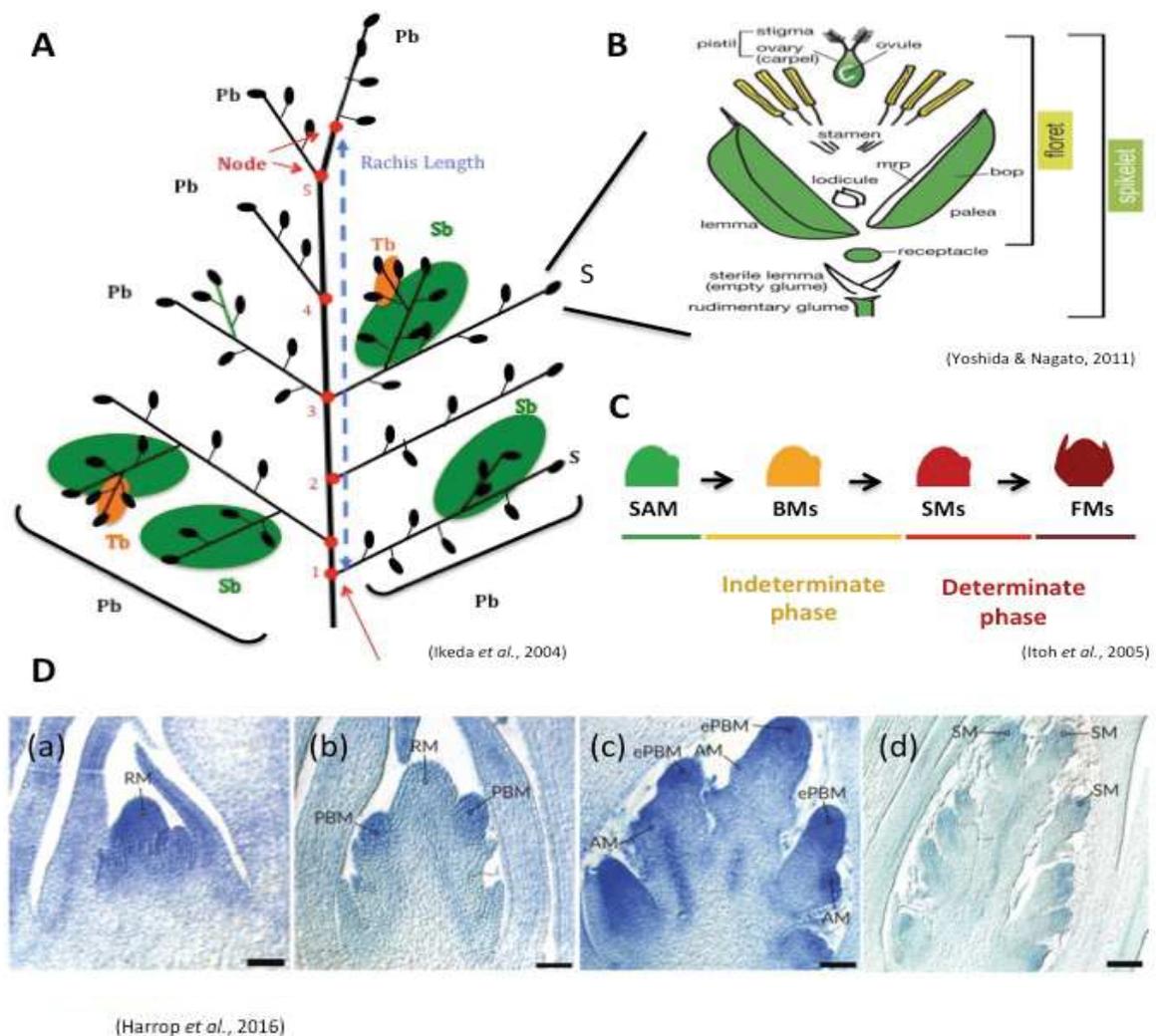
**Figure 19. The single-origin model of rice domestication.** From Cheng et al. (2019).

Meanwhile, *O. glaberrima* was domesticated from *Oryza barthii* in West Africa only about 3000 years ago (Linares, 2002a; Li, Zheng and Ge, 2011; Cubry *et al.*, 2018). The domestication of African rice originated in areas of the Upper Niger and Sahelian River (Sweeney *et al.*, 2006; Li *et al.*, 2011; Wang *et al.*, 2014; Cubry *et al.* 2018). Compared to Asian rice, the genetic diversity of African rice is considerably lower (Ishii, Xu, & McCouch, 2001; Wang *et al.*, 2014). It can be explained by the hypothesis that *O. glaberrima* is the result of a double evolutionary bottleneck. The first was associated with the divergence from Asian *Oryza*, ancestors of *O. barthii* were likely introduced from Asia to Africa. The second was caused by events happening during African rice domestication (Nabholz *et al.*, 2014).

Asian rice is currently the world's second worldwide food crop, with over 90 percent of the world's rice being produced in and mainly consumed by 6 Asian countries (China, India, Indonesia, Bangladesh, Vietnam and Japan) comprising 80% of the world's production and consumption (Abdullah, Ito and Adhana, 2006). In contrast, African rice is primarily cultivated in tropical West Africa. Due to its different origin, African rice differs from its Asian counterpart in many qualitative and quantitative traits (Vaughan, Lu and Tomooka, 2008). Nonetheless, African rice varieties possess other unique and useful traits, for instance, weed competitiveness, tolerance to various abiotic stresses (acidity, salinity and drought) and resistance to pathogens (Sarla and Swamy, 2005). The differences between Asian and African rice can be used as good sources of germplasm in breeding programs. For example, new varieties, named "New Rice for Africa" (NERICA), are derived from hybridization between *O. glaberrima* and *O. sativa*. Thanks to the hardiness of the African species and the high productivity of the Asian species, new varieties with higher yield, strong resistance to abiotic and biotic stress and the ability to adapt with the growing conditions of west Africa can now be introduced into cultivation (Linares, 2002a; Sarla and Swamy, 2005).

### 3.3. RICE PANICLE ARCHITECTURE

Inflorescence architecture is initially determined during the reproductive phase, by meristematic activities that define the branching pattern along with flower positioning (Yamburenko, Kieber and Schaller, 2017). Rice panicle development involves four types of inflorescence meristems: rachis, branch, spikelet and floral meristems. During the early stages of reproductive development, the shoot apical meristem (SAM) is converted into rachis meristem (RM). Subsequently, some cells differentiate into primary branch meristems (PBMs) in the axils of freshly developed bracts. The bract growth then terminates and the primary branches elongate (ePBMs). At the same time, during elongation, the PBM can generate axillary meristems (AMs), which may develop into secondary and higher-order branches or be directly transformed into spikelet meristems (SMs). Both PBM and secondary branch meristem (SBM) eventually form a terminal SM. Each SM produces one floral meristem (FM), which forms a single floret (**Fig. 20A-D**). The floret consists of a pair of lemmas and paleas, lodicules (equivalent to eudicot petals), stamens and a carpel (**Fig. 20B**). The RM, PBM and ePBM/AM stages are indeterminate (i.e. meristematic cells are maintained) whereas the SM is determinate in that the stem cells lose their activity (Ikeda, Sunohara and Nagato, 2004) (**Fig. 20C**).



**Figure 20. Architecture of mature panicle and early developmental stages.** (A) Schematic view of the inflorescence architecture. Pb, primary branch; Sb, secondary branch; Tb, tertiary branch, S, spikelet. (B) The spikelet structure. mrp, marginal region of palea; bop, body of palea. (C) The inflorescence meristem transition. SAM, shoot apical meristem; BMs, branching meristems; SMs, spikelet meristems; FMs, floral meristems. (D) Morphology of inflorescence development at the early stages (Harrop *et al.*, 2016). Toluidine blue-stained sections of developing panicles at (a) rachis meristem (RM), (b) primary branch meristem (PBM), (c) elongating primary branch meristem with axillary meristem (ePBM/AM), and (d) spikelet meristem (SM) stages of differentiation. Scale bars represent 50  $\mu\text{m}$  (a–c), 100  $\mu\text{m}$  (d). From Ikeda *et al.* (2014, Yoshida and Nagato (2011), Itoh *et al.* (2005) and Harrop *et al.* 2016.

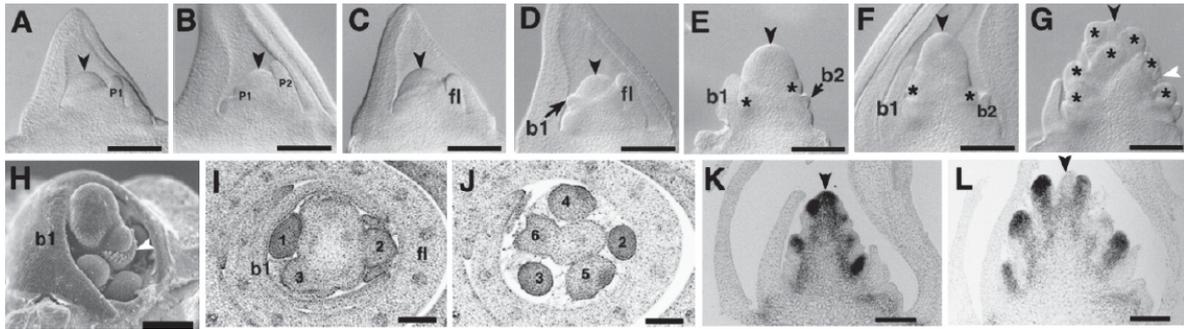
The panicle structure or complexity are modified in domesticated species and it can be seen that from a wild panicle with few primary and secondary branches bearing relatively few grains, human selection has produced a highly branched panicle carrying larger numbers of seeds than the wild ancestors. Despite the independent domestication histories of *O. sativa* and

*O. glaberrima*, most varieties of both species share a similar panicle phenotype with a higher seed number and more complex branching compared to the wild ancestors (Linares, 2002; Yamaki *et al.*, 2011; Ta *et al.*, 2017). During the independent domestications of African and Asian rice, artificial selection for improved yield led to the convergent evolution of inflorescence architectures that in turn affect grain output. Nevertheless, little is known about the molecular mechanisms underlying the phenotypic convergence between the domesticated species.

### 3.4. THE DEVELOPMENT OF THE RICE PANICLE

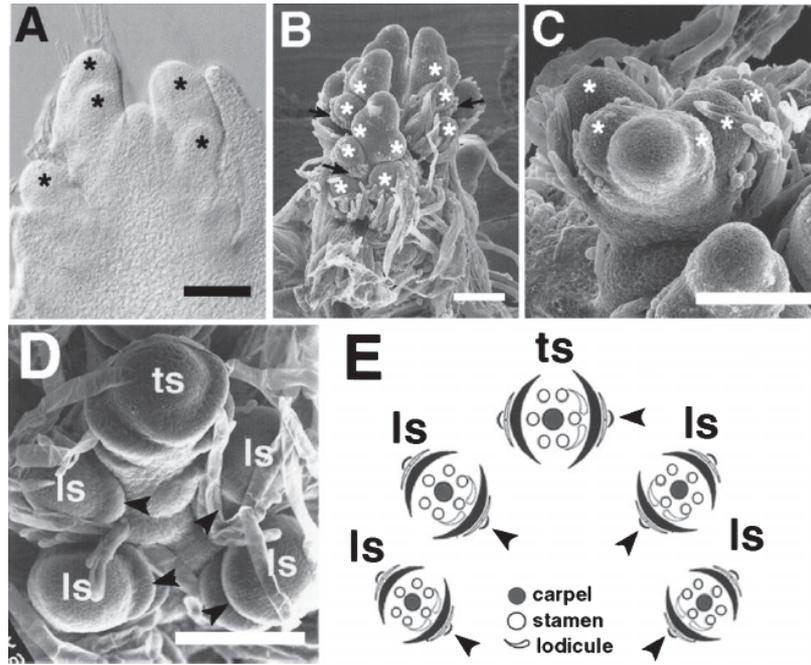
The transition from the vegetative to the reproductive phase in rice occurs when appropriate signals such as shorter day length and higher temperature have been received from the environment. After producing the last foliage leaf, the SAM is transformed into a rachis meristem. Compared with the vegetative meristem, the rachis meristem is taller and wider when producing the first bract (**Fig. 21A-C**). In the SAM, the leaf primordium 2 (P2) is longer than the shoot meristem and covers more than half of it while the primordium is formed (**Fig. 21B**). However, when the first bract primordium is formed, the tip of the flag leaf primordium is longer than the rachis meristem (**Fig. 21D**).

After the bract 2 and first primary branch are established, ten or more bracts and primary branches are rapidly initiated in spiral arrangement (**Fig. 21E-H**). In the early reproductive phase, a dramatic change from  $\frac{1}{2}$  alternate to spiral phyllotaxy occurs. This change does not happen suddenly but occurs step-by-step. The angles of the first two bracts are slightly smaller than  $180^\circ$  (**Fig. 21I**), about  $160^\circ - 170^\circ$ ; they gradually converge to  $144^\circ$  (**Fig. 21J**). The direction of spiral is either clockwise or counter-clockwise, depending on which margin of flag leaf primordium becomes inside. Usually, the rachis meristem is aborted at an early stage after producing ten or more primary branch primordia. In **Fig. 21K**, mRNAs of the rice ortholog of the *A. thaliana* *SHOOTMERISTEMLESS* (*STM*) gene *OSHI* were detected in the rachis meristem during the production of primary branch primordia (Suzaki *et al.*, 2004). After the last primary branch, *OSHI* expression is seen to disappear from the rachis meristem (**Fig. 21L**). These results indicate that the rachis meristem loses its activity after producing a cultivar-specific number of primary branches (Ikeda, Sunohara and Nagato, 2004).



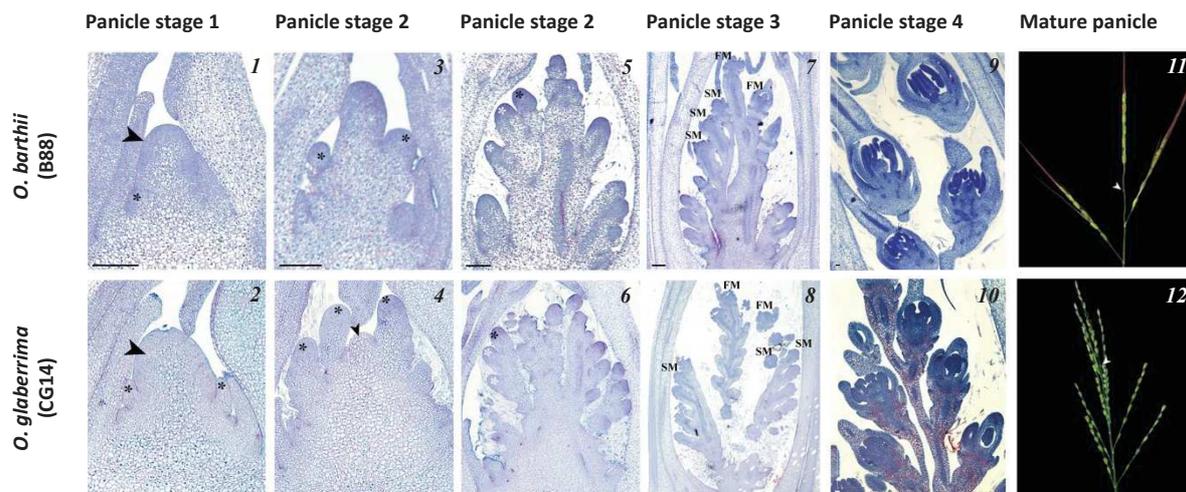
**Figure 21. Early stages of panicle development in rice.** (A)-(G) Cleared shoot and inflorescence apices, (H) SEM image of young inflorescence, (I)-(J) Rachis apex, (K)-(L) Expression of *OSHI* in rachis apex \*: primary branch meristem; fl: flag leaf, b1: first bract, b2: second bract. Bar=150 $\mu$ m for (A) to (G), 100 $\mu$ m for (H) to (L). Adapted from (Ikeda, Sunohara and Nagato, 2004).

Once all branch primordia have been initiated, the rachis meristem loses its activity and aborts. Then all branch meristems are simultaneously elongated (Fig. 22A). After the elongation of primary branches, secondary branches are formed in the basal regions (Fig. 22B). The secondary branches may produce tertiary branches. The branch meristem identity depends on that of the rachis meristem. Eventually it will be converted into a spikelet meristem (or terminal spikelet) and form rudimentary glumes while the lateral meristems become lateral spikelets. Lateral meristems of primary branches are arranged in a biased distichous phyllotaxy (Fig. 22D-E). During spikelet differentiation, two rudimentary glumes are formed, followed by the lemma and palea. Then the remaining spikelet meristem converts into a flower (floret) meristem to form floral organs (two lodicules, six stamens and one carpel) (Fig. 22E). The inflorescence remains short (<4 cm) at this stage (Itoh *et al.*, 2005). But subsequently the rachis and branches start rapid elongation after floral organ primordia are differentiated. Maturation of anthers and ovules takes place during rapid branch elongation.



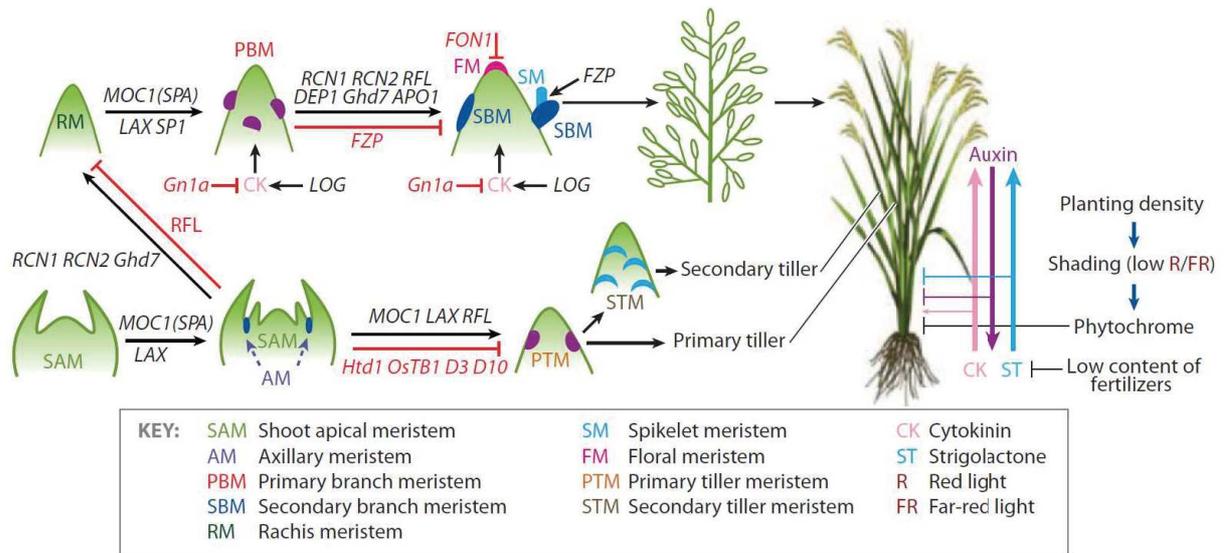
**Figure 22. Later stages of panicle development in rice.** (A) Elongation of primary branches (black star), (B) Differentiation of secondary branch primordial (white star), (C) Top view of (B) showing biased distichously phyllotaxy of secondary branch (white star), (D) Spikelet formation in primary branch apex, (E) Schematic presentation of terminal spikelet and lateral spikelet disposition in the primary branch apex. Terminal spikelet is distinguished from lateral spikelets by its direction of insertion and the position of glumes (arrowhead indicate lemma site). ts: terminal spikelet, ls: lateral spikelet, bar: 100 $\mu$ m. From (Ikeda, Sunohara and Nagato (2004).

Similarly to *O. sativa* (Ikeda, Sunohara and Nagato, 2004), the panicle development of African species was also divided into 4 stages (i.e. stage 1 to 4) (Fig. 23) with stage 1 corresponding to vegetative shoot apical meristem (SAM) shortly before phase transition. Rachis meristem (RM) formation is considered as beginning of stage 2, with a RM taller and slightly wider than vegetative SAM. In turn, the RM produces primary branch meristems (PBMs) leading to the formation of primary branches (PBs). At stage 3, PBs elongate and contribute to the higher order branches through the establishment of secondary branch meristems (SBMs). At stage 3, the spikelets and floret meristems are differentiated from all branch and axillary meristems in the panicle. At stage 4, floret organs are developed (Ta *et al.*, 2016)



**Figure 23. Histological description of selected developmental stages of African rice panicles.** *O. barthii*: 1,3,5,7,9,11; *O. glaberrima*: 2,4,6,8,10,12; stage 1: unbranched stage with elongation of rachis meristem (arrowheads) and formation of primary branch meristems (\*) (1,2); stage 2: early branching stage with rachis meristem (arrowheads) and elongating primary branches (3,4). At the end of this stage, secondary branches (white \*) are initiated from PBs (\*) (5,6); stage 3: late branching stage with elongated secondary branch and spikelet meristem (SM) and floret meristem (FM) differentiation (7,8); stage 4: floret organ differentiation/development (9,10); mature stage: 11 and 12. White arrowhead: vestige of aborted rachis meristem. Scale bar: 100  $\mu$ m. Adapted from Ta *et al.*, (2016).

To understand the structural organisation of the panicle, it is essential to understand the basic biological process of panicle development, as well as the differentiation of meristems into spikelet and floret. The development of the rice panicle in particular and grass inflorescences in general involves a variety of stage-specific landmark events (Ainsworth, 2006; Xing and Zhang, 2010; Zhang and Yuan, 2014a) (**Fig. 24**). As described above, there are three main stages of panicle development: transition, branching and spikelet differentiation. These different stages are associated with specific gene expression patterns relating to molecular mechanisms involved in the formation of meristems, cell proliferation, which in turn affects meristem size and thus eventually regulates the rate of spikelet differentiation (**Fig. 24**). The genes in question have been identified in *O. sativa* mainly through the characterisation of mutants. However, in some cases, these genes were identified by QTL characterization.



**Figure 24. A schematic representation of genes involved in panicle formation.** From Xing & Zhang (2010).

### 3.5. GENETIC CONTROL OF TILLERING

On the main stem of the plant, the tiller is a specialized grain-bearing branch formed on the un-elongated basal internodes. Primary and early secondary tillers play a role in grain yield, while the contribution of tertiary and late secondary tillers is less significant, although they also consume nutrients, water, and photosynthates (Li *et al.*, 2003). To improve the yield of cereal crops, there has been a tendency to breed for fewer panicles but more branches and therefore more grains on dwarf plants having low tiller number with therefore fewer unproductive tillers (Khush, 2001; Peng *et al.*, 2008).

Among grass species, although patterns of tillering vary, tiller formation involves a common process divided into two steps: firstly the initiation of the axillary bud; and secondly, its outgrowth (Wang and Li, 2011). In the first step, plants of the rice *monoculm1/small panicle (moc/spa)* mutant were found to form only the main shoot without any axillary meristem. *MOC1* encodes a transcription factor of the GRAS family orthologous to *LAS* of *A. thaliana* and *LATERAL SUPPRESSOR (LS)* of tomato (Li *et al.* 2003). Moreover, *moc1* and *ls* mutants reduce the number of branches and spikelets, by affecting the branching process of inflorescence development. Comparison of the *las* and *moc1* mutants suggested that *MOC1/LS/LAS* play a conserved role in the initiation and maintenance of tillers (McSteen and Leyser, 2005).

There are three classes of hormones implicated in the regulation of bud outgrowth: auxins, cytokinins and strigolactones (**Fig. 24**). These hormones are transported throughout the plant, auxin is produced mostly in the young expanding leaves of growing shoot apices and is transported basipetally down the site, through the polar auxin transport (PAT) stream and indirectly inhibits tiller initiation. Strigolactones and cytokinins are mainly produced in the root, but also locally in the shoot, and are transported acropetally in the xylem (Domagalska and Leyser, 2011). Brewer (2009) suggested that strigolactones act downstream of auxin and directly inhibit axillary bud outgrowth (Brewer et al. 2009). However, in recent study on *A. thaliana*, *max* mutants showed that strigolactones may control the outgrowth of axillary buds through PAT (Crawford *et al.*, 2010).

In maize, the *Teosinte branched1* (*TBI*) gene has been identified as a major contributor to evolutionary changes in maize that occurred during its domestication from teosinte. This gene acts to repress the growth of axillary organs and to enable the formation of the female inflorescence (Doebley, Stec and Hubbard, 1997). The rice *TBI* gene (*OsTBI*), homolog of the maize *TBI*, encodes putative transcription factor carrying a basic helix–loop–helix type of DNA-binding motif, named TCP domain. Transgenic rice plants overexpressing *OsTBI* show greatly reduced lateral branching without affecting the initiation of axillary buds, whereas a loss-of-function mutant of *OsTBI* exhibits enhanced lateral branching, indicating that *OsTBI* functions as a negative regulator for lateral branching in rice (Takeda *et al.*, 2003) (**Fig. 24**). Moreover, in the *moc1* mutant, Li and al. (2003) observed that the expression of *OsTBI* was significantly reduced suggesting that *OsTBI* acts downstream of *MOC1*. The various examples above suggest a general conservation of mechanisms controlling axillary bud activity between monocot and eudicot species. Surprisingly however, none of the vegetative branching mutants studied in grass species were reported as being affected in their inflorescence branching pattern, suggesting that these two processes have at least partly diverged in terms of their regulation.

### **3.6. AXILLARY MERISTEM INITIATION DURING INFLORESCENCE DEVELOPMENT**

The complex overall architecture of plants is mainly controlled by patterns of axis formation. Concerning the generation of axillary meristems (AMs) and the growth of branch meristems, several key regulatory genes have been identified through mutant analyses. In rice, beside *monoculm1*, other mutants such as *lax panicle1* (*lax1*), *lax panicle2* (*lax2*), and *frizzy panicle* (*fzp*) as well as two genes involved in the control of auxin signaling, *OsPIN1* and *OsPID*, have been shown to affect the patterning of AMs and also panicle development

(Komatsu *et al.*, 2001, 2003; Xu *et al.*, 2005; Woods, Hope and Malcomber, 2011)

*LAX PANICLE1 (LAX1)*, encoding a basic helix-loop-helix (bHLH) transcription factor, is essential for the formation of all AMs during vegetative development and all lateral structures during panicle development (Komatsu *et al.*, 2003; Oikawa and Kyojuka, 2009). The absence of lateral and terminal spikelet meristems in the *lax1* mutant, along with a decrease in the number of branch meristems, clearly indicates that *LAX1* is required for the initiation and maintenance of lateral meristems and terminal spikelet meristems in the rice panicle (Komatsu *et al.*, 2001, 2003). In addition, the accumulation of the LAX1 protein during axillary meristem formation is subject to a two-step regulatory process linked to a non-cell autonomous mode of action. In the first step, the *LAX1* gene is expressed in the axils of leaves at plastochron stage 4 (P4), then in the second step the LAX1 protein is trafficked to the whole axillary meristem (Oikawa & Kyojuka, 2009). Tabuchi *et al.* (2011) reported that the *lax panicle2 (lax2)* mutant displays altered AM formation. *LAX2* encodes a nuclear protein that contains a plant-specific conserved domain and physically interacts with *LAX1* and plays a role as a novel factor that acts together with *LAX1* in rice to regulate the process of AM formation. Similarly *BARREN STALK1 (BA1)*, the *LAX1* ortholog in maize, is required for the formation of all types of axillary meristems throughout plant organogenesis. The *ba1* mutants lack tillers and female inflorescence branches (ears), and the male inflorescence (tassel) is unbranched, shorter than wild-type, and almost completely devoid of spikelets (Ritter, Padilla and Schmidt, 2002; Gallavotti *et al.*, 2004). By using phylogenomic and comparative expression analyses, Woods *et al.* (2011) studied the expression of genes belonging to the *BA1/LAX1* clade in both monocots and eudicots. The results obtained suggested a conserved regulatory mechanism involving *BA1/LAX1* genes during AM formation and inflorescence development in diverse flowering plants, but with differential timing of expression between monocots and eudicots (Woods *et al.*, 2011).

Another important signal in the initiation of AMs and lateral organs is auxin. The distribution of auxin is controlled by a combination of polar auxin transport (PAT) and localized auxin biosynthesis. PAT requires polar localization of the *PINFORMED (PIN)* family of auxin efflux carriers (Zažímalová *et al.*, 2007). In rice, *OsPIN1* also functions in PAT, and the over-expression or suppression of *OsPIN1* expression through a transgenic approach was found to result in changes to tiller numbers and shoot/root ratio (Xu *et al.*, 2005). The serine/threonine protein kinase PINOID (PID) has been shown to regulate the localization of PIN proteins in *A. thaliana* (Friml *et al.*, 2004). It was reported that *PID* carries out its function

through the control of the sub-cellular localization of PIN proteins, which direct the flow of active auxin transport. Morita & Kyoizuka (2007) demonstrated that over-expression of *OsPID* caused a variety of abnormalities in rice development that could be mimicked by NPA treatment, suggesting that the defect was probably caused by disturbance of PAT and that *OsPID* is involved in the control of auxin fluxes. Mutants with defects in *PIN1* or *PID* genes have similar phenotypes (Morita and Kyoizuka, 2007).

In another mutant named *frizzy panicle (fzp)*, the formation of florets is replaced by sequential rounds of branching, such that several rudimentary glumes are formed in place of the spikelet (Komatsu et al. 2001). All meristems remain undifferentiated during early development. Therefore, the degeneration of AMs may occur during the maturation stage when internodes elongate. Not all AMs of the *fzp* mutant do not develop into a branch, for reasons that remain unclear. The *fzp* mutant phenotype suggests that *FZP* is required to prevent the formation of AMs within the spikelet meristems and that it permits the subsequent formation of branch meristem identity. *FZP* encodes a protein from the APETALA2/ ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR (AP2/ERF) family and is the rice ortholog of the maize *BRANCHED SILKLESS1 (BD1)* gene, which controls spikelet meristem formation in this species. The AP2 domain region is strongly conserved between different grass species and the corresponding genes are expressed in a distinct domain of the spikelet meristem. Its expression pattern suggests that signalling pathways regulate meristem identity from the lateral domains of the spikelet meristem (Komatsu *et al.*, 2003).

### **3.7. AXILLARY MERISTEM OUTGROWTH DURING INFLORESCENCE DEVELOPMENT**

During grass inflorescence development, the transition from inflorescence/branch meristems to spikelet meristems is one of the key events that determines the final inflorescence architecture. The regulators of this transition in rice include the *ABERRANT PANICLE ORGANIZATION 1 (APO1)*, *APO2* and *TAWAWAI (TAW1)* genes. The *aberrant panicle organization1 (apo1)* mutant forms small panicles with reduced numbers of branches and spikelets. In addition, the *apo1* mutant exhibits abnormal floral organ identity and a loss of floral determinacy (Ikeda, Sunohara and Nagato, 2004). The phenotype of the *apo1* mutant indicates that the *APO1* gene suppresses precocious conversion of branch meristems to spikelet meristems, thus ensuring a sufficient number of spikelets in the final structure. The overexpression of *APO1* genes causes large panicles with an increased number of spikelets, the panicle size being highly correlated with the expression level of *APO1*. This difference is

caused by different rates of cell proliferation. Collectively, these results suggest that the level of *APO1* activity regulates the panicle architecture through control of cell proliferation in the meristem (Ikeda *et al.*, 2007). In addition, *APO1* also controls the plastochron and the formation of floral organs. *APO1* encodes a F-box protein, an ortholog of *UNUSUAL FLORAL ORGAN (UFO)* from *A. thaliana*. *UFO* is proposed to activate *LEAFY (LFY)* to promote FM fate through direct protein-protein interaction in a proteasome-dependent manner (Chae *et al.*, 2008).

The *RFL* gene (renamed *APO2*) was identified as the ortholog of *LEAFY/FLORICAULA* gene (*LFY/FLO*) from *A. thaliana* and *A. majus* respectively. As in *A. thaliana*, *APO2* interacts with *APO1* at the protein level to cooperatively play important roles in panicle development by regulation of transition meristem fate. However, *APO2/APO1* and *LFY/UFO* act oppositely on inflorescence development. While *APO2/APO1* are expressed in incipient lateral branch primordial and suppress the transition from inflorescence meristem to floral meristem (Kyoizuka *et al.*, 1998; Rao *et al.*, 2008; Ikeda-Kawakatsu *et al.*, 2012), *LFY/UFO* promote the initiation of floral meristem to determining *A. thaliana* morphology (McKim and Hay, 2010). The finding suggests that genetic mechanisms for controlling inflorescence architecture have evolutionarily diverged between rice (monocots) and *A. thaliana* (eudicots).

*TAW1* encodes a nuclear protein belonging to the ALOG [*Arabidopsis* LIGHT-DEPENDENT SHORT HYPOCOTYLS 1 (LSH1) and *Oryza* LONG STERILE LEMMA 1 (G1)] family, conserved in monocots and eudicots. Similarly to *APO1* and *APO2*, *TAW1* regulates rice inflorescence shape by suppressing the transition from inflorescence/branch meristem to spikelet meristem (Yoshida *et al.*, 2013). The dominant gain-of-function mutant *tawawa1-D* exhibits prolonged inflorescence meristem activity and delayed spikelet specification, causing prolonged branch development and increased spikelet numbers. In addition, Yoshida *et al.*, (2013) indicated that *TAW1* induces the expression of members of the *SVP* subfamily of MADS-box genes, including *OsMADS22*, *OsMADS47* and *OsMADS55*. Although the protein has no known functional domains, *TAW1* may function as a unique transcription regulator in promoting inflorescence meristem activities and limiting the phase change to spikelet meristem.

According to the concept of apical dominance, the relationship between shoot growth and branching is regulated by a balance between auxin (which inhibits the growth of axillary buds) and cytokinin (which relieves the inhibition) (Barazesh & McSteen, 2008; Zhang and

Yuan, 2014a). Such phytohormone balance also regulates panicle branching. Molecular cloning and analysis of a QTL for grain number, *GRAIN NUMBER1 (Gn1a)*, demonstrated the role of cytokinin in controlling panicle size (Ashikari *et al.*, 2005). *Gn1a* encodes a cytokinin oxidase/dehydrogenase (*OsCKX2*), an enzyme that degrades cytokinin. Reduced expression of *OsCKX2* allows cytokinin accumulation, leading to an increase in spikelet number and thus an enhancement of grain yield. Similarly, homologs of *OsCKX* in barley, wheat, and triticale are associated with the regulation of inflorescence size and spikelet numbers (Zalewski *et al.*, 2010; Zhang *et al.*, 2012). In contrast, the *lonely guy (log)* mutant, that has a defect in synthesis of active cytokinins, produces a much smaller panicle than the wild type (Kurakawa *et al.*, 2007). The *LONELY GUY (LOG)* gene encodes an enzyme that catalyses the final step of cytokinin biosynthesis within the meristem. *LOG*, which is expressed in 2 or 3 layers of cells at the top of the meristem, is thought to regulate shoot meristem maintenance. These findings reflect a conserved role of cytokinin in regulating reproductive meristem size and activity, and an indirect effect on branching in monocots and eudicots. Moreover, they suggest that cytokinin metabolism and signalling in grasses contribute to grain yield and they offer a strategy for breeding programs to improve crop yield (Ashikari *et al.*, 2005).

Two other genes characterized as QTLs associated with grain yield are *DENSE AND ERECT PANICLE (DEP1)* and *IDEAL PLANT ARCHITECTURE1 (IPA1)/WEALTHY FARMER'S PANICLE (WFP)*. *DEP1* encodes a protein of unknown function containing a PEBP (phosphatidylethanolamine-binding protein) domain and is pleiotropically responsible for three key traits (dense panicle, high grain number per panicle and erect panicle). The dominant allele at the *DEP1* locus is a gain-of-function mutation causing truncation of the *DEP1* protein, resulting in enhanced meristematic activity and cell proliferation through *OsCKX2* (Huang *et al.*, 2009), leading to a reduction of inflorescence internode length and an increase in the number of panicle branches (Huang *et al.*, 2009). *IPA1/WFP* corresponds to *OsSPL14*, an SBP-box (SQUAMOSA promoter binding protein-like) protein-encoding gene that is the target of the microRNAs *Osa-miR156* and *Osa-miR529*. Higher level expression of *OsSPL14* during the reproductive stage promotes panicle branching and higher grain yield in rice (Miura *et al.* 2010; Jeong *et al.* 2011). The *ipa1* allele harbours a point mutation within the *Osa-miR156/Osa-miR529* target site and thus perturbs the microRNA-directed regulation of *IPA1* in rice plants (Jiao *et al.*, 2010). The *wfp* allele harbours a mutation in its promoter region (Miura *et al.*, 2010). But in both cases, the resulting consequence of these mutations is a higher level of *OsSPL14* transcript accumulation leading to increased panicle branching.

### 3.8. THE TIMING OF SPIKELET DIFFERENTIATION

*TERMINAL FLOWER1 (TFL1)/CENTRORADIALIS (CEN)*-like genes play important roles in the determination of plant architecture, mainly by controlling the timing of phase transition in *A. thaliana* and *A. majus* respectively (Conti and Bradley, 2007). Mutation of *TFL1* and *CEN* converts branch meristems into terminal flowers. In contrast to these loss of function phenotypes, ectopic overexpression of *TFL1/CEN*-like genes confers reverse effects, leading to late flowering and more branches in *A. thaliana* (Nakagawa, Shimamoto and Kyozyuka, 2002). In rice, the orthologous genes of *TFL1/CEN* are *REDUCED CULM NUMBER1* and *2 (RCN1, RCN2)* which when overexpressed cause a delay in the transition to the reproductive phase of up to 2 months compared with wild-type plants (Nakagawa, Shimamoto and Kyozyuka, 2002). Detailed observations of panicle structure revealed that the phase change from the branch shoot to the floral meristem state was also delayed, leading to the generation of higher-order panicle branches. In contrast, knocking down the expression of *RCN* genes resulted in much smaller panicle size with reduced branches (Liu *et al.*, 2013). These results suggest that *RCN* genes coordinate panicle development and flowering time (Nakagawa, Shimamoto and Kyozyuka, 2002)

Through cloning and molecular analysis of a QTL for grain number, Xue *et al.* (2008) showed that *Ghd7* has large pleiotropic effects, including grain number per panicle, heading date and plant height. *Ghd7* encodes a CCT-domain protein that has crucial roles in regulating processes such as photoperiodic flowering, vernalization, circadian rhythms and light signalling. The *Ghd7* gene controls heading date under long-day conditions, through its enhanced expression, thus delaying flowering. Detailed examination of the panicle revealed that *Ghd7* changes the numbers of both primary and secondary branches. As a result, the effect on panicle complexity of the *Ghd7* mutation is associated with a change in the duration of panicle differentiation.

Members of the *AP2/ERF* gene family, such as *INDETERMINATE SPIKELET 1 (IDS1)* in maize, are important for determining the degree of ramification in branch meristems, by regulating spatial-temporal expression of spikelet meristem genes. In rice, two *AP2* genes *SUPERNUMERARY BRACT (SNB)* and *INDETERMINATE SPIKELET 1 (OsIDS1)* were established to play important roles in the determination of panicle architecture and the formation of the floral meristem (FM). *SNB* and *OsIDS1* synergistically control inflorescence meristem architecture and FM establishment (Lee and An, 2012). It was revealed that accumulation of *AP2* mRNAs is fine-tuned by *Osa-miR172* miRNA-mediated regulation to

establish the correct spatial arrangement of the floral meristem (Chuck *et al.*, 2007). *SNB* and *OsIDS1* are expressed throughout the branch and spikelet meristems. When these meristems are initiated, *Osa-miR172* expression is recruited, resulting in a depletion of *SNB* and *OsIDS1* transcripts, ensuring spikelet development at the correct position and time (Lee and An, 2012). These data demonstrate the importance of the interaction between *Osa-miR172* and *AP2/ERF* family genes in the determination of FMs.

The two rice *API-like* genes, named *OsMADS14* and *OsMADS18*, play a role in specifying floral meristem identity (Jeon *et al.*, 2000; Fornara *et al.*, 2004). Yet the distinct mechanism through which these genes act is poorly understood. Other members of the MADS box protein family, belonging to the SEPALLATA (SEP) clade, have been studied (Fornara *et al.*, 2004; Malcomber and Kellogg, 2004; Gao *et al.*, 2010). All five SEP subfamily genes in rice are expressed exclusively during panicle development, while their spatial and temporal expression patterns vary. *PAP2/OsMADS34* expression starts the earliest among the five *SEP* genes and a low but significant level of *PAP2* mRNA is detected in the panicle meristem, in branch meristems immediately after the transition, and in glume primordia, consistent with it playing a role in the early development of the spikelet. It has been reported that *PAP2/OsMADS34* controls the transition to spikelet meristems (Kobayashi *et al.*, 2009; Gao *et al.*, 2010). Mutations in *PAP2* cause a disorganized pattern of panicle branching and a reduction in competency to become an SM, resulting in the transformation of early arising spikelets in to branch meristems. These phenotypes are consistent with its expression pattern starting from the early stages of rachis meristem development and suggest a role as a positive regulator of spikelet meristem identity (Kobayashi *et al.*, 2009). Although knockdown of the three *API-like* genes (*OsMADS14*, *OsMADS15* and *OsMADS18*) does not significantly affect inflorescence development, elimination of *PAP2* function in the triple knockdown plants was found to severely impede the transition of the SAM to the IM, suggesting a combined action of the three *API-like* genes with *PAP2*. Furthermore, the precocious flowering phenotype caused by the overexpression of *Hd3a*, a rice florigen gene, was weakened in *pap2-1* mutants. Collectively, these data suggest that *PAP2* and the three *API-like* genes act redundantly in the meristem to specify the identity of the IM downstream of the florigen signal.

### 3.9. FLORAL ORGAN PATTERNING

From the 1980s onwards, studies of the earliest flower mutants in the two model eudicot species, *A. thaliana* and *A. majus*, allowed the conception of a molecular model of specification of floral meristem and organ identity, called the ABC model (Coen and Meyerowitz, 1991).

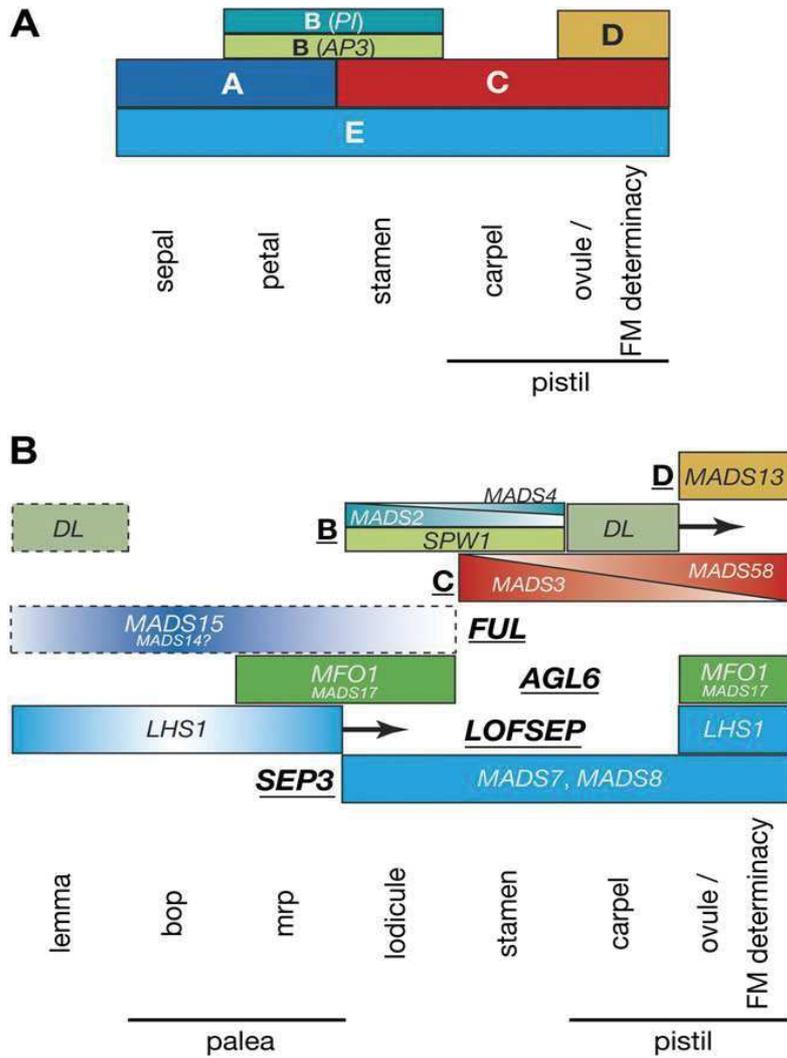
The ABC model assigned overlapping domains of activity to three different classes of genes. Sepals and petals were affected by the expression of class A genes, petals and stamens by class B genes, whereas stamens and carpels were identified by class C genes (Carpenter and Coen, 1990; Coen and Meyerowitz, 1991) (**Fig. 25A**). Subsequently the model was further extended to the ABCDE (or ABCE) model, to include D-class genes proposed as ovule identity genes and E class genes corresponding to *SEP*- and *AGL6*-like genes. E class genes function broadly across the floral meristem to support the function of ABC class genes (Theissen and Saedler, 2001; Causier, Schwarz-Sommer and Davies, 2010; Rijpkema *et al.*, 2010) (**Fig. 25A**).

In *A. thaliana*, the A class genes are represented by *APETALA1* (*API*) and *APETALA2* (*AP2*), which determine the initiation of sepal and petal. However, little is known about the *API* and *AP2* homologs of rice. The B class genes, *APETALA3* (*AP3*) and *PISTILLATA* (*PI*), are responsible for the establishment of petal and stamen identity in the second and third whorls, respectively. Functional studies of *SUPERWOMANI* (*SPW1*), the *AP3* ortholog in rice, and *OsMADS2* and *MADS4*, the *PI* orthologs in rice, indicated that these genes are necessary for lodicule and stamen identity (Nagasawa *et al.*, 2003; Kanno *et al.*, 2007; Soltis *et al.*, 2007). Similar results have been obtained from genetic analysis of the *AP3* ortholog in maize. Taken together, the finding suggests that the functions of *AP3* and *PI* in class B are strongly conserved among grass species (Causier, Schwarz-Sommer and Davies, 2010; Rijpkema *et al.*, 2010).

The *AGAMOUS* C class gene in *A. thaliana* (*AG*) is necessary for stamen and carpel identity, but is also required to specify the determinacy of the floral meristem (Kramer, Jaramillo and Di Stilio, 2004). One of the paralogs that has been studied in rice is *OsMADS3*, which is expressed in developing stamens and carpels and promotes the determinacy of the floral meristem (Kyoizuka and Shimamoto, 2002; Li *et al.*, 2011). The other gene in rice is *DROOPING LEAF* (*DL*), a member of YABBY gene family, involved in lemma specification whereas its homolog in *A. thaliana*, namely *CRABS CLAW* (*CRC*), determines carpel identity (Nagasawa *et al.*, 2003; Li *et al.*, 2011), raising the possibility that organ identity functions can shift between non-homologous loci.

The D class in *A. thaliana* corresponds to the gene *SEEDSTICK* (*STK*) while the E class is comprised of a set of four paralogs known as *SEPALATA1* (*SEP1*), *SEP2*, *SEP3*, and *SEP4*, encoding proteins that are cofactors in complexes with other MADS box factors that determine floral organ identities and meristem determinacy (Pelaz *et al.*, 2000). The expression patterns of *SEP* genes are diverse and highly variable, and functional data has been difficult to obtain, most likely due to extensive redundancy. However, the rice *LEAFY HULL STERILE*

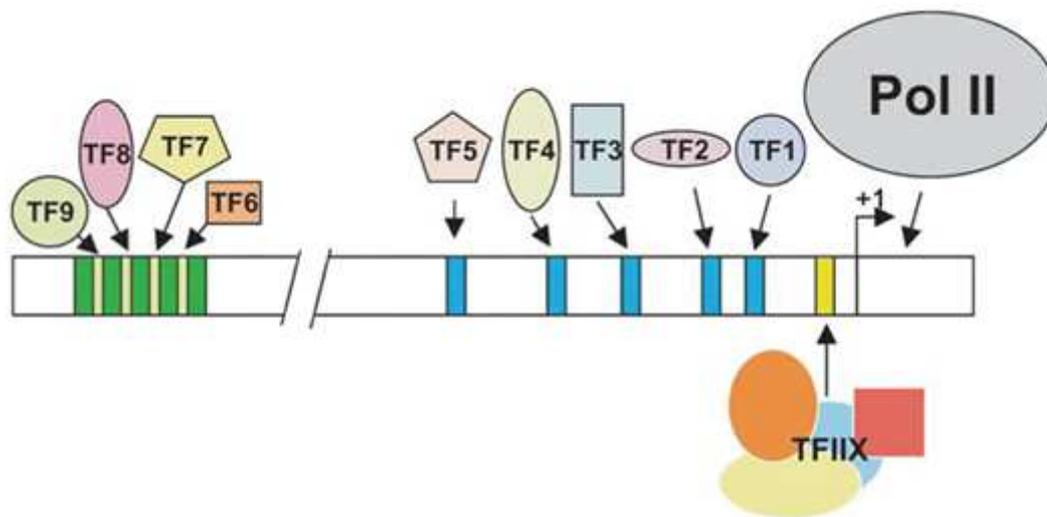
*(LHS1)/OsMADS1* gene, belonging to a subgroup of *LOF-SEP* genes, has been shown to contribute to regulation of the identity of the palea and lemma as well as to meristem determinacy and to the structure of the inflorescence (**Fig. 25B**) (Prasad, Parameswaran and Vijayraghavan, 2005; Jeon, Lee and An, 2008). (Khanday, Ram Yadav and Vijayraghavan, 2013) indicated that *LHS1/OsMADS1* integrates transcriptional and signalling pathways to promote rice floret specification and development by negatively regulating *PAP2/OsMADS34*. In early stages of panicle development, *LHS1/OsMADS1* promotes the transition from branch meristem to spikelet meristem. *LHS1* also regulates auxin transport, auxin signalling, auxin-dependent expression and three cytokinin A-type response regulators (Khanday, Ram Yadav and Vijayraghavan, 2013). A broad comparative study of the expression of *LHS1* orthologs across the grasses revealed a high degree of variability in their individual patterns, both within and between florets (Malcomber and Kellogg, 2004).



**Figure 25. Genetic models for flower development in *A. thaliana* and rice.** (A) Combinatorial actions of A-, B-, C-, D-, and E-function specify sepal, petal, stamen, carpel, and ovule identities in *A. thaliana*. Heterodimerization of *AP3* and *PI* is necessary for B-function. (B) Combinatorial actions of A-, B-, C-, D-, and E-function specify sepal, petal, stamen, carpel, and ovule identities in rice. Underlined bold letters represent gene classes/clades. Arrows indicate non-autonomous effects of *DL* and *LHS1* on FM determinacy and lodicule development, respectively. The dark blue-dotted box indicates potential function of FUL-like genes although their role in lemma, palea, and lodicule development has not completely been explored except for that of *OsMADS15* in palea development. The light green dotted box indicates expression of *DL* in lemma, although its function in lemma development is unknown. mrp, marginal region of palea; bop, body of palea; FM, floral meristem. From (Yoshida and Nagato (2011).

#### 4. TRANSCRIPTION FACTORS IN THE PLANT KINGDOM

In eukaryotes, the transcription of nuclear genes is carried out by various different RNA polymerases (Roeder and Rutter, 1969). More specifically, RNA polymerase II is responsible for the transcription of not only all protein-coding genes but also of some that specify non-coding RNAs (Kornberg, 2007; Cramer *et al.*, 2008). RNA polymerase II promoters typically consist of different discrete DNA sequences (also named boxes or elements) of functional importance, usually found upstream of the transcription start site (Lenhard *et al.*, 2012). Based on their position or function, these sequences can be classified as basal promoter elements, upstream promoter elements, and enhancers. To affect the transcription of a gene, the corresponding transcription factors bind to RNA polymerase II promoters as described in **Figure 26**.



**Figure 26. The structure of eukaryotic promoters.** Eukaryotic gene promoters are composed of discrete binding sites for multiple transcription factors dispersed over long distances (usually several thousands of base pairs). General transcription factors (TFIIX) for RNA polymerase II (Pol II) interact with sequences located near the transcription start site (yellow). Specific transcription factors (TF1 to TF9) recognize particular sequences located in proximal promoter regions (blue; at hundreds of base pairs of the start site) or in enhancers (green; at thousands of base pairs of the start site). The transcriptional activity of a gene will be defined by the nature of the transcription factors bound in different regions of its promoter. The transcription start site is indicated by +1. Figure adapted from Gonzalez, (2016).

Basal promoter elements are usually found near the transcription start site (Juven-Gershon *et al.*, 2008). They play a role by promoting the binding of RNA polymerase II, along with the general transcription factors involved in the expression of most genes (Li *et al.*, 1994; Orphanides *et al.*, 1996; Roeder, 1996; Conaway and Conaway, 1997; Reese, 2003). A well-known basal promoter element is the TATA box, recognized by TATA-box binding protein

(TBP) (Peterson *et al.*, 1990; Burley, 1996), which is a component of the general transcription factor II D (TFIID) (Horikoshi *et al.*, 1990).

The upstream promoter elements are located further upstream from the basal elements (up to several hundred base pairs from the transcription start site). During transcription, they are recognized by specific transcription factors depending on the type of elements present in each gene (Mitchell and Tjian, 1989; Ptashne and Gann, 1997; Lee and Young, 2000). In most genes, when the interaction between the general transcription factors with the basal promoter is not sufficient for the assembly of a stable transcription complex, the specific promoter elements will bring additional proteins into contact that provide the necessary stability (Gill, 1996; Stargell and Struhl, 1996; Struhl *et al.*, 1998). In addition, many of these specific promoter elements are required for the transcriptional regulation of gene expression under different circumstances, thus receiving the name of response elements.

The sequences that affect the expression of particular genes linked to them are called enhancers (Stadhouders *et al.*, 2012; Smallwood and Ren, 2013; Levine *et al.*, 2014). Enhancers contain groups of response elements and have the peculiarity of acting at long distances (up to several thousand base pairs away from the transcribed region), through the formation of loops in the DNA. Furthermore, interactions of promoter elements with their corresponding binding proteins are also affected by the chromatin structure of the genomic region. This constitutes a further source of complexity that affects transcription (Li *et al.*, 2007; Cairns, 2009; Venters and Pugh, 2009; Voss and Hager, 2014).

#### **4.1. STRUCTURE OF TRANSCRIPTION FACTORS**

Transcription factors are defined as proteins that influence the transcription of genes by binding to defined regions of the genome (Latchman, 1997). The most basic element within a transcription factor is the DNA-binding domain that recognize specific sequences within the promoter regions of the gene(s) that it regulates (Kummerfeld and Teichmann, 2006). By binding to these sequences, the transcription factor becomes either an activator or a repressor that increases or decreases the transcription of the target gene(s) accordingly. This ability mostly depends on other domains in the protein that act independently from the DNA-binding domain. Therefore, transcription factors possess a modular structure and the possibility to acquire new properties by domain mixing or shuffling, a process occurring and used by researchers to generate new mechanisms of transcriptional regulation (Gossen and Bujard, 1992; Morgenstern and Atchley, 1999; Beerli *et al.*, 2000; Ansari and Mapp, 2002; Traven *et al.*, 2006; Liu *et al.*, 2013).

## 4.2. DNA RECOGNITION BY TRANSCRIPTION FACTORS

Transcription factors can recognize specific DNA sequences by establishing interactions between the side chains of amino acids of the DNA binding domain with the nucleotides of the target site. For specific recognition, interactions must be established with the nucleotide bases that are located inside the DNA double-helical structure. Therefore, most transcription factors generate connections with DNA by binding to the major groove. However, some interactions through the minor groove were also reported; for instance, through hydrogen bonding and van der Waals contacts (weak attractions that occur between molecules in close proximity to each other) (Shimoni and Glusker, 1995; Suzuki *et al.*, 1995; Luscombe *et al.*, 2001; Rohs *et al.*, 2010). The specificity, strength or affinity of an interaction can be defined by unspecific contacts established with the sugar phosphate backbone, including ionic interactions between DNA phosphates and positively charged residues of the DNA-binding domain. The topology of the DNA around the transcription factor-binding site is also another factor affecting the specificity and strength of an interaction (Pan *et al.*, 2010). In addition, curvatures in DNA are often required by transcription factors to bind their target genes efficiently (Rohs *et al.*, 2009), and some transcription factors induce DNA bending upon binding (van der Vliet and Verrijzer, 1993), thus leading to changes that facilitate other processes, like DNA melting or the binding of additional proteins.

## 4.3. DNA-BINDING DOMAINS

The structures of DNA-binding domain are very diverse. Based on the mechanism of interaction with DNA sequence, there are three classes of DNA-binding domain structure, which are alpha helices, beta sheets or disordered regions (Pabo and Sauer, 1992). Usually, the DNA-binding domain forms a module that can be separated from the rest of the transcription factor structure without losing activity. DNA-binding domains are named based on their structural characteristics. Since it is possible for a DNA-binding domain to interact with different transcription factors, a transcription factor will be classified and receive the name corresponding to the DNA-domain with which it interacts (**Table 1**, Stegmaier *et al.*, 2004; Vaquerizas *et al.*, 2009; Charoensawan *et al.*, 2010).

**Table 1. Classification of Transcription Factors\***

- 
- 1. Superclass: Basic domains**
    - 1.1 Class: Leucine zipper factors (bZIP)
    - 1.2 Class: Helix–loop–helix factors (bHLH)
    - 1.3 Class: Helix–loop–helix/leucine zipper factors (bHLH-ZIP)
    - 1.4 Class: NF-1
    - 1.5 Class: RF-X
    - 1.6 Class: bHSH
  - 2. Superclass: Zinc-coordinating DNA-binding domains**
    - 2.1 Class: Cys4 zinc finger of nuclear receptor type
    - 2.2 Class: Diverse Cys4 zinc fingers
    - 2.3 Class: Cys2His2 zinc-finger domain
    - 2.4 Class: Cys6 cysteine–zinc cluster
    - 2.5 Class: Zinc fingers of alternating composition
  - 3. Superclass: Helix–turn–helix**
    - 3.1 Class: Homeodomain
    - 3.2 Class: Paired box
    - 3.3 Class: Fork head/winged helix
    - 3.4 Class: Heat shock factors
    - 3.5 Class: Tryptophan clusters
    - 3.6 Class: TEA domain
  - 4. Superclass: beta-Scaffold factors with minor groove contacts**
    - 4.1 Class: Rel homology region (RHR)
    - 4.2 Class: STAT
    - 4.3 Class: p53
    - 4.4 Class: MADS box
    - 4.5 Class: beta-Barrel alpha helix transcription factors
    - 4.6 Class: TATA-binding proteins
    - 4.7 Class: HMG
    - 4.8 Class: Heteromeric CCAAT factors
    - 4.9 Class: Grainyhead
    - 4.10 Class: Cold-shock domain factors
    - 4.11 Class: Runt
  - 0. Superclass: Other transcription factors**
    - 0.1 Class: Copper fist proteins
    - 0.2 Class: HMGI(Y)
    - 0.3 Class: Pocket domain
    - 0.4 Class: E1A-like factors
    - 0.5 Class: AP2/EREBP-related factors
- 

\* From <http://www.gene-regulation.com/pub/databases/transfac/clSM.html>

Transcription factors sharing the same type of DNA-binding domain (in other words, transcription factors from the same family) tend to have more similar DNA-binding specificities than those that belong to different families. In some cases, variations in DNA-binding specificity are often observed within the same family, and these are most often due to changes in specific residues of the DNA-binding domain (Berger *et al.*, 2008; Noyes *et al.*, 2008; Badis *et al.*, 2009). Thus, changes in the amino acid residues of the DNA-binding domain also occur during evolution and are used by researchers to create transcription factors with novel DNA-binding characteristics (Blancafort *et al.*, 2004; Amoutzias *et al.*, 2007; Joung and Sander, 2013).

#### **4.4. REGULATION OF TRANSCRIPTION FACTOR FUNCTION**

The activity of transcription factors usually impacts upon the expression of the genes involved. The distribution of transcription factors within the cell depends on the presence of proteins that either maintain them in the cytosol or assist them to move to the nucleus (Whiteside and Goodbourn, 1993; Kaffman and O'Shea, 1999). The movement of proteins to the nucleus is controlled by the presence of nuclear localization signals (NLS, usually stretches of positively charged residues exposed to the protein surface) in their structure, which are recognized by the nuclear transport machinery (Stewart, 2007).

Although most transcription factors include NLS in their structure, some of them recruit a complex with an NLS-containing partner for nuclear localization. In contrast, some transcription factors containing NLS are kept in the cytosol by partners that block their NLS or their interaction with the import machinery. Interactions with ligands or post-transcriptional modifications disturbing these barriers can lead to proteolysis of the inhibitory protein, allowing migration to the nucleus and binding to the target gene(s).

Interactions with ligands or post-transcriptional modifications likely determine the stability of transcription factors, which can then be targeted for degradation by the proteasome with or without these signaling events (Geng *et al.*, 2012; Yao and Ndoja, 2012). Post-translational modifications can also directly regulate the activity of a transcription factor that is always present inside the nucleus. These modifications including phosphorylation, acetylation, glycosylation, ubiquitination, sumoylation, and redox-dependent changes, have been characterized (Jackson and Tjian, 1988; Bohmann 1990; Bannister and Miska, 2000; Gill, 2003; Liu *et al.*, 2005; Ndoja *et al.*, 2014). Modifications provide scope for regulation via effects on the DNA-binding ability of the transcription factor or its capacity to control transcription, acting on its interaction with other proteins when bound to DNA (Yamasaki,

2016).

The general characteristics of transcription in plants are similar to those of other eukaryotes. To be more specific, many plant transcription factor families are also present in fungi and animals, suggesting that they are ancient acquisitions. However, there are still many differences between the transcription factor families of plants and other eukaryotes. To understand the characteristics of the different plant transcription factor families, as well as their roles in transcriptional regulation of different processes, it is necessary to carry out detailed studies of individual groups. The knowledge gained from such studies can shed light on how plants function at the molecular level (Yamasaki, 2016; Gonzalez, 2016).

#### **4.5. DEFINITION, CLASSIFICATION, STRUCTURE OF AP2/ERF TRANSCRIPTION FACTORS**

##### **4.5.1. Definition**

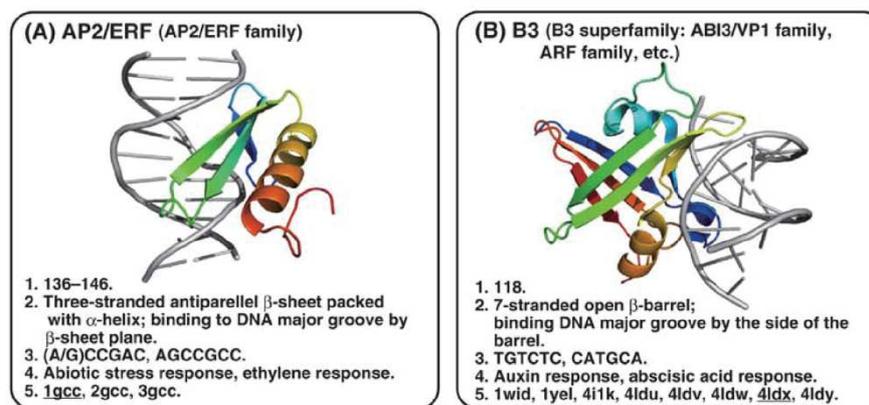
The APETALA2/ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR (AP2/ERF) domain was originally identified as a highly conserved DNA-binding domain (DBD) found in transcription factors (TFs) involved in floral development and ethylene response (Jofuku *et al.*, 1994; Ohme-Takagi and Shinshi, 1995). Approximately 150 genes in the *A. thaliana* genome share this highly basic motif encompassing ~60 amino acids, which characterises the largest of the plant transcription factor families. The family can be classified into four subfamilies: AP2, ERF, dehydration-responsive element-binding protein (DREB), and RELATED TO ABI3/VP1 (RAV) and some soloist (unclassified genes in the previous subfamilies) (Nakano *et al.*, 2006; Sharoni *et al.*, 2011; Rashid *et al.*, 2012). The AP2/ERF TFs function in a diverse range of plant-specific functions (e.g. responses to abiotic stresses, such as cold, dehydration, heat shock, and mechanical stress, ethylene response, and the development of flowers, roots, embryos, and seeds) (Jofuku *et al.*, 1994; Ohme-Takagi and Shinshi, 1995; Fowler and Thomashow, 2002; Aida *et al.*, 2004; Chandler *et al.*, 2007; Kagaya and Hattori, 2009; Dietz *et al.*, 2010; Mizoi *et al.*, 2012; Yamasaki, 2016).

##### **4.5.2. Classification**

DREB and ERF subfamily members possess a single AP2/ERF domain, whereas AP2 subfamily members possess two repeated domains. RAV subfamily members possess an N-terminal AP2/ERF domain and a C-terminal B3 domain (Kagaya *et al.*, 1999). The CBF1 protein in the DREB subfamily also requires amino acids outside the AP2/ERF domain for DNA binding (Canella *et al.*, 2010). Variations in the combination of DBDs, as well as in

sequences inside and outside the AP2/ERF domain, bring about differences in the DNA sequences recognized by the proteins. The AP2/ERF proteins possessing a single AP2/ERF domain are subdivided into the ERF and DREB subfamilies, based on their differential binding affinity to DNA sequences. Members of the ERF and DREB subfamilies bind to GC-rich motifs but ERF subfamily proteins bind to the core motif AGCCGCC; whereas DREB subfamily proteins containing C-repeats recognize the cis-acting element, A/GCCGAC (Dinh *et al.*, 2012) (**Fig. 27**).

The AP2/ERF domain of the RAV subfamily TF recognizes the CAACA sequence independently of the B3 domain, which recognizes the CACCTG sequence (Kagaya *et al.*, 1999). The AP2 subfamily includes members characterized by a tandem repetition of two AP2 domains (AP2-R1 and AP2-R2) and a small number of proteins with a single AP2 domain, showing higher similarity to the one contained in double-AP2 proteins than to the AP2 domain of the ERF and DREB proteins (Licausi, Ohme-Takagi and Perata, 2013). AP2-R2 specifically binds *in vitro* to the TTTGTT or AACAAA motif (Dinh *et al.*, 2012). Additionally, members of the AP2 subfamily possessing a double AP2 domain do not recognize a tandem repetition of ERF- or DREB-binding elements (Dinh *et al.*, 2012).



**Figure 27. Structures of the DNA-binding domains (DBDs) that characterize the AP2/ERF plant-specific transcription factors (TFs).** (A) The AP2/ERF domain in complex with DNA, (B) the B3 domain in complex with DNA. Information on (1) number of proteins in *Arabidopsis thaliana*, including those predicted in genome, (2) brief structural description, (3) representative recognition sequences, (4) representative functions of family members, and (5) Protein Data Bank (PDB) entry codes (those used in the figure are underlined) are also provided. Figure adapted from Yamasaki (2016).

#### 4.5.3. Structure

The structure of the complex of the AP2/ERF domain of AtERF1 and the GCC-box

DNA has been determined by NMR spectroscopy, whereby the AP2/ERF domain was described as a “GCC-box binding domain” in the original report (Allen *et al.*, 1998). The structure of the AP2/ERF domain consists of a three-stranded antiparallel beta-sheet and an alpha-helix. The beta-sheet moiety fits into the major groove of the DNA, in which the plane of the sheet is nearly parallel to the helical axis of the DNA. This DNA-binding mode is considered atypical because the majority of DBDs, such as the Zn finger domain in combination with an antiparallel beta-sheet and an alpha-helix, place the alpha-helix in the major groove of the DNA (**Fig. 27**). For sequence recognition, three arginine and two tryptophan residues, which are highly conserved in the sequence motif, directly contact bases. Three-dimensional structures and DNA-binding modes of all AP2/ERF proteins strikingly similar to those of AtERF1 have been reported (Wojciak *et al.*, 1999; Fadeev *et al.*, 2009) that possess a three-stranded antiparallel beta-sheet and an alpha-helix, where the b-sheet fits into the DNA major groove.

Determinant residues to illustrate the difference in sequence specificity between subfamilies have been analysed mainly by site-directed mutagenesis experiments, at least partly in terms of the AtERF1 structure (Hao *et al.*, 2002; Sakuma *et al.*, 2002; Krizek, 2003; Liuet *et al.*, 2006; Yang *et al.*, 2009). The data revealed that determinant residues between DREB and ERF TFs, which recognize slightly different sequences, are not necessarily those in direct contact with bases; those likely to structurally influence the orientation of base-contacting residues can be the determinant for specificity. For the AP2 domain, which recognizes greatly different DNA sequences from other subfamilies, the DNA-binding interface was suggested to be very different from that of AtERF1 (Krizek, 2003). The DNA-binding mode shown is applicable to ERF domains, but not strictly speaking necessarily to AP2 domains. Mainly existing in plants but the AP2/ERF have also been identified in ciliate, cyanobacteria, and bacteriophage (Magnani *et al.*, 2004; Wuitschick *et al.*, 2004). Moreover, a large number of AP2/ERF-like domains have been identified in apicomplexa (Balaji *et al.*, 2005; Painter *et al.*, 2011).

#### **4.6. AP2/ERF TRANSCRIPTION FACTORS IN ARABIDOPSIS AND RICE**

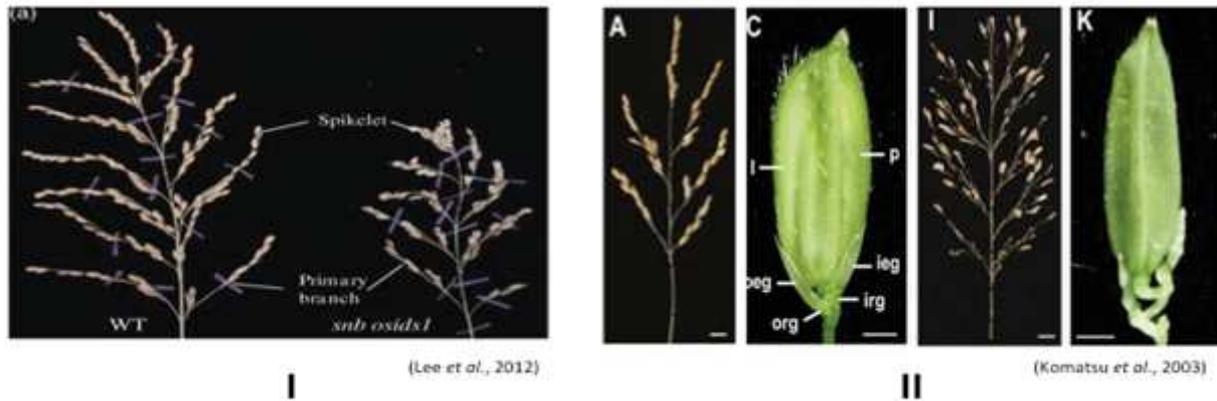
In *A. thaliana*, the AP2/ERF genes from the 4 subfamilies are present in the genome. The RAV subfamily consists of 6 genes encoding a single AP2/ERF domain with one B3 domain. The DREB subfamily consists of 57 genes (Sakuma *et al.*, 2002) belonging to group I to IV of AP2/ERF family (Nakano *et al.*, 2006). The ERF subfamily consists of 65 genes (Sakuma *et al.*, 2002) belonging to group V to X of AP2/ERF family (Nakano *et al.*, 2006).

The AP2 subfamily consists of 18 genes then divided into 3 groups based on the motifs conserved in and outside AP2 domain previously described in Kim *et al.*, 2005 and detailed more in the following parts of my study: euAP2, euANT/PLT and basalANT groups. The remaining member, At4g13040, with an AP2/ERF domain sequence but low homology in comparison with the other *AP2/ERF* genes, was separated as a soloist.

One hundred seventy *AP2/ERF* family genes were identified by phylogenetic analysis of the rice genome (*Oryza sativa ssp. japonica*) including the four subfamilies (AP2, ERF, DREB and RAV), then they were classified into ten clades or groups, and two soloists depending on the number of AP2/ERF domains in the protein and the existence of other DNA binding domains (Rashid *et al.*, 2012; Sakuma *et al.*, 2002). The RAV subfamily comprises 4 genes encoding a single AP2/ERF domain with one B3 domain. The DREB subfamily consists of 55 genes and the ERF subfamily includes 80 genes. The AP2 subfamily amounts to 27 genes encoding proteins with two or one AP2 domain.

There is increasing evidence that some *AP2/ERF* family genes are involved in panicle meristem identity determination by regulating the expression of spikelet meristem identity genes in grass species (Zhang and Yuan, 2014). For example, in rice, both the *frizzy panicle (fzp)* and *supernumerary bract (snb)* mutants showed repetitive production of rudimentary glumes (Yoshida and Nagato, 2011) (**Fig. 28**). *FZP*, an ortholog of maize *BD1*, encodes an ERF transcription factor (Chuck, 2002; Komatsu *et al.*, 2003; Lee *et al.*, 2007). Analyses of the phenotypes of both *fzp* and *bd1* mutants in rice revealed the presence of sterile lemmas similar to the glumes found in other grass species. The formation of florets was replaced by sequential rounds of branching with the production of rudimentary glumes in *fzp* mutants. Hence, *FZP* was hypothesised to suppress axillary meristem formation on the SM or to enhance FM identity. In contrast, *snb* mutants did not display repetitive branching, but sporadically formed incomplete florets subtended by repetitive rudimentary glumes (Lee *et al.*, 2007; Lee and An, 2012). *SNB*, encoding a nuclear protein with two AP2/ERF domains and another AP2 gene, *OsIDS1*, considered as orthologs of maize INDETERMINATE SPIKELET1 (*IDS1*) and SISTER OF INDETERMINATE SPIKELET1 (*SID1*) respectively (Chuck, Meeley and Hake, 1998, 2008; Malcomber *et al.*, 2006; Lee *et al.*, 2007) apparently perform redundant functions in specifying FM identity in rice (Yoshida and Nagato, 2011). In fact, *SNB* impacts floral organ identity and floral determinacy (Lee *et al.*, 2007) while *OsIDS1* operates in FM establishment and in the formation of the floral organs (Lee and An, 2012). Both *SNB* and *OsIDS1* are regulated by miR172, overexpression of miR172 resulting in the elongation of the lemma/palea and loss of floral determinacy (Dai *et al.*, 2016).

Although many AP2/ERF transcription factors have been extensively studied in various plants, further efforts are still needed to identify other novel AP2/ERF genes involved in plant development, especially during panicle development. More specifically, it is of great interest to determine to what extent these genes may have played a role in determining panicle diversity.

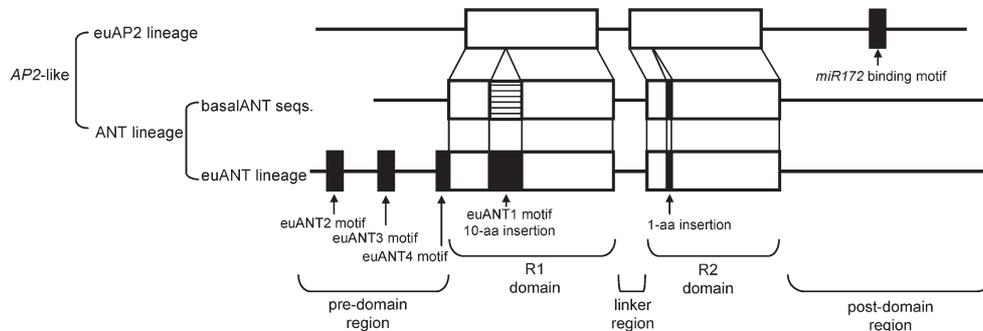


**Figure 28. Comparison of inflorescence architecture** (I) between wild type (WT) and *snb osids1* mutant (Lee *et al.*, 2012) (II) between wild type (WT) and *fzp* mutant (Komatsu *et al.*, 2003) (A) The WT inflorescence. (C) WT terminal spikelets. (I) The *fzp-3* inflorescence. (C) *fzp-3* terminal spikelets. org, outer rudimentary glume; irg, inner rudimentary glume; oeg, outer empty glume; ieg, inner empty glume; l, lemma; p, palea. Scale bars: 1 cm (A, I); 1 mm (C, K).

As mentioned before, AP2 subfamily genes may be differentiated into 3 groups: euAP2, euANT/PLT and basalANT (**Fig. 29**). The euAP2 or APETALA2-like group members possess a recognition site for the microRNA *miR172* in transcript sequence. This group of genes in *A. thaliana* have been properly studied in the context of their role in floral ontogeny, such as floral meristem identity and flowering time (Zhang *et al.*, 2015; Wang *et al.*, 2016; Huang *et al.*, 2017; Gras *et al.*, 2018). Besides, their 2 orthologs in rice *OsSNB* and *OsIDS1* were well characterized in the context of flower identity, as mentioned above (Zhu *et al.*, 2009; Lee and An, 2012). Other AP2-like rice orthologs *OsSHAT1* and *OsRSR1* control seed shattering and starch synthesis in rice grain, respectively (Fu and Xue, 2010; Zhou *et al.*, 2012).

The ANT lineage comprises 2 groups: basalANT (or WRINKLED-like/WRI) group and euANT/PLT (or PLETHORA-like) group. The euANT/PLT group members have additional 3 motifs conserved in pre-domain sequences compared to basalANT group members. While the basalANT group has been functionally described in *A. thaliana* (To *et al.*, 2012; Park, Go and Suh, 2016), two *WR11* genes, *OsWR11-1* and *OsWR11-2*, and two additional *WR11* homologs, *OsWR13* and *OsWR14* were identified in the rice genome. *OsWR11* was found to be ubiquitously expressed in rice plants, including in developing seeds (Ap *et al.*, 2019).

However, the importance of these genes remains unclear in rice, with the exception of *SMALL ORGAN SIZE1* (*ERF142*), another *WRI* gene, which was shown to control organ size and to modulate root meristem size (Aya *et al.*, 2014; Li *et al.*, 2015; Hirano *et al.*, 2017).



**Figure 29. Detailed structure of AP2 subfamily genes along with euAP2, basalANT and euANT/PLT groups.** Open boxes indicate AP2 domains, and black boxes indicate lineage-specific motifs or insertions. Hatched region indicates that portion of the euANT1 motif that is not conserved in the basalANT sequences. From Kim *et al.* (2005)

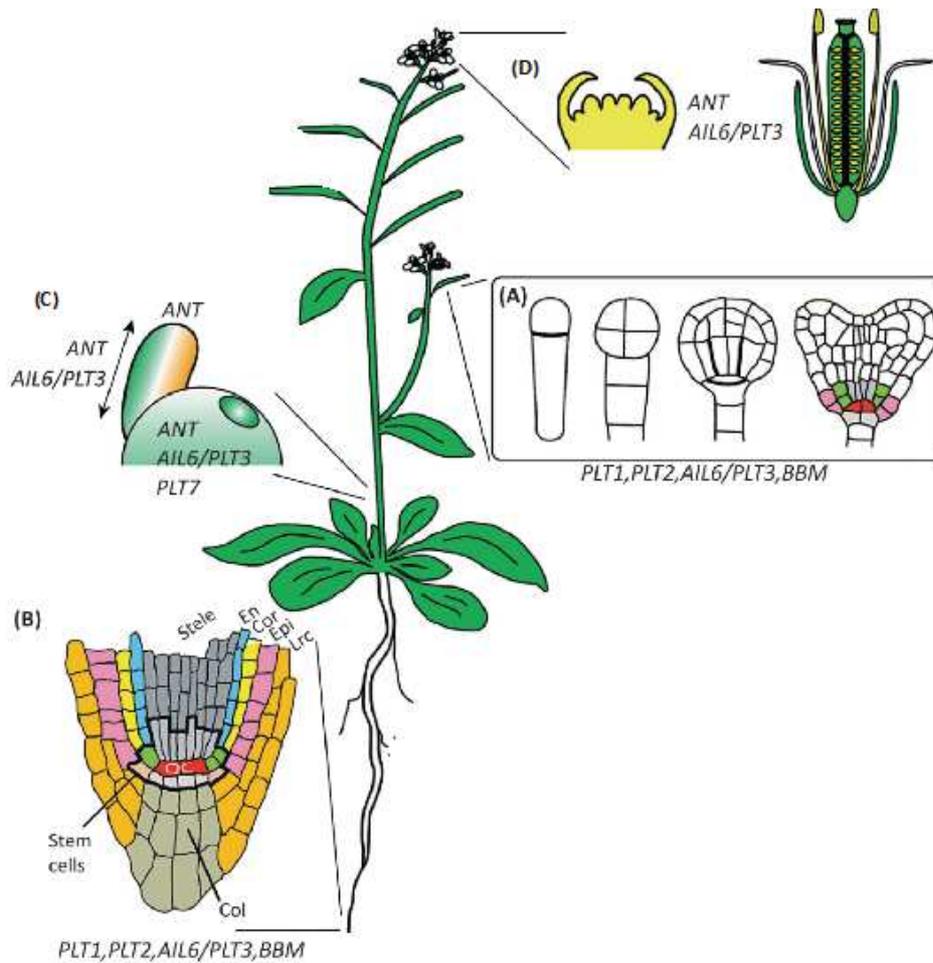
The euANT/PLT group members, *PLETHORA1* (*PLT1*), *PLETHORA2* (*PLT2*), *AINTEGUMENTA AINTEGUMENTA-LIKE6/PLETHORA3* (*AIL6/PLT3*), and *BABYBOOM* (*BBM*) genes play a major role in the basal patterning of the embryo (**Fig. 30A**). *PLT1* and *PLT2* gene expression has been described from the octant stage onwards, in the lower tier of the embryo proper. Early embryonic *AIL6/PLT3* and *BBM* expression has not been reported. Later in embryogenesis, *PLT1* expression becomes restricted to the quiescent centre (QC) and surrounding stem cells, whereas the *PLT2*, *AIL6/PLT3*, and *BBM* expression domains are slightly expanded to include the ground tissue and provascular cells. Post-embryonically, the *AIL6/PLT3* expression maximum is in the columella stem cells, in contrast to the QC peak expression observed for *PLT1*, *PLT2*, and *BBM* in the root meristem. Only combinations of *plt1*, *plt2*, *ail6/plt3*, and *bbm* mutants show embryonic abnormalities. The *plt1* and *plt2* seedlings show defective root development, confirming that *PLT1* and *PLT2* are required for stem cell niche specification. The *plt1/plt2/ail6* triple mutant embryos show aberrant organization of the embryonic root pole and seedlings are rootless. *plt2/bbm* double mutants fail to develop past the early embryo stage, demonstrating the importance of *PLT2* and *BBM* in embryogenesis. Expression patterns set up during embryogenesis are maintained during post-embryonic root development (**Fig. 30B**), where they show partly overlapping gradients of protein accumulation, with the highest protein levels in the stem cell niche (Horstman *et al.*, 2014; Scheres and Krizek, 2018).

The *AINTEGUMENTA* (*ANT*), *AIL6/PLT3*, and *AINTEGUMENTA-*

*LIKE7/PLETHORA7 (AIL7/PLT7)* genes are expressed in distinct but overlapping domains in the inflorescence and vegetative shoot meristems. *AIL6/PLT3* and *PLT7* are expressed throughout the meristem, but their expression is increased in the peripheral zones and in the central zone, respectively. In contrast, *ANT* is expressed exclusively in the peripheral zones and marks the cryptic bract region of the floral meristem. The *ant/ail6/plt7* triple mutant shoot stops growing after the production of a few leaves owing to reduced cell divisions in the meristem and differentiation of the meristematic cells, showing that the *ANT*, *AIL6/PLT3*, and *AIL7/PLT7* genes are required for shoot apical meristem (SAM) maintenance (**Fig. 30C**). The *ant* and *ail6/plt3* mutations enhance the *wus* and shoot meristemless (*stm*) phenotypes, whereas the *plt7* mutation partially rescues these phenotypes, indicating that *ANT*, *AIL6/PLT3*, and *PLT7* do not function in a strictly redundant fashion. Furthermore, the euANT/PLT proteins are well-known to determine stem cell fate, to boost organ growth and to suppress the differentiation in meristem function within the shoot of *A. thaliana* (Horstman *et al.*, 2014; Scheres and Krizek, 2018). Both the *ant* and *ap2* single mutants have fewer cells in their floral meristems, resulting in a reduced number of floral organs (**Fig. 30D**).

Ten euANT/PLT genes were previously reported in *O. sativa ssp. japonica* (*OsPLT1* to *OsPLT10*; Li and Xue, 2011) compared to eight members of the euANT/PLT group in *A. thaliana* (Mähönen *et al.*, 2014; Scheres and Krizek, 2018). Functional analyses have shown that some of these genes related to *BBM* may also function in embryogenesis and may act as pluripotency factors (Khanday *et al.* (2019; Lowe *et al.*, 2016). Moreover, the *BBM* and *ANT*-related genes in rice may be associated with root development, as illustrated by the *crowrootless5* mutant affected in *OsPLT8* gene (Li and Xue, 2011; Kitomi *et al.* 2011).

The *ANT* gene was specifically considered to be a key regulator of shoot and floral development in *A. thaliana*. However, there is a missing part of the *ANT*-related gene role in rice in the context of panicle architecture development, raising questions about the importance of euANT/PLT genes in the regulation of panicle meristem identities and activities. Consequently, the euANT/PLT gene group was chosen for further characterization in order to address key questions about the molecular mechanisms regulating panicle architecture and development. In the following chapters, I will discuss how the euANT/PLT genes from rice were subjected to further analysis.



**Figure 30. Sites of AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT) function during development of *A. thaliana*.** (A) Embryogenesis. (B) Root development. (C) Organ polarity. (D) Floral organ identity and ovule development. BBM, BABYBOOM; ANT, AINTEGUMENTA; PLT, PLETHORA. Figure adapted from (Horstman *et al.*, 2014).

**REFERENCES (INCLUDING IN THE END OF MATERIALS AND METHODS)**

## 5. OBJECTIVES

Studies of genes involved in panicle development will improve our understanding of the regulatory processes underlying the architecture of this important plant structure. The broad goal of this PhD project was to identify and to characterize *AP2/ERF* genes regulating panicle development and its diversity in the genus *Oryza*, then to determine whether the corresponding genes were impacted by the two independent domestications that occurred in the same genus. Within this context, it was aimed to improve existing knowledge of the molecular mechanisms involved in the establishment of distinct meristem fates during panicle development, this being a key element affecting rice yield and therefore human nutrition. Ultimately, an approach of this type should contribute towards the identification of key factors regulating rice panicle development that could be targeted for improvement in breeding programs related to yield and environmental adaptation, one of the today's key agronomical research priorities.

My 3-year PhD project was initiated in this context in November 2016, with the aim of answering the questions below:

- a What role(s) do AP2/ERF transcription factors play in early rice panicle development? Do they influence the panicle architecture diversity and were they targeted in the domestication of rice (crop vs. wild relative)?
- b What are the roles of *PLETHORA/AINTEGUMENTA* genes in relation to panicle structure development?

In the context of my PhD, I aimed to address these questions or at least to obtain results to partly address them. I decided to proceed in three steps:

1. I focused on the identification and *in silico* analysis of key *AP2/ERF* genes associated with early panicle development through transcriptome profiling of Asian and African rice species during reproductive development and benefited from the ongoing analyses. An *in-silico* analysis of *AP2/ERF* genes detected in different RNA-seq datasets produced in the hosting lab was carried out.
2. The next step was to profile the corresponding gene expression patterns during panicle development by using large-scale qRT-PCR (Fluidigm® technology) and investigate the function of those genes potentially governing the inter-specific diversity of rice panicle development at early stages so as to study the possible involvement of these genes in the evolution of panicle structure.

3. Based on the results obtained from Step 1 and Step 2, several *AP2/ERF* candidate genes were selected for further functional analysis to investigate their possible contributions to panicle development and its diversity. For those *AP2/ERF* genes for which mutants were available, a phenotypic analysis of rice panicles at early stages was carried out. In the case of *AP2/ERF* genes of interest for which no mutant was available, I generated transgenic plants using CRISPR/Cas9 methodology.

My work focused mainly on rice species available in IRD-Montpellier. The PhD project was carried out with the support of French and Vietnamese scientists in Montpellier and Hanoi in the framework of collaborations between the Evo-Devo of the Inflorescence (EDI) group (DIADE, IRD Montpellier) and the International Joint Laboratory (LMI RICE) in the Agricultural genetics Institute (AGI) in Ha Noi (Viet Nam).

I received financial support from the French Embassy in Vietnam (Programme de Bourses d'Excellence 2016) for 36 months from October 2016 to September 2019 and the "Allocations de recherche pour une thèse au Sud" (ARTS) program from IRD for 18 months from April 2017 to October 2018. According to ARTS rules, I worked for three months in Vietnam in 2017 and 2018.

This PhD thesis is organized in five chapters:

- **Chapter I** presents a set of AP2-like genes associated with inflorescence branching and architecture in domesticated rice
- **Chapter II** describes the *in silico* analysis and expression profiling during panicle development of *euANT/PLT* genes in rice
- **Chapter III** explores the functional involvement of *euANT/PLT* genes in panicle architecture determination in *Oryza sativa*
- **Chapter IV** provides a general discussion and perspectives resulting from this work.
- **Chapter V** describes the materials and methods used in the work.

According to GAIA Doctoral School rules from University of Montpellier, Chapter 2 and Chapter 3 were written in article format and Chapter I includes a published scientific article.

# **CHAPTER I**

**A set of AP2-like genes is associated with inflorescence  
branching and architecture in domesticated rice**

## GENERAL INTRODUCTION

Understanding the events during panicle development is crucial for the sustainable improvement of rice yield potential using targeted breeding programmes. Although progress has been made in the characterization of individual genes and their interactions, a complete understanding of the mechanisms controlling panicle morphology and grain yield requires mechanistic studies that can explain the interactions of gene regulatory network components with each other and with the environment (Azpeitia et al., 2013). To achieve this goal, studies needed to be performed to examine differences in gene expression between different meristem types during development, with particular regard to branch formation. The laboratory at IRD Montpellier has developed two sets of Illumina® RNA-seq data: one relating to the different types of meristems (RM, PBM, ePBM/AM and SM) in early stage panicles of *O. sativa* ssp. Japonica (Harrop et al., 2016), and another corresponding to early panicles at the branching stage in domesticated and wild species of Asian and African rice (*O. sativa*, *O. rufipogon*, *O. glaberrima* and *O. barthii*) (Harrop et al., 2019). The analysis of RNAs derived from both protein-encoding genes, as well as non-coding RNAs, is ongoing for these two datasets, with the aim of characterizing the underlying gene regulatory networks. Interestingly, a large number of *AP2/ERF* (*APETALA2/Ethylene Responsive Factor*) genes were identified in the two datasets, suggesting roles for members of this family in the regulation of differential meristem fates in the panicle and a possible involvement in the determination of inter-specific inflorescence diversity.



RESEARCH PAPER

# A set of *AP2*-like genes is associated with inflorescence branching and architecture in domesticated rice

Thomas W. R. Harrop<sup>1</sup>, Otho Mantegazza<sup>2</sup>, Ai My Luong<sup>2</sup>, Kevin Béthune<sup>2</sup>, Mathias Lorieux<sup>3</sup>, Stefan Jouannic<sup>2</sup> and Hélène Adam<sup>2,\*</sup> 

<sup>1</sup> Laboratory for Evolution and Development, Department of Biochemistry, University of Otago, PO Box 56, Dunedin 9054, Aotearoa New Zealand

<sup>2</sup> University of Montpellier, DIADE, IRD, France

<sup>3</sup> Rice genetics and Genomics Laboratory, International Center for Tropical Agriculture, Cali 6713, Colombia

\* Correspondence: [helene.adam@ird.fr](mailto:helene.adam@ird.fr)

Received 17 April 2019; Editorial decision 15 July 2019; Accepted 15 July 2019

Editor: Frank Wellmer, Trinity College Dublin, Ireland

## Abstract

**Rice yield is influenced by inflorescence size and architecture, and inflorescences from domesticated rice accessions produce more branches and grains. Neither the molecular control of branching nor the developmental differences between wild and domesticated rice accessions are fully understood. We surveyed phenotypes related to branching, size, and grain yield across 91 wild and domesticated African and Asian accessions. Characteristics related to axillary meristem identity were the main phenotypic differences between inflorescences from wild and domesticated accessions. We used whole transcriptome sequencing in developing inflorescences to measure gene expression before and after the transition from branching axillary meristems to determinate spikelet meristems. We identified a core set of genes associated with axillary meristem identity in Asian and African rice, and another set associated with phenotypic variability between wild and domesticated accessions. *AP2/EREBP*-like genes were enriched in both sets, suggesting that they are key factors in inflorescence branching and rice domestication. Our work has identified new candidates in the molecular control of inflorescence development and grain yield, and provides a detailed description of the effects of domestication on phenotype and gene expression.**

**Keywords:** *AP2/EREBP*-like genes, development, domestication, inflorescence, rice.

## Introduction

Rice produces grains on complex raceme inflorescences called panicles, which consist of a series of branches of different orders. The rachis is the main axis, and primary, secondary, and tertiary branches form higher-order axes (see [Supplementary Fig. S1](#) at *JXB* online). Grain yield in rice is linked to inflorescence branching, because the number of spikelets produced on the higher-order branches determines the number of grains per panicle ([Doebley \*et al.\*, 2006](#); [Xing and Zhang, 2010](#); [Olsen](#)

and [Wendel, 2013](#)). Inflorescence size and architecture are key targets in selective breeding for improved rice grain yield and quality ([Doust, 2007](#); [Wang and Li, 2011](#)), and these characteristics are different between domesticated and wild accessions.

Inflorescence branching and the number of grains per inflorescence vary between clades, species, and accessions of rice in the genus *Oryza* ([Yamaki \*et al.\*, 2010](#); [Taguchi-Shiobara \*et al.\*, 2011](#)). Asian and African clades diverged around 2 million years

ago and were domesticated independently (Zhu and Ge, 2005). Domesticated Asian rice, *Oryza sativa*, diverged 10 000 years ago via a complex domestication process from the wild Asian rice species, *Oryza rufipogon*. The *O. sativa* species complex involves a network of subspecies, including *O. sativa* ssp. *indica* and *O. sativa* ssp. *japonica* (Huang *et al.*, 2012; Choi *et al.*, 2017; Stein *et al.*, 2018). Domesticated African rice, *Oryza glaberrima*, has a simpler history. It was domesticated from its wild relative, *Oryza barthii*, 3000 years ago, without subsequent introgression (Vaughan *et al.*, 2008; Cubry *et al.*, 2018). In both African and Asian clades, domesticated species produce inflorescences that are more complex and ramified than those of their wild relatives (Linares, 2002; Yamaki *et al.*, 2010; Ta *et al.*, 2017).

Panicle structure is established early after floral transition. The vegetative shoot meristem develops into a reproductive rachis meristem, which produces axillary meristems until its abortion. The axillary meristems on the rachis all develop as primary branches, which themselves produce a variable number of axillary meristems. Axillary meristems on the primary branches can differentiate into secondary branches, some of which may also produce tertiary branch meristems. Axillary meristems on the primary branches can also differentiate directly into spikelet meristems. Finally, all axillary meristems and terminal meristems differentiate gradually from the top to the base of the panicle into spikelet meristems and then florets. Each rice grain is produced from a single spikelet (reviewed by Itoh *et al.*, 2005).

The complexity of panicle branching and the number of grains are determined by two basic developmental outcomes: the number of axillary meristems produced along all panicle axes; and the rate of meristem fate transition, which determines whether an axillary meristem grows into a higher-order branch or differentiates into a spikelet and gives rise to a grain. The maturation rate and identity of reproductive and axillary meristems are controlled by transcriptional regulators, some of which have been identified in *O. sativa* (Xing and Zhang, 2010; Wang and Li, 2011). Genes including *G1 LIKE PROTEIN 5* (*TAWAWA1/TAW1*), *LAX PANICLE 1* (*LAX1*), and *LAX PANICLE 2* (*LAX2*) control the branching phase via establishment and activity of indeterminate meristems (Komatsu *et al.*, 2001; Tabuchi *et al.*, 2011; Yoshida *et al.*, 2013). Genes such as *SUPERNUMERARY BRACT GENE* (*SNB*), *FRIZZY PANICLE* (*FZP*), and *LEAFY HULL STERILE 1* (*LHS1*) regulate the transition from indeterminate meristems to determinate spikelet meristems and the subsequent transition from spikelet to floret meristem (Jeon *et al.*, 2000; Komatsu *et al.*, 2001; Agrawal *et al.*, 2005; Chen *et al.*, 2006; Lee *et al.*, 2007; Khanday *et al.*, 2013). Regulatory changes in some genes, including *FZP*, *WEALTHY FARMER'S PANICLE* (*IDEAL PLANT ARCHITECTURE 1*, *WFP/IPA1/SPL14*), *GRAIN SIZE 2* (*G22/GRF4*), and *GRF6*, are also associated with modified panicle phenotype (Jiao *et al.*, 2010; Miura *et al.*, 2010; Bai *et al.*, 2016; Duan *et al.*, 2015; Huang *et al.*, 2018).

Despite advances in understanding molecular mechanisms that regulate panicle branching in rice (Furutani *et al.*, 2006; Harrop *et al.*, 2016), and characterization of individual genes associated with variation in panicle branching (Bai *et al.*, 2012; Ikeda *et al.*, 2013), the transcriptional control of

axillary meristem identity that underlies the phenotypic diversity of branching across rice species is not fully understood. The independent domestications of African and Asian rice in the genus *Oryza* provide a comparative context to study the evolution of agronomic traits such as panicle architecture, the mechanisms underlying parallel evolution of phenotype, and the basic molecular control of inflorescence branching. Here, we phenotyped panicles from domesticated and wild accessions of Asian and African rice, and used whole-transcriptome sequencing (RNAseq) to reveal gene expression patterns associated with the diversity of inflorescence architecture. Our analysis reveals an association between expression of *AP2/EREBP*-like genes, panicle architecture and domestication.

## Materials and methods

### Plant material and growth conditions

Panicle morphological traits were measured in 91 accessions of *O. rufipogon*, *O. sativa*, *O. glaberrima*, and *O. barthii*, grown in Cali, Colombia and Montpellier, France (Supplementary Table S1). At panicle maturity, we collected the three main panicles from three plants per accession, per replicate (i.e. 18 panicles per accession). We used four accessions for expression analysis: *O. sativa* ssp. *indica* IR64, *O. rufipogon* W1654, *O. glaberrima* Tog5681, and *O. barthii* B88. These accessions were grown in a greenhouse in Montpellier, France, in June 2014, under long day conditions (14 h light–10 h dark). After 6–8 weeks they were transferred to short day conditions (11 h light–13 h dark) to induce flowering. To confirm panicle phenotypes in the growth conditions used for RNAseq, we evaluated panicle traits for nine panicles from each accession, which were grown in the greenhouse under the same growth conditions. The *arl5* and *smos1-3* mutants (Kitomi *et al.*, 2011; Aya *et al.*, 2014) were grown in a greenhouse in Montpellier, France, in October 2017 under short day conditions (11 h light–13 h dark). At least 18 panicles were used for panicle phenotyping. All greenhouse plants were grown at 28 °C with 80% relative humidity. For phenotyping analyses, each panicle was spread out and fixed on white paper using adhesive tape. Panicles were photographed and the images were used for panicle structure and seed number analysis with P-TRAP software (AL-Tam *et al.*, 2013).

### Tissue collection and RNA sequencing

For expression analysis, we collected 15 immature panicles each from at least 10 plants per accession, per stage, collected from 4 to 15 d after floral induction (i.e. transfer to short day conditions) for each biological replicate. For sample collection, leaves surrounding the young panicle were removed by hand and the reproductive tissue was cut with a sharp blade under a Stemi 508 (Zeiss, Germany) stereo microscope to identify developmental stage. The reproductive tissues were immediately frozen in liquid nitrogen, and total RNA including small RNA was extracted using the RNeasy Plant Mini kit with RLT and RWT buffers (Qiagen, Germany). DNase treatments were performed using the RNase-free DNase set (Qiagen, Germany). RNA integrity numbers of the extracted RNA, measured using a 2100 Bioanalyzer (Agilent, USA), were between 8.6 and 10. Stage specificity was validated with quantitative real-time RT-PCR (qPCR) using stage-specific marker genes (Supplementary Table S2); 400 ng of total RNA was used for each sample for RNAseq library preparation with the TruSeq Stranded Total RNA with Ribo-Zero Plant kit (Illumina, USA). After quantification with a 2100 Bioanalyzer, 125-base paired-end reads were generated on a HiSeq 2500 (Illumina) by the GeT platform (Toulouse, France).

### qPCR

cDNA was synthesized from 1 µg of DNase-treated total RNA using the SuperScript III First-Strand Synthesis System (Invitrogen, USA).

A Biomark HD Microfluidic Dynamic Array (Fluidigm, USA) was used for large-scale qPCR. A 96×96 Dynamic Array Integrated Fluidic Circuit (Fluidigm, USA) was loaded with cDNA and primer combinations after 15 cycles of specific target amplification and exonuclease I treatment. A fast cycling protocol with EvaGreen dye (Bio-Rad Laboratories, USA) was used for amplification. Three biological replicates were performed for each sample. Data were normalized using four genes (*LOC\_Os06g11170*, *LOC\_Os06g48970*, *LOC\_Os01g16970*, and *LOC\_Os03g61680*). Gene expression relative to the normalization factors was estimated using the  $2^{-\Delta\Delta CT}$  method without a calibrator sample (Livak and Schmittgen, 2001). Primer sequences are listed in Supplementary Table S2.

### Data analysis

We trimmed reads and removed adaptors with cutadapt (Martin, 2011), before mapping to the MSU v7 annotation of the *Oryza sativa* ssp. *japonica* cv. Nipponbare reference genome (Ouyang et al., 2007) using STAR in two-pass mode (Dobin et al., 2013). To generate per-library gene expression cutoffs, we used the 95th percentile of reads that mapped to intergenic regions of the genome, as described previously (Harrop et al., 2016). We used DESeq2 (Love et al., 2014) for differential expression analysis of genes that passed the cutoff. We used annotations from PlnTFDB v3.0 (Pérez-Rodríguez et al., 2010) and PlantTFDB v4.0 (Jin et al., 2017) to analyse expression of transcription factors. Soft clustering of transcription factor genes was performed with Mfuzz (Kumar and Futschik, 2007), and enrichment of transcription factor family genes was tested with the GSEA method using the FGSEA package (Subramanian et al., 2005; Sergushichev, 2016, Preprint).

### Reproducibility and data availability

Raw sequence data are hosted at the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) under accession PRJNA518559. The code we used to analyse the RNAseq data and panicle phenotype data is hosted at <https://github.com/tomharrop/5acc>, and the code for qPCR analysis is at <https://github.com/othomantegazza/5acc-qpcr>. We used snakemake (Köster and Rahmann, 2012) to arrange analysis steps into workflows and monitor dependencies, and Singularity (Kurtzer et al., 2017) to capture the computing environment. The final results and all intermediate steps can be exactly reproduced from the raw data with a single command using snakemake and Singularity. The source for this manuscript is hosted at <https://github.com/tomharrop/ird-5acc-paper>.

## Results

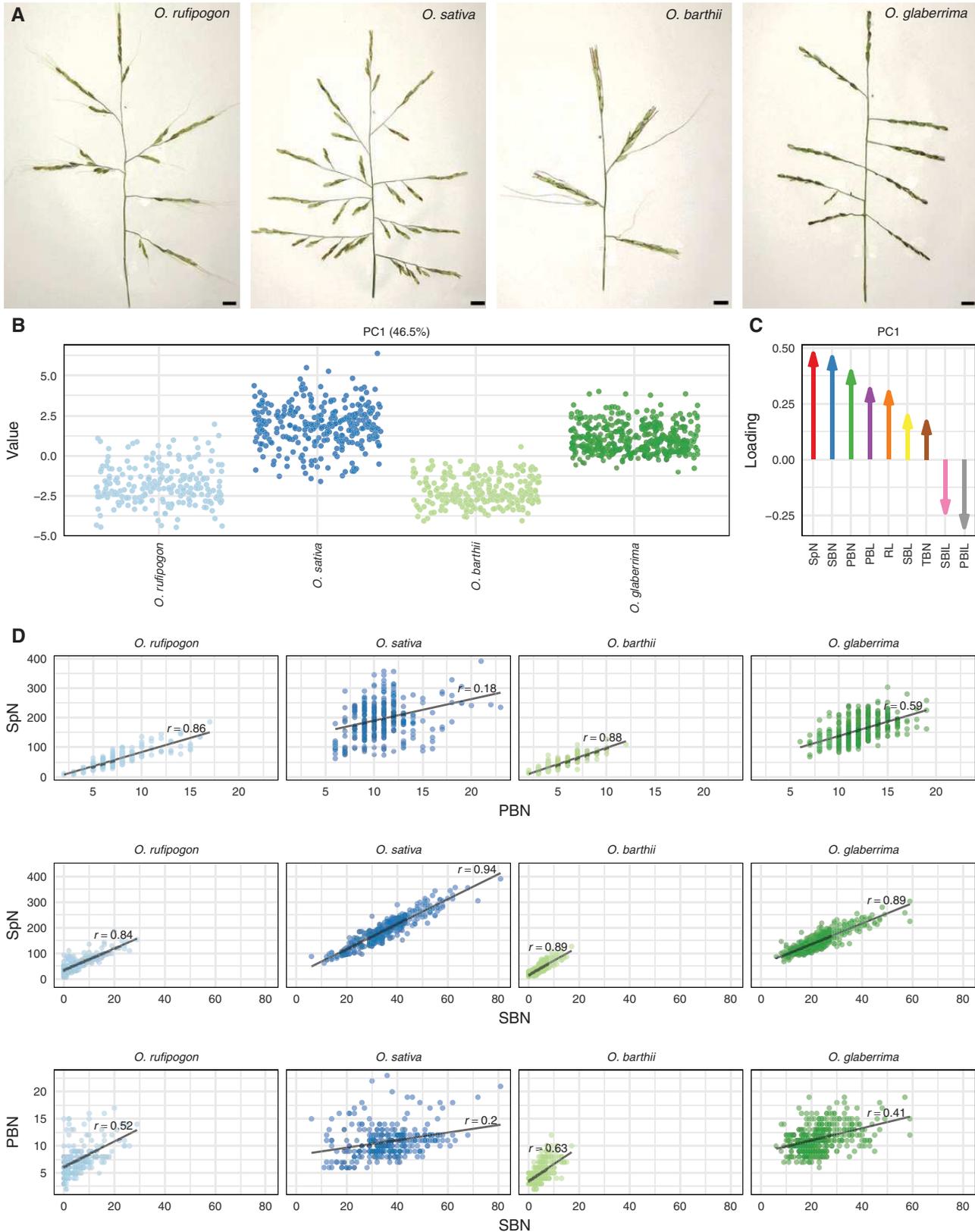
### Parallel changes in panicle architecture between wild and domesticated accessions

To measure the diversity of panicle architecture, we phenotyped 91 rice accessions (Supplementary Table S1), including wild Asian rice (*Oryza rufipogon*), domesticated Asian rice (*Oryza sativa*), wild African rice (*Oryza barthii*), and domesticated African rice (*Oryza glaberrima*), using P-TRAP image analysis software for automated measurement of traits (Fig. 1A; Supplementary Fig. S1; Supplementary Tables S1, S3; AL-Tam et al., 2013). Principal components analysis (PCA) of the phenotyping data identified a major coordinate (PC1) that accounts for 46.5% of variability (Fig. 1B). PC1 separates domesticated and wild accessions, but not Asian and African accessions, and is the only component that separates panicles from different accessions (Supplementary Fig. S2). Spikelet number, secondary branch number, and primary branch number have the highest loadings on PC1, whereas length traits have lower absolute loading on PC1 (Fig. 1C). For all species, spikelet number correlates more with secondary branch number than

it does with primary branch number. Primary branch number correlates with spikelet number more in wild than in domesticated species, but this correlation is weaker in Asian species than in African species. Primary and secondary branch numbers do not correlate, suggesting they are controlled by different genetic mechanisms (Fig. 1D). Our phenotypic analysis indicates similar changes in panicle architecture between wild and domesticated accessions in the independent African and Asian domestication processes, which presumably result from parallel, artificial selection on panicle architecture. Spikelet number, secondary branch number and primary branch number are the main contributors to these differences in panicle architecture, and these phenotypes are all related to axillary meristem formation and fate transition (Teo et al., 2014; Zhang and Yuan, 2014).

### Measurement of gene expression in developing panicles

We investigated gene expression changes associated with the diversity of panicle architecture and differences between the Asian and African domestication processes via RNAseq. We used a single accession each of domesticated Asian rice (*O. sativa* ssp. *indica* IR64) and its wild relative (*O. rufipogon* W1654), and domesticated African rice (*O. glaberrima* Tog5681) and its wild relative (*O. barthii* B88). Based on the extensive phenotyping described above, the chosen accessions are consistent with species-wide patterns of panicle architecture (Supplementary Fig. S3). To measure whole-transcriptome gene expression in these accessions, we collected immature panicles from each accession at four developmental stages: rachis meristem (RM); indeterminate meristem (IM), including panicles displaying primary branch initiation, elongation of primary branches, and axillary meristem initiation; determinate meristem (DM), including panicles wherein axillary meristems had transitioned into early spikelet differentiation; and floret meristem (FM), with early differentiation of floral organs (Supplementary Fig. S4A). We first confirmed staging of the panicles by extracting RNA from pooled immature panicles at each stage and measuring expression of markers of panicle development by quantitative real-time RT-PCR (qPCR) (Supplementary Fig. S4B). Because branching complexity is related to branch meristem establishment and meristem fate transition (Kyoizuka et al., 2014), and secondary branch number and spikelet number contribute to differences between wild and domesticated accessions (Fig. 1), we chose the IM and DM stages for RNAseq. cDNA libraries for sequencing were constructed from rRNA-depleted RNA samples from three biological replicates at both stages for all four accessions. Using the *O. sativa* ssp. *japonica* cv. Nipponbare reference genome to map reads (Ouyang et al., 2007), we obtained an average of more than 20 million uniquely mapped reads within exons for each accession, including African rice species (Supplementary Table S4). Our analysis was limited to transcripts from the four studied accessions that have homologs in the reference annotation that are similar enough for reads to map unambiguously. To allow for differences in mapping between accessions, we compared genes using between-stage read count differences within accessions, rather than read count



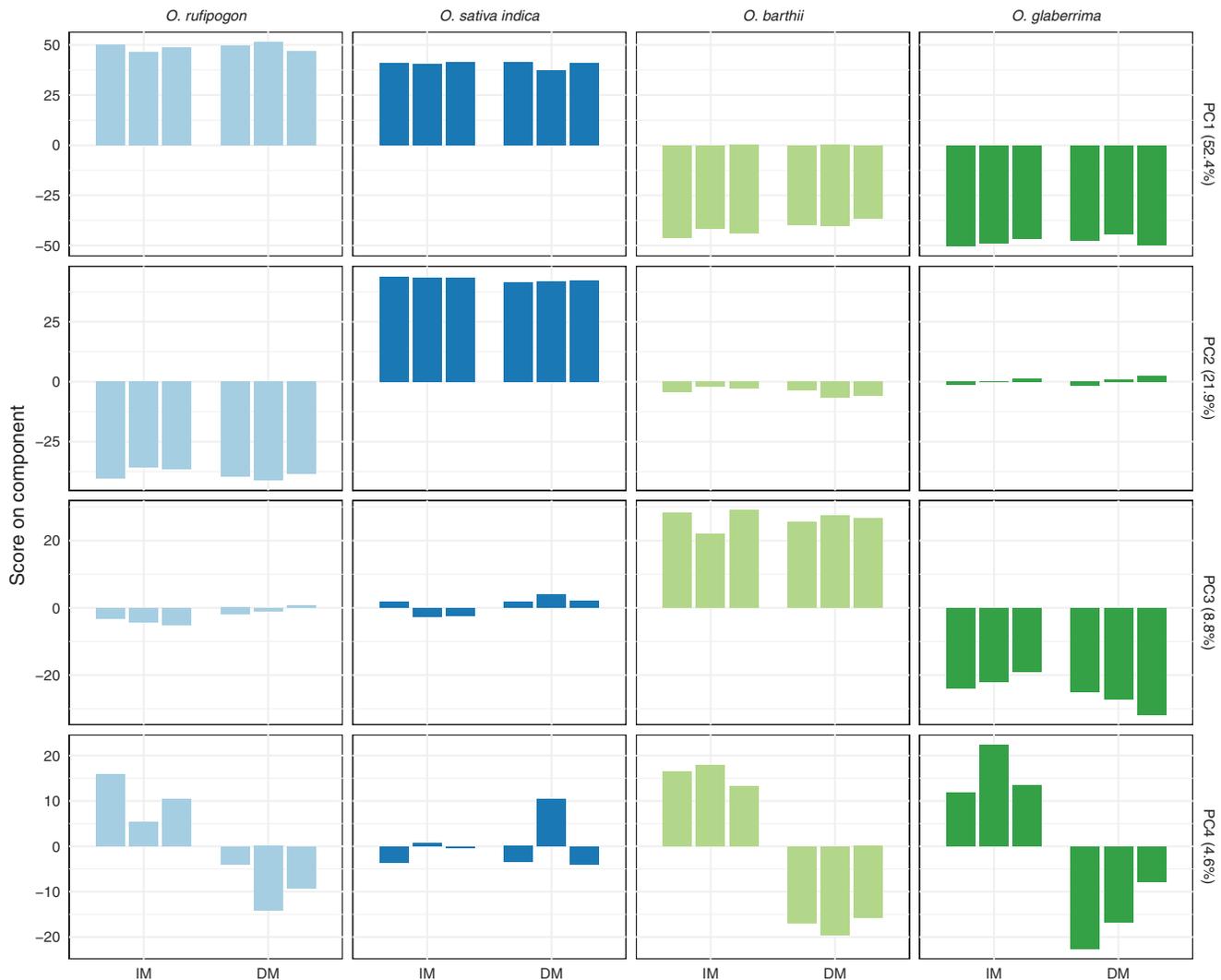
**Fig. 1.** Panicle complexity of 91 rice accessions. The main component of variability in panicle phenotypes splits accessions by domestication status, and is related to spikelet number, secondary branch number, and primary branch number. (A) We measured traits using spread panicles from *O. rufipogon*, *O. sativa*, *O. barthii*, and *O. glaberrima*. (B) The first principal component (PC1) in the panicle phenotype data accounts for 46.5% of variability and separates wild and domesticated accessions independently of continent. (C) Spikelet number (SpN), secondary branch number (SBN), and primary branch number (PBN) have the highest loadings on PC1. (D) Correlation between the main panicle traits that contribute to panicle architecture diversity. Primary branch number and spikelet number correlate in wild species. Secondary branch number and spikelet number correlate more in cultivated species than in wild species. Primary and secondary branch numbers do not correlate. PBIL, primary branch internode length; PBL, primary branch length; RL, rachis length; SBIL, secondary branch internode length; SBL, secondary branch length; TBN, tertiary branch number.

differences between accessions. Pairwise distances between samples, calculated on the number of reads per gene from all detected genes, grouped samples first by stage, then by accession, and then by continent (Supplementary Fig. S5). We did not observe grouping by domestication status, suggesting that transcriptome-wide changes during domestication are subtle compared with differences between species.

We used PCA on transformed raw counts to investigate general patterns of variation in the transcriptomes (Fig. 2). PC1 separates African and Asian accessions, and PC2 and PC3 separate wild and domesticated accessions in Asian and African samples, respectively. PC1–PC3 may relate to species-specific differences unrelated to panicle architecture, or mapping biases introduced by mapping all libraries against the *O. sativa* ssp. *japonica* reference. While the first three PCs separate different combinations of accessions, PC4 separates developmental stages across all four accessions, although separation is weaker in *O. sativa* ssp. *indica*. This suggests a conserved mechanism that controls the transition from indeterminate to determinate phase of axillary meristem development in all accessions.

### AP2/EREBP-like transcription factors are core regulators of panicle branching

To identify the core set of genes involved in axillary meristem determination in all four accessions, we used differential expression (DE) tests to find genes that were up- or down-regulated between stages across all accessions. Positive log<sub>2</sub>-fold change values (L<sub>2</sub>FCs) indicate higher expression in DM than in IM. There were 153 genes that had at least 1.5-fold DE between stages in all species at a false-discovery rate of 0.1, including 88 genes up-regulated in DM samples and 65 genes down-regulated in DM samples (Supplementary Table S5). There was an enrichment of transcription factor (TF) genes in the list of 153 differentially expressed genes (37 TF genes;  $P=6.0\times 10^{-12}$ , hypergeometric test), including *LHS1*, *LAX1*, *PANICLE PHYTOMER2 (PAP2)*, and *MOSAIC FLORAL ORGANS 1 (MFO1)*, which regulate inflorescence architecture or meristem fate transition in rice (Komatsu *et al.*, 2001; Ohmori *et al.*, 2009; Kobayashi *et al.*, 2010, 2012; Khanday *et al.*, 2013). This indicates that RNAseq of developing panicles at the IM and



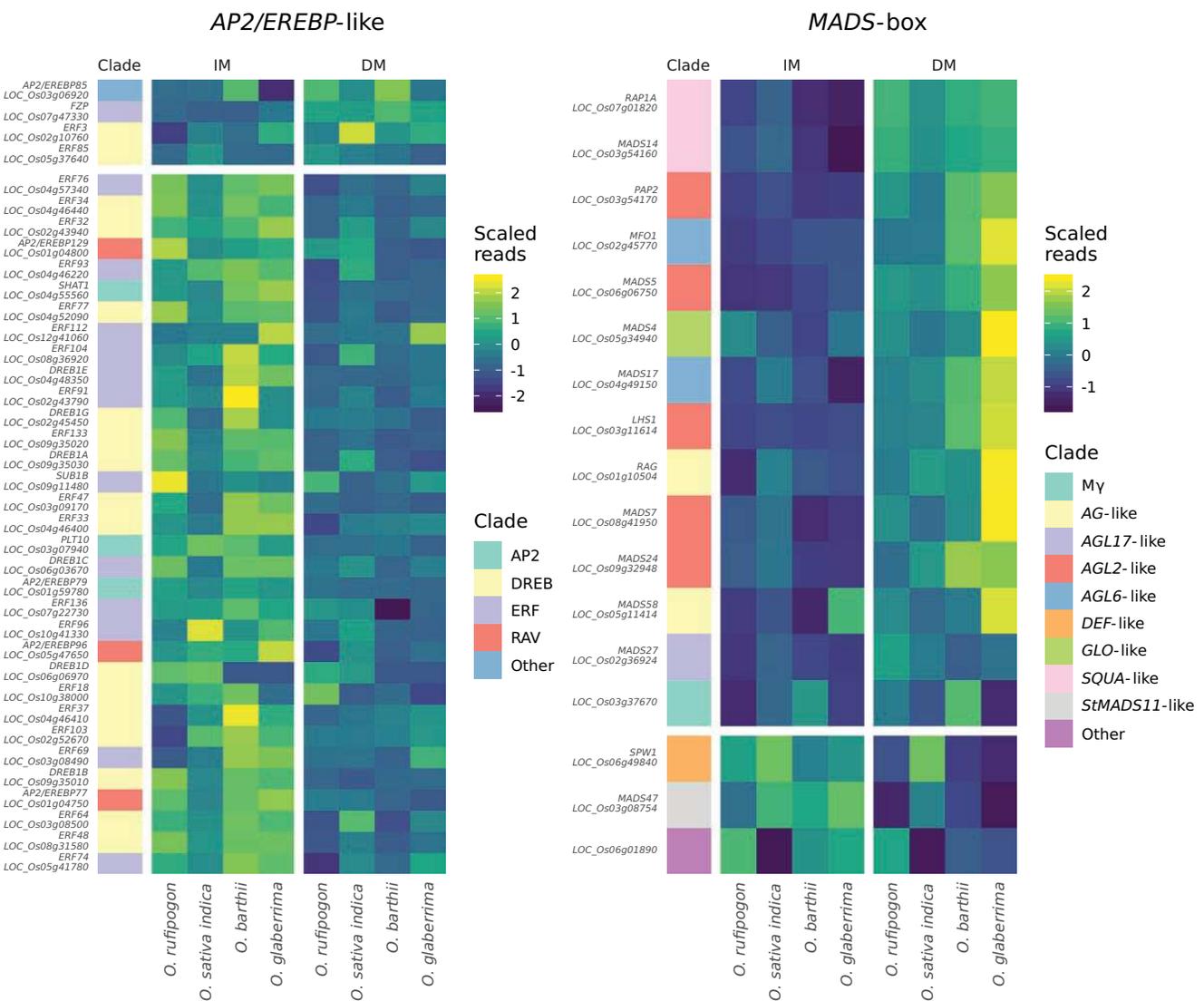
**Fig. 2.** Principal components analysis of transformed read counts for each library. Principal component 4 (PC4) separates RNAseq samples by developmental stage, and explains 4.6% of total variability. The first three components explain 83.1% of variability, and separate RNAseq samples by species. Bars show single samples (three replicates per accession per stage).

DM stage identifies genes that control branching, and suggests that transcription factors are prominent among these genes.

We used gene set enrichment analysis (GSEA; Subramanian *et al.*, 2005; Sergushichev, 2016, Preprint) on genes ranked by  $L_2FC$  to test for family-level enrichment among transcription factors. *AP2/EREBP*-like genes and MADS-box genes were both enriched ( $P_{adj}=3.5\times 10^{-5}$  and  $P_{adj}=4.0\times 10^{-4}$ , respectively, GSEA permutation  $t$ -test; Supplementary Table S6). Correspondingly, the list of 153 differentially expressed genes included 10 MADS-box genes and seven *AP2/EREBP*-like genes (Supplementary Table S5). Most differentially expressed MADS-box genes are more highly expressed at the DM stage, whereas most differentially expressed *AP2/EREBP*-like genes are more highly expressed at the IM stage. The majority of *AP2/EREBP*-like genes that have higher expression at the IM stage are from the ERF and DREB clades (Fig. 3). The DE

results are consistent with the role of transcriptional regulation in panicle branching, and highlight a set of candidate core regulators of axillary meristem determination and branching that are conserved between rice species. The pattern of expression of *AP2/EREBP*-like genes may indicate a role in the promotion of indeterminate axillary meristem identity or suppression of the transition from axillary meristem to spikelet meristem. MADS-box genes may have an inverse role as promoters of determinate meristem. Co-regulation of members of TF families, sometimes at the clade level, highlights the redundant or overlapping functions of TF families in meristem establishment and fate transition.

To test the role of *AP2/EREBP*-like genes in the control of panicle architecture, we phenotyped panicles from two loss-of-function mutants. The *crown rootless5 (crl5)* mutant of the *AP2*-like gene *PLETHORA 8 (PLT8)*; Kitomi *et al.*, 2011) produced

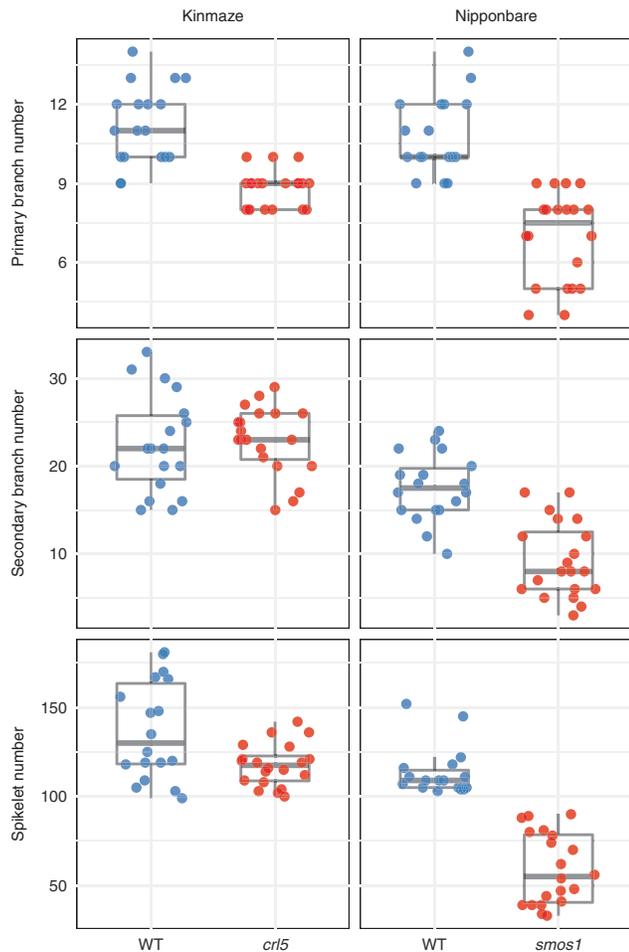


**Fig. 3.** *AP2/EREBP*-like and MADS-box transcription factors change expression between IM and DM. For each family, we plotted genes that were in the top 10% of all genes by absolute  $L_2FC$  between IM and DM, without filtering on adjusted  $P$ -value. Genes in the upper panels had higher expression at the DM stage, whilst genes in the lower panels had higher expression at the IM stage. Most *AP2/EREBP*-like genes that pass the 10% cutoff were more highly expressed in the IM. *AP2/EREBP*-like genes that were more highly expressed in IM mainly belong to the ERF and DREB clades. In contrast, most MADS-box genes that pass the cutoff were more highly expressed at the DM stage. Clades for *AP2/EREBP*-like genes are from the Plant Transcription Factor Database v4.0 (Jin *et al.*, 2017) and Sharoni *et al.* (2011). MADS-box clades were manually tabulated from Arora *et al.* (2007).

panicles with a shorter rachis with fewer primary branches (Fig. 4; Supplementary Table S7), consistent with a peak of *PLT8* expression in rachis meristem tissues from *O. sativa* ssp. *japonica* (Supplementary Fig. S6; Harrop *et al.*, 2016). Panicles from the *small organ size1* (*smos1*) mutant of *ERF142* (Aya *et al.*, 2014) have a reduced number of primary and secondary branches, and fewer spikelets (Fig. 4; Supplementary Table S7). *ERF142* expression is highest in primary branch and elongating primary branch meristem tissues in *O. sativa* ssp. *japonica* (Supplementary Fig. S6; Harrop *et al.*, 2016). Although neither gene was differentially expressed in our RNAseq dataset, these phenotypes support the involvement of AP2/EREBP-like genes in control of panicle architecture.

#### AP2/EREBP-like gene expression is associated with panicle diversity and domestication

To identify common patterns of expression of transcriptional regulators related to variation in panicle phenotypes, we used soft clustering of scaled L<sub>2</sub>FCs between IM and DM on the subset of annotated TF genes that were detected in our

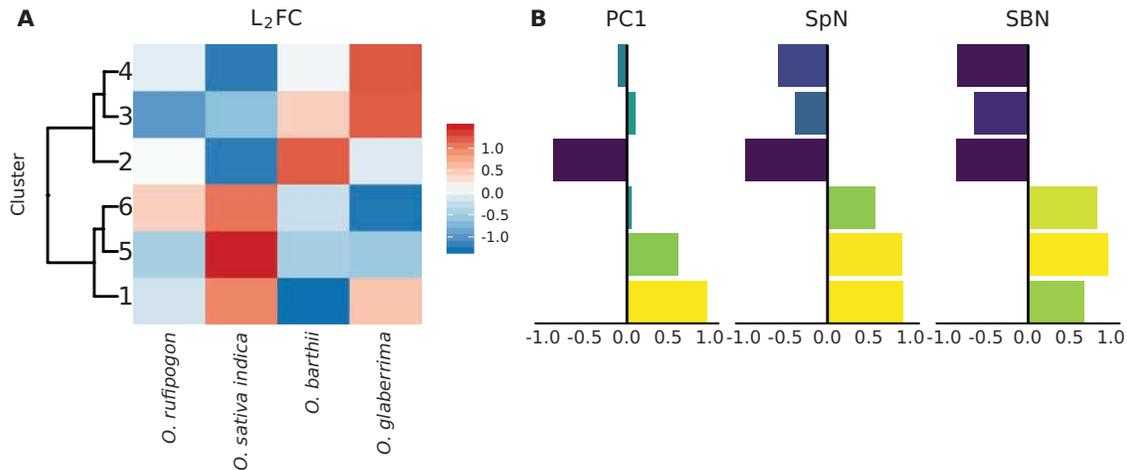


**Fig. 4.** Mutants in two AP2/EREBP-like genes, *PLT8* and *ERF142*, have defects in panicle architecture compared with their background accessions. The *crl5* mutant of *PLT8* (*LOC\_Os07g03250*) produced fewer primary branches and spikelets, and the *smos1-3* mutant of *ERF142* (*LOC\_Os05g32270*) produced fewer primary branches, secondary branches, and spikelets.

RNAseq dataset. We recovered six clusters comprising a total of 119 genes (Fig. 5; Supplementary Table S8). To determine which clusters were related to domestication, we calculated correlations between the mean L<sub>2</sub>FC value of genes in each cluster and PC1 in the phenotyping data, and between mean L<sub>2</sub>FC and the number of secondary branches and spikelets from repeat panicle phenotyping for the accessions used for RNAseq (Figs 1, 5B; Supplementary Fig. S7).

Clusters 3, 4, and 6 correlated with spikelet number (SpN) and secondary branch number (SBN), but not with PC1, meaning that the L<sub>2</sub>FC of genes in those clusters does not correlate with the phenotypic differences between wild and domesticated accessions. Clusters of genes with high L<sub>2</sub>FC in *O. sativa* ssp. *indica* have a positive correlation with SBN and SpN, whereas clusters of genes with low L<sub>2</sub>FC in *O. sativa* ssp. *indica* have a negative correlation with SBN and SpN. Cluster 4 had an enrichment of HB genes (9 out of 31 genes;  $P_{adj}=2.5\times 10^{-4}$ ). It also contained three MIKC<sup>C</sup>-type MADS-box genes (*LHS1*, *MFO1*, and *MADS14*), which promote spikelet meristem determination (Jeon *et al.*, 2000; Ohmori *et al.*, 2009; Kobayashi *et al.*, 2012), and three AP2/EREBP-like genes including *OsINDETERMINATE SPIKELET 1* (*OsIDS1*), which also controls inflorescence architecture (Chuck *et al.*, 2008; Lee and An, 2012). L<sub>2</sub>FCs of genes in this cluster are low in *O. sativa* ssp. *indica*, high in *O. glaberrima* and intermediate in the two wild accessions. Although these genes may be involved in regulation of panicle complexity, their expression did not appear to have changed in parallel in the two domestications. L<sub>2</sub>FCs of genes in clusters 3 and 6 change between accessions from the two continents. In cluster 3, L<sub>2</sub>FCs are higher in African species than in Asian species, meaning that the genes are more highly expressed in DM stages in African species. Genes in cluster 6 have the opposite pattern, with higher L<sub>2</sub>FCs in Asian species compared with African species. Cluster 3 contained *LAX1* and *ABERRANT PANICLE ORGANIZATION 2* (*FLO-LFY HOMOLOG OF RICE*, *AP02/RFL*), which are involved in axillary meristem establishment and outgrowth and promotion of indeterminate meristematic activity in rice, respectively (Komatsu *et al.*, 2001; Ikeda-Kawakatsu *et al.*, 2012). Their higher expression at the DM stage in panicles from both wild and domesticated African accessions could be associated with a reduced number of spikelets.

Clusters 1, 2, and 5 correlated with the main principal component (PC1) in the phenotyping data, which separates wild and domesticated species independently of continent. Clusters 1 and 5 are also positively correlated with spikelet number and secondary branch number, whereas cluster 2 is negatively correlated. The correlation of L<sub>2</sub>FC patterns with PC1 suggests that genes in these clusters may be associated with changes in panicle architecture between wild and domesticated species. Cluster 1 and cluster 5 both had a positive correlation with PC1. L<sub>2</sub>FCs are higher in domesticated accessions for genes in cluster 1, meaning that they are more highly expressed at the DM stage in domesticated accessions. Genes in cluster 2 have lower L<sub>2</sub>FCs in domesticated accessions, meaning that they are more highly expressed at the DM stage in wild accessions. Cluster 2 also had the strongest negative correlation with PC1, and low L<sub>2</sub>FCs in *O. sativa* ssp.



**Fig. 5.** Gene expression clusters correlate with the main component of diversity of panicle architecture (PC1) and the number of secondary branches (SBN) and spikelets (SpN). Clusters contained 19–31 genes each (Supplementary Table S8). (A) Mean, scaled log<sub>2</sub>-fold change (L<sub>2</sub>FC) of genes by cluster and accession. (B) Pearson correlation with PC1, SBN, and SpN. PC1 is the main principal component that separates panicles from domesticated and wild accessions of rice (Fig. 1). Correlations with SBN and SpN are based on repeat panicle phenotyping for the accessions used for RNAseq in greenhouse conditions (Supplementary Fig. S8; Supplementary Table S10).

*indica*. The lower L<sub>2</sub>FCs in domesticated accessions could implicate cluster 2 genes in promotion of determinate meristem fate, because their lower expression at the DM stage may result in more activity of indeterminate axillary meristems. Eight *GROWTH-REGULATING FACTOR1* (GRF) family genes, which are involved in the regulation of cell proliferation (Kim and Tsukaya, 2015), were detected in our dataset, and three of them were present in cluster 2 ( $P_{\text{adj}}=7.6\times 10^{-3}$ , hypergeometric test). In contrast to cluster 2, most of the genes in cluster 5 have L<sub>2</sub>FCs close to zero in *O. sativa ssp. indica*, and negative L<sub>2</sub>FCs in the other accessions (Supplementary Fig. S8), meaning that the expression of these genes decreases at the DM stage in all accessions except *O. sativa ssp. indica*. The lack of repression of cluster 5 genes and to a lesser extent cluster 1 genes at the DM stage in *O. sativa ssp. indica* could result in more branching via the promotion of indeterminate axillary meristem identity. Cluster 5, which had the highest mean L<sub>2</sub>FC in *O. sativa ssp. indica*, had an enrichment of *AP2/EREBP*-like genes (6 out of 17 genes;  $P_{\text{adj}}=7.6\times 10^{-3}$ , hypergeometric test), and cluster 1 also contains 3 *AP2/EREBP*-like genes (Supplementary Table S8). We used qPCR to confirm these patterns in all four stages of each accession for all *AP2/EREBP*-like genes in cluster 5 (Supplementary Fig. S9). The enrichment of *AP2/EREBP*-like genes in cluster 5 and the presence of three *AP2/EREBP*-like genes in cluster 1 suggest that their pattern of expression is associated with differences in panicle architecture across wild and domesticated accessions.

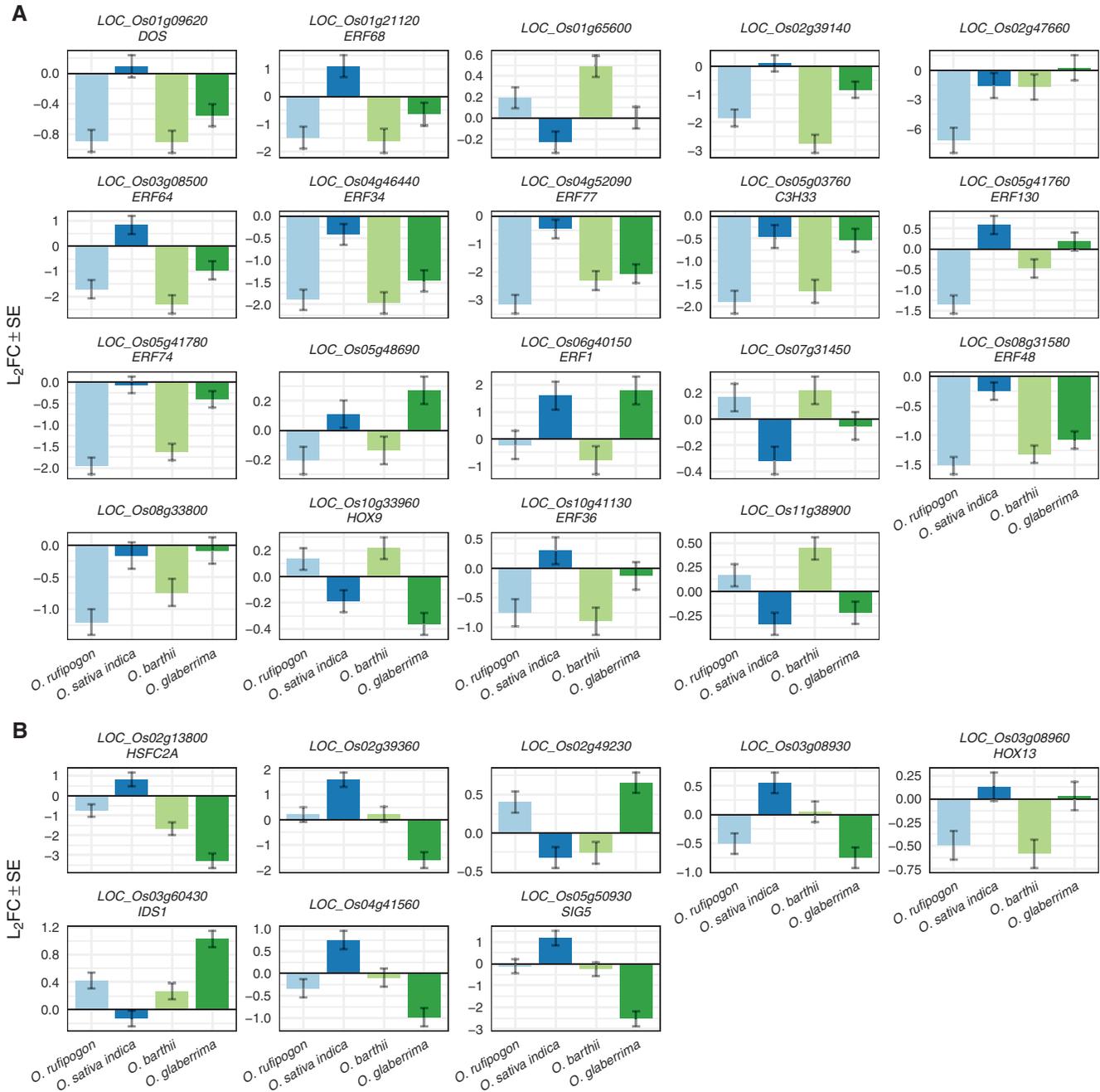
To find TF genes associated with parallel changes in panicle architecture during domestication, we tested the stage $\times$ domestication interaction for *O. rufipogon*, *O. sativa ssp. indica*, *O. barthii* and *O. glaberrima* at a false discovery rate of 0.1 (Supplementary Table S9). We detected 19 genes with a stage $\times$ domestication interaction, including nine *AP2/EREBP*-like genes ( $P=4.4\times 10^{-7}$ , hypergeometric test; Fig. 6A). These genes are putative targets of parallel selection on panicle architecture that occurred during domestication. *AP2/EREBP*-like genes were also prominent when we tested the stage $\times$ accession

interaction separately for each domestication (12 out of 85 genes in Asian accessions; 8 out of 50 genes in African accessions; Supplementary Table S9). Consistent with its presence in cluster 4, *OsIDS1* was also differentially expressed in both Asian and African domestications, although the direction of change was different (Fig. 6B). Genes with this pattern of expression in the four accessions may have also been targets of selection on panicle architecture, but evolved divergently.

The prominence of *AP2/EREBP*-like genes among putative core regulators of branching in all four *Oryza* species, and among genes associated with differences between wild and cultivated accessions, suggest that they were key targets of artificial selection for improvement in panicle architecture, and were involved in changes to the regulatory network controlling branching that occurred during domestication.

## Discussion

The purpose of this study was to identify genetic factors underlying the diversity of panicle architecture, which influences grain number in rice. Our transcriptomic comparison of panicles at indeterminate and determinate stages of axillary meristem development revealed a core set of transcription factors associated with axillary meristem phase transition in wild and domesticated African and Asian rice (Fig. 2). Some of these transcription factors, including *AP2/EREBP*-like and MIKCC-type MADS-box genes, appear to be co-regulated at the family or clade level (Fig. 3). Our phenotypic survey of 91 accessions showed that characteristics related to axillary meristem formation and fate transition are the main factors separating wild and domesticated rice (Fig. 1). Three clusters of gene expression correlated with the major component of phenotypic variability between wild and domesticated accessions, containing enrichments of *GRF* and *AP2/EREBP*-like transcription factor genes (Fig. 5). We observed a correlation between expression of *AP2/EREBP*-like transcription factors, domestication status, and derived phenotypes, suggesting



**Fig. 6.** Parallel and divergent evolution of gene expression associated with domestication. (A) Expression of genes with a stage×domestication interaction when all four accessions were tested together. We used this test to identify genes where the change in  $L_2FC$  between indeterminate (IM) and determinate (DM) stages changed in the same direction in both African and Asian domesticated accessions. (B) Expression of genes with a stage×domestication interaction when tested separately for African and Asian accessions. These genes have divergent changes between wild and domesticated accessions. The genes plotted in (B) had an interaction in both of the separate tests, but not in the single test used to identify the genes in (A).

that expression of these genes has changed as a result of artificial selection during domestication. As well as basic insights into molecular control of branching in rice, this work provides an overview of the outcome of the domestication process at phenotypic and whole-transcriptome levels.

A set of AP2/EREBP-like genes decrease in expression over the course of wheat spike development (Li *et al.*, 2018), and the molecular function of some individual AP2/EREBP-like genes has been reported in relation to inflorescence or root development. *FZP* represses the formation of axillary

meristem and induces transition from spikelet to floral meristem (Komatsu *et al.*, 2001). Along with the AP2/EREBP-like gene *FZP*, AP2/EREBP85, *ERF3*, and *ERF85* were more highly expressed at the DM stage across all four accessions, consistent with a role in the regulation of axillary meristem identity (Fig. 3). Conversely, most of the AP2/EREBP-like genes that change expression between stages had lower expression in DM than in IM, including mainly ERF and DREB clade genes and three RAV-like genes. In Arabidopsis, RAV orthologs repress flowering genes (Matías-Hernández *et al.*,

2014), and *ERF3* interacts with the HB gene *WOX11* to promote crown root development (Zhao *et al.*, 2015). Our results suggest additional roles for these genes in promoting reproductive axillary meristem initiation. This family- and clade-level view of their expression suggests that *AP2/EREBP*-like play a central role in the regulation of phase transition from indeterminate to determinate state.

*AP2/EREBP*-like genes are also associated with domestication. The *AP2*-like wheat domestication gene *Q* regulates inflorescence and glume shape and spike length, and has variation in binding sites for microRNA miR172 between wild and domesticated species (Simons *et al.*, 2006; Debernardi *et al.*, 2017; Greenwood *et al.*, 2017). Recently, characterization of a quantitative trait locus associated with grain yield and panicle branching and variation in the promoter region showed that *FZP* regulates panicle architecture and is associated with rice domestication (Bai *et al.*, 2017; Fujishiro *et al.*, 2018; Huang *et al.*, 2018). Our analysis indicates a modified expression pattern of some *AP2/EREBP*-like genes between wild and domesticated rice accessions. One of the clusters of genes with expression patterns that correlate to the main phenotypic differences between wild and domesticated accessions contained an enrichment of *AP2/EREBP*-like genes. In addition, nine out of the 19 genes with a stage $\times$ domestication interaction across the entire dataset were *AP2/EREBP*-like genes (Fig. 6A). Eight of these nine genes have lower expression in panicles at the DM stage, and may be involved in promotion of indeterminate axillary meristem and/or suppression of the transition from axillary meristem to spikelet meristem. Among these genes, *ERF130* (*MULTI-FLORET SPIKELET1*, *MFS1*) regulates the timing of the transition of spikelet meristems to terminal spikelets and positively regulates the expression of *OsIDS1* and *SNB* (Ren *et al.*, 2013). Taken with the roles of *AP2* mutants in modification of rice panicle architecture (this study; Komatsu *et al.*, 2001; Lee and An, 2012), our results implicate several *AP2/EREBP*-like genes as putative targets of artificial selection during rice domestication resulting in parallel evolution of expression.

*OsIDS1* controls panicle branching in rice, and positively regulates branch meristem identity by repressing genes that specify spikelet identity (Lee and An, 2012). We found a stage $\times$ domestication interaction in expression of *OsIDS1*, but the change was divergent between African and Asian accessions (Fig. 6B). This suggests that even if genes are involved in similar regulatory networks, their expression may diverge during selection. Parallel morphological evolution during Asian and African rice domestication is sometimes associated with different genomic modifications (Furuta *et al.*, 2015; Cubry *et al.*, 2018; Hu *et al.*, 2018; Lv *et al.*, 2018), illustrating that the interaction between genetic variation and the diverse selection pressures associated with domestication can result in diverse genomic outcomes.

Our interspecific analysis showed that parallel domestication of African and Asian rice resulted in similar modifications in panicle architecture, with changes in branch number contributing to derived phenotypes more than branch length traits (Fig. 1B, C). Secondary branch number correlated with spikelet number in all species, but primary branch number

only correlates with spikelet number in wild species (Fig. 1D). Yield improvements seem to have occurred mainly through increased ramification of branching, rather than, for example, elongation of the primary axis. This suggests that increased production of secondary branches from axillary meristems rather than direct differentiation into spikelets was important in both domestications.

The molecular control of phase transition from indeterminate to determinate axillary meristems is not fully understood in rice panicle development. Microarray analysis of developing panicles has identified a small set of differentially expressed genes, enriched for TFs (Furutani *et al.*, 2006). In microdissected meristem tissues, gene expression changes gradually during transition in axillary meristem identity (Harrop *et al.*, 2016), similar to the gradual meristem maturation during tomato inflorescence development (Park *et al.*, 2012). We also observed a small set of genes that were differentially expressed between IM and DM consistently in all accessions (153 of 25 229 tested), with an enrichment of TFs (37 of 153 genes). More genes were upregulated in DM than down-regulated, as in wheat spike development (Wang *et al.*, 2017). These results suggest that the gene regulatory network controlling phase transition from IM to DM is controlled by a small subset of core, conserved genes in Asian and African rice.

Because of redundant and overlapping activities of TFs, gene regulatory networks are robust to perturbations in single genes, allowing them to produce a stable transcriptional output in variable cellular and environmental conditions (Gitter *et al.*, 2009; Dai *et al.*, 2009; Wu and Lai, 2015). *AP2/EREBP*-like and MADS-box genes were enriched among differentially expressed genes. Expression of *AGL6*-like, *AP1/FUL*-like, and four of the five *SEPALLATA*-like MADS-box genes all increased between indeterminate and determinate stage, as expected given their roles in spikelet determination and inflorescence development (Yoshida and Nagato, 2011; Zhang and Yuan, 2014). This suggests family-level or clade-level co-regulation of TFs in rice panicle development. Our observation that some genes appear to be co-expressed at the family or clade level may indicate partial functional redundancy between groups of homologous TFs in the rice genome that act in the gene regulatory network controlling axillary meristem determination.

This work has revealed the core set of genes that are associated with the determination of axillary meristem identity across Asian and African rice accessions. We have also discovered candidate targets of artificial selection on panicle phenotype during domestication. In particular, the interspecific expression pattern of *AP2/EREBP*-like genes suggests that they regulate axillary meristem determination and have evolved during domestication. Along with efforts to characterize the molecular function of the candidate genes we have identified and to measure their expression in other accessions, further investigation of the gene regulatory network controlling meristem identity in developing panicles would continue to drive an understanding of the process of inflorescence development in rice and its relationship to grain yield.

## Supplementary data

Supplementary data are available at *JXB* online.

Fig. S1. Spread mature rice panicle.

Fig. S2. Principal component analysis (PCA) of panicle phenotyping data showing components 1–4.

Fig. S3. The accessions used for RNAseq are consistent with species-wide patterns of panicle architecture.

Fig. S4. Early stages of rice panicle development used for gene expression analysis.

Fig. S5. Heatmap of pairwise distances between RNAseq samples.

Fig. S6. Expression of *AP2/EREBP*-like genes in *O. sativa* spp. *japonica* cv. Nipponbare meristems (data from Harrop *et al.*, 2016).

Fig. S7. Phenotyping of the five *Oryza* accessions used for RNAseq.

Fig. S8. Most genes in cluster 5 have negative  $L_2FCs$  between IM and DM in *O. rufipogon*, *O. barthii*, and *O. glaberrima*, but  $L_2FCs$  in *O. sativa* spp. *indica* are closer to zero.

Fig. S9. Expression analysis along early panicle development of *AP2/EREBP*-like genes present in cluster 5.

Table S1. Rice accessions used in this study.

Table S2. Sequences of primers used.

Table S3. Quantification of panicle traits in 91 accessions from wild and domesticated Asian and African rice species.

Table S4. Read and mapping statistics for all RNAseq samples.

Table S5. Differential expression test results between stages across all species; we used an arbitrary differential expression threshold of 1.5-fold change in expression and adjusted *P*-value (false discovery rate) less than 0.1.

Table S6. Transcription factor family enrichment by  $L_2FC$ .

Table S7. Quantification of panicle traits from *crl5* and *smos1* mutants.

Table S8. Clustered genes.

Table S9. Differential expression test results for the stage $\times$ accession interaction in Asian and African accessions.

Table S10. Detailed quantification of panicle traits from rice accessions used for sequencing analysis.

## Acknowledgements

Dr Ko Hirano and Prof. Yoshiaki Inukai from Nagoya University, Japan, kindly provided *smos1-3* and *crl5* mutant seeds, respectively. We thank Lady Johanna Arbelaez Rivera (CIAT, Colombia) for plant care and Céline Cardi and Hélène Vignes (Grand plateau technique régional de génotypage, CIRAD, Montpellier) for support with high-throughput qPCR and cDNA library preparation. Sophie Chéron and Harold Chrestin assisted with plant care and phenotyping, and Christine Tranchant-Dubreuil, Axel Verdier, and the IRD bioinformatics platform assisted with technical support in Montpellier. We also thank James Tregear for feedback on the manuscript and Alain Ghesquière for discussion on the project.

## Funding

This research was funded by Agropolis Foundation through the *Investissements d'avenir* programme (ANR-10-LABX-0001-01),

Fondazione Cariplo (EVOREPRICE 1201-004), and the CGIAR Research Program on Rice. We also received support from the Excellence Scholarship Program of the Embassy of France in Vietnam (LAM).

## Author contributions

HA and SJ designed the research with input from TWRH and OM; HA, TWRH, OM, KB, and AML performed the research; TWRH, OM, HA, and ML performed data analysis, collection, and interpretation; TWRH, SJ, and HA wrote the manuscript.

## References

- Agrawal GK, Abe K, Yamazaki M, Miyao A, Hirochika H. 2005. Conservation of the E-function for floral organ identity in rice revealed by the analysis of tissue culture-induced loss-of-function mutants of the *OsMADS1* gene. *Plant Molecular Biology* **59**, 125–135.
- Al-Tam F, Adam H, Anjos Ad, Lorieux M, Larmande P, Ghesquière A, Jouannic S, Shahbazzkia HR. 2013. P-TRAP: a panicle trait phenotyping tool. *BMC Plant Biology* **13**, 122.
- Arora R, Agarwal P, Ray S, Singh AK, Singh VP, Tyagi AK, Kapoor S. 2007. MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genomics* **8**, 242.
- Aya K, Hobo T, Sato-Izawa K, Ueguchi-Tanaka M, Kitano H, Matsuoka M. 2014. A novel AP2-type transcription factor, SMALL ORGAN SIZE1, controls organ size downstream of an auxin signaling pathway. *Plant & Cell Physiology* **55**, 897–912.
- Bai X, Huang Y, Hu Y, Liu H, Zhang B, Smaczniak C, Hu G, Han Z, Xing Y. 2017. Duplication of an upstream silencer of *FZP* increases grain yield in rice. *Nature Plants* **3**, 885–893.
- Bai X, Huang Y, Mao D, Wen M, Zhang L, Xing Y. 2016. Regulatory role of *FZP* in the determination of panicle branching and spikelet formation in rice. *Scientific Reports* **6**, 19022.
- Bai X, Wu B, Xing Y. 2012. Yield-related QTLs and their applications in rice genetic improvement. *Journal of Integrative Plant Biology* **54**, 300–311.
- Chen ZX, Wu JG, Ding WN, Chen HM, Wu P, Shi CH. 2006. Morphogenesis and molecular basis on *Naked Seed Rice*, a novel homeotic mutation of *OsMADS1* regulating transcript level of *AP3* homologue in rice. *Planta* **223**, 882–890.
- Choi JY, Platts AE, Fuller DQ, Hsing YI, Wing RA, Purugganan MD. 2017. The rice paradox: multiple origins but single domestication in Asian rice. *Molecular Biology and Evolution* **34**, 969–979.
- Chuck G, Meeley R, Hake S. 2008. Floral meristem initiation and meristem cell fate are regulated by the maize *AP2* genes *ids1* and *sid1*. *Development* **135**, 3013–3019.
- Cubry P, Tranchant-Dubreuil C, Thuillet AC, *et al.* 2018. The rise and fall of African rice cultivation revealed by analysis of 246 new genomes. *Current Biology* **28**, 2274–2282.e6.
- Dai Z, Dai X, Xiang Q, Feng J. 2009. Robustness of transcriptional regulatory program influences gene expression variability. *BMC Genomics* **10**, 573.
- Debernardi JM, Lin H, Chuck G, Faris JD, Dubcovsky J. 2017. microRNA172 plays a crucial role in wheat spike morphogenesis and grain threshability. *Development* **144**, 1966–1975.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
- Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* **127**, 1309–1321.
- Doust A. 2007. Architectural evolution and its implications for domestication in grasses. *Annals of Botany* **100**, 941–950.
- Duan P, Ni S, Wang J, Zhang B, Xu R, Wang Y, Chen H, Zhu X, Li Y. 2015. Regulation of *OsGRF4* by *OsmiR396* controls grain size and yield in rice. *Nature Plants* **2**, 15203.

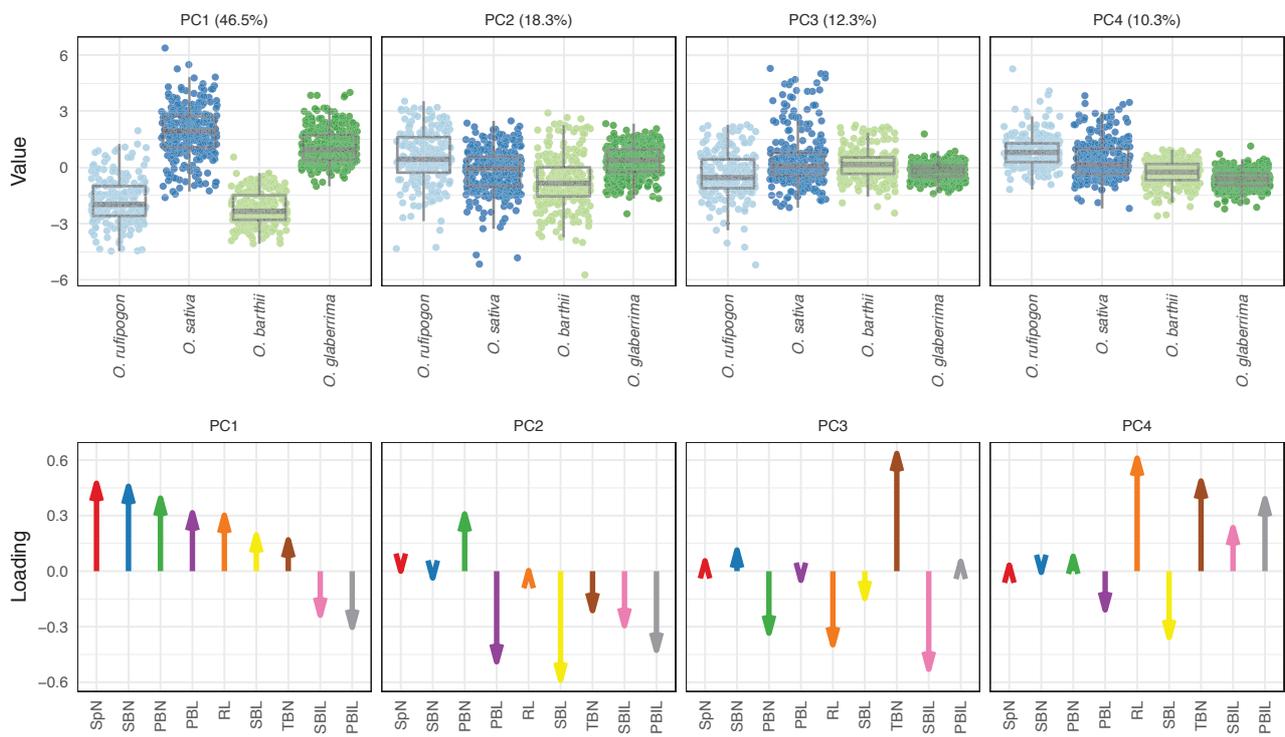
- Fujishiro Y, Agata A, Ota S, Ishihara R, Takeda Y, Kunishima T, Ikeda M, Kyozyuka J, Hobo T, Kitano H. 2018. Comprehensive panicle phenotyping reveals that *qSm7/FZP* influences higher-order branching. *Scientific Reports* **8**, 12511.
- Furuta T, Komeda N, Asano K, *et al.* 2015. Convergent loss of awn in two cultivated rice species *Oryza sativa* and *Oryza glaberrima* is caused by mutations in different loci. *G3* **5**, 2267–2274.
- Furutani I, Sukegawa S, Kyozyuka J. 2006. Genome-wide analysis of spatial and temporal gene expression in rice panicle development. *The Plant Journal* **46**, 503–511.
- Gitter A, Siegfried Z, Klutstein M, Fornes O, Oliva B, Simon I, Bar-Joseph Z. 2009. Backup in gene regulatory networks explains differences between binding and knockout results. *Molecular Systems Biology* **5**, 276.
- Greenwood JR, Finnegan EJ, Watanabe N, Trevaskis B, Swain SM. 2017. New alleles of the wheat domestication gene *Q* reveal multiple roles in growth and reproductive development. *Development* **144**, 1959–1965.
- Harrop TW, Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, Kater MM. 2016. Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *The Plant Journal* **86**, 75–88.
- Hu M, Lv S, Wu W, *et al.* 2018. The domestication of plant architecture in African rice. *The Plant Journal* **94**, 661–669.
- Huang X, Kurata N, Wei X, *et al.* 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**, 497–501.
- Huang Y, Zhao S, Fu Y, *et al.* 2018. Variation in the regulatory region of *FZP* causes increases in secondary inflorescence branching and grain yield in rice domestication. *The Plant Journal* **96**, 716–733.
- Ikeda M, Miura K, Aya K, Kitano H, Matsuoka M. 2013. Genes offering the potential for designing yield-related traits in rice. *Current Opinion in Plant Biology* **16**, 213–220.
- Ikeda-Kawakatsu K, Maekawa M, Izawa T, Itoh J, Nagato Y. 2012. *ABERRANT PANICLE ORGANIZATION 2/RFL*, the rice ortholog of Arabidopsis *LEAFY*, suppresses the transition from inflorescence meristem to floral meristem through interaction with *APO1*. *The Plant Journal* **69**, 168–180.
- Itoh J, Nonomura K, Ikeda K, Yamaki S, Inukai Y, Yamagishi H, Kitano H, Nagato Y. 2005. Rice plant development: from zygote to spikelet. *Plant & Cell Physiology* **46**, 23–47.
- Jeon JS, Jang S, Lee S, *et al.* 2000. *Leafy Hull Sterile1* is a homeotic mutation in a rice *MADS* box gene affecting rice flower development. *The Plant Cell* **12**, 871–884.
- Jiao Y, Wang Y, Xue D, *et al.* 2010. Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. *Nature Genetics* **42**, 541–544.
- Jin J, Tian F, Yang DC, Meng YQ, 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**, D1040–D1045.
- Khanday I, Yadav SR, Vijayraghavan U. 2013. Rice *LHS1/OsMADS1* controls floret meristem specification by coordinated regulation of transcription factors and hormone signaling pathways. *Plant Physiology* **161**, 1970–1983.
- Kim JH, Tsukaya H. 2015. Regulation of plant growth and development by the GROWTH-REGULATING FACTOR and GRF-INTERACTING FACTOR duo. *Journal of Experimental Botany* **66**, 6093–6107.
- Kitomi Y, Ito H, Hobo T, Aya K, Kitano H, Inukai Y. 2011. The auxin responsive AP2/ERF transcription factor *CROWN ROOTLESS5* is involved in crown root initiation in rice through the induction of *OsRR1*, a type-A response regulator of cytokinin signaling. *The Plant Journal* **67**, 472–484.
- Kobayashi K, Maekawa M, Miyao A, Hirochika H, Kyozyuka J. 2010. *PANICLE PHYTOMER2 (PAP2)*, encoding a SEPALLATA subfamily MADS-box protein, positively controls spikelet meristem identity in rice. *Plant & Cell Physiology* **51**, 47–57.
- Kobayashi K, Yasuno N, Sato Y, Yoda M, Yamazaki R, Kimizu M, Yoshida H, Nagamura Y, Kyozyuka J. 2012. Inflorescence meristem identity in rice is specified by overlapping functions of three *AP1/FUL*-like MADS box genes and *PAP2*, a *SEPALLATA* MADS box gene. *The Plant Cell* **24**, 1848–1859.
- Komatsu M, Maekawa M, Shimamoto K, Kyozyuka J. 2001. The *LAX1* and *FRIZZY PANICLE 2* genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Developmental Biology* **231**, 364–373.
- Köster J, Rahmann S. 2012. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522.
- Kumar L, E Futschik M. 2007. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* **2**, 5–7.
- Kurtzer GM, Sochat V, Bauer MW. 2017. Singularity: Scientific containers for mobility of compute. *PLoS ONE* **12**, e0177459.
- Kyozyuka J, Tokunaga H, Yoshida A. 2014. Control of grass inflorescence form by the fine-tuning of meristem phase change. *Current Opinion in Plant Biology* **17**, 110–115.
- Lee DY, An G. 2012. Two AP2 family genes, *SUPERNUMERARY BRACT (SNB)* and *OsINDETERMINATE SPIKELET 1 (OsIDS1)*, synergistically control inflorescence architecture and floral meristem establishment in rice. *The Plant Journal* **69**, 445–461.
- Lee DY, Lee J, Moon S, Park SY, An G. 2007. The rice heterochronic gene *SUPERNUMERARY BRACT* regulates the transition from spikelet meristem to floral meristem. *The Plant Journal* **49**, 64–78.
- Li Y, Fu X, Zhao M, Zhang W, Li B, An D, Li J, Zhang A, Liu R, Liu X. 2018. A genome-wide view of transcriptome dynamics during early spike development in bread wheat. *Scientific Reports* **8**, 15338.
- Linares OF. 2002. African rice (*Oryza glaberrima*): History and future potential. *Proceedings of the National Academy of Sciences, USA* **99**, 16360–16365.
- Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  method. *Methods* **25**, 402–408.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550.
- Lv S, Wu W, Wang M, *et al.* 2018. Genetic control of seed shattering during African rice domestication. *Nature Plants* **4**, 331–337.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12.
- Matías-Hernández L, Aguilar-Jaramillo AE, Marín-González E, Suárez-López P, Pelaz S. 2014. *RAV* genes: regulation of floral induction and beyond. *Annals of Botany* **114**, 1459–1470.
- Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M. 2010. *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nature Genetics* **42**, 545–549.
- Ohmori S, Kimizu M, Sugita M, Miyao A, Hirochika H, Uchida E, Nagato Y, Yoshida H. 2009. *MOSAIC FLORAL ORGANS1*, an *AGL6*-like MADS box gene, regulates floral organ identity and meristem fate in rice. *The Plant Cell* **21**, 3008–3025.
- Olsen KM, Wendel JF. 2013. Crop plants as models for understanding plant adaptation and diversification. *Frontiers in Plant Science* **4**, 290.
- Ouyang S, Zhu W, Hamilton J, *et al.* 2007. The TIGR rice genome annotation resource: improvements and new features. *Nucleic Acids Research* **35**, D883–D887.
- Park SJ, Jiang K, Schatz MC, Lippman ZB. 2012. Rate of meristem maturation determines inflorescence architecture in tomato. *Proceedings of the National Academy of Sciences, USA* **109**, 639–644.
- Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LG, Rensing SA, Kersten B, Mueller-Roeber B. 2010. *PlnTFDB*: updated content and new features of the plant transcription factor database. *Nucleic Acids Research* **38**, D822–D827.
- Ren D, Li Y, Zhao F, *et al.* 2013. *MULTI-FLORET SPIKELET1*, which encodes an AP2/ERF protein, determines spikelet meristem fate and sterile lemma identity in rice. *Plant Physiology* **162**, 872–884.
- Sergushichev A. 2016. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv* 060012. [Preprint].
- Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Sasaya T, Choi IR, Omura T, Kikuchi S. 2011. Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant & Cell Physiology* **52**, 344–360.
- Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai YS, Gill BS, Faris JD. 2006. Molecular characterization of the major wheat domestication gene *Q*. *Genetics* **172**, 547–555.
- Stein JC, Yu Y, Copetti D, *et al.* 2018. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nature Genetics* **50**, 285–296.

- Subramanian A, Tamayo P, Mootha VK et al.** 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences, USA* **102**, 15545–15550.
- Ta KN, Adam H, Staedler YM, Schöenberger J, Harrop T, Tregear J, Do NV, Gantet P, Ghesquière A, Jouannic S.** 2017. Differences in meristem size and expression of branching genes are associated with variation in panicle phenotype in wild and domesticated African rice. *EvoDevo* **8**, 2.
- Tabuchi H, Zhang Y, Hattori S, et al.** 2011. *LAX PANICLE2* of rice encodes a novel nuclear protein and regulates the formation of axillary meristems. *The Plant Cell* **23**, 3276–3287.
- Taguchi-Shiobara F, Kojima Y, Ebitani T, Yano M, Ebana K.** 2011. Variation in domesticated rice inflorescence architecture revealed by principal component analysis and quantitative trait locus analysis. *Breeding Science* **61**, 52–60.
- Teo ZW, Song S, Wang YQ, Liu J, Yu H.** 2014. New insights into the regulation of inflorescence architecture. *Trends in Plant Science* **19**, 158–165.
- Vaughan DA, Lu B-R, Tomooka N.** 2008. The evolving story of rice evolution. *Plant Science* **174**, 394–408.
- Wang Y, Li J.** 2011. Branching in rice. *Current Opinion in Plant Biology* **14**, 94–99.
- Wang Y, Yu H, Tian C, Sajjad M, Gao C, Tong Y, Wang X, Jiao Y.** 2017. Transcriptome association identifies regulators of wheat spike architecture. *Plant Physiology* **175**, 746–757.
- Wu WS, Lai FJ.** 2015. Functional redundancy of transcription factors explains why most binding targets of a transcription factor are not affected when the transcription factor is knocked out. *BMC Systems Biology* **9 Suppl 6**, S2.
- Xing Y, Zhang Q.** 2010. Genetic and molecular bases of rice yield. *Annual Review of Plant Biology* **61**, 421–442.
- Yamaki S, Miyabayashi T, Eiguchi M, Kitano H, Nonomura K-I, Kurata N.** 2010. Diversity of panicle branching patterns in wild relatives of rice. *Breeding Science* **60**, 586–596.
- Yoshida A, Sasao M, Yasuno N et al.** 2013. *TAWAWA1*, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proceedings of the National Academy of Sciences, USA* **110**, 767–772.
- Yoshida H, Nagato Y.** 2011. Flower development in rice. *Journal of Experimental Botany* **62**, 4719–4730.
- Zhang D, Yuan Z.** 2014. Molecular control of grass inflorescence development. *Annual Review of Plant Biology* **65**, 553–578.
- Zhao Y, Cheng S, Song Y, Huang Y, Zhou S, Liu X, Zhou DX.** 2015. The interaction between rice ERF3 and WOX11 promotes crown root development by regulating gene expression involved in cytokinin signaling. *The Plant Cell* **27**, 2469–2483.
- Zhu Q, Ge S.** 2005. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* **167**, 249–265.

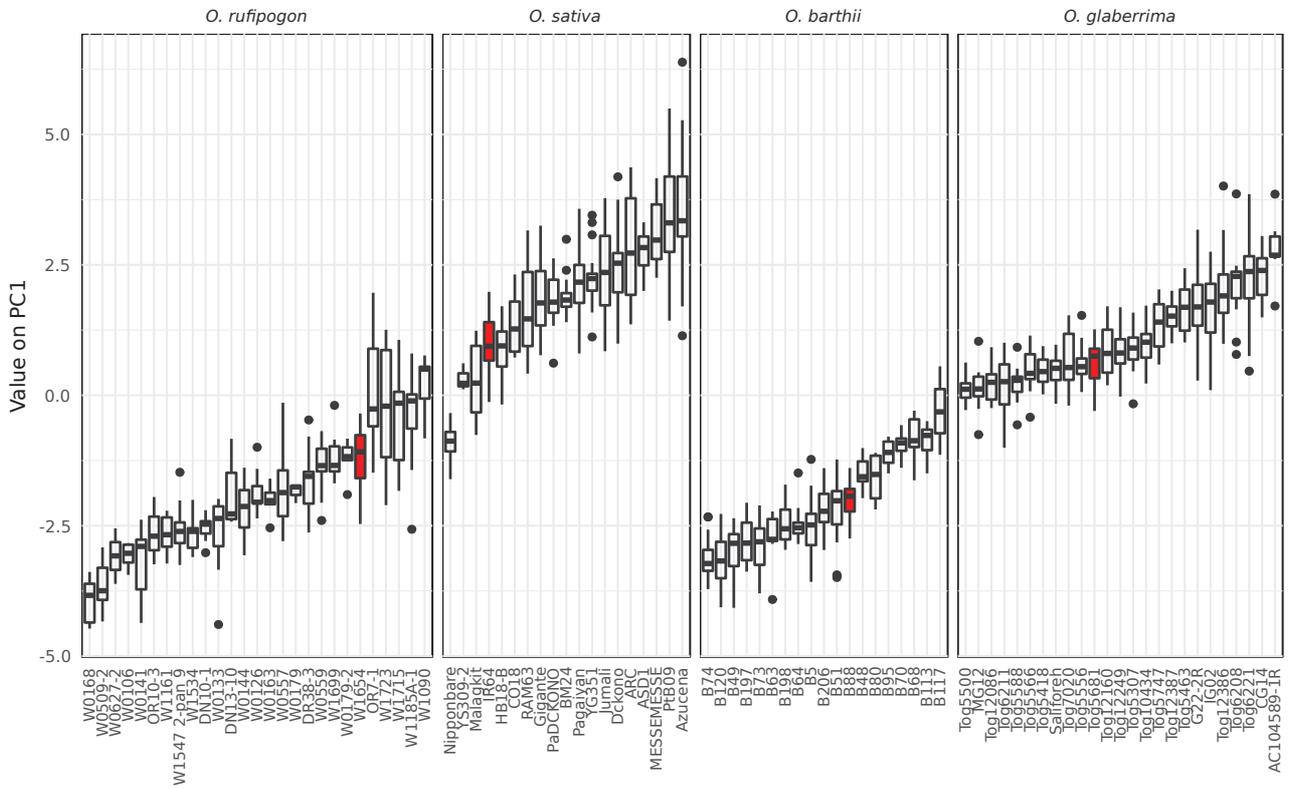
## Supporting information



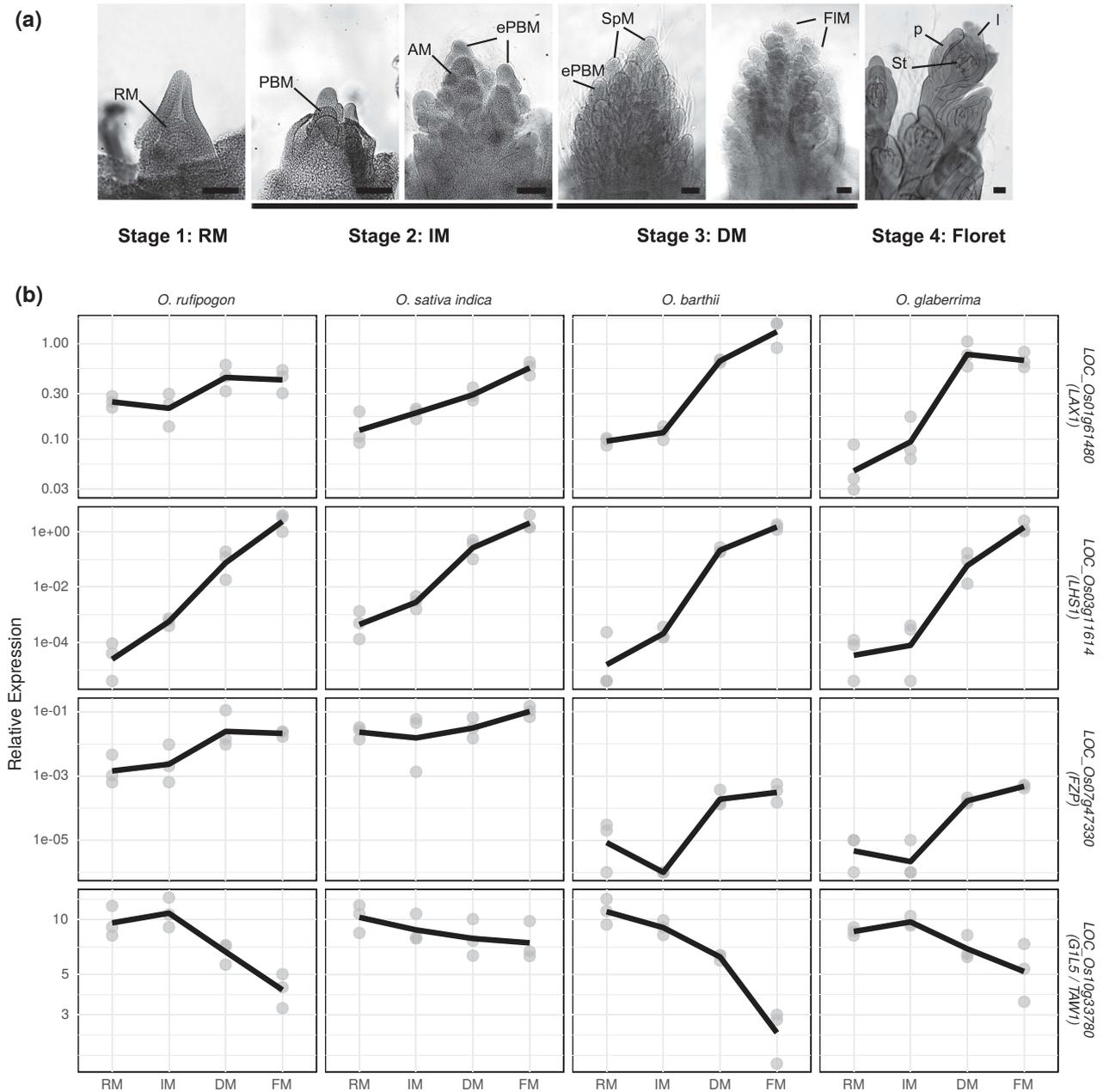
**Figure S1.** Spread mature rice panicle. PB: Primary branch; PBIL: Primary branch internode length; PBL: Primary branch length; RL: Rachis length; SB: Secondary branch; SBIL: Secondary branch internode length; SBL: Secondary branch length; Sp: Spikelet.



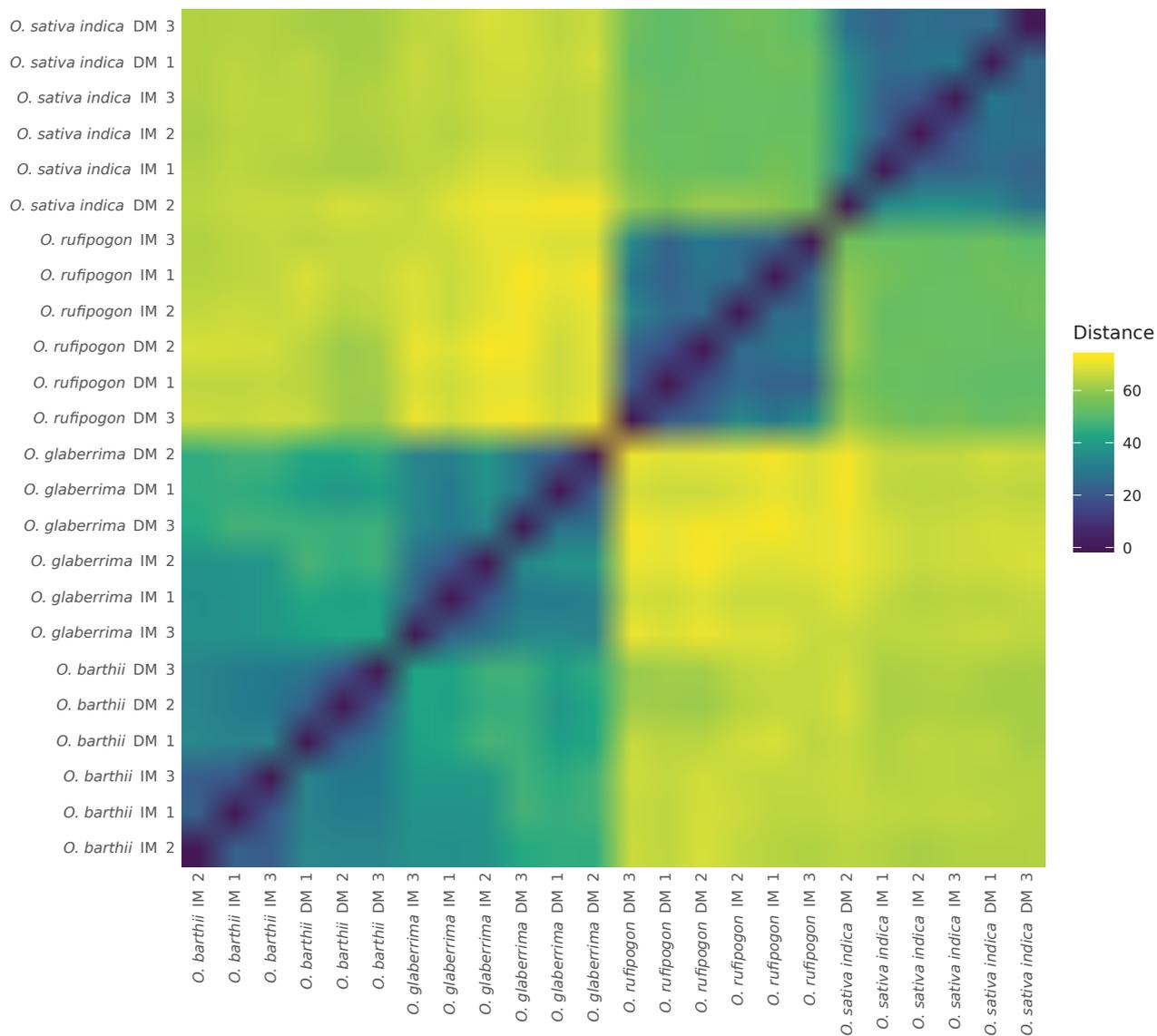
**Figure S2.** Principal component analysis (PCA) of panicle phenotyping data showing components 1–4. PC1 accounts for 46.5% of variability and separates panicles from domesticated and wild accessions. The lower ordinated do not separate panicles by species.



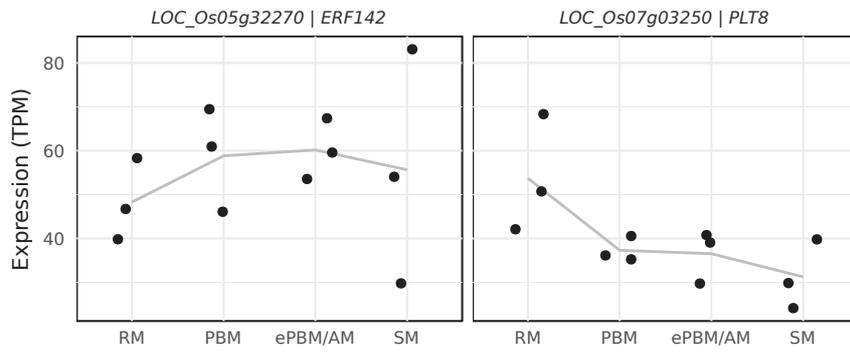
**Figure S3.** The accessions used for RNAseq are consistent with species-wide patterns of panicle architecture. The *y*-axis shows the projection of each panicle on principal component 1 (PC1), which separates wild and domesticated accessions (Fig. 1). The accessions chosen for RNAseq are shown in red. Accessions used for phenotyping are listed in supporting information Table S1.



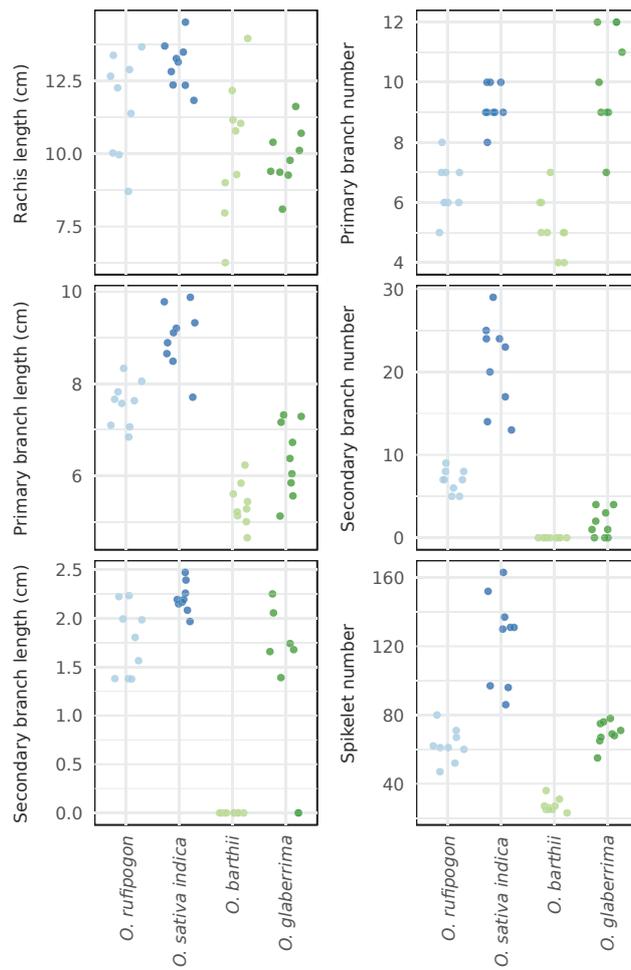
**Figure S4.** Early stages of rice panicle development used for gene expression analysis. **(a)** Developmental stages of immature panicles collected for expression analysis. Stage 1: rachis meristem; Stage 2: indeterminate meristem (IM) stage with formation of primary branch meristems, elongation of primary branch meristem and formation of axillary meristem; Stage 3: determinate meristem (DM) stage with spikelet meristem and floret differentiation; Stage 4: floret displaying early floral organ differentiation. The scale bar indicates 100  $\mu$ m. **(b)** Quantitative RT-PCR using meristem stage-specific marker genes for validation of staging. AM: axillary meristem; ePBM: elongating primary branch meristem; FIM: floret meristem; l: lemma; p: palea; PBM: primary branch meristem; RM: Rachis meristem; SpM: spikelet meristem; St: stamen.



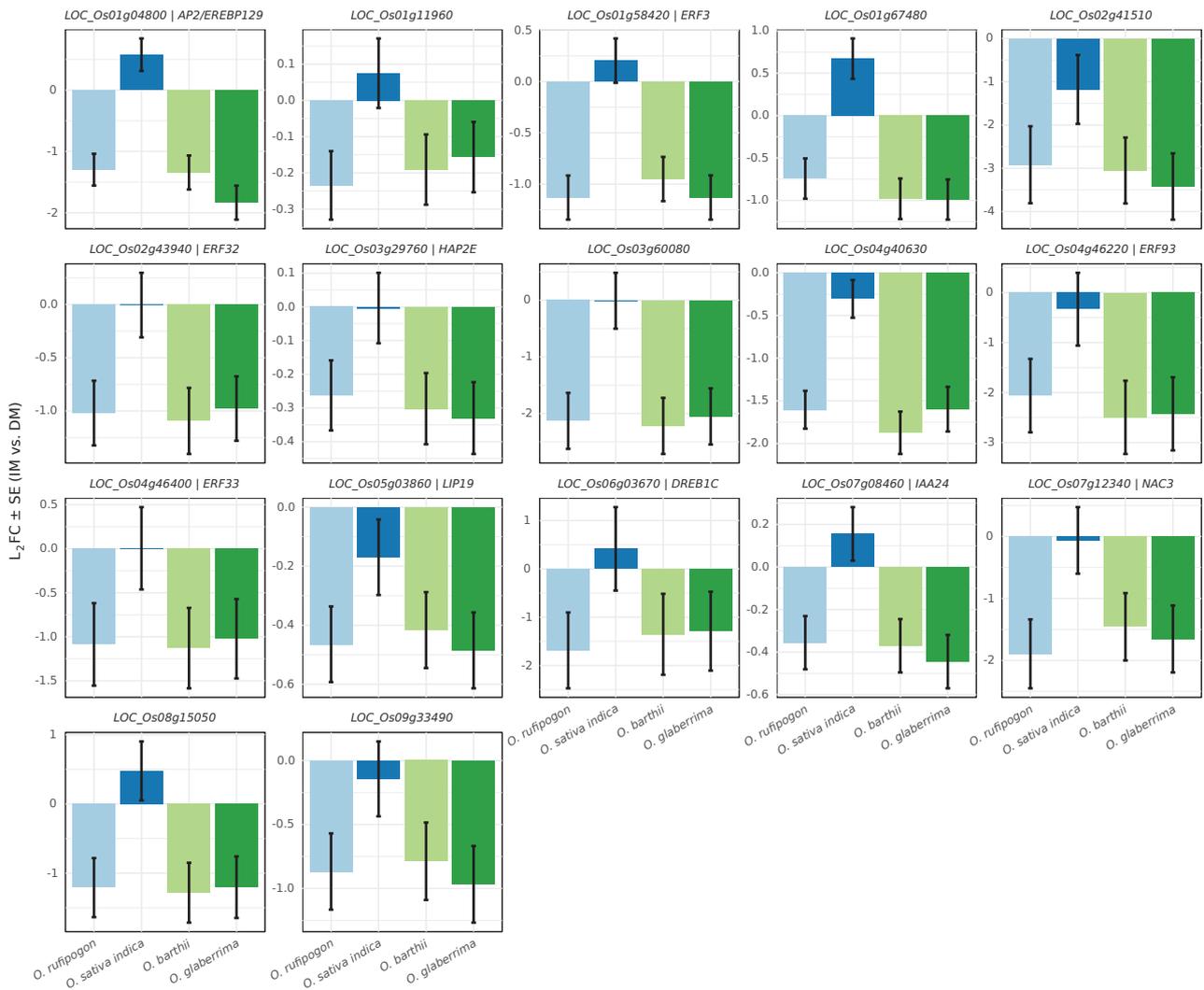
**Figure S5.** Heatmap of pairwise distances between RNAseq samples. Samples group by stage, species and continent. The numbers indicate single samples (three replicates per accession per stage). The axes are ordered by hierarchical clustering of Minkowski distances between samples.



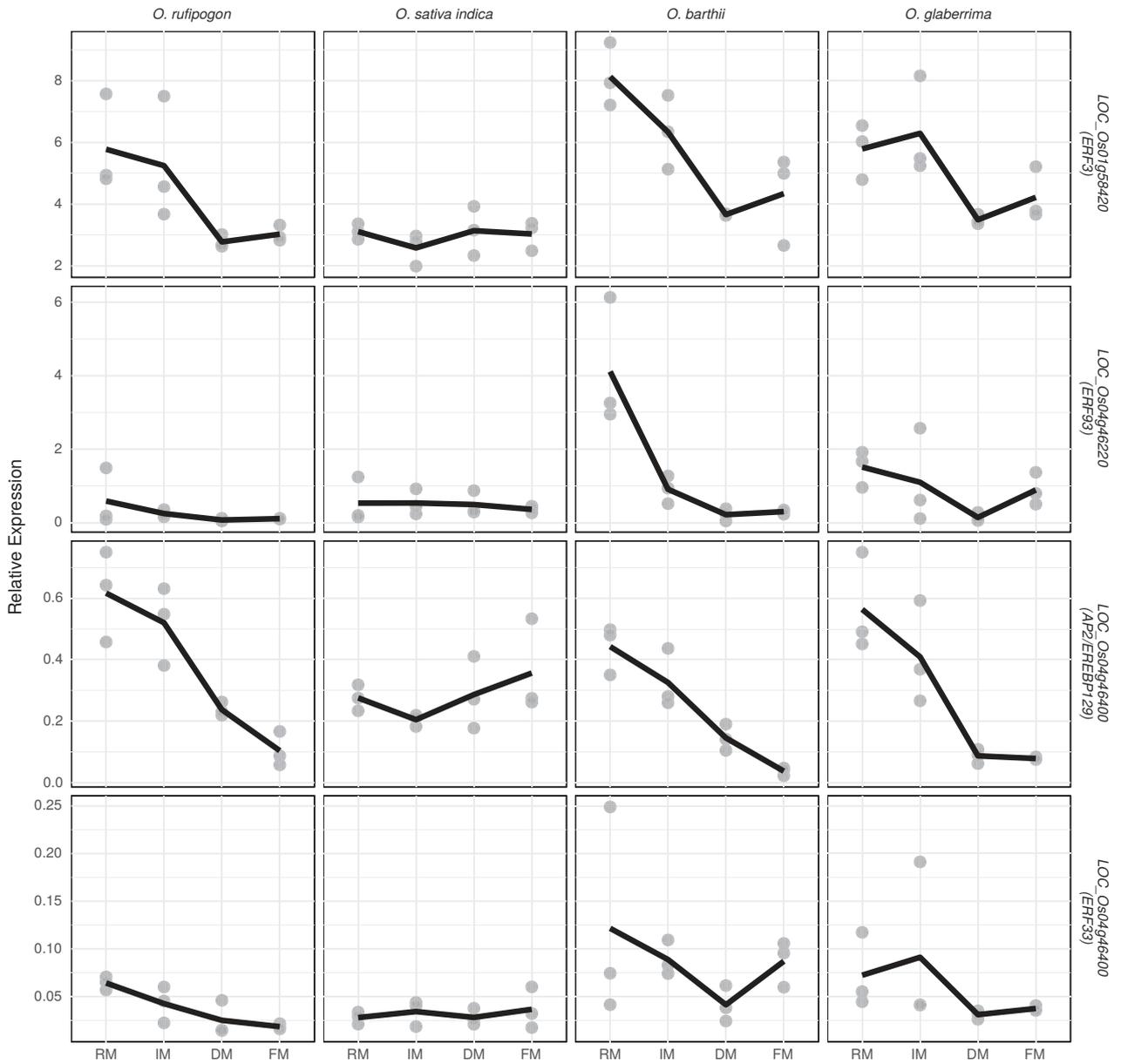
**Figure S6.** Expression of *AP2/EREBP*-like genes in *O. sativa japonica* cv. Nipponbare meristems (data from Harrop *et al.*, 2016). Both genes are expressed at all stages. *PLT8* expression peaks in RM. RM, rachis meristem; PBM, primary branch meristem; ePBM/AM, extending primary branch meristem and axillary meristem; SM, spikelet meristem.



**Figure S7.** Phenotyping of the four *Oryza* accessions used for RNAseq. These plants were grown at the same time and in the same conditions as the plants used for gene expression analysis. The domesticated accessions produce more spikelets and secondary branches than their wild relatives. The domesticated accessions have a similar number of primary branches, but the Asian domesticated species has more secondary branches and spikelets than the domesticated African species.



**Figure S8.** Most genes in cluster 5 have negative  $L_2FC$ s between IM and DM in *O. rufipogon*, *O. barthii* and *O. glaberrima*, but  $L_2FC$ s in *O. sativa indica* are closer to zero. This cluster has an enrichment of *AP2/EREBP*-like genes.



**Figure S9.** Expression analysis along early panicle development of *AP2/EREBP*-like genes present in cluster 5. DM: determinate meristem; FM: floret meristem; IM: indeterminate meristem; RM: rachis meristem.

## GENERAL CONCLUSIONS AND PERSPECTIVES

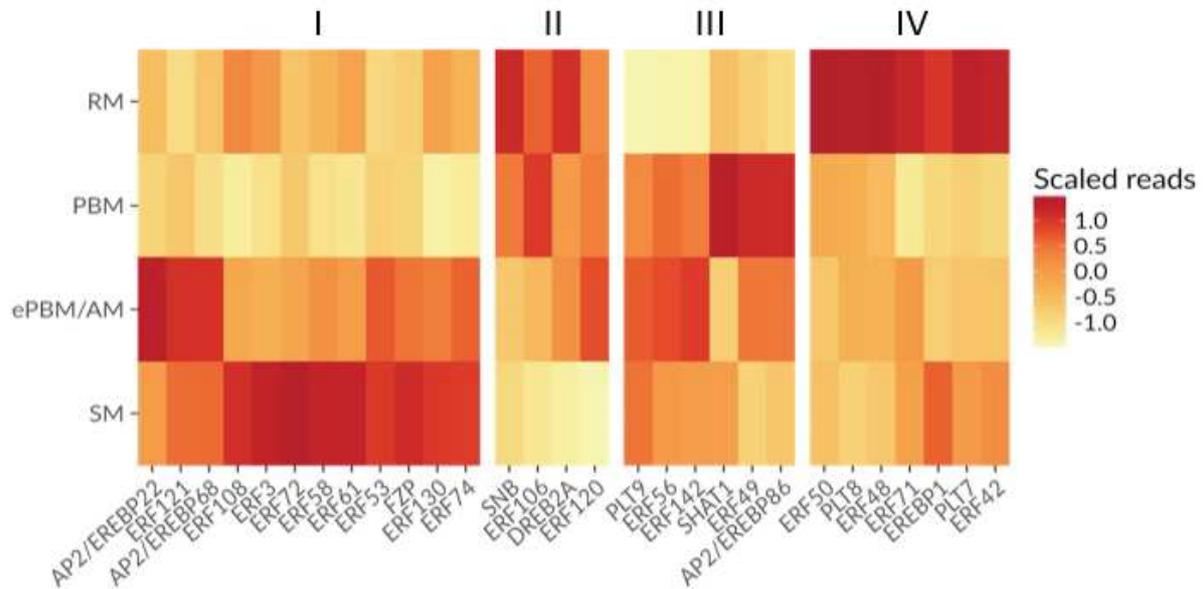
The analysis of a multi-species RNA-seq dataset (5-acc dataset) indicated a significant enrichment for *AP2/ERF* genes in the sets of genes associated with axillary meristem identity and phenotypic variability between wild and domesticated accessions in Asian and African rice respectively. Moreover, the panicle phenotypes of *cr15* and *erf142* mutants reveal a role for these genes in the control of panicle architecture. Along with the genes showing significantly differential expression between stages and species, numerous additional AP2/ERF genes were identified in this RNA-seq dataset. More precisely, 82 of the 170 *AP2/ERF* genes identified in the *O. sativa ssp. japonica cv. Nipponbare* genome were detected in the 5-acc dataset. Similarly, in the meristem-specific RNA-seq dataset (LMD dataset) from *O. sativa ssp. japonica* (Harrop *et al.*, 2016), 29 *AP2/ERF* genes were found to be differentially expressed between different types of panicle meristems. It can be noted that 2 genes (*AP2/EREBP86* and *FZP*) that were not detected in the Multi-species dataset were identified in the LMD study. Collectively, 84 *AP2/ERF* genes were reported in the two RNA-seq studies (**Table S1.1**).

A phylogenetic tree based on 84 protein sequences was constructed to offer a broad view of the diversity of *AP2/ERF* genes expressed in panicle meristems (**Fig. 1.1**). The classification of *AP2/ERF* clades was based on previous studies (Nakano *et al.*, 2006; Rashid *et al.*, 2012). A bibliographic survey was carried out to obtain general information on all 84 genes. Of the 84 genes in the list, several have been previously functionally characterized. Those characterized *AP2/ERF* genes for which mutants were available from previous publications are indicated in the tree (**Fig. 1.1**). Some of them were already demonstrated to be involved in flower or inflorescence development. For instance, the *ERF130* (*MULTI-FLORET SPIKELET1*, *MFS1*) gene controls the transition of spikelet meristems to terminal spikelets and positively regulates the expression of the *SNB* and *OsIDS1* genes (Ren *et al.*, 2013). *SUPERNUMERARY BRACT* (*SNB*) and *Oryza sativa INDETERMINATESPIKELET 1* (*OsIDS1*) have been well characterized, studies suggesting that they play a crucial role in regulating the spatio-temporal expression of B- and E-function floral organ identity genes in lodicules and may also be involved in determining inflorescence branching complexity (Lee and An, 2012a; Zhu *et al.*, 2009). The *FRIZZY PANICLE* (*FZP*) gene is highly expressed during axillary branch and spikelet formation rather than at the rachis stage, in agreement with its proposed function in preventing axillary branching (Komatsu *et al.*, 2003; Bai *et al.*, 2016). It should be noted that *SNB*, *OsIDS1*, *RICE STARCH REGULATOR 1* (*RSR1*), *SHATTERING ABORTION1* (*SHAT1*) and *AP2/EREBP59* are all regulated by the microRNA *miR172* (Lee

and An, 2012b; Zhu *et al.*, 2009). Other *AP2/ERF* genes identified in both RNA-seq datasets were *ERF3* and *ERF71* involved in rice root development (Zhao *et al.*, 2015; Lee *et al.*, 2016), *EREBP1* and *ERF48* participating in stress tolerance (Jisha *et al.*, 2015; Jung *et al.*, 2017) and *ERF142* (*SMALL ORGAN SIZE1*, *SMOS1*) controlling organ size and associated with auxin and brassinosteroid signalling (Hirano *et al.*, 2017). Collectively, these data illustrate that *AP2/ERF* genes are of great significance for studies of the molecular mechanisms regulating plant development, notably with regard to panicle architecture development in the context of meristem identity and domestication. The diverse expression patterns that they display during early development raise questions about how and when they act, and whether they work redundantly or individually.



The phylogenetic tree illustrated that *AP2/ERF* genes expressed during panicle development did not belong to any specific subfamily. 84 genes included 2 Soloists, 34 ERFs, 26 DREBs, 3 RAVs and 19 AP2-subfamily genes. These genes are characterized by divergent expression between species and/or meristem types. The different expression patterns of genes recognized in the LMD dataset (Harrop *et al.*, 2016) can be categorized into four groups (**Fig. 1.2**). Group I contains AP2/ERF genes highly expressed in meristem types corresponding to later stages (ePBM & SM). Group II genes are those expressed from early stages (RM, PBM and ePBM) but not in SM. Group III genes are those expressed in the later stages but not, or at lower level, in RM. Group IV consists of genes that are highly expressed at the rachis stage. Most of the genes differentially expressed between meristem types can be found in the ERF, DREB and AP2 subfamilies. It was particularly striking that 19 of the 27 AP2 subfamily genes were detected in panicle meristems and 9 of them were differentially expressed between meristem types including all four categories of expression patterns indicated above. Additionally, two AP2-subfamily genes (*OsIDS1* and *OsPLT9*) that belong to the same cluster (4) showed expression correlation with spikelet number (SpN) and secondary branch number (SBN) in the multi-species RNA-seq dataset (5-acc dataset) (Harrop *et al.*, 2019). In the light of all these observations, the AP2-subfamily genes are clearly of great interest for studies of the regulation of panicle architecture in the context of meristem identity and domestication.



**Figure 1.2.** Heatmap showing expression changes of twenty-nine AP2/ERF genes in different meristem types in *O. sativa ssp. japonica* panicles. Left panel: RM, rachis meristem; PBM, primary branch meristem; ePBM, elongated PBM; AM, axillary meristem; SM, spikelet meristem. I, group 1 with genes highly expressed in later stages (ePBM/AM and SM); II, group 2 with genes strongly expressed from early stages (RM/ PBM/ ePBM/ AM) not in SM; III, group 3 with genes expressed in the later stage (PBM/ ePBM/ AM/ SM) but not in RM; IV, group 4 with genes strongly expressed at the rachis stage (RM).

The AP2 subfamily includes members characterized by a tandem repeat of two AP2 domains (AP2-R1 & AP2-R2) and a small number of proteins with a single AP2 domain, showing higher similarity to those contained in double-AP2 proteins than to the AP2 domain of the ERF proteins (Licausi *et al.*, 2013). In *Arabidopsis*, the AP2 subfamily consists of 18 genes divided into 3 groups based on the motifs conserved in and outside the AP2 domain previously described by Kim *et al.* (2005) and detailed more in my study: namely, the euAP2, basalANT and euANT/PLT groups. The basalANT and euAP2 lineage genes encode proteins with either one or two AP2 domains while all of euANT/PLT lineage encode proteins with two AP2 domains. In *Oryza sativa ssp. japonica*, the AP2 subfamily is composed of 27 genes (Table 1.1).

**Table 1.1.** Summary of AP2/ERF transcription factors of *Arabidopsis thaliana* and *Oryza sativa ssp. Japonica*

<i>Arabidopsis thaliana</i>					
This study			Nakano <i>et al.</i> , 2006		
Classification	Group	Number	Classification	Group	Number
AP2 sub-family		<b>18</b>	AP2 sub-family		<b>18</b>
	euAP2	6		Double AP2/ERF domain	14
	basalANT	4		Single AP2/ERF domain	4
	euANT/PLT	8	ERF sub-family	Groups I to IV	57
DREB sub-family		57		Groups V to X	58
ERF sub-family		65		Groups VI-L and Xb-L	7
RAV sub-family		6	RAV sub-family		6
Soloist	At4g13040	1	Soloist	At4g13040	1
	<b>Total</b>	<b>147</b>		<b>Total</b>	<b>147</b>

<i>Oryza sativa ssp. japonica</i>					
This study			Rashid <i>et al.</i> , 2012		
Classification	Group	Number	Classification	Group	Number
AP2 sub-family		<b>27</b>	AP2 sub-family		<b>27</b>
	euAP2	6		Group Ic	6
	basalANT	9		Group Ia	10
	euANT/PLT	12		Group Ib	11
DREB sub-family		55	DREB sub-family	Group IV	56
ERF sub-family		80	ERF sub-family	Group III	80
RAV sub-family		4	RAV sub-family	Group II	5
Soloist		4	Soloist		2
	<b>Total</b>	<b>170</b>		<b>Total</b>	<b>170</b>

Sharoni <i>et al.</i> , 2011			Nakano <i>et al.</i> , 2006		
Classification	Group	Number	Classification	Group	Number
AP2 sub-family		<b>24</b>	AP2 sub-family	no information	
DREB sub-family	Groups A1 to A6	57	ERF sub-family	Groups I to IV	56
ERF sub-family	Groups B1 to B7	77		Groups V to XIV	83
RAV sub-family		5	RAV sub-family	no information	
	<b>Total</b>	<b>163</b>		<b>Total</b>	<b>139</b>

The euAP2 genes have been well characterized in both *Arabidopsis* and rice. The euAP2 or APETALA2-like group contains 6 members in *A. thaliana* and 6 members in rice, including a miR172-binding sequence within the mRNA transcript sequence. The six genes in *A. thaliana* have been extensively studied in the context of their role in floral ontogeny, including floral meristem identity and flowering time (Huang *et al.*, 2017; Gras *et al.*, 2018; Wang *et al.*, 2016; Zhang *et al.*, 2015). Moreover, in rice, 2 other euAP2 genes, *SNB* and *OsIDS1*, have been well characterized in the context of flower identity as mentioned before. Two other AP2-like rice genes, *SHAT1* and *RSR1*, control respectively seed shattering and starch synthesis in the rice grain, (Zhou *et al.*, 2012; Fu and Xue, 2010; Jiang *et al.*, 2019).

In the same way, the basalANT group consisting of four *WRINKLED (WRI)* genes (*WRI1*, *WRI2* or *ADAP*, *WRI3*, and *WRI4*) has been described in *Arabidopsis* (To *et al.*, 2012). *WRI1* is involved in the regulation of seed storage metabolism in this species (Cernac and Benning, 2004). In maize, a duplication event has created two *WRI1* genes referred to as *ZmWRI1a* and *ZmWRI1b*. While the *ZmWRI1a* gene clearly regulates oil accumulation in seeds, and the *ZmWRI1b* gene rescues *Arabidopsis wri1* mutants, differential expression patterns suggest a unique role for each of the two genes (Pouvreau *et al.*, 2011). There are no reports about the function of *WRI2*; however *SMOS1*, the rice ortholog of *WRI2*, controls organ size and is associated with hormone signalling (Hirano *et al.*, 2017). *WRI3* or *ADAP* is a positive regulator of the ABA response and is also involved in regulating seedling growth (Lee *et al.*, 2009). *WRI3* and *WRI4* are expressed more highly in flowers, stems and roots, and are thought to play a role in the tissue-specific synthesis of fatty acids (To *et al.*, 2012). No studies implicating *WRI* genes in the regulation of inflorescence or flower structure have been reported to date. Nevertheless, euANT/PLT genes are known to be involved in the determination of stem cell fate, the boosting of organ growth and the suppression of differentiation in the framework of shoot meristem function in *Arabidopsis* (Scheres and Krizek, 2018; Horstman *et al.*, 2014).

The AINTEGUMENTA-LIKE/PLETHORA (euANT/PLT) group is a small group in the AP2 subfamily. Six of the 10 *PLT* genes identified in *O. sativa ssp. japonica* (Li and Xue, 2011) were detected in the RNA-seq datasets: *OsPLT1*, *OsPLT4*, *OsPLT7*, *OsPLT8*, *OsPLT9* and *OsPLT10*. However, only *OsPLT7*, *OsPLT8* and *OsPLT9* were found in both RNA-seq datasets. The phylogenetic tree identified two other genes closely related to the *PLT* genes in rice: *AP2/EREBP22* and *AP2/EREBP86*. In the LMD RNA-seq dataset, the expression patterns of *OsPLT7* and *OsPLT8* were classified into group IV (**Fig. 1.2**), these genes being highly expressed in the rachis meristem. Moreover, they show expression patterns which are reversed

with respect to *OsPLT9* (group III) and *AP2/EREBP22* (group I) which have lower expression in the rachis meristem.

The *AINTEGUMENTA* (*ANT*) gene from *A. thaliana* was considered, within the group of *AIL/PLT* genes, to be the main regulator of shoot development (Scheres and Krizek, 2018). In rice, the *OsPLT7*, *OsPLT8* and *OsPLT9* genes were classified into the same group as *ANT* due to high similarities shared between the respective sequences from rice and *A. thaliana* (Li and Xue, 2011). One of the rice genes, *OsPLT8/CROWN ROOTLESS5* (*CRL5*), was functionally characterised in the context of root development (Kitomi *et al.*, 2011) revealing a key role for this gene in crown root initiation. However, information on the possible roles of *PLT* genes in rice panicle architecture development has been lacking to date. In the light of results obtained from panicle RNA-seq studies, it is of particular interest to evaluate the possible importance of *PLT* genes in the regulation of inflorescence meristem identities and activities. Consequently, for the present study, the *PLT* gene group was chosen for further functional characterization so as to address questions concerning the molecular mechanisms that regulate panicle architecture and development.

In the following chapters, I focus on *PLETHORA* genes and present the results of analyses performed on this group. An *in-silico* analysis of *PLETHORA* genes detected in the two RNA-seq datasets was carried out. The next step was to profile the corresponding gene expression patterns and to investigate the function of candidate *PLETHORA*-related genes potentially governing the early stages of rice panicle development, with the aim of evaluating the importance of these genes in the evolution of panicle structure.

## REFERENCES

- Bai, X., Huang, Y., Mao, D., Wen, M., Zhang, L. and Xing, Y. (2016) Regulatory role of FZP in the determination of panicle branching and spikelet formation in rice. *Scientific reports*, **6**, 19022. Available at: <http://www.nature.com/srep/2016/160108/srep19022/full/srep19022.html>.
- Cernac, A. and Benning, C. (2004) WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. *Plant Journal*, **40**, 575–585.
- Fu, F.-F. and Xue, H.-W. (2010) Coexpression analysis identifies Rice Starch Regulator1, a rice AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant physiology*, **154**, 927–938.
- Gras, D.E., Vidal, E.A., Undurraga, S.F., et al. (2018) SMZ/SNZ and gibberellin signaling are required for nitrate-elicited delay of flowering time in Arabidopsis thaliana. *Journal of Experimental Botany*, **69**, 619–631.
- Harrop, T.W.R., Mantegazza, O., Luong, A.M., Béthune, K., Lorieux, M., Jouannic, S. and Adam, H. (2019) A set of AP2-like genes is associated with inflorescence branching and architecture in domesticated rice. *Journal of Experimental Botany*.
- Harrop, T.W.R., Ud Din, I., Gregis, V., Osnato, M., Jouannic, S., Adam, H. and Kater, M.M. (2016) Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *Plant Journal*, **86**, 75–88.
- Hirano, K., Yoshida, H., Aya, K., et al. (2017) SMALL ORGAN SIZE 1 and SMALL ORGAN SIZE 2/DWARF AND LOW-TILLERING Form a Complex to Integrate Auxin and Brassinosteroid Signaling in Rice. *Molecular Plant*, **10**, 590–604. Available at: <http://dx.doi.org/10.1016/j.molp.2016.12.013>.
- Horstman, A., Willemsen, V., Boutilier, K. and Heidstra, R. (2014) AINTEGUMENTA-LIKE proteins: Hubs in a plethora of networks. *Trends in Plant Science*, **19**, 146–157. Available at: <http://dx.doi.org/10.1016/j.tplants.2013.10.010>.
- Huang, Z., Shi, T., Zheng, B., Yumul, R.E., Liu, X., You, C., Gao, Z., Xiao, L. and Chen, X. (2017) APETALA2 antagonizes the transcriptional activity of AGAMOUS in regulating floral stem cells in Arabidopsis thaliana. *New Phytologist*, **215**, 1197–1209.
- Jiang, L., Ma, X., Zhao, S., et al. (2019) The APETALA2-Like Transcription Factor SUPERNUMERARY BRACT Controls Rice Seed Shattering and Seed Size. *The Plant Cell*, **31**, 17–36.
- Jisha, V., Dampanaboina, L., Vadassery, J., Mithöfer, A., Kappara, S. and Ramanan, R. (2015) Overexpression of an AP2/ERF type transcription factor OsEREBP1 confers biotic and abiotic stress tolerance in rice. *PLoS ONE*, **10**, 1–24.
- Jung, H., Chung, P.J., Park, S.-H., Redillas, M.C.F.R., Kim, Y.S., Suh, J.-W. and Kim, J.-K. (2017) Overexpression of *OsERF48* causes regulation of *OsCML16*, a calmodulin-like protein gene that enhances root growth and drought tolerance. *Plant Biotechnology Journal*, 1–14. Available at: <http://doi.wiley.com/10.1111/pbi.12716>.
- Kim, S., Soltis, P.S., Wall, K. and Soltis, D.E. (2005) Phylogeny and Domain Evolution in the APETALA2 -like Gene Family.
- Komatsu, M., Chujo, A., Nagato, Y., Shimamoto, K. and Koyzuka, J. (2003) FRIZZY PANICLE is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development*, **130**, 3841–3850. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12835399>.
- Lee, D.-K., Jung, H., Jang, G., Jeong, J.S., Kim, Y.S., Ha, S.-H., Choi, Y. Do and Kim, J.-K. (2016) Overexpression of the OsERF71 transcription factor Alters Rice Root Structure and Drought Resistance. *Plant Physiology*, pp.00379.2016. Available at: <http://www.plantphysiol.org/lookup/doi/10.1104/pp.16.00379>.
- Lee, D.Y. and An, G. (2012a) Two AP2 family genes, SUPERNUMERARY BRACT (SNB) and OsINDETERMINATE SPIKELET 1 (OsIDS1), synergistically control inflorescence architecture and floral meristem establishment in rice. *Plant Journal*, **69**, 445–461.
- Lee, D.Y. and An, G. (2012b) Two AP2 family genes, SUPERNUMERARY BRACT (SNB) and OsINDETERMINATE SPIKELET 1 (OsIDS1), synergistically control inflorescence architecture and floral meristem establishment in rice. *Plant Journal*, **69**, 445–461.
- Lee, S.J., Cho, D.I., Kang, J.Y. and Kim, S.Y. (2009) An ARIA-interacting AP2 domain protein is a novel component of ABA signaling. *Molecules and Cells*, **27**, 409–416.
- Li, P. and Xue, H. (2011) Structural characterization and expression pattern analysis of the rice PLT gene family. *Acta Biochimica et Biophysica Hungarica*, **43**, 688–697. Available at: <http://abbs.oxfordjournals.org/cgi/doi/10.1093/abbs/gmr068>.
- Licausi, F., Ohme-Takagi, M. and Perata, P. (2013) APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: Mediators of stress responses and developmental programs. *New Phytologist*, 639–649.
- Nakano, T., Suzuki, K., Fujimura, T. and Shinshi, H. (2006) Genome-Wide Analysis of the ERF Gene Family. *Plant Physiology*, **140**, 411–432.

- Pouvreau, B., Baud, S., Vernoud, V., et al.** (2011) Duplicate Maize Wrinkled1 Transcription Factors Activate Target Genes Involved in Seed Oil Biosynthesis. *Plant Physiology*, **156**, 674–686.
- Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J. and Xu, Y.** (2012) AP2/ERF transcription factor in rice: Genome-wide analysis and phylogenetic relationships between monocots and dicots. *Evolutionary Bioinformatics*, **2012**, 321–355.
- Ren, D., Li, Y., Zhao, F., et al.** (2013) MULTI-FLORET SPIKELET1, which encodes an AP2/ERF protein, determines spikelet meristem fate and sterile lemma identity in rice. *Plant physiology*, **162**, 872–84. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23629832>.
- Scheres, B. and Krizek, B.A.** (2018) Coordination of growth in root and shoot apices by AIL/PLT transcription factors. *Current Opinion in Plant Biology*, **41**, 95–101. Available at: <http://dx.doi.org/10.1016/j.pbi.2017.10.002>.
- Sharoni, A.M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., Choi, I.R., Omura, T. and Kikuchi, S.** (2011) Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant and Cell Physiology*, **52**, 344–360.
- To, A., Joubes, J., Barthole, G., Lecureuil, A., Scagnelli, A., Jasinski, S., Lepiniec, L. and Baud, S.** (2012) WRINKLED Transcription Factors Orchestrate Tissue-Specific Regulation of Fatty Acid Biosynthesis in Arabidopsis. *The Plant Cell*, **24**, 5007–5023.
- Wang, P., Cheng, T., Lu, M., Liu, G., Li, M., Shi, J., Lu, Y., Laux, T. and Chen, J.** (2016) Expansion and Functional Divergence of AP2 Group Genes in Spermatophytes Determined by Molecular Evolution and Arabidopsis Mutant Analysis. *Frontiers in Plant Science*, **7**, 1–15. Available at: <http://journal.frontiersin.org/Article/10.3389/fpls.2016.01383/abstract>.
- Zhang, B., Wang, L., Zeng, L., Zhang, C. and Ma, H.** (2015) Arabidopsis TOE proteins convey a photoperiodic signal to antagonize CONSTANS and regulate flowering time. *Genes and Development*, **29**, 975–987.
- Zhao, Y., Cheng, S., Song, Y., Huang, Y., Zhou, S., Liu, X. and Zhou, D.-X.** (2015) The Interaction between Rice ERF3 and WOX11 Promotes Crown Root Development by Regulating Gene Expression Involved in Cytokinin Signaling. *The Plant cell*, **27**, 2469–83. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4815106&tool=pmcentrez&rendertype=abstract>.
- Zhou, Y., Lu, D., Li, C., et al.** (2012) Genetic control of seed shattering in rice by the APETALA2 transcription factor shattering abortion1. *The Plant cell*, **24**, 1034–48. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3336138&tool=pmcentrez&rendertype=abstract>.
- Zhu, Q.-H., Upadhyaya, N.M., Gubler, F. and Helliwell, C. a** (2009) Over-expression of miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice (*Oryza sativa*). *BMC plant biology*, **9**, 149.

## SUPPLEMENTAL FIGURES AND TABLES

**Table S1.1. *AP2/ERF* genes reported in the two available RNA-seq datasets: 5\_acc (Harrop *et al.*, 2019) and LMD (Harrop *et al.*, 2016)**

MSUID	RAPDB	Symbols	5acc_dataset	LMD_dataset	
			detected	Detected	
LOC_Os01g04020	Os01g0131600	ERF108	TRUE	TRUE	group 1
LOC_Os01g04750	Os01g0140700	AP2/EREBP77	TRUE	FALSE	
LOC_Os01g04800	Os01g0141000	AP2/EREBP129	TRUE	FALSE	
LOC_Os01g07120	Os01g0165000	DREB2A	TRUE	TRUE	group 2
LOC_Os01g12440	Os01g0224100	ERF53	TRUE	TRUE	group 1
LOC_Os01g21120	Os01g0313300	ERF68	TRUE	FALSE	
LOC_Os01g46870	Os01g0657400	ERF54	TRUE	FALSE	
LOC_Os01g49830	Os01g0693400	AP2/EREBP127	TRUE	FALSE	
LOC_Os01g58420	Os01g0797600	ERF3	TRUE	TRUE	group 1
LOC_Os01g59780	Os01g0813300	AP2/EREBP79	TRUE	FALSE	
LOC_Os02g06330			TRUE	FALSE	
LOC_Os02g09650	Os02g0189500		TRUE	FALSE	
LOC_Os02g13710	Os02g0231000	ERF38	TRUE	FALSE	
LOC_Os02g29550	Os02g0499000	ERF140	TRUE	FALSE	
LOC_Os02g42585	Os02g0638650	ERF141	TRUE	FALSE	
LOC_Os02g43790	Os02g0654700	ERF91	TRUE	FALSE	
LOC_Os02g43940	Os02g0656600	ERF32	TRUE	FALSE	
LOC_Os02g43970	Os02g0657000	ERF35	TRUE	FALSE	
LOC_Os02g45420	Os02g0676800	ERF20	TRUE	FALSE	
LOC_Os02g45450	Os02g0677300	DREB1G	TRUE	FALSE	
LOC_Os02g51300	Os02g0747600	AP2/EREBP22	TRUE	TRUE	group 1
LOC_Os02g51670	Os02g0752800	ERF49	TRUE	TRUE	group 3
LOC_Os02g54160	Os02g0782700	EREBP1	TRUE	TRUE	group 4
LOC_Os02g55380	Os02g0797100	ERF127	TRUE	FALSE	
LOC_Os03g06920	Os03g0165200	AP2/EREBP85	TRUE	FALSE	
LOC_Os03g07940	Os03g0176300	PLT10	TRUE	FALSE	
LOC_Os03g08500	Os03g0183300	ERF64	TRUE	FALSE	
LOC_Os03g09170	Os03g0191900	ERF47	TRUE	FALSE	
LOC_Os03g12950	Os03g0232200	PLT9	TRUE	TRUE	group 3
LOC_Os03g15660	Os03g0263000	ERF9	TRUE	FALSE	
LOC_Os03g19900	Os03g0313100	AP2/EREBP86	FALSE	TRUE	group 3
LOC_Os03g56050	Os03g0770700	PLT7	TRUE	TRUE	group 4
LOC_Os03g60120	Os03g0815800	ERF58	TRUE	TRUE	group 1
LOC_Os03g60430	Os03g0818800	IDS1	TRUE	FALSE	

LOC_Os04g42570	Os04g0504500	PLT4	TRUE	FALSE	
LOC_Os04g44670	Os04g0529100	ERF45	TRUE	FALSE	
LOC_Os04g46220	Os04g0546800	ERF93	TRUE	FALSE	
LOC_Os04g46400	Os04g0549700	ERF33	TRUE	FALSE	
LOC_Os04g46440	Os04g0550200	ERF34	TRUE	FALSE	
LOC_Os04g48350	Os04g0572400	DREB1E	TRUE	FALSE	
LOC_Os04g52090	Os04g0610400	ERF77	TRUE	FALSE	
LOC_Os04g55520	Os04g0648900	ERF8	TRUE	FALSE	
LOC_Os04g55560	Os04g0649100	SHAT1	TRUE	TRUE	group 3
LOC_Os04g55970	Os04g0653600	PLT1	TRUE	FALSE	
LOC_Os04g57340	Os04g0669200	ERF76	TRUE	FALSE	
LOC_Os05g03040	Os05g0121600	RSR1	TRUE	FALSE	
LOC_Os05g25260	Os05g0316800	ERF56	TRUE	TRUE	group 3
LOC_Os05g27930	Os05g0346200	ERF42	TRUE	TRUE	group 4
LOC_Os05g29810	Os05g0361700	ERF61	TRUE	TRUE	group 1
LOC_Os05g32270	Os05g0389000	ERF142	TRUE	TRUE	group 3
LOC_Os05g41760	Os05g0497200	ERF130	TRUE	TRUE	group 1
LOC_Os05g41780	Os05g0497300	ERF74	TRUE	TRUE	group 1
LOC_Os06g06970	Os06g0165600	DREB1D	TRUE	FALSE	
LOC_Os06g07030	Os06g0166400	ERF7	TRUE	FALSE	
LOC_Os06g08340	Os06g0181700	ERF2	TRUE	FALSE	
LOC_Os06g09390	Os06g0194000	ERF71	TRUE	TRUE	group 4
LOC_Os06g11860	Os06g0222400	ERF120	TRUE	TRUE	group 2
LOC_Os06g36000	Os06g0553700	ERF122	TRUE	FALSE	
LOC_Os06g40150	Os06g0604000	ERF1	TRUE	FALSE	
LOC_Os06g43220	Os06g0639200	AP2/EREBP59	TRUE	FALSE	
LOC_Os06g47590	Os06g0691100	ERF121	TRUE	TRUE	group 1
LOC_Os07g03250	Os07g0124700	PLT8	TRUE	TRUE	group 4
LOC_Os07g12510	Os07g0227600	ERF57	TRUE	FALSE	
LOC_Os07g13170	Os07g0235800	SNB	TRUE	TRUE	group 2
LOC_Os07g42510	Os07g0617000	ERF65	TRUE	FALSE	
LOC_Os07g47330	Os07g0669500	FZP	FALSE	TRUE	group 1
LOC_Os08g07440	Os08g0171100	AP2/EREBP68	TRUE	TRUE	group 1
LOC_Os08g31580	Os08g0408500	ERF48	TRUE	TRUE	group 4
LOC_Os08g34360	Os08g0442400	AP2/EREBP109	TRUE	FALSE	
LOC_Os08g36920	Os08g0474000	ERF104	TRUE	FALSE	
LOC_Os08g42550	Os08g0537900	ERF106	TRUE	TRUE	group 2
LOC_Os09g11460	Os09g0286600	SUB1C	TRUE	FALSE	
LOC_Os09g11480	Os09g0287000	SUB1B	TRUE	FALSE	
LOC_Os09g13940	Os09g0309700	ERF109	TRUE	FALSE	
LOC_Os09g20350	Os09g0369000	ERF50	TRUE	TRUE	group 4

LOC_Os09g25600	Os09g0423800	AP2/EREBP153	TRUE	FALSE	
LOC_Os09g26420	Os09g0434500	ERF72	TRUE	TRUE	group 1
LOC_Os09g35010	Os09g0522000	DREB1B	TRUE	FALSE	
LOC_Os09g35020	Os09g0522100	ERF133	TRUE	FALSE	
LOC_Os09g35030	Os09g0522200	DREB1A	TRUE	FALSE	
LOC_Os10g41130	Os10g0560700	ERF36	TRUE	FALSE	
LOC_Os12g03290	Os12g0126300	AP2/EREBP155	TRUE	FALSE	
LOC_Os12g41030			TRUE	FALSE	
LOC_Os12g41060	Os12g0603300	ERF112	TRUE	FALSE	
<b>Total</b>	84		82	29	

# CHAPTER II

## ***In silico* analysis and expression profiling during panicle development of *euANT/PLT* genes in rice**

Luong Ai My, Adam Hélène, Béthune Kevin, Jouannic Stefan

University of Montpellier, DIADE, IRD, France

## INTRODUCTION

Agricultural progress is crucial to assure food production and security for a growing population, particularly in developing countries (Borlaug, 2007). Rice is a staple food for over half of the world's population and around one billion people depend economically on rice cultivation. A sustainable increase in rice production against a backdrop of climate change, diminishing water and land availability requires the plant to have an improved grain output. Rice yield is directly influenced by the architecture of the inflorescence (Ikeda *et al.*, 2004; Tanaka *et al.*, 2013). Furthermore, panicle development is genetically controlled and considerable diversity in panicle architecture exists at both the inter- and intra-specific levels. The genus *Oryza* has two independently domesticated species: cultivated Asian rice (*Oryza sativa*) and cultivated African rice (*Oryza glaberrima*) (Khush, 1997). Domestication of *O. sativa* from its wild progenitor, *Oryza rufipogon*, is presumed to have been initiated about 10000 years ago (Huang *et al.*, 2012; Choi *et al.*, 2017; Stein *et al.*, 2018). In contrast, *O. glaberrima* was domesticated from *Oryza barthii* in West Africa more recently, about 3500 years ago (Vaughan *et al.*, 2008; Cubry *et al.*, 2018). Despite the independent domestication histories of *O. sativa* and *O. glaberrima*, most varieties of both species share a similar panicle phenotype with a higher seed number and more complex branching compared to the wild ancestors (Yamaki *et al.*, 2011; Ta *et al.*, 2017; Linares, 2002). Nevertheless, little is known about the molecular mechanisms underlying the phenotypic convergence observed between the two domesticated species.

Previous studies used genome-wide expression profiling to compare either the different reproductive meristem types in the *O. sativa* panicle or between 4 distinct species at the branching stage of development. The results obtained revealed a large number of *AP2/ERF* (*APETALA2/Ethylene Responsive Factor*) genes displaying differential expression patterns between the different reproductive meristems and/or between different species (Harrop *et al.* 2016, Harrop *et al.* 2019), suggesting roles in rice panicle architecture development and the determination of its diversity. AP2/ERF transcription factors are a family of proteins defined by a conserved domain containing about sixty to seventy amino acids. They can be subdivided into four sub-families (AP2, ERF, DREB and RAV) depending on the number of AP2/ERF domains in the protein and the existence of other DNA binding domains (Rashid *et al.*, 2012; Nakano *et al.*, 2006; Sharoni *et al.*, 2011). The *RAV* subfamily genes encode proteins possessing a single AP2/ERF domain and one B3 domain. Proteins of the ERF-like and DREB-like subfamilies possess a single AP2/ERF domain. They are

subdivided into two subfamilies (ERF and CBF/DREB) based on DNA binding specificity. The AP2 subfamily includes members characterized by a tandem repetition of two AP2 domains (AP2-R1 and AP2-R2) plus a small number of proteins containing a single AP2 domain (Licausi *et al.*, 2013). There is increasing evidence that certain AP2 sub-family genes are involved in panicle meristem identity determination by regulating the expression of spikelet meristem identity genes. For example, in rice, both *osids1* and *snb* mutants showed a significant decrease in branch and spikelet number within a panicle (Lee and An, 2012), whereas *SMALL ORGAN SIZE1 (ERF142)* controls organ size and modulates root meristem size in rice without any apparent role in floral organ development (Aya *et al.*, 2014; Hirano *et al.*, 2017). AP2 sub-family genes may be differentiated into 3 groups: euAP2, euANT/PLT and basalANT. The euAP2 genes, which incorporate a microRNA172-binding sequence in within their protein coding sequence, have been properly studied in the context of their role in floral ontogeny, such as in floral meristem identity and flowering time (Huang *et al.*, 2017; Gras *et al.*, 2018; Wang *et al.*, 2016; Zhang *et al.*, 2015). The ANT lineage comprises 2 groups: the basalANT (or WRINKLED-like) group and the euANT (or PLETHORA-like) group. The euANT/PLT group is distinguished from the basalANT group by the presence of a long pre-domain region and by 4 characteristic motifs: a 10-amino acid insertion located in the first AP2 domain and 3 other motifs in the pre-domain region.

The *AINTEGUMENTA/PLETHORA (euANT/PLT)* proteins are well-known for their involvement in determining stem cell fate, in the promotion of organ growth and in the suppression of differentiation in the context of shoot meristem function in *Arabidopsis thaliana* (Scheres and Krizek, 2018; Horstman *et al.*, 2014). The *AINTEGUMENTA (ANT)* gene of *A. thaliana* was considered to be the main regulator of shoot development within the group of *PLT* genes found in this species (Scheres and Krizek, 2018). However, little is known about the role of *PLT* genes in rice panicle development. We report here on the characterization and detailed expression analysis of rice *PLT* genes, providing insights for the elucidation of their biological functions.

## MATERIALS AND METHODS

### Plant materials and growth conditions

For Fluidigm qPCR experiments, five accessions were used: *O. sativa ssp. japonica* cv. Nipponbare, *O. sativa ssp. indica* cv. IR64, *O. rufipogon* acc. W1654, *O. glaberrima* acc. Tog5681 and *O. barthii* acc. B88. For harvesting samples, panicle meristems were collected from 15 plants of each accession. Plants were grown in a greenhouse at IRD Montpellier (France), under long day conditions (14h light/10h dark) at 28°C-30°C, and humidity at 60%. After 6 to 8 weeks, flowering induction was carried out under a cycle of 10h of daylight. Panicles were collected at 4 different stages: stage 1, rachis and primary branch meristem; stage 2, elongated primary branch and secondary meristems; stage 3, spikelet differentiation; stage 4, young flowers with differentiated organs.

### *In silico* identification of *euANT/PLT* genes in *Arabidopsis* and various rice genomes

For exhaustively analyzing *euANT/PLT* gene structure, AP2 subfamily genomic and protein sequences were retrieved from the TAIR database for *Arabidopsis thaliana* (Berardini *et al.*, 2015) and from the databases MSU (Kawahara *et al.*, 2013) and RAPDB (Sakai *et al.*, 2013) for *Oryza sativa ssp. japonica*. The *OsPLT* sequences of *O. sativa ssp. japonica* were then used for BLAST searches on Gramene (<http://www.gramene.org>) to identify orthologs between the genomes, for African wild (*O. barthii*, accession IRGC 105608) and domesticated (*O. glaberrima* AGI1.1, accession IRGC 96717 & *O. glaberrima* CG14 – UMR DIADE, data in preparation), Asian wild (*O. rufipogon*, OR\_W1943) and domesticated (*O. sativa indica*, ASM465v1), using the online tool BioMart version 0.7. To obtain information about orthologs of *euANT/PLT* genes in other rice species, 3kb of upstream genomic sequence was obtained for each *OsPLT* ortholog, as was the corresponding protein sequence, using the Gramene resource (<http://www.gramene.org>). Protein alignments and phylogeny trees were built using MEGA 7.0 to determine relationships between the *euANT/PLT* genes in these rice species. Combined un-rooted neighbor-joining (NJ) trees were generated using MEGA 7.0 with the following default parameters: p-distance correction, pairwise deletion and bootstrap (1000 replicates). The visual phylogeny tree was built using EvolView version 3.0 (<https://www.evolgenius.info>) (Subramanian *et al.*, 2019). Manual searches were performed in order to identify potential hormone-responsive elements within promoters. The following motifs were investigated: auxin response element (AuxRE) (TGTCTC, GAGACA, CTCTGT,

ACTTTA) (Hagen and Guilfoyle, 2002; Baumann *et al.*, 1999), jasmonate-responsive element (JARE) (AACGTG) (Boter *et al.*, 2004), abscisic acid responsive element (ABRE) (MACGYGB, ACGTG, ACGT, AACGTT, WAACCA and YAACKG) (Simpson *et al.*, 2003; Ross *et al.*, 2004; Abe *et al.*, 2003), GA-responsive element (GARE) (TAACAAR, TGAC(C/T)) (Ogawa *et al.*, 2003; Zhang *et al.*, 2004), cytokinin response element (CKRE) (AGATT; GATCTT) (Ross *et al.*, 2004), and ethylene response element (GCC box) (Ross *et al.*, 2004) using PlantPAN3.0 (<http://plantpan.itps.ncku.edu.tw>) (Chow *et al.*, 2019).

### ***OsPMT* gene sequence validation**

In order to validate the annotation of *OsPMT* genes in *O. sativa* according to the data from the three rice genome databases: MSU (LOC\_OsXXX) and RAPDB (OsXXX) for *Oryza sativa ssp. japonica* cv. Nipponbare, and Phytozome (OsKitaakeXXX) for *Oryza sativa ssp. japonica* cv. Kitaake v3.1), primers were design using Primer3Plus (<https://primer3plus.com>) (Untergasser *et al.*, 2007) and NCBI Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>) web facilities. The specificity of primers was then checked again by BLAST on Gramene (<http://www.gramene.org>) and Rice Genome Annotation Project (Kawahara *et al.*, 2013) (<http://rice.plantbiology.msu.edu/index.shtml>). Primer sequences are listed in **Table S2.2**. Sanger sequencing of the PCR products were carried out (Genewiz, England) for validation of the *O. sativa* annotation genes.

### **RNA extraction and Fluidigm® high throughput qPCR**

Total RNA was extracted from panicles at 4 different developmental stages using the RNeasy Plant Mini kit with RLT and RWT buffers (QIAGEN, Germany). DNase treatments were performed using the RNase-free DNase kit (QIAGEN, Germany). cDNA was synthesized from 250 ng of DNase-treated total RNA using the SuperScript III First-Strand Synthesis System (Invitrogen, USA). A Biomark HD Microfluidic Dynamic Array (Fluidigm, USA) was used for large-scale qPCR. Before performing qPCR, the sample mixture and assay mixture were prepared individually. A 96 × 96 Dynamic Array Integrated Fluidic Circuit (Fluidigm, USA) was loaded with cDNA and primer combinations after 15 cycles of specific target amplification and exonuclease I treatment according to the manufacturer's instructions. A fast cycling protocol with EvaGreen dye (Bio-Rad Laboratories, USA) was used for amplification. Three biological replicates were performed for each sample. Data were

normalized using 3 genes (*ACT2-LOC\_Os11g06390*, *HK04-LOC\_Os01g16970*, *HK09-LOC\_Os03g61680*). Gene expression relative to the normalization standards was estimated using the Normalized relative quantification method with multiple housekeeping genes (Pfaffl, 2001; Pfaffl, 2002). Primer sequences are listed in **Table S2.3**. The fluidgr R Package (<https://github.com/othomantegazza/fluidgr>) was used to load, normalize, scale and visualize Fluidigm qPCR data.

## RESULTS

### The euANT/PLT group in *O. sativa*

In *A. thaliana*, the AP2 subfamily consists of 18 genes divided into 3 groups, namely the euAP2, basalANT and euANT/PLT groups. In *O. sativa ssp. japonica*, the AP2 subfamily contains 27 genes. A phylogenetic tree was generated from the alignment of protein sequences of AP2 subfamily members retrieved from these two species (**Fig. 2.1**). Ten *PLT* genes were previously reported in *O. sativa ssp. japonica* (*OsPLT1* to *OsPLT10*; Li and Xue, 2011) compared to eight members of the *PLT* group in *A. thaliana* (Mähönen *et al.*, 2014; Scheres and Krizek, 2018). According to the tree in Fig. 2.1, two other proteins might belong to the same cluster as the PLT proteins in rice, namely AP2/EREBP22 and AP2/EREBP86, thus increasing the total number of *OsPLT* genes to 12 members. The two aforementioned proteins show a close relationship with AtAIL6 and AtAIL7 while OsPLT7, OsPLT8 and OsPLT9 were found to cluster with AtANT proteins. Our previously described RNA-seq datasets reported numerous AP2 subfamily genes that were expressed in rice panicle meristems (Harrop *et al.*, 2019; Harrop *et al.*, 2016), including 5 *euAP2*, 6 *basalANT* and 8 *PLT* genes.



**Gene and protein structure of *euANT/PLT* group members in *O. sativa***

Gene and protein structure of *euANT/PLT* group members varied considerably between the two rice databases RAPDB and MSU (**Table 2.1**). For all members of the group, the genomic and protein sequences are in conflict between two databases, from small variations to major differences. In the cases of *OsPLT4*, 5 and 6, protein and CDS lengths were similar in both databases. Both databases indicated that the genomic length of *OsPLT8* was 4702 bp but the corresponding protein and CDS lengths were different. In the case of the *OsPLT3* and *OsPLT7*, the RAPDB protein sequences lacked a second AP2 domain (AP2-R2) whereas the MSU sequences contained two complete domains. In a similar way the MSU protein sequences of *OsPLT2*, *OsPLT8*, *OsPLT9* and *AP2/ERE2BP22* were found to be missing one of the two AP2 domains but the corresponding RAPDB sequences contained two AP2 repeats.

Further analyses revealed that the number of exons that *OsPLT* genes contained was totally disparate between the RAP-DB and MSU databases (**Fig. 2.2**). Compared to RAPDB, the MSU protein coding sequences lacked a 9-bp mini-exon, this applied for all *OsPLT* genes except *OsPLT3* and *OsPLT7* (**Fig. 2.2**). Additionally, the number of transcripts varied between two databases. *OsPLT1* has two transcripts (LOC-Os04g55970.1 & .2) attributed to it in the MSU database but only one in the RAP-DB database. *OsPLT2* has two transcripts (Os06g0657500-01 & -02) according to the RAPDB database and only a single one in MSU.

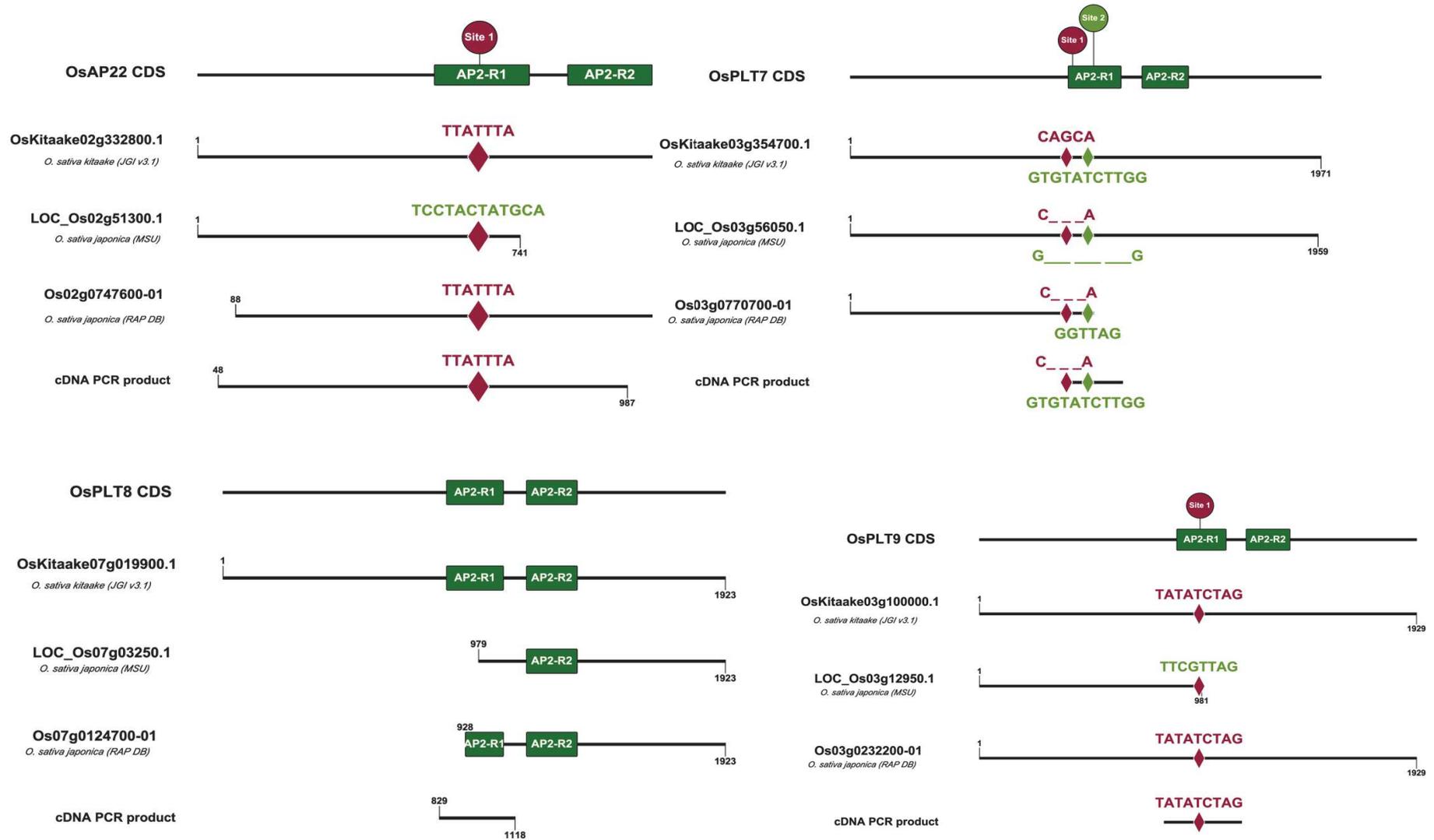
**Table 2.1. Features of *euANT/PLT* genes and proteins in RAPDB and MSU databases.**

<b>Gene Symbol</b>	<b>RAPDB ID</b>	<b>Gene length (bp)</b>	<b>Protein length (aa)</b>	<b>CDS length (bp)</b>	<b>No. Exon</b>	<b>AP2 R1 (aa)</b>	<b>AP2 R2 (aa)</b>	<b>MSU-ID</b>	<b>Gene length (bp)</b>	<b>Protein length (aa)</b>	<b>CDS length (bp)</b>	<b>No. Exon</b>	<b>AP2 R1 (aa)</b>	<b>AP2 R2 (aa)</b>
<i>OsPLT1</i>	Os04g0653600	4551	495	1488	9	73	65	LOC_Os04g55970	4585	492	1479	8	70	65
<i>OsPLT2</i>	Os06g0657500	3855	469	1410	9	73	65	LOC_Os06g44750	3903	263	792	8	65	
<i>OsPLT3</i>	Os02g0614300	2094	321	966	4	65		LOC_Os02g40070	4627	700	2103	8	70	65
<i>OsPLT4</i>	Os04g0504500	4264	658	1977	9	73	65	LOC_Os04g42570	4475	658	1977	8	73	65
<i>OsPLT5</i>	Os01g0899800	4343	695	2088	9	73	65	LOC_Os01g67410	4350	695	2088	8	73	65
<i>OsPLT6</i>	Os11g0295900	5214	559	1680	9	73	65	LOC_Os11g19060	5459	559	1680	6	73	65
<i>OsPLT7</i>	Os03g0770700	3683	339	1020	8	31		LOC_Os03g56050	3581	652	1959	8	70	65
<i>OsPLT8</i>	Os07g0124700	4702	331	996	7	49	65	LOC_Os07g03250	4702	314	945	6		65
<i>OsPLT9</i>	Os03g0232200	4303	642	1929	9	73	65	LOC_Os03g12950	4306	326	981	8	34	
<i>OsPLT10</i>	Os03g0176300	2330	169	510	8	73	65	LOC_Os03g07940	4689	413	1242	8	92	65
<i>AP2/EREBP22</i>	Os02g0747600	3624	338	1017	8	73	65	LOC_Os02g51300	4273	246	741	8	40	
<i>AP2/EREBP86</i>	Os03g0313100	5168	431	1296	8	73	65	LOC_Os03g19900	5170	490	1473	9	132	65



In order to validate the annotation of *euANT/PLT* candidate genes for further analysis (see Chapter 3), the sequences of *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* candidate genes and CDS retrieved from the three databases were compared to panicle-derived cDNA PCR products from Nipponbare and Kitaake cultivars (**Fig. 2.3**). For *OsAP2/EREBP22*, sequences retrieved from the three databases, (i.e. *OsKitaake02g332800.1*, *LOC\_Os02g51300.1* and *Os02g0747600-01* locii), were different in CDS length, with 1104 bp, 741 bp and 1017 bp, respectively. The 9-bp mini-exon including the conflict site (-TTATTTA-) was missing in MSU sequence while it was available in two other sequences. The PCR fragments from nucleotides 48 to 987 of the CDS in both Nipponbare and Kitaake cultivars cDNAs confirmed that the *OsKitaake02g332800.1* with the length of 1104 bp and the presence of 9-bp exon was the right annotation of *OsAP2/EREBP22* in rice panicle. For *OsPLT7*, sequences retrieved from three databases, (i.e. *OsKitaake03g354700.1*, *LOC\_Os03g56050.1* and *Os03g0770700-01* locii), were different in length of CDS, with 1971 bp, 1959 bp and 1020 bp, respectively. The 9-bp mini-exon including the conflict site (-GTGTATCTTGG-) and three nucleotides (-AGC-) were present in Phytozome *Kitaake* sequence while it was missing in two other sequences. The PCR fragments including the two conflict sites in both Nipponbare and Kitaake cultivar cDNAs confirmed the presence of the 9-bp exon. However, the PCR product sequences indicated that the panicle-derived cDNA sequences of *OsKitaake03g354700.1* did not include -AGC- nucleotides. For *OsPLT8*, sequences retrieved from three databases, (i.e. *OsKitaake07g019900.1*, *LOC\_Os07g03250.1* and *Os07g0124700-01* locii), were different in length of CDS, with 1923 bp, 945 bp and 996 bp, respectively. The first AP2 domain was totally missing in MSU sequence and partly missing in RAPDB sequences while it was available in Kitaake sequences. The primers designed to amplify the first AP2 domain sequence in both Nipponbare and Kitaake cultivar cDNAs confirmed that the *OsKitaake07g019900.1* sequence with the length of 1923 bp was the right annotation of *OsPLT8* in rice panicle. Finally, For *OsPLT9* sequences retrieved from three databases, *OsKitaake03g100000.1* and *LOC\_Os03g12950.1*, were different in length of CDS, with 1929 bp and 981 bp, respectively, while *OsKitaake03g100000.1* and *Os03g0232200-01* were exactly the same. The 9-bp mini-exon including the doubtful site (-TATATCTAG-) was missing in MSU sequence while it was available in two other sequences. The PCR fragment sequence including the conflict site in both Nipponbare and Kitaake cultivar cDNAs confirmed that the *OsKitaake03g100000.1* and *Os03g0232200-01* with the

length of 1929 bp and the presence of 9-bp exon was the right annotation of *OsPLT9* in rice panicle.

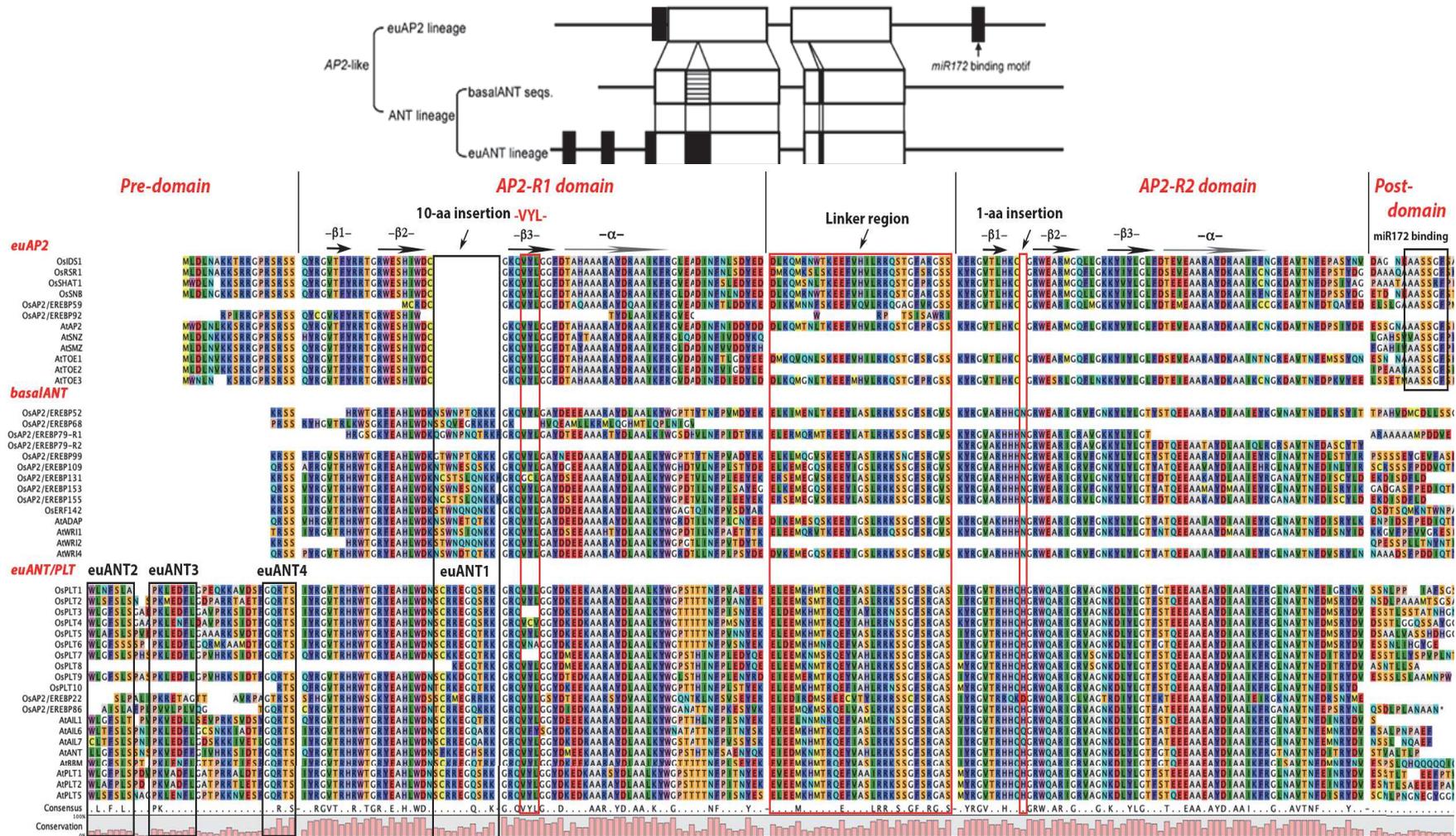


**Figure 2.3. Structure validation of *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22*.** *O. sativa ssp. japonica* CDS sequences were retrieved from different databases: *Oryza sativa ssp. japonica* cv. *Kitaake* v3.1 DOE-JGI (*OsKitaakeXXX*), *O. sativa ssp. japonica* cv. *Nipponbare* MSU (*LOC\_OsXXX*) and RAPDB databases (*OsXXX*). The green boxes indicate the AP2-domain coding sequences. The length of CDS was indicated in the tail of each sequence. The site marks (red or green) indicate the conflicts between CDS sequences of different databases.

**O. *sativa* and *A. thaliana* euANT/PLT genes relationship**

The full protein sequence encoded by each *PLT* gene, as well those of some other members of the AP2-subfamily, were exhaustively re-analysed. Protein sequences including a complete AP2 domain were used for alignment (**Table S1**). From the alignment of the two AP2 domain sequences and connecting linker region, it can be seen that the AP2 domain sequences were highly conserved between *A. thaliana* and rice within the different groups (**Fig. 2.4**). Clearly, all PLT genes encoded two AP2 domains while euAP2 and basalANT can have a single AP2. Each AP2 domain confers a typical three-dimensional conformation organized into a layer of three antiparallel beta-sheets followed by a parallel alpha-helix (Allen *et al.*, 1998). Despite some differences in amino acid identities between the three groups, three beta-sheets and one alpha helix sequence were maintained within the AP2 subfamily except for some distinctions in the cases of AP2/EREBP59 and AP2/EREBP92 (euAP2 group), AP2/EREBP52, AP2/EREBP68, AP2/EREBP79 and AtWRI2 (basalANT group) and OsPLT8 (PLT group). The -VYL- motif encoded by a 9bp mini-exon was highly conserved in  $\beta$ 3 sheet of AP2-R1. The euANT1 motif reported as being the signature of the ANT lineage was observed. Both the basalANT and euANT/PLT group display a 10 amino acid (aa) insertion in AP2-R1 domain whereas euAP2 proteins lack this motif. In the AP2-R2 domain, a single aa insertion is detected in *ANT* genes compared to euAP2 (Kim *et al.*, 2005).

Outside the AP2 domains, the linker between two repeats, which includes a constant number of 29 aa were also seen to be highly conserved. Compared to euAP2 and basalANT, most of the PLT proteins possessed three unique motifs, namely euANT2 (WL-FSLS), euANT3 (PK-EDFL) and euANT4 (GQRTS) in the pre-domain sequence, with some variations noted in the cases of OsAP2/EREBP22 and OsAP2/EREBP86 and abnormality in the sequences of OsPLT8 and OsPLT10. While the post-domain sequences of euAP2 members contained the miR172 binding motif (which in the mRNA encodes the amino acid sequence AASSGF), the post-domain of PLTs were largely divergent.



**Figure 2.4.** Alignment of AP2 subfamily protein sequences from *O. sativa ssp. japonica* and *A. thaliana*. The alignment was generated by MEGA7. In the top schema (adapted from (Kim *et al.*, 2005)), black boxes indicate specific motifs, blank boxes indicate AP2 domain, hatched box indicates the 10-aa insertion that is not conserved in the basalANT sequences. The AP2-R1 of euANT/PLT group includes a conserved 10-amino acids (aa) insertion, called as motif euANT1. Three conserved motifs euANT2, euANT3, euANT4 were specific in the pre-domain region of euANT/PLT group. The AP2-R2 of basal ANT and euANT/PLT group consisted of a 1-aa insertion compared to euAP2 group. The miR172-binding motif of euAP2 sequences located in the post-domain. A consensus covered the residues appeared in more than 90% of sequences was shown below the alignment. From left to right, the pre-domain, first AP2 domain AP2-R1, linker region between two domains, second AP2-R2 and post-domain sequence

### Orthologs of *euANT/PLT* genes in other rice species

In order to carry out a comparison of *euANT/PLT* genes in the different rice species and to prepare for high-throughput qRT-PCR, a multi-species investigation of *PLT* orthologs was performed. All of the *PLT* orthologs studies are listed in **Table 2.2**. Orthologs of all *PLT* genes were present in both African and Asian species of wild and domesticated rice. Furthermore, in *O. barthii*, a duplication of *OsPLT2* was observed in chromosome 6 within a region that included the *OsPLT2* ortholog (OBART06G23380) and a repeated segment thereof in reverse orientation (OBART06G23420). A duplication was also observed for *OsPLT6* in *O. glaberrima*, with the *OsPLT6* ortholog (ORGLA11G0090500) being found on chromosome 11 and a duplicated locus on chromosome 3 (ORGLA03G0393500).

An alignment of 59 amino acid sequences of the PLT group from 5 different rice species was carried out to study interspecific structural variation (**Fig. 2.5**). The alignment did not include the sequences of *OsPLT2*, *OsPLT8* and *AP2/EREBP86* orthologs in *O. glaberrima* due to incomplete information being available from the database, which is in progress. In general, the PLT sequences were highly conserved between species. The African domesticated rice *O. glaberrima* sequences obtained in UMR DIADE (Montpellier, France, data not published) confirmed the presence of *OsPLT2*, *OsPLT8* and *AP2/EREBP86* orthologs in this species, which have not been reported to date in the Gramene database. It was noteworthy that the 3 conserved motifs in the pre-domain of the euANT/PLT lineage were completely conserved between wild and domesticated species. Additionally, the 10 aa euANT1 motif inserted in the AP2-R1 domain and the 29 aa linker sequence between two AP2 domains was unchanged between the different species. Nonetheless, some variations in sequences between species were detected, for instance, the VYL motif on the C-terminal side of the euANT1 region was present in domesticated species (*O. sativa ssp. japonica* cv. Nipponbare, *O. glaberrima* or *O. sativa ssp. indica*) but never appeared in wild species (*O. barthii* and *O. rufipogon*). According to Nakano *et al.*, 2006, motif VYL occurs inside the  $\beta$ -3 sheet, near the connection between the  $\beta$ -3 sheet and  $\alpha$ -helix in the 3D structure of the AP2 protein; thus changing this motif might result in the modification of the spatial structure of the protein. However, further sequencing confirmation should be carried out in order to confirm that this motif is missing in wild rice species.

A phylogenetic tree constructed from the euANT/PLT sequences of the five species displayed a number of nodes grouping together the different orthologs of a given PLT protein, indicating that *euANT/PLT* genes evolved before the domestication of rice and were

maintained in domesticated species during evolution (**Fig. 2.6**). Our analysis revealed the presence of two copies of *OsPLT2* in African wild rice *O. barthii* and two copies of *OsPLT6* in African domesticated *O. glaberrima*. *OsPLT2* was previously observed to be highly expressed in young roots but weakly expressed at the seed stage whereas *OsPLT6* was strongly expressed in seeds but weakly expressed at other stages (Li and Xue, 2011). Differences in *OsPLT2* and *OsPLT6* copy number between wild and domesticated accessions suggest that the *OsPLT* family may have undergone evolutionary changes during rice domestication.

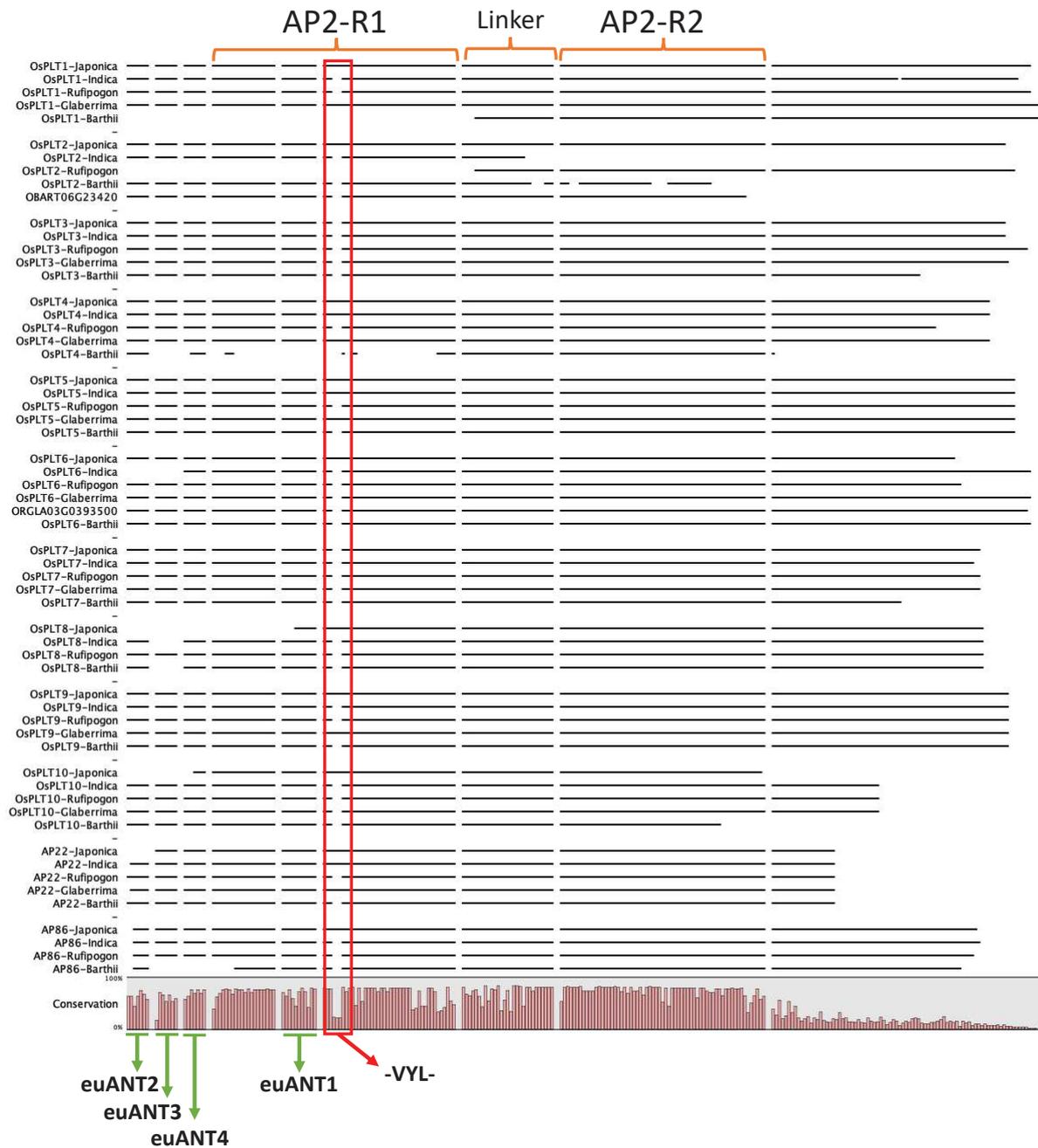
**Table 2.2. List of euANT/PLT orthologs in 5 rice species**

Symbol	Orientation	Location	<i>O. sativa ssp. japonica</i>	<i>O. sativa ssp. indica</i>	<i>O. rufipogon</i>	<i>O. glaberrima</i>	<i>O. barthii</i>
<i>OsPLT1</i>	Forward	Chromosome 04	Os04g0653600	BGOSGA017263	ORUFI04G29440	ORGLA04G0240800	OBART04G27840
<i>OsPLT2</i>	Forward	Chromosome 06	Os06g0657500	BGOSGA023402	ORUFI06G25070	Unavailable*	OBART06G23380**
<i>OsPLT3</i>	Forward	Chromosome 02	LOC_Os02g40070	BGOSGA008619	ORUFI02G25430	ORGLA02G0207800	OBART02G24000
<i>OsPLT4</i>	Forward	Chromosome 04	LOC_Os04g42570	BGOSGA016726	ORUFI04G19390	ORGLA04G0150400	OBART04G17940
<i>OsPLT5</i>	Reverse	Chromosome 01	Os01g0899800	BGOSGA000316	ORUFI01G43970	ORGLA01G0345800	OBART01G40650
<i>OsPLT6</i>	Forward	Chromosome 11	LOC_Os11g19060	BGOSGA035153	ORUFI11G10800	ORGLA11G0090500**	OBART11G10210
<i>OsPLT7</i>	Forward	Chromosome 03	LOC_Os03g56050	BGOSGA013636	ORUFI03G37050	ORGLA03G0326600	OBART03G35620
<i>OsPLT8</i>	Forward	Chromosome 07	LOC_Os07g03250	BGOSGA025077	ORUFI07G01630	Unavailable*	OBART07G01750
<i>OsPLT9</i>	Reverse	Chromosome 03	Os03g0232200	BGOSGA011170	ORUFI03G09610	ORGLA03G0090300	OBART03G09490
<i>OsPLT10</i>	Reverse	Chromosome 03	Os03g0176300	BGOSGA011350	ORUFI03G05420	ORGLA03G0053700	OBART03G05640
<i>AP2/EREBP22</i>	Reverse	Chromosome 02	Os02g0747600	BGOSGA005624	ORUFI02G34150	ORGLA02G0280000	OBART02G32450
<i>AP2/EREBP86</i>	Reverse	Chromosome 03	Os03g0313100	BGOSGA010867	ORUFI03G15620	Unavailable*	OBART03G14970

\* unavailable in GRAMENE but available in *O. glaberrima* CG14 – UMR DIADE, France – data not published yet.

\*\* A paralog is ORGLA03G0393500

\*\*\* A paralog is OBART06G23420



**Figure 2.5. Structure and alignment of euANT/PLT orthologs protein sequences from 5 rice species.** Alignment of the amino sequences generated by MEGA7. The prefix Japonica, Rufipogon and Indica, Barthii and Glaberrima indicated *O. sativa ssp. japonica* cv. Nipponbare, Asian wild (*O. rufipogon*, OR\_W1943) and domesticated (*O. sativa ssp. indica*, ASM465v1), African wild (*O. barthii*, accession IRGC 105608) and domesticated (*O. glaberrima* AGI1.1, accession IRGC 96717), respectively; AP22 is AP2/EREBP22, AP86 is AP2/EREBP86. OBART06G23420 is paralog of OsPLT2 in *O. barthii*, ORGLA03G0393500 is paralog of OsPLT6 in *O. glaberrima*. A graph covered the residue conservation between sequences was shown below the alignment. From left to right, euANT2, 3, 4, conserved motifs of euANT/PLT groups in pre-domain, euANT1 motif in the first AP2-R1 domain, the linker between two domains, second AP2-R2 domain and the post-domain. The red box indicates the -VYL- motif of AP2-R1 domain in ortholog sequences of different rice species.



### ***In silico* expression profiling analysis of the *euANT/PLT* group in *O. sativa japonica***

Using data contained in the RiceXPro database (<http://ricexpro.dna.affrc.go.jp>) (Sato *et al.*, 2013), the expression of the 12 rice *euANT/PLT* genes was profiled in different tissues or organs throughout the various stages of development of the plant in the field (**Fig. S1**). On the basis of this profiling, the expression patterns of the *euANT/PLT* genes could be classified into 3 types: seed-preferential (*OsPLT1*, 2, 6, 10 and *AP2/EREBP86*), root-preferential (*OsPLT3*, 4 and 5), and inflorescence-preferential (*OsPLT7*, 8 and 9), the exception being *OsAP2/EREBP22*, which was not specifically more expressed in any particular organ. Gene expression profiles were further investigated by comparing the aforementioned data to qRT-PCR results obtained using RNAs extracted from primary root, crown root, old crown root, seeding, stem, stem-base, leaf, flower, and seed (Li and Xue, 2011) (**Fig. S2**). In the latter study, *OsPLT7*, *OsPLT8* and *OsPLT9* were revealed to be strongly expressed in the stem and the stembase of young seedlings. The strong expression of these three genes in the young inflorescence, as observed in the RiceXpro database, was coherent with their identification in the two RNA-seq datasets obtained from rice panicle and its meristems (Harrop *et al.*, 2019; Harrop *et al.*, 2016).

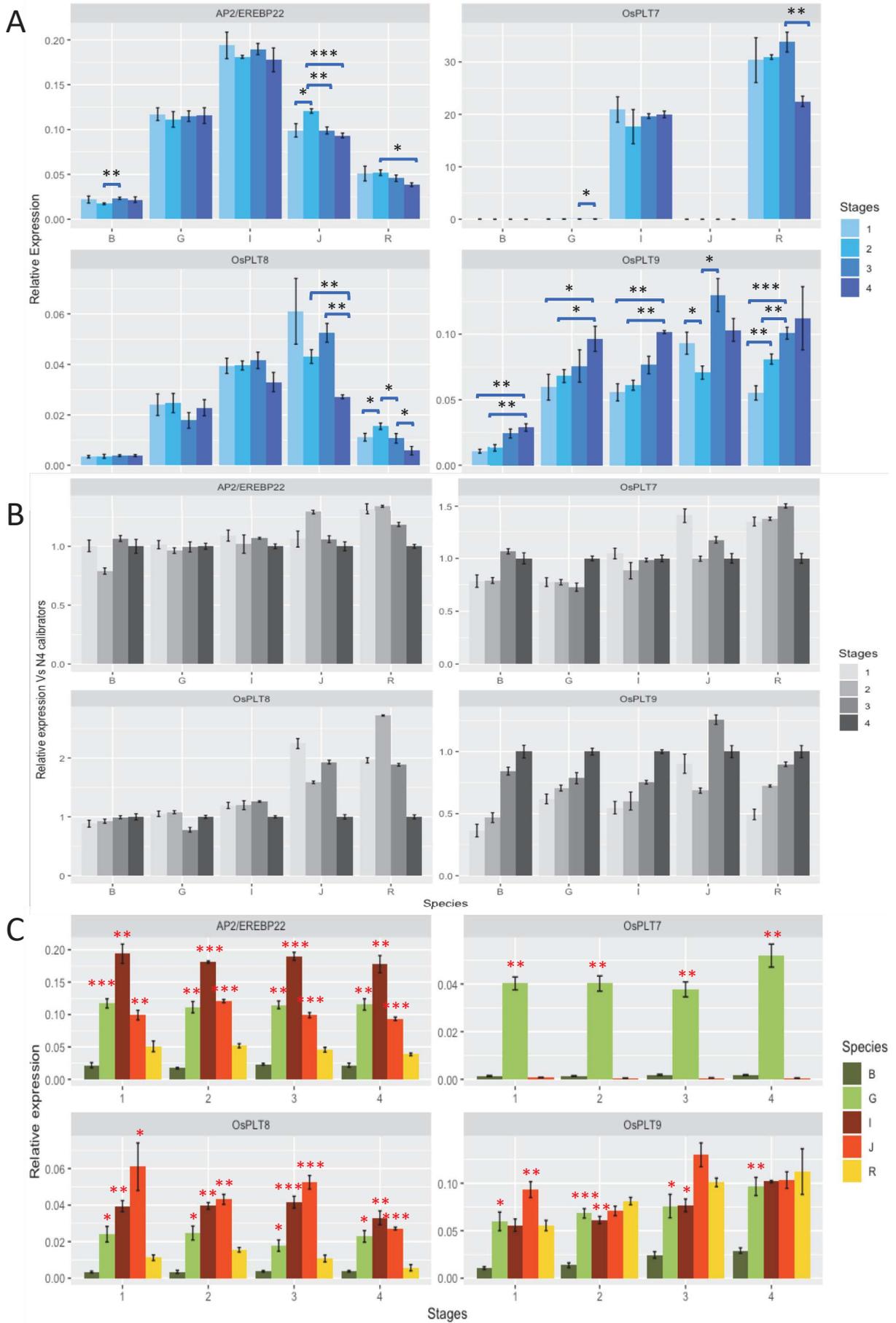
### **Expression patterns in different rice species of *euANT/PLT* genes preferentially expressed in the panicle**

We studied in more detail the expression pattern of the 3 *euANT/PLT* genes preferentially expressed in the panicle (*OsPLT7*, *OsPLT8* and *OsPLT9*) as well as the *OsAP2/EREBP22*. Analyses were performed to investigate panicle development at early stages in different rice species. High throughput qRT-PCR based on the Fluidigm® system was performed using *O. sativa ssp. japonica* cv. *Nipponbare*, *O. sativa ssp. indica* cv. IR64, *O. rufipogon* acc. W1654, *O. glaberrima* acc. Tog5681 and *O. barthii* acc. B88 in order to test the expression of these genes during early panicle development from four stages designated as N1 to N4. N1 corresponded to the rachis meristem stage, N2 to the primary to secondary branching stage, N3 to axillary branching or spikelet formation and N4 to spikelet & floral establishment.

The qPCR data obtained corroborated the expression profiles described earlier for the four genes in the *O. sativa ssp. japonica* LMD RNA-seq dataset (Harrop *et al.*, 2016). *AP2/EREBP22* was observed to be more highly expressed in the branch meristem compared with the other stages and the qPCR results confirmed that it is most strongly expressed at the N2 stage in *O. sativa ssp. japonica*. Both *OsPLT7* and *OsPLT8* were more highly expressed at rachis stages in comparison with other stages. Furthermore, *OsPLT9* displayed

progressively increasing expression levels from N1 to N4, not only in *O. sativa sssp. Japonica* but also in other rice species, with a high level of expression at spikelet formation (**Fig. 7A-B**).

*OsPLT9* was differentially expressed not only between different stages but also between wild and domesticated rice. *OsPLT9* was highly expressed in the African domesticated species compared to the wild one, significantly at all four stages. Similarly, qPCR results illustrated that both *AP2/EREBP22* and *OsPLT8* showed differential expression during panicle development between domesticated and wild type rice of both the Asian and African species. Surprisingly, the transcript abundance of *OsPLT7* was extremely high in *O. sativa ssp. indica* (IR64) (domesticated Asian rice) and *O. rufipogon* (W1654) (wild Asian rice) while much lower in *O. sativa ssp. japonica* (domesticated Asian rice). Additionally, *OsPLT7* was more highly expressed in *O. glaberrima* (domesticated African rice) compared with *O. barthii* (African wild rice) (**Fig. 7C**). The differential expression of *OsPLT* genes of interest between different stages of panicle development and between different rice species suggests the possibility that *OsPLT7-9* and *AP2/EREBP22* might play important roles in rice panicle development and be of significance in the context of domestication.



**Figure 2.7 Relative gene expression of *AP2/EREBP22*, *OsPLT7*, *OsPLT8* and *OsPLT9* through panicle development at early stages in five species.** In the qRT-PCR experiment, there were 3 panicle biological replicates from each accession, at early stages (N1 to N4). N1 was corresponding to rachis meristem stage, N2 was primary-secondary branching stage, N3 was axillary branching or spikelet formation and N4 was at spikelet & floral establishment. The prefix of species B and G, R and I, J indicated African wild (*O. barthii* acc. B88) and domesticated (*O. glaberrima* acc. Tog5681), Asian wild (*O. rufipogon* acc. W1654) and domesticated (*O. sativa ssp. indica* acc. IR64, *O. sativa ssp. japonica* cv. *Nipponbare*) respectively. **(A)** Relative gene expression of each gene between 4 stages in 5 species without calibrator. Black star indicates the test between stages in each species. **(B)** Relative gene expression of each gene between species with N4 stage of each species as calibrator. **(C)** Comparison of gene expression through different stages between wild and domesticated species. Red stars above the domesticated species indicate the test between the wild and domesticated in each stage (R vs. J or I - Asian rice; B vs. G - African rice). Student t-test, \* $P \leq 0.05$ , \*\*  $P \leq 0.01$ , \*\*\*  $P \leq 0.001$ .

### **Cis-element analysis of promoter regions of *OsPLT* genes**

Since the significant differences in expression levels observed between wild and domesticated species for the *OsPLT7-9* and *AP2/EREBP22* genes could potentially be attributable to interspecific variability in *cis*-regulatory sequences in their promoter regions, we investigated this possibility. Promoter regions of *OsPLT1-6* were previously analyzed for known consensus *cis*-regulatory elements of the auxin response element (TGTCTC, GAGACA, CTCTGT), ABA response element (MACGYGB), cytokinin response element (AGATT) and ethylene response element (GCC-box) (Li and Xue, 2011) (**Fig. S3**). Expression studies of *OsPLT1-6* genes in 7-day-old roots (including primary root, crown root, and lateral root) under treatment with auxin, ABA, cytokinin, and ethylene by qRT-PCR revealed that all of these genes were induced by auxin. However, analyses of the promoter region of *OsPLT7-10*, *AP2/EREBP22* and *AP2/EREBP86* were not reported. A search for hormone-responsive motifs in the 3kb genomic region upstream of these genes was carried out using PlantPAN3.0 (Chow *et al.*, 2019). Analysis of the promoter regions of *OsPLT7-10*, *AP2/EREBP22* and *AP2/EREBP86* indicated the presence of motifs implicated in the response to auxin (AuxREs), jasmonic acid (JAREs), ABA (ABREs), Gibberellin (GAREs), cytokinin (CKREs) and ethylene (EtREs) (**Table 2.3**). The most abundant motif observed was the ABA responsive element, especially in the case of *AP2/EREBP22* for which nearly 70 predicted ABREs were observed. Moreover, the *OsPLT7*, *OsPLT8* and *OsPLT9* promoter regions were also found to possess numerous predicted ABREs. In summary, all of the *OsPLT* genes displayed motifs for potential regulation by auxin, ABA and cytokinin while only some of them contained promoter sequences suggesting response to JA (*OsPLT9* and *AP2/EREBP22*) and GA (*OsPLT7*, *OsPLT8* and *AP2/EREBP86*). Most of *OsPLTs* contained a GCC-box (EtRE) in promoter region, except *OsPLT2*, *OsPLT5*, *OsPLT7* and *OsPLT8*. Moreover, RiceXPro database also indicated that expression of *OsPLT7*, 8 and 9 in the shoot was induced in each case by auxin treatment after 3-6 hours (**Fig. S4**). Overall, the promoter regions of the different *OsPLT* genes in *O. sativa* were divergent with different numbers of hormone response elements.

In the next step of this study, the promoters of the panicle-preferentially expressed genes *OsPLT7*, 8, 9 and *AP2/EREBP22* and their orthologs in other rice species were retrieved and aligned. The results indicated that a number of SNPs were distributed between wild and domesticated rice species along with insertions and deletions that were specific to either the Asian or the African rice domestication (**Fig. 2.8**). Among these four genes, the promoter sequences of *O. barthii* (African wild) and *O. glaberrima* (African domesticated)

shared great similarity, and presented much difference compared to Asian species. However, there were some distinct features in the promoter sequences of *OsPLT7*, 8 and 9 that allowed differentiation between African wild and domesticated types. Likewise, Asian domesticated *O. sativa ssp. indica* displayed considerable variation in promoter sequences compared to the Asian wild *O. rufipogon* and Asian domesticated *O. sativa ssp. japonica*, which might help explain the divergent expression patterns observed with *O. sativa ssp. indica*. Changes in promoter sequences between orthologs may have consequences for hormone response, through alterations to the number and location of regulatory *cis*-elements in a given species.

**Table 2.3.** Auxin-, jasmonic acid (JA)-, abscisic acid (ABA)-, Gibberellin (GA)-, cytokinin- (CK)-, and ethylene-regulated *cis*-elements in 3kb-upstream promoters of *OsPLT7-10*, *AP2/EREBP22* and *AP2/EREBP86* in *O. sativa ssp. japonica cv. Nipponbare*.

Gene	AuxRE	JA	ABARE	GARE	CKRE	EtRE
<i>AP2/EREBP22</i>	3	3	68	0	3	20
<i>AP2/EREBP86</i>	15	0	18	4	6	1
<i>OsPLT7</i>	1	0	18	1	15	1
<i>OsPLT8</i>	3	0	12	1	6	0
<i>OsPLT9</i>	9	1	23	0	9	0
<i>OsPLT10</i>	10	0	34	0	6	9

**AuxRE**, auxin response element; **JA**, Jasmonic acid; **ABRE**, ABA response elements; **GARE**, Gibberellin response elements; **CKRE**, cytokinin response element and **EtRE**, ethylene response elements

**(AuxRE)** (TGTCTC, GAGACA, CTCTGT, ACTTTA) (Hagen and Guilfoyle, 2002; Baumann *et al.*, 1999)

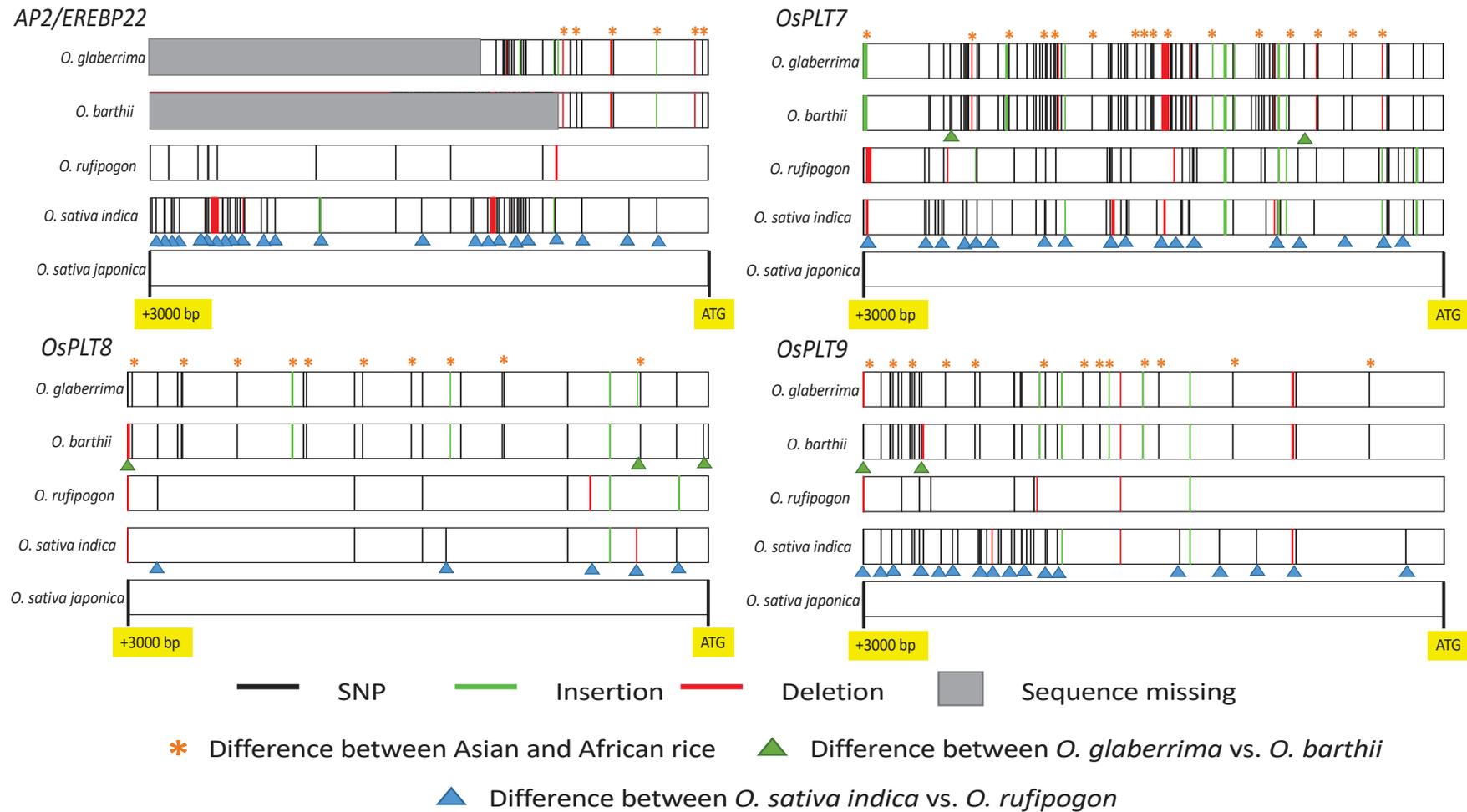
**(JARE)** (AACGTG) (Boter *et al.*, 2004)

**(ABRE)** (MACGYGB, ACGTG, ACGT, AACGTT, WAACCA and YAACKG) (Simpson *et al.*, 2003; Ross *et al.*, 2004; Abe *et al.*, 2003)

**(GARE)** (TAACAAR, TGAC(C/T)) (Ogawa *et al.*, 2003; Zhang *et al.*, 2004)

**(CKRE)** (AGATT; GATCTT) (Ross *et al.*, 2004)

**(EtRE)** (GCCGCC) (Ross *et al.*, 2004)



**Figure 2.8.** Distribution of sequence polymorphisms (SNPs) (black lines), insertions (light green lines) and deletions (red lines) in 3000 bp-upstream promoter sequences of *OsPLT7*, *8*, *9* and *AP2/EREBP22* of other rice species in comparison with reference *O. sativa ssp. japonica*. Green arrows indicates the difference between wild (*O. barthii*) and domesticated (*O. glaberrima*) of African rice. Blue arrows indicates the difference between wild (*O. rufipogon*) and domesticated (*O. sativa ssp. indica*) of Asian rice. Orange mark indicates the difference between Asian and African rice. Bp, Base pairs; Grey blocks indicate the missing sequence.

## DISCUSSION

### ***euANT/PLT* gene sequences and annotations show divergences between the two common rice databases**

The rice genome annotation project maintained by the Michigan State University (Kawahara *et al.*, 2013) and the Rice Annotation Project Database (RAPDB) (Sakai *et al.*, 2013) greatly facilitate a comprehensive analysis of genome structure and gene functions in rice. The MSU and the RAPDB, which use the same reference genome (*O. sativa ssp japonica* cv. Nipponbare) are separate projects, generating different pseudomolecules and different methods/criteria are used to identify gene models between the two groups. While much of the annotation is similar between the two groups, there will inevitably be differences due to the alternative methods used (Kawahara *et al.*, 2013). This may however lead in some instances to uncertainties about genomic structure and functional interpretation. For example, *SMALL ORGAN SIZE1 (SMOS1)/SHOEBOX (SHB)/ERF142* was well characterized as *Os05g0389000* containing 9 exons encoding a 425 amino acid protein with an AP2 domain (Aya *et al.*, 2014) and then as *LOC\_Os05g32270* containing only 8 exons (Li *et al.*, 2015). Another AP2/ERF gene, *CROWN ROOTLESS5 (CRL5/OsPLT8)* was classified as *Os07g0124700* (i.e. in RAPDB) with 8 exons encoding a predicted protein of 642 amino acid residues with two AP2 domains (Kitomi *et al.*, 2011) and then reported as *LOC\_Os07g03250* (i.e. in MSU) with 7 exons encoding a predicted protein of 314 amino acid residues with only one AP2 domain (Li and Xue, 2011). Finally, we confirmed that it was an alternative annotation that was right. The validation of the annotation of the *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* genes revealed that none of the 3 databases are better. These examples within the *APETALA2* family illustrate a general need for comprehensive validation of gene structure to be carried out so as to facilitate the functional characterization of rice genes.

### **OsPLT proteins are highly conserved between *Arabidopsis* and different rice species**

The most regions observed to be the most highly conserved within the AP2 proteins studied were two AP2 domains (designated R1 and R2) and the region between these two repeats (Kim *et al.*, 2005; Zumajo-Cardona and Pabón-Mora, 2016). The AP2 domains are essential for biological activity (Jofuku *et al.*, 1994) and the conserved linker is critical for DNA binding through direct contact with the genomic DNA or positioning of the two AP2 repeats thereon (Krizek, 2003; Nole-Wilson, 2000). Proteins of the euANT/PLT lineage contain two AP2 domains serving this function while euAP2 and basalANT groups may have

a single AP2 domain, raising questions about how the two AP2 domains cooperate in the case of the euANT/PLT group. In analyses of OsPLT orthologs from different rice species, two AP2 repeats were conserved in almost all protein sequences, with the exception of *O. barthii* orthologs of OsPLT1 and OsPLT4 and the *O. rufipogon* ortholog of OsPLT2 which lack the first domain in their protein sequences. Further experiments should be carried out to check whether these sequences are indeed lacking or whether the database sequences are incomplete. In our study, complete sequences of two AP2 repeats and the linker of all AP2 subfamily members were retrieved and extensively analyzed, confirming the high level of sequence conservation between *A. thaliana* and other rice species. Furthermore, the observed conservation of the three beta-sheets and alpha-helix of AP2 domains is an indication that 3D-structure also is maintained between the proteins of the different species.

Interestingly, a motif of three amino acids -VYL- within AP2-R1 is observed in domesticated rice species (*O. glaberrima* or *O. sativa ssp. indica* or *O. sativa ssp. japonica*) but never in wild ones (*O. barthii* and *O. rufipogon*). The latter motif is highly conserved between AP2 proteins in both *A. thaliana* and rice species. It is encoded by an independent small 9-bp exon in *Arabidopsis* (Ma *et al.*, 2013) and castor bean (*Ricinus communis*) (Ji *et al.*, 2018). It has been demonstrated that mutations in -VYL- motif reduced AtANT DNA-binding activity (Krizek, 2003). Likewise, site-directed mutagenesis of amino acids within -VYL- failed to restore the full oil contents of *wri1-1* seeds, indicating that the -VYL- sequence is also crucial for *AtWRI1* function (Ma *et al.*, 2013). In rice, the presence of mini-exons was reported (**Fig. 2.2**). While the -VYL- motif was replaced by -GCL- in OsWRI1 (OsAP2/EREBP131, **Fig. 2.4**), both OsWRI1 and AtWRI1 almost lost their activity when GCL or VYL, respectively, were deleted, suggesting that the sequences are important for their activity (Ap *et al.*, 2019). However, the results were different from those obtained with castor bean WRI1. Both splice variants of castor bean WRIs (RcWRI1-A containing -VYL- and RcWRI1-B lacking -VYL-) restore the full oil content of *wri1-1*, even though RcWRI1-B lacks VYL (Ji *et al.*, 2018). These contradictory results suggest that the importance of the VYL motif was species-dependent. The apparent lack of VYL in wild rice species should be confirmed by experiments before drawing any conclusions about the importance of the VYL motif in OsPLT activity in the context of the domestication process.

In the present study, the previously reported four motifs euANT1-4 conserved in the euANT/PLT lineage (Kim *et al.*, 2005; Dipp-Álvarez and Cruz-Ramírez, 2019) were confirmed in different rice species (**Fig. 2.4**). In rice, 10 members of the OsPLT lineage were first identified by Li and Xue, 2011 but no comments were made about these motifs. Through

the alignment of whole sequences from the AP2-subfamily, it was observed that the euAP2 and basalANT proteins did not contain the three motifs euANT2-4 in their pre-domain and that the 10-aa insertion in basalANT sequences was different compared to the euANT1 motif in PLT group. Based on euANT1 sequence identities, we can now classify two other proteins (OsAP2/EREBP22 and OsAP2/EREBP86) into the euANT/PLT group despite the incompletely conserved motifs in pre-domain sequence. The euANT1 motif is present in all sequences from the different rice species; however further work will be required to investigate the missing sections corresponding to the *O. barthii* orthologs of OsPLT1 and OsPLT4 and the *O. rufipogon* ortholog of OsPLT2. Intriguingly, one of the euANT sequences from *Sorghum bicolor*, *Sobic.007G056700*, has a divergent 10-aa insertion in the AP2-R1 domain that does not include the euANT1 motif, nor does it possess the other euANT pre-domain motifs (Dipp-Álvarez and Cruz-Ramírez, 2019).

The pre-domains of OsPLT proteins were mostly conserved between different rice species. However, in some cases not all of the three domains were conserved. The euANT2-3 motifs were seen to be missing in the OsPLT6 ortholog of *O. sativa ssp. indica* while it was conserved in OsPLT6 of other species, raising the questions about the quality of annotation. It was reported that 69% of euANT sequences possessed the euANT2 motif, 79% the euANT3 motif and 90% the euANT4 motif, so the rule of the pre-domain is not universally applicable. Relatively little is known about the function of the pre-domain, apart from a study of BABYBOOM (BBM) orthologs. In *A. thaliana*, BBM was observed to cooperatively function in the regulation of somatic embryogenesis and embryo development with a functional importance demonstrated for the euANT2 motif (Ouakfaoui *et al.*, 2010). Deletion of euANT2 prevented the somatic embryos from generating shoots even after prolonged periods in culture.

In the present study, all members of the euANT/PLT lineage were identified in each rice species, indicating that the euANT/PLT group was maintained during rice domestication. However for *OsPLT2* in the African species, two possible scenarii might have occurred: *i.* a duplication before domestication in African wild species and the lost of one of the two copies after domestication, *ii.* a duplication in the wild species after the radiation of the domesticated one. In parallel, a duplication subsequent to domestication may have occurred for *OsPLT6* in *O. glaberrima*.

***OsPLTs* gene acting in hormone signaling pathways**

In *A. thaliana*, PLT proteins function in auxin signalling pathways throughout plant development and at different levels, via AUXIN RESPONSE FACTORS (ARFs) and PIN feedback loops (Horstman *et al.*, 2014; Scheres and Krizek, 2018). Local auxin synthesis is essential for meristem function and stem elongation. Auxin plays a key role in regulating the formation, activity, and fate of meristems, thereby shaping plant architecture (Gallavotti, 2013). Prolonged high auxin levels were found to be required to promote PLT activities; moreover positive feedback from PLT to auxin biosynthesis and transport plays a role during the generation of new primordia (Mähönen *et al.*, 2014). For instance, two PLT proteins in *A. thaliana*, PLT1 and PLT2, function downstream of auxin in the specification of the root stem cell niche (Horstman *et al.*, 2014). Similarly, in *O. sativa ssp. japonica*, the role of OsPLT8 (CROWN ROOTLESS5, CRL5) was characterized in the crown root, in which it was proven to act downstream of the ARF-mediated auxin signaling pathways. *OsPLT8* expression was regulated by ARFs through the hormone responsive TGTC motif (GACACTGACA) in its upstream regulatory sequence (Kitomi *et al.*, 2011). In addition, *OsPLT1-6* were found to be up-regulated in roots after 3-12 hours of auxin treatment (Li and Xue, 2011) and *OsPLT7-9* were strongly expressed in shoots after 6 hours of auxin treatment (**Fig. S4**). In this context, the abundance of potential auxin response *cis*-elements located in the promoter regions of *OsPLT* genes suggests that *OsPLT* functions could be mediated by auxin distribution and response via interaction with auxin-responsive upstream sequences.

Multiple hormones including auxin (Vernoux *et al.*, 2010; Gallavotti, 2013) and cytokinin (Bartrina *et al.*, 2011; Han *et al.*, 2014) have been shown to participate in the regulation and transition of flower development. Cytokinin has been shown to be a positive regulator of cell proliferation and to perform a crucial role in shoot apical meristem (SAM) function. In contrast to effects observed with cytokinin, gibberellin was found to induce cell differentiation by promoting longitudinal cell expansion and cytoskeletal rearrangement (Ha *et al.*, 2010). Additionally, ethylene may promote the reproductive transition in rice (Iqbal *et al.*, 2017) whereas ABA inhibits the floral transition in *Arabidopsis* (Wang *et al.*, 2013). *PLT* genes were observed to be stimulated by auxin, cytokinin, gibberellin, jasmonic acid and ethylene (Vernoux *et al.*, 2010; Kim, 2016; Abiri *et al.*, 2017; Qi *et al.*, 2011; Gu *et al.*, 2017). In our study, the presence of multiple hormone responsive elements in *OsPLT* promoters probably indicates that PLT genes participate in many processes, including flower development and meristem specification, via hormone-mediated signalling pathways.

***OsPLT7, OsPLT8, OsPLT9 and AP2/EREBP22 are differentially expressed during the early stages of panicle development in rice***

In *A. thaliana*, *ANT*, *AIL6* and *AIL7* were found to be expressed in a distinct but redundant fashion in the inflorescence and vegetative shoot meristems (Horstman *et al.*, 2014). The *ant/ail6/ail7* triple mutant shoot ceased growing after producing a few leaves due to reduced cell divisions in the meristem and differentiation of the meristematic cells, proving that the *ANT*, *AIL6* and *AIL7* genes were required for SAM maintenance. The *ant* and *ail6* mutants resulted in shoot-meristemless phenotypes, while *ail7* partially overlapped these phenotypes, demonstrating that *ANT*, *AIL6* and *AIL7* do not cooperate in a completely redundant mode (Mudunkothge and Krizek, 2012). Both *ANT* and *AIL6* are known to be involved in many aspects of flower development, some of which are regulated by auxin (Krizek, 2011b; Krizek, 2011a). Nonetheless, the phylogenetic tree presented here (**Fig. 2.1**) illustrates that *OsPLT7*, *OsPLT8* and *OsPLT9* share a close relationship with *ANT*, and that *AP2/EREBP86* and *AP2/EREBP22* belong to the same cluster as *AIL6* and *AIL7*.

It is interesting to note that eight of the twelve *OsPLT* genes were previously identified in our Laser Microdissection (LMD) and/or Multiple-species RNA-seq datasets (Harrop *et al.*, 2019; Harrop *et al.*, 2016), namely: *OsPLT1*, *OsPLT4*, *OsPLT7-10*, *AP2/EREBP22* and *AP2/EREBP86*, indicating that they are expressed in the rice panicle meristem. Moreover, *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* were present in both RNA-seq datasets. In the LMD RNA-seq dataset, *OsPLT7* and *OsPLT8* were found to be highly expressed in the rachis meristem while *OsPLT9* was preferentially expressed in the axillary meristem along with *AP2/EREBP22*, with a lower expression in the rachis meristem. *AP2/EREBP22*, *AP2/EREBP86*, *OsPLT7*, *OsPLT8* and *OsPLT9* were detected in the LMD RNA-seq dataset but not the other *OsPLTs* indicating that the latter genes are expressed within the meristems whereas the others may be expressed in other parts of the panicle and/or at a lower level within the meristems. Our high throughput qPCR studies also confirmed the expression patterns of *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* in *O. sativa ssp. japonica*. Moreover, it was demonstrated that these four genes display differential expression between wild and domesticated rice species. Consequently, *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* are of particular interest for studying the molecular mechanisms that regulate the development and architecture of the rice panicle in the context of meristem identity and domestication.

## **ACKNOWLEDGEMENTS**

Dr François Sabot from UMR DIADE, IRD Montpellier kindly provided genomic sequences of *OsPLT2*, *OsPLT8* and *AP2/EREBP86* of *O. glaberrima* and bioinformatic support. We thank Céline Cardé and H  l  ne Vignes (Grand plateau technique r  gional de g  notypage, CIRAD, Montpellier) for support with high-throughput qPCR and cDNA library preparation. Sophie Ch  ron and Harold Chrestin assisted with plant care. We also thank James Tregear for feedback on the manuscript and discussion on the project.

## **FUNDING**

This research was funded by Agropolis Fondation through the ‘‘Investissements d’Avenir’’ programme (ANR-10-LABX-0001-01), Fondazione Cariplo (EVOREPRICE 1201-004), and the CGIAR Research Program on Rice. We also received support from the Excellence Scholarship Program of the Embassy of France in Vietnam (LAM) and the Allocations de recherche pour une th  se au Sud (ARTS) program from Institut de Recherche pour le D  veloppement (IRD).

## REFERENCES

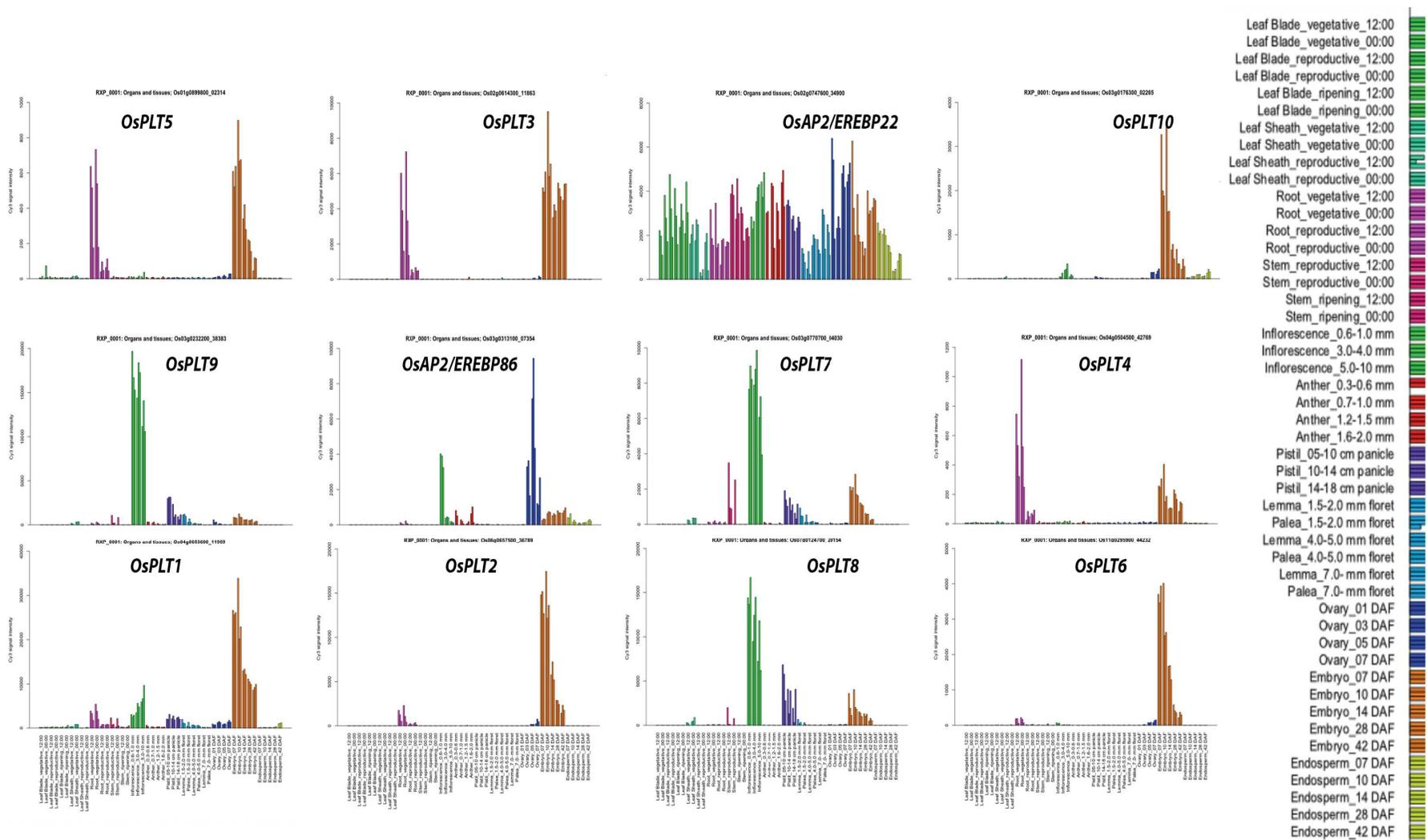
- Abe, H., Urao, T., Ito, T., Seki, M. and Shinozaki, K.** (2003) Transcriptional Activators in Abscisic Acid Signaling. *Society*, **15**, 63–78.
- Abiri, R., Shaharuddin, N.A., Maziah, M., et al.** (2017) Role of ethylene and the APETALA 2/ethylene response factor superfamily in rice under various abiotic and biotic stress conditions. *Environmental and Experimental Botany*, **134**, 33–44.
- Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M. and Suzuki, M.** (1998) A novel mode of DNA recognition by a  $\beta$ -sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO Journal*, **17**, 5484–5496.
- Ap, A., Transcription, D., Mano, F., Aoyanagi, T. and Kozaki, A.** (2019) Atypical Splicing Accompanied by Skipping. , 1–12.
- Aya, K., Hobo, T., Sato-Izawa, K., Ueguchi-Tanaka, M., Kitano, H. and Matsuoka, M.** (2014) A novel AP2-type transcription factor, SMALL ORGAN SIZE1, controls organ size downstream of an auxin signaling pathway. *Plant and Cell Physiology*, **55**, 897–912.
- Bartrina, I., Otto, E., Strnad, M., Werner, T. and Schmülling, T.** (2011) Cytokinin Regulates the Activity of Reproductive Meristems, Flower Organ Size, Ovule Formation, and Thus Seed Yield in Arabidopsis thaliana. *The Plant Cell*, **23**, 69–80.
- Baumann, K., Paolis, A. De, Costantino, P. and Gualberti, G.** (1999) The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the rolB oncogene in plants. *Plant Cell*, **11**, 323–333.
- Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. and Huala, E.** (2015) The arabidopsis information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis*, **53**, 474–485.
- Borlaug, N.** (2007) Feeding a Hungry World. *Science*, **318**, 359–359. Available at: <http://www.sciencemag.org/cgi/doi/10.1126/science.1151062>.
- Boter, M., Ruíz-Rivero, O., Abdeen, A. and Prat, S.** (2004) Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and Arabidopsis. *Genes and Development*, **18**, 1577–1591.
- Choi, J.Y., Platts, A.E., Fuller, D.Q., Hsing, Y.I., Wing, R.A., Purugganan, M.D. and Kim, Y.** (2017) The rice paradox: Multiple origins but single domestication in Asian Rice. *Molecular Biology and Evolution*, **34**, 969–979.
- Chow, C.N., Lee, T.Y., Hung, Y.C., Li, G.Z., Tseng, K.C., Liu, Y.H., Kuo, P.L., Zheng, H.Q. and Chang, W.C.** (2019) Plantpan3.0: A new and updated resource for reconstructing transcriptional regulatory networks from chip-seq experiments in plants. *Nucleic Acids Research*, **47**, D1155–D1163.
- Cubry, P., Tranchant-Dubreuil, C., Thuillet, A.C., et al.** (2018) The Rise and Fall of African Rice Cultivation Revealed by Analysis of 246 New Genomes. *Current Biology*, **28**, 2274–2282.e6.
- Dipp-Álvarez, M. and Cruz-Ramírez, A.** (2019) A Phylogenetic Study of the ANT Family Points to a preANT Gene as the Ancestor of Basal and euANT Transcription Factors in Land Plants. *Frontiers in Plant Science*, **10**, 1–12.
- Gallavotti, A.** (2013) The role of auxin in shaping shoot architecture. *Journal of Experimental Botany*, **64**, 2593–2608.
- Gras, D.E., Vidal, E.A., Undurraga, S.F., et al.** (2018) SMZ/SNZ and gibberellin signaling are required for nitrate-elicited delay of flowering time in Arabidopsis thaliana. *Journal of Experimental Botany*, **69**, 619–631.
- Gu, C., Guo, Z.H., Hao, P.P., Wang, G.M., Jin, Z.M. and Zhang, S.L.** (2017) Multiple regulatory roles of AP2/ERF transcription factor in angiosperm. *Botanical Studies*, **58**, 1–8.
- Ha, C.M., Jun, J.H. and Fletcher, J.C.** (2010) Shoot apical meristem form and function. *Current Topics in Developmental Biology*, **91**, 103–140.
- Hagen, G. and Guilfoyle, T.** (2002) Auxin-responsive gene expression: Genes, promoters and regulatory factors. *Plant Molecular Biology*, **49**, 373–385.
- Han, Y., Yang, H. and Jiao, Y.** (2014) Regulation of inflorescence architecture by cytokinins. *Frontiers in Plant Science*, **5**, 1–4.
- Harrop, T.W.R., Mantegazza, O., Luong, A.M., Béthune, K., Lorieux, M., Jouannic, S. and Adam, H.** (2019) A set of AP2-like genes is associated with inflorescence branching and architecture in domesticated rice. *Journal of Experimental Botany*.
- Harrop, T.W.R., Ud Din, I., Gregis, V., Osnato, M., Jouannic, S., Adam, H. and Kater, M.M.** (2016) Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *Plant Journal*, **86**, 75–88.
- Hirano, K., Yoshida, H., Aya, K., et al.** (2017) SMALL ORGAN SIZE 1 and SMALL ORGAN SIZE 2/DWARF AND LOW-TILLERING Form a Complex to Integrate Auxin and Brassinosteroid Signaling in

- Rice. *Molecular Plant*, **10**, 590–604. Available at: <http://dx.doi.org/10.1016/j.molp.2016.12.013>.
- Horstman, A., Willemsen, V., Boutilier, K. and Heidstra, R.** (2014) AINTEGUMENTA-LIKE proteins: Hubs in a plethora of networks. *Trends in Plant Science*, **19**, 146–157. Available at: <http://dx.doi.org/10.1016/j.tplants.2013.10.010>.
- Huang, X., Kurata, N., Wei, X., et al.** (2012) A map of rice genome variation reveals the origin of cultivated rice. *Nature*, **490**, 497–501. Available at: <http://dx.doi.org/10.1038/nature11532>.
- Huang, Z., Shi, T., Zheng, B., Yumul, R.E., Liu, X., You, C., Gao, Z., Xiao, L. and Chen, X.** (2017) APETALA2 antagonizes the transcriptional activity of AGAMOUS in regulating floral stem cells in *Arabidopsis thaliana*. *New Phytologist*, **215**, 1197–1209.
- Ikeda, K., Sunohara, H. and Nagato, Y.** (2004) Developmental Course of Inflorescence and Spikelet in Rice. *Breeding Science*, **54**, 147–156.
- Iqbal, N., Khan, N.A., Ferrante, A., Trivellini, A., Francini, A. and Khan, M.I.R.** (2017) Ethylene Role in Plant Growth, Development and Senescence: Interaction with Other Phytohormones. *Frontiers in Plant Science*, **08**, 1–19.
- Ji, X.J., Mao, X., Hao, Q.T., Liu, B.L., Xue, J.A. and Li, R.Z.** (2018) Splice variants of the castor WRI1 gene upregulate fatty acid and oil biosynthesis when expressed in tobacco leaves. *International Journal of Molecular Sciences*, **19**.
- Jofuku, K.D., Boer, B.G.W. Den, Montagu, M. Van and Okamoto, J.K.** (1994) Control of arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell*, **6**, 1211–1225.
- Kawahara, Y., la Bastide, M. de, Hamilton, J.P., et al.** (2013) Improvement of the oryza sativa nipponbare reference genome using next generation sequence and optical map data. *Rice*, **6**, 3–10.
- Khush, G.S.** (1997) Origin, dispersal, cultivation and variation of rice. *Oryza: From Molecule to Plant*, 25–34. Available at: [http://link.springer.com/10.1007/978-94-011-5794-0\\_3](http://link.springer.com/10.1007/978-94-011-5794-0_3).
- Kim, J.** (2016) CYTOKININ RESPONSE FACTORS Gating Environmental Signals and Hormones. *Trends in Plant Science*, **21**, 993–996. Available at: <http://dx.doi.org/10.1016/j.tplants.2016.10.004>.
- Kim, S., Soltis, P.S., Wall, K. and Soltis, D.E.** (2005) Phylogeny and Domain Evolution in the APETALA2 - like Gene Family.
- Kitomi, Y., Ito, H., Hobo, T., Aya, K., Kitano, H. and Inukai, Y.** (2011) The auxin responsive AP2/ERF transcription factor CROWN ROOTLESS5 is involved in crown root initiation in rice through the induction of OsRR1, a type-A response regulator of cytokinin signaling. *Plant Journal*, **67**, 472–484.
- Komatsu, M., Maekawa, M., Shimamoto, K. and Kyozuka, J.** (2001) The LAX1 and FRIZZY PANICLE 2 genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Developmental biology*, **231**, 364–373.
- Krizek, B.A.** (2011a) Aintegumenta and Aintegumenta-Like6 regulate auxin-mediated flower development in Arabidopsis. *BMC Research Notes*, **4**, 176. Available at: <http://www.biomedcentral.com/1756-0500/4/176>.
- Krizek, B.A.** (2003) AINTEGUMENTA utilizes a mode of DNA recognition distinct from that used by proteins containing a single AP2 domain. *Nucleic Acids Research*, **31**, 1859–1868.
- Krizek, B.A.** (2011b) Auxin regulation of Arabidopsis flower development involves members of the AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT) family. *Journal of Experimental Botany*, **62**, 3311–3319.
- Lee, D.Y. and An, G.** (2012) Two AP2 family genes, SUPERNUMERARY BRACT (SNB) and OsINDETERMINATE SPIKELET 1 (OsIDS1), synergistically control inflorescence architecture and floral meristem establishment in rice. *Plant Journal*, **69**, 445–461.
- Li, J., Zhao, Y., Chu, H., et al.** (2015) SHOEBOX Modulates Root Meristem Size in Rice through Dose-Dependent Effects of Gibberellins on Cell Elongation and Proliferation. *PLoS Genetics*, **11**, 1–21.
- Li, P. and Xue, H.** (2011) Structural characterization and expression pattern analysis of the rice PLT gene family. *Acta Biochimica et Biophysica Hungarica*, **43**, 688–697. Available at: <http://abbs.oxfordjournals.org/cgi/doi/10.1093/abbs/gmr068>.
- Licausi, F., Ohme-Takagi, M. and Perata, P.** (2013) APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: Mediators of stress responses and developmental programs. *New Phytologist*, 639–649.
- Linares, O.F.** (2002) Pnas\_99\_25\_16360.Pdf.
- Ma, W., Kong, Q., Arondel, V., Kilaru, A., Bates, P.D., Throver, N.A., Benning, C. and Ohlrogge, J.B.** (2013) WRINKLED1, A Ubiquitous Regulator in Oil Accumulating Tissues from Arabidopsis Embryos to Oil Palm Mesocarp. *PLoS ONE*, **8**, 1–13.
- Mähönen, A.P., Tusscher, K. Ten, Siligato, R., et al.** (2014a) PLETHORA gradient formation mechanism separates auxin responses. *Nature*, **515**, 125–129.
- Mähönen, A.P., Tusscher, K. Ten, Siligato, R., et al.** (2014b) PLETHORA gradient formation mechanism separates auxin responses. *Nature*, **515**, 125–129.
- Mudunkothge, J.S. and Krizek, B.A.** (2012) Three Arabidopsis AIL/PLT genes act in combination to regulate

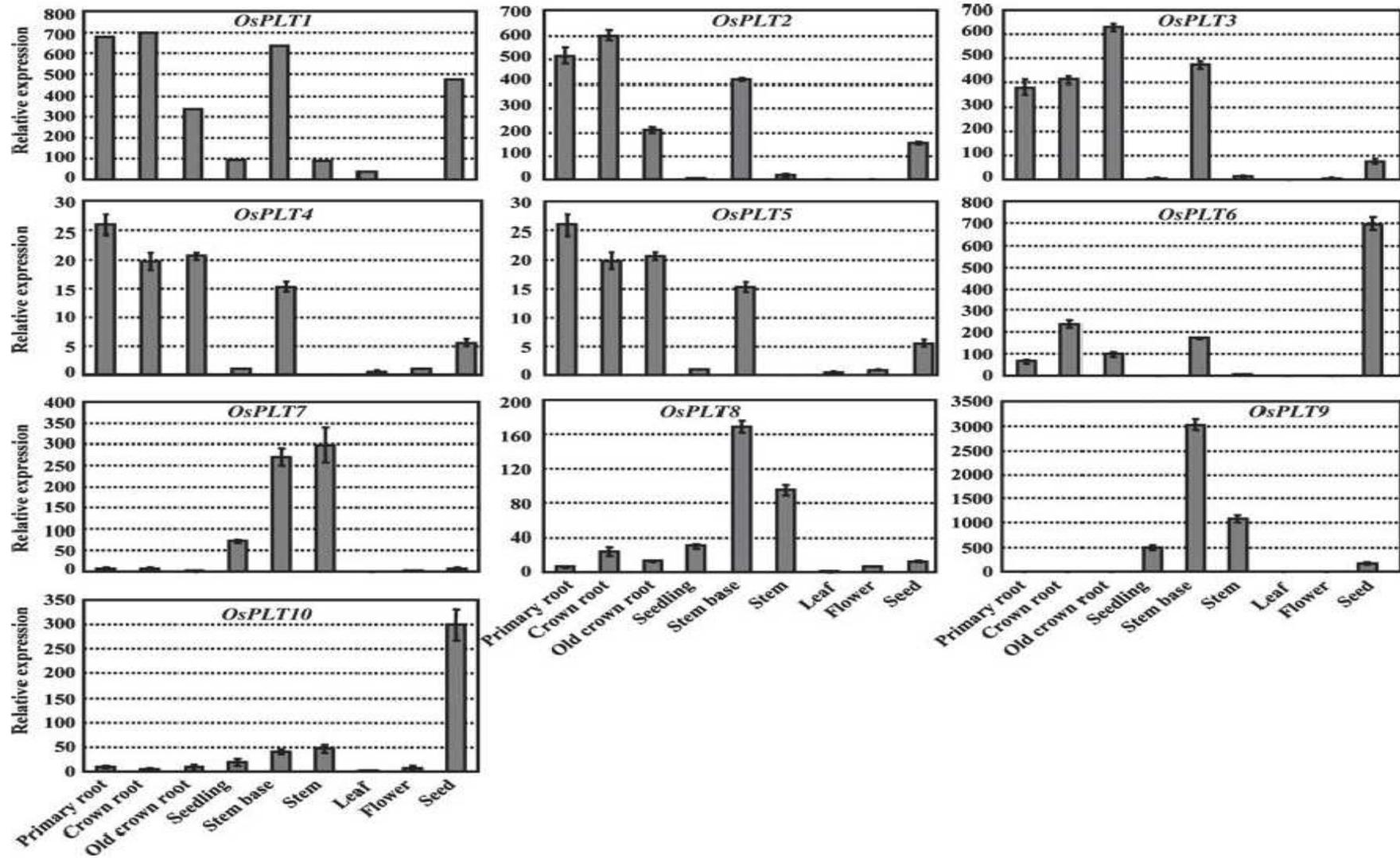
- shoot apical meristem function. *Plant Journal*, **71**, 108–121.
- Nakano, T., Suzuki, K., Fujimura, T. and Shinshi, H.** (2006) Genome-Wide Analysis of the ERF Gene Family. *Plant Physiology*, **140**, 411–432.
- Nole-Wilson, S.** (2000) DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Research*, **28**, 4076–4082.
- Ogawa, M., Hanada, A., Yamauchi, Y., Kuwahara, A., Kamiya, Y. and Yamaguchi, S.** (2003) Gibberellin biosynthesis and response during Arabidopsis seed germination. *Plant Cell*, **15**, 1591–1604.
- Ouakfaoui, S. El, Schnell, J., Abdeen, A., Colville, A., Labbé, H., Han, S., Baum, B., Laberge, S. and Miki, B.** (2010) Control of somatic embryogenesis and embryo development by AP2 transcription factors. *Plant Molecular Biology*, **74**, 313–326.
- Pfaffl, M.W.** (2001) <Nucleic acids research 2001 Pfaffl A new mathematical model for.pdf>. , **29**, 16–21.
- Pfaffl, M.W.** (2002) Relative expression software tool (REST(C)) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research*, **30**, 36e – 36.
- Qi, W., Sun, F., Wang, Q., Chen, M., Huang, Y., Feng, Y.Q., Luo, X. and Yang, J.** (2011) Rice ethylene-response AP2/ERF factor OsEATB restricts internode elongation by down-regulating a gibberellin biosynthetic gene. *Plant Physiology*, **157**, 216–228. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21753115> <http://www.plantphysiol.org/content/157/1/216.full.pdf>.
- Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J. and Xu, Y.** (2012) AP2/ERF transcription factor in rice: Genome-wide analysis and synteny relationships between monocots and dicots. *Evolutionary Bioinformatics*, **2012**, 321–355.
- Ross, E.J.H., Stone, J.M., Elowsky, C.G., Arredondo-Peter, P., Klucas, R. V. and Sarath, G.** (2004) Activation of the *Oryza sativa* non-symbiotic haemoglobin-2 promoter by the cytokinin-regulated transcription factor, ARR1. *Journal of Experimental Botany*, **55**, 1721–1731.
- Sakai, H., Lee, S.S., Tanaka, T., et al.** (2013) Rice annotation project database (RAP-DB): An integrative and interactive database for rice genomics. *Plant and Cell Physiology*, **54**, 1–11.
- Sato, Y., Takehisa, H., Kamatsuki, K., et al.** (2013) RiceXPro Version 3.0: Expanding the informatics resource for rice transcriptome. *Nucleic Acids Research*, **41**, 1206–1213.
- Scheres, B. and Krizek, B.A.** (2018) Coordination of growth in root and shoot apices by AIL/PLT transcription factors. *Current Opinion in Plant Biology*, **41**, 95–101. Available at: <http://dx.doi.org/10.1016/j.pbi.2017.10.002>.
- Sharoni, A.M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., Choi, I.R., Omura, T. and Kikuchi, S.** (2011) Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant and Cell Physiology*, **52**, 344–360.
- Simpson, S.D., Nakashima, K., Narusaka, Y., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K.** (2003) Two different novel cis-acting elements of *erd1*, a *clpA* homologous Arabidopsis gene function in induction by dehydration stress and dark-induced senescence. *Plant Journal*, **33**, 259–270.
- Stein, J.C., Yu, Y., Copetti, D., et al.** (2018) Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nature Genetics*, **50**, 285–296. Available at: <http://dx.doi.org/10.1038/s41588-018-0040-0>.
- Subramanian, B., Gao, S., Lercher, M.J., Hu, S. and Chen, W.** (2019) OUP accepted manuscript. *Nucleic Acids Research*, 1–6.
- Ta, K.N., Adam, H., Staedler, Y.M., et al.** (2017) Differences in meristem size and expression of branching genes are associated with variation in panicle phenotype in wild and domesticated African rice. *EvoDevo*, **8**, 1–14.
- Tanaka, W., Pautler, M., Jackson, D. and Hirano, H.Y.** (2013) Grass meristems II: Inflorescence architecture, flower development and meristem fate. *Plant and Cell Physiology*, **54**, 313–324.
- Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A.M.** (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, **35**, 71–74.
- Vaughan, D.A., Lu, B.R. and Tomooka, N.** (2008) The evolving story of rice evolution. *Plant Science*, **174**, 394–408.
- Vernoux, T., Besnard, F. and Traas, J.** (2010) Auxin at the shoot apical meristem. *Cold Spring Harbor perspectives in biology*, **2**, 1–14.
- Wang, P., Cheng, T., Lu, M., Liu, G., Li, M., Shi, J., Lu, Y., Laux, T. and Chen, J.** (2016) Expansion and Functional Divergence of AP2 Group Genes in Spermatophytes Determined by Molecular Evolution and Arabidopsis Mutant Analysis. *Frontiers in Plant Science*, **7**, 1–15. Available at: <http://journal.frontiersin.org/Article/10.3389/fpls.2016.01383/abstract>.
- Wang, Y., Li, L., Ye, T., Lu, Y., Chen, X. and Wu, Y.** (2013) The inhibitory effect of ABA on floral transition is mediated by ABI5 in Arabidopsis. *Journal of Experimental Botany*, **64**, 675–684.
- Yamaki, S., Miyabayashi, T., Eiguchi, M., Kitano, H., Nonomura, K.-I. and Kurata, N.** (2011) Diversity of

- panicle branching patterns in wild relatives of rice. *Breeding Science*, **60**, 586–596.
- Yoshida, H. and Nagato, Y.** (2011) Flower development in rice. *Journal of Experimental Botany*, **62**, 4719–4730.
- Zhang, B., Wang, L., Zeng, L., Zhang, C. and Ma, H.** (2015) Arabidopsis TOE proteins convey a photoperiodic signal to antagonize CONSTANS and regulate flowering time. *Genes and Development*, **29**, 975–987.
- Zhang, Z.L., Xie, Z., Zou, X., Casaretto, J., Ho, T.H.D. and Shen, Q.J.** (2004) A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells 1[w]. *Plant Physiology*, **134**, 1500–1513.
- Zumajo-Cardona, C. and Pabón-Mora, N.** (2016) Evolution of the APETALA2 Gene Lineage in Seed Plants. *Molecular biology and evolution*, **33**, 1818–1832.

## SUPPLEMENTAL FIGURES AND TABLES

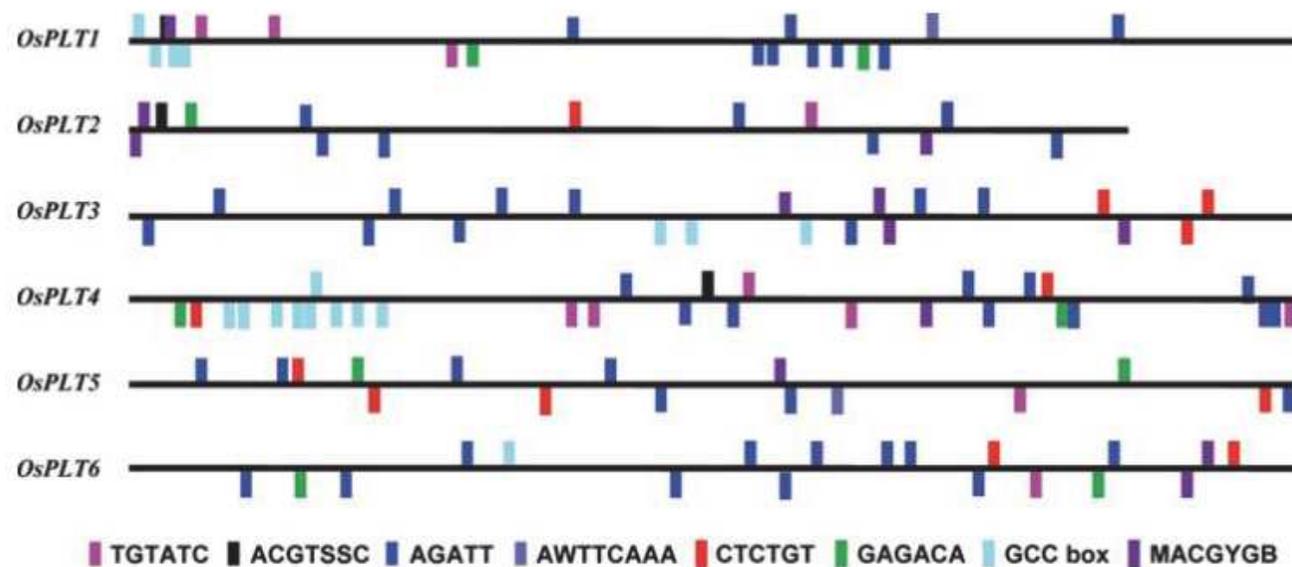


**Figure S2.1. Expression profiling of rice *euANT/PLT* genes in the RiceXPro database (Sato *et al.*, 2013). Spatio-temporal gene expression of various tissues/organs throughout entire growth in the field. DAF, day after flowering.**

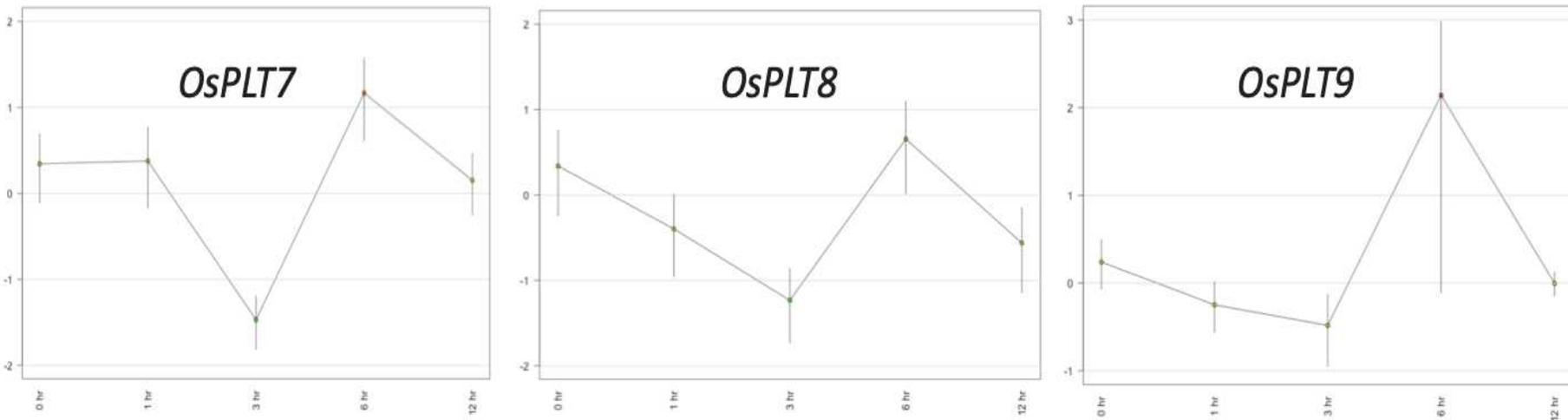


**Figure S2.2. Expression of *OsPLT* genes in various tissues.** qRT-PCR analysis revealed the high expression of *OsPLT1-5* in primary root, crown root, old crown root, stem-base, and seed, and high expression of *OsPLT7/2/9* in seedling, stem-base, and stem. *OsPLT6* and *OsPLT10* were predominantly expressed

Gene	TGTATC	CTCTGT	GAGACA	MACGYGB	AGATT	GCC BOX
<i>OsPLT1</i>	3	0	2	1	8	4
<i>OsPLT2</i>	1	1	1	3	7	0
<i>OsPLT3</i>	0	3	0	4	10	3
<i>OsPLT4</i>	5	2	2	1	10	9
<i>OsPLT5</i>	1	4	2	1	7	0
<i>OsPLT6</i>	1	2	2	2	11	1



**Figure S2.3. Analysis of hormone response elements in the promoter of *OsPLT1-6*.** Computational analyses of promoter sequences were performed using analysis tools from the PLACE website, [http://www.dna.affrc.go.jp/htdocs/ PLACE/](http://www.dna.affrc.go.jp/htdocs/PLACE/)) and manual searches for cis-regulatory elements are performed by TOUCAN software. Numbers are *cis*-elements are listed in the upper table. The position of the *cis*-elements in the 3kb region upstream the start codon of *OsPLT1-6* is illustrated by the lower panel. Auxin-response elements: TGTATC, CTCTGT, GAGACA; ABA-response element: MACGYGB and ACGTSSC; Cytokinin-response element: AGATT; Ethylene-response element: GCC box and AWTTCAAA. Figure modified from (Li and Xue, 2011).



**Figure S2.4. Expression of *OsPLT7*, *8* and *9* in rice shoot samples.** The shoot samples were collected at 0 min (pretreatment), and at 1 h, 3 h, 6 h and 12 h of incubation after auxin treatment (10 $\mu$ M IAA). Y-axis indicates the gene expression ratio (log<sub>2</sub>) of treatment sample *versus* the control (Treatment/Control). Data retrieved from RiceXPro database (<http://ricexpro.dna.affrc.go.jp>) (Sato *et al.*, 2013).

Table S2.1. Members of the AP2 sub-family of *Oryza sativa* ssp. japonica (prefix OsXXX or LOC\_OsXXX) and *Arabidopsis thaliana* (prefix AtXXX)

Classification	Gene Symbol	Representative Transcript (MSU)	Representative Transcript (RAPDB)	Gene Symbol	Representative Transcript
euAP2	OsDS1	LOC_Os03g60430.2	Os03t0818800-01	AtAP2	AT4G36920.1
euAP2	OsRSR1	LOC_Os05g03040.1	Os05t0121600-01	AtSNZ	AT2G39250.1
euAP2	OsSHAT1	LOC_Os04g55560.2	Os04t0649100-03	AtSMZ	AT3G54990.1
euAP2	OsSNB	LOC_Os07g13170.2	Os07t0235800-01	AtTOE1	AT2G28550.3
euAP2	OsAP2/EREBP59	LOC_Os06g43220.1	Os06t0639200-01	AtTOE2	AT5G60120.2
euAP2	OsAP2/EREBP92	LOC_Os05g28800.1	Os05t0356201-00	AtTOE3	AT5G67180.1
baa ANT	OsAP2/EREBP52	LOC_Os05g45954.1	Os05t0536250-00	AtADAR	AT1G16060.1
baa ANT	OsAP2/EREBP68	LOC_Os08g07440.1	Os08t0171100-01	AtWR1	AT3G54320.1
baa ANT	OsAP2/EREBP79	LOC_Os01g59780.1	Os01t0813300-00	AtWR2	AT2G41710.3
baa ANT	OsAP2/EREBP99	LOC_Os06g05340.1	Os06t0145700-00	AtWR4	AT1G79700.2
baa ANT	OsAP2/EREBP109	LOC_Os08g34360.1	Os08t0442400-01		
baa ANT	OsAP2/EREBP131	LOC_Os11g03540.1	Os11t0129700-01		
baa ANT	OsAP2/EREBP153	LOC_Os09g25600.1	Os09t0423800-01		
baa ANT	OsAP2/EREBP155	LOC_Os12g03290.1	Os12t0126300-00		
baa ANT	OsERF142	LOC_Os05g32270.1	Os05t0389000-02		
euANT/PLT	OsPLT1	LOC_Os04g55970.2	Os04t0653600-01	AtAIL1	AT1G72570.1
euANT/PLT	OsPLT2	LOC_Os06g44750.1	Os06t0657500-02	AtAIL6	AT5G10510.3
euANT/PLT	OsPLT3	LOC_Os02g40070.1	Os02t0614300-01	AtAIL7	AT5G65510.1
euANT/PLT	OsPLT4	LOC_Os04g42570.1	Os04t0504500-01	AtANT	AT4G37750.1
euANT/PLT	OsPLT5	LOC_Os01g67410.1	Os01t0899800-01	AtBBM	AT5G17430.1
euANT/PLT	OsPLT6	LOC_Os11g19060.1	Os11t0295900-01	AtPLT1	AT3G20840.1
euANT/PLT	OsPLT7	LOC_Os03g56050.1	Os03t0770700-01	AtPLT2	AT1G51190.1
euANT/PLT	OsPLT8	LOC_Os07g03250.1	Os07t0124700-01	AtPLT5	AT5G57390.1
euANT/PLT	OsPLT9	LOC_Os03g12950.1	Os03t0232200-01		
euANT/PLT	OsPLT10	LOC_Os03g07940.1	Os03t0176300-00		
euANT/PLT	OsAP2/EREBP22	LOC_Os02g51300.1	Os02t0747600-01		
euANT/PLT	OsAP2/EREBP86	LOC_Os03g19900.1	Os03t0313100-01		

**Bold** transcripts- sequences from rice used in alignment of Figure 2. Total: 46 sequences (28 from *O. sativa* ssp. japonica and 18 from *A. thaliana*)

**Table S2.2.** Primers used for validation of the CDS of *PLETHORA* genes in rice

Gene name	Oligo Name	Sequence (5'-3')
LOC_Os03g56050	HA-PLT7-V-F	AGTCCATTGACACGTTCCGGC
	HA-PLT7-V-R	CAGCTCCTCCTGGTAGTCCT
LOC_Os07g03250	HA-PLT8-V-F	ATCGACACGTTCCGGTCAGAG
	HA-PLT8-V-R	GCCACATACTCCTGCCTTGT
	HA-PLT8-V2-R	TCTGGCCTTCCTTCTTGCAG
	HA-PLT8-V3-F	CCAGGAAAGGGAGGCAAGTT
LOC_Os03g12950	HA-PLT8-V3-R	CCCGTGAGAAACCGCTACTT
	HA-PLT9-V-F	CCAAAAGCAGCCTGTTCACC
	HA-PLT9-V-R	CCGCTGCTTCTCCTTCTCAA
	HA-PLT9-V2-F	TGACGGAAGCAGAGATGCTG
LOC_Os02g51300	HA-PLT9-V2-R	CTAACGAACTTGCCTTCCCTTC
	HA-AP22-V-F	TGATGAGAAGGCGGAGAGGA
	HA-AP22-V-R	AGCAGCAATGTCGTAAGCCT

**Table S2.3. Primers used for Fluidigm high-throughput qPCR**

Gene	ID	Oligo Name	Sequence (5'-3')
OsPLT7	LOC_Os03g56050	PLT7-F	TGGCTCACCTCAGAAGGAA
		PLT7-R	TCCTCCTGCGTGCTGAATGT
OsPLT8	Os07g0124700	PLT8-F	TAGGGTTCTTGGTTGCTCGG
		PLT8-R	CGGAGAAGAAGGAAAGGTGG
OsPLT9	Os03g0232200	PLT9-F	CGAGAGAGCAACGCAAGAAC
		PLT9-R	AGAGCGAGAAGCCTAACCAG
AP2/EREBP22	Os02g0747600	EREBP22-F	CTCGGTAGTTGATTCCCTCCC
		EREBP22-R	CACCAATCAATCGCTCTACCC
HK04	LOC_Os01g16970	HA-HK04-F	AGTTCGTCAAGTCTCCATTCC
		HA-HK04-R	CAGAGACTGATTCCAAGCC
HK09	LOC_Os03g61680	HA-HK09-F	TCAAGATAGTCACAGAGAGCC
		HA-HK09-R	AGCATCGGGAAGAGAACAGG
ACT2	LOC_Os11g06390	LOC_Os11g06390-F	ACCAGTAGGAGGAAATGGCTGACGG
		LOC_Os11g06390-R	TGCCCCATACCAACCATCACACC

# **CHAPTER III**

**Functional involvement of *euANT/PLT* genes in panicle architecture determination in *Oryza sativa***

## **Functional involvement of *euANT/PLT* genes in panicle architecture determination in *Oryza sativa***

Luong Ai My<sup>1</sup>, Khong Ngan Giang<sup>2</sup>, Le Quang Hoa<sup>3</sup>, Tregear James<sup>1</sup>, Adam H el ene<sup>1</sup>,  
Jouannic Stefan<sup>1</sup>

1. University of Montpellier, DIADE, IRD, France
2. LMI RICE, Agricultural genetics institute, University of Montpellier, IRD, CIRAD, University of science and Technology of Hanoi, Vietnam
3. Hanoi University of science and technology, Vietnam

## INTRODUCTION

The architecture of the mature inflorescence is initially determined during the reproductive phase, through meristematic activities that define the branching pattern along with flower positioning (Yamburenko *et al.*, 2017). Rice panicle development is governed by the activities of four different types of inflorescence meristem: rachis, branch, spikelet and floral meristems. At the outset of reproductive development, the shoot apical meristem (SAM) is converted into a rachis meristem (RM). Subsequently, some cells differentiate to form primary branch meristems (PBMs) in the axils of newly developed bracts. Bract growth then terminates and the primary branches elongate (ePBMs). At the same time, during elongation, the PBM can generate axillary meristems (AMs), which may develop into secondary and higher-order branches or be directly transformed into lateral spikelet meristems (SMs). Both PBM and secondary branch meristem (SBM) eventually form a terminal SM. The RM, PBM and ePBM/AM stages are indeterminate (i.e. meristematic cells are maintained) whereas the SM is determinate in that the stem cells lose their activity (Ikeda *et al.*, 2004). Grain yield in rice is affected by inflorescence branching, as the number of spikelets produced on the higher-order branches determines the number of grains per panicle (Doebley *et al.*, 2006; Xing and Zhang, 2010; Olsen and Wendel, 2013).

A number of genes affecting branching complexity have been characterised in rice (see Wang and Li, 2011 for a review). Several of these genes encode AP2/ERF transcription factors, which are a family of proteins defined by a conserved domain containing about sixty to seventy amino acids. AP2/ERF transcription factors may be classified into four sub-families (AP2, ERF, DREB and RAV) depending on the number of AP2/ERF domains in the protein and the presence of other DNA binding domains (Rashid *et al.*, 2012; Nakano *et al.*, 2006; Sharoni *et al.*, 2011). There is increasing evidence that certain AP2 sub-family genes are involved in panicle meristem identity determination by regulating the expression of spikelet meristem identity genes. For example, in rice, both the *osids1* and *snb* mutants showed a significant decrease in branch and spikelet number within a panicle (Lee and An, 2012), whereas *SMALL ORGAN SIZE1 (ERF142)* controls organ size and modulates root meristem size in rice without any abnormalities in floral organ (Aya *et al.*, 2014; Hirano *et al.*, 2017).

The *AINTEGUMENTA/PLETHORA (euANT/PLT)* proteins belong to the AP2 sub-family and are well-known to play a role in determining stem cell fate, boosting organ growth and suppressing cell differentiation in the context of shoot meristem function in

*Arabidopsis thaliana* (Scheres and Krizek, 2018; Horstman et al., 2014). The *AINTEGUMENTA* (*ANT*) gene from *A. thaliana* was considered, within the group of *PLT* genes, to be the main regulator of shoot development (Scheres and Krizek, 2018). In rice, previous studies demonstrated that the *ANT*-related genes *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* were expressed in panicle meristems with differential expression between the different type of meristems or between wild and cultivated species from Asia and Africa (Harrop et al., 2016; Harrop et al., 2019). Panicle meristem-specific RNA-seq analysis revealed that *OsPLT7* and *OsPLT8* were highly expressed in the rachis meristem while *OsPLT9* was preferentially expressed in the axillary meristems along with *AP2/EREBP22*, with a lower expression in the rachis meristem (Harrop et al., 2016). High throughput qPCR studies confirmed the expression patterns of *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* during panicle meristem development in *O. sativa ssp. japonica* (Chapter 2), suggesting that the latter genes are potential molecular regulators of panicle architecture. For the present study, we generated CRISPR-Cas9 edited mutants of these *OsPLT* genes in order to study their possible roles in the development of the rice panicle and the regulation of its structure.

## MATERIALS AND METHODS

### Plant materials and growth condition

For *in situ* hybridization samples, *O. sativa ssp. japonica* cv. Nipponbare panicle meristems were collected from 15 plants of each accession. Plants were grown in the greenhouse at IRD Montpellier (France), under long day conditions (14h light/10h dark) at 28°C-30°C, and humidity at 60%. After 6 to 8 weeks, flowering was induced under a cycle of 10h of daylight. Panicles were collected at 4 different stages: stage 1, rachis and primary branch meristem; stage 2, elongated primary branch and secondary meristems; stage 3, spikelet differentiation; stage 4, young flowers with differentiated organs.

The *cr15* mutant and associated wild type line, kindly provided by Prof. Yoshiaki Inukai from Nagoya University (Kitomi *et al.*, 2011), were grown in the greenhouse in Montpellier, France, in October 2017 under long day conditions (14h light:10h dark). After 6 to 8 weeks they were transferred to short day conditions (11h light:13h dark) to induce flowering. The humidity at 60%.

The mutant lines generated from *Oryza sativa ssp. japonica* cv. kitaake were grown in the greenhouse in May 2018 (knock-out mutant, T0 generation), in October 2018 (knock-out mutant, T1 generation) and in March 2019 (knock-out mutant, T2 generation) under long day conditions (14h light:10h dark). After 6 to 8 weeks they were transferred to short day conditions (11h light:13h dark) to induce flowering. The humidity at 60%.

### Generation and selection of CRISPR-Cas9 edited plants

The plasmid vector pRGEB32 was transformed into *Agrobacterium tumefaciens* strain EHA105. CRISPR-Cas9 plasmid vector constructions were carried out using the polycistronic gRNA-tRNA plasmid system according to (Xie *et al.*, 2015). Primers used for vector constructions were listed in **Table S3.1**. The *O. sativa ssp. japonica* cv. kitaake plants were genetically transformed as previously described (Sallaud *et al.*, 2003). Primary transformants (T0 generation) produced T1 seeds that were analysed for segregation of antibiotic resistance (Sallaud *et al.*, 2003).

Polymerase chain reactions (PCRs) were performed with GoTaq DNA Polymerase Reaction Buffer (Promega, USA), according to manufacturer's instruction. DNA fragments were analyzed by electrophoresis on an ethidium bromide stained agarose gel. The primers used are listed in **Table S3.1**. The genome editing device in the regenerated plants was detected with primers specific for the hygromycin resistance gene and the Cas9 gene. The

selected regenerated plants in the T0 generation were those carrying the full-length T-DNA. CRISPR-Cas9 induced genomic deletions were detected by PCR with primers flanking the two target sites of each gene. PCR cycling conditions were as follow: 95 °C for 2 min (1 cycle) and 95 °C for 30 s, an annealing step at various temperatures depending on the T<sub>m</sub> of the primers used (typically T<sub>m</sub> -5 °C), for 30 s, and 72 °C for 1 min (35 cycles) with a 5 min final extension step at 72 °C. The integrity and size distribution of the PCR fragments were checked by ethidium bromide stained agarose gel electrophoresis.

Sanger sequencing of selected PCR products was carried out (Genewiz, UK) to determine the specific mutation. The results were visualized using 4Peaks (<http://www.mekentosj.com>). Double peaks were resolved using the Degenerate Sequence Decoding (Liu *et al.*, 2015) and CRISP-ID (Dehairs *et al.*, 2016) web tools.

### **Plant phenotyping**

For the *crl5*, *erf142* and *erf48* mutants and associated wild type genetic backgrounds (*Oryza sativa ssp. japonica* cv. Kimaze or cv. Nipponbare), at least 18 panicles of each line were used for panicle phenotyping. For the T2 generation of CRISPR-Cas9 induced mutant lines generated from *Oryza sativa ssp. japonica* cv. kitaake, at least 20 to 30 panicles were collected for panicle analysis. For phenotyping analyses, each panicle was spread out and fixed on white paper using adhesive tape. Panicles were photographed and the images were used for panicle structure and seed number analysis with the aid of P-TRAP software (ALTam *et al.*, 2013). Morphological traits of the panicles measured were: rachis length; number of spikelets per panicle; number of primary, secondary or tertiary branches per panicle; and the length of associated internodes (**Fig. S3.1**). Other trait values were recorded during the development of the plants, including flowering time, plant height, the number of tillers and the number of efficient tillers (i.e. tiller producing a panicle) per plant. Plants were photographed using a Canon PowerShot G12 camera.

### **RNA *in situ* hybridization**

Panicle samples were collected at different developmental stages as described in Harrop *et al.* (2019) and were embedded in Paraplast X-TRA (Sigma-Aldrich, USA) as described by Huijser *et al.* (1992). Digoxigenin-labelled antisense and sense RNA probes were generated with the DIG RNA Labelling Kit SP6/T7 (Roche, Germany) according to the manufacturer's instructions. The primers used to generate the probes are listed in **Table S3.2**. Hybridization was performed as described by Adam *et al.* 2011 using a VECTOR Blue

Alkaline Phosphatase Substrate Kit (Vector Laboratories, Burlingame, CA, USA) for detection.

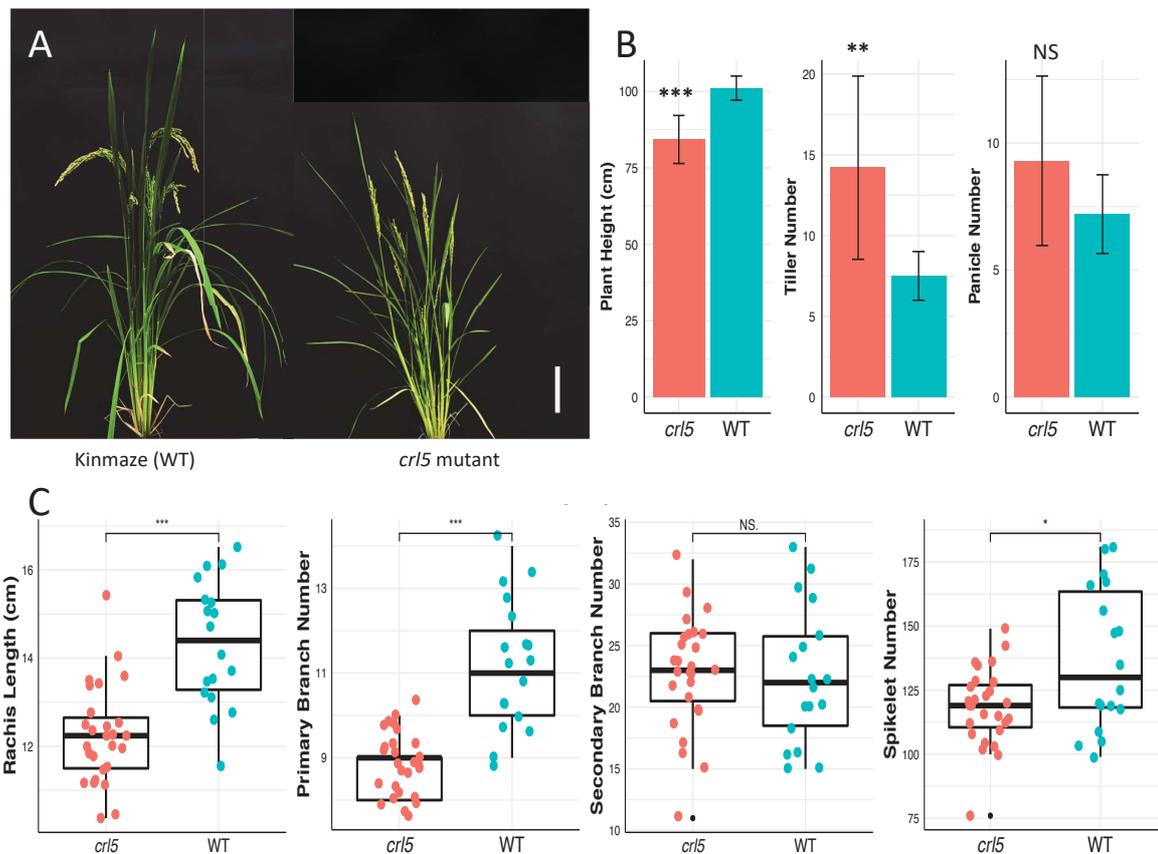
## RESULTS

### Panicle phenotyping of AP2/ERF-related mutants

A phenotypic analysis of early stage rice panicles was carried out for AP2/ERF mutants that were already available at the beginning of this study. Following requests to other laboratories, we received seeds for the loss of function mutant of *ERF142* (LOC\_Os05g32270) (Aya *et al.*, 2014), the loss of function mutant of *OsPLT8/CRL5* (LOC\_Os07g03250) (Kitomi *et al.*, 2011) and the overexpressing line of *ERF48* (LOC\_Os08g31580) (Jung *et al.*, 2017).

Panicle phenotyping data obtained for the *erf142* mutant and the *ERF48* overexpressing line are shown in **Fig. S3.2**. The *erf142* mutant showed a decrease in the number of primary branches, secondary branches and spikelet number. This result, taken on its own, did not allow us to identify which stage(s) of panicle development were affected by the gene knockout. However, previous RNA-seq results indicated low expression of *OsERF142* in the rachis meristem and higher expression during branching and spikelet formation (Harrop *et al.*, 2016), so this phenotype might be explained by a role for *OsERF142* in branch formation, via regulatory activities in cell division and cell expansion. In this connection, *ERF142* has been well characterized as a factor controlling organ size and modulating root meristem size in rice (Aya *et al.*, 2014) (Li *et al.*, 2015). Earlier studies did not however provide evidence for the involvement of the *ERF48* gene in inflorescence development, since an overexpressing transgenic line did not show any difference in panicle phenotype compared to wild type.

The *crl5* mutant was characterised by a lower number of shorter tillers compared to wild-type (WT) plants while the number of efficient tillers and panicles was the same in mutants and wild type (**Fig. 3.1A and B**). The rachis length of *crl5* panicles was significantly shorter than that of wild type. Interestingly, the *crl5* mutant showed a significant reduction in the total number of primary branches compared to the wild type (**Fig. 3.1C and Fig. S3.3**). The total number of secondary branches was conserved, but alongside the decrease in primary branch number, the number of secondary branches per primary branch was increased in the *crl5* mutant compared to WT. Overall, the total number of spikelets per panicle is not significantly affected in the mutant.



**Figure 3.1. Phenotype of the *crl5* mutant.** **A**, Seventy-five-day-old wild-type (WT) Kinmaze (left) and *crl5* (right) plants. Scale bar, 10 cm. **B**, Comparison of plant height, tiller number and panicle number between *crl5* and WT. **C**, Comparison of panicle phenotype (rachis length, the number of primary branches, the number of secondary branches and the number of spikelets per panicle, respectively from left to right) between *crl5* and WT. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.

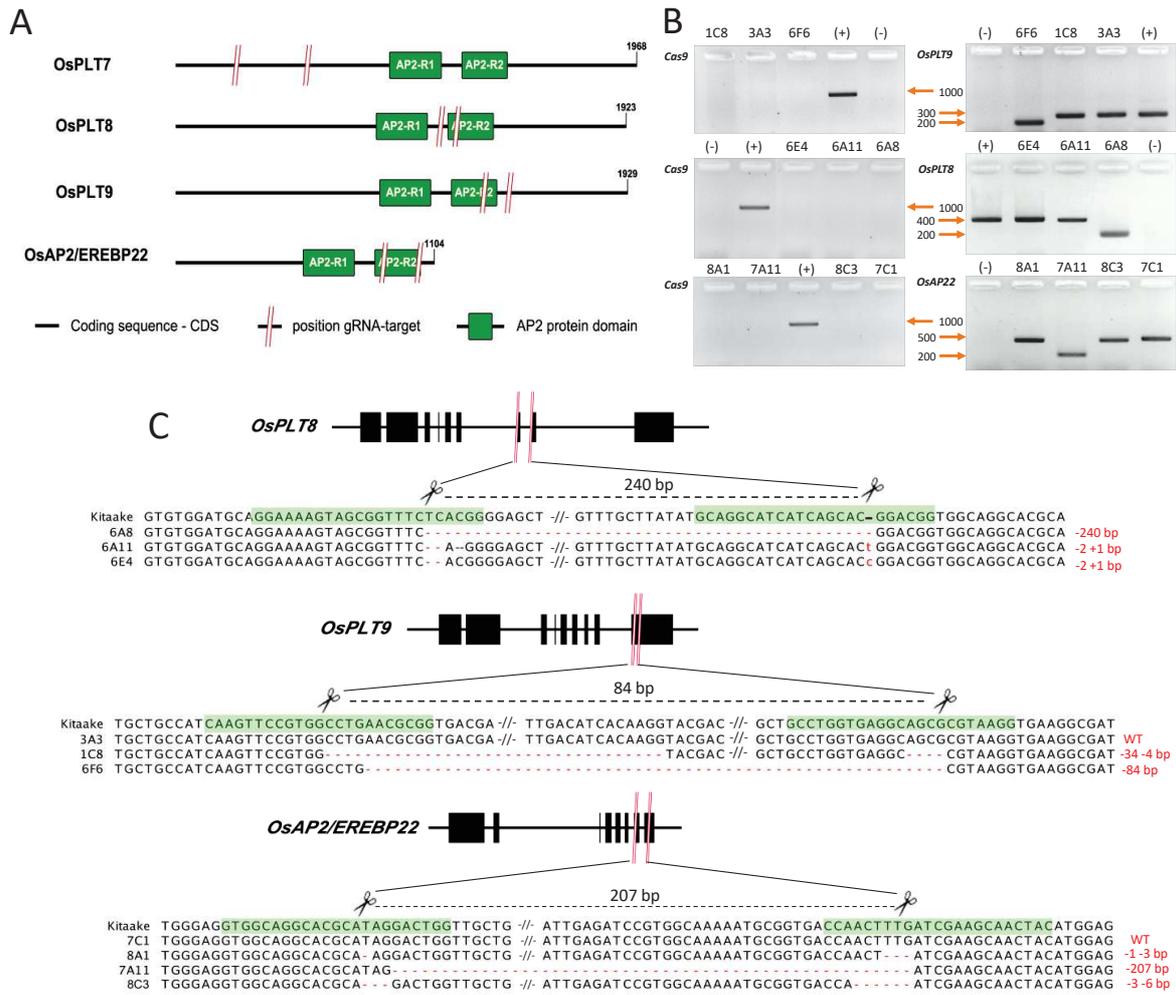
### Generation of CRISPR-Cas9 edited mutants of rice *ANT*-related genes

Based on the results of previous studies, the *OsPLT7*, *OsPLT8*, *OsPLT9* and *OsAP2/EREBP22* genes were selected to generate mutants for further functional analysis in to characterize their contribution to panicle development. To create the knock-out mutants, the CRISPR-Cas9 method with the aid of the polycistronic gRNA-tRNA system was used (Xie *et al.*, 2015). The gRNAs were designed to target genes as illustrated by **Fig. 3.2A**, in favour of creating mutations at the second AP2 domain, with two gRNAs per gene for individual gene targeting (*OsPLT7*, *OsPLT8*, *OsPLT9* & *AP2/EREBP22*) and double targeting for *OsPLT7* and *OsPLT8* genes due to their being closely related. Each transformation resulted in the production of plants that were 100% hygromycin positive. Initial genotyping of T0 plants allowed the selection of plants with homozygous or

heterozygous deletions in the genomic sequence of the *OsPLT8*, *OsPLT9* and *AP2/EREBP22* genes along with sister plants (i.e. transgenic plants without mutations in the targeted genes). The genotyping of *OsPLT7*-targeted plants and of plants expected to have double mutations in *OsPLT7* and *OsPLT8* genes was delayed due to difficulties experienced with the primers designed for the corresponding PCR amplifications. For this reason, these two categories of knockouts were not analysed in the framework of this PhD.

The seeds obtained from selected T0 plants carrying a homozygous deletion in *OsPLT8*, *OsPLT9* and *AP2/EREBP22* genes were sown to obtain the T1 generation. Screening for plants lacking a Cas9 insertion was carried out on these T1 plants in order to obtain the T-DNA-free homozygous mutants .

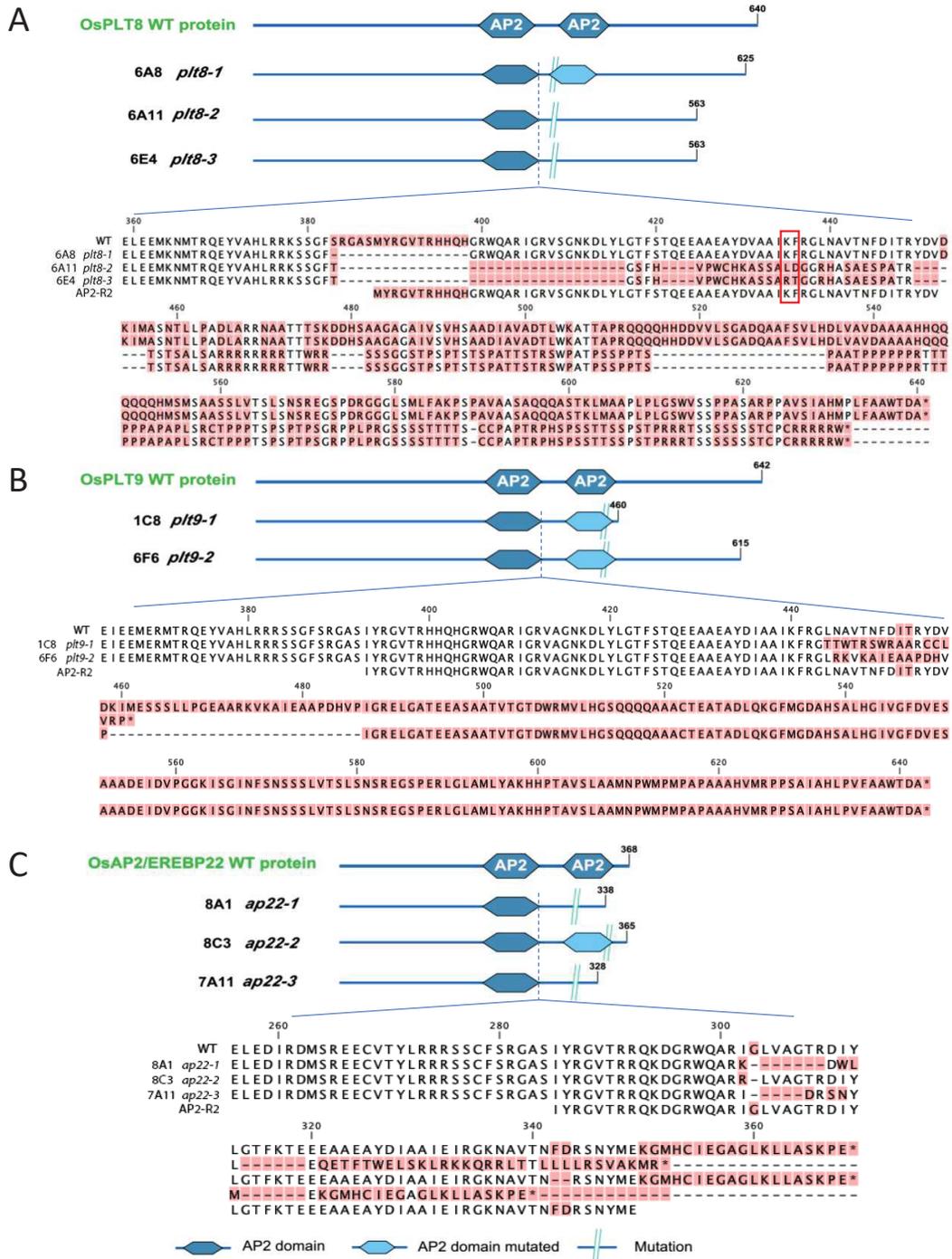
Using the aforementioned approach, we obtained three T-DNA-free homozygote mutant alleles for *OsPLT8*, two for *OsPLT9* and three for *OsAP2/EREBP22*, with at least two independent transgenic lines for each allele (**Fig. 3.2B-C**). Sister plants were identified for the *OsPLT9* and *AP2/EREBP22* genes (3A3 and 7C1 lines respectively), but not for the *OsPLT8* gene. For *OsPLT8*, three allelic lines were selected: *plt8-1* (line 6A8 containing a 240 bp deletion), *plt8-2* (line 6A11 containing a 2 bp deletion and a 1bp insertion) and *plt8-3* (line 6E4 containing a 2 bp deletion and a 1bp insertion). For *OsPLT9*, two allelic lines were selected, namely *plt9-1* (line 1C8 containing a 38 bp deletion) and *plt9-2* (line 6F6 containing a 84 bp deletion). For *AP2/EREBP22*, three allelic lines were selected, namely *ap22-1* (line 8A1 containing a 4 bp deletion), *ap22-2* (line 8C3 containing a 9 bp deletion) and *ap22-3* allele (line 7A11 containing a 207 bp deletion).



**Figure 3.2. Generating the edited mutants of *OsPLT* genes.** **A**, schematic view of the position of the gRNA-target positions. The black lines indicate the coding sequence (CDS) with the length of CDS was shown in the tail. The red strips represent the gRNA-target positions. The green box represents the AP2-binding domain. **B**, Gel-electrophoresis of PCR products amplified from genomic DNA of mutant lines. Agarose gel 2%, ladder 1kb. **C**, Representative structure and sequence of genomic DNA deletion aligned with that of wild-type (WT). Black boxes indicate exons. The red strips represent the gRNA-target positions. The gRNA paired region is labelled with green colour. The numerals at the end indicate numbers of deleted (-) or inserted (+) bases between two Cas9 cuts. The total length between two Cas9 cut sites (labelled with scissor) is indicated on the top. Red letters (in *OsPLT8*) indicate the inserted nucleotides. Red dashes in the aligned sequences indicate deletions. The scratch marks (---) resume the normal sequence. bp, base pairs.

Based on the genomic deletion observed in each line, the resulting changes in protein sequence can be predicted (**Fig. 3.3**). For *OsPLT8*, the predicted protein of *plt8-1* contained a deletion at the beginning of the second AP2-domain without changing the downstream sequence of the protein. In contrast, the predicted protein sequences for *plt8-2* and *plt8-3* were shorter in length, due to a totally lacking second AP2-domain and a completely changed C-terminal region. For *OsPLT9*, the predicted protein of *plt9-1* carried a mutation at the end of the second AP2-domain and also lacked the C-terminal region. In contrast, the predicted protein of *plt9-2* contained a deletion at the end of the second AP2-domain but retained its C-terminal region. For *AP2/EREBP22*, the predicted proteins of *ap22-1* and *ap22-3* lacked their second AP2-domain and possessed a modified C-terminal region while that of *ap22-2* contained only a 3 amino acid deletion in the second AP2-domain.

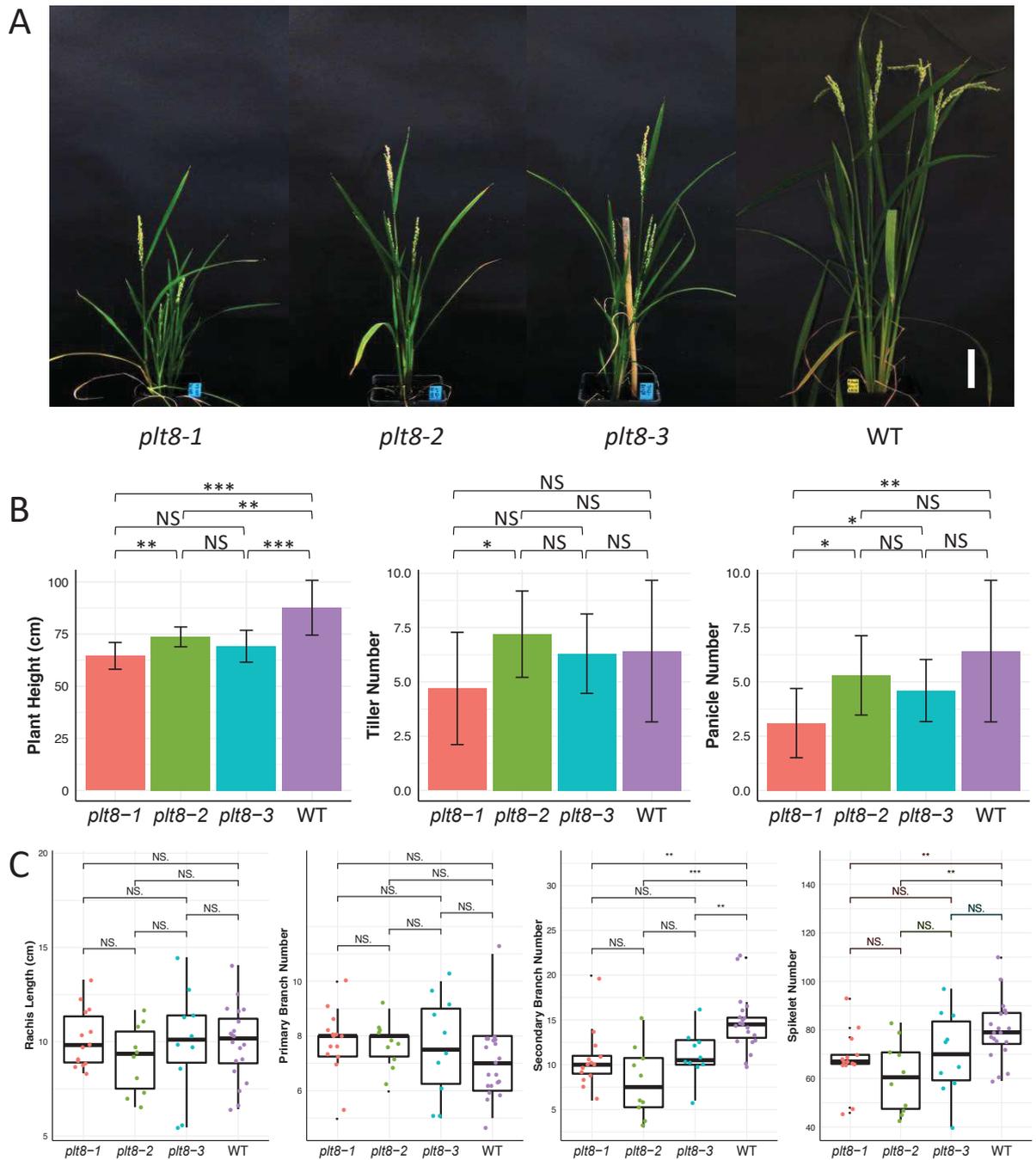
The T2 plants from selected mutant lines were grown for phenotyping and seed multiplication with 10 plants per line studied for *OsPLT8* and *AP2/EREBP22* and 12 plants per line for *OsPLT9*. Due to limited seed numbers obtained from T1 plants, the line containing the *ap22-3* allele (7A11 line with a 207 bp deletion) was grown in the T2 generation only for seed multiplication and not for phenotyping. The wild-type cultivar Kitaake plants were grown as control plants for *plt8* mutant lines. The 3A3 and 7C1 sister lines were grown as the wild-type control plants for the *plt9* and *ap2/erebp22* mutant lines respectively. The T2 seedling genotypes were confirmed for each mutant and sister line.



**Figure 3.3. Predicted protein sequence changes in the edited mutants of *OsPLT* genes.** Schemas describing the protein structure with protein domain for each mutant line of *OsPLT8* (A), *OsPLT9* (B) and *AP2/EREBP22* (C), respectively, along with representative sequences of the protein deletion aligned with that of wild-type (WT). The blue lines indicate the protein sequence with the length of protein was shown in the tail. The light blue strips represent the mutation positions. The dark blue box represents the normal AP2-binding domain. The light blue box represents the mutant AP2-binding domain. *OsPLT8* mutant lines: 6A8, *plt8-1*, 6A11, *plt8-2*, 6E4, *plt8-3*. *OsPLT9* mutant lines: 1C8, *plt9-1*, 6F6, *plt8-2*. *OsAP2/EREBP22* mutant lines: 8A1, *ap22-1*, 8C3, *ap22-2*, 7A11, *ap22-3*. AP2-R2, second AP2 domain sequence. Pink background of the alignment indicate the protein changes in the mutant compared to WT.

**The *plt8* mutants produce fewer secondary branches**

The mean height of *plt8-1*, *plt8-2* and *plt8-3* plants 75 days after sowing was much shorter than that of the wild-type (WT) (**Fig. 3.4A**). It was noted that the above-ground part of some mature *plt8* plants had collapsed. This may be a consequence of an unstable root system as observed in the *cr15* mutant. Although the number of tillers in the *plt8-1* mutant remained the same as WT, the number of panicles in plants of this genotype was significantly reduced. In contrast, numbers of tillers and panicles in plants of *plt8-2* and *plt8-3* lines were equal to those of the WT (**Fig. 3.4B**). The *plt8-2* and *plt8-3* lines were not found to display any significant difference in phenotype. In spite of the rachis length of all mutant lines being equivalent to that of the WT, the lengths of both primary branches and secondary branches of *plt8* panicles, along with secondary branch internode length, were found to be shorter than those of WT panicles (**Fig. S3.4**). Although the number of primary branches did not change, the number of secondary branches decreased in all *plt8* mutants compared to that of the WT, resulting in a reduction in spikelet number per panicle in *plt8* mutants (**Fig. 3.4C**). This result suggested that *OsPLOT8* might play a role in the formation of secondary branches and in branch elongation.



**Figure 3.4. Phenotype of the edited mutants of *OsPLT8*.** **A**, Seventy-five-day-old *plt8-1*, *plt8-2*, *plt8-3* and wild-type (WT) plants. Scale bar, 10 cm. **B**, Comparison of plant height, tiller number and panicle number between *plt8* and WT plants. **C**, Comparison of panicle phenotype (rachis length, number of primary branches, number of secondary branches and number of spikelets per panicle, respectively from left to right) between *plt8* mutants and WT. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.

Several additional defects were observed, including a delay in flowering time (**Table 3.1**), and a lower maturation rate compared to that of the WT, as previously observed in the *crl5* mutant (Kitomi et al., 2011). The phenotyping of *plt8* plants was carried out on a limited number of panicles due to the low yield of seeds from the T1 generation. This phenotyping will therefore need to be confirmed in the T3 generation using a larger panel of plants and panicles.

**Table 3.1.** Summary of the flowering time of rice mutants and wild-type relatives

Genotype	1st flowering (DAG)	50% flowering (DAG)	100% flowering (DAG)
<b>Spring 2018</b>			
<i>crl5</i>	78	83	96
Kinmaze	83	84	96
<b>Spring 2019</b>			
<i>plt9-1</i>	45	48	54
<i>plt9-2</i>	45	48	54
<i>plt9</i> sister-line	48	50	56
<i>plt8-1</i>	62	66	68
<i>plt8-2</i>	53	59	63
<i>plt8-3</i>	59	65	67
Kitaake	48	54	61
<i>ap22-1</i>	46	53	56
<i>ap22-2</i>	45	46	52
<i>ap22</i> sister-line	52	54	61

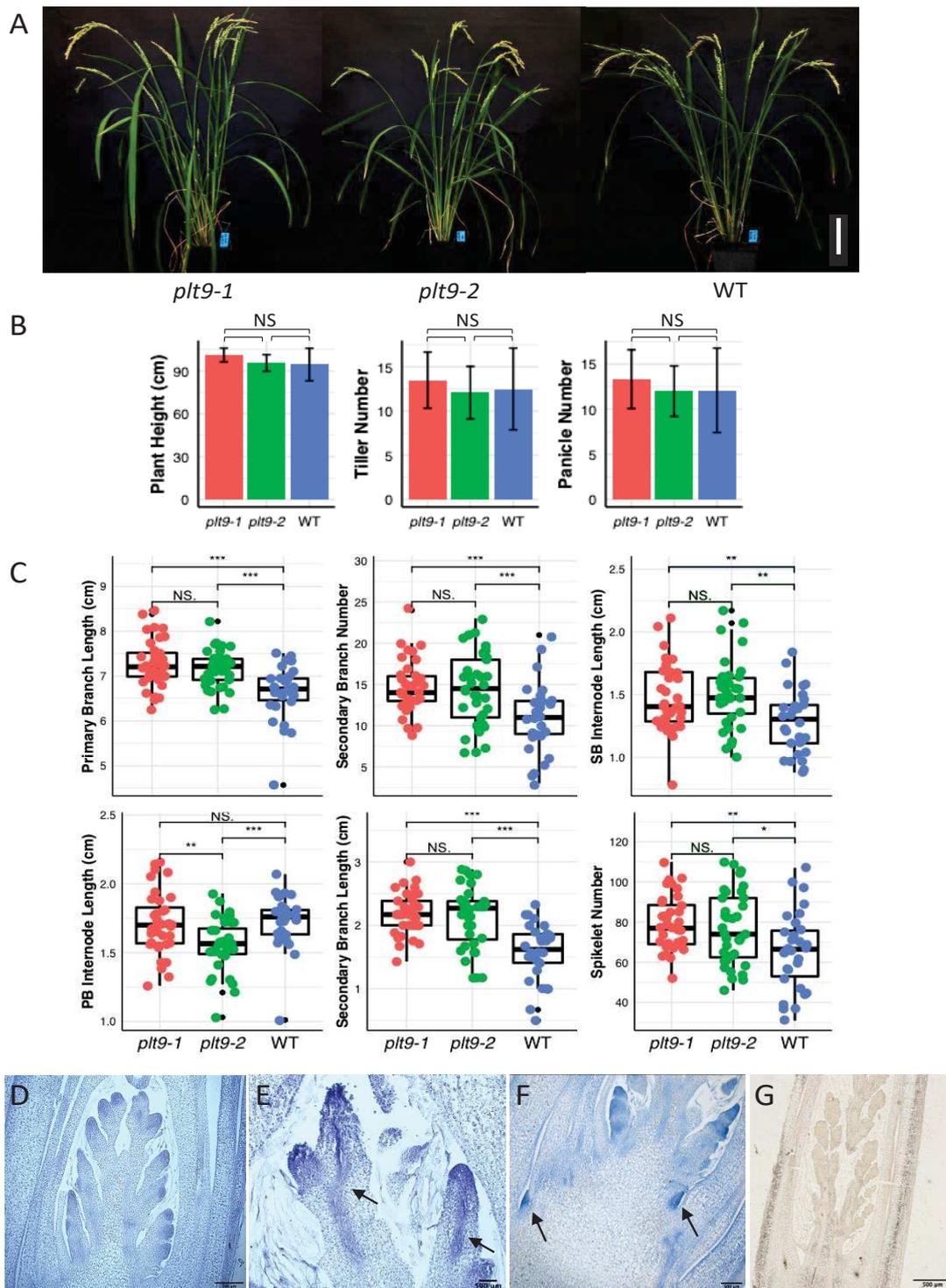
**DAG:** Days after germination

**Larger numbers of secondary branches produced in the *plt9* mutant lines**

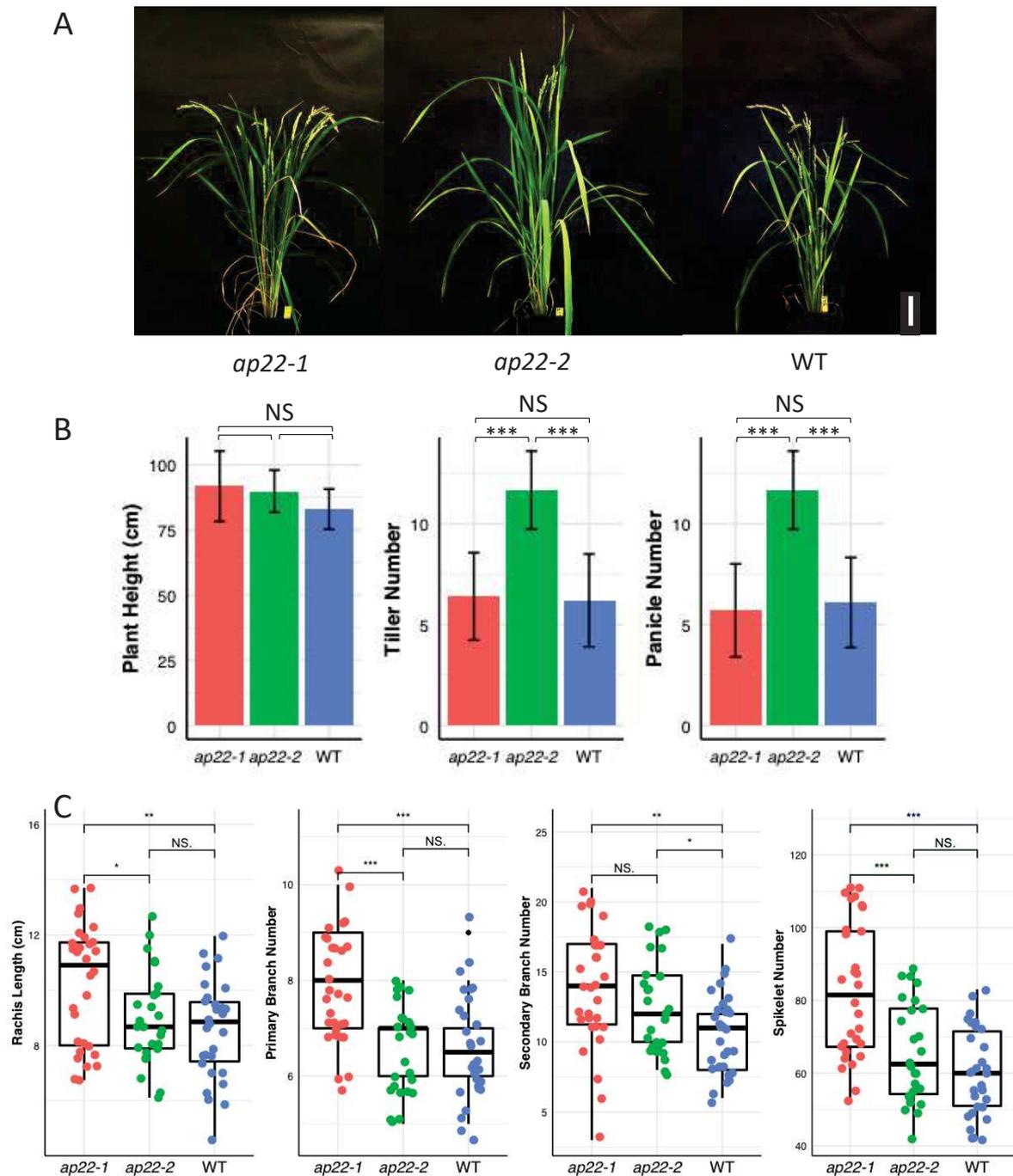
The overall phenotypes of *plt9-1*, *plt9-2* and the sister line plants 75 days after sowing were much alike with a similar height, a comparable flowering time and similar numbers of tillers and panicles (**Fig. 3.5A and B, Table 3.1**). The panicle phenotypes of *plt9-1* and *plt9-2* plants were similar in terms of rachis length, primary and secondary branch length, secondary branch internode length, numbers of primary and secondary branches and spikelet number. An exception to this general pattern was primary branch internode length (**Fig. S3.5**). The internode length of the *plt9-1* panicle primary branch was equivalent to that of the sister-line. In contrast, although no change was detected in the number of primary branches in *plt9-2* panicles, their length was significantly greater than that of the sister line. Additionally, both *plt9-1* and *plt9-2* panicles showed a significant increase in lengths of secondary branches and secondary branch internodes, along with a higher number of secondary branches compared to those of the sister line, thereby resulting in an increased spikelet number per panicle (**Fig. 3.5C and S3.5**).

**The *ap22/erebp22* mutants produce more primary and secondary branches**

While *ap22-1* and sister line plants were similar in terms of height, tiller number and panicle number at 75 days after sowing, the *ap22-2* line plants produced significantly more tillers and panicles than the sister line with a similar plant height (**Fig. 3.6A and B**). The lengths of primary and secondary branches and the lengths of primary and secondary internodes were similar in the *ap22-1* mutant and sister lines (**Fig. S3.6**). Nevertheless, rachis length, primary and secondary branch numbers and spikelet number were considerably higher in the *ap22-1* mutant compared to the sister line (**Fig. 3.6C**), implying a role for *OsAP2/EREBP22* in panicle branch meristem function. Meanwhile, with the exception of secondary branch length, no difference was observed between *ap22-2* and the sister line for all the other scored traits (**Fig. 3.6C and S3.6**). With the same spikelet number per panicle but a larger number of panicles, the *ap22-2* plant was able to produce more seeds than the sister line. Finally, both the *ap22-1* and *ap22-2* mutant lines showed an earlier flowering time of about one week compared to the sister-line (**Table 3.1**).



**Figure 3.5. Phenotype of the edited mutants of *OsPLT9*.** **A**, Seventy-five-day-old *plt9-1*, *plt9-2* and wild-type (WT) (sister-line) plants. Scale bar, 10 cm. **B**, Comparison of plant height, tiller number and panicle number between *plt9* and WT plants. No difference between plant height, tiller number and panicle number between mutants and WT. **C**, Comparison of panicle phenotype between *plt9* and WT plants. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant. **D-G**, *in situ* hybridizations of *OsPLT9* used as a probe in panicle (D-E) and tiller (F) meristems. The arrows indicate signals in vascular tissue; the result obtained using a sense probe is shown in G. Scale bar, 500  $\mu\text{m}$ .



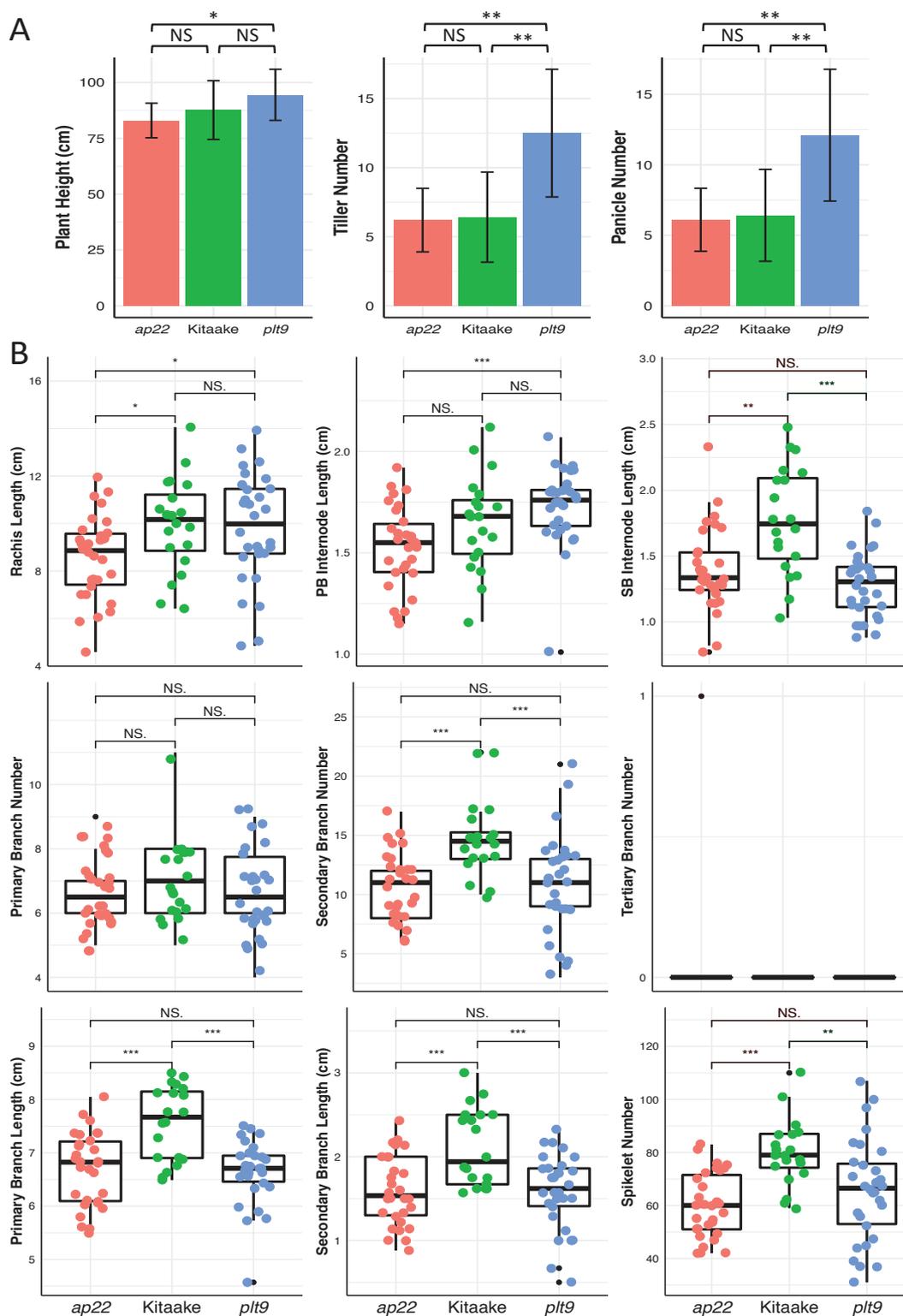
**Figure 3.6. Phenotype of the edited mutants of *OsAP2/EREBP22*.** **A**, Seventy-five-day-old *ap22-1*, *ap22-2* and wild-type (WT) (sister-line) plants. Scale bar, 10 cm. **B**, Comparison of plant height, tiller number and panicle number between *ap22* and WT plants. **C**, Comparison of panicle phenotype (rachis length, number of primary branches, number of secondary branches and number of spikelets per panicle, respectively from left to right) between *ap22* and WT plants. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.

### Comparison between the sister lines and the natural wild-type plants

Sister lines generated during the rice transformation but without any CRISPR-Cas9 induced mutation in the targeted gene are generally considered to be the best control for assessing the characteristics of mutants. However, due to the high efficiency of the CRISPR-Cas9 system, the obtaining of sister-lines may often be difficult. In our study, experiments carried out to generate mutants of *OsPLT9* and *AP2/EREBP22* genes provided sister lines without any mutation in the targeted gene; however, for *OsPLT8* no sister line was obtained. In this case, the natural wild-type was used as the control.

It should be noted that some differences were observed between the sister-lines and the natural wild-type (WT) of *O. sativa ssp. japonica* cv. Kitaake. Although the *ap-22* sister line and the natural WT shared comparable height, tiller number and panicle number (**Fig. 3.7A**), panicles of the *ap22-2* sister line were smaller than those of natural WT with shorter rachis, primary branch, secondary branch and secondary branch internode lengths, causing a significant reduction in spikelet number per panicle compared to that of WT (**Fig. 3.7B**). In contrast, the *plt9* sister line produced significantly more tillers and panicles per plant than the WT (**Fig. 3.7A**). However, the natural WT plants produced larger panicles with more secondary branches than the *plt9* sister line panicle, leading to a significantly larger number of spikelets per panicle compared with the *plt9* sister line (**Fig. 3.7B**).

This finding might suggest an impact of the tissue culture process on plant development after regeneration, even after several generations. The tissue culture-induced remobilisation of transposable elements through epigenetic modifications of the genome (Sabot *et al.*, 2011) may cause changes to plant phenotype. However such changes might be also the consequence of the presence of CRISPR-Cas9 induced “off target” mutations in other loci, even if the specificity of the gRNAs used was sufficient. This might explain the different phenotypes observed between the sister plants, even if the two transformations were conducted in parallel. Overall, the differences in phenotype observed between the sister lines and the natural wild-type indicate that the best negative control is still the sister line, considering the fact that these plants experience the same tissue culture-induced genomic stress and were both potentially subject to “off target” mutations. In this light, any conclusions drawn concerning the *plt8* mutants need to be considered as preliminary.



**Figure 3.7. Phenotype comparison of the sisterlines and kitaake WT.** **A**, Comparison of plant height, tiller number and panicle number between *ap22*-sisterline, WT Kitaaake and *plt9*-sisterline plants. **B**, Comparison of panicle phenotype between sisterlines and WT plants. PB, Primary branch; SB, Secondary branch. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.

### Spatial expression profiling of *OsPLT* genes in early stages of panicle development

*OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* transcripts were detected in the panicle with differential patterns using a meristem-specific RNA-seq-based approach (Harrop et al. 2016). These data prompted us to examine spatial expression of these genes in panicle meristems using *in situ* hybridization. RNA *in situ* hybridization revealed that *OsPLT9* was highly expressed in the axillary branch and spikelet meristems during early development in the panicle of *O. sativa ssp. japonica* cv. Nipponbare (**Fig. 3.6D**). Moreover, *OsPLT9* was also strongly expressed in the vascular tissue of elongating branches (**Fig. 3.6E**) and in vegetative meristems generating new tillers (**Fig. 3.6F**), suggesting that *OsPLT9* might act not only in the functioning of axillary branching meristems but also in tiller development.

Unfortunately, it was not possible to get robust *in situ* hybridization data for the *OsPLT7*, *OsPLT8* and *AP2/EREBP22* genes. Two distinct probes were designed for each of these genes but led to similar results. In the case of *OsPLT7* the sense probes produced a stronger signal than the antisense ones. No transcript was detectable using the *OsPLT8* probes. A generalised coloration was observed throughout the sample, suggesting a non-specific background signal. The use of more stringent *in situ* hybridization conditions did not improve the quality of the signal.

## DISCUSSION

### Second AP2 domain mutations affect the function of *OsPLT* candidate genes

The most highly conserved regions within the AP2 proteins are the two AP2 domains and the region between these two repeats (Kim *et al.*, 2005; Zumajo-Cardona and Pabón-Mora, 2016). The AP2 domains are essential for biological activity (Jofuku *et al.*, 1994) and the conserved linker is crucial for DNA binding through direct contact with the genomic DNA or positioning of the two AP2 repeats thereon (Krizek, 2003; Nole-Wilson, 2000). Changes to the AP2-binding domain result in functional modifications. For *OsPLT8*, the *plt8-1* allele almost preserved the two domains while *plt8-2* and *plt8-3* lost the second one. The *plt8-2* and *plt8-3* alleles code for near-identical proteins except for 2 amino acids, which is consistent with the similar phenotypes observed for these two mutant lines. The lack of the second AP2 domain in the latter two mutant alleles led to a reduction in panicle number. However, flowering time was not affected, in contrast with the phenotype of the *plt8-1* allelic line.

For the *OsPLT9* gene, the *plt9-1* mutant allele encoded a truncated protein form lacking a C-terminal domain with the second AP2 domain partially modified, whereas the protein specified by the *plt9-2* allele is characterised by a short deletion in the C-terminal region after the second AP2 domain. The similar phenotypes observed between *plt9-1* and *plt9-2* mutant plants indicate that the C-terminal region is not essential for *OsPLT9* activities in panicle meristems. However, both *plt9-1* and *plt9-2* mutant alleles are associated with larger panicles (i.e. longer branches and internodes) with larger numbers of secondary branches. This result suggests that the presence of a functional second AP2 domain in the PLT9 protein is important for its biological activity.

For *OsAP2/EREBP22*, the *ap22-1* mutant allele encoded a truncated form of the protein lacking the second AP2 domain and led to the formation of more highly branched panicle with an increase in spikelet number. The *ap22-2* allele specified a protein lacking only 3 amino acids in the second AP2 domain. This more minor structural alteration might explain why the panicle showed little modification compared to that of the sister line. Nevertheless, these smaller changes in the structure of the second AP2 domain led to a significant increase in tiller and panicle production.

Overall, the diverse phenotypic effects associated with the different protein modifications specified by the mutant alleles suggest differential effects on DNA-binding specificity and/or protein-protein interactions, the latter most likely involving N- and C-terminal regions, leading to the alteration of different developmental processes. The present data indicate that the AP2 domain, specifically the second one, plays an important and distinct role in the activity of OsPLT proteins in the panicle meristem and directly impacts upon the development and architecture of the rice panicle. Perturbations to OsPLT protein functions may result in a reduction in spikelet number, such as in the case of *OsPLT8*, or an increase in spikelet yield as observed for *OsPLT9* and *OsAP2/EREBP22*.

### ***OsPLT* candidate genes impact on the rice panicle architecture**

We generated the *osplt8*, *osplt9* and *osap2/erebp22* edited mutants in the same genetic background (*O. sativa ssp. japonica* cv. Kitaake) in order to assess the function of these genes in the development of the rice panicle. The *plt8* mutants showed no change in primary branch number but a significant reduction in the total number of secondary branches. The panicle phenotype of the *plt8* mutants raises questions about the role of the AP2 domain in

the biological activity of OsPLT8. Panicle phenotyping of the *crl5* mutant corresponding to the *OsPLT8* gene revealed a significant decrease in the total number of primary branches while the total number of secondary branches was unchanged and spikelet number only slightly reduced. This implies that the number of secondary branches generated from each primary branch increased in the mutant, suggesting that *OsPLT8* might act as a suppressor of axillary branching like the *FRIZZY PANICLE (FZP)* gene encoding an ERF subfamily protein (Komatsu *et al.*, 2003). In the developing rice inflorescence, the rachis meristem (RM), which is derived from the shoot apical meristem (SAM), eventually aborts after producing several primary branches, and is left as a vestige at the base of the terminal primary branch. It is interesting to note that both RNA-seq and qPCR results confirmed the higher expression of *OsPLT8* at the RM stage (see Chapter 2), suggesting that this gene might function in the regulation of primary branch formation via the promotion of RM activity, in a role resembling that of TAWAWA1 (Yoshida *et al.*, 2013).

The apparently conflicting results obtained, in terms of panicle phenotypes, between the *crl5* and *plt8* mutants might be attributable to differences between the forms of OsPLT8 protein produced in the different mutants. The *crl5* mutant was characterised by a single nucleotide substitution in the *OsPLT8* gene, which resulted in a non-sense mutation (leading to a truncated protein missing the two AP2 domains). In contrast, the *plt8* mutants might retain partial OsPLT8 biochemical function since they contain only changes in the second AP2 domain, the N-terminal domain and first AP2 domain being preserved. *OsPLT8/CRL5* was also shown to be induced by auxin and involved in the *de novo* pathway to initiate crown roots in rice, as well as being highly expressed in the early stages of inflorescence development (Kitomi *et al.*, 2011). Additionally, the *plt8* mutants also displayed a deficient root system, confirming the involvement of *OsPLT8* in crown root initiation as observed for *crl5* (Kitomi *et al.*, 2011).

The *plt9* mutations exerted a strong impact on the elongation of branches and branch internodes along with an increase in total secondary branch number and an increased spikelet number. The strongly localised expression of *OsPLT9* in the panicle axillary meristem observed by *in situ* hybridization confirmed earlier results obtained by meristem-specific RNA-seq (Harrop *et al.*, 2016), suggest a role for this gene in the control of the activity of the axillary meristems. This gene is also expressed in the vascular tissue of elongating branches, which might help explain its apparent importance in branch elongation. Some *AP2/ERF* genes were described previously as being involved in internode elongation. In *A. thaliana*,

ERF11 promotes plant internode elongation by activating GA biosynthesis, and the expression of *GA3ox1* and *GA20oxs* are increased in *ERF11*-overexpressing plants (Zhou *et al.*, 2016). In rice, *OsEATB* gene (for ERF protein Associated with Tillering and panicle Branching) restricts internode elongation by down-regulating a gibberellin biosynthetic gene (Qi *et al.*, 2011). The *OsEATB* over-expressing mutant produced more secondary branches, more spikelets per panicle and more panicles per plant. Additionally, a large body of physiological research has shown that ethylene and gibberellin are involved in the plant elongation process. Cross-talk mediated by *OsEATB* between ethylene and GA underlies differences observed in rice internode elongation (Qi *et al.*, 2011). However, the promoter region of *OsPLT9* does not contain any GA- or ethylene-responsive elements and plant height in the *plt9* mutant was the same as for the sister line, raising questions about the relationship between the panicle internode and plant internode elongation and their differential regulation. However, little is known about how these two characters interact. A study by (Sunohara *et al.*, 2003) indicated that the elongation of lower internodes was regulated independently of that of the upper ones and that culm elongation, especially for the upper internodes, was affected by the early developmental mode of panicles. The strong expression of *OsPLT9* in vascular tissues and in the pre-meristem initiating the new tiller suggests a potential function for this gene in tiller development.

The strong effect of *osap2/erebp22* mutations on secondary branch development and the numbers of tillers and panicles can be explained by the functioning of the AP2 domain, as described above. While *OsPLT8* and *OsPLT9* are closely related to *ANT* of *A. thaliana*, *OsAP2/EREBP22* groups into the same cluster as the *AIL6* and *AIL7* proteins. In *A. thaliana*, the *PLT* transcription factors form a gradient that controls stem cell identity, meristem identity, cell expansion, and cell differentiation (Santuari *et al.*, 2016). The *ANT*, *AIL5*, *AIL6/PLT3* and *AIL7* genes show partially overlapping, but distinct, expression patterns within the inflorescence meristem and in developing flowers (Horstman *et al.* 2014). Moreover, it has been shown that *euANT/PLT* genes are of importance in *de novo* meristem formation. Induction of the *AIL6/PLT3*, *PLT5* and *PLT7* genes is among the earliest transcriptional responses to cytokinin in the context of *de novo* shoot and root formation from callus in *A. thaliana*. It is also relevant to note that the *plt3 plt5 plt7* triple mutants are defective in shoot regeneration (Ikeuchi *et al.*, 2016). Collectively, the latter observations suggest that *OsPLT8*, *OsPLT9* and *OsAP2/EREBP22* might act redundantly in governing the branching of the rice panicle. Furthermore, the *crl5*, *osplt8*, *osplt9* and *osap2/erebp22*

mutants have been shown to display an altered flowering time (this study, Kitomi et al. 2011) suggesting an additional role in the control of the transition from the vegetative to the reproductive phase. Further characterization of panicle development and meristem size at very early stages (ie. during the branching phase) in relation to the *in situ* expression patterns of *OsPLT8*, *OsAP2/EREBP22* and other developmental landmark genes, should help clarify some of the questions emanating from the present study. More widely, it will be of interest to address the role of these *OsPLTs* in other aspects of rice development, including root phenotype and flower shape.

## CONCLUSIONS

This work is the first report on the involvement of the *PLETHORA* genes *OsPLT8*, *OsPLT9* and *OsAP2/EREBP22*, in the the regulation of the reproductive phase (i.e. flowering time and panicle architecture) in rice. We successfully generated loss of function mutants of these genes by use of the CRISPR-Cas9 system and described the phenotypes of these mutants. Overall, the panicle phenotypes observed for mutants of the *OsPLT* genes suggest that they regulate branch meristem determination, their activity depending on the presence of a functional AP2 domain at the protein level. Future work should focus on investigating the regulatory networks in which the *OsPLT* genes operate to control meristem identity in the developing panicle. Further studies of the expression of these genes in other rice accessions and in different conditions of stress and hormone response should help improve our understanding of their role in inflorescence development and provide a driving force for breeding programs aimed at improving yield and environmental adaptation.

## ACKNOWLEDGEMENTS

We thank Prof. Yoshiaki Inukai from Nagoya University for providing the *crl5/plt8* and Kinmaze seeds, Miyako Ueguchi-Tanaka from Nagoya University for providing the *erf142* and Nipponbare seeds, and Ju-kon Kim from Seoul National University for providing *oserf48* and Illmibyeo seeds. We also thank Sophie Chéron and Harold Chrestin for their strong support for plant care and seed harvesting.

## **FUNDING**

This research was funded by Agropolis Fondation through the “Investissements d’Avenir” programme (ANR-10-LABX-0001-01), Fondazione Cariplo (EVOREPRICE 1201-004), and the CGIAR Research Program on Rice. We also received support from the Excellence Scholarship Program of the Embassy of France in Vietnam (LAM) and the “Allocations de Recherche pour une Thèse au Sud” (ARTS) program from Institut de Recherche pour le Développement (IRD).

## REFERENCES

- AL-Tam, F., Adam, H., Anjos, A.D., Lorieux, M., Larmande, P., Ghesquière, A., Jouannic, S. and Shahbazkia, H.R. (2013) P-TRAP: A Panicle Trait Phenotyping tool. *BMC Plant Biology*, **13**.
- Aya, K., Hobo, T., Sato-Izawa, K., Ueguchi-Tanaka, M., Kitano, H. and Matsuoka, M. (2014) A novel AP2-type transcription factor, SMALL ORGAN SIZE1, controls organ size downstream of an auxin signaling pathway. *Plant and Cell Physiology*, **55**, 897–912.
- Dehairs, J., Talebi, A., Cherifi, Y. and Swinnen, J. V. (2016) CRISP-ID: Decoding CRISPR mediated indels by Sanger sequencing. *Scientific Reports*, **6**, 1–5. Available at: <http://dx.doi.org/10.1038/srep28973>.
- Harrop, T.W.R., Mantegazza, O., Luong, A.M., Béthune, K., Lorieux, M., Jouannic, S. and Adam, H. (2019) A set of AP2-like genes is associated with inflorescence branching and architecture in domesticated rice. *Journal of Experimental Botany*.
- Harrop, T.W.R., Ud Din, I., Gregis, V., Osnato, M., Jouannic, S., Adam, H. and Kater, M.M. (2016) Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *Plant Journal*, **86**, 75–88.
- Hirano, K., Yoshida, H., Aya, K., et al. (2017) SMALL ORGAN SIZE 1 and SMALL ORGAN SIZE 2/DWARF AND LOW-TILLERING Form a Complex to Integrate Auxin and Brassinosteroid Signaling in Rice. *Molecular Plant*, **10**, 590–604. Available at: <http://dx.doi.org/10.1016/j.molp.2016.12.013>.
- Ikeda, K., Sunohara, H. and Nagato, Y. (2004) Developmental Course of Inflorescence and Spikelet in Rice. *Breeding Science*, **54**, 147–156.
- Ikeuchi, M., Ogawa, Y., Iwase, A. and Sugimoto, K. (2016) Plant regeneration: cellular origins and molecular mechanisms. *Development*, **143**, 1442–1451.
- Jofuku, K.D., Boer, B.G.W. Den, Montagu, M. Van and Okamoto, J.K. (1994) Control of arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell*, **6**, 1211–1225.
- Jung, H., Chung, P.J., Park, S.H., Redillas, M.C.F.R., Kim, Y.S., Suh, J.W. and Kim, J.K. (2017) Overexpression of OsERF48 causes regulation of OsCML16, a calmodulin-like protein gene that enhances root growth and drought tolerance. *Plant Biotechnology Journal*, **15**, 1295–1308.
- Kim, S., Soltis, P.S., Wall, K. and Soltis, D.E. (2005) Phylogeny and Domain Evolution in the APETALA2 - like Gene Family.
- Kitomi, Y., Ito, H., Hobo, T., Aya, K., Kitano, H. and Inukai, Y. (2011) The auxin responsive AP2/ERF transcription factor CROWN ROOTLESS5 is involved in crown root initiation in rice through the induction of OsRR1, a type-A response regulator of cytokinin signaling. *Plant Journal*, **67**, 472–484.
- Komatsu, M., Chujo, A., Nagato, Y., Shimamoto, K. and Kyoizuka, J. (2003) FRIZZY PANICLE is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development*, **130**, 3841–3850. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12835399>.
- Komatsu, M., Maekawa, M., Shimamoto, K. and Kyoizuka, J. (2001) The LAX1 and FRIZZY PANICLE 2 genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Developmental biology*, **231**, 364–373.
- Krizek, B.A. (2003) AINTEGUMENTA utilizes a mode of DNA recognition distinct from that used by proteins containing a single AP2 domain. *Nucleic Acids Research*, **31**, 1859–1868.
- Lee, D.Y. and An, G. (2012) Two AP2 family genes, SUPERNUMERARY BRACT (SNB) and OsINDETERMINATE SPIKELET 1 (OsIDS1), synergistically control inflorescence architecture and floral meristem establishment in rice. *Plant Journal*, **69**, 445–461.
- Li, J., Zhao, Y., Chu, H., et al. (2015) SHOEBOX Modulates Root Meristem Size in Rice through Dose-Dependent Effects of Gibberellins on Cell Elongation and Proliferation. *PLoS Genetics*, **11**, 1–21.
- Liu, W., Xie, X., Ma, X., Li, J., Chen, J. and Liu, Y.-G. (2015) DSDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations. *Molecular Plant*, **8**, 1431–1433. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1674205215002609>.
- Nakano, T., Suzuki, K., Fujimura, T. and Shinshi, H. (2006) Genome-Wide Analysis of the ERF Gene Family. *Plant Physiology*, **140**, 411–432.
- Nole-Wilson, S. (2000) DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Research*, **28**, 4076–4082.
- Qi, W., Sun, F., Wang, Q., Chen, M., Huang, Y., Feng, Y.Q., Luo, X. and Yang, J. (2011) Rice ethylene-response AP2/ERF factor OsEATB restricts internode elongation by down-regulating a gibberellin biosynthetic gene. *Plant Physiology*, **157**, 216–228. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21753115> <http://www.plantphysiol.org/content/157/1/216.full.pdf>.
- Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J. and Xu, Y. (2012) AP2/ERF transcription factor in rice: Genome-wide analysis and yntenic relationships between monocots and udicots. *Evolutionary*

- Bioinformatics*, **2012**, 321–355.
- Sabot, F., Picault, N., El-Baidouri, M., et al.** (2011) Transpositional landscape of the rice genome revealed by paired-end mapping of high-throughput re-sequencing data. *The Plant Journal*, **66**, 241–246.
- Sallaud, C., Meynard, D., Boxtel, J. Van, et al.** (2003) Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theoretical and Applied Genetics*, **106**, 1396–1408.
- Santuari, L., Sanchez-Perez, G.F., Luijten, M., et al.** (2016) The PLETHORA gene regulatory network guides growth and cell differentiation in Arabidopsis roots. *Plant Cell*, **28**, 2937–2951.
- Sharoni, A.M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., Choi, I.R., Omura, T. and Kikuchi, S.** (2011) Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant and Cell Physiology*, **52**, 344–360.
- Sunohara, H., Satoh, H. and Nagato, Y.** (2003) Mutations in panicle development affect culm elongation in rice. *Breeding Science*, **53**, 109–117.
- Xie, K., Minkenberg, B. and Yang, Y.** (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences*, **112**, 3570–3575. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1420294112>.
- Yamburenko, M. V., Kieber, J.J. and Schaller, G.E.** (2017) Dynamic patterns of expression for genes regulating cytokinin metabolism and signaling during rice inflorescence development. *PLoS ONE*, **12**, 1–18.
- Yoshida, A., Sasao, M., Yasuno, N., et al.** (2013) TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proceedings of the National Academy of Sciences*, **110**, 767–772. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1216151110>.
- Yoshida, H. and Nagato, Y.** (2011) Flower development in rice. *Journal of Experimental Botany*, **62**, 4719–4730.
- Zhou, X., Zhang, Z.L., Park, J., et al.** (2016) The ERF11 transcription factor promotes internode elongation by activating gibberellin biosynthesis and signaling. *Plant Physiology*, **171**, 2760–2770.
- Zumajo-Cardona, C. and Pabón-Mora, N.** (2016) Evolution of the APETALA2 Gene Lineage in Seed Plants. *Molecular biology and evolution*, **33**, 1818–1832.

## SUPPLEMENTAL FIGURES AND TABLES

Table S3.1. Primers used for plasmid constructions, qRT-PCRs and genotyping

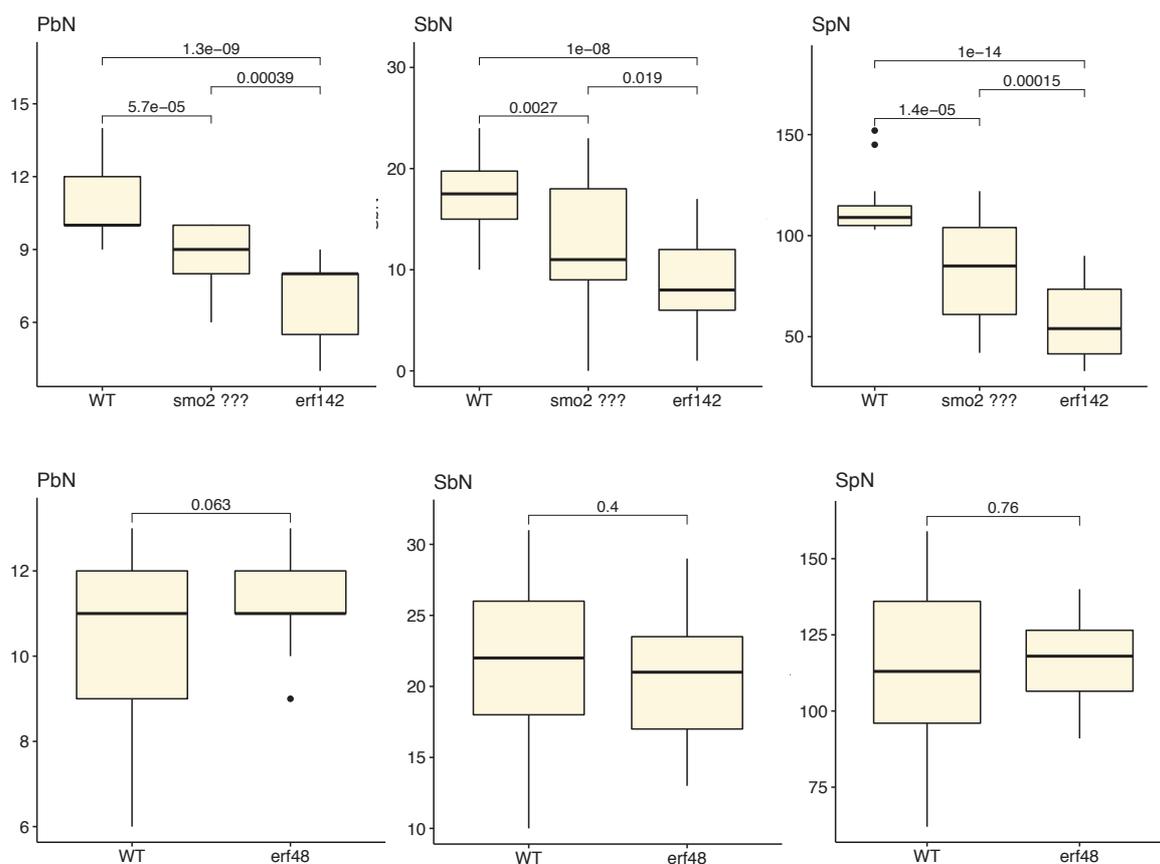
Oligo Name	Sequence (5'-3')	Purpose
HA_PLT7-Geno-F	TGGAGGTGGGCTGTTCTACA	Genotyping and sequencing of <i>PLT7</i>
HA_PLT7-Geno-R	GTGTCTGTCACTATGGCGCT	
HA_PLT7-Seq-F	GTTCTACAACCCTGCCGCC	
HA_PLT7-Seq-R	TATGGCGCTGCTAGCTACTAC	
HA_PLT8-Geno-F	GTAAACGTTCTACCGGCCA	Genotyping and sequencing of <i>PLT8</i>
HA_PLT8-Geno-R	TCGTAAGTAAAGTGCCGAGG	
HA_PLT8-Seq-F	TGAAGAAGTCTGAACATGTGTGT	
HA_PLT8-Seq-R	TAGAGGTCCTTGTTGCCGGA	
HA_PLT9-Geno-F	ACGGAGCGTTTCCATTGGTT	Genotyping and sequencing of <i>PLT9</i>
HA_PLT9-Geno-R	GTGCCCGTGACAGTAGCAG	
HA_PLT9-Seq-F	TTTATGCAATACAGGCACTCAAG	
HA_PLT9-Seq-R	CTCGCTTCCCTCGGTCCG	
HA_AP22-Geno-F	CTTCAGTTTGTGCCAAGGCT	Genotyping and sequencing of <i>AP2-EREBP22</i>
HA_AP22-Geno-R	CCCTAATATGCGATGCGGCT	
HA_AP22-Seq-F	CTCTGCTAAACCATGCCCT	
HA_AP22-Seq-R	CTTCAAGCCTGCCCTTCTAT	
HA_Hygro-F	GCTCCAGTCAATGACCGCTG	Hygromycin phosphotransferase II detection
HA_Hygro-R	CTCGGAGGGCGAAGAATCTC	
HA_UBI-F	GCTTGTGCGTTTCGATTTGA	Detection of <i>Cas9</i>
HA_Cas9-R	CCGCTCGTGCTTCTTATCCT	
HA_control-F	CGCTGCCACTCTCCACTGA	Control for DNA quality
HA_control-R	AGCTGCTTCCACTCGTTCCA	
HA_UGW-U3-F	GACCATGATTACGCCAAGCTTAAGGAATC TTTAAACATACG	Detection of U3:PTG
HA_UGW-gRNA-R	GGACCTGCAGGCATGCACGCGCTAAAAAC GGACTAGC	
L5AD5-F	CGGGTCTCAGGCAGGATGGGCAGTCTGGG <u>CA</u> ACAAAGCACCAAGTGG	PTG synthesis and cloning <i>FokI</i> site ( <u>underlined</u> ) was used to generate compatible overhangs (labeled with red color) for cloning into pRGEB32
L3AD5-R	TAGGTCTCCAAACGGATGAGCGACAGCAA <u>AC</u> AAAAAAAAAAAA GCACCGACTCG	
S5AD5-F	CGGGTCTCAGGCAGGATGGGCAGTCTGGG <u>CA</u>	
S3AD5-R	TAGGTCTCCAAACGGATGAGCGACAGCAA <u>AC</u>	

**Table S3.2. Primers used for *in situ* hybridization probes**

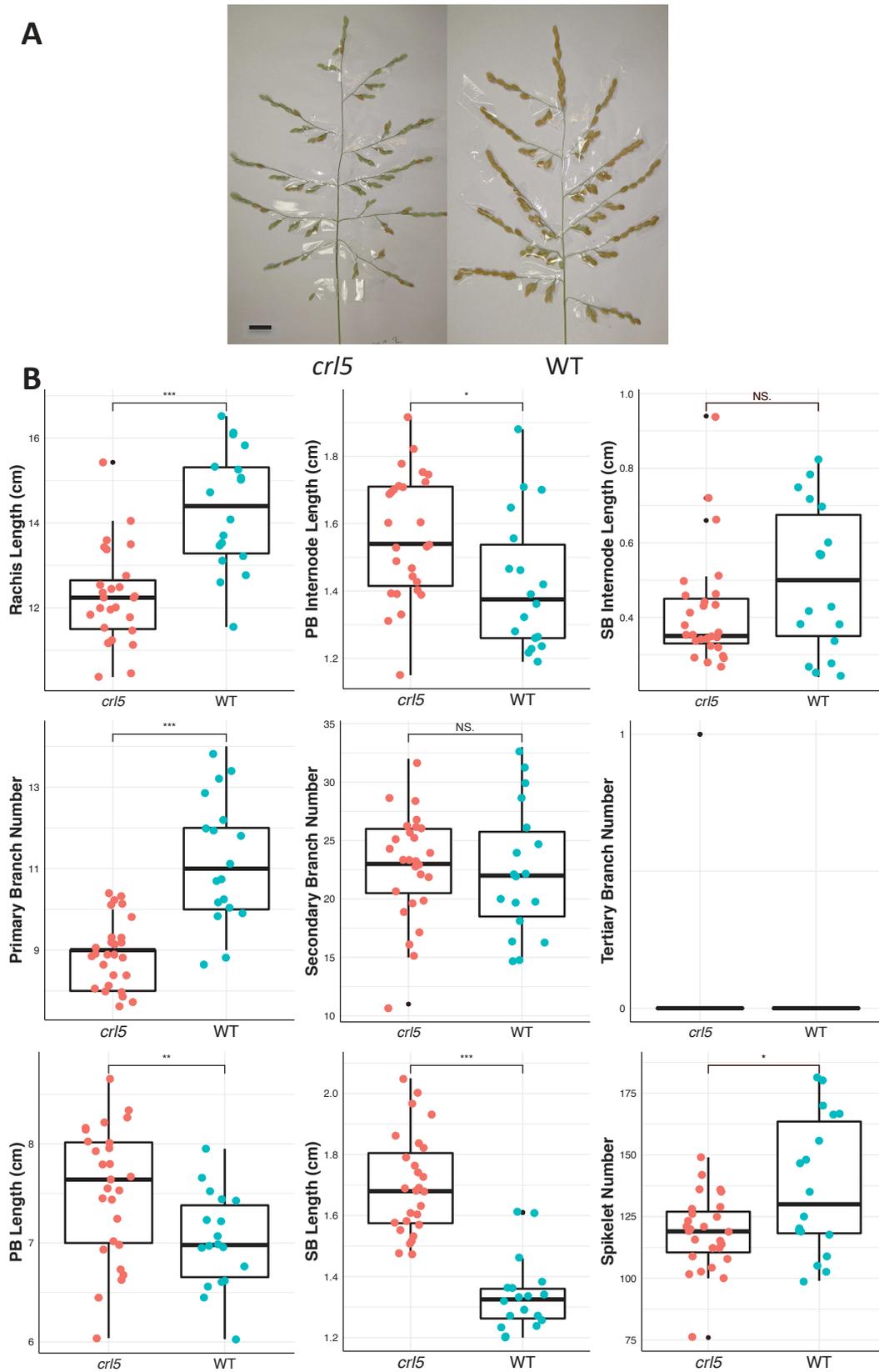
Oligo Name	Sequence (5'-3')
HA-CRL5-insitu-F	ACGACGTGGCGGCGATCAAG
HA-CRL5-insitu-R	TTAGGCGTCGGTCCAGGCGG
HA-CRL5-insitu-T7-F	GCGAAATTAATACGACTCACTATAGGGCGAAACGACGTGGCGGCGATCAAG
HA-CRL5-insitu-T7-R	GCGAAATTAATACGACTCACTATAGGGCGAATTAGGCGTCGGTCCAGGCGG
HA-PLT8-pro-2F	ATTCTCCGTCCTCCACGACCTC
HA-PLT8-T7-2F	GCGAAATTAATACGACTCACTATAGGGCGAAATTCTCCGTCCTCCACGACCTC
HA-PLT7-new-pro-F	ACCGCCGGAGGTGTTCG
HA-PLT7-new-pro-R	TTGCTGTTCTTGGTCGCCTC
HA-PLT7-new-pro-T7-F	GCGAAATTAATACGACTCACTATAGGGCGAAACCGCCGGAGGTGTTCG
HA-PLT7-new-pro-T7-R	GCGAAATTAATACGACTCACTATAGGGCGAATTGCTGTTCTTGGTCGCCTC
HA-AP22-new-pro-F	ACCACGAAGGGAGTTGAGTC
HA-AP22-new-pro-R	GCCGTAGTTCAGCAGTACC
HA-AP22-new-pro-T7-F	GCGAAATTAATACGACTCACTATAGGGCGAAACCACGAAGGGAGTTGAGTC
HA-AP22-new-pro-T7-R	GCGAAATTAATACGACTCACTATAGGGCGAAGCCGTAGTTCAGCAGTACC
HA_PLT9-Sens	TATAGCCAGGGCCAAGAAGC
HA_PLT9-ASens	GCTGGTGACGAGGAAGTTCT
HA_PLT9-pT7sens	GCGAAATTAATACGACTCACTATAGGGCGAATATAGCCAGGGCCAAGAAGC
HA_PLT9-pT7AS	GCGAAATTAATACGACTCACTATAGGGCGAAGCTGGTGACGAGGAAGTTCT
HA-T7-HIS	GCGAAATTAATACGACTCAC



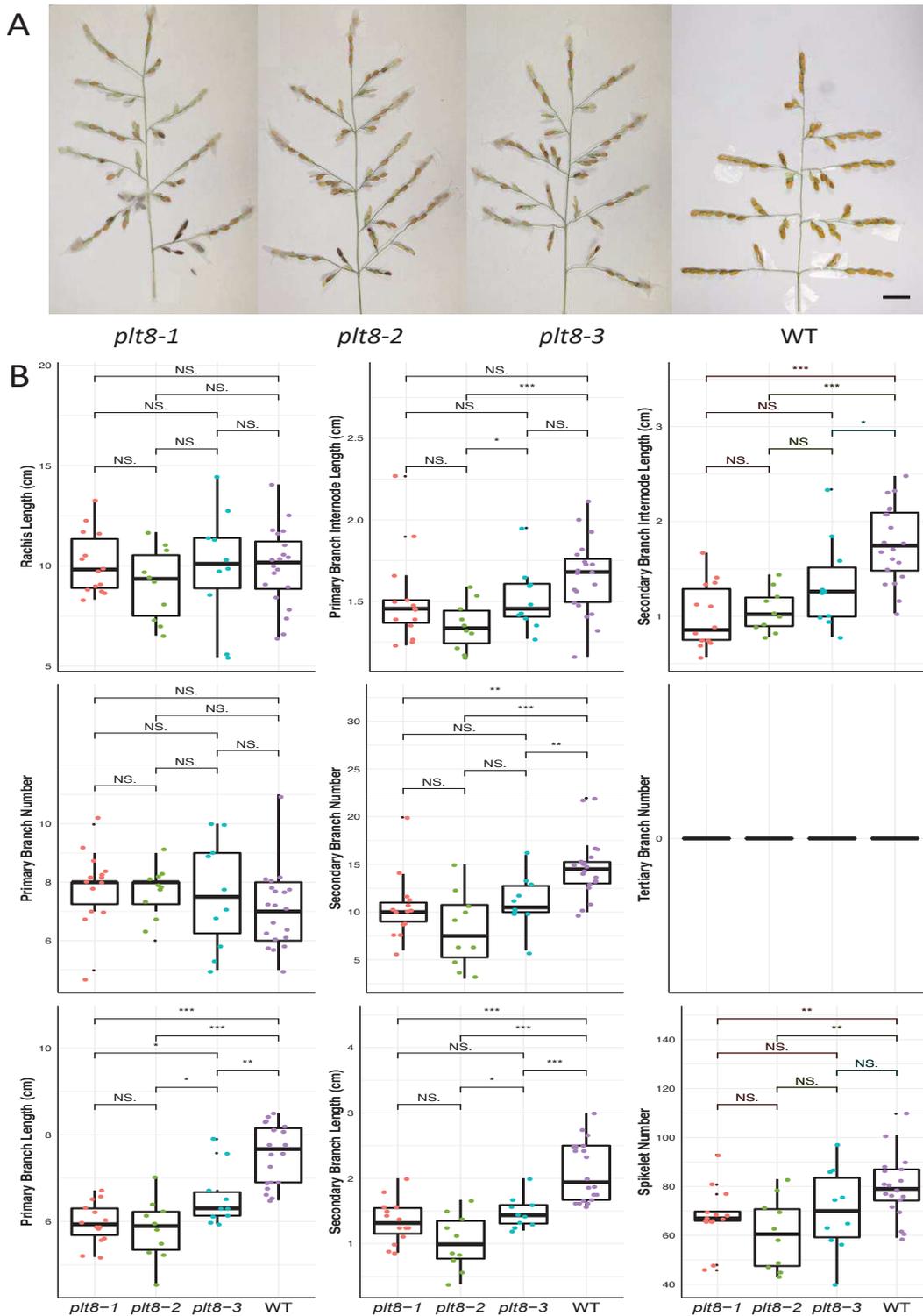
**Figure S3.1.** Spread mature rice panicle. PB: Primary branch; PBIL: Primary branch internode length; PBL: Primary branch length; RL: Rachis length; SB: Secondary branch; SBIL: Secondary branch internode length; SBL: Secondary branch length; Sp: Spikelet. (Harrop *et al.*, 2019)



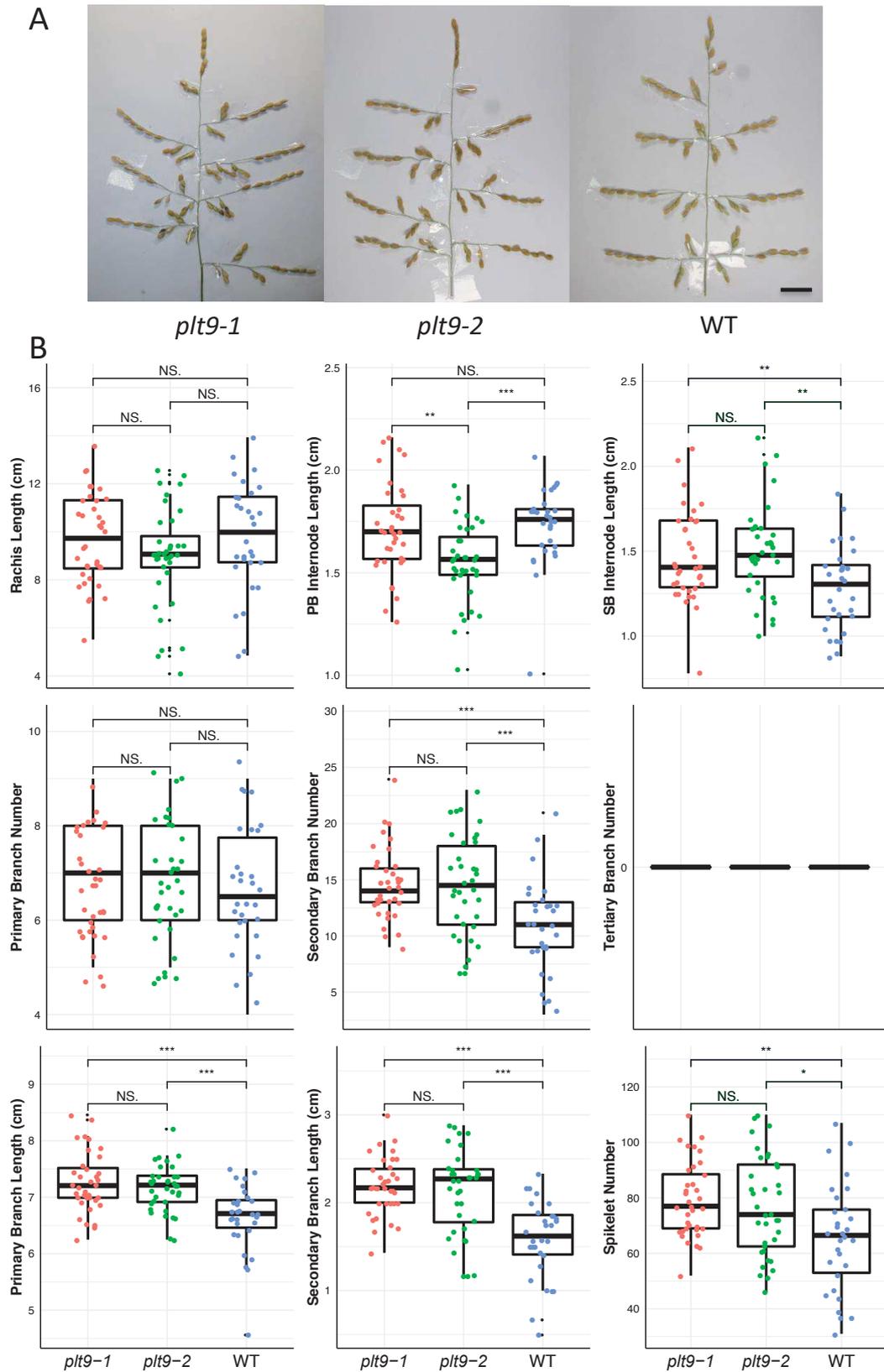
**Figure S3.2. Panicle phenotype of *erf142* and *erf48* mutants.** PbN, Primary branch Number, SbN, Secondary branch Number, SpN, Spikelet Number. Panicle trait scoring was carried out using P-TRAP software (AL Tam et al., 2013)



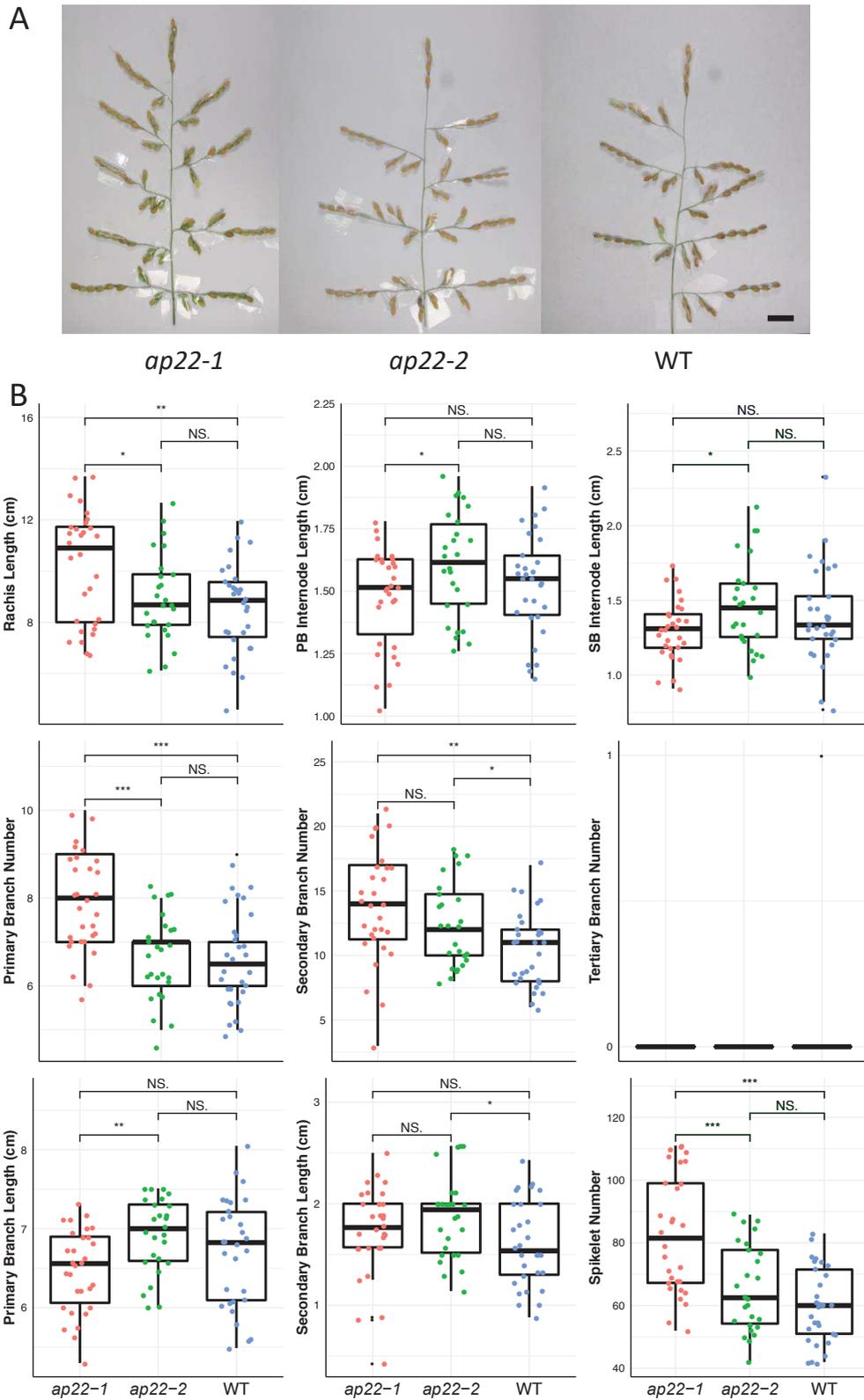
**Figure S3.3. Panicle phenotype of *crl5* mutant.** **A**, Panicle of seventy-five-day-old *crl5* (left) and wild-type (WT) Kinmaze (right) plant. Scale bar, 2 cm. **B**, Comparison of panicle phenotype between *crl5* and WT. PB, Primary branch; SB, Secondary branch. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.



**Figure S3.4. Panicle phenotype of edited mutants of *OsPLT8*.** **A**, Panicle of seventy-five-day-old *plt8* and wild-type (WT) (*O. sativa ssp. japonica* Kitaake) plants. Scale bar, 2 cm. **B**, Comparison of panicle phenotype between *plt8* and WT plants. PB, Primary branch; SB, Secondary branch. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.



**Figure S3.5. Panicle phenotype of edited mutants of *OsPLT9*.** **A**, Panicle of seventy-five-day-old *plt9* and wild-type (WT) (sister-line) plants. Scale bar, 2 cm. **B**, Comparison of panicle phenotype between *plt9* and WT plants. PB, Primary branch; SB, Secondary branch. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.



**Figure S3.6. Panicle phenotype of edited mutants of *OsAP2/EREBP22*.** **A**, Panicle of seventy-five-day-old *ap22* and wild-type (WT) (sister-line) plants. Scale bar, 2 cm. **B**, Comparison of panicle phenotype between *ap22* and WT plants. PB, Primary branch; SB, Secondary branch. Student t-test, significance \* P value  $\leq 0.05$ , \*\* P value  $\leq 0.01$ , \*\*\* P value  $\leq 0.001$ , NS non-significant.

# **Chapter IV**

## **GENERAL DISCUSSION AND PERSPECTIVES**

The overall objective of this PhD project was to identify and characterize *AP2/ERF* genes involved in the development of the panicle and the determination of its structural diversity in the genus *Oryza*, then to evaluate whether these genes were impacted by the two independent domestications that occurred in this genus. Most of the work was focused on genes encoding *PLETHORA (PLT)* transcription factors, a small group within the *AP2/ERF* superfamily. We first showed that *AP2/ERF* genes were key factors in inflorescence branching and rice domestication. Subsequently, we carried out the characterization and detailed expression analysis of rice *PLT* genes in different rice species, providing insights for the elucidation of their biological functions. Lastly, we generated CRISPR-Cas9 edited mutants of *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* to study their involvement in rice panicle development and architectural diversity. Through the phenotyping of panicle traits in mutants, we gathered a strong body of evidence supporting the hypothesis that these *OsPLT* genes function in panicle architecture establishment.

In each chapter, I have discussed in detail the relevance of our findings in the context of current knowledge. Here, I make a summary of key findings and provide some broader perspectives.

### ***PLETHORA* genes control meristem development**

In *A. thaliana*, the *PLETHORA1–3 (PLT1–3)* and *BABYBOOM (BBM/PLT4)* genes have been reported as master regulators of root meristem initiation and maintenance (Horstman *et al.*, 2014; Aida *et al.*, 2004; Galinha *et al.*, 2007a). Root stem cells are maintained by a small group of slowly dividing organizer cells collectively called the quiescent center (QC). The *plt1/plt2* double mutants display a significant decrease in root meristem size and depletion of QC markers (Aida *et al.*, 2004). Over-expression of *PLT* genes induces the accumulation of stem cells in the root meristem and can cause the production of ectopic roots from the shoot apex (Galinha *et al.*, 2007). Moreover, *PLT* proteins accumulate in a spatially differential fashion in the root apex to form a gradient that guides the transition of stem cells towards differentiation (Galinha *et al.*, 2007a; Mähönen *et al.*, 2014a; Santuari *et al.*, 2016). *ANT*, *AIL6/PLT3*, and *AIL7/PLT7* are expressed throughout the meristematic regions of the inflorescence and vegetative shoot. The shoot of *ant/ail6/plt7* triple mutant stops developing after the production of a few leaves due to reduced cell divisions in the meristem and differentiation of the meristematic cells, indicating that *ANT*, *AIL6/PLT3*, and *AIL7/PLT7* genes are required for shoot apical meristem (SAM) maintenance (Horstman *et al.*, 2014).

Within the shoot meristem, AIL/PLT proteins promote organ development and repress differentiation (Horstman *et al.*, 2014). *AIL6/PLT3*, *AIL5/PLT5* and *AIL7/PLT7* genes are among the earliest transcriptional responses induced during *de novo* shoot and root formation from callus in *A. thaliana* and the *plt3 plt5 plt7* triple mutants are defective in shoot regeneration (Ikeuchi *et al.*, 2016; Kareem *et al.*, 2015). *AIL6/PLT3* and *PLT7* are expressed throughout the meristem, but their expression is increased in the peripheral zones and in the central zone, respectively. By contrast, ANT is expressed exclusively in the peripheral zones and marks the cryptic bract region of the floral meristem. 20 years of *PLT* research in *Arabidopsis* has shown that these transcription factors are crucial for stem cell niche specification, meristem maintenance, organ initiation, and growth (Horstman *et al.*, 2014). Moreover, in *Populus spp.*, *PLT* genes have been shown to control meristem activity during adventitious rooting (Rigal *et al.*, 2012). Similarly, in the moss *Physcomitrella patens*, four *PLT* genes are crucial for the formation of apical gametophore stem cells (Aoyama *et al.*, 2012).

In rice, little was known until recently about the function of *PLT* genes in meristem functioning. Recently, in research carried out by Khanday *et al.* (2019), triple mutants were generated for *BBM1-3* (*OsPLT6*, *OsPLT5*, *OsPLT3*), allowing a demonstration that male genome-derived expression of *BBM1/OsPLT6* triggers the embryonic program in the fertilized egg cell and that this gene acts redundantly with other *BBM* genes. Additionally, *BBM-like* genes promote regeneration from tissue culture, suggesting that they function as pluripotency factors (Lowe *et al.*, 2016). In earlier studies, expression profiling of *OsPLT1-OsPLT6* in rice roots indicated that these genes are all expressed in the primodium of crown root, and mostly in the initial cells adjacent to QC of primary, crown, and lateral roots (Li and Xue, 2011). Moreover, *in situ* hybridization experiments confirmed that *OsPLT1* and *OsPLT2* are expressed in the stele initial cells of the meristemic zone, and that *OsPLT3* and *OsPLT5* are highly expressed in QC and surrounding initial cells (Li and Xue, 2011). *OsPLT8/CRL5* was also shown to be induced by auxin and involved in the *de novo* pathway to initiate crown roots in rice, as well as being highly expressed in the early stages of inflorescence development (Kitomi *et al.*, 2011). Additionally, the *plt8* mutants also displayed a deficient root system, confirming the involvement of *OsPLT8* in crown root initiation as observed for *crl5* (Kitomi *et al.*, 2011).

In the first stage of my PhD project, exhaustive *in silico* analysis of the AP2-subfamily was performed and twelve members of the PLT group were identified in rice. Six *OsPLT* genes (*OsPLT4*, *OsPLT7*, *OsPLT8*, *OsPLT9*, *OsPLT10*, *AP2/EREBP22* and

*AP2/EREBP86*) were identified in rice panicle meristems. The genes *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* were found to be differentially expressed through different stages of panicle development and between different rice species, suggesting that they might play important roles in panicle meristem development and be of significance in the context of domestication. RNA *in situ* hybridization revealed that *OsPLT9* was highly expressed not only in the axillary branch and spikelet meristems during early panicle development but also in vegetative meristems generating new tillers. Taken together, our results provide strong evidence that *PLT* genes are involved in regulating meristem activity during development and that they display structural and functional conservations between different species of flowering plants.

### ***OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* may have played a role in panicle architecture evolution**

*OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* transcripts were all detected in the panicle by RNA-seq with differential patterns in the meristem-specific study (Harrop *et al.*, 2016). They were also detected in the panicle branching meristem in the multiple-species RNA-seq dataset (Harrop *et al.*, 2019). During my PhD, further expression profiling of these genes was carried out at other stages of panicle development and in different rice species. The results obtained added weight to the hypothesis of a potential role for these genes in the determination of panicle architecture in the context of domestication. Subsequently, the CRISPR-Cas9 edited mutants of *OsPLT7*, *OsPLT8*, *OsPLT9* and *AP2/EREBP22* in *Oryza sativa ssp. japonica* cv. Kitaake were generated and panicle phenotyping of available mutants were performed, revealing that mutations in *OsPLT8*, *OsPLT9* and *AP2/EREBP22* provoke changes in panicle traits. More specifically, the *plt8* mutants produce fewer secondary branches than wild type while the *plt9* mutants produce more secondary branches and the *ap22/erebp22* mutants produce more branches of both the primary and secondary types.

However, there are still several questions about the roles of these *PLT* genes in panicle development that remain to be addressed. The first point is that our study needs to be completed by further functional and genetic analyses, in order to determine more precisely whether alterations in *PLT* gene expression lead to the differences in panicle architecture observed between different rice ecotypes and species. In the case of inter-specific panicle architecture variations, the development and use of genetic transformation techniques for other species of *Oryza* will clearly be important. Moreover, from a functional standpoint it will be also be important to identify which other factors interact with *PLT* genes (e.g.

hormones, other genes, etc.) and how they are coordinated during rice panicle development. Below I propose some priorities for the short term and long term to address these questions.

### **The short/mid-term perspectives**

Firstly, in order to finalize the genotyping of CRISPR-Cas9 edited mutants, of which the single *plt7* and double *plt7plt8* mutants still remain to be isolated, I will design new primers for genotyping and sequencing. It must be remembered that the region of *OsPLT7* targeted for mutation by gRNA (nearly 500bp) was much larger than those of *OsPLT8*, *OsPLT9* and *AP2/EREBP22* (around 200 bp). Moreover, the strong sequence conservations between these genes create difficulties for designing specific primers. Each pair of primers designed for the present study was tested by PCR and sequencing in order to confirm specificity. I have already tested three different pairs of primers for amplifying the region targeted for mutation in the *OsPLT7* gene. However, due to the low efficiency of the designed primers, the genotyping of *OsPLT7*-targeted plants was postponed for both the single and double mutant lines.

Secondly, expression profiling of the different *PLT* genes in the available CRISPR-Cas9 mutants could be carried out, as a first step in young seedlings and as a second step in the panicle at early developmental stages. This would allow us to confirm whether these mutants are knockout or not, and to evaluate whether there are interactions between the different genes in terms of their transcriptional regulation. Moreover, expression profiling of these genes alongside other developmental landmark genes involved in the initiation and/or maintenance of lateral meristems could provide a means to identify other candidate genes that might be involved in the same signalling cascade(s). Examples of developmental landmark genes include *Oryza sativa homeobox1 (OSH1)*, which is associated with meristematic cell fate control in angiosperms (Tsuda and Hake, 2015). Similarly the *LAX PANICLE1 (LAX1)* and *SQUAMOSA promoter binding protein-like14 (SPL14)* genes were found to be involved in axillary meristem establishment and outgrowth during *O. sativa* panicle development (Miura *et al.*, 2010; Komatsu *et al.*, 2003). A second group of marker genes associated with meristem fate control includes the *SEPALLATA-like* gene *LEAFY HULL STERILE1/OsMADS1 (LHS1)*, which promotes the transition from branch meristems to spikelet meristems (Jeon *et al.*, 2000; Khanday *et al.*, 2013), as well as the *LEAFY* ortholog *ABERRANT PANICLE ORGANIZATION2 (APO2)* and *TAWAWAI (TAW1)*, both reported to act as suppressors of the transition from branch meristems to spikelet meristems (Ikeda-Kawakatsu *et al.*, 2012; Yoshida *et al.*, 2013). Analysis of the *in situ* expression patterns of

*OsPLT7*, *OsPLT8* and *OsAP2/EREBP22* in *O. sativa ssp. japonica Nipponbare* will be performed again with new designed probes, so as to describe the spatial patterns of expression of these genes in the panicle meristem. Moreover, the *in situ* expression patterns of other developmental landmark genes in selected mutant backgrounds could provide more information concerning the regulatory networks in which the *OsPLTs* participate during rice development.

It will be important to characterise, at the histological level, any possible differences between the mutant and wild-type plants at very early stages of panicle development at the time of branching establishment, in order to establish a link with the observed phenotype of the mature stage panicle. Additionally, studies of meristem size and flower structure in the available mutants could be carried out using electron-microscopy. In order to obtain a wider view of *OsPLT* gene function, a simple experiment can be carried out to measure the root development of young rice seedlings grown in transparent medium *in vitro*, in order to examine any possible effects of mutations on crown root initiation and/or elongation as already reported for the *OsPLT8/CRL5* gene (Kitomi *et al.*, 2011).

Lastly, in parallel with the deletion mutants, over-expressing *OsPLT* mutants were also generated during my PhD. The seeds of T0 generation plants were collected and sown to obtain the next generation. The T1 over-expressing mutants will be tested by genotyping (for T-DNA presence) and processed by qRT-PCR in order to select lines with different expression levels of the transgenes. Subsequently, the phenotyping of these mutant lines will be carried out. By combining these data with the phenotypes of CRISPR-Cas9 mutants, it should be possible to obtain a deeper insight into the functions of the *PLT* genes in rice panicle development.

### **The long-term perspectives**

Because of the time limitations of a PhD project, it was not possible to analyze other factors that might regulate the activity of *OsPLT* genes in rice panicle development. For instance, auxin plays a key role in regulating the formation, activity and fate of meristems, thereby shaping plant architecture (Gallavotti, 2013). Prolonged high auxin levels were found to be required to promote PLT activities; moreover positive feedback from PLT to auxin biosynthesis and transport plays a role during the generation of new primordia (Mähönen *et al.*, 2014). Although a few genes have been identified as regulators of auxin transport in *O. sativa* (e.g. *OsPIN1*, *LAX*), the exact role of auxin and its mode of action during the early stages of panicle development in rice remain to be elucidated (Morita and Kyojuka 2007;

McSteen 2009). It will be of great interest to analyze auxin marker lines in *O. sativa* such as DR5-GUS. DR5 is a synthetic promoter allowing the visualisation of auxin response. Moreover, we can also apply the complementary types of auxin biosensor constructs (*DR5-VENUS*), one derived from the Aux/IAA-based biosensor *DII-VENUS* but constitutively driven by maize ubiquitin-1 promoter, in which a synthetic auxin-responsive promoter (*DR5<sub>rev</sub>*) was used to drive expression of the yellow fluorescent protein (YFP) (Yang *et al.*, 2017). The use of such a construction in both wild-type and mutant backgrounds should provide information on relationship between *OsPLT* genes and auxin signalling. Besides, auxin-induced *CRL5/OsPLT8* promotes crown root initiation through repression of cytokinin signaling by positively regulating type-A RR, OsRR1 (Kitomi *et al.*, 2011). Cytokinins were shown as playing a role in rice panicle development, as reported by (Ashikari *et al.*, 2005), through the characterization of the *Gn1a* QTL related to the cytokinin oxidase/dehydrogenase gene *OsCKX2*. In this sense, it will be interesting to analyze cytokinin-markers lines in *O. sativa* in both wild-type and mutant backgrounds.

On a related note, the *in situ* expression patterns of *OsPLT7*, *OsPLT8*, *OsPLT9*, *OsAP2/EREBP22* in panicle meristem of other rice species should be determined, depending on the availability of samples of these species. If genetic transformation of *O. glaberrima* (domesticated African rice) can be successfully and efficiently optimized, knock-out or over-expressing mutants could be obtained for this species as for Asian rice. This will allow comparisons to be made regarding the importance of *OsPLT* genes in determining panicle phenotype in the two domesticated species.

In the context of the intra-specific diversity of panicle structure, an analysis of the expression patterns of the *OsPLT* genes could be carried out using the panel (or a sub-panel) of *O. sativa* Vietnamese varieties developed by LMI RICE and AGI (Phung *et al.*, 2014) that has been subjected to intensive panicle phenotyping over the last few years in Vietnam for genome wide association studies (Ta *et al.*, 2018). It is envisaged that *PLT* group gene expression will be studied by qRT-PCR in accessions that display contrasting panicle architecture complexity. This approach should make it possible to establish a link between intra-specific panicle architecture diversity and the expression patterns of the *OsPLT* genes.

## REFERENCES

- Aida, M., Beis, D., Heidstra, R., et al.** (2004) The PLETHORA genes mediate patterning of the Arabidopsis root stem cell niche. *Cell*, **119**, 109–120.
- Aoyama, T., Hiwatashi, Y., Shigyo, M., Kofuji, R., Kubo, M., Ito, M. and Hasebe, M.** (2012) AP2-type transcription factors determine stem cell identity in the moss *Physcomitrella patens*. *Development (Cambridge)*, **139**, 3120–3129.
- Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R. and Scheres, B.** (2007a) PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. *Nature*, **449**, 1053–1057.
- Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R. and Scheres, B.** (2007b) PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. *Nature*, **449**, 1053–1057. Available at: <https://doi.org/10.1038/nature06206>.
- Gallavotti, A.** (2013) The role of auxin in shaping shoot architecture. *Journal of Experimental Botany*, **64**, 2593–2608.
- Harrop, T.W.R., Mantegazza, O., Luong, A.M., Béthune, K., Lorieux, M., Jouannic, S. and Adam, H.** (2019) A set of AP2-like genes is associated with inflorescence branching and architecture in domesticated rice. *Journal of Experimental Botany*.
- Harrop, T.W.R., Ud Din, I., Gregis, V., Osnato, M., Jouannic, S., Adam, H. and Kater, M.M.** (2016) Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *Plant Journal*, **86**, 75–88.
- Horstman, A., Willemsen, V., Boutilier, K. and Heidstra, R.** (2014) AINTEGUMENTA-LIKE proteins: Hubs in a plethora of networks. *Trends in Plant Science*, **19**, 146–157. Available at: <http://dx.doi.org/10.1016/j.tplants.2013.10.010>.
- Ikeda-Kawakatsu, K., Maekawa, M., Izawa, T., Itoh, J.I. and Nagato, Y.** (2012) ABERRANT PANICLE ORGANIZATION 2/RFL, the rice ortholog of Arabidopsis LEAFY, suppresses the transition from inflorescence meristem to floral meristem through interaction with APO1. *Plant Journal*, **69**, 168–180.
- Ikeuchi, M., Ogawa, Y., Iwase, A. and Sugimoto, K.** (2016) Plant regeneration: Cellular origins and molecular mechanisms. *Development (Cambridge)*, **143**, 1442–1451.
- Jeon, J.S., Jang, S., Lee, S., et al.** (2000) leafy hull sterile 1 is a homeotic mutation in a rice MADS box gene affecting rice flower development. *Plant Cell*, **12**, 871–884.
- Kareem, A., Durgaprasad, K., Sugimoto, K., et al.** (2015) PLETHORA genes control regeneration by a two-step mechanism. *Current Biology*, **25**, 1017–1030. Available at: <http://dx.doi.org/10.1016/j.cub.2015.02.022>.
- Khanday, I., Ram Yadav, S. and Vijayraghavan, U.** (2013) Rice LHS1/UsMADS1 controls floret meristem specification by coordinated regulation of transcription factors and hormone signaling pathways. *Plant Physiology*, **161**, 1970–1983.
- Khanday, I., Skinner, D., Yang, B., Mercier, R. and Sundaresan, V.** (2019) A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature*, **565**, 91–95. Available at: <http://dx.doi.org/10.1038/s41586-018-0785-8>.
- Komatsu, K., Maekawa, M., Ujiie, S., Satake, Y., Furutani, I., Okamoto, H., Shimamoto, K. and Kyojuka, J.** (2003) LAX and SPA: Major regulators of shoot branching in rice. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 11765–11770.
- Li, P. and Xue, H.** (2011) Structural characterization and expression pattern analysis of the rice PLT gene family. *Acta Biochimica et Biophysica Hungarica*, **43**, 688–697. Available at: <http://abbs.oxfordjournals.org/cgi/doi/10.1093/abbs/gmr068>.
- Lowe, K., Wu, E., Wang, N., et al.** (2016) Morphogenic regulators Baby boom and Wuschel improve monocot transformation. *Plant Cell*, **28**, 1998–2015.
- Mähönen, A.P., Tusscher, K. Ten, Siligato, R., et al.** (2014a) PLETHORA gradient formation mechanism separates auxin responses. *Nature*, **515**, 125–129.
- Mähönen, A.P., Tusscher, K. Ten, Siligato, R., et al.** (2014b) PLETHORA gradient formation mechanism separates auxin responses. *Nature*, **515**, 125–129.
- Miura, K., Ikeda, M., Matsubara, A., Song, X.-J., Ito, M., Asano, K., Matsuoka, M., Kitano, H. and Ashikari, M.** (2010) OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nature Genetics*, **42**, 545–549. Available at: <https://doi.org/10.1038/ng.592>.
- Phung, N.T.P., Mai, C.D., Mournet, P., et al.** (2014) Characterization of a panel of Vietnamese rice varieties using DArT and SNP markers for association mapping purposes. *BMC Plant Biology*, **14**, 371. Available at: <http://bmcplantbiol.biomedcentral.com/articles/10.1186/s12870-014-0371-7>.
- Rigal, A., Yordanov, Y.S., Perrone, I., et al.** (2012) The AINTEGUMENTA LIKE1 Homeotic Transcription Factor PtAIL1 Controls the Formation of Adventitious Root Primordia in Poplar. *Plant Physiology*, **160**, 1996–2006.

- Santuari, L., Sanchez-Perez, G.F., Luijten, M., et al.** (2016) The PLETHORA gene regulatory network guides growth and cell differentiation in Arabidopsis roots. *Plant Cell*, **28**, 2937–2951.
- Tsuda, K. and Hake, S.** (2015) Diverse functions of KNOX transcription factors in the diploid body plan of plants. *Current Opinion in Plant Biology*, **27**, 91–96. Available at: <http://dx.doi.org/10.1016/j.pbi.2015.06.015>.
- Yoshida, A., Sasao, M., Yasuno, N., et al.** (2013) TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proceedings of the National Academy of Sciences*, **110**, 767–772. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1216151110>.

# CHAPTER V

## MATERIALS & METHODS

The materials and methods was briefly presented in each chapter. Here is more details about each experiment that I carried out during my PhD.

### 1. Chemicals and kits

All molecular biology grade chemicals and organic solvents were purchased from Fluka, Sigma-Aldrich®, Heraeus Kulzer (Germany), Duchefa (Netherland), BIO-RAD, Labonord (France) and Carlo Ebra (Italy). Kits for DNA and RNA extraction were purchased from Qiagen (France). SuperScript III cDNA First-strand synthesis system and restriction endonucleases were purchased from Invitrogen (USA), and Promega (USA). SYBRGreen I kit was provided by Roche (France). GoTaq DNA polymerase was purchased from Promega (USA). Phusion High-fidelity Polymerase was purchased from ThermoFisher Scientific (USA). PrimeSTAR HS DNA Polymerase was purchase from TaKaRa Bio (Japan). All the enzymes were supplied and used with their buffers. Biomark highthroughput qPCR kit was purchased from Fluidigm (USA). The pGEM®-T Easy Vector cloning systems (Promega), which allows to direct ligation of PCR-amplified fragments without enzyme treatment, was used for the cloning of PCR products according to manufacturer's instructions. The pGTR plasmid and the pRGEB32 binary vector, which were used to generate Knock-Out mutant lines, were provided by Dr. Hoa Le (Hanoi University of Science and Technology- Vietnam, purchased from AddGene - USA). *E. coli* competent TOP10 was purchased from ThermoFisher Scientific (USA), *Agrobacterium tumefaciens* EHA105 (UMR IPME, IRD Montpellier, France). DNA ladder 100bp and 1kb were purchased from Promega (USA).

### 2. Plant materials and growth condition

For *in situ* hybridization samples and Fluidigm qPCR experiments, five accessions were used: *O. sativa ssp. japonica* cv. Nipponbare, *O. sativa ssp. indica* cv. IR64, *O. rufipogon* acc. W1654, *O. glaberrima* acc. Tog5681 and *O. barthii* acc. B88. For collecting samples for *in situ* hybridization, panicle meristems were collected from 15 plants of each accession. Plants were grown in green house at IRD Montpellier (France), under long day conditions (14h light/10h dark) at 28°C-30°C, and humidity at 60%. After 6 to 8 weeks, flowering induction is under the cycle of 10h of light. Panicles were collected at 4 different stages: stage 1, rachis and primary branch meristem; stage 2, elongated primary branch and secondary meristems; stage 3, spikelet differentiation; stage 4, young flowers with differentiates organs.

The *crl5*, *smos1-3* and *erf48* mutants and relative wild type genetic background kindly provided by..... (Kitomi *et al.*, 2011; Jung *et al.*, 2017; Aya *et al.*, 2014) were grown in a greenhouse in Montpellier, France, in October 2017 under short day conditions (11h light/13h dark) and humidity at 60%.

The mutant lines generated from *Oryza sativa ssp. japonica* cv. *kitaake* were grown in greenhouse in May 2018 (knock-out mutant, T0 generation), in October 2018 (knock-out

mutant, T1 generation) and in March 2019 (knock-out mutant, T2 generation and over-expressing mutant, T0 generation) under long day conditions (14h light/10h dark) at 28°C-30°C, and humidity at 60%.

### **3. *In Silico* analysis of AP2/ERF family**

#### **3.1. Identification of AP2/ERF genes in rice genome**

Firstly, rice *AP2/ERF* genes were identified in the genome of *Oryza sativa ssp. japonica* cv. Nipponbare using ESTs and cDNA sequences. The data were downloaded from various public repositories, including National Centre for Biotechnology Information (NCBI), Database of Rice Transcription Factors (DRTF), MSU Rice Genome Annotation Project Database and Plant Genome Database (PlantGDB). Next, all retrieved sequences were subjected to the BLAST on Gramene (<http://www.gramene.org>) and Rice Genome Annotation Project (Kawahara *et al.*, 2013) to find homologous sequences in the rice genome. Moreover, the results were compared to the data already published (Rashid *et al.*, 2012; Sharoni *et al.*, 2011; Nakano *et al.*, 2006). In addition, Simple Modular Architecture Research Tool (SMART) was used to confirm the presence of the AP2/ERF domain in the resulting sequences.

#### **3.2. Phylogenetic and MEME motif analysis**

The AP2/ERF domain-comprising protein sequences obtained from various sources were aligned using MEGA 7.0 software (Kumar *et al.*, 2016) and redundant entries were removed. A combined un-rooted neighbor-joining (NJ) tree was generated in MEGA 7.0 with the following default parameters: p-distance correction, pairwise deletion and bootstrap (1000 replicates). The visual phylogeny tree was built by using EvolView version 2.0 (He *et al.*, 2016). To obtain more informations about protein sequences, conserved motifs in rice AP2/ERF protein sequences were identified using a motif-based sequence analysis tool, MEME Suite version 5.0.5, with following parameters: optimum width 6–200 amino acids, any number of repetitions of a motif, and maximum number of motifs set at 12.

#### **3.3. *In silico* gene expression analysis**

To further investigate *AP2/ERF* gene expression, the rice expression profile database RiceXPro (Sato *et al.*, 2013) which is a repository of gene expression data derived from microarray experiments encompassing the entire life cycle of the rice plant from germination, seedling, tillering, stem elongation, panicle initiation, booting, heading, flowering, and ripening stage, was used. This tool generates a heat map of normalized signal intensity values for each plant tissue for each gene and provides a quantitative measure of the transcript of particular genes. In addition, I used the Rice Expression Database (RED), a sub-project of

IC4R (Information Commons for Rice; <http://ic4r.org>), which integrates expression profiles derived entirely from NGS RNA-Seq data of *Oryza sativa ssp. japonica* cv. Nipponbare and Rice eFP Browser ((Waese *et al.*, 2017); [http://bar.utoronto.ca/eplant\\_rice/](http://bar.utoronto.ca/eplant_rice/)) - a robust research platform that incorporates all data and provides integrated search, analysis, and visualization features through a single portal. The data obtained from this investigation were analyzed along with the results obtained from gene expression of RNA-seq dataset and Fluidigm high-throughput qPCR presented in the following part.

### **3.4. *In silico* identification of AP2 sub-family genes in African and Asian wild and domesticated rice and *Arabidopsis* genomes**

The *Arabidopsis AP2/ERF* genes were retrieved from the supplementary file provided by a genome wide survey of ERF family genes in *Arabidopsis* and rice (Nakano *et al.*, 2006), and were confirmed using a database of genetic and molecular biology for the model plant *Arabidopsis thaliana* TAIR (Berardini *et al.*, 2015).

To obtain information about orthologs of *AP2/ERF* gene family in other rice species, *AP2/ERF* genomic, protein, 3kb upstream sequences were retrieved in African wild (*O. barthii*, accession IRGC 105608) and domesticated (*O. glaberrima* AG11.1, accession IRGC 96717 & *O. glaberrima* CG14 – UMR DIADE, data in preparation), Asian wild (*O. rufipogon*, OR\_W1943) and domesticated (*O. sativa indica*, ASM465v1) from Gramene (<http://www.gramene.org>) using available online tool BioMart version 0.7. The protein alignment and phylogeny tree were built by MEGA 7.0 to determine the AP2 sub-family genes and PLETHORA genes in these rice species. The sequences were then double-checked by BLAST on *O. sativa ssp. japonica* genome on Gramene (<http://www.gramene.org>) and Rice Genome Annotation Project (Kawahara *et al.*, 2013) to identify orthologs between the genomes.

## **4. Primer Design**

For expression analysis, primers were designed to specifically amplify 80-150 bp-long fragments of cDNA using Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) (Untergasser *et al.*, 2007) and NCBI Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>) web facilities. The specificity of primers was then checked again by BLAST on Gramene (<http://www.gramene.org>) and Rice Genome Annotation Project (Kawahara *et al.*, 2013) <http://rice.plantbiology.msu.edu/index.shtml>.

For genotyping of knock-out transgenic mutants, primers were designed to specifically amplify the genomic region flanking both target sites using Primer3Plus and NCBI Primer-BLAST. The specificity of primers was then checked again by BLAST on Gramene and Rice Genome Annotation Project.

List of all primers used in this PhD project was listed in **Table S5.1-S5.7**.

## **5. *AP2/ERF* gene expression at inflorescence branching in wild and domesticated rice**

### **5.1. Tissue collection and RNA sequencing**

For expression analysis, three biological replicates of approximately 15 immature panicles each from at least 10 plants per accession, per stage, from 4 days to 15 days after floral induction, were collected. For sample collection, leaves surrounding the young panicle were removed by hand and these productive tissues were cut with a sharp blade under a Stemi508 (Zeiss, Germany) stereo microscope. The reproductive tissues were immediately frozen in liquid nitrogen, and total RNA including small RNA was extracted using the RNeasy Plant Mini kit with RLT and RWT buffers (QIAGEN, Germany). DNase treatments were performed using the RNA easy-free DNase set (QIAGEN, Germany). RNA integrity numbers of the extracted RNA, measured using a 2100 Bioanalyzer (Agilent, USA.), were between 8.6 and 10. Stage specificity was validated with quantitative real-time RT-PCR (qPCR) using stage-specific marker genes. 400 ng of total RNA was used for each sample for RNAseq library preparation with the TruSeq Stranded Total RNA with Ribo-Zero Plant kit (Illumina, USA.). After quantification with a 2100 Bioanalyzer, 125-base paired-end reads were generated on a HiSeq 2500 (Illumina, USA) by the GeT Platform (Toulouse, France). This work has been done corresponding to the two Illumina® RNA-seq datasets: inter-specific RNA-seq dataset from panicles of five species and meristem-specific RNA-seq dataset from Laser micro-dissected meristems (LMD) in *O. sativa ssp. Japonica* (Harrop *et al.*, 2016).

For the Multi-species dataset, 5 accessions were selected according to their contrasting panicle architectures: *O. sativa ssp. japonica* cv. Nipponbare, *O. sativa ssp. indica* cv. IR64, *O. rufipogon* acc. W1654, *O. glaberrima* acc. Tog5681 and *O. barthii* acc. B88. RNA-seq was performed at stages 2 (primary branch formed and elongated) and stage 3 (spikelets and floret meristems differentiated from all branches and axillary meristems in the panicle) for the 5 accessions based on the description of panicle development in *O. sativa* (Ikeda *et al.*, 2004).

For the Meristem-specific dataset, reproductive meristems from *O. sativa ssp. Japonica* cv. *Nipponbare* panicles were sampled by laser dissection microscopy and single-end sequencing was performed. Four meristem identities were considered in this case: rachis meristem (RM), primary branch meristem (PBM), elongated primary branch with axillary meristems (ePBM/AM) and spikelet meristems (SM). Part of this analysis has already been published (Harrop *et al.*, 2016).

## 5.2. Fluidigm high throughput qPCR

The Fluidigm 96x96 Dynamic Array (Fluidigm, San Francisco, California, USA) is an integrated fluidic circuit that accepts 96 samples and 96 primer-probe sets and assembles them into 9,216 qPCR reactions. cDNA was synthesized from 1 µg of DNase-treated total RNA using the SuperScript III First-Strand Synthesis System (Invitrogen, USA). A Biomark HD Microfluidic Dynamic Array (Fluidigm, USA) was used for large-scale qPCR. Before performing qPCR, the sample mixture and assay mixture were prepared individually. A 96 × 96 Dynamic Array Integrated Fluidic Circuit (Fluidigm, USA) was loaded with cDNA and primer combinations after 15 cycles of specific target amplification and exonuclease I treatment. A fast cycling protocol with EvaGreen dye (Bio-Rad Laboratories, USA) was used for amplification. Three biological replicates were performed for each sample. Data were normalized using 3 genes (ACT2-LOC\_Os11g06390, HK04-LOC\_Os01g16970, HK09-LOC\_Os03g61680). Gene expression relative to the normalization factors was estimated using the  $2^{-\Delta\Delta CT}$  method without a calibrator sample (Livak and Schmittgen, 2004). Primer sequences are listed in **Table S5.1**.

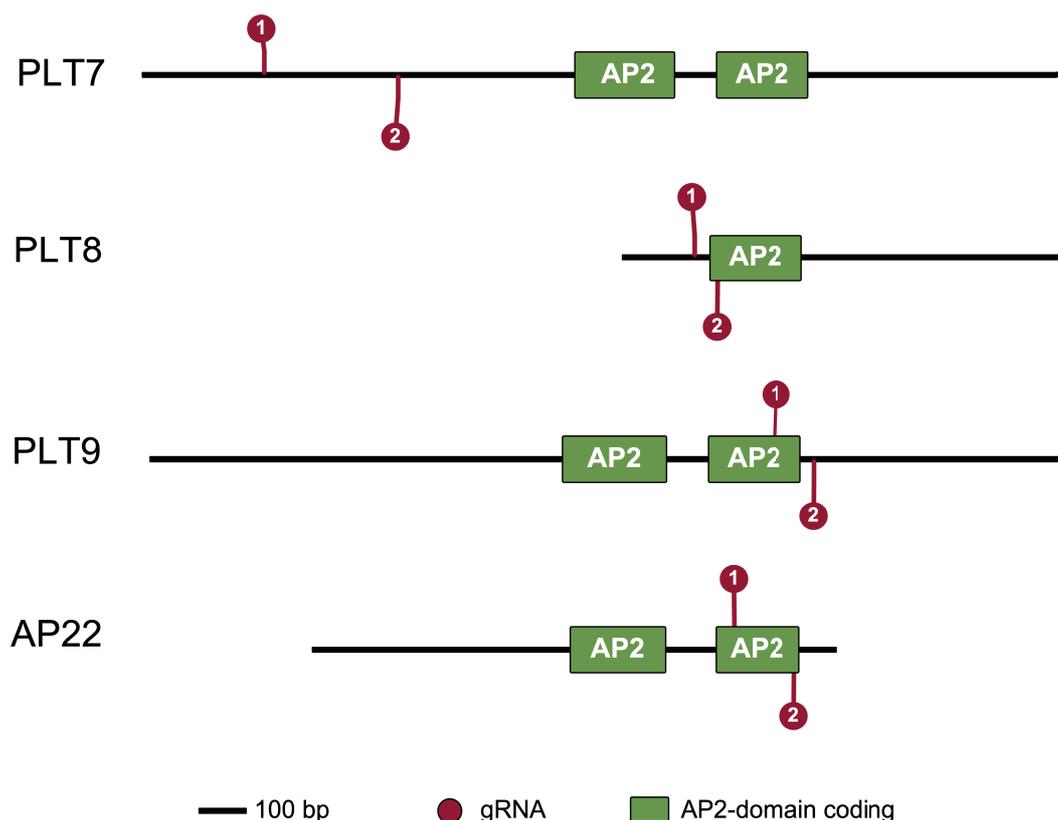
Code for Fluidigm qPCR analysis is at <https://github.com/othomantegazza/fluidgr>, an R Package to Analyze Fluidigm qPCR Data for Gene Expression developed based on the  $\Delta\Delta$  Ct method (Livak and Schmittgen, 2004). This package allows us to load, normalize, scale and visualize Fluidigm qPCR data. A dedicated Excel macro program and JASP (Version 0.9, <https://jasp-stats.org>) was used to manage and analyze qPCR Data.

## 6. Generating the transgenic lines of *PLETHORA* genes

### 6.1. Generating Knock-Out mutants of *PLETHORA* genes by using CRISPR/Cas9 system

#### 6.1.1. Design of gRNAs to target the specific genes

CDS sequences of *PLT7* (LOC\_Os03g56050), *PLT8* (LOC\_Os07g03250), *PLT9* (Os03g0232200) and *AP2/EREBP22* (Os02g0747600) were used as query to design specific guide RNAs (gRNAs), a short synthetic RNA composed of a scaffold sequence necessary for Cas-binding and a user-defined ~20 nucleotide spacer that defines the genomic target to be modified, by using CRISPR-direct (Naito *et al.*, 2015) and Benchling (<https://www.benchling.com/crispr>) web tools. The gRNAs targeting the exon regions, especially the AP2-domain coding region (**Fig. 5.1**) with high specificity and high efficiency value were selected and the specificity of gRNA was checked again by BLAST on Gramene (<http://www.gramene.org>) and Rice Genome Annotation Project (<http://rice.plantbiology.msu.edu/index.shtml>).



**Figure 5.1. Schematic diagram of rice *PLT* genes targeted by gRNAs: *PLT7* (LOC\_Os03g56050), *PLT8* (LOC\_Os07g03250), *PLT9* (Os03g0232200) and *AP2/EREBP22* (Os02g0747600; indicated as AP22).** The black lines represent the CDS of *PLT* genes. The green rectangles indicate AP2 domain coding region. The relative location of gRNA targeting site is shown as red marks.

### 6.1.2. Construction of vectors generating Knock-Out mutants

The polycistronic tRNA-gRNA (PTG) constructs have been synthesized based on the principle of Golden Gate (GG) assembly which is broadly used to assemble DNA parts like customized transcription activator-like effector (TALE). The assembly approach allows synthesizing PTGs with different combinations of gRNAs using the same components as previously described (Xie *et al.*, 2015). PTGs with no more than 6 gRNAs could be synthesized by one step GG assembly, whereas PTGs with more than 6 gRNAs require two or more steps of GG assembly. The schematic diagrams of PTG synthesis approach are shown in **Fig. 5.2**, and details of primer design, GG assembly and plasmid construction are described below.

To ligate multiple DNA parts in a specific order, GG assembly requires distinct 4-bp overhangs, and could not be 5'-GGCA-3' or 5'-AAAC-3' which are used in terminal adaptors for cloning to pRGEB32 (**Fig. 5.3**), to ligate two DNA parts after digestion with *BsaI* (or other type II endonucleases such as *AarI*, *BbsI*, *BsmAI*, *BsmBI*). The gRNA spacer is the only unique sequence in PTG, thus PTGs should be divided into DNA parts within the gRNA

spacer region. A gRNA spacer was split into two parts with 4 bp overlap and each half of the spacer was synthesized within oligo primers with a *BsaI* site. The details of oligo primers are described in **Table S5.2**.

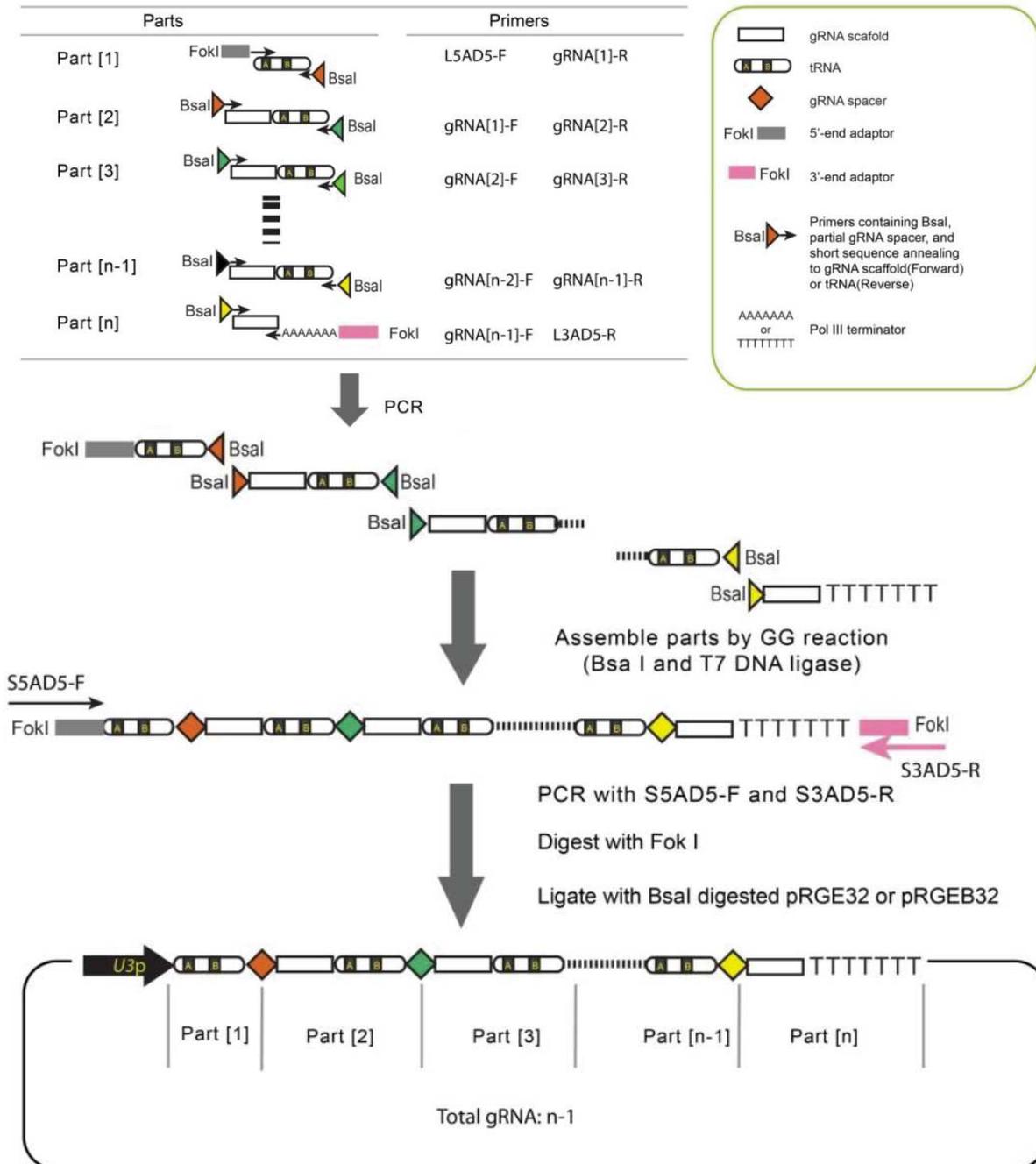
Each fragment of PTG constructs was amplified using Phusion High-fidelity Polymerase Reaction Buffer (ThermoFisher Scientific, USA). The primers used were listed in **Table S5.2 – S5.3**. 50  $\mu\text{L}$  PCR reaction consists of 0.1 ng pGTR plasmid (purchased from AddGene, USA, Plasmid #63143), 10  $\mu\text{L}$  5X Phusion HF buffer, 1  $\mu\text{L}$  dNTPs (10mM), 2.5  $\mu\text{L}$  of each primer (10 $\mu\text{M}$ ), Phusion Taq 0.5  $\mu\text{L}$  and sterile dH<sub>2</sub>O. PCR cycling conditions were: 98 °C for 2 min (1 cycle) and 98 °C for 10 s, an annealing step at various temperatures depending on the T<sub>m</sub> of the primers used (typically T<sub>m</sub> -5 °C), for 20 s, and 72 °C for 30s (35 cycles) with a 5 min final extension step at 72 °C. The integrity and size distribution of the genes were checked by agarose-gel electrophoresis.

The PCR products were then purified with QIAquick PCR Purification Kit (Qiagen, France). The concentrations of DNA samples were determined using a Nano Drop ND-1000 Spectro apparatus. Individual parts were ligated together by GG assembly as follow **Table S5.4**. The 20  $\mu\text{L}$  reaction consists of 25-50 ng of samples (equal amount for each part), 10  $\mu\text{L}$  2X T7 DNA ligase buffer (NEB), 2  $\mu\text{L}$  Bovine Serum Albumin (1mg/mL), 0.5  $\mu\text{L}$  *BsaI* (10 U/ $\mu\text{L}$ , NEB) and 0.5  $\mu\text{L}$  T7 DNA ligase (3000 U/ $\mu\text{L}$ , NEB). GG reactions were performed in a thermal cycler (Bio-Rad) by incubation at 37 °C for 5 min and 20 °C for 10 min during 30-50 cycles, then held at 20 °C for 1 hour.

The GG reactions products were diluted with 180  $\mu\text{L}$  sterile dH<sub>2</sub>O, then were amplified with S5AD5-F and S3AD5-R primers. 50  $\mu\text{L}$  PCR reaction consists of 1  $\mu\text{L}$  GG ligation product, 10  $\mu\text{L}$  5X Phusion HF buffer, 1  $\mu\text{L}$  dNTPs (10mM), 2.5  $\mu\text{L}$  of each primer (10  $\mu\text{M}$ ), Phusion Taq 0.5  $\mu\text{L}$  and sterile dH<sub>2</sub>O. PCR cycling conditions were: 98 °C for 2 min (1 cycle) and 98 °C for 10 s, an annealing step at various temperatures depending on the T<sub>m</sub> of the primers used (typically T<sub>m</sub> -5 °C), for 20 s, and 72 °C for 30s (35 cycles) with a 5 min final extension step at 72 °C. The integrity and size distribution of the genes were checked by agarose-gel electrophoresis. The PCR products were then purified with QIAquick PCR Purification Kit (Qiagen, France). The concentrations of DNA samples were determined using a Nano Drop ND-1000 Spectro apparatus.

The purified PCR products were then digested with enzyme *FokI* (NEB). The *FokI* digested products were separated in 1% agarose gel. The DNA bands with the expected size from the gel were excised from the gel and purified with QIAquick Gel Extraction Kit (Qiagen, France). The concentrations of DNA samples were determined using a Nano Drop ND-1000 Spectro apparatus. The purified products were ligated into the *BsaI*-digested pRGEB32 binary vector (AddGene, USA, Plasmid #63142) by GG assembly with T4 DNA ligase (NEB). The products were purified and transformed by heat shock into *E. coli* TOP10

competent cells (ThermoFisher Scientific (USA). The colonies were selected on LB agar plate containing Ampicillin antibiotic (100 µg/mL) and confirmed by PCR. The selected plasmids were isolated and sequenced for confirmation.

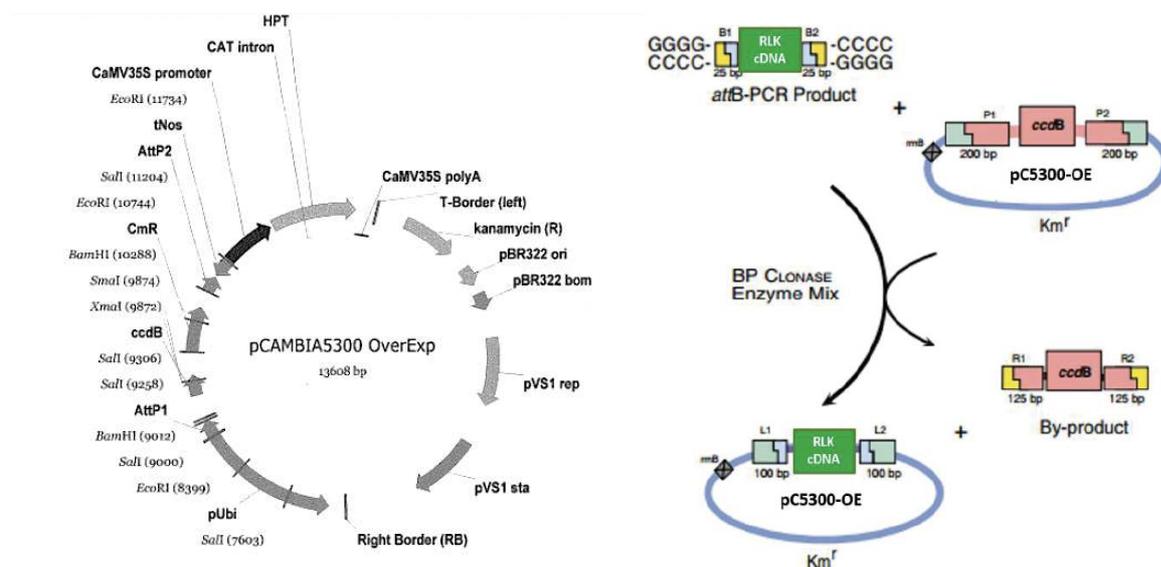


**Figure 5.2. Schematic diagrams for GG assembly to synthesize PTGs from PCR parts and clone them into plasmid vectors pRGE32.** A PTG with n-1 gRNAs are divided into n parts (Part[1] - Part[n], see the bottom). Each part was amplified with spacer-specific primers containing *BsaI* adaptor, except two terminal parts using gRNA spacer primer and terminal specific primers containing *FokI* site (L5AD5-F and L3AD5-R). These PCR parts were ligated together using GG assembly to produce the PTG with complete gRNA spacers. The assembled product was amplified with short terminal specific primers (S5AD5-F and S3AD5-R). After Fok I digestion, the PTG fragment was inserted into the *BsaI* digested pRGE32. See Table S5.2 and Table S5.3 for primer sequences and SI Methods for details (figure modified from Xie, Minkenberg and Yang, 2015).



## 6.2. Generating the Over-Expressing mutants of *PLETHORA* genes

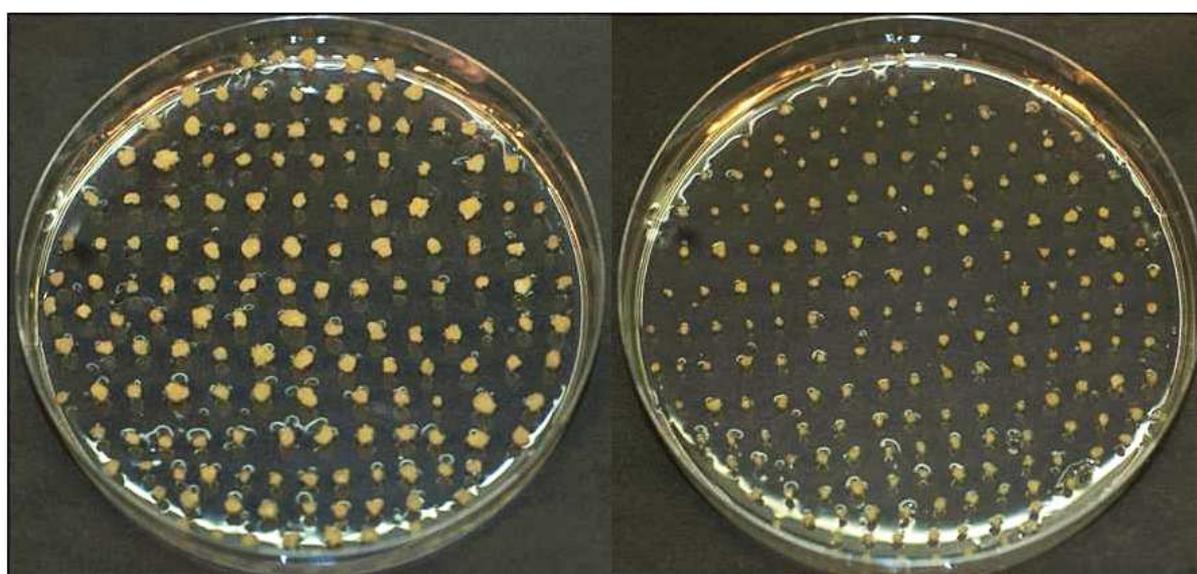
The CDS of *PLT7* (LOC\_Os03g56050), *PLT8* (LOC\_Os07g03250), *PLT9* (Os03g0232200) and *AP2/EREBP22* (Os02g0747600) were retrieved from rice database and modified with BP tailed. Gene synthesis of the CDS with BP sequence adaptors was carried out by a private company (Eurofins Genomics, Belgium). *PLT8* (1003 bp), *PLT9* (1987 bp) and *AP2/EREBP22* (1162 bp) with BP tailed (PLT CDS) were successfully synthesized and ligated into pEX-A258 plasmid (Eurofins Genomics, Belgium). The CDS fragments in pEX-A258 were transfer into pCAMBIA 5300-OE binary vector by using BP clonase reaction (Gateway BP recombinase Gateway® BP Clonase™ II Enzyme Mix, Invitrogen, USA). The pCAMBIA 5300-OE was a modified binary vector in which the *ccdB* gene surrounded by the BP recombination sites were cloned between the constitutive promoter of ubiquitin gene from maize and the terminator of the nopaline syntase gene from *A. tumefaciens* (Fig. 5.4). The PC5300-OE vector was obtained from J-C. Breitler, CIRAD, France. After cloning, the presence and the integrity of the PLT CDS were verified by sequencing. The PC5300.OE-PLT plasmids were transferred into electro-competent *Agrobacterium tumefaciens* strain EHA105.



**Figure 5.4.** (Left) Schematic diagram of pCAMBIA5300-Overexpression circle map (JC. Breitler, CIRAD, France). From right to left, Kanamycin (R), kanamycin resistance gene; HPT, hygromycin phosphotransferase; CaMV35S-promoter, cauliflower mosaic virus 35S promoter; pUbi, rice ubiquitin promoter. (Right) Principle of cloning PCR product by BP reaction (Gateway technologies, Promega, USA).

### 6.3. *Agrobacterium tumefaciens*-mediated rice transformation

Binary vectors were transformed via electroporation into the *Agrobacterium tumefaciens* strain EHA105. Genetic transformation of *O. sativa ssp. japonica* cv. Kitaake was conducted according to a previously described protocol (Hiei and Komari, 2008) with modification of Hoge's group (Institute of Molecular Plant Science, University of Leiden, Netherland) and DAR group (CIRAD, Montpellier, France). The protocol was summarized in **Table S5.5**. The composition of the different media used is detailed in the Material and Methods part 9. Mature seeds were induced to generate the rice callus in NB medium plate at 28°C in dark for 20-24 days at 60% of humidity. The single rice calli derived from seeds (**Fig. 5.5**) were then transferred to new NB medium for multiplication.



**Figure 5.5.** The first day of multiplication of rice calli derived from mature seeds of *O. sativa ssp. japonica* cv. Kitaake.

After 7-10 days of multiplication, these embryogenic calli are the most reactive to genetic transformation. To be ready for rice calli transformation, *Agrobacterium* carrying binary vectors were inoculated in AB medium plates containing Kanamycine (50 µg/mL) and Rifampicine (20 µg/mL) antibiotics 2-3 days at 28°C in dark, then were diluted in medium R2-CL liquid containing 100 µM acetosyringone at  $OD_{600} = 0.1$ . The *Agrobacterium* R2-CL liquid prepared was poured to plate containing 25-30 rice calli and kept for 10 min. The bacteria liquid was removed and the calli were then dried and placed in R2-CS medium plates for 3 days at 25°C in dark with humidity of 60%. After 3 days of co-cultivation, all calli were moved to selection medium I (R2-S) plates for 1 week. The selected calli were moved to selection medium II (NB-S) plates for 14-18 days, then selection III Pre-generation (PRN) medium for 7-10 days. All three selection stages were done at 28°C in dark with humidity of

60%. Selected calli were then transferred to regeneration (RN) medium for 18-22 days at 28°C in light (12 light hours per day) with humidity of 60%. The regenerated green plantlets of about 3 cm were transferred in the rooting medium (MS) tube for 3-4 weeks at 28°C in light (12 light hours per day) with humidity of 60% (**Fig. 5.6**). After that, DNA was extracted from plant leaves and genotyped as described below. All transgenic plants detected were transferred to greenhouse.



**Figure 5.6.** The regeneration stage of rice calli transformed (at day 20<sup>th</sup> of regeneration) (left picture). The rooting stage of plantlets transformed (at day 20<sup>th</sup> of rooting) (right picture).

**Table S5.5. Summary about culture condition of rice transformation**

Step	Medium	Culture condition	Time duration
Callus induction	NB	28°C, humidity 60%, dark	20 to 24 days
Callus multiplication	NB	28°C, humidity 60%, dark	7 to 10 days
Co-culture with <i>Agrobacterium</i>	R2-CS	25°C, humidity 60%, dark	3 days
Selection I	R2-S	28°C, humidity 60%, dark	7 days
Selection II	NB-S	28°C, humidity 60%, dark	14 to 18 days
Selection III	PRN	28°C, humidity 60%, dark	7 to 10 days
Regeneration	RN	28°C, humidity 60%, light 12 hours per day	18 to 22 days
Rooting	MS	28°C, humidity 60%, light 12 hours per day	20 to 25 days
<i>In vivo</i> acclimation	MS	green house, 28°C, humidity 60%, light 12 hours per day	2 months

## 7. Genotyping and Phenotyping Knock-Out transgenic lines

### 7.1. Genomic DNA isolation

DNA samples were quickly extracted from leaves of 14-day-old seedlings following Edward's method (Edwards *et al.*, 1991). About 50 mg leaf were grinded with ball bearings in 2 mL microfuge tube, then homogenized in 400 $\mu$ L Edward's buffer (200 mM Tris-HCl pH 8, 250 mM NaCl, 25 mM EDTA, 0.5% SDS), vortexed well and centrifuged at 14,000 rpm for 10 min at 4°C. The upper aqueous layer was transferred to a new 1.5 mL microfuge tube which contained of 300 $\mu$ L cold iso-propanol, well mixed and gently inverted. The tubes were centrifuged at 14,000 rpm for 5 minutes at 4°C. Supernatant was discarded and washed by adding 500 $\mu$ L of ethanol (EtOH) 70% (v/v). The tubes were centrifuged at 14,000 rpm for 10 min and the EtOH was discarded. The pellet was dried and resuspended in 50 $\mu$ L sterile dH<sub>2</sub>O. The concentrations of DNA samples were determined using a Nano Drop ND-1000 Spectro apparatus.

### 7.2. Genotyping of genome-edited progeny

Polymerase chain reactions (PCRs) were performed with GoTaq DNA Polymerase Reaction Buffer (Promega, USA), according to manufacturer's instruction. DNA fragments were analyzed by electrophoresis with agarose gel stained with ethidium bromide. The primers used are listed in **Table S5.3**. The genome editing device in the regenerated plants was detected with primers specific for the 600bp Hygromycine resistant gene fragment and 1kb Cas9 gene fragment (**Fig. 5.3**). The selected regenerated plants in T<sub>0</sub> generation were those carrying the full T-DNA. Genomic deletions were detected by PCR with primers flanking the two target sites of each gene. PCR cycling conditions were: 95 °C for 2 min (1 cycle) and 95 °C for 30 s, an annealing step at various temperatures depending on the T<sub>m</sub> of the primers used (typically T<sub>m</sub> -5 °C), for 30 s, and 72 °C for 1 min (35 cycles) with a 5 min final extension step at 72 °C. The integrity and size distribution of the genes were checked by agarose-gel electrophoresis.

Selected PCR products were sent to Genewiz, England for Sanger sequencing to determine the specific mutation. The results were visualized using 4Peaks (by A. Griekspoor and Tom Groothuis, mekentosj.com). Double peaks were resolved using degenerate sequence decoding (Liu *et al.*, 2015) and CRISP-ID (Dehairs *et al.*, 2016) web tools.

### 7.3. Plant phenotyping

For *cr15*, *erf48* and *smos1-3* mutants and related wild type genetic background, at least 18 panicles of each line were used for panicle phenotyping. For mutant lines generated from *Oryza sativa ssp. japonica* cv. kitaake, at least 20 panicles were collected for panicle analysis. For phenotyping analyses, each panicle was spread out and fixed on white paper using adhesive tape. Panicles were photographed and the images were used for panicle structure and

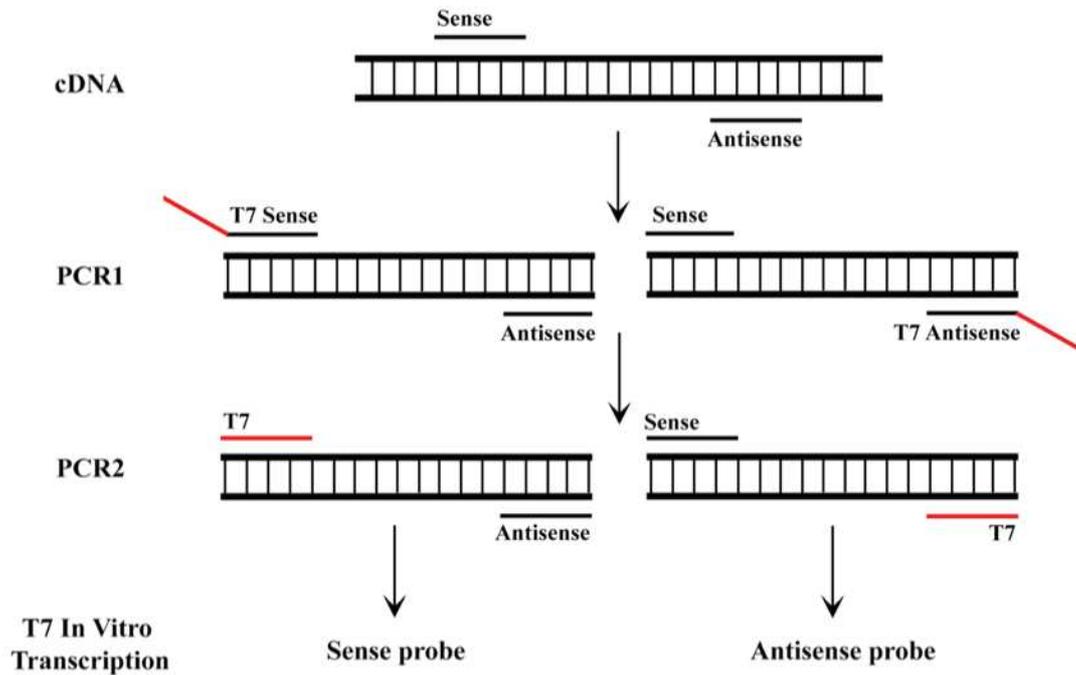
seed number analysis with P-TRAP software (AL-Tam *et al.*, 2013). Then, ptrapr R-script (<https://othomantegazza.github.io/ptrapr>), was used to load the P-TRAP output to visualize and make the topology of multiple panicles. Morphological traits of the panicles were measured: rachis length, the number of spikelets per panicle, the number of primary, secondary or tertiary branches per panicle, and the number of nodes associated. Other traits were collected during the development of the plants, including flowering time, plant height, the number of tillers and efficient tillers (i.e. tiller producing a panicle) per plant. The plants were photographed using a Canon PowerShot G12 camera. Flowers of *O. sativa* cv. kitaake wild type and mutant plants were cut to observe the shape of palea, lemma, and pistil and stamen phenotype. The flowers were photographed using a Q-capture pro 7 imaging system in conjunction with a binocular Leica MZFLII.

## 8. *In situ* hybridization

### 8.1. Preparation of sense and antisense RNA probes

RNA probes for *in situ* hybridizations were obtained by using PCR-amplified fragments including a T7 RNA polymerase promoter sequence at one end. cDNA synthesized from total RNAs of different stages during panicle development was used to prepare the probes. PCR amplifications were performed using specific primers using a two steps PCR amplification (see **Table S5.6**). Each 50  $\mu$ L PCR reaction of PCR1 consists of 25-50 ng of cDNA, 10  $\mu$ L 5X Phusion HF buffer, 1  $\mu$ L dNTPs (10mM), 2.5  $\mu$ L of both sense and antisense gene specific primers (10 $\mu$ M), Phusion Taq 0.5  $\mu$ L and sterile dH<sub>2</sub>O. PCR amplifications were performed using following conditions: 98 °C for 2 min (1 cycle) and 98 °C for 10 s, an annealing step at various temperatures depending on the T<sub>m</sub> of the primers used (typically T<sub>m</sub> -5 °C) for 30 s, and 72 °C for 30s (35 cycles) with a 5 min final extension step at 72 °C.

Then, PCR2 reaction used PCR1 products (1/200 dilution) as template with 2 reactions (T7 primer +Primer Antisense; Primer Sense + T7 primer) (**Fig. 5.7**). PCR1 and PCR2 conditions are similar (see above). PCR2 product was purified by EtOH precipitation, and eluted in 50 $\mu$ L RNase sterile dH<sub>2</sub>O. The concentrations of DNA samples were determined using a NanoDrop ND-1000 Spectro apparatus. PCR2 products (1 $\mu$ g/ $\mu$ L) were used directly as template for sense and antisense probe transcription. The RNA probes were synthesized using UTP–digoxigenin (Roche, France) as the label in conjunction with a T7 Maxi Script kit (Ambion, France) according to the manufacturer's instructions. The size distribution and labeling efficiency of the RNA probes were evaluated by agarose gel electrophoresis (2% (w/v)) and by dot-blot hybridization (according to manufacturer's instructions), respectively.



**Figure 5.7. Schematic diagram showing sense and antisense probe assays.** DNA fragments from PCR reaction Sense+T7 provided antisense RNA probes while the DNA fragments from PCR reaction T7+Antisense provided sense RNA probes. Red marks indicate T7 oligo sequence.

## 8.2. Fixation of tissues

Samples were placed in the fixation solution (4% (v/v) paraformaldehyde, 0.1 M phosphate buffer, pH7) in a 50 mL tube and vacuum was applied several times until the samples fall down at the bottom of the tubes. Samples were incubated in this solution overnight at 4°C. Tissues were rinsed 3 times (15 min each rinse) with first rinsed-solution (PBS 1X, glycine 0.1M) and one time for 30 min in PBS 1X and then in PBS 1X overnight at 4°C.

Tissues were dehydrated through a graded EtOH series (30, 50, 70, 80, 90, 95% (v/v)) for 1 hour each. The dehydration was finished with 3 washes (30 min each) in 100% (v/v) EtOH and tissues were incubated in this solution overnight at 4°C. For a long-term storage, the samples have to be kept in 70% (v/v) EtOH at 4°C.

## 8.3. Impregnation in paraplast

Fixed tissues were incubated in solution of EtOH 50% (v/v) and butanol 50% (v/v) for 1 hour at room temperature. Tissues were rinsed twice in butanol 100% for 1 hour and kept in butanol 100% 48 hours at 4°C.

Samples were embedded in paraffin by gradual change of solutions from butanol to histo-clear (HC) and then from HC to paraffin following this protocol: samples were incubated in a

series of HC: butanol solutions (1:3; 1:1; 3:1) at 4°C for 1 hour each. Then, the samples were washed twice with 100% HC and kept overnight at 4°C in 100% HC. HC was then replaced by paraffin using following protocol: samples were incubated in a series of HC: paraffin solutions (3:1; 1:1; 1:3) at 60°C for 3 hours each, then several times in 100% paraffin, and incubated in paraffin in special block made using plastic mold (i.e. culin) adapted to the size of the sample overnight at 60°C. After that, blocks were incubated in room temperature until paraffin become solid, stored at 4°C for short period storage but at -20°C for long-term period.

The slides of sample were prepared one day before *in situ* hybridization experiment. The plastic block of paraffin containing the samples was dissected and histological sections of 7 µm of thickness were made using a microtome (Leica GmbH D6907 instrument Nussloch, Model Jung RM 2055 -Germany), spread on glass slides (Silanized Slides VWR) and then dried at 36°C overnight. The slide could be stored at 4°C for 3 months.

#### **8.4. Probe hybridization and labelling**

In the first step, the tissues slides were dewaxed by histo-clear (3 times for 10 min) and hydrated through a graded EtOH series, (100°- 2 times for 10 min, 70° and 50° for 5 min). Then, the second step, the slides were treated by proteinase K in 200 mL 1X Proteinase K buffer (added 134 µL proteinase K 0.1U/ml) at 37°C for 15 min and washed by TRIS 1X (2 times) for 5 min, PBS at 0.2% of glycine for 2 min, PBS 1X pH 7.4 (2 times) at 2 min. In step 3, the slide was dehydrated by cleaned through a graded EtOH series (50°, 70°, 100° x 2 times) for 1 min per solution. Hybridization mix was prepared at step 4 including 50 µL formamide 100%, 10 µL SSC 20X buffer, 20 µL Sulfate Dextran 50% (w/v), 4µL Denhardt 50X, 1 µL tRNA (11ng/mL), 1.5 µL probe (200ng/µL) and DEPC water to have final volume 100 µL per slide. Mix was denatured at 65°C for 5 min and kept in ice. Hybridization chamber was stick in the slide and loaded with the hybridization mix (100 µL per slide). The hybridization was done in a humidified box (Thermo, plate Omni slide thermo cycler) at 42°C overnight. In order to accomplish step 5, the slide was washed to remove non hybridized single-stranded RNA probe: SSC 2X buffer (one time for 5 min in room temperature and one time for 45min at 50°C) ; NTE 1X buffer (2 times for 5 min) ; NTE 1X buffer included 400 µL RNase A (10g/L) for 30 min at 37°C ; SSC 2X buffer for 15 min, and PB 1X (2 times for 5 min). In the last step, the slide was incubated in 700 µL of 1% (w/v) blocking solution in PBS 1X buffer for 1 hour in humid chamber. Then, this solution was replaced by 500 µL of 1% (w/v) blocking solution with 1/500 antibody anti-digoxigenin in humid chamber. The slide was washed by PBS 1X buffer (3 times for 10 min) and revelation buffer 1X (2 times for 10 min). The hybridization was revealed by using the Vector Blue Alkaline Phosphatase Substrate Kit III (Vector Laboratories) according to the supplier's instructions. The sections

were observed using a Leica (Leitz DMRB) microscope and photographs were taken with a Q-capture pro 7 imaging system.

## 9. MEDIAS, SOLUTIONS AND BUFFERS

### Solutions used for *in situ* hybridization and rice transformation

All the *in situ* hybridization solutions were *RNAse* free and prepared using DEPC-treated water.

All the transformation media were autoclaved. For the medium added antibiotics, the antibiotics were sterile filtered and added when the media were autoclaved and cooled down.

<b>DEPC water (1L)</b>		<b>NTE 10X buffer (1L)</b>	
DEPC	100µL	Tris-HCL (100mM)	12.114g
Water (miliQ)	900 µL	NaCl (5M)	292.2g
Mixed well and incubated 4 hours before autoclave		EDTA (10mM)	3.671g
		pH	7.5
<b>Proteinase K 10X buffer (500mL)</b>		<b>Revelation 10X buffer</b>	
Tris-HCL (1M)	60.57g	Tris-HCL (1M)	121,14g
EDTA (0.5M)	84.05g	pH	8.2
pH	8		
<b>TRIS 10X buffer (500mL)</b>		<b>RNase A (10g/L)</b>	
Tris-HCL (1M)	60.57g	RNase power	0.04g
pH	8.2	NTE1X	4mL
		Boiled for 5 min at 100°C, then aliquoted in 1.5mL microfuge tube.	
<b>PBS 10X buffer (1L)</b>		<b>Blocking solution 10 % (100mL)</b>	
Na <sub>2</sub> HPO <sub>4</sub> (70mM)	9.94g	Blocking reagent	10g
NaH <sub>2</sub> PO <sub>4</sub> (30mM)	3.6g	Maleic acid (C <sub>4</sub> H <sub>4</sub> O <sub>4</sub> )	1.161g
NaCl (1.2M)	70.128g	Heated by microwave to solubilise, mixed well, autoclaved and aliquoted. Solution was stored at -20°C.	
KCl (27mM)	2g		
pH	7.4		
<b>SSC 20X buffer (1L)</b>		<b>Denhardt 50X (500mL)</b>	
NaCl (3M)	175.5g	NaCl	0.8766g
Na <sub>3</sub> C <sub>6</sub> H <sub>5</sub> O <sub>7</sub> (300mM)	88g	pH	7.5
pH	7		
<b>Medium LB (1L)</b>		<b>Medium LB agar</b>	
Bacto®-tryptone	10g	Bacto®-tryptone	10g
Bacto®-yeast extract	5g	Bacto®-yeast extract	5g
NaCl	10g	NaCl	10g
In deionised water		Bacto-Agar	1,5% (w/v)



## SUPPLEMENTAL FIGURES AND TABLES

**Table S5.1. Primers used for Fluidigm high-throughput qPCR**

Gene name	Oligo Name	Sequence (5'-3')
LOC_Os05g41760	HA_MS F1-F	CGGCTCGTGATCTCGACACGTAC
	HA_MS F1-R	CACAGCCGGACCAGTGCTCTC
LOC_Os03g60430	HA-OsIDS1-F	ATAGCCAGGGTGAAGCAGAAGG
	HA-OsIDS1-R	CCAACCAAACACTTCCACCC
LOC_Os04g55970	HA-PLT1-F	GCAAGGGACGCCAAGTGT
	HA-PLT1-R	GAAAGTTGGTTGTGGTGCT
LOC_Os03g07940	HA-PLT10-F	AAGGATTGGAAGGGTCCGAG
	HA-PLT10-R	CGATGAGGTGGGTGCTGGA
LOC_Os04g42570	HA-PLT4-02-F	AAGAAGATAAGGCGGCTCGG
	HA-PLT4-02-R	GCAATGTACTCCTGCCTGGT
LOC_Os05g03040	HA-RSR1-F	AGCGAAGTAGAGGCTGCAAG
	HA-RSR1-R	CAGTAGGCAGCTCACCATCA
LOC_Os07g13170	HA-SNB-F03	AGTTCTGTTGTTACTCAGCC
	HA-SNB-R03	GAGGTTCAAAAGAAGCCC
LOC_Os03g56050	PLT7-F	TGGCTCACCTCAGAAGGAA
	PLT7-R	TCCTCCTGCGTGCTGAATGT
LOC_Os07g03250	PLT8-F	TAGGGTTCTTGGTTGCTCGG
	PLT8-R	CGGAGAAGAAGGAAAGGTGG
LOC_Os03g12950	PLT9-F	CGAGAGAGCAACGCAAGAAC
	PLT9-R	AGAGCGAGAAGCCTAACCAG
LOC_Os02g51300	EREBP22-F	CTCGGTAGTTGATTCCTCCC
	EREBP22-R	CACCAATCAATCGCTCTACCC
LOC_Os07g47330	HA-FZP-F01	CTCCGACTCCTACTCTCCATTC
	HA-FZP-R01	CAGAGGCAAAGTGCGTGATTAG
LOC_Os04g55560	SHAT1-F	ACCGTTTCTTCCTCTTTCGTTC
	SHAT1-R	AGTTCGTCAACGCATCCTTG
LOC_Os06g11170	RT 418 F	GGAATGTGGACGGTGACACT
	RT 419 R	TCAA AATAGAGTCCAGTAGATTTGTCA
LOC_Os06g48970	RT 420 F	GGCGCTTAAAGA ACTTAAGAGGA
	RT 421 R	TGCATCGTAGCCCCTGTAAT
LOC_Os01g16970	HA-HK04-F	AGTTCGTCAAGTCTCCATTCC
	HA-HK04-R	CAGAGACTGATTCCAAGCC
LOC_Os03g61680	HA-HK09-F	TCAAGATAGTCACAGAGAGCC
	HA-HK09-R	AGCATCGGGAAGAGAACAGG

Table S5.2. Detail about primers designed to amplify *PTG* genes (gRNA-tRNA parts)

gRNA Target	Spacer/Protospacer (5'->3') <sup>(a)</sup>	Oligo ID	Sequence (5'->3') <sup>(b)</sup>
PLT8a	GGAAAAGT <u>AGCG</u> GTTTCTC A	PLT8a-F	TA <i><b>GGTCTCC</b></i> <u>AGCG</u> GTTTCTCA gtttagagctagaa
		PLT8a-R	AT <i><b>GGTCTCA</b></i> <u>CGCTACTTTTCC</u> tgcaccagccgggaa
PLT8b	GCAGGCAT <u>CATC</u> AGCACGG A	PLT8b-F	TA <i><b>GGTCTCC</b></i> <u>CATCAGCACGGA</u> gtttagagctagaa
		PLT8b-R	AT <i><b>GGTCTCA</b></i> <u>GATGATGCCTGC</u> tgcaccagccgggaa
PLT9a	CAAGTTCC <u>GTGG</u> CCTGAAC G	PLT9a-F	TA <i><b>GGTCTCC</b></i> <u>GTGGCCTGAACG</u> gtttagagctagaa
		PLT9a-R	AT <i><b>GGTCTCA</b></i> <u>CCACGGA</u> ACTTG tgcaccagccgggaa
PLT9b	GCCTGGTG <u>AGGC</u> AGCGCGT A	PLT9b-F	TA <i><b>GGTCTCC</b></i> <u>AGGCAGCGCGTA</u> gtttagagctagaa
		PLT9b-R	AT <i><b>GGTCTCA</b></i> <u>GCCTCACCAGGC</u> tgcaccagccgggaa
PLT7a	GCCG <u>GCGG</u> CGGATCGTACT A	PLT7a-F	TA <i><b>GGTCTCC</b></i> <u>GCGGCGGATCGTACTA</u> gtttagagctagaa
		PLT7a-R	AT <i><b>GGTCTCA</b></i> <u>CCGCCGGC</u> tgcaccagccgggaa
PLT7b	CGGGGATG <u>ACAG</u> CGCACTG C	PLT7b-F	TA <i><b>GGTCTCC</b></i> <u>ACAGCGCACTGC</u> gtttagagctagaa
		PLT7b-R	CG <i><b>GGTCTCA</b></i> <u>CTGTCATCCCCG</u> tgcaccagccgggaa
AP2-22a	GTGGCAGG <u>CACG</u> CATAGGA C	AP2-22a-F	TA <i><b>GGTCTCC</b></i> <u>CACGCATAGGAC</u> gtttagagctagaa
		AP2-22a-R	CG <i><b>GGTCTCA</b></i> <u>CGTGCCTGCCAC</u> tgcaccagccgggaa
AP2-22b	GTAGTTGC <u>TTCG</u> ATCAAAG T	AP2-22b-F	TA <i><b>GGTCTCC</b></i> <u>TTCGATCAAAGT</u> gtttagagctagaa
		AP2-22b-R	CG <i><b>GGTCTCA</b></i> <u>CGAAGCAACTAC</u> tgcaccagccgggaa

(a) The RED UNDERLINED letters indicate the overhang sequences in Golden Gate assembly.

(b) The first two letters are randomly added nucleotides. *ITALIC BOLD* sequences indicate the *BsaI* sites (5'-GGTCTCN-3', N indicates any nucleotide), UNDERLINED sequences are part of gRNA spacer whereas RED UNDERLINED sequences are overhangs after *BsaI* digestion. Sequences in lower case are specific for gRNA scaffold (5'-gtttagagctagaa-3', in forward primers) or tRNA (5'-tgcaccagccggg-3', in reverse primers).

**Table S5.3. Primers used for plasmid constructions, qRT-PCRs, and genotyping**

Oligo Name	Sequence (5'-3')	Purpose
HA_PLT7-Geno-F	TGGAGGTGGGCTGTTCTACA	Genotyping and sequencing of <i>PLT7</i>
HA_PLT7-Geno-R	GTGTCTGTCACTATGGCGCT	
HA_PLT7-Seq-F	GTTCTACAACCCTGCCGCC	
HA_PLT7-Seq-R	TATGGCGCTGCTAGCTACTAC	
HA_PLT8-Geno-F	GTAAACGTTCTACCGGCCA	Genotyping and sequencing of <i>PLT8</i>
HA_PLT8-Geno-R	TCGTACTGAAAGTGCCGAGG	
HA_PLT8-Seq-F	TGAAGAAGTCTGAACATGTGTGT	
HA_PLT8-Seq-R	TAGAGGTCCTTGTTGCCGGA	
HA_PLT9-Geno-F	ACGGAGCGTTTCCATTGGTT	Genotyping and sequencing of <i>PLT9</i>
HA_PLT9-Geno-R	GTGCCCGTGACAGTAGCAG	
HA_PLT9-Seq-F	TTTATGCAATACAGGCACTCAAG	
HA_PLT9-Seq-R	CTCGCTTCCTCGGTCGC	
HA_AP22-Geno-F	CTTCAGTTTGTTGCCAAGGCT	Genotyping and sequencing of <i>AP2-EREBP22</i>
HA_AP22-Geno-R	CCCTAATATGCGATGCGGCT	
HA_AP22-Seq-F	CTCTGCTAAACCATGCCCT	
HA_AP22-Seq-R	CTTCAAGCCTGCCCTTCTAT	
HA_Hygro-F	GCTCCAGTCAATGACCGCTG	Hygromycin phosphotransferase II detection
HA_Hygro-R	CTCGGAGGGCGAAGAATCTC	
HA_UBI-F	GCTTGTGCGTTTCGATTTGA	Detection of <i>Cas9</i>
HA_Cas9-R	CCGCTCGTGCTTCTTATCCT	
HA_control-F	CGCTGCCACTCTCCACTGA	Control for DNA quality
HA_control-R	AGCTGCTTCCACTCGTTCCA	
HA_UGW-U3-F	GACCATGATTACGCCAAGCTTAAGGAATCTTTAAA CATACG	Detection of U3:PTG
HA_UGW-gRNA-R	GGACCTGCAGGCATGCACGCGCTAAAAACGGACT AGC	
L5AD5-F	CGGGTCTCAGGCAGGATGGGCAGTCTGGGCAACA AAGCACCAGTGG	PTG synthesis and cloning <i>FokI</i> site ( <u>underlined</u> ) was used to generate compatible overhangs (labeled with red color) for cloning into pRGEB32
L3AD5-R	TAGGTCTCCAAACGGATGAGCGACAGCAAACAAA AAAAAAA GCACCGACTCG	
S5AD5-F	CGGGTCTCAGGCAGGATGGGCAGTCTGGGCA	
S3AD5-R	TAGGTCTCCAAACGGATGAGCGACAGCAAAC	

**Table S5.4. Detail about construction of PTG cassettes**

PCR ID	Forward Primer	Reverse primer	Part Symbol
P1	L5AD5-F	PLT7a-R	L5AD-PLT7a
P2	PLT7b-F	L3AD5-R	PLT7b-L3AD
P3	PLT7a-F	PLT7b-R	PLT7a-PLT7b
P4	L5AD5-F	PLT8a-R	L5AD-PLT8a
P5	PLT8b-F	L3AD5-R	PLT8b-L3AD
P6	PLT8a-F	PLT8b-R	PLT8a-PLT8b
P7	L5AD5-F	PLT9a-R	L5AD-PLT9a
P8	PLT9b-F	L3AD5-R	PLT9b-L3AD
P9	PLT9a-F	PLT9b-R	PLT9a-PLT9b
P10	L5AD5-F	AP2-22a-R	L5AD-AP22a
P11	AP2-22b-F	L3AD5-R	AP22b-L3AD
P12	AP2-22a-F	AP2-22b-R	AP22a-AP22b
P13	PLT7b-F	PLT8a-R	PLT7b-PLT8a

Construct	Encoding gRNA	PCR parts used
PLT7	PLT7a, PLT7b	P1, P3, P2
PLT8	PLT8a, PLT8b	P4, P6, P5
PLT9	PLT9a, PLT9b	P7, P9, P8
AP22	AP22a, AP22b	P10, P12, P11
PLT7+8	PLT7a, PLT7b PLT8a, PLT8b	P1, P3, P13, P6, P5

**Table S5.6. Primers used for *in situ* hybridization probes**

Oligo Name	Sequence (5'-3')
HA-CRL5-insitu-F	ACGACGTGGCGGCGATCAAG
HA-CRL5-insitu-R	TTAGGCGTCGGTCCAGGCGG
HA-CRL5-insitu-T7-F	GCGAAATTAATACGACTCACTATAGGGCGAAACGACGTGGCGGCGATCA AG
HA-CRL5-insitu-T7-R	GCGAAATTAATACGACTCACTATAGGGCGAATTAGGCGTCGGTCCAGGC GG
HA-PLT8-pro-2F	ATTCTCCGTCCTCCACGACCTC
HA-PLT8-T7-2F	GCGAAATTAATACGACTCACTATAGGGCGAAATTCTCCGTCCTCCACGAC CTC
HA-PLT7-new-pro-F	ACCGCCGGAGGTGTCTG
HA-PLT7-new-pro-R	TTGCTGTTCTTGGTCGCCTC
HA-PLT7-new-pro-T7-F	GCGAAATTAATACGACTCACTATAGGGCGAAACCGCCGGAGGTGTCTG GCGAAATTAATACGACTCACTATAGGGCGAATTGCTGTTCTTGGTCGCCT C
HA-PLT7-new-pro-T7-R	C
HA-AP22-new-pro-F	ACCACGAAGGGAGTTGAGTC
HA-AP22-new-pro-R	GCCGTAGTTCCAGCAGTACC
HA-AP22-new-pro-T7-F	GCGAAATTAATACGACTCACTATAGGGCGAAACCACGAAGGGAGTTGAG TC
HA-AP22-new-pro-T7-R	GCGAAATTAATACGACTCACTATAGGGCGAAGCCGTAGTTCCAGCAGTA CC
HA_PLT9-Sens	TATAGCCAGGGCCAAGAAGC
HA_PLT9-ASens	GCTGGTGACGAGGAAGTTCT
HA_PLT9-pT7sens	GCGAAATTAATACGACTCACTATAGGGCGAATATAGCCAGGGCCAAGAA GC
HA_PLT9-pT7AS	GCGAAATTAATACGACTCACTATAGGGCGAAGCTGGTGACGAGGAAGTT CT
HA-T7-HIS	GCGAAATTAATACGACTCAC

**Table S5.7. Primers used for validating the annotation CDS of interested *PLETHORA* genes**

Gene name	Oligo Name	Sequence (5'-3')
LOC_Os03g56050	HA-PLT7-V-F	AGTCCATTGACACGTTCGGC
	HA-PLT7-V-R	CAGCTCCTCCTGGTAGTCCT
LOC_Os07g03250	HA-PLT8-V-F	ATCGACACGTTCGGTCAGAG
	HA-PLT8-V-R	GCCACATACTCCTGCCTTGT
	HA-PLT8-V2-R	TCTGGCCTTCCTTCTTGCAG
	HA-PLT8-V3-F	CCAGGAAAGGGAGGCAAGTT
LOC_Os03g12950	HA-PLT8-V3-R	CCCGTGAGAAACCGCTACTT
	HA-PLT9-V-F	CCAAAAGCAGCCTGTTCACC
	HA-PLT9-V-R	CCGCTGCTTCTCCTTCTCAA
	HA-PLT9-V2-F	TGACGGAAGCAGAGATGCTG
LOC_Os02g51300	HA-PLT9-V2-R	CTAACGAACTTGCCTTCCCTTC
	HA-AP22-V-F	TGATGAGAAGGCGGAGAGGA
	HA-AP22-V-R	AGCAGCAATGTCGTAAGCCT

## GENERAL REFERENCES FOR GENERAL INTRODUCTION AND MATERIALS &amp; METHODS

- AL-Tam, F., Adam, H., Anjos, A.D., Lorieux, M., Larmande, P., Ghesquière, A., Jouannic, S. and Shahbazkia, H.R.** (2013) P-TRAP: A Panicle Trait Phenotyping tool. *BMC Plant Biol.*, **13**.
- Abdullah, A. Bin, Ito, S. and Adhana, K.** (2006) Estimate of rice consumption in Asian countries and the world towards 2050. In *Proceedings for Workshop and Conference on Rice in the World at Stake*. pp. 28–43.
- Ahearn, K.P., Johnson, H.A., Weigel, D. and Wagner, D.R.** (2001) NFL1, a *Nicotiana tabacum* LEAFY-like gene, controls meristem initiation and floral structure. *Plant Cell Physiol.*, **42**, 1130–1139.
- Ainsworth, C.C.** (2006) *Flowering and its manipulation*, Wiley Online Library.
- Arite, T., Umehara, M., Ishikawa, S., Hanada, A., Maekawa, M., Yamaguchi, S. and Koyuzuka, J.** (2009) d14, a strigolactone-insensitive mutant of rice, shows an accelerated outgrowth of tillers. *Plant Cell Physiol.*, **50**, 1416–1424.
- Ashikari, M., Sakakibara, H., Lin, S., et al.** (2005) Cytokinin oxidase regulates rice grain production. *Science* (80-. ), **309**, 741–745.
- Aya, K., Hobo, T., Sato-Izawa, K., Ueguchi-Tanaka, M., Kitano, H. and Matsuoka, M.** (2014) A novel AP2-type transcription factor, SMALL ORGAN SIZE1, controls organ size downstream of an auxin signaling pathway. *Plant Cell Physiol.*, **55**, 897–912.
- Barazesh, S. and McSteen, P.** (2008) Hormonal control of grass inflorescence development. *Trends Plant Sci.*, **13**, 656–662. Available at: <https://doi.org/10.1016/j.tplants.2008.09.007>.
- Benlloch, R., Berbel, A., Ali, L., Gohari, G., Millán, T. and Madueño, F.** (2015) Genetic control of inflorescence architecture in legumes. *Front. Plant Sci.*, **6**, 1–14.
- Benlloch, R., Berbel, A., Serrano-Mislata, A. and Madueño, F.** (2007) Floral initiation and inflorescence architecture: A comparative view. *Ann. Bot.*, **100**, 659–676.
- Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. and Huala, E.** (2015) The arabidopsis information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis*, **53**, 474–485.
- Bergthorsson, U., Adams, K.L., Thomason, B. and Palmer, J.D.** (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature*, **424**, 197.
- Betran, E., Emerson, J.J., Kaessmann, H. and Long, M.** (2004) Sex chromosomes and male functions: where do new genes go? *Cell Cycle*, **3**, 871–873.
- Betrán, E. and Long, M.** (2002) Expansion of genome coding regions by acquisition of new genes. *Genetica*, **115**, 65–80.
- Bleckmann, A., Weidtkamp-Peters, S., Seidel, C.A.M. and Simon, R.** (2010) Stem cell signaling in Arabidopsis requires CRN to localize CLV2 to the plasma membrane. *Plant Physiol.*, **152**, 166–176.
- Bortiri, E., Chuck, G., Vollbrecht, E., Rocheford, T., Martienssen, R. and Hake, S.** (2006) ramosa2 encodes a LATERAL ORGAN BOUNDARY domain protein that determines the fate of stem cells in branch meristems of maize. *Plant Cell*, **18**, 574–585.
- Bowman, J.L. and Eshed, Y.** (2000) Formation and maintenance of the shoot apical meristem. *Trends Plant Sci.*, **5**, 110–115.
- Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R. and Coen, E.** (1997) Inflorescence commitment and architecture in Arabidopsis. *Science* (80-. ), **275**, 80–83.
- Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M. and Simon, R.** (2000) Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. *Science* (80-. ), **289**, 617–619.
- Brewer, P.B., Dun, E.A., Ferguson, B.J., Rameau, C. and Beveridge, C.A.** (2009) Strigolactone acts downstream of auxin to regulate bud outgrowth in pea and Arabidopsis. *Plant Physiol.*, **150**, 482–493.
- Brukhin, V. and Morozova, N.** (2011) Plant growth and development - Basic knowledge and current views. *Math. Model. Nat. Phenom.*, **6**, 1–53.
- Caetano-Anollés, G., Yafremava, L.S., Gee, H., Caetano-Anollés, D., Kim, H.S. and Mittenthal, J.E.** (2009) The origin and evolution of modern metabolism. *Int. J. Biochem. Cell Biol.*, **41**, 285–297.
- Carpenter, R. and Coen, E.S.** (1990) Floral homeotic mutations produced by transposon-mutagenesis in *Antirrhinum majus*. *Genes Dev.*, **4**, 1483–1493.
- Causier, B., Schwarz-Sommer, Z. and Davies, B.** (2010) Floral organ identity: 20 years of ABCs. In *Seminars in cell & developmental biology*. Elsevier, pp. 73–79.
- Chae, E., Tan, Q.K.-G., Hill, T.A. and Irish, V.F.** (2008) An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development*, **135**, 1235–1245.
- Charlesworth, D., Liu, F.L. and Zhang, L.** (1998) The evolution of the alcohol dehydrogenase gene family by loss of introns in plants of the genus *Leavenworthia* (Brassicaceae). *Mol. Biol. Evol.*, **15**, 552–559.
- Chen, Y.H., Gols, R. and Benrey, B.** (2015) Crop Domestication and Its Impact on Naturally Selected Trophic Interactions. *Annu. Rev. Entomol.*, **60**, 35–58.

- Choi, J.Y., Platts, A.E., Fuller, D.Q., Hsing, Y.I., Wing, R.A., Purugganan, M.D. and Kim, Y.** (2017) The rice paradox: Multiple origins but single domestication in Asian Rice. *Mol. Biol. Evol.*, **34**, 969–979.
- Chu, H., Qian, Q., Liang, W., et al.** (2006) The FLORAL ORGAN NUMBER4 gene encoding a putative ortholog of Arabidopsis CLAVATA3 regulates apical meristem size in rice. *Plant Physiol.*, **142**, 1039–1052.
- Chuck, G., Cigan, A.M., Saetern, K. and Hake, S.** (2007) The heterochronic maize mutant Corngrass1 results from overexpression of a tandem microRNA. *Nat. Genet.*, **39**, 544.
- Clark, R.M., Linton, E., Messing, J. and Doebley, J.F.** (2004) Pattern of diversity in the genomic region near the maize domestication gene *tb1*. *Proc. Natl. Acad. Sci.*, **101**, 700–707.
- Clark, S.E., Williams, R.W. and Meyerowitz, E.M.** (1997) The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell*, **89**, 575–585.
- Coen, E.S. and Meyerowitz, E.M.** (1991) The war of the whorls: genetic interactions controlling flower development. *Nature*, **353**, 31.
- Cong, B., Liu, J. and Tanksley, S.D.** (2002) Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. *Proc. Natl. Acad. Sci.*, **99**, 13606–13611.
- Conti, L. and Bradley, D.** (2007) TERMINAL FLOWER1 is a mobile signal controlling Arabidopsis architecture. *Plant Cell*, **19**, 767–778.
- Crawford, S., Shinohara, N., Sieberer, T., Williamson, L., George, G., Hepworth, J., Müller, D., Domagalska, M.A. and Leyser, O.** (2010) Strigolactones enhance competition between shoot branches by dampening auxin transport. *Development*, **137**, 2905–2913.
- Daum, G., Medzihradzky, A., Suzuki, T. and Lohmann, J.U.** (2014) A mechanistic framework for noncell autonomous stem cell induction in Arabidopsis. *Proc. Natl. Acad. Sci.*, **111**, 14619–14624.
- Dehairs, J., Talebi, A., Cherifi, Y. and Swinnen, J. V.** (2016) CRISP-ID: Decoding CRISPR mediated indels by Sanger sequencing. *Sci. Rep.*, **6**, 1–5. Available at: <http://dx.doi.org/10.1038/srep28973>.
- DeYoung, B.J., Bickle, K.L., Schrage, K.J., Muskett, P., Patel, K. and Clark, S.E.** (2006) The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in Arabidopsis. *Plant J.*, **45**, 1–16.
- Doebley, J.** (1993) Genetics, development and plant evolution. *Curr. Opin. Genet. Dev.*, **3**, 865–872.
- Doebley, J., Stec, A. and Hubbard, L.** (1997) The evolution of apical dominance in maize. *Nature*, **386**, 485.
- Doebley, John F., Gaut, B.S. and Smith, B.D.** (2006) The molecular genetics of crop domestication. *Cell*, **127**, 1309–1321.
- Doebley, John F., Gaut, B.S. and Smith, B.D.** (2006) The Molecular Genetics of Crop Domestication. *Cell*, **127**, 1309–1321.
- Domagalska, M.A. and Leyser, O.** (2011) Signal integration in the control of shoot branching. *Nat. Rev. Mol. Cell Biol.*, **12**, 211.
- Drouin, G. and Dover, G.A.** (1990) Independent gene evolution in the potato actin gene family demonstrated by phylogenetic procedures for resolving gene conversions and the phylogeny of angiosperm actin genes. *J. Mol. Evol.*, **31**, 132–150.
- Edwards, K., Johnstone, C. and Thompson, C.** (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Plant Mol. Biol. Rep.*, **19**, 1991.
- Eyre-Walker, A., Gaut, R.L., Hilton, H., Feldman, D.L. and Gaut, B.S.** (1998) Investigation of the bottleneck leading to the domestication of maize. *Proc. Natl. Acad. Sci.*, **95**, 4441–4446.
- Fernie, A.R. and Yan, J.** (2019) De Novo Domestication: An Alternative Route toward New Crops for the Future. *Mol. Plant*, **12**, 615–631. Available at: <https://doi.org/10.1016/j.molp.2019.03.016>.
- Fletcher, J.C.** (2018) The CLV-WUS stem cell signaling pathway: a roadmap to crop yield optimization. *Plants*, **7**, 87.
- Fornara, F., Pařenicová, L., Falasca, G., et al.** (2004) Functional characterization of OsMADS18, a member of the API/SQUA subfamily of MADS box genes. *Plant Physiol.*, **135**, 2207–2219.
- Frary, Anne, Nesbitt, T.C., Frary, Amy, et al.** (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science (80-. )*, **289**, 85–88.
- Freeling, M.** (2009) Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu. Rev. Plant Biol.*, **60**, 433–453.
- Friml, J., Yang, X., Michniewicz, M., et al.** (2004) A PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. *Science (80-. )*, **306**, 862–865.
- Gallavotti, A., Zhao, Q., Kozuka, J., Meeley, R.B., Ritter, M.K., Doebley, J.F., Pe, M.E. and Schmidt, R.J.** (2004) The role of *barren stalk1* in the architecture of maize. *Nature*, **432**, 630.
- Gao, X., Liang, W., Yin, C., et al.** (2010) The SEPALLATA-like gene OsMADS34 is required for rice inflorescence and spikelet development. *Plant Physiol.*, **153**, 728–740.
- Gilad, Y., Rifkin, S.A. and Pritchard, J.K.** (2008) Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet.*, **24**, 408–415.

- Gonzalez, D.H.** (2016) Chapter 1 - Introduction to Transcription Factor Structure and Function. In D. H. Gonzalez, ed. *Plant Transcription Factors*. Boston: Academic Press, pp. 3–11. Available at: <http://www.sciencedirect.com/science/article/pii/B9780128008546000014>.
- Guo, Y., Han, L., Hymes, M., Denver, R. and Clark, S.E.** (2010) CLAVATA2 forms a distinct CLE-binding receptor complex regulating Arabidopsis stem cell specification. *Plant J.*, **63**, 889–900.
- Ha, C.M., Jun, J.H. and Fletcher, J.C.** (2010) Shoot apical meristem form and function. *Curr. Top. Dev. Biol.*, **91**, 103–140.
- Hake, S.** (2008) Inflorescence architecture: the transition from branches to flowers. *Curr. Biol.*, **18**, R1106–R1108.
- Han, Y., Yang, H. and Jiao, Y.** (2014) Regulation of inflorescence architecture by cytokinins. *Front. Plant Sci.*, **5**, 1–4.
- Harrop, T.W.R., Ud Din, I., Gregis, V., Osnato, M., Jouannic, S., Adam, H. and Kater, M.M.** (2016) Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *Plant J.*, **86**, 75–88.
- He, Z., Zhai, W., Wen, H., et al.** (2011) Two evolutionary histories in the genome of rice: The roles of domestication genes. *PLoS Genet.*, **7**, 1–10.
- He, Z., Zhang, H., Gao, S., Lercher, M.J., Chen, W.H. and Hu, S.** (2016) Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.*,
- Hiei, Y. and Komari, T.** (2008) Agrobacterium-mediated transformation of rice using immature embryos or calli induced from mature seed. *Nat. Protoc.*, **3**, 824–834.
- Huang, X., Kurata, N., Wei, X., et al.** (2012) A map of rice genome variation reveals the origin of cultivated rice. *Nature*, **490**, 497–501. Available at: <http://dx.doi.org/10.1038/nature11532>.
- Huang, X., Qian, Q., Liu, Z., et al.** (2009) Natural variation at the DEP1 locus enhances grain yield in rice. *Nat. Genet.*, **41**, 494.
- Ikeda-Kawakatsu, K., Maekawa, M., Izawa, T., Itoh, J.I. and Nagato, Y.** (2012) ABERRANT PANICLE ORGANIZATION 2/RFL, the rice ortholog of Arabidopsis LEAFY, suppresses the transition from inflorescence meristem to floral meristem through interaction with APO1. *Plant J.*, **69**, 168–180.
- Ikeda, K., Ito, M., Nagasawa, N., Kyoizuka, J. and Nagato, Y.** (2007) Rice ABERRANT PANICLE ORGANIZATION 1, encoding an F-box protein, regulates meristem fate. *Plant J.*, **51**, 1030–1040.
- Ikeda, K., Sunohara, H. and Nagato, Y.** (2004) Developmental Course of Inflorescence and Spikelet in Rice. *Breed. Sci.*, **54**, 147–156.
- Ikeda, M., Mitsuda, N. and Ohme-Takagi, M.** (2009) Arabidopsis WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *Plant Cell*, **21**, 3493–3505.
- Ishii, T., Xu, Y. and McCouch, S.R.** (2001) Nuclear-and chloroplast-microsatellite variation in A-genome species of rice. *Genome*, **44**, 658–666.
- Itoh, J.I., Nonomura, K.I., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., Kitano, H. and Nagato, Y.** (2005) Rice plant development: From zygote to spikelet. *Plant Cell Physiol.*, **46**, 23–47.
- Izawa, T., Konishi, S., Shomura, A. and Yano, M.** (2009) DNA changes tell us about rice domestication. *Curr. Opin. Plant Biol.*, **12**, 185–192.
- Jack, T.** (2004) Molecular and genetic mechanisms of floral control. *Plant Cell*, **16**, S1–S17.
- Jeon, J.-S., Lee, S. and An, G.** (2008) Intragenic control of expression of a rice MADS box gene OsMADS1. *Mol. Cells (Springer Sci. Bus. Media BV)*, **26**.
- Jeon, J.-S., Lee, S., Jung, K.-H., Yang, W.-S., Yi, G.-H., Oh, B.-G. and An, G.** (2000) Production of transgenic rice plants showing reduced heading date and plant height by ectopic expression of rice MADS-box genes. *Mol. Breed.*, **6**, 581–592.
- Jeong, D.-H., Park, S., Zhai, J., Gurazada, S.G.R., Paoli, E. De, Meyers, B.C. and Green, P.J.** (2011) Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant Cell*, **23**, 4185–4207.
- Jiao, Y., Wang, Y., Xue, D., et al.** (2010) Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.*, **42**, 541.
- Jin, Y.-K. and Bennetzen, J.L.** (1994) Integration and nonrandom mutation of a plasma membrane proton ATPase gene fragment within the Bsl retroelement of maize. *Plant Cell*, **6**, 1177–1186.
- Jung, H., Chung, P.J., Park, S.H., Redillas, M.C.F.R., Kim, Y.S., Suh, J.W. and Kim, J.K.** (2017) Overexpression of OsERF48 causes regulation of OsCML16, a calmodulin-like protein gene that enhances root growth and drought tolerance. *Plant Biotechnol. J.*, **15**, 1295–1308.
- Kanno, A., Nakada, M., Akita, Y. and Hirai, M.** (2007) Class B gene expression and the modified ABC model in nongrass monocots. *Sci. World J.*, **7**, 268–279.
- Kawahara, Y., la Bastide, M. de, Hamilton, J.P., et al.** (2013) Improvement of the oryza sativa nipponbare reference genome using next generation sequence and optical map data. *Rice*, **6**, 3–10.

- Kellogg, E.A.** (2007) Floral displays: genetic control of grass inflorescences. *Curr. Opin. Plant Biol.*, **10**, 26–31.
- Khanday, I., Ram Yadav, S. and Vijayraghavan, U.** (2013) Rice LHS1/UsMADS1 controls floret meristem specification by coordinated regulation of transcription factors and hormone signaling pathways. *Plant Physiol.*, **161**, 1970–1983.
- Khush, G.S.** (2001) Green revolution: the way forward. *Nat. Rev. Genet.*, **2**, 815.
- Khush, G.S.** (1997) Origin, dispersal, cultivation and variation of rice. *Oryza From Mol. to Plant*, 25–34. Available at: [http://link.springer.com/10.1007/978-94-011-5794-0\\_3](http://link.springer.com/10.1007/978-94-011-5794-0_3).
- Kimura, M.** (1991) The neutral theory of molecular evolution: a review of recent evidence. *Japanese J. Genet.*, **66**, 367–386.
- Kinoshita, A., Betsuyaku, S., Osakabe, Y., et al.** (2010) RPK2 is an essential receptor-like kinase that transmits the CLV3 signal in Arabidopsis. *Development*, **137**, 3911–3920.
- Kirchoff, B.K.** (2003) Shape matters: Hofmeister’s rule, primordium shape, and flower orientation. *Int. J. Plant Sci.*, **164**, 505–517.
- Kitomi, Y., Ito, H., Hobo, T., Aya, K., Kitano, H. and Inukai, Y.** (2011) The auxin responsive AP2/ERF transcription factor CROWN ROOTLESS5 is involved in crown root initiation in rice through the induction of OsRR1, a type-A response regulator of cytokinin signaling. *Plant J.*, **67**, 472–484.
- Kobayashi, K., Maekawa, M., Miyao, A., Hirochika, H. and Kyozyuka, J.** (2009) PANICLE PHYTOMER2 (PAP2), encoding a SEPALLATA subfamily MADS-box protein, positively controls spikelet meristem identity in rice. *Plant cell Physiol.*, **51**, 47–57.
- Koes, R.** (2008) Evolution and development of virtual inflorescences. *Trends Plant Sci.*, **13**, 1–3.
- Komatsu, M., Chujo, A., Nagato, Y., Shimamoto, K. and Kyozyuka, J.** (2003) FRIZZY PANICLE is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development*, **130**, 3841–3850. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12835399>.
- Komatsu, M., Maekawa, M., Shimamoto, K. and Kyozyuka, J.** (2001) The LAX1 and FRIZZY PANICLE 2 genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Dev. Biol.*, **231**, 364–373.
- Kondo, T., Sawa, S., Kinoshita, A., Mizuno, S., Kakimoto, T., Fukuda, H. and Sakagami, Y.** (2006) A plant peptide encoded by CLV3 identified by in situ MALDI-TOF MS analysis. *Science (80-. )*, **313**, 845–848.
- Konishi, S., Izawa, T., Lin, S.Y., Ebana, K., Fukuta, Y., Sasaki, T. and Yano, M.** (2006) An SNP caused loss of seed shattering during rice domestication. *Science (80-. )*, **312**, 1392–1396.
- Kramer, E.M., Jaramillo, M.A. and Stilio, V.S. Di** (2004) Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. *Genetics*, **166**, 1011–1023.
- Kumar, S., Stecher, G. and Tamura, K.** (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.*, **33**, 1870–1874. Available at: <https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msw054>.
- Kurakawa, T., Ueda, N., Maekawa, M., Kobayashi, K., Kojima, M., Nagato, Y., Sakakibara, H. and Kyozyuka, J.** (2007) Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature*, **445**, 652.
- Kyozyuka, J., Konishi, S., Nemoto, K., Izawa, T. and Shimamoto, K.** (1998) Down-regulation of RFL, the FLO/LFY homolog of rice, accompanied with panicle branch initiation. *Proc. Natl. Acad. Sci.*, **95**, 1979–1982.
- Kyozyuka, J. and Shimamoto, K.** (2002) Ectopic expression of OsMADS3, a rice ortholog of AGAMOUS, caused a homeotic transformation of lodicules to stamens in transgenic rice plants. *Plant Cell Physiol.*, **43**, 130–135.
- Laux, T., Mayer, K.F.X., Berger, J. and Jürgens, G.** (1996) The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development*, **122**, 87–96.
- Lee, D.Y. and An, G.** (2012) Two AP2 family genes, SUPERNUMERARY BRACT (SNB) and OsINDETERMINATE SPIKELET 1 (OsIDS1), synergistically control inflorescence architecture and floral meristem establishment in rice. *Plant J.*, **69**, 445–461.
- Lenhard, M. and Laux, T.** (1999) Shoot meristem formation and maintenance. *Curr. Opin. Plant Biol.*, **2**, 44–50.
- Li, C., Zhou, A. and Sang, T.** (2006) Rice domestication by reducing shattering. *Science (80-. )*, **311**, 1936–1939.
- Li, H., Liang, W., Yin, C., Zhu, L. and Zhang, D.** (2011) Genetic interaction of OsMADS3, DROOPING LEAF, and OsMADS13 in specifying rice floral organ identities and meristem determinacy. *Plant Physiol.*, **156**, 263–274.
- Li, X., Qian, Q., Fu, Z., et al.** (2003) Control of tillering in rice. *Nature*, **422**, 618–621.
- Li, Z.M., Zheng, X.M. and Ge, S.** (2011) Genetic diversity and domestication history of African rice (*Oryza glaberrima*) as inferred from multiple gene sequences. *Theor. Appl. Genet.*, **123**, 21–31.

- Lin, H., Wang, R., Qian, Q., et al.** (2009) DWARF27, an iron-containing protein required for the biosynthesis of strigolactones, regulates rice tiller bud outgrowth. *Plant Cell*, **21**, 1512–1525.
- Linares, O.F.** (2002a) African rice (*Oryza glaberrima*): history and future potential. *Proc. Natl. Acad. Sci.*, **99**, 16360–16365.
- Linares, O.F.** (2002b) Pnas\_99\_25\_16360.Pdf.
- Lippman, Z. and Tanksley, S.D.** (2001) Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom. *Genetics*, **158**, 413–422.
- Lippman, Zachary B., Cohen, O., Alvarez, J.P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y. and Zamir, D.** (2008) The making of a compound inflorescence in tomato and related nightshades. *PLoS Biol.*, **6**, 2424–2435.
- Lippman, Zachary B., Cohen, O., Alvarez, J.P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y. and Zamir, D.** (2008) The making of a compound inflorescence in tomato and related nightshades. *PLoS Biol.*, **6**, e288.
- Liu, C., Teo, Z.W.N., Bi, Y., Song, S., Xi, W., Yang, X., Yin, Z. and Yu, H.** (2013) A conserved genetic pathway determines inflorescence architecture in Arabidopsis and rice. *Dev. Cell*, **24**, 612–622.
- Liu, W., Xie, X., Ma, X., Li, J., Chen, J. and Liu, Y.-G.** (2015) DSDDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations. *Mol. Plant*, **8**, 1431–1433. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1674205215002609>.
- Livak, K.J. and Schmittgen** (2004) Relative quantification. *A-Z Quant. PCR*, **55**, 87–112. Available at: <http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btg157%5Cnhttp://www.gene-quantification.de/efficiency.html%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/11328886%5Cninternal-pdf://dorak-book-real-time-pcr-2006-2998043649/dorak-book-real-t>.
- Naito, Y., Hino, K., Bono, H. and Ui-Tei, K.** (2015) CRISPRdirect: Software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*, **31**, 1120–1123.
- Nakano, T., Suzuki, K., Fujimura, T. and Shinshi, H.** (2006) Genome-Wide Analysis of the ERF Gene Family. *Plant Physiol.*, **140**, 411–432.
- Londo, J.P., Chiang, Y.C., Hung, K.H., Chiang, T.Y. and Schaal, B.A.** (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 9578–9583.
- Long, M., Betrán, E., Thornton, K. and Wang, W.** (2003) The origin of new genes: glimpses from the young and old. *Nat. Rev. Genet.*, **4**, 865.
- Malcomber, S.T. and Kellogg, E.A.** (2004) Heterogeneous expression patterns and separate roles of the SEPALLATA gene LEAFY HULL STERILE1 in grasses. *Plant Cell*, **16**, 1692–1706.
- Mallo, M., Wellik, D.M. and Deschamps, J.** (2010) Hox genes and regional patterning of the vertebrate body plan. *Dev. Biol.*, **344**, 7–15.
- Martin, C., Ellis, N. and Rook, F.** (2010) Do transcription factors play special roles in adaptive variation? *Plant Physiol.*, **154**, 506–511.
- Mayer, K.F.X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G. and Laux, T.** (1998) Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell*, **95**, 805–815.
- McKim, S. and Hay, A.** (2010) Patterning and evolution of floral structures—marking time. *Curr. Opin. Genet. Dev.*, **20**, 448–453.
- McSteen, P.** (2006) Branching out: the ramosa pathway and the evolution of grass inflorescence morphology. *Plant Cell*, **18**, 518–522.
- McSteen, P. and Leyser, O.** (2005) Shoot branching. *Annu. Rev. Plant Biol.*, **56**, 353–374.
- Miura, K., Ikeda, M., Matsubara, A., Song, X.-J., Ito, M., Asano, K., Matsuoka, M., Kitano, H. and Ashikari, M.** (2010) OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat. Genet.*, **42**, 545–549. Available at: <https://doi.org/10.1038/ng.592>.
- Molina, J., Sikora, M., Garud, N., et al.** (2011) Molecular evidence for a single evolutionary origin of domesticated rice. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 8351–8356.
- Molinero-Rosales, N., Jamilena, M., Zurita, S., Gómez, P., Capel, J. and Lozano, R.** (1999) FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. *Plant J.*, **20**, 685–693.
- Morgante, M., Brunner, S., Pea, G., Fengler, K., Zuccolo, A. and Rafalski, A.** (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.*, **37**, 997.
- Morita, Y. and Kyojuka, J.** (2007) Characterization of OsPID, the rice ortholog of PINOID, and its possible involvement in the control of polar auxin transport. *Plant Cell Physiol.*, **48**, 540–549.
- Moyroud, E., Kusters, E., Monniaux, M., Koes, R. and Parcy, F.** (2010) LEAFY blossoms. *Trends Plant Sci.*, **15**, 346–352.
- Müller, R., Bleckmann, A. and Simon, R.** (2008a) The receptor kinase CORYNE of Arabidopsis transmits the

- stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *Plant Cell*, **20**, 934–946.
- Müller, R., Bleckmann, A. and Simon, R.** (2008b) The receptor kinase CORYNE of Arabidopsis transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *Plant Cell*, **20**, 934–946.
- Muños, S., Ranc, N., Botton, E., et al.** (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. *Plant Physiol.*, **156**, 2244–2254.
- Murray, J.A.H., Jones, A., Godin, C. and Traas, J.** (2012) Systems analysis of shoot apical meristem growth and development: Integrating hormonal and mechanical signaling. *Plant Cell*, **24**, 3907–3919.
- Nagasawa, N., Miyoshi, M., Kitano, H., Satoh, H. and Nagato, Y.** (1996) Mutations associated with floral organ number in rice. *Planta*, **198**, 627–633.
- Nagasawa, N., Miyoshi, M., Sano, Y., Satoh, H., Hirano, H., Sakai, H. and Nagato, Y.** (2003) SUPERWOMAN1 and DROOPING LEAF genes control floral organ identity in rice. *Development*, **130**, 705–718.
- Nardmann, J. and Werr, W.** (2006) The shoot stem cell niche in angiosperms: expression patterns of WUS orthologues in rice and maize imply major modifications in the course of mono- and dicot evolution. *Mol. Biol. Evol.*, **23**, 2492–2504.
- Nei, M., Rogozin, I.B. and Piontkivska, H.** (2000) Purifying selection and birth-and-death evolution in the ubiquitin gene family. *Proc. Natl. Acad. Sci.*, **97**, 10866–10871.
- Nesbitt, T.C. and Tanksley, S.D.** (2002) Comparative sequencing in the genus *Lycopersicon*: implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics*, **162**, 365–379.
- Nurminsky, D.I., Nurminskaya, M. V., Aguiar, D. De and Hartl, D.L.** (1998) Selective sweep of a newly evolved sperm-specific gene in *Drosophila*. *Nature*, **396**, 572.
- Ochman, H.** (2001) Lateral and oblique gene transfer. *Curr. Opin. Genet. Dev.*, **11**, 616–619.
- Ogawa-Ohnishi, M., Matsushita, W. and Matsubayashi, Y.** (2013) Identification of three hydroxyproline O-arabinosyltransferases in *Arabidopsis thaliana*. *Nat. Chem. Biol.*, **9**, 726.
- Ogawa, M., Shinohara, H., Sakagami, Y. and Matsubayashi, Y.** (2008) Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. *Science (80- )*, **319**, 294.
- Oikawa, T. and Kyozuka, J.** (2009) Two-step regulation of LAX PANICLE1 protein accumulation in axillary meristem formation in rice. *Plant Cell*, **21**, 1095–1108.
- Olsen, K.M., Caicedo, A.L., Polato, N., McClung, A., McCouch, S. and Purugganan, M.D.** (2006) Selection under domestication: evidence for a sweep in the rice waxy genomic region. *Genetics*, **173**, 975–983.
- Patthy, L.** (1996) Exon shuffling and other ways of module exchange. *Matrix Biol.*, **15**, 301–310.
- Patthy, L.** (1999) Genome evolution and the evolution of exon-shuffling—a review. *Gene*, **238**, 103–114.
- Pelaz, S., Ditta, G.S., Baumann, E., Wisman, E. and Yanofsky, M.F.** (2000) B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature*, **405**, 200.
- Peng, S., Khush, G.S., Virk, P., Tang, Q. and Zou, Y.** (2008) Progress in ideotype breeding to increase rice yield potential. *F. Crop. Res.*, **108**, 32–38.
- Pina, S. Della, Souer, E. and Koes, R.** (2014) Arguments in the evo-devo debate: say it with flowers! *J. Exp. Bot.*, **65**, 2231–2242.
- Prasad, K., Parameswaran, S. and Vijayraghavan, U.** (2005) OsMADS1, a rice MADS-box factor, controls differentiation of specific cell types in the lemma and palea and is an early-acting regulator of inner floral organs. *Plant J.*, **43**, 915–928.
- Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, Lawrence D. and Coen, E.** (2007) Evolution and development of inflorescence architectures. *Science (80- )*, **316**, 1452–1456.
- Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, Lawrence D. and Coen, E.** (2007) Evolution and development of inflorescence architectures. *Science (80- )*, **316**, 1452–1456.
- Rao, N.N., Prasad, K., Kumar, P.R. and Vijayraghavan, U.** (2008) Distinct regulatory role for RFL, the rice LFY homolog, in determining flowering time and plant architecture. *Proc. Natl. Acad. Sci.*, **105**, 3646–3651.
- Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J. and Xu, Y.** (2012) AP2/ERF transcription factor in rice: Genome-wide analysis and synteny relationships between monocots and dicots. *Evol. Bioinforma.*, **2012**, 321–355.
- Rastogi, S. and Liberles, D.A.** (2005) Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol. Biol.*, **5**, 28.
- Rebocho, A.B., Blik, M., Kusters, E., Castel, R., Procissi, A., Roobeek, I., Souer, E. and Koes, R.** (2008) Role of EVERGREEN in the development of the cymose petunia inflorescence. *Dev. Cell*, **15**, 437–447.
- Remizowa, M. V., Rudall, P.J., Choob, V. V and Sokoloff, D.D.** (2012) Racemose inflorescences of monocots: structural and morphogenetic interaction at the flower/inflorescence level. *Ann. Bot.*, **112**, 1553–1566.
- Rensing, S.A.** (2014) Gene duplication as a driver of plant morphogenetic evolution. *Curr. Opin. Plant Biol.*, **17**, 43–48. Available at: <http://dx.doi.org/10.1016/j.pbi.2013.11.002>.
- Rijkema, A.S., Vandenbussche, M., Koes, R., Heijmans, K. and Gerats, T.** (2010) Variations on a theme:

- changes in the floral ABCs in angiosperms. In *Seminars in cell & developmental biology*. Elsevier, pp. 100–107.
- Ritter, M.K., Padilla, C.M. and Schmidt, R.J.** (2002) The maize mutant barren stalk1 is defective in axillary meristem development. *Am. J. Bot.*, **89**, 203–210.
- Robert, W.** (1982) Inflorescence Architecture : How Flower Number , Arrangement , and Phenology Affect Pollination and Fruit-Set '. *Am. J. Bot.*, **69**, 585–594.
- Ross-Ibarra, J., Morrell, P.L. and Gaut, B.S.** (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proc. Natl. Acad. Sci.*, **104**, 8641 LP – 8648. Available at: [http://www.pnas.org/content/104/suppl\\_1/8641.abstract](http://www.pnas.org/content/104/suppl_1/8641.abstract).
- Samach, A., Klenz, J.E., Kohalmi, S.E., Risseuw, E., Haughn, G.W. and Crosby, W.L.** (1999) The UNUSUAL FLORAL ORGANS gene of Arabidopsis thaliana is an F-box protein required for normal patterning and growth in the floral meristem. *Plant J.*, **20**, 433–445.
- Sarla, N. and Swamy, B.P.M.** (2005) Oryza glaberrima: a source for the improvement of Oryza sativa. *Curr. Sci.*, 955–963.
- Sato, Y., Takehisa, H., Kamatsuki, K., et al.** (2013) RiceXPro Version 3.0: Expanding the informatics resource for rice transcriptome. *Nucleic Acids Res.*, **41**, 1206–1213.
- Satoh-Nagasawa, N., Nagasawa, N., Malcomber, S., Sakai, H. and Jackson, D.** (2006) A trehalose metabolic enzyme controls inflorescence architecture in maize. *Nature*, **441**, 227.
- Schmitz, G. and Theres, K.** (2005) Shoot and inflorescence branching. *Curr. Opin. Plant Biol.*, **8**, 506–511.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F.X., Jürgens, G. and Laux, T.** (2000) The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell*, **100**, 635–644.
- Second, B.Y.G.** (1982) Origin of the genic diversity. *Africa (Lond)*.
- Shannon, S. and Meeks-Wagner, D.R.** (1991) A mutation in the Arabidopsis TFL1 gene affects inflorescence meristem development. *Plant Cell*, **3**, 877–892.
- Sharoni, A.M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., Choi, I.R., Omura, T. and Kikuchi, S.** (2011) Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol.*, **52**, 344–360.
- Shinohara, H. and Matsubayashi, Y.** (2015) Reevaluation of the CLV 3-receptor interaction in the shoot apical meristem: dissection of the CLV 3 signaling pathway from a direct ligand-binding point of view. *Plant J.*, **82**, 328–336.
- Slack, F. and Ruvkun, G.** (1998) Heterochronic genes in development and evolution. *Biol. Bull.*, **195**, 375–376.
- Soltis, D.E., Ma, H., Frohlich, M.W., Soltis, P.S., Albert, V.A., Oppenheimer, D.G., Altman, N.S., depamphilis, C. and Leebens-Mack, J.** (2007) The floral genome: an evolutionary history of gene duplication and shifting patterns of gene expression. *Trends Plant Sci.*, **12**, 358–367.
- Souer, E., Krol, A. van der, Kloos, D., Spelt, C., Blied, M., Mol, J. and Koes, R.** (1998) Genetic control of branching pattern and floral identity during Petunia inflorescence development. *Development*, **125**, 733–742.
- Souer, E., Rebocho, A.B., Blied, M., Kusters, E., Bruin, R.A.M. de and Koes, R.** (2008) Patterning of inflorescences and flowers by the F-Box protein DOUBLE TOP and the LEAFY homolog ABERRANT LEAF AND FLOWER of petunia. *Plant Cell*, **20**, 2033–2048.
- Stebbins, G.L.** (1974) *Flowering plants: evolution above the species level*.
- Stein, J.C., Yu, Y., Copetti, D., et al.** (2018) Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. *Nat. Genet.*, **50**, 285–296. Available at: <http://dx.doi.org/10.1038/s41588-018-0040-0>.
- Suzaki, T., Ohneda, M., Toriba, T., Yoshida, A. and Hirano, H.-Y.** (2009) FON2 SPARE1 redundantly regulates floral meristem maintenance with FLORAL ORGAN NUMBER2 in rice. *PLoS Genet.*, **5**, e1000693.
- Suzaki, T., Sato, M., Ashikari, M., Miyoshi, M., Nagato, Y. and Hirano, H.-Y.** (2004) The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to Arabidopsis CLAVATA1. *Development*, **131**, 5649–5657.
- Suzaki, T., Sato, M., Ashikari, M., Miyoshi, M., Nagato, Y. and Hirano, H.-Y.** (2004) The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to Arabidopsis CLAVATA1. *Development*, **131**, 5649–5657.
- Suzaki, T., Toriba, T., Fujimoto, M., Tsutsumi, N., Kitano, H. and Hirano, H.-Y.** (2006) Conservation and diversification of meristem maintenance mechanism in Oryza sativa: function of the FLORAL ORGAN NUMBER2 gene. *Plant Cell Physiol.*, **47**, 1591–1602.
- Suzaki, T., Yoshida, A. and Hirano, H.-Y.** (2008) Functional diversification of CLAVATA3-related CLE proteins in meristem maintenance in rice. *Plant Cell*, **20**, 2049–2058.
- Sweeney, M.T., Thomson, M.J., Pfeil, B.E. and McCouch, S.** (2006) Caught red-handed: Rc encodes a basic

- helix-loop-helix protein conditioning red pericarp in rice. *Plant Cell*, **18**, 283–294.
- Ta, K.N., Adam, H., Staedler, Y.M., et al.** (2017) Differences in meristem size and expression of branching genes are associated with variation in panicle phenotype in wild and domesticated African rice. *Evodevo*, **8**, 1–14.
- Ta, K.N., Sabot, F., Adam, H., Vigouroux, Y., Mita, S. De, Ghesquière, A., Do, N. V., Gantet, P. and Jouannic, S.** (2016) miR2118-triggered phased siRNAs are differentially expressed during the panicle development of wild and domesticated African rice species. *Rice*, **9**.
- Tabuchi, H., Zhang, Y., Hattori, S., et al.** (2011) LAX PANICLE2 of rice encodes a novel nuclear protein and regulates the formation of axillary meristems. *Plant Cell*, **23**, 3276–3287.
- Takeda, T., Suwa, Y., Suzuki, M., Kitano, H., Ueguchi-Tanaka, M., Ashikari, M., Matsuoka, M. and Ueguchi, C.** (2003) The OsTB1 gene negatively regulates lateral branching in rice. *Plant J.*, **33**, 513–520.
- Tanaka, W., Ohmori, Y., Ushijima, T., Matsusaka, H., Matsushita, T., Kumamaru, T., Kawano, S. and Hirano, H.-Y.** (2015) Axillary Meristem Formation in Rice Requires the WUSCHEL Ortholog TILLERS ABSENT1. *Plant Cell*, **27**, 1173–84. Available at: <http://www.plantcell.org/content/27/4/1173.full>.
- Tang, H., Sezen, U. and Paterson, A.H.** (2010) Domestication and plant genomes. *Curr. Opin. Plant Biol.*, **13**, 160–166.
- Theissen, G. and Saedler, H.** (2001) Plant biology: floral quartets. *Nature*, **409**, 469.
- Vaughan, D.A., Lu, B.R. and Tomooka, N.** (2008) The evolving story of rice evolution. *Plant Sci.*, **174**, 394–408.
- Veron, A.S., Kaufmann, K. and Bornberg-Bauer, E.** (2006) Evidence of interaction network evolution by whole-genome duplications: a case study in MADS-box proteins. *Mol. Biol. Evol.*, **24**, 670–678.
- Vollbrecht, E., Springer, P.S., Goh, L., Buckler IV, E.S. and Martienssen, R.** (2005) Architecture of floral branch systems in maize and related grasses. *Nature*, **436**, 1119.
- Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A.M.** (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.*, **35**, 71–74.
- Waese, J., Fan, J., Pasha, A., et al.** (2017) ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell*, **29**, 1806–1821.
- Wambugu, P.W., Brozynska, M., Furtado, A., Waters, D.L. and Henry, R.J.** (2015) Relationships of wild and domesticated rices (*Oryza* AA genome species) based upon whole chloroplast genome sequences. *Sci. Rep.*, **5**, 1–9. Available at: <http://dx.doi.org/10.1038/srep13957>.
- Wang, H., Nussbaum-Wagler, T., Li, B., Zhao, Q., Vigouroux, Y., Faller, M., Bomblies, K., Lukens, L. and Doebley, J.F.** (2005) The origin of the naked grains of maize. *Nature*, **436**, 714.
- Wang, M., Yu, Y., Haberer, G., et al.** (2014) The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nat. Genet.*, **46**, 982–988. Available at: <http://www.nature.com/doi/10.1038/ng.3044> <http://dx.doi.org/10.1038/ng.3044>.
- Wang, R.-L., Stec, A., Hey, J., Lukens, L. and Doebley, J.** (1999) The limits of selection during maize domestication. *Nature*, **398**, 236.
- Wang, W., Brunet, F.G., Nevo, E. and Long, M.** (2002) Origin of sphinx, a young chimeric RNA gene in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.*, **99**, 4448–4453.
- Wang, Y. and Li, J.** (2011) Branching in rice. *Curr. Opin. Plant Biol.*, **14**, 94–99.
- Wang, Y. and Li, J.** (2008) Molecular basis of plant architecture. *Annu. Rev. Plant Biol.*, **59**, 253–279.
- Wang, Z., Zheng, F., Shen, G., Gao, J., Snustad, D.P., Li, M., Zhang, J. and Hong, M.** (1995) The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene. *Plant J.*, **7**, 613–622.
- Waters, M.T., Brewer, P.B., Bussell, J.D., Smith, S.M. and Beveridge, C.A.** (2012) The Arabidopsis ortholog of rice DWARF27 acts upstream of MAX1 in the control of plant development by strigolactones. *Plant Physiol.*, **159**, 1073–1085.
- Weberling, F.** (1989) Morphology of flowers and inflorescences. (Cambridge University Press: Cambridge, UK).
- Weigel, D. and Meyerowitz, E.M.** (1994) The ABCs of floral homeotic genes. *Cell*, **78**, 203–209.
- Whitewoods, C.D., Cammarata, J., VENZA, Z.N., et al.** (2018) CLAVATA was a genetic novelty for the morphological innovation of 3D growth in land plants. *Curr. Biol.*, **28**, 2365–2376.
- Woods, D.P., Hope, C.L. and Malcomber, S.T.** (2011) Phylogenomic analyses of the BARREN STALK1/LAX PANICLE1 (BA1/LAX1) genes and evidence for their roles during axillary meristem development. *Mol. Biol. Evol.*, **28**, 2147–2159.
- Wray, G.A.** (2007) The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.*, **8**, 206.
- Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M. V and Romano, L.A.** (2003) The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.*, **20**, 1377–1419.
- Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D. and Gaut, B.S.** (2005) The effects of artificial selection on the maize genome. *Science (80- )*, **308**, 1310–1314.

- Wright, S.I. and Gaut, B.S.** (2004) Molecular population genetics and the search for adaptive evolution in plants. *Mol. Biol. Evol.*, **22**, 506–519.
- Xie, K., Minkenberg, B. and Yang, Y.** (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci.*, **112**, 3570–3575. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1420294112>.
- Xing, Y. and Zhang, Q.** (2010) Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol.*, **61**, 421–442.
- Xu, C., Liberatore, K.L., MacAlister, C.A., et al.** (2015) A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nat. Genet.*, **47**, 784.
- Xu, M., Zhu, L., Shou, H. and Wu, P.** (2005) A PIN1 family gene, OsPIN1, involved in auxin-dependent adventitious root emergence and tillering in rice. *Plant Cell Physiol.*, **46**, 1674–1681.
- Yadav, R.K., Perales, M., Gruel, J., Girke, T., Jönsson, H. and Reddy, G.V.** (2011) WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. *Genes Dev.*, **25**, 2025–2030.
- Yamaki, S., Miyabayashi, T., Eiguchi, M., Kitano, H., Nonomura, K.-I. and Kurata, N.** (2010) Diversity of panicle branching patterns in wild relatives of rice. *Breed. Sci.*, **60**, 586–596.
- Yamasaki, K.** (2016). Structures, Functions, and Evolutionary Histories of DNA-Binding Domains of Plant-Specific Transcription Factors. In *Plant Transcription Factors* (pp. 57-72). Academic Press.
- Yamburenko, M. V., Kieber, J.J. and Schaller, G.E.** (2017) Dynamic patterns of expression for genes regulating cytokinin metabolism and signaling during rice inflorescence development. *PLoS One*, **12**, 1–18.
- Yoshida, A., Sasao, M., Yasuno, N., et al.** (2013) TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proc. Natl. Acad. Sci.*, **110**, 767–772. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1216151110>.
- Yoshida, H. and Nagato, Y.** (2011) Flower development in rice. *J. Exp. Bot.*, **62**, 4719–4730.
- Zalewski, W., Galuszka, P., Gasparis, S., Orczyk, W. and Nadolska-Orczyk, A.** (2010) Silencing of the HvCKX1 gene decreases the cytokinin oxidase/dehydrogenase level in barley and leads to higher plant productivity. *J. Exp. Bot.*, **61**, 1839–1851.
- Zažimalová, E., Křeček, P., Skůpa, P., Hoyerova, K. and Petrášek, J.** (2007) Polar transport of the plant hormone auxin—the role of PIN-FORMED (PIN) proteins. *Cell. Mol. Life Sci.*, **64**, 1621–1637.
- Zhang, D. and Yuan, Z.** (2014a) Molecular control of grass inflorescence development. *Annu. Rev. Plant Biol.*, **65**, 553–578.
- Zhang, D. and Yuan, Z.** (2014b) Molecular Control of Grass Inflorescence Development.
- Zhang, L. Bin, Zhu, Q., Wu, Z.Q., Ross-Ibarra, J., Gaut, B.S., Ge, S. and Sang, T.** (2009) Selection on grain shattering genes and rates of rice domestication. *New Phytol.*, **184**, 708–720.
- Zhang, L., Zhao, Y., Gao, L., Zhao, G., Zhou, R., Zhang, B. and Jia, J.** (2012) TaCKX6-D1, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. *New Phytol.*, **195**, 574–584.
- Zhu, Y., Wang, Y., Li, R., Song, X., Wang, Q., Huang, S., Jin, J.B., Liu, C. and Lin, J.** (2010) Analysis of interactions among the CLAVATA3 receptors reveals a direct interaction between CLAVATA2 and CORYNE in Arabidopsis. *Plant J.*, **61**, 223–233.

## LIST OF PUBLICATIONS

**Harrop, T.W.R., Mantegazza, O., Luong, A.M., Béthune, K., Lorieux, M., Jouannic, S. and Adam, H.** (2019) A set of AP2-like genes is associated with inflorescence branching and architecture in domesticated rice. *Journal of Experimental Botany*. doi: 10.1093/jxb/erz340

## SCIENTIFIC COMMUNICATION

- 16th International Symposium on Rice Functional Genomics, September 5~7, 2018, Tokyo University Agriculture, Setagaya, Tokyo, Japan, Poster presentation
- PhD Day in IRD Montpellier, 2017, 2018, 2019, oral presentations
- Journée "Des avancées sur les recherches en biologie et génétique au service de la riziculture" organisée par l'Académie d'Agriculture de France, 06 juin 2018, Agropolis, Montpellier, Poster presentation

# Récapitulatif de participation aux Formations Ai My LUONG

**Doctorat** : Génétique et amélioration des plantes

**Ecole Doctorale** : GAIA - Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

**Etablissement** : Université de Montpellier

Date de la 1<sup>ere</sup> inscription en thèse : 1 novembre 2016 (4 A en 2019)

**Directeur de thèse** : Stéphane JOUANNIC

**Sujet de thèse** : Impact des facteurs de transcription AP2 / ERF sur l'évolution de la structure de la panicule de riz

## **Formations suivies**

### **Catégorie : Enseignement à distance**

➤ MOOC Doctorat et poursuite de carrière (07 novembre 2018) En ligne  
24 heures

➤ MOOC Ethique de la recherche (06 septembre 2018) En ligne  
20 heures

Total du nombre d'heures pour la catégorie Enseignement à distance : 44 h

### **Catégorie : Formations scientifiques interdisciplinaires**

➤ Statistique pour expérimentateurs (03 avril 2017) salle 167 campus St priest Bat 2  
25 heures enregistrées par : I2S - Information, Structures, Systèmes.

Total du nombre d'heures pour la catégorie Formations scientifiques interdisciplinaires : 25 h

### **Catégorie : Journée Doctorale**

➤ Présentation de l'ED GAIA pour les 1<sup>ieres</sup> années (12 janvier 2017) Bât 23 salle 23.01  
2 heures enregistrées par : GAIA - Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau.

Total du nombre d'heures pour la catégorie Journée Doctorale : 2 h

### **Catégorie : Langues vivantes**

➤ FLE - Français Langue Étrangère (session Septembre 2018) (01 octobre 2018) Université de Montpellier, Département des Langues  
Bâtiment 5  
22 heures Note : B2

➤ Conversation en Français pour les doctorants étrangers (23 janvier 2018) Université de Montpellier Site Nord Triolet bâtiment 3, salle  
collège doctoral  
24 heures

➤ Cours Allemand Niveau débutant (22 janvier 2018) Université de Montpellier Site Nord Triolet Bâtiment 3, salle collège doctoral  
13.3 heures

Total du nombre d'heures pour la catégorie Langues vivantes : 59.3 h

### **Catégorie : Méthodes, outils, langages**

➤ FLE - Français Langue Étrangère 2017 (07 février 2017) Université de Montpellier, Département des Langues Bâtiment 5  
30 heures enregistrées par : Collège Doctoral Languedoc Roussillon.

➤ Pour des visuels de présentation orale pertinents et percutants (24 mars 2017) salle 050 - Bât.1 - Campus st Priest, 860 rue de Saint Priest,  
34090 Montpellier  
7 heures enregistrées par : Collège Doctoral Languedoc Roussillon.

Total du nombre d'heures pour la catégorie Méthodes, outils, langages : 37 h

### **Catégorie : Outils et méthodes**

➤ Writing and defense a PhD (11 février 2019) salle de conférence Institut Européen des Membranes  
9 heures

Total du nombre d'heures pour la catégorie Outils et méthodes : 9 h

### **Catégorie : Ouverture scientifique et culturelle**

➤ Rentrée du Collège Doctoral Languedoc-Roussillon (15 décembre 2016) Maison des Etudiants Aimé Schoenig, Site universitaire  
RICHTER, Montpellier

5 heures enregistrées par : Collège Doctoral Languedoc Roussillon.

Total du nombre d'heures pour la catégorie : 5 h

**Total participation : 181.3 heures / 11 modules**



# Impact of AP2/ERF transcription factors on rice panicle development



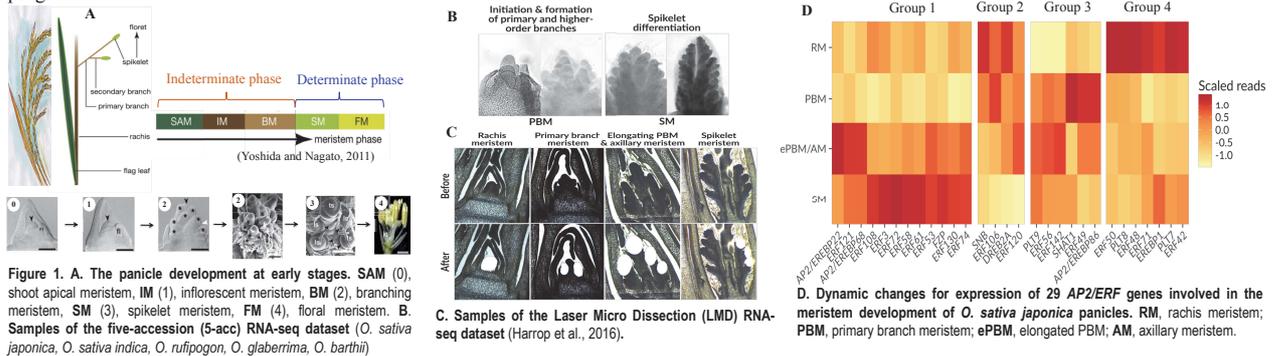
LUONG Ai My<sup>(1)</sup>, MANTEGAZZA Otho<sup>(1)</sup>, HARROP Thomas<sup>(1)</sup>, TREGEAR James<sup>(1)</sup>, ADAM Helene<sup>(1)</sup>, KHONG Ngan Giang<sup>(2)</sup>, LEBRUN Michel<sup>(2,3)</sup>, JOUANNIC Stefan<sup>(1)</sup>



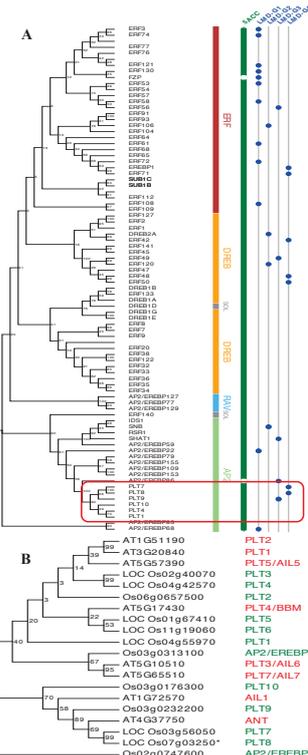
<sup>(1)</sup> EDI team, UMR DIADE, IRD Montpellier, France; <sup>(2)</sup> LMI RICE, Vietnam; <sup>(3)</sup> UMR LSTM, IRD; ✉ stephane.jouannic@ird.fr



**Introduction** In plants, *AP2/ERF* genes encode transcriptional regulators involved in many developmental and physiological processes. An *in silico* analysis of two RNA-seq datasets from different rice species led to the identification of 85 *AP2/ERF* genes expressed during rice panicle development at early stages. Some of these genes might play a role in determining the inter-specific diversity of rice panicle development. For a better understanding of panicle development regulation to improve the rice yield, we characterize *AP2/ERF* genes of potential interest. Another purpose will be to enhance the knowledge of molecular mechanisms involved in the meristem establishment during panicle development, this being a key determinant of panicle architecture. The study results will allow identifying the key factors of rice panicle growth that could be targeted for improvement in breeding programs.

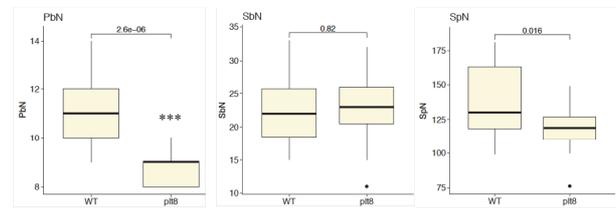


## 1. In silico analysis of AP2/ERF genes expressed in the early development of panicle



- ✓ 85 *AP2/ERF* genes expressed in both RNA-seq datasets
- ✓ The *AP2/ERF* genes expressed in panicle do not belong to any specific subfamily
- ✓ These genes were differentially expressed in five *Oryza* species and in four types of panicle meristems
- ✓ Six *PLETHORA*-like (*PLT*) genes expressed in rice panicle
- ✓ No report about the effects of *PLT* genes on rice panicle architecture and development

## 2. Panicles of crown rootless 5 (*clr5*) mutant

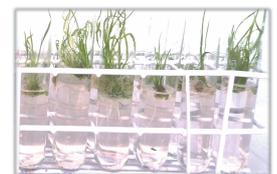


**Figure 3.** Panicle phenotype of *clr5/plt8* mutant. The panicle morphological traits were scored using P-TRAP software (Al Tam et al. 2013). Pbn, Primary branch Number, SbN, Secondary branch Number, SpN, Spikelet Number. Scale bar = 10 cm.

- ✓ *CRL5/PLT8* gene promotes crown root initiation (Kitomi et al., 2011)
- ✓ *CRL5/PLT8* affects also the number of primary branches but not the number of secondary branches

## 3. Functional analysis of PLT genes

- ✓ We generate knock-out mutants of *PLT7*, *PLT8*, *PLT9* and *AP2/EREBP22* genes by using CRISPR/Cas9, as well as over-expressing lines.
- ✓ The functional characterization of the transgenic plants is on going.



**Figure 4.** The mutant plants at the rooting stage

## Conclusion & Perspectives

Our work will determine to what extent *AP2/ERF* and in particular *PLT* genes may be involved in panicle development and its branching diversity and how they were impacted by domestication in Asia and Africa. One of the perspectives in short-term is to finalize the functional analysis of the interesting genes and to study how these genes contribute to rice yield and its plasticity under changing environment. A longer-term perspective is to integrate this knowledge to breeding programs.

**Acknowledgements** We thank LMI-RICE and Le Quang Hoa at HUST (Hanoi, Vietnam) for helping to generate the vector constructs of CRISPR/Cas9 system, Jean-Luc Verdeil and Genevieve Conjevo for assistance with laser microdissection at the PHIV<sup>2</sup> imaging platform (CIRAD, Montpellier, France) and Prof. Yoshiaki Inukai from Nagoya University for providing the *clr5/plt8* and kinmaze seeds. Ai-My Luong's Ph.D. fellowship was funded by the French Embassy of France in Vietnam and IRD (ARTS program).

**References** (1) Yoshida and Nagato, 2011 *J Exp Bot*, 62:4719-4730. (2) Harrop et al., 2016 *Plant J*, 86:75-88. (3) Al Tam et al., 2013 *BMC Plant Biol*, 13:122. (4) Kitomi et al., 2011. *Plant J*, 67:472-484.