Functional diversity of genes involved in rice panicle architecture

Kim Nhung Ta

Collège Doctoral du Languedoc-Roussillon

THÈSE

Pour obtenir le grade de

Docteur

Délivré par **Université Montpellier 2**

Préparée au sein de l'école doctorale SIBAGHE
Et de l'unité de recherche DIADE

Spécialité : **Biologie Intégrative des Plantes**

Présentée par **TA Kim Nhung**

Diversité fonctionnelle des gènes impliqués dans le contrôle de l'architecture paniculaire chez le riz

Soutenue le 05 décembre 2014 devant le jury composé de

Mr Patrick LAUFS, DR, INRA	Rapporteur
Mr Michiel VANDENBUSSCHE, CR, CNRS	Rapporteur
Mme Catherine DAMERVAL, DR, CNRS	Examinateur
Mr Michel LEBRUN, Professeur, UM2	Examinateur
Mme Anne DIEVART, CR, CIRAD	Invitée
Mr Stéphane JOUANNIC, CR, IRD	Co-directeur de thèse
Mr Pascal GANTET, Professeur, UM2	Directeur de thèse

UNIVERSITÉ MONTPELLIER 2
SCIENCES ET TECHNIQUES

UNIVERSITÉ MONTPELLIER 2

Ecole Doctorale : Systémes intégrés en Biologie, Agronomie, Géosciences, Hydrosciences et Environnement Spécialité

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER 2

Spécialité : BIP – Biologie Intégrative des Plantes – UM2

Kim Nhung TA

Functional diversity of genes involved in rice panicle architecture

Laboratoire : IRD 1097 – DIADE – Diversité et Adaptation et Développement des plates
Equipe Genetique et genomique du riz, DIADE
BP 64501, 911 avenue Agropolis, F-34394 Montpellier Cedex 5, France

Soutenue publiquement le 05 decembre 2014 devant le juri composé de :

Dr. Patrick LAUFS	**INRA**	**Rapporteur**
Dr. Michiel VANDENBUSSCHE	**CNRS**	**Rapporteur**
Prof. Michel LEBRUN	**Université Montpellier 2**	**Examinateur**
Dr. Catherine DAMERVAL	**CNRS**	**Examinateur**
Dr. Anne DIEVART	**CIRAD**	**Invitée**
Dr. Stéphane JOUANNIC	**IRD**	**Co-directeur de Thèse**
Prof. Pascal GANTET	**Université Montpellier 2**	**Directeur de Thèse**

Acknowledgments

First and foremost, I would like to show my greatest appreciation to my supervisor, Dr. Stéphane 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. Hélène Adam for her constant encouragement, and expertise. Her willingness and optimistic cheer me up from the first day I went to IRD. I feel motivated and encouraged every time I attend the meeting with Stéphane and Hélène. Without their encouragement and guidance this project would not have been possible. I also wish to express my gratitude to professor Pascal Gantet for his supports and the opportunity he gave me to do my PhD in UM2 and IRD. I am indebted to the IRD centre, Vietnamese Government (program 322) and the Global Rice Science Partnership (GRiSP) for providing research facilities and funding this work.

Also, I would like to take this opportunity to thank to the Genome and Development of Rice (GDR-IRD France) team 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. Francoi Sabot for his help in bio-informatic analysis. I want to thank Shopie Cherron and Harold Chrestin for plant care in greenhouse. I also need to thank to M.Colin, a kindhearted woman who not only give me the technical advices for histological work but also help me get acquainted with IRD laboratory from the beginning. Besides, we would like to thank to Dr. Brigitte Courois for her enthusiasm and generous advice to do phenotyping analysis. The guidances 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 Nguyen Thi Thu Thuy, Tran Tuan Tu, Phung Thi Phuong Nhung, Nguyen Vu Phong, Nguyen Van Anh, Tran Van Giang, Nguyen Kim Hanh, Hoang Thanh Minh, Amandine Doud and other friends. I am very happy for all the smiles and tears 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.

Abstract

Rice panicle architecture is one of the most important morphological traits specifying rice yield potential, which was under selection during rice domestication. A panicle is a branched structure composed of a rachis, primary branches, higher order branches (*i.e.* secondary and tertiary branches) and finally spikelets. This morphology, depending on the activity of axillary meristems during its development, shows a wide diversity in both inter-specific (i.e. crops vs. wild-relatives) and intra-specific (Asian or/and African rice) levels. Several important genes/QTLs have been characterized in *Oryza sativa* as controlling panicle architecture by regulating meristem fate, cell division and hormone signaling. However the mechanisms related to rice panicle diversity and its evolution in the context of domestication are still largely unknown. During my PhD, I mainly investigated the histological and molecular bases of panicle diversity between the African species *Oryza glaberrima* and its wild-relative *Oryza barthii*. I analyzed the expression patterns of orthologs of *O. sativa* landmark genes related to panicle development and was involved in small RNA transcriptomic analysis in early stages of panicle development. This work revealed a high conservation of the spatial expression pattern of the landmarks genes studied but have highlighted a differential timing and level of the expression of these genes during the panicle development between two species. The genes promoting meristem activity were upper-accumulated over a longer period during the panicle development in the crop species, whereas the gene promoting spikelet/floret meristem fate behaved in opposite way. This work also has shown similar heterochronic alteration of the expression of members of the *miR2118*-triggered 21-nt phased siRNA pathway, known to be involved in male gametogenesis. Together, these findings suggest that variation of panicle complexity in African rice may rely on heterochronic changes in branching activity as well as spikelet/floret meristem determinacy.

Key words: panicle, branching, meristem fate, African rice, domestication, evo-devo

Résumé

L'architecture de la panicule de riz est l'un des caractères morphologiques majeurs du potentiel de rendement, sélectionné lors de sa domestication. Une panicule est une structure ramifiée, composée d'un axe principal (rachis), de branches primaires, et d'ordres supérieurs de branchement (branches secondaires et tertiaires) et enfin les épillets. Cette structure, qui dépend de l'activité des méristèmes axillaires au cours du développement de la panicule, montre une grande diversité à la fois inter-spécifique (espèces cultivées vs espèces sauvages apparentées) et intra-spécifiques (riz asiatique et / ou africain). Plusieurs gènes/QTL importants ont été caractérisés chez *Oryza sativa* pour le contrôle de l'architecture de la panicule en régulant l'identité des méristèmes, la division cellulaire et la signalisation hormonale. Cependant, les mécanismes liés à la diversité de la panicule de riz et son évolution dans le contexte de la domestication sont encore largement inconnus. Durant ma thèse, j'ai principalement contribué à l'étude des bases histologique et moléculaires de la diversité de la panicule entre l'espèce africaine *Oryza glaberrima* et *Oryza barthii*, l'espèce sauvage apparentée. J'ai analysé les profils d'expression d'orthologues à des gènes de *O. sativa* liés au développement de la panicule et participé à l'analyse transcriptomique de petits ARN dans les premiers stades de développement de la panicule. Ce travail a révélé une différence de période d'initiation et de niveau d'expression de ces gènes au cours du développement de la panicule entre les deux espèces, conjointement avec une forte conservation de leurs domaines d'expression. Les gènes qui favorisent l'activité des méristèmes sont sur-accumulés sur une période plus longue au cours du développement de la panicule chez l'espèce cultivée, tandis que les gènes liés au développement des épillets se comportent de manière opposée. Ces travaux ont également montré une altération similaire de l'expression des membres de la voie de siRNA phasés initiés par *miR2118*, voie connue pour être impliquée dans la gamétogenèse mâle. L'ensemble de ces résultats suggère que la diversité de complexité de la panicule chez les riz africains reposerait sur des altérations hétérochroniques de l'activité de ramification et de déterminisme des méristèmes d'épillets.

Mots-clés : panicule, ramification, identité méristématique, riz africain, domestication, evo-devo

TABLE OF CONTENTS

1.

INTRODUCTION

1. INTRODUCTION

1.1 Evolution – developmental biology

During evolution, plants give rise to a staggering complexity of morphological structures with different shape, colors, and functions. However, all plants have a common ancestor: a single eukaryotic, which acquired a photosynthetic cyanobacterium as an endosymbiont (the ancestral plastid). The plant kingdom could 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) (Figure 1.1). The first land plant (liverworth) appeared around 450 million years ago in the Orodovician period. In early Devonian-age rocks, 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). The 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 angiosperms plant (*Phylum Anthophyta*) arose (Figure 1.1). Angiosperms also were known as flowering plant because they defined typically characteristic differences from other groups of 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 eudicots and monocots species) became the world's dominant plants that today occupy almost every habitats on earth with approximately 235 000 species (Figure 1.1) (Edwards 2000; Bowman et al. 2007; Soltis et al. 2008). This species diversification makes angiosperms evolution is the most fascinating question in biology.

To gain insights into the morphological diversity of angiosperm, it is essential to understand the evolution of mechanisms underlying the developmental process that was known as "Evo-Devo" – evolutionary developmental biology. The key question in Evo-Devo is how DNA sequence changes are related 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 at a global scale. Numerous theories were proposed to explain diversification and speciation (Slack and Ruvkun 1998; Arthur 2002; Koes 2008; Carroll 2008). The main statements widely accepted were: (i) the functional divergence of duplicated gene (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).

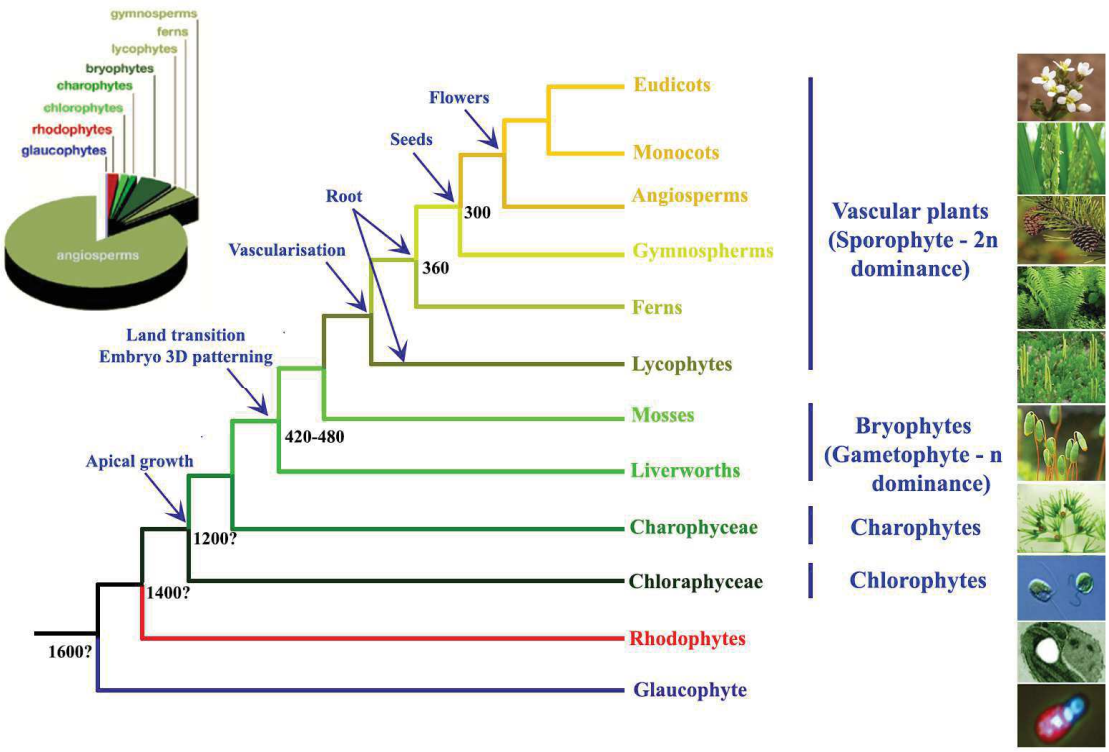


Figure 1.1 Phylogenetic and morphological innovations among plants. Depicted are relationships among the three lineages of plants: glaucophytes (freshwater algae; blue), rhodophytes (red algae; red), and the green plants (chlorophytes, charophytes, and land plants; from green to orange). Estimated dates 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 are angiosperms (Adapted from Bowman et al. 2007).

1.1.1 Gene duplication as a driving force for evolution

Gene duplication (*i.e.* paralogous genes) has been indicated as an important process in the generation of evolution novelty. The importance of gene duplication by doubling a chromosomal band in a mutant of the fruit fly *Drosophila melanogaster*, which exhibited extreme reduction in eye size has been recognized since 1936 by Bridges (Bridges 1936). In 1970, Ohno et al. hypothesized a significant role for gene duplication in the evolution of biological complexity. Genes can duplicate at single-gene, chromosome, and whole genome level (Freeling 2009). Many innovations in metabolic networks come from duplications of genes encoding enzymes (Caetano-Anollés et al. 2009). Whereas, a whole genome duplicate might create larger-scale change in molecular network than a single-gene duplication might. One example of whole genome duplication is MADS-box proteins that present the

evolution of a protein–protein interaction network of transcription factors in several plant species (Veron et al. 2007).

After a duplication event, genes can either be lost or retained in the population of the species. If a new allele contains duplicate genes 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 functions of the duplicate genes. The birth and death of genes are a common theme in gene family and genome evolution (Nei et al. 2000) with those genes involved in the physiologies that vary greatly among species (*i.e.* immunity, reproduction and sensory systems) probably having high rates of gene birth and death. Pseudogenization or non-functionalization is a purely neutral process that ultimately eliminates one of the duplicated copies as a functional gene 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 population. Neo-functionalization is an adaptive process where 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 constraint (Kramer et al. 2004; Rastogi and Liberles 2005; Freeling 2009) (Figure 1.2).

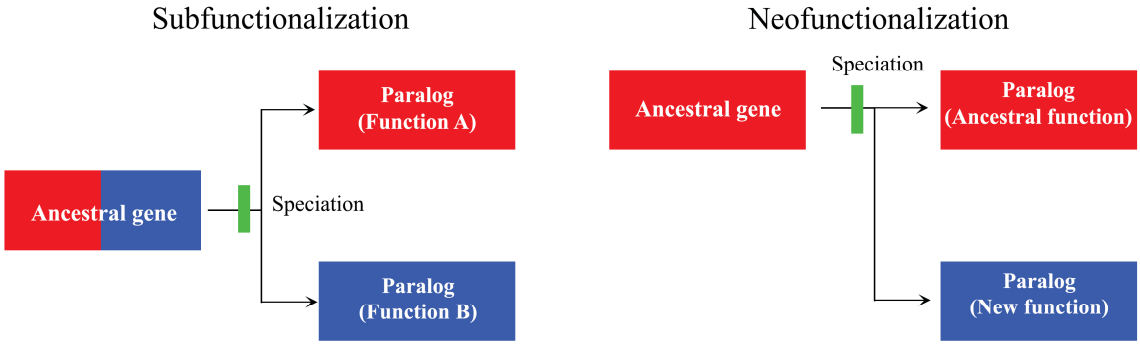


Figure 1.2 Schematic representation biogenesis of sub- and neo-functionalization

1.1.2 Expression pattern divergence of conserved genes

Although it is widely accepted that morphological variation between organisms arose from genetic alterations, the molecular mechanisms supporting these variations remain poorly understood. Nevertheless in many examples it was found to be due mostly to the variation of expression pattern of functionally conserved genes than through the emergence of new genes and functions (Wray et al.

2003; Martin et al. 2010). For example, despite 700 million years of evolutionary separation, mammalian HOX proteins, a conserved homeodomain transcription factors family found in vertebrates, can still functionally replace their *Drosophila* homologues (Mallo et al. 2010). A similar conservation was observed for extracellular proteins such as HEDGEHOG and WNT and their signaling pathways involved in embryo and various organ patterning (De Robertis 2008). In plant, 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. This class of genes are highly conserved in terms of sequence and function in many species notably Arabidopsis, Antirrhinum and rice (Weigel and Meyerowitz 1994; Nagasawa et al. 2003; Kanno et al. 2007). However, AP3/PI homologs in some species (i.e. petunia, maize, tulips, lilies, etc.) are expressed broadly across the floral meristem, indicating divergence of expression domains of conserved genes during evolution among angiosperm species, which is related to modification of floral organ identities (Soltis et al. 2007; Rijpkema et al. 2010a). These findings suggested that the variation in genes expression is an important source of phenotypic diversity.

Gene expression patterns are governed by complex gene regulatory networks that are described as *cis*-regulator and *trans*-regulatory elements. Consistent with the original definitions, *cis*-regulatory elements have an allele-specific effect on gene expression, and mapped near the target gene whereas *trans* elements affect the expression of distant genes, through the regulation of several alleles. *Trans*-regulatory elements work through an intermolecular interaction between different molecules to regulate the target genes such as transcription factors or insulators that regulate transcription initiation or small interfering RNA that regulates RNA stability. Whereas *cis*-regulatory are physicaly and genetically linked with the gene (or mRNA) they regulate (*i.e.* in a gene or an adjacent regulatory element near the target genes) such as 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 et al. 2008).

Although several researches reported that variation in *cis*-regulatory elements play important role in Evo-Devo biology (*i.e*, teosinte branched in maize, Ultrabithorax and yellow in fruitfly,…) we still know little about trans-regulator element (Wray 2007) Figure 1.3 represents several cases of potential mutation in *cis*-regulatory elements (CREs) that could affect transcription process, and as a result, could lead to morphological variation.

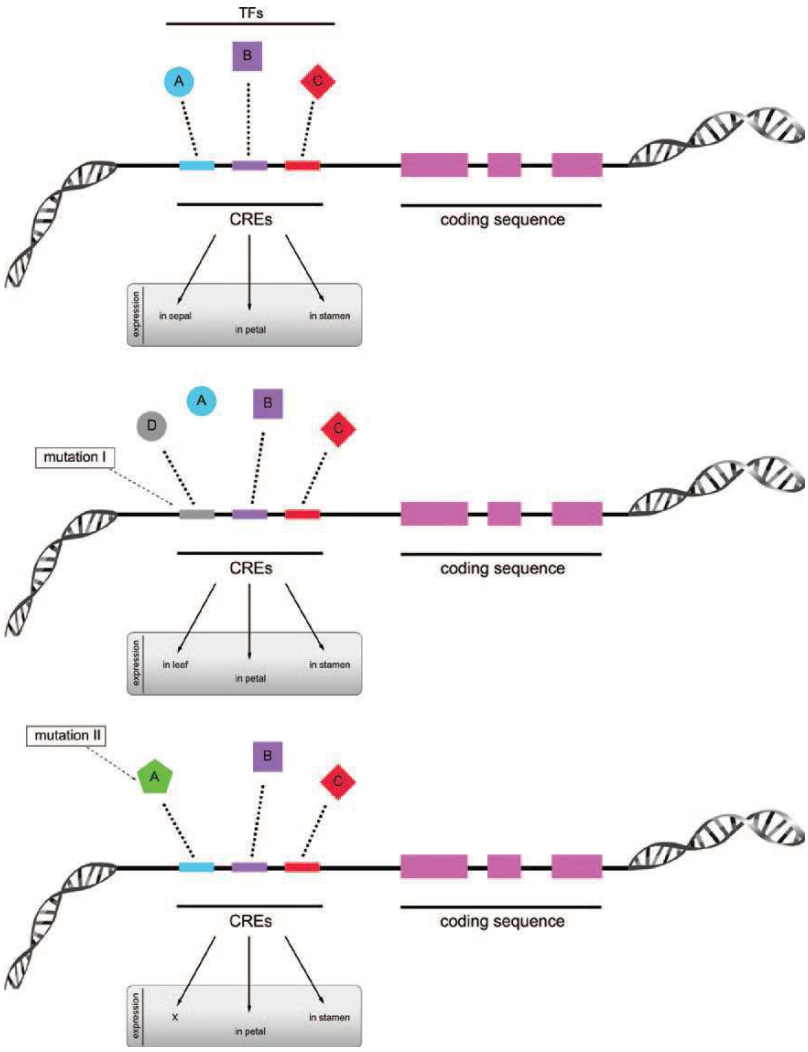


Figure 1.3: **Schematic representation of a gene with its *cis*-regulatory elements (CREs) and the potential mutations that can affect transcriptional processes**. CREs (A, B, C) together with their respective transcription factors (TFs) allow expression of a gene in a specific organ (or tissue). Middle 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) (From Pina et al. 2014).

1.1.3 "*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 (Figure 1.4) (Long et al. 2003).

Exon shuffling created around 19% of exons in eukaryotic genes by making an ectopic recombination of exons and domain from distinct genes (Patthy 1996; Patthy 1999). Morgante et al.

(2005) indicated that some genic insertions occurring in Maize shared the structural hallmarks of Helitron rolling-circle transposons. DNA segments defined by Helitron termini contained multiple gene-derived fragments that located in multiple genomic locations. Some of these produced transcripts containing segments of different genes, supporting the idea that these transposition events have a role in exon shuffling and the ceation of new proteins (Morgante et al. 2005).

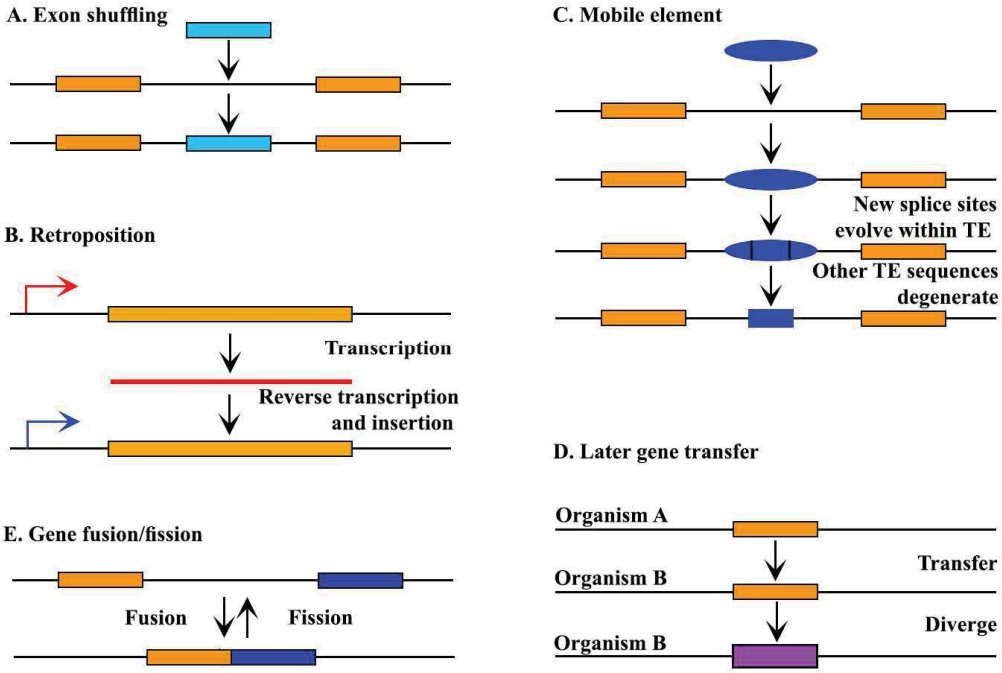


Figure 1.4: *De novo* formation of novel protein-coding genes. (A) exon shuffling, (B) retroposition, (C) Mobile element, (D) later gene transfer, (E) gene fusion/fission (Adapted from Long et al. 2003).

Retroposition is a mechanism related to functional retrogenes when new duplicated genes are created in new genomic positions by reverse transcription or other process (Wang et al. 2002; Betrán and Long 2002). New functional retrogenes have been reported in various organisms, especially mammals and *Drosophila melanogaster* (Long et al. 2003; Betrán et al. 2004). In plants, beside the few retrogenes have been identified in the actin gene family of potato (*Solanum tuberosum*), the alcohol dehydrogenase gene family in *Leavenworthia,* the *Bs1* retrotransposon in maize (Drouin and Dover 1990; Jin and Bennetzen 1994; Charlesworth et al. 1998). Wang et al (2006) showed 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 among organisms (Ochman 2001; Bergthorsson et al. 2003). The model propose that two adjacent genes can fuse into a single genes or that a single gene can split into two genes through the deletion, insertion or mutation of the translation stop codon and the transcription termination signal in the upstream gene could create a new gene function (Nurminsky et al. 1998) (Figure 1.4).

Almost new functional genes were created from ancient genes, thus they have continuous changes in sequence and structures to establish a further diverged function. The new genes seem to be necessary for adequate function, and only one or a few changes leading to new functions might be the exception. In contrast, the *de novo* gene origination process that a whole protein-coding gene created from a fragment of non-coding sequence is really seldom (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 non-coding region in genome create new function and the role of non-coding region in genome is still an open question.

1.1.4 Domestication

1.1.4.1 Domestication process

Evolution process leads to increase adaptation of organism with changing environment, whereas domestication process leaded to increase adaptation of plants and animals to cultivation or rearing and utilization by human. However, understanding domestication has been a tremendous help in understanding evolution. The domestication began when the agriculture began to encourage the growth of edible wild plants around 10 000 years ago. From hunting-gathering period, human selected and re-sow the favor grass from previous season for the next season. When the process was repeated several times, the population of plants that had desirable traits would be increased in the field. By 4000 years ago, ancient peoples had completed the domestication of all major crop species upon which human survival is dependent. It is about 500 angiosperm species that have been subject to at least some attempts at domestication which are distributed in six different centers of agriculture origin: Mesoamerica, the Andes of South America (including their piedmonts), Southwest Asia (the Fertile Crescent), Africa (Ethiopia and the Sahel), Southern China, and Southeast Asia (Harlan 1992; Smartt and Simmonds 1995).

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 of 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" (Hammer 1984).

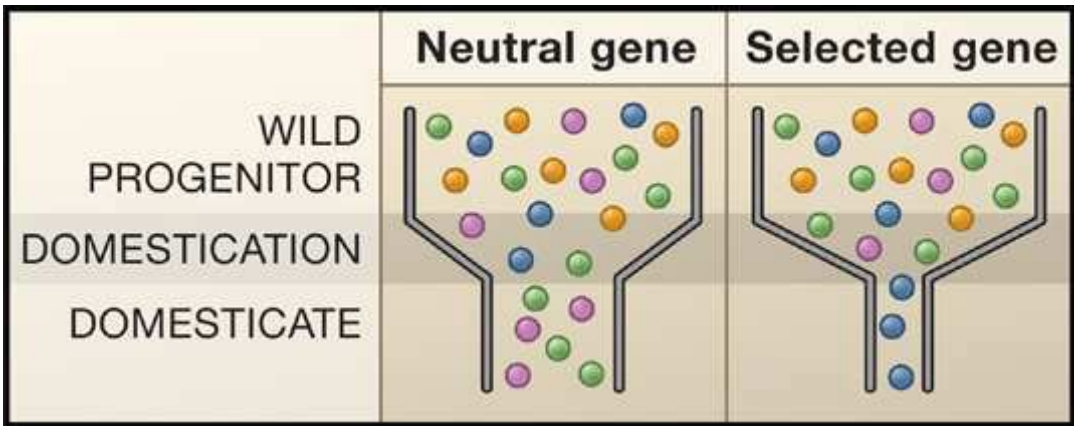


Figure 1.5: The Effects of the Domestication Bottleneck on Genetic Diversity. (Left) Population bottlenecks are a common important demographic event during domestication. Genetic diversity is represented by shaded balls; the bottleneck reduces diversity in neutral genes, as shown by the loss of the orange and blue variants. (Right) Selection decreases diversity beyond that caused by the bottleneck, as shown by the loss of all but one genetic variant in the domesticated species. Note, however, that an exceptionally strong domestication bottleneck could leave little variation in neutral genes. In that case, it may be very difficult to distinguish selected from neutral loci (From Doebley et al. 2006).

During the domestication, these early agricultural practices have 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 winnowing caused a genetic bottleneck, which reduced genetic diversity throughout the genome (Figure 1.5) (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). Notably, all genes in the genome did not experience loss in diversity equally. For genes that do not influence favored phenotypes (which are called neutral genes), the loss in diversity is resulting only of the strength of the bottleneck in terms of the population size and duration (Figure 1.5). However, genes that influence desirable phenotypes experienced a more drastic loss of diversity,

namely domestication genes (Figure 1.5). This was a consequence that plants carrying favored alleles contributed the progeny to each subsequent generation and that other alleles were eliminated from the population (Wright et al. 2005).

One unknown in the domestication process is the extent to which new mutations *versus* preexisting genetic variation in the wild species contributed to the evolution of crop phenotypes. For example, in a few cases, crops possess alleles of major genes that disrupt seed shattering (Li et al. 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, although at lower frequencies. Given the large store of genetic variation in the progenitor species, it seems most reasonable that domestication largely involved filtering out the best alleles from standing allelic variation in crop ancestors, although new mutations in key developmental pathways may have been instrumental for some traits.

1.1.4.2 Domestication genes

Several genes that were targeted during domestication or crop improvement have been identified in pathways governing fruit size and shape, seed dispersal, tillering, seed color, and many other traits (Doebley et al. 2006; Izawa et al. 2009). Because the traits involved are mostly quantitative in nature, the path to identify these genes consist of 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. Although the list of well-documented domestication genes is short, some generalities are beginning to appear (Paterson 2002; Chapman et al. 2008; Pearl et al. 2014). Examples of QTL characterization from various crops will be used after, to illustrate the nature of plant domestication-related genes and the corresponding selected mutations.

One of the most important domestication traits in rice is the loss of shattering with the main allele located on chromosome 4 (*sh4*). *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 et al 2006). *SH4* encodes a gene with homology to MYB3 transcription factors. A single amino changes in the predicted DNA binding domain converts plants from shattering to non-shattering (Li et al 2006). Interestingly, the non-shattering allele was present in all the *O. sativa* varieties surveyed, including members of *indica*, tropical and temperate *japonica* subpopulations, but not in *Oryza rufipogon* its wild-relative, leading to the hypothesis that this mutation was fixed very early during the domestication process but was not

present in the wild progenitors. 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 shattering zone, thus preventing shattering (Konishi et al. 2006). It has been also 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* gene. *Rc* encodes a bHLH transcription factor that plays a role in changing from red pericarp (in wild rice) to white pericarp (in most cultivated rice cultivars). The gene's function is impaired in the ancestor by a 14-bp frame-shift deletion that truncates the protein before the bHLH domain, thus produce white pericarp. This mutation is common within all *O. sativa* sub-populations (Sweeney et al. 2006; Sweeney and McCouch 2007). *Waxy* gene encodes a granule bound starch synthase (GBSS), an altered introns splice donor site in the gene lead to glutinous ("sticky") varieties lack 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 et al. 1997; Doebley, 2004). *tb1* represses the outgrowth of the axillary meristems and branch elongation via its repressive effect on the cell cycle in maize, thus maize plants typically have a single stalk with short branches tipped by ears, whereas teosinte plants are more highly branched (Doebley et al. 1997; Wang et al. 1999). Another QTL in maize is *Teosinte glume architecture1* (*tga1*) belonging to the squamosa-promoter binding protein (SBP) family of transcriptional regulators. The effects of *tga1* explain for the differences in glume induration between maize and teosinte (Dorweiler and Doebley 1997). The difference in function between the maize and teosinte alleles of *tga1* appears to be the result of a single amino acid change (Wang et al. 2005).

Q is a major gene involved in wheat domestication that was identified as a member of the AP2 family of plant-specific transcriptional regulators (Simons et al. 2006). The *Q* gene governs the free-threshing character and square spike phenotype and play important role in domestication of wheat. The cultivated (*Q*) allele is expressed at a higher level than the wild (*q*) allele and two alleles differ by a single amino acid change that affects protein dimerization, suggesting that both regulatory and protein function changes could be involved (Simons et al. 2006).

In tomato, *Fruitweight2.2* (*fw2.2*) and *SUN* are two domestic genes regulating the fruit shape. *fw2.2* that inhibits the cell division in the fruit, was identified to support mainly the QTL controlling 30% of the difference in fruit mass between wild and cultivated tomato (Frary et al. 2000; Cong et al. 2002). However, the large- and small-fruited alleles have no differences in protein sequence, supporting the hypothesis that changes in gene regulation of *fw2.2* underlie the evolution of tomato fruit size (Nesbitt and Tanksley 2002). Whereas, Xiao et al (2008) showed that overexpression of *SUN* elongate the fruit shape. Sun increase expression causes a gene duplication event mediated by the long terminal repeat retrotransposon (Figure 1.6).

As illustrated above, the form and nature of the mutations associated with domestication process is highly variable (Figure 1.6). Obviously understanding domestication has already been and will continue to be a tremendous help in understanding evolution mechanisms. Since crop domestication started just more than 10 000 years ago, it would be a good model system for understanding these mechanisms. This knowledge will offer also a solid foundation for crop engineering in the near future.

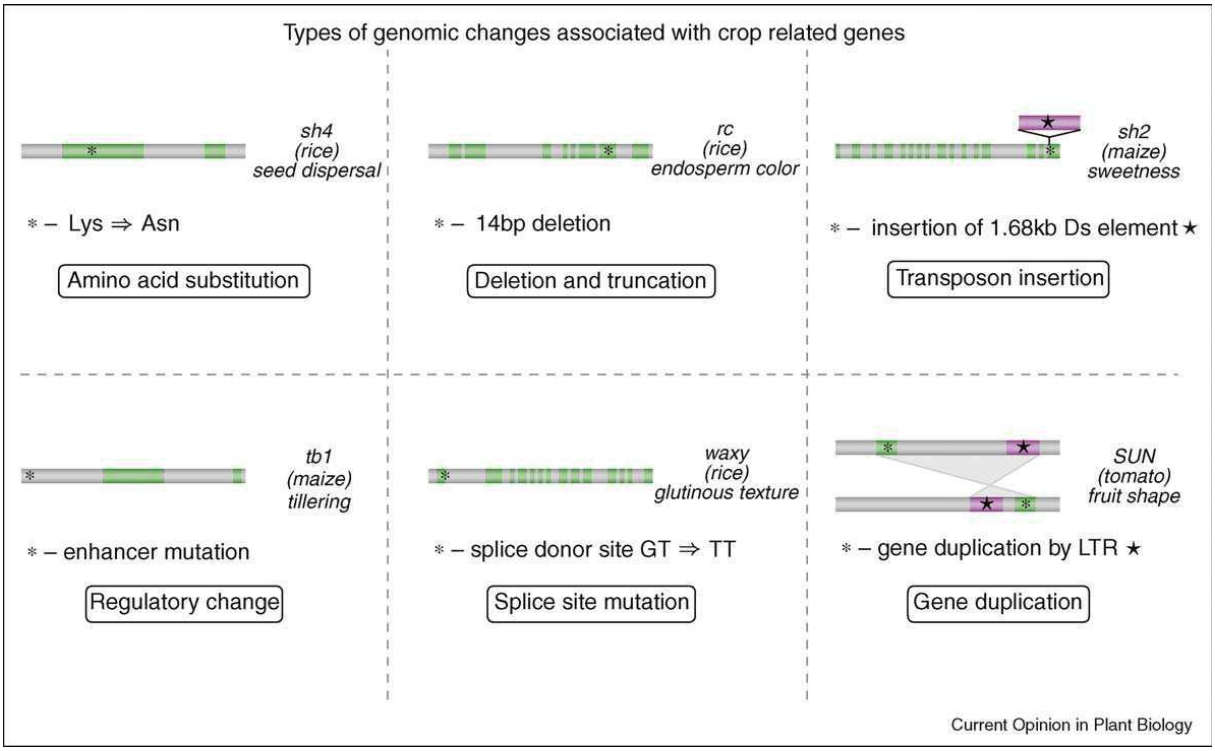


Figure 1.6 Types of changes associated with crop-related genes. One specific example is given for each type of genomic change: amino acid substitution (*sh4* in rice), deletion and truncation (*rc* in rice), transposon insertion (*sh2* in maize), regulatory change (*tb1* in maize), splice site mutation (*waxy* in rice) and gene duplication (*Sun* in tomato) (From by Tang et al. 2010).

1.1.5 Plant small RNAs

An increasing number of studies show that regulatory non-coding RNAs as well as protein-coding gene changes have been a driving force of morphological evolution of plants. Plant small RNAs constitute a family of regulatory non-coding RNAs of 21-24 nucleotides (nt) that play important role in a variety of biological regulation processes, such as development, plant defense, and epigenetic modifications. Base on distinct origin and biogenesis, with functions at both transcriptional and post-transcriptional levels, small RNAs in plants can be categorized into two major classes: the small interfering RNAs (siRNAs) which are derived from double-stranded RNA (dsRNA) precursors and microRNAs (miRNAs) which are derived from single-stranded precursors with a hairpin structure (a self complementary) (Figure 1.7) (Axtell 2013).

Although plant small RNAs are highly diverse, all small RNAs have a sets of RNA-dependent RNA polymerase (RDR), DICER-LIKE (DCL), and ARGONAUTE (AGO) family members required for their biogenesis, function and unique size distributions. RDRs synthesize second-strand RNA using an RNA template, resulting in the production of dsRNAs. DCL endonucleases process helical RNA precursors (either dsRNA or the helical regions of stem-loop single-stranded RNAs) to release short double-stranded duplexes, 20 to 24-nt long, with 2-nt 3' overhangs. AGOs then engage these duplexes, retaining only one of the two possible strands and discarding the other. AGO-loaded small RNAs serve as specificity determinants to select RNA targets based on small-RNA/target complementarily. Target binding is followed by repressive activities orchestrated by the associated AGO protein such as direct translational repression, chromatin modifications, and slicer-independent destabilization of target mRNAs. RDRs, DCLs and AGOs are all encoded by multigenic families in plants with conserved clades. Each clade is often specialized for the production or use of a certain class of small RNAs (Margis et al. 2006; Vaucheret 2008). Furthermore, the defining features and biogenesis requirements for miRNAs and siRNAs are known to be conserved and to remain distinct from one another in multiple plant species (Axtell 2013)

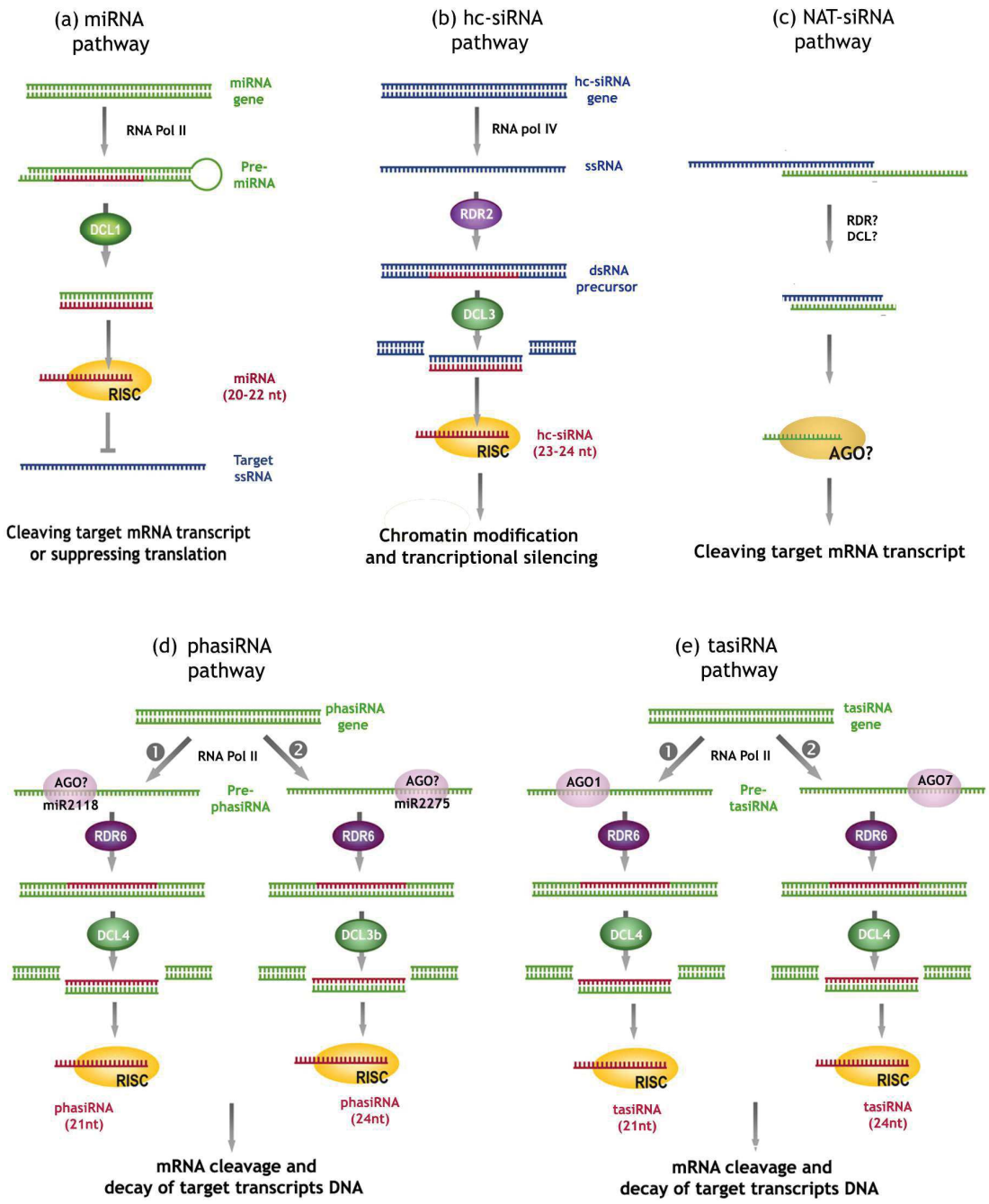


Figure 1.7: A schematic overview of plant small RNAs, their biogenesis pathways, and their modes of action. MiRNAs are generated from stem-loop precursors whereas siRNAs (including hc-siRNA, NAT-siRNA, phasiRNA, tasiRNA) are processed from long dsRNAs. (a) MiRNA genes are transcribed by RNA polymerase II to generate the primary transcripts (pre-miRNAs). Dicer like protein1 (DCL1) participates in the second step of processing (dicing) to produce miRNA duplexes. The duplex is separated and usually one strand is selected as mature miRNAs, whereas the other strand is degraded. The final products act as guide molecules in translational control or cleavage of certain mRNAs. (b) Heterochromatic siRNAs (hc-siRNAs) are generated from double-stranded precursors, which are transcribed from heterochromatic regions by Pol IV and converted to dsRNA by

RNA-DEPENDENT RNA POLYMERASE2 (RDR2). The 24-ht hc-siRNAs are processed from the long dsRNA precursor by DCL3. (c) Natural siRNAs (NAT-siRNAs) dsRNA precursors of NAT-siRNAs are thought to arise from the hybridization of separately transcribed, complementary RNAs. The biogenesis of NAT-siRNA is still unclear. (d-e) Phased si-RNAs (phasiRNAs) are generated from long dsRNA precursors converted from single-stranded RNAs of TAS genes or PHAS genes by RDR. Either 21-nt or 22-nt miRNAs, bound by AGO7 or AGO1, respectively, are required as triggers to initiate the conversion of the ssRNA to dsRNA precursors. In the two-hit model of trans-acting siRNAs or phasiRNAs biogenesis, one of two 21-nt miRNA binding sites is cleaved, whereas in the single-hit model there is one miRNA binding site, for a 22 nt miRNA. In both cases, the long dsRNA is cleaved into phased 21-nt siRNAs by DCL4, or phased 24-nt siRNAs by DCL3b. TasiRNAs are loaded into AGO1 or AGO7 and direct cleavage of mRNA targets. Rice panicle-specific 21-nt phasiRNAs are bound by an unknown AGO and presumably (based on activities of other 21 nt siRNAs) direct cleavage of mRNA targets, while 24-nt phasiRNAs presumably (based on activities of other 24 nt siRNAs) direct DNA methylation at target genes (also via an unknown AGO) (Adapted from Ghildiyal and Zamore 2009 and Arikit et al. 2013).

1.1.5.1 MicroRNAs (miRNAs)

MicroRNAs (20 to 22-nt long) are typically processed from a hairpin-like secondary structure of a noncoding mRNA (ncRNA), with a precursor mRNA generated by RNA polymerase II (RNA Pol II). The RNase III enzyme DCL1 is responsible for the biogenesis of the mature miRNA *via* processing of the mRNA precursor (Figure 1.7) (Valencia-Sanchez et al. 2006; Voinnet 2009). Loss-of-function *dcl1* mutants of *Arabidopsis thaliana* resulted in decreasing miRNA levels and ectopically increased the expression of miRNA target genes (Kasschau et al. 2003; Reinhart et al. 2002).

In plant, miRNAs were known as key post-transcriptional regulators in plants that normally suppress gene expression by (i) cleaving their target mRNA transcript at highly specific sites or (ii) suppressing translation; these modes of action depend largely on the miRNAs complementarity with target sequences (Mallory and Vaucheret 2006; Valencia-Sanchez et al. 2006; Voinnet 2009) (Figure 1.7). In general, most target mRNAs only contain one single miRNA-complementary site, and most corresponding miRNAs perfectly complement these sites and cleave the target mRNAs (Kidner and Martienssen 2005). However, some miRNAs, such as *miR172*, regulate gene expression by repressing gene translation, although they can perfectly complement the target mRNAs (Chen 2004).

Concerning plant development, miRNA defects caused many developmental deficiencies, such as delaying flower timing, over-proliferation of shoot meristems and embryogenic suspensor cells, and converting normally determinate floral meristems to indeterminate meristems (Kim 2005; Yang et al 2007). Some miRNA genes are also involved in hormone signaling and environmental stress (Sunkar and Zhu 2004; Yang et al. 2007). Concerning plant evolution, highly conserved miRNAs predate the divergence of gymnosperms and angiosperms 305 million years ago, and the divergence between vascular plants and mosses 490 million years ago (Axtell and Bowman 2008). However, several studies have shown that a minority of miRNA families are conserved between plant families, while the

majority are family- or species-specific, suggesting that most known miRNA genes arose relatively recently in evolutionary time (Zhang et al. 2006; Cuperus et al. 2011). Notably, some miRNA families deriving from common ancestor genes show different patterns of expression across the species, reflecting episodes of gene duplication followed by lineage specific functional diversification such as *miR159*/*miR319* (Palatnik et al. 2007) or complete loss in some taxonomic groups, as in the example of *miR529 versus miR156* (Cuperus et al. 2011). Montes et al. (2014) indicated a miRNA superfamily including *miR390*, *miR1432* and several other miRNAs related in sequence exhibits the most diversified pattern of taxonomic distribution suggesting a complex evolutionary history. The presence of miRNA families across the phylogeny of terrestrial species is presented in Figure 1.8

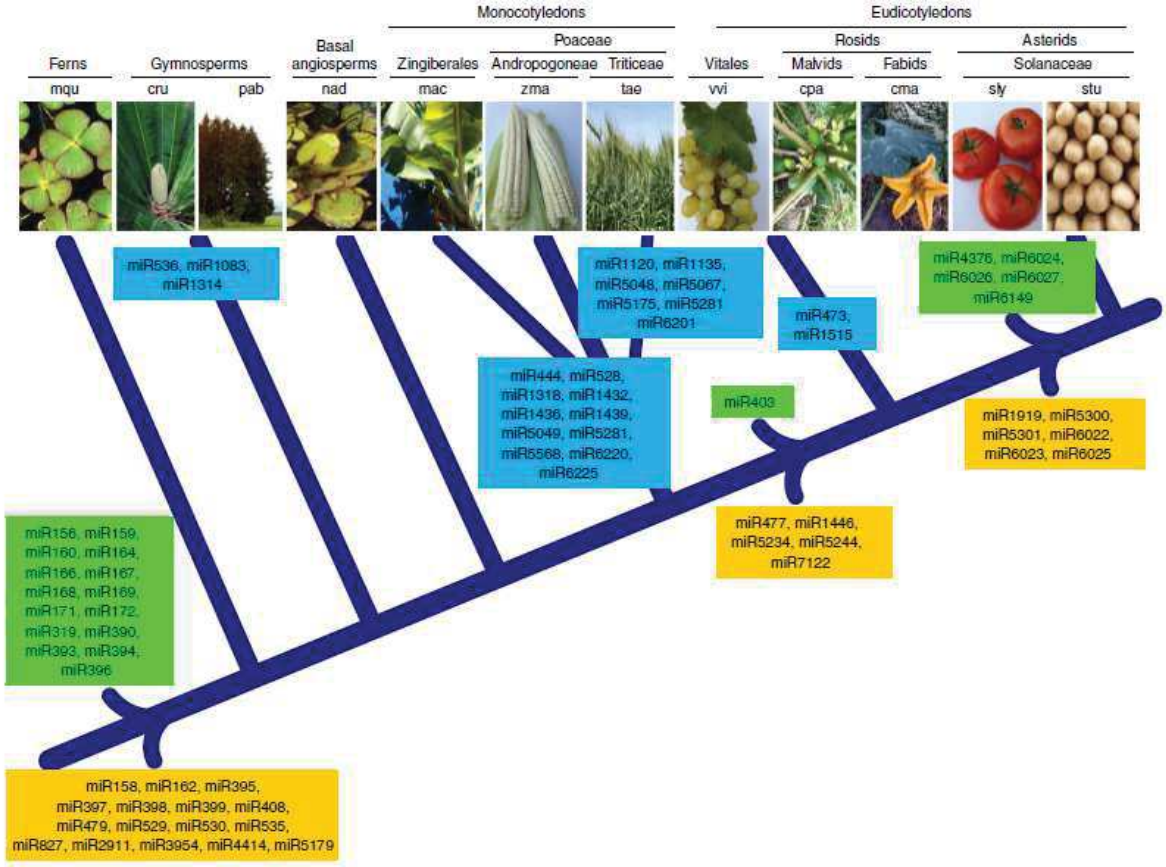


Figure 1.8: miRNA family emergences across the phylogeny of terrestrial plant species. Families colored in green are conserved across virtually all corresponding species. Families colored in orange are conserved, although missing in a few corresponding species. Families colored in blue appear to be specific to a particular group of species (From Montes et al. 2014).

A well-described case is related to *miR164* regulating the expression of *NO APICAL MERISTEM* (*NAM*) genes belonging to the NAC family of transcription factors. This post-transcriptional regulation of these genes is necessary for normal embryonic, stem development and floral development (Laufs et al. 2004; Mallory et al. 2004; Peaucelle et al. 2007). Homologues of

miR164 have been reported from all angiosperms, gymnosperms and ferns but not in moss (Mallory et al. 2002; Reinhart et al. 2002). Similarly, a potential *miR164*-binding site is present in *NAM*-related genes of core and basal angiosperms and gymnosperms (Adam et al. 2010). At least two of them, *NAM* and *CUP*, have a similar role to the *Arabidopsis* CUC genes in petunia and *Antirrhinum,* respectively (Souer et al. 1996; Weir et al. 2004) suggesting that the regulator mechanism in *Arabidopsis* is likely to be evolutionary conserved (Laufs et al. 2004; Kidner and Martienssen 2005). Another example is *miR172* regulating *APETALLA* (*AP1* and *AP2*) gene expression to regulate floral organ identity and flowering time (Chen 2004). The target site of *miR172* is significantly conserved in gymnosperm *AP2* homologs suggesting a highly conserved regulatory function over the 300 million years since the divergence of gymnosperms and flowering plant lineages (Chen 2004; Shigyo et al. 2006).

1.1.5.2 Small interference RNAs (siRNAs)

These siRNAs typically range in size between 21- and 24-nt long in plant. They are associated with both post-transcriptional forms of RNA interference (RNAi) and transcriptional silencing involving chromatin modification by cleaving and decaying their target mRNA transcript, DNA methylation and histone modifications of target loci, respectively (Finnegan and Matzke 2003; Xie et al. 2004; Bonnet et al. 2006). In contrast to miRNAs, siRNAs are processed from precursors containing extensive or exclusive double-stranded RNA (dsRNA) structure, such as transcripts containing inverted repeats or intermediates formed during RNA virus replication (Hannon 2002). Moreover, the formation of siRNAs depends on the multiple DCL activities or pathways to provide the small-sized (approximately 21 nucleotides) and large-sized (approximately 24 nucleotides) classes (Tang et al. 2003). The siRNAs can be subdivided into heterochromatic siRNAs (hc-siRNAs), phased or secondary siRNAs (phasiRNAs), *trans*-acting siRNA (ta-siRNAs) and natural antisense transcript siRNAs (NAT-siRNAs) (Figure 1.7) (Axtell 2013).

Heterochromatic siRNAs (hc-siRNAs)

Most heterochromatic siRNAs are 23- or 24-nt long, derived from intergenic and/or repetitive genomic regions and are associated with the *de novo* deposition of repressive chromatin modifications at target DNA loci (Matzke et al. 2009; Law and Jacobsen 2010). The function of hc-siRNAs is largely to maintain genome integrity, by maintenance of suppressive levels and types of DNA methylation on transposable elements. Heterochromatic siRNAs depend specifically on an alternative DNA-dependent RNA polymerase (RNA Pol IV), RDR2, DCL3 and CLASSY1, a protein with a possible role in chromatin remodeling for their biogenesis (Kasschau et al. 2007; Chen 2009; Mosher

et al. 2009) and on members of the AGO4 clade of AGOs (AGO4, AGO6, and AGO9 in *Arabidopsis*) for their function (Figure 1.7) (Henderson et al. 2006; Qi et al. 2006).

Like miRNAs, hc-siRNAs as a distinct class of endogenous plant small RNAs are clearly conserved in multiple species. For instance, most small RNAs in immature maize ears are 24-nt long and are dependent on *mop1*, a maize *RDR2* homolog, for their accumulation (Nobuta et al. 2008). Similarly, 24-nt small RNAs dependent on *OsDCL3a* and *OsRDR2* dominate the small RNA profile of wild-type rice (Wu et al. 2010). However, the conservation of heterochromatic siRNAs is nonetheless quite distinct from that of miRNAs: in the case of miRNAs, individual miRNAs themselves can be conserved across multiple species. In contrast, individual heterochromatic siRNA loci do not appear to be conserved even between closely related species (Ma et al. 2010), even though the pathway itself is conserved. Many hc-siRNA loci overlap with transposons or transposon fossils, and there are likely to be characterized by rapid birth and death of individual heterochromatic siRNA loci in response to the rapid changes in transposon position and copy number that occur during plant evolution (Axtell 2013). Beyond flowering plants, the analysis of hc-siRNA pathway in several conifers and some other lineages indicates that hc-siRNAs lost functional in the conifers (Dolgosheina et al. 2008) but this pathway could be deployed in a tissue-specific manner of some lineages (*i.e. Selaginella)* (Banks et al. 2011). Synthesizing the available data, it appears that the hc-siRNA pathway is ancestral within the land plants.

Secondary siRNAs

Almost secondary siRNAs namely phased siRNAs (or phasiRNAs), derive from an mRNA converted to dsRNA by RDR6 and processed by DCL4 (Figure 1.7). Some secondary siRNAs are also capable of acting in *trans* to direct repression of distinct mRNA targets - hence the term *trans*-acting siRNAs (*ta*-siRNA). Since phasiRNAs and *ta*-siRNAs often apply to the same locus, many of the known *ta*-siRNAs are also phased. Compared with miRNAs, secondary siRNAs are well conserved, present in flowering plants as well as more diverged lineages (Talmor-Neiman et al. 2006; Axtell et al. 2006). In addition, some individual secondary siRNA genes themselves are conserved between different plant species to varying degrees. Taken together, these consistent traits indicate that secondary siRNAs are a robust, distinct, and biologically meaningful class of small RNA genes (Axtell 2013).

Deep analysis indicated that plant *ta*-siRNAs (mostly 21-nt long) are generated by the convergence of the miRNA and siRNA pathways. In *Arabidopsis*, four *Trans-Acting siRNA* (*TAS1–TAS4*) transcripts are initially targeted and cleaved by the RISCs containing AGO1-*miR173* (for *TAS1*

and *TAS2*), AGO7-*miR390* (for *TAS3*) and AGO1-*miR828* (for *TAS4*), respectively (Allen et al. 2005; Montgomery et al. 2008). Subsequently, RDR6 and the RNA binding protein SUPPRESSOR OF GENE SILENCING 3 (SGS3) may be recruited to cleavage sites where RDR6 converts the single-stranded RNAs into double-stranded RNAs, which are cleaved by DCL4 into phased 21-nt siRNAs (Axtell 2013). Among them, the *TAS3*-associated *ta*-siRNAs, triggered by the microRNA *miR390*, target mRNAs of *Auxin Response Factor3* (*ARF3*) and *ARF4* genes involved in various developmental processes (such as organ polarity, meristem identity, and developmental timing) and is widely conserved (Allen et al. 2005; Axtell et al. 2006; Song et al. 2012a; Song et al. 2012b). The model of *ta*-siRNA biogenesis in rice from *TAS3* was described in Figure 1.7.

The phasiRNA biogenesis in eudicot species is initiated from disease resistance genes belonging to the Nucleotide-Binding Site–Leucine-Rich Repeat (NBS-LRR) superfamily that is triggered by the *miR482*/*miR2118* superfamily of miRNAs in multiple species (Li et al. 2012; Shivaprasad et al. 2012). The authors showed that both viral and bacterial infections of tomato correlate with reductions in *miR482* accumulation and increases in NBS-LRR disease-resistance mRNA accumulation. Interestingly, the targets of *miR2118* family members in rice have not been reported to encode NBS-LRR protein but long non-coding RNAs (Johnson et al. 2009; Song et al. 2012a). Interestingly, *miR2118 -triggered phasiRNAs* were shown to be panicle-specific 21-nt small RNAs generated from thousand loci over rice genome respectively (Johnson et al. 2009; Song et al. 2012a; Komiya et al. 2014). A similar phasiRNA pathway occur in rice panicle but triggered by *miR2275* and producing 24-nt phasiRNAs from a reduced number of loci (Song et al., 2012a). Similarly to *ta*-siRNA biogenesis, phasiRNA biogenesis is also a RDR6-dependent pathway, but after double-strand RNAs through *OsRDR6*, the processing of 21- and 24-nt phasiRNAs required *OsDCL4* and *OsDCL3b*, respectively (Figure 1.7) (Song et al. 2012a,b; Komiya et al. 2014). Although the function of 21- and 24-nt phasiRNAs and their targets is still largely unknown (Johnson et al. 2009; Song et al. 2012a,b), *miR2118* and *miR2275* are preferentially expressed in rice and maize stamens (Song et al. 2012a). Moreover, the action of the 21-nt phasiRNAs seems to rely on their interaction with the gamete-specific AGO protein, MEIOSIS ARRESTED AT LEPTOTENE1 (MEL1) (Komiya et al. 2014). However, their function during male gametogenesis is still unclear. The panicle-enriched 21- and 24-nt phasiRNAs were also reported in *Brachypodium distachyon* and in maize (Vogel et al. 2010; Song et al. 2012a), indicating the conservation of these two secondary siRNA pathways in grasses in addition to the *ta*-siRNA pathway related to *TAS3* loci.

Natural siRNAs (NAT-siRNAs)

NAT-siRNAs are a narrowly described, unusual, and perhaps questionable category of small RNAs purportedly derived from two distinct, homologous, and interacting mRNAs (Borsani et al. 2005). In contrast to the other types of siRNAs, which rely on an RDR to synthesize the precursor dsRNA, the dsRNA precursors of NAT-siRNAs are thought to arise from the hybridization of separately transcribed, complementary RNAs (Figure 1.7). The separate RNAs can be complementary because they were transcribed from opposite strands of the same locus; these are the *cis*-NAT-siRNAs. Alternatively, the hybridizing RNAs can arise from genes that possess no overlap; these are the *trans*-NAT-siRNAs. Only *cis*-NAT-siRNAs have been described in plants; *trans*-NAT-siRNAs remain only a hypothetical possibility (Axtell et al. 2006). Several *cis*-NAT-siRNAs have been functionally analyzed that regulated development and stress resistant in *Arabidopsis* such as salt-stress-induced, antiviral defense, and regulate the reproductive by controlling sperm function during double fertilization (Borsani et al. 2005; Katiyar-Agarwal et al. 2006; Ron et al. 2010) . Genome-wide analyses of *cis*-NAT gene pairs in *Arabidopsis* as well as the presence of siRNAs in many of the *cis*-NATs suggest that siRNA regulation of cis-NATs *via* the RNAi pathway is an important gene regulatory mechanism for at least a subgroup of *cis*-NATs (Jin et al. 2008).

Several study indicated that cis-NAT genes were not strong sources of small-RNAs, only 6% to 16% of *Arabidopsis* and rice *cis*-NAT pairs, respectively, were associated with appreciable amounts of small RNA accumulation (Henz et al. 2007; Zhang et al. 2012). In addition, the authors also found a significant enrichment of small RNA accumulation within the overlapped regions of *cis*-NAT gene pairs, relative to non-overlapping positions in the two genes (Henz et al. 2007; Zhang et al. 2012). It seems that a *cis*-NAT gene configuration by itself is not generally predictive of *cis*-NAT-siRNA formation, and this suggests that *cis*-NAT-siRNAs may not play a major role in the regulation of most of the *cis*-NAT genes observed in plants. Concerning the biogenesis pathway, *cis*-NAT-siRNA production is strikingly heterogeneous that require individualized subsets of RDRs, DCLs, and other factors for their accumulation. Many of the *cis*-NAT-siRNAs investigated to date depend on an RDR for their accumulation; however the mechanism of RDR pathway was still unknown. Further study about *cis*-NAT-siRNA biogenesis and their function is clearly needed.

1.2 Inflorescences and Meristems

The inflorescence is the structure of the plant bearing the flowers (*i.e.* the structure bearing the reproductive organs). A huge diversity of this structure is observed within the angiosperms with singled-flowered species to highly branched structures with different organization and complexity levels. The development of these structures depends on the activity of both apical and axillary meristems regulated by different regulatory gene networks. In the following sections, current knowledge on meristem functioning as well as related regulatory gene networks will be illustrated. Moreover, the diversity of inflorescence architecture and its modeling to explain its evolution will be developed.

1.2.1 Meristem functioning

The post-embryogenic growth of plants depends on the continuous function of the tissue containing undifferentiated cells called "meristem" established during embryogenesis. The development of an apical-basal axis is defined by the root apical meristem (RAM) at the one and the shoot apical meristem (SAM) at the other (McSteen and Leyser 2005). The apical meristems of both RAM and SAM are undifferentiated (indeterminate) meristem. While the RAM localized at the root top harbors stem cells that divide asymmetrically and generate initial cells for all the cell types in the root, the SAM is responsible for all the post-embryonic aerial organs such as leaves, stems and flowers. These two meristems are both established during embryonic development.

Shoot apical meristems (SAMs) have very similar structures in different plant species including cell layers and concentric zones. Cells in the outermost layer called L1 divide in anticlinal orientation and contribute to the epidermal layer. The L2 layer is internal to the L1 and gives rise mostly to mesophyll tissue. The interior of the meristem is made up of the L3 and comprises multiple cell layers which help in the formation of the internal tissues, mesophyll and vascular (Figure 1.9) (Furner et al. 1992; Xie et al. 2009). In addition, three distinct zones of the SAM are defined by cytoplasmic densities and cell division rates: the peripheral zone (PZ), the central zone (CZ) comprising the organizing center (OC) and the rib zone (RZ) (Figure 1.9 A). These zones might represent a functional subdivision of the SAM: lateral organs are produced from cells recruited from the PZ whereas stem tissue is derived from cells recruited from the RZ. The CZ acts as a reservoir of stem cells, which replenish both the peripheral and rib zones, as well as maintaining the integrity of the central zone (Lenhard and Laux 1999).

As mention above, SAM generates above-ground aerial organs throughout the lifespan of land plants. In order to fulfill this function, the meristem must maintain a balance between the self-renewal of a reservoir of central stem cells and organ initiation from peripheral cells (Lenhard and Laux 1999; Doerner 2003). Involve to the self-maintaining function, the activity of the pluripotent stem cell population in the SAM is dynamically controlled by complex, overlapping signaling networks that include the feedback regulation of meristem maintenance genes and the signaling of plant hormones (*i.e.* cytokinin) (Lenhard and Laux 1999; Bowman and Eshed 2000; Pautler et al. 2013) (Figure 1.9 C,D). In *Arabidopsis*, cells in the OC express the transcription factor WUSCHEL (WUS), which promotes the expression of CLAVATA3 (CLV3), a small peptide that moves into the surrounding tissue (Kondo et al. 2006; Mülleret al. 2008). In the CZ, CLV3 interacts with the receptor-like kinase CLAVATA1 (CLV1), inhibiting *WUS* expression and promoting stem cell fate (Schoof et al. 2000). The cytokinin phytohormone (CK) is also required to establish and maintain the CZ. The transcription factor SHOOTMERISTEMLESS (STM) up-regulates the expression of *ISOPENTENYL TRANSFERASE* (*IPT*) genes that are rate limiting for cytokinin biosynthesis (Jasinski et al. 2005; Yanai et al. 2005). Cytokinin activates A-type transcriptional regulator (ARRs) *via* a phosphorelay system. Then, A-type ARRs stimulate downstream cytokinin responses but also down-regulate the expression of *WUS* (Hwang et al. 2012). WUS inhibits the expression of A-type ARRs, creating a negative feedback loop that regulates size and position of the OC and, thus, of the CZ (Leibfried et al. 2005).

Organ initiation likewise requires the function of multifactor gene regulatory networks, as well as instructive cues from the plant hormone auxin and reciprocal signals from the shoot meristem. The high local auxin concentrations of auxin was required for the initiation of a new organ and the PIN-FORMED (PIN) transporters that create fluxes of auxin through the tissues are required for the creation of such auxin maxima (Figure 1.9) (Benková et al. 2003; Bayer et al. 2009; Besnard et al. 2011). The other set of transporters associated with auxin distribution at the SAM is the family of *AUX/LIKE AUX (LAX)* influx carriers (Péret et al. 2012). The *AUX1* gene is expressed at the L1 cell layer of the SAM (Bennett et al. 1996; Reinhardt et al. 2003) and might concentrate auxin at the meristem surface and act to facilitate organ positioning (Reinhardt et al. 2003; Besnard et al. 2011; Murray et al. 2012).

During the vegetative phase of growth, the SAM generates leaf primordial on its periphery, and then develops the secondary shoots, or tillers. Upon receiving the appropriate environmental and developmental signals, plants switch to the reproductive phase. The vegetative converts into inflorescence meristem (IM) including rachis meristem that then produce branch meristem, and floral

meristems. During this period, the meristem changes its fate and converts from indeterminate meristem in apical and axillary meristems (*i.e.* self-maintaining activity on) into determinate meristem for the floral meristems (*i.e.* self-maintaining activity off, but organogenesis on). This transition relies on different regulatory-gene networks leading to an inactivation of *WUS*-like activity within the meristem.

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.

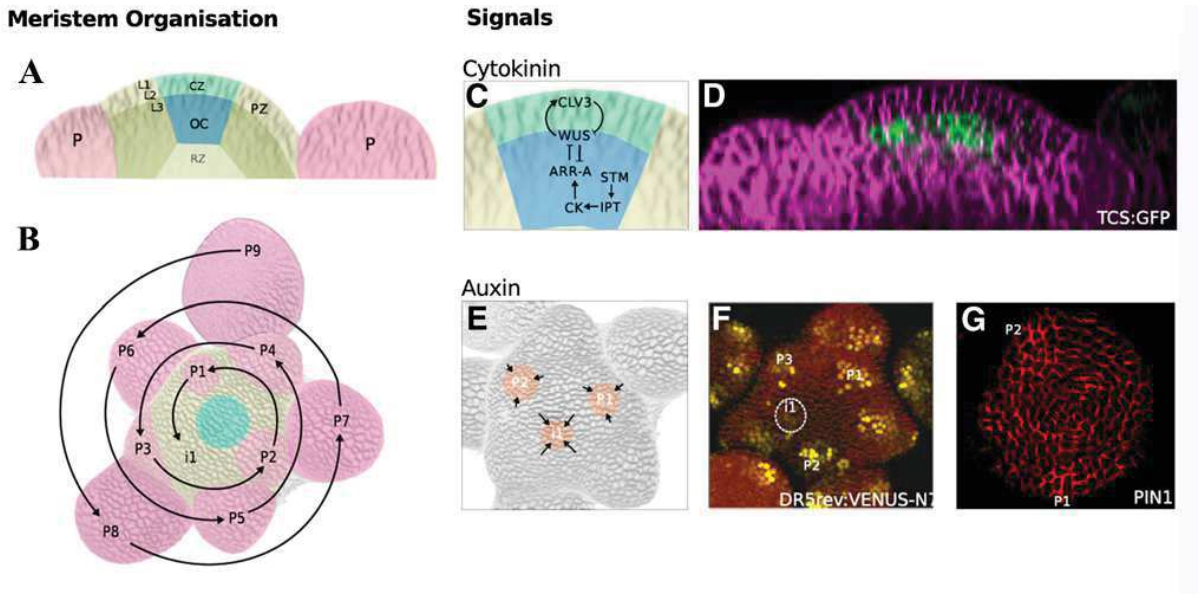


Figure 1.9 Structural and functional organization of the SAM in *Arabidopsis thaliana*. (A) The distance zone and layer of SAM. (B) Primordia are spaced according to a regular pattern or phyllotaxis. P9 indicates the oldest primordium and P1 the youngest, (i1) is the next primordium. (C) Meristem maintenance genes and the signaling of plant hormones control self-maintaining of SAM. (D) High level of cytokinin was detected in the organizing center. (E) to (G) The positions of primordia are determined by auxin maxima (orange) that are created by self-organizing patterns of auxin transport (arrows) (E). The auxin response reporter DR5:3xVENUS-N7 is detected in primordia before outgrowth begins (see circled i1 in [F]). Directional movement of auxin is produced by the activity of PIN1 proteins, which transport auxin out of cells and are polarly localized. Immuno-localization of PIN1 in the L1 of the SAM shows that PIN1 proteins are oriented toward the auxin maxima (G). (Adapted from Murray et al. 2012).

1.2.2 Inflorescences architecture

In general, inflorescence architecture was made from repetitive units including bract (the terminal leaves associated with a flower), pedicel (the stalk bearing a flower) and flower. Based on the

complexity of the iterative nature of plants, inflorescence types are diverse and range from an unbranched main axis terminating with the production of a single flower, as in a tulip, to more- or less-complex branching systems that produce numerous flowers over an extended period of time, such as tomato, rice, etc (Figure 1.10) (Weberling 1989).

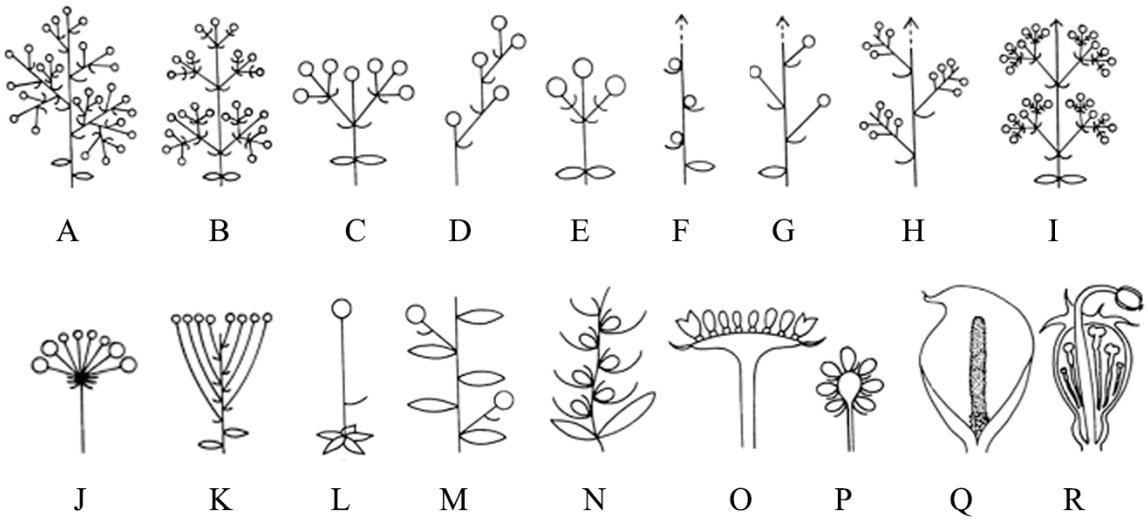


Figure 1.10: Diagrams of different types of inflorescences. (A–E) Determinate inflorescences: (A-B) panicle type (A: panicle; B thyrsoid); (C-E) cyme type (C: dichasium; D: monochasium, E: triad); (F-Q) Indeterminate inflorescences (the raceme type) (F) spike; (G) raceme; (H) panicle-like; (I) thyrse; (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.rbgsyd.nsw.gov.au/).

Inflorescences can be grouped into three main types, namely raceme-type (*i.e. Arabidopsis*), cyme-type (*i.e.* tomato, petunia) and panicle-type (*i.e.* rice), based on the termination events on the inflorescence meristems of various order (Figure 1.10) (Benlloch et al. 2007; Prusinkiewicz et al. 2007). In raceme-type of *Arabidopsis*, main inflorescence meristem grows indefinitely and generates either floral meristems (FMs) or primary branch meristems (PBMs) that continuous produce either secondary branch meristems (SBMs) or FMs (Figure 1.10 G) (Remizowa et al. 2013). The panicle-type inflorescences are largely characteristic of grasses such as rice (*O. sativa*) and oat (*Avena sativa*). 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 (Figure 1.10 A) (Yamaki et al. 2010). In cymose inflorescences the apex also transforms into a terminal flower, but growth of the inflorescence continues through lateral axes produced below the

terminal flower (Fig 1.10 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). Cymose inflorescence display variation, a simple form of *Silene latifolia*, tobacco (*Nicotiana tabacum*) to a sympodium of tomato (*Solanum lycopersicum*) (Benlloch et al. 2007; Teo et al. 2014).

In general, the variation among inflorescence could be attributed by three main factors: (i) determinacy or indeterminacy of meristem within the shoot system; (ii); extent of growth in each of three dimensions of stem and stem-like structures (*i.e.* internodes length) and (iii) relative positions of lateral shoots and/or flowers (*i.e.* phyllotaxy). Theses different points will be illustrated in the following sub-sections.

Based on the determinacy of shoot meristems, inflorescence architecture could be divided into basic types: indeterminate and determinate (Figure 1.10 and Figure 1.11). If the inflorescence meristems are considered determinate, they will produce floral meristem (was known as determinate meristem) and it is no longer able to establish new lateral meristem. In contrast, inflorescence meristems are considered indeterminate, as they continue to initiate new branch meristems or lateral meristems (were known as indeterminate meristem) (Weberling 1989; Benlloch et al. 2007).

1.2.2.1 Indeterminate and determinate inflorescence architecture

In determinate species, all shoot meristems in the inflorescence eventually become floral meristems (Figure 1.11 B-C). In this case, the inflorescence structures were 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. A variation of the cyme pattern is found from simple-cyme type such as in Petunia to more complex cyme-type called sympodium such as in tomato (*Solanum lycopersicum*). This type of inflorescence could be terminated by a flower (as in pepper or petunia), five to six flowers (as in tomato), or dozens of flowers (as in the Chilean potato vine) (Hake 2008). The other determinate inflorescence architecture is the panicle type. In contrast to the cyme, in this type of inflorescence a clear main shoot axis exists but this is terminated by a floret meristem (Benlloch et al. 2007; Lippman et al. 2008).

On the contrary, in indeterminate species, the apical meristem remains indeterminate and produces lateral meristems that become flowers (Figure 1.11 A). The inflorescences where 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 branches meristem are called compound racemes such as in the *Leguminosae* species pea (*Pisum sativum*), *Medicago truncatula* or *Lotus japonicas*.

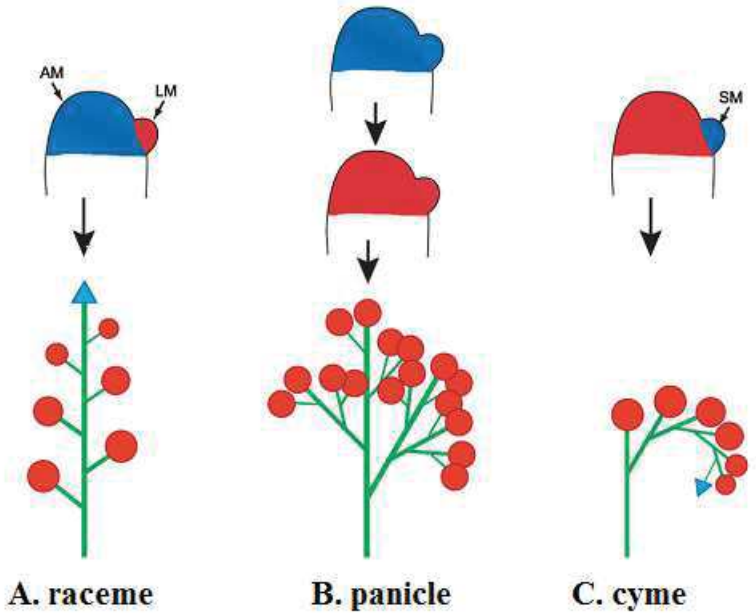


Figure 1.11 Schematic representation of the development and architecture of the three major inflorescence types. Top: diagrams showing the relative position and developmental fate or identity of apical and lateral meristems in distinct inflorescences. Red color indicates floral identity, blue color non-floral or shoot identity. Bottom: diagrams of fully developed inflorescences. Flowers are indicated by red circles, meristems by blue triangles. am, apical meristem; lm, lateral meristem; sm, (lateral) sympodial meristem. (From Rijpkema et al. 2010)

1.2.2.2 Internode length affect to the inflorescence architecture

Nodes and internodes were divided from stem – main structural axes of a vascular plant. The nodes is the area of a plant's stem from which one or more organs such as leaves, roots and spikelet meristem (in inflorescence architecture) grow; whereas the internode is the distance of one node from the other. In inflorescence architecture, internodes length is contributory factor to diverse inflorescence typologies. In rosette plants such as *Arabidopsis*, the transition from vegetative stage to reproductive stage is accompanied by internode elongation. The marked difference in internodes length between the vegetative and reproductive stems distinguishes the height of plant and whether the flowers are presented to pollinators. However, rather than the vegetative and reproductive portions of the plant, relative internode and pedicel lengths within the reproductive portion of the plant play important role to distinguish inflorescence typologies (Figure 1.12). A raceme is characterized by lateral flowers, with pedicels forming in sequential axils. They are separated by visibly identifiable

internodes. Pedicel length can reduce completely leading to sessile flowers and to an inflorescence called a spike. A radial increase of the growth stem in this case leads to a convertion of a spike to a spadix. Spadix type could be converted to capitulum type if the internode lengths reduce. In contrary, the elongation of pedicel length interact with the length of internode could make corymble and umble type of inflorescence architecture.

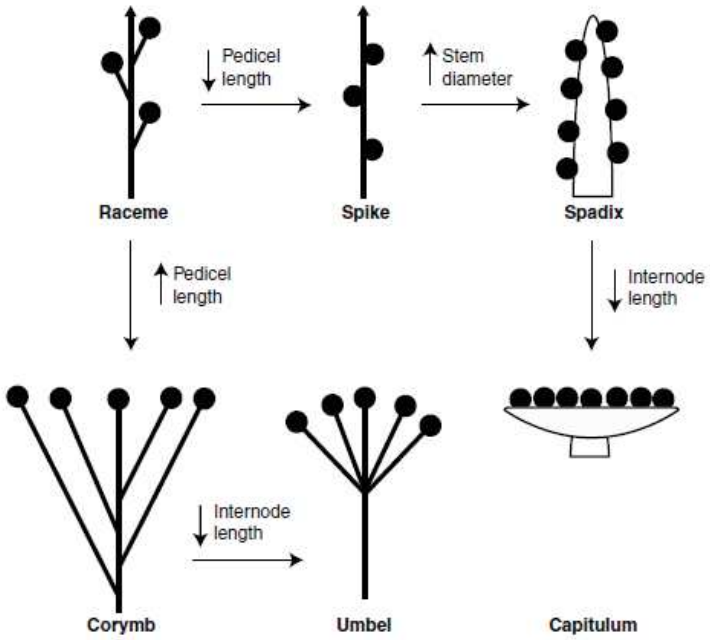


Figure 1.12: The length and diameter of stems and pedicels determine inflorescence architecture. (From Ainsworth 2008).

1.2.2.3 Phyllotaxy of inflorescence architecture

The term of *phyllotaxis* means "leaf arrangement" in Greek and was coined in 1754 by Charles Bonnet. 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. Similarity with the leave arrangement, the phyllotatic changes could produce the additional patterns structure in the 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). It suggests that the correspondence between changes in apex growth affecting whether an inflorescence is a raceme or a capitulum and parallel changes in the complexity of the spiral

phyllotactic pattern (Figure 1.13 A-B). Otherwise, Kirchoff has shown that Hofmeister's rule (1868) to demonstrate the early development of flower in *Phenakospermum guyannense* (Strelitziaceae) and *Heliconia latispatha* (Heliconiaceae). Both species have a cyme with three sepals and symmetric flowers. The cymes arise in the axils of primary bract and produce prophyll and successive flowers that mirror images of each other (Figure 1.13 C-D). The observation indicated that phyllotaxy not only applied to leaf arrangement, but also to organ placement in inflorescence and inflorescence architecture (Kirchoff 2003).

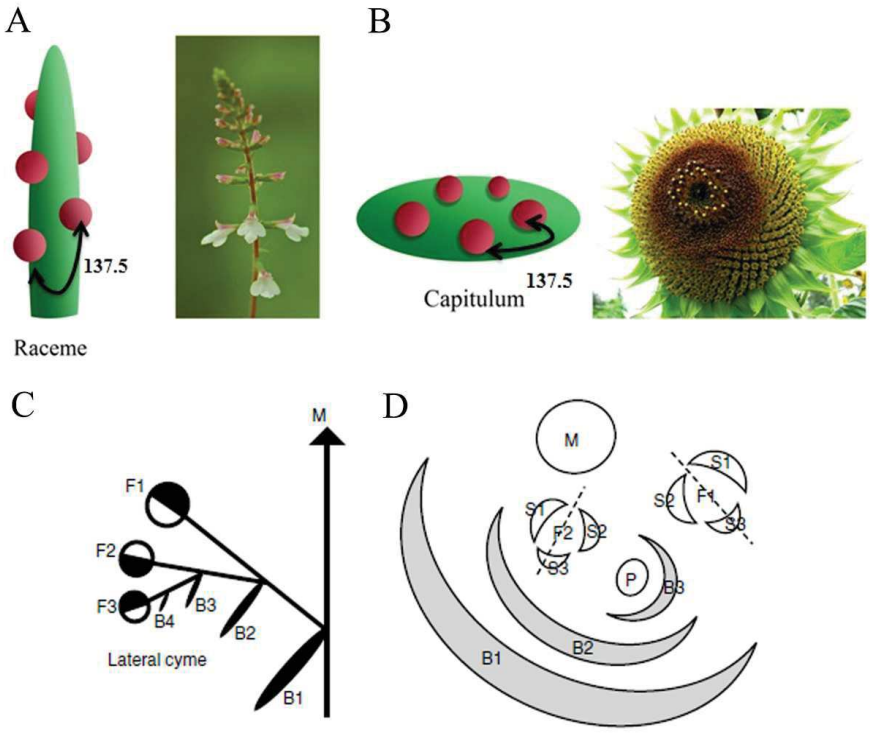


Figure 1.13: Inflorescence architecture is affected by phyllotaxy of axillary meristems. (A-B) Inflorescence architecture of raceme and capitulum are affected by phyllotaxy of axillary meristems initiation, respectively. (C) Schematic of a lateral cyme of *P. guyannens.* (D) Cross section of the lateral 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 2008).

1.2.3 Modeling of inflorescence architecture evolution

As mention in the previous section, the inflorescence architecture depends on when and where floral meristem identify. In the discipline of evolution and development, evolution changes were regulated by developmental time or "heterochrony" and were proposed to explain much of

morphological diversity (Slack and Ruvkun 1998). In case of inflorescence architecture, Prusinkiewicz et al. (2007) have presented a model of inflorescence evolution based on the difference in the time required for apical and lateral meristems to acquire floral fate.

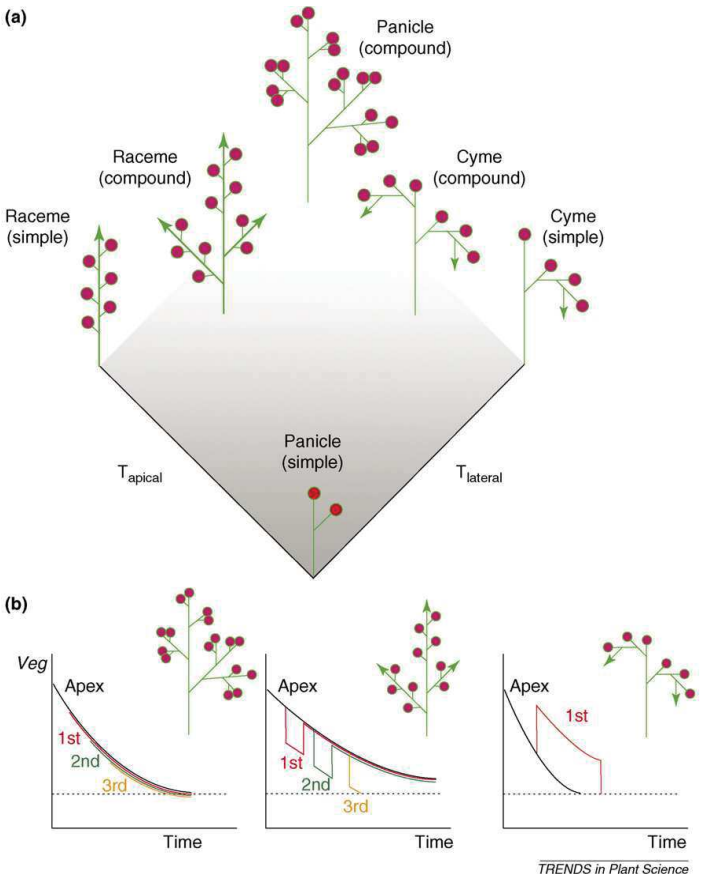


Figure 1.14: Development of distinct virtual inflorescence structures. (a) Structure of inflorescences and position in morphospace. Flowers are indicated by red circles and meristems by green arrows. The inflorescence types are positioned in a 2D morphospace defined by the time required for apical (T_{apical}) 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 (From Koes 2008).

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*" drop below a certain threshold, the meristems convert to the floret meristem. Depending on the tendency of "*veg*" timing during inflorescence development, panicle, cymes or racemes would be determined (Figure 1.14). The model was supported and improved by molecular genetic analysis, which identified several genes from

different species (*i.e. Arabidopsis*, petunia, tomato and rice) as factors related to *veg* parameter. This will be detailed in the following section.

1.2.4 Molecular bases of inflorescence architecture

According to Prusinkiewicz's model, the main types of inflorescences (ie. raceme, cyme and panicles) could be explaining by differential expression of a parameter "*veg"*. Among the molecular mechanisms that may be comprised by this parameter, is the conserved activity of *LEAFY* (*LFY*) and *UNUSUAL FLORAL ORGANS* (*UFO*) orthologous genes, which are regulated in a species-specific manner (Erik Souer et al. 2008; Moyroud et al. 2010).

In *Arabidopsis*, *LFY* encodes a transcription factor that promotes the transition from inflorescence meristem (indeterminate fate) to floral meristem (determinate fate) (Jack 2004; Irish 2010), whereas *TERMINAL FLOWER1 (TLF1)* suppresses this process (Shannon and Meeks-Wagner 1991; Bradley 1997). Considering Prusinkiewicz's model, *LFY* represses *veg* factor while *TLF1* promotes *veg*. Because *LFY* and *TFL1* down-regulate each other (Jack 2004), *TFL1* expressed in the apical meristem (Conti and Bradley 2007) whereas *LFY* 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 *Arabidopsis* 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 et al. 2007). The results suggest that expression of *LFY* and *TFL1* during inflorescence development lead to the raceme-type of *Arabidopsis*.

In contrast, the petunia *LFY* ortholog, *ABERRANT LEAF AND FLOWER* (*ALF*) expressed in both vegetative and reproductive stages (Souer et al. 1998; Molinero-Rosales 2000). 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, there is the ortholog of *Arabidopsis UFO* gene in petunia, namely *DOUBLE-TOP* (*DOT*), which plays an important role to identify FM in petunia. *DOT* and *UFO* encode F-box proteins that interact with *ALF* and *LFY* 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 seem to have been a key factor in the evolution of the distinct raceme type in *Arabidopsis* and cyme type in petunia inflorescence (Figure 1.15) (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 eudicot species. In contrast to the function of *LFY* and *UFO*,

their orthologs in grasses suppress the transition from IM to FM to determined the inflorescence morphology (Ikeda et al. 2007; Rao et al. 2008). This finding suggests a conserved mechanism of those genes among grass species (McKim and Hay 2010).

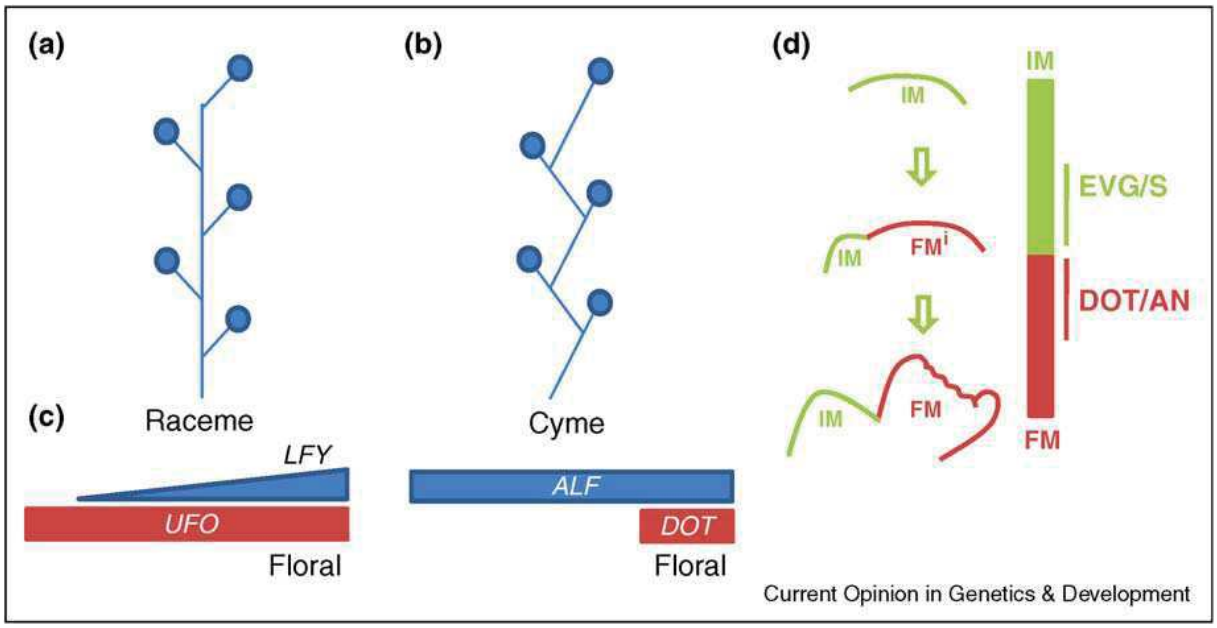


Figure 1.15: Modulation of a conserved mechanism regulates diverse flowering architectures. (a) Flowers (shown as blue circles) arise laterally from an apical IM in racemes. (b) In cymes (petunia), each flower originates as a lateral IM that transits into a FM after producing a new lateral IM, which 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 *Arabidopsis* racemes by *LFY* expression, and in Petunia cymes 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).

However, how the conserved mechanisms diverged during evolution, and to what extent that contributed to the evolution of distinct architectures? And how they produce the determinate inflorescence with zig-zag branching patterns in cymes? Recent work has identified *EVERGREEN* (*EVG*), which encodes a WOX (WUSCHEL-related homeobox) transcription factor which is essential for activation of *DOT* and specification of floral identity (Laux et al. 1996; Rebocho et al. 2008). Compare with *WOX* orthologs in *Arabidopsis*, *EVG* is expressed exclusively in inflorescence meristem and switching 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) expresses in inflorescence meristem until *ANANTHA* (*AN*, the *DOT* ortholog) expresses to identify the floral meristem (Lippman et al. 2008) (Figure 1.15 d). Because *EVG*/*S* and

DOT/*AN* were 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 those genes' expression provides a flexible mechanism to modulate the duration of the inflorescence phase before commitment to floral fate, and hence 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 that present long branches at the base of a main spike with spikelet pairs covering the long branch and main spike in contrary with rice where long branches bearing a single spikelet (Figure 1.16 A and 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) architecture has a transformation of spikelet pairs from short branches into long branches bearing single or paired spikelets (Vollbrecht et al. 2005) (Figure 1.16 C), whereas the ear produced more higher-order branched in *ra1* mutants, resulting in reduced fertility (Vollbrecht et al. 2005; McSteen 2006). Compare with *ra1*, *ra2* is expressed earlier during branching process in maize. The cross-examination 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 (Figure 1.16 B) (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 transcription of target genes.

In the context of inflorescence architecture evolution, the 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. The *RA1* orthologs were found in these species (Vollbrecht et al. 2005). Interestingly, deep analysis in early stage 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 (Kellogg 2007; McSteen 2006; Vollbrecht et al. 2005).

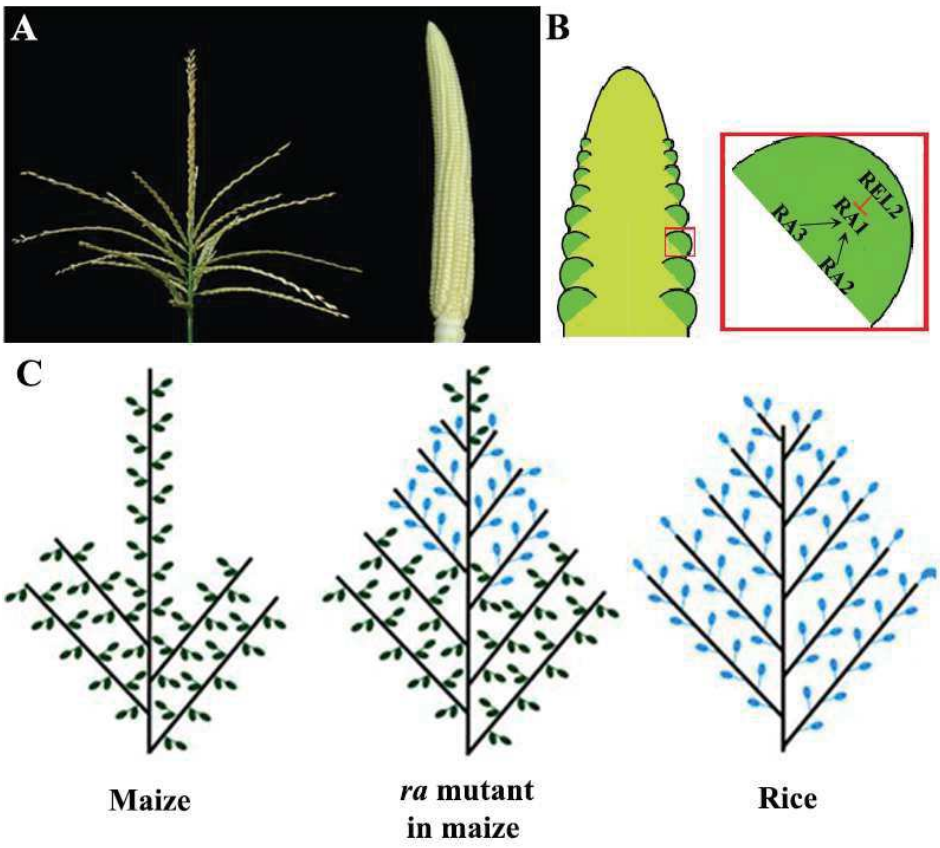


Figure 1.16: Inflorescence architecture of maize and ramose pathway. (A) Maize tassel (left) and ear (right); (B) *ramosa* pathway controlling the maize inflorescence architecture; (C) Simplified schematic of inflorescence morphology in maize and rice compared with the phenotype of the *ra* mutants in 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).

1.3 Rice panicle development

1.3.1 Importance of rice and rice domestication

The Poaceae family (Grasses) includes over 10.000 species displaying huge diversity of morphology. This family contains many important domesticated species, including: *Oryza sativa* (rice), *Triticum aestivum* (wheat), *Zea mays* (maize) and *Sorghum bicolor* (Sorghum). These crops feed much of the world through the grains produced by their inflorescence (Barazesh and McSteen 2008). Among grass family, rice provides the staple food for over half the world's population and about one billion depend on rice cultivation for their livelihoods. In 2013-2014, rice accounts for over 20 percent of global calorie intake and beyond providing approximately 20.4% produce of world cereal (FAO, http://www.fao.org/), rice also plays an important cultural role. Products of the rice plants are used for a number of different purposes, such as fuel, thatching, industrial starch, and artwork. In some countries like China and Viet Nam, it has been suggested that rice has been cultivated for 10.000 years, where it gradually rose to become an important part of aristocratic life. 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 cannot follow the increase of world population. In 2009, nearly one billion people were living in poverty, including 640 millions 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 lever rise, and other problems (pest, urbanization…) threaten rice product.

The rice genus (*Oryza*) comprises approximately 23 species and is represented cytogenetically by 10 genome groups: the A-, B-, C-, BC-, CD-, E-, F-, G-, HJ- and HK-genomes (Vaughan et al. 2003). The A-genome group, also called the *Oryza sativa* complex, consists of 8 diploid species. In this group, there are two cultivated species: Asian rice (*Oryza sativa* L.) and African rice (*Oryza glaberrima* Steud.) which were domesticated independently in Asia and Africa (Second, 1982). While Asian rice currently spreads worldwide as the world's largest food crop, African rice is grown primarily in tropical West Africa.

The Asian rice included two main subspecies (spp), *indica* and *japonica*, which were domesticated from the wild rice *Oryza rufipogon* more than 10 000 years ago. However the story of Asian rice domestication is still in debate. Through in-depth analyses of the domestication sweeps and genome-wide patterns, Huang et al. (2012) reveal that *O. sativa* ssp *japonica* rice was first

domesticated from a specific population of *O. rufipogon* around the middle area of the Pearl River in southern China, and that *O. sativa* spp *indica* rice was subsequently developed from crosses between *japonica* rice and local wild rice as the initial cultivars spread into South East and South Asia (Huang et al. 2012). Although, molecular phylogenetic analyses indicated that *indica* and *japonica* originated independently (He et al. 2011; Londo et al. 2006), the well-characterized domestication genes in rice were found to be fixed in both subspecies with the same alleles and genome-wide survey of SNP polymorphisms provided the strongest support for a single domestication origin of *O. sativa* (Li et al. 2006; Zhang et al. 2009; Molina et al. 2011; Huang et al. 2012).

Meanwhile, the cultivated African rice (*O. glaberrima)* derived from the wild species *O. barthii* from 2000 to 3000 years ago. The domestication of African rice is a single origin in areas of Upper Niger and Sahelian River (Sweeney and McCouch 2007; Li et al. 2011; Wang et al. 2014). However, compare with Asian rice, the genetic diversity in African rice is substantially lower (Ishii et al. 2001; Nabbotz et al. 2014; Orjuela et al. 2014; Wang et al. 2014). The fact was supported by the hypothesis that *O. glaberrima* is the product of a double evolutionary bottleneck. The first was associated with the divergence from Asian Oryza, perhaps ancestors of *O. barthii* were introduced to Africa from Asia. The second was due to African domestication (Li et al. 2011).

Asian rice currently spreads worldwide as the world's largest food crop, with over 90 percent of the world's rice is produced and consumed in the Asian Region by 6 countries (China, India, Indonesia, Bangladesh, Vietnam and Japan) comprising 80% of the world's production and consumption (Abdullah et al. 2006). Meanwhile African rice only grows primarily in tropical West Africa. With its distinct origin, African rice differs from its Asian counterpart in many qualitative and quantitative traits (Vaughan et al. 2008). Another interesting feature is that African rice varieties have many unique and useful traits such as weed competitiveness, tolerance to various abiotic stress (acidity, salinity and drought) and resistance to diseases/pests (Sarla and Swamy 2005). The difference between characters of Asian and African rice can be used as good sources of useful traits in the breeding program. For instance, the new varieties, named "New Rice for Africa" (NERICA), are a cross between *O. glaberrima* and *O. sativa.* They combine the hardiness of the African species with the productivity of the Asian species to provide new varieties that are high-yielding, drought- and pest-resistant and adapted to the growing conditions of west Africa (Linares 2002; Sarla and Swamy 2005)

1.3.2 Rice panicle architecture

Rice panicle has three types of inflorescence meristems: rachis meristem, branch meristem and floral meristem. The panicle structure consists of a main rachis with some primary branches, higher order branches (*i.e.* secondary branch, tertiary branch) and spikelets that are the primary unit of the grass inflorescence comprising glume (bract-like organs) and florets. The floret consists of a pair of lemma and palea, lodicules (equivalent to eudicot petals), stamens and a carpel (Figure 1.17). In rice, one spikelet meristem produces one floral meristem, therefore the number of spikelet meristem will determine the number of grain per panicle (Yoshida and Nagato 2011). In the panicle, the basal end of the rachis meristem is easy found by the presence of a trace of bract. In rice, the rachis meristem is not converted to a spikelet meristem, but aborted and remains as a vestige that is detected near the base of the uppermost primary branch.

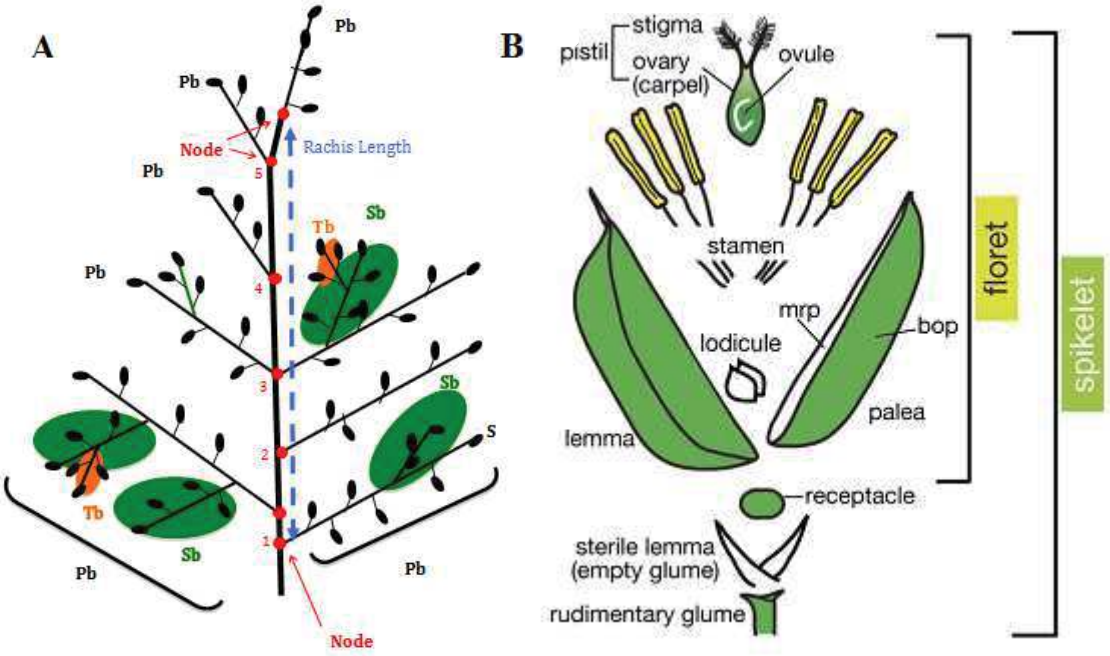


Figure 1.17: Schematic representations showing the structure of the rice inflorescence (A) and spikelet (B). Pb, primary branch; Sb, secondary branch; tb, tertiary branch; mrp, marginal region of palea; bop, body of palea. (Adapted from Yoshida and Nagato 2011).

The panicle structure or complexity changed in domesticated species from a panicle with few primary and secondary branches bearing relatively few grains, to a highly branched panicle carrying larger numbers of seeds than the wild ancestors. The cultivated rice species *O. sativa* and *O. glaberrima* tend to generate higher-order branching and more mature seeds compare to their wild relatives *O. rufipogon* and *O. barthii*, respectively (Yamaki et al. 2010). The difference of panicle

structure is also significant between the two domesticated species, *O. glaberrima* showing an intermediate panicle complexity between *O. sativa* and the two wild-relatives *O. rufipogon* and *O. barthii*. However, no deep description has been done yet for the wild rice species and for the cultivated African rice.

1.3.3 Developmental time-course of Asian rice panicle

The transition from vegetative to reproductive phase in *O. sativa* occurs in response to an environmental change such as short day length and high temperature. After producing the last foliage leaf, the SAM is converted to a rachis meristem. Compare with vegetative meristem, rachis meristem becomes taller and slightly wider when producing the first bract (Figure 1.18 A-C). In the SAM, the leaf primordium 2 (P2) is higher than the shoot meristem and covers more than half of it; when the primordium is just formed (Figure 1.18 B). However, when the first bract primordium is formed, the tip of the flag leaf primordium is higher than the rachis meristem (Fig 1.18 D).

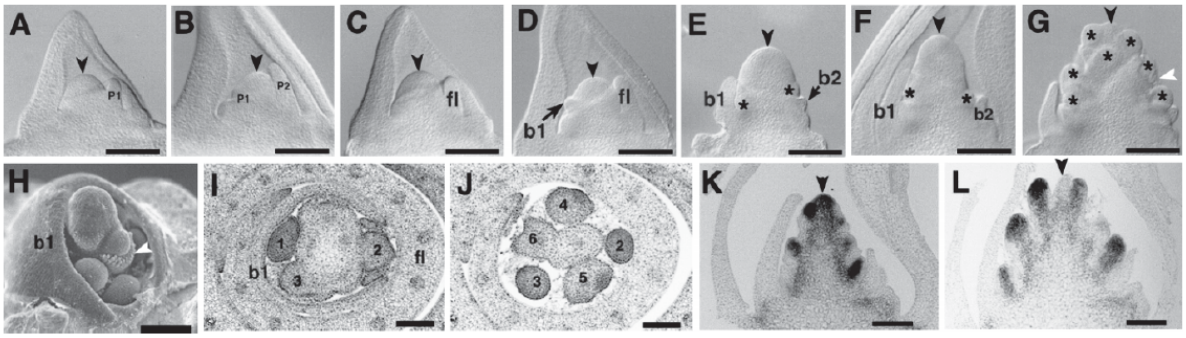


Figure 1.18: Early stages of panicle development in rice. (A)-(G) Cleared shoot and inflorescence apices (H): SEM image of young inflorescence (I)-(J): Rachis apex (K)-(L): Expression of *OSH1* in rachis apex *: primary branch meristem; fl: flag leaf, b1: first bract, b2: second bract. Bar=150µm for (A) to (G), 100µm for (H) to (L) (from Ikeda et al., 2004).

After the bract 2 and first primary branch are formed, ten or more bracts and primary branches are rapidly produced in spiral arrangement (Figure 1.18 E-H). In the early reproductive phase, a drastic change from ½ alternate to spiral phyllotaxy occurs. This change does not occur abruptly but proceeds rather gradually. The divergence angles of the first two bracts are slightly smaller than 180° (Figure 1.15 I), about 160° - 170°; they gradually converge to 144° (Figure 1.18 J). The direction of spiral is either clockwise, depending on which margin of flag leaf primordium becomes inside. Normally, rachis meristem is abort at early stage after producing ten or more primary branch primordial. In Figure 1.18 K, mRNAs of the ortholog of Arabidopsis *STM* gene namely *OSH1* were detected in rachis meristem during the production of primary branch primordial. However, soon after the last

primary branch, the *OSH1* expression disappeared from the rachis meristem (Figure 1.18 L). The results indicate that the rachis meristem lost its activity after producing a cultivar-specific number of primary branches (Ikeda et al. 2004).

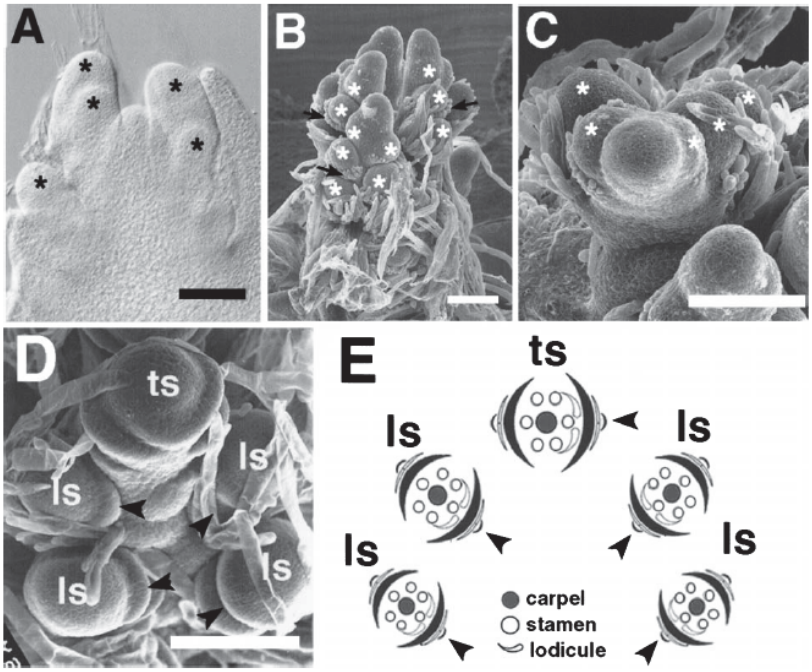


Figure 1.19: Later stages of panicle development in rice. (A) Elongation of primary branches (black star). (B) Differentiation of secondary branch primordial (white star). (C) Top view of (B) showing biased distichously phyllotaxy of secondary branch (white star). (D) Spikelet formation in primary branch 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: 100µm (From Ikeda et al. 2004).

When all branch primordia are formed, the rachis meristem loses its activity and aborts. Then all branch meristems elongate almost simultaneously (Figure 1.19 A). When primary branches elongate to some extent, secondary branches are formed in the basal regions (Figure 1.19 B). The secondary branches might produce tertiary branches. The fate of branch meristem differs from that of a rachis meristem as they are invariably 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 with a divergence angle of about 110° (Figure 1.19 D-E). Since spikelets are differentiated, two rudimentary glumes are formed, followed by lemma and palea formation. Then the remaining spikelet meristem converts into a flower (floret) meristem to form floral organs (two lodicules, six stamens and one carpel) (Figure 1.19 E). The inflorescence

remains short (<4 cm) at this stage (Itoh et al. 2005). But then the rachis and branches start rapid elongation after floral organ primordial are differentiated. Maturation of anthers and ovules takes place during rapid branch elongation.

To understand the making of 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 developmental course of rice panicle in particular and grass inflorescence in general has a variety of stage-specific landmark events (Ainsworth 2008; Xing and Zhang 2010; Zhang and Yuan 2014) (Figure 1.20). As described above, there are three main stages of panicle development: transition, branching and spikelet differentiation. These different stages are associated with specific gene expressions relative 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 (Figure 1.20). These genes have been evidenced in *O. sativa* mainly through the characteristic of mutants. However, in some cases, these genes were evidenced by QTL characterization.

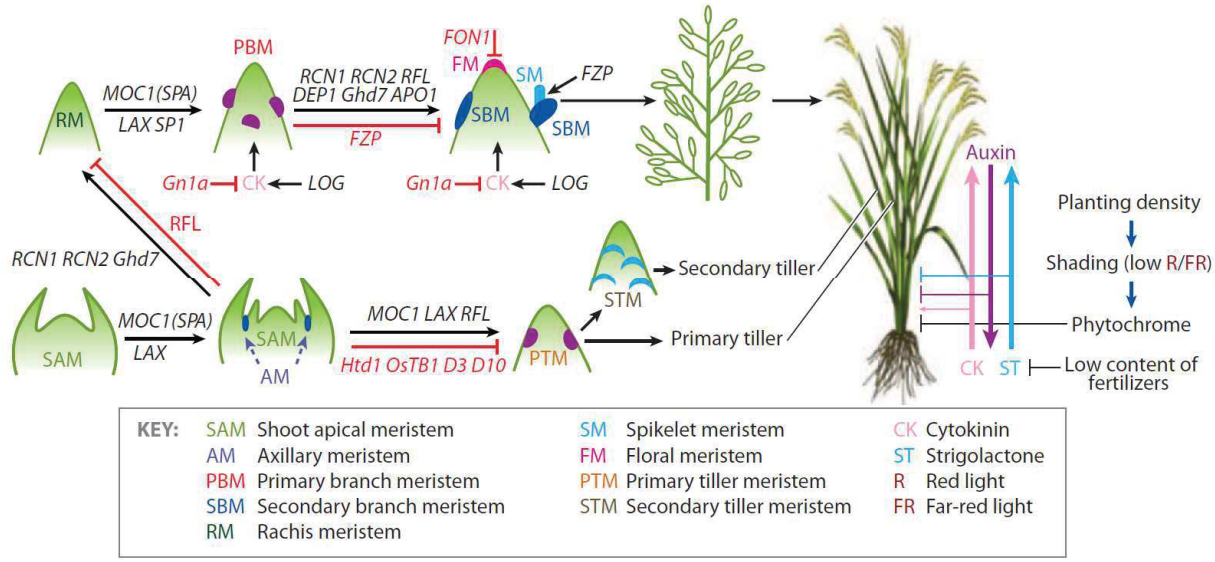


Figure 1.20: A schematic representation of genes involved in tillering and panicle formation (from Xing and Zhang 2010).

1.3.4 Genetic control of tillering

Tiller is a specialized grain-bearing branch formed on the un-elongated basal internodes on the main stem of the plant. From an agronomic point of view, grain yield is usually contributed by the primary tillers and some early secondary tillers, whereas tertiary and late secondary tillers make little

contribution, although they also consume nutrients, water, and photosynthates (Li et al. 2003). In breeding programs, there is a trend to breed for dwarf plant, low tiller number with few unproductive tillers, less panicle but more branches with more grains to achieve high yield (Khush 2001; Peng et al. 2008). Among grass species, although the rice tillering pattern is obviously different from that of *Arabidopsis*, the formation of a tiller undergoes the similar process as the tillering of *Arabidopsis*. This corresponds to a two stages process: the initiation of axillary bud at each leaf and its outgrowth (Wang and Li 2011), and after the formation, tillers produce short internodes, leaves with juvenile characteristic, and activate the axillary meristems during vegetative development. The tillers are responsible for the number of reproductive inflorescences and yield of grass such as rice, wheat and barley. In other grass, as maize, plants were selected during domestication to produce few tillers to concentrate resources in the main shoot

Concerning the initiation of tillering in rice, *monoculm1/small panicle (moc/spa)* mutant plants formed only the main shoot without any axillary meristem. *MOC1* encodes a transcription factor of the GRAS family orthologous to *LAS* of *Arabidopsis* and *LATERAL SUPPRESSOR (LS*) of tomato (Li et al. 2003). Moreover, *moc1* and *ls* mutants have defect during branching process of inflorescence development, which lead to reduction in the number of branches and spikelets. Comparison of the *las* and *moc1* mutants implied that *MOC1*/*LS*/*LAS* play a conserved role in the initiation and maintaining of the tillers (McSteen and Leyser 2005). The other mutations reported in rice and *Arabidopsis* are *more axillary branching* (*max*) and *dwarf* (d) involved in controlling the outgrowth of axillary meristems (Zou et al. 2005; Ishikawa et al. 2005). Both mutants are characterized by an increase of the number of branches and reduced plant height. The parallel studies on rice and *Arabidopsis* revealed that the phenotype of such mutants were related by the signaling of strigolactones (Gomez-Roldan et al. 2008; Umehara et al. 2008). In *Arabidopsis*, at least four genes were reported that involved to the strigolactone-mediated shoot branching: *MAX1, MAX2* and *MAX4* are related in the biosynthesis of strigolactones whereas *MAX2* is responsible for the strigolactones signal (Gomez-Roldan et al. 2008; Booker et al. 2005). Excepted *MAX1* which was not identified in rice, the homologs of *MAX2, MAX3* and *MAX4* in rice are *DWARF3 (D3), DWARF17* (*D17*)/*HTD1* and *DWARF10* (*D10*), respectively (Ishikawa et al. 2005; Zou et al. 2005).

In addition, other genes identified in rice as involved in strigolactone-dependent pathway are *DWARF27* (*D27*) and *DWARF14* (*D14*). D27 encodes a novel iron-containing protein that related the biosynthesis and *D14* encodes for a hydrolase/esterase downstream of strigolactones synthesis and may participate in the conversion of strigolactones to their bioactive form (Arite et al. 2009; Lin et al. 2009; Wang and Li 2011; Waters et al. 2012). With the study about function of strigolactones during

tillering, there are three classes of hormones implicated in the regulation of bud outgrowth: auxin, cytokinin and strigolactones (Figure 1.17). These hormones are transported through out the plant, auxin is produced mostly in 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 the axillary bud outgrowth (Brewer et al. 2009). However, in recent study on *Arabidopsis*, *max* mutants showed that strigolactones may control the outgrowth of axillary buds through PAT (Crawford et al. 2010). This model seems to be similar in rice since in the loss of function mutant *d27*, PAT was significantly increased (Lin et al. 2009).

In maize, *Teosinte branched1* (*tb1*) gene has been identified as a major contributor to the evolutionary changes in maize during its domestication from teosinte. This gene acts to repress the growth of axillary organs and to enable the formation of the female inflorescences (Doebley et al. 1997). The rice *TB1* gene (*OsTB1*), homolog of the maize *TB1,* encodes putative transcription factor carrying a basic helix–loop–helix type of DNA-binding motif, named TCP domain. Transgenic rice plants overexpressing *OsTB1* show greatly reduced lateral branching without affecting the initiation of axillary buds, whereas a loss-of-function mutant of *OsTB1* exhibits enhanced lateral branching, indicating that *OsTB1* functions also as a negative regulator for lateral branching in rice (Takeda et al. 2003). Moreover, in the *moc1* mutant, Li and al. (2003) investigated that the expression of *OsTB1* is significantly reduced suggesting that *OsTB1* is acting downstream of *MOC1*. These examples indicate the conservation of mechanisms for controlling axillary bud activity between monocot and eudicot plants. However, surprisingly, none of these vegetative branching mutants in grass species were reported as affected in inflorescence branching pattern, suggesting that these two processes may have at least partly diverged in term of regulation.

1.3.5 Axillary meristem initiation during inflorescence development

The complex architecture of plants is mainly controlled by the pattern of axis formation. Concerning the generation of axillary meristems (AMs) and the growth of branch meristem in a basic process, several genes have been reported through analysis of mutants. In rice, beside *monoculm1,* other mutants such as *lax panicle1 (lax1), lax panicle2 (lax2),* and *frizzy panicle (fzp)* as well as two gene involves control auxin signal *OsPIN1* and *OsPID* have been shown to affect the patterning of

AMs and the panicle development (Komatsu et al. 2001; Komatsu et al. 2003; Xu et al. 2005; Woods et al. 2011).

LAX PANICLE1 (LAX1) encodes 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 Kyozuka 2009). BA1/LAX Absence of lateral and terminal spikelet meristems and decrease of branch meristems in *lax1* mutant 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; Komatsu et al. 2003). In addition, the accumulation of *LAX1* protein in axillary meristem formation is subjected to a two-step regulation related to a non-cell autonomous mode of action. In the first step, *LAX1* gene is expressed in the axils of leaves at stage 4 of plastochoron (P4), and then *LAX1* protein is trafficked to the whole axillary meristem around this stage (Oikawa and Kyozuka 2009). In a recent research, Tabuchi et al. (2011) reported that *lax panicle2* (*lax2)* mutant has altered in AM formation. *LAX2* encodes a nuclear protein that contains a plant-specific conserved domain and physically interacts with *LAX1* and plays role as a novel factor that acts together with *LAX1* in rice to regulate the process of AM formation. Similarly with *LAX1*, its ortholog in maize, *BARREN STALK1 (BA1)*, 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 et al. 2002; Gallavotti et al. 2004). In addition, by using phylogenomic and comparative expression analyses, Woods et al. (2011) reported that *BA1/LAX1* clade expresses in both of monocots and eudicots. This suggests a conserved mechanism of *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. 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 (Zazimalova et al. 2007). In rice, *OsPIN1* also functions in PAT, and over-expression or suppression of the *OsPIN1* expression through a transgenic approach resulted in changes of 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 *Arabidopsis* (Friml et al. 2004). It was reported that *PID* carries out its function through the control of sub-cellular localization of PIN proteins, which direct the flow of active auxin transport. Morita and Kyozuka (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 were probably caused by disturbance of PAT and that *OsPID* is involved in the control of auxin fluxes. Mutants with defects in *PIN1* or *PID* have similar (Morita and Kyozuka 2007).

In another mutant named *frizzy panicle (fzp),* the formation of florets is replaced by sequential rounds of branching, such as several rudimentary glumes are formed in place of the spikelet (Komatsu et al. 2001). In addition, all meristems remain undifferentiated during the early development. Therefore, the degeneration of AMs might occur during the maturation stage when internodes elongate. Thus, all AMs of *fzp* mutant do not develop into a branch remain unclear. The *fzp* mutant phenotype suggests that *FZP* is required to prevent the formation of AMs within the spikelets meristems and permits the subsequent formation of branch meristem identity. *FZP* encodes a protein containing the ethylene-responsive element binding factor (ERF) domain and is the rice ortholog of the maize *BRANCHED SILKLESS1* (*BD1*) gene, which controls spikelet meristems formation in this species (Komatsu et al. 2003). Interestingly, ERF domain is conserved in different grasses and is expressed in a distinct domain of the spikelet meristem. Its expression pattern suggests that signaling pathways regulate meristem identity from lateral domains of the spikelet meristem.

1.3.6 Axillary meristem outgrowth during inflorescence development

During inflorescence development, the transition from inflorescence/branch meristems to spikelet meristems is one of the key events during the establishment of the inflorescence architecture in grasses. The regulators of this transition in rice include *ABERRANT PANICLE ORGANIZATION 1 (APO1), APO2* and *TAWAWA1 (TAW1).* The *aberrant panicle organization1* (*apo1*) mutant forms small panicle with reduced numbers of branches and spikelets. In addition, *apo1* mutant exhibits abnormal floral organ identity and a loss of floral determinacy (Ikeda et al. 2005). The phenotype of *apo1* mutant indicated that the *APO1* gene suppresses precocious conversion of branch meristems to spikelet meristems, thus ensuring a number of spikelets. 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 in meristem size 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 plastochoron and the formation of floral organs. *APO1* encodes a F-box protein, an ortholog of *UNUSUAL FLORAL ORGAN* (*UFO*) from *Arabidopsis* (see section 1.3.1). In *Arabidopsis*, *UFO* is proposed to activate *LEAFY (LFY)* to promote FM fate through direct protein interaction in a proteasome-dependent manner (Chae et al. 2008).

The *RFL* gene (recently renamed *APO2*) was identified as the ortholog of *LEAFY/FLORICAULA* gene (*LFY/FLO*) from *Arabidopsis thaliana* and *Antirrhinum majus* respectively. As in *Arabidopsis*, *APO2* interacts with *APO1* at the molecular 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 (J Kyozuka et al. 1998; Rao et al. 2008; Ikeda-Kawakatsu et al. 2012), LFY/UFO promote the initiation of floral meristem to determining Arabidopsis morphology (McKim and Hay 2010). The finding suggests that genetic mechanisms for controlling inflorescence architecture have evolutionarily diverged between rice (monocots) and Arabidopsis (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. 2012). 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 (2012) 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 and McSteen 2008; Zhang and Yuan 2014). Such phytohormone balance also regulates the panicle branching. The 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 spikelets number and thus enhancing 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, *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). *LONELY GUY (LOG)* gene encodes an enzyme that catalysis the final step of

cytokinin biosynthesis within 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 indirectly affect branching in monocots and eudicots. Otherwise, they suggest cytokinin metabolism and signaling in grasses contribute to grain yield and provide a strategy for breeding programs to improve crop yield (Ashikari et al. 2005)

Two remaining genes characterized as QTLs related to grain yield have been isolated are *DENSE AND ERECT PANICLE* (*DEP1*) and *IDEAL PLANT ARCHITECTURE1* (*IPA1*)*/WEALTHY FARMER'S PANICLE* (*WFP*). *DEP1* encodes an unknown protein containing the PEBP (phosphatidylethanolamine-binding protein) domain, and this gene is pleiotropically responsible for all three 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 increased 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 *Osa-miRNA156*. Higher expression of *OsSPL14* in the reproductive stage promotes panicle branching and higher grain yield in rice (Miura et al. 2010). The *ipa1* allele harbors a point mutation within the *Osa-miR156* target site and thus perturbing the *Osa-miR156*-directed regulation of *IPA1* in rice plants (Jiao et al. 2010). The *wfp* allele harbors 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 a higher level of panicle branching. Interestingly, another microRNA named *Osa-miR529* showed sequence similarity with the *Osa-miR156* families and share common targeting of *OsSPL14*. Moreover, it was shown that *Osa-miR529* is in fact the major contributor of *OsSPL14* cleavage in panicle (Jeong et al. 2012).

1.3.7 The duration of spikelet differentiation

The duration of panicle development refers to the period from the first bract primordium to heading. Morphologically, the size of a panicle is determined to a large extent by the appearance of the terminal primary branch. *TERMINAL FLOWER1 (TFL1)/CENTRORADIALIS (CEN)-like* genes play important roles in determining plant architecture, mainly by controlling the timing of phase transition in *Arabidopsis* and *Antirrhinum*, respectively (Conti and Bradley 2007). Mutation of *TFL1* and *CEN* converts branch meristems into terminal flower. In contrast to theses loss of function phenotypes, ectopic overexpression of *TLF1/CEN-like* genes conferred basically opposite effects, leading to late

flowering and more branches in *Arabidopsis* (Nakagawa et al. 2002). In rice, ortholog genes of *TLF1/CEN* are *Reduced Culm Number1* and *2* (*RCN1, RCN2*) of which overexpression delayed the transition of reproductive phase up to 2 months compared with wild-type plants (Nakagawa et al. 2002). Detailed observations of the 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 contrary, knocking down RCN genes resulted in much smaller panicle with reduced branches (Liu et al. 2013). These results suggest that *RCN* coordinate panicle development and flowering time (Nakagawa et al. 2002).

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

Members of the *AP2* 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 panicle architecture and formation of 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 floral meristem (Chuck et al. 2007; Lee et al. 2010). *SNB* and *OsIDS1* are expressed throughout the branch and spikelet meristems. When those meristems were initiated, *Osa-miR172* expression was recruited, and subsequently this depleted *SNB* and *OsIDS1* transcripts, ensuring spikelet development at the correct position and time (Lee and An 2012). These data show the importance of a balance between *Osa-miR172* and *AP2* family genes in the determination of FMs.

The two rice *AP1-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 member of 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 was detected in the panicle meristem, in branch meristems immediately after the transition, and in glumes primordial, consistent with its role in the early development of spikelet formation. Recently, it has been reported that *PAP2/OsMADS34* controls the transition to spikelet meristems (Gao et al. 2010; Kobayashi 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. 2010). In addition, the knockdown of the three *AP1*-like genes (*OsMADS14*, *OsMADS15* and *OsMADS18*) did not significantly affect inflorescence development. On the other hand, the elimination of PAP2 function in the triple knockdown plants severely impeded transition of the SAM to the IM suggests a combined action of the three *AP1*-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. The result proposes that *PAP2* and the three *AP1*-like genes coordinately act in the meristem to specify the identity of the IM downstream of the florigen signal.

1.3.8 Floral organ patterning

From 1980s, forward the earlier flower mutants in two model eudicot species, *Arabidopsis* and Antirrhinum, Coen suggested a molecular models of how floral meristem and organ identity be specified, called ABC model (Coen et al. 1991). The ABC model indicated the overlapping domains of three classes, the sepal and petal were affected by class A genes, the petal and stamen were affected by class B genes, whereas the stamen and carpel was identified by class C genes (Carpenter and Coen 1990; Coen and Meyerowitz 1991) (Figure 1.21 A). Then the model was further extended to the ABCDE (or ABCE) model, including 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 for the function of ABC class genes (Theissen and Saedler 2001; Rijpkema et al. 2010a; Causier, Schwarz-Sommer, and Davies 2010) (Figure 1.21A).

In *Arabidopsis*, the A class genes are represented by *APETALA1* (*AP1*) and *APETALA2* (*AP2*), which are determined the initiation of sepal and petal. However, a little is known about the *AP1* and *AP2* homologs of Oryza. 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 in *SUPERWOMAN1* (*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). The similar results have been obtained from genetic analysis of *AP3* ortholog in maize. Taken together, the finding suggests that *AP3* and *PI* in class B are deeply conserved among grass species (Causier et al. 2010; Rijpkema et al. 2010).

The *AGAMOUS* C class gene in *Arabidopsis* (*AG*) is necessary for stamen and carpel identity, but is also required to specify the determinacy of the floral meristem (Kramer et al. 2004). One of the paralogs that has been estimated in *Oryza* is *OsMADS3* which is expressed in the developing stamens and carpels and promotes the determinancy of floral meristem (Kyozuka 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 *Arabidopsis*, namely *CRABS CLAW* (*CRC*), determines the 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 *Arabidopsis* corresponds to the gene *SEEDSTICK* (*STK*) while the E class is comprised of a set of four paralogs known as *SEPALLATA1* (*SEP1*)*, SEP2, SEP3*, and *SEP4*, that are cofactors in complexes with other MADS box factors that determine floral organ identities and meristem determinacy (Pelaz et al. 2000; Ditta et al. 2004). 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. Notably, the *Oryza LEAFY HULL STERILE* (*LHS1*)/*OsMADS1*, from a subgroup of *LOF-SEP* genes, contributes to the identity of the palea and lemma as well as to meristem determinacy and the structure of the inflorescence (Figure 1.21 B) (Prasad et al. 2005; Jeon et al. 2008). Khanday et al. indicated that *OsMADS1* integrates transcriptional and signaling pathways to promote rice floret specification and development by negatively regulating *PAP2*/*OsMADS34*. In early stage of panicle development, *LHS1* promotes for the transition from branch meristem to spikelet meristem. *LHS1* also regulates auxin transport, signaling, auxin-dependent expression and three cytokinin A-type response regulators (Khanday et al. 2013). In addition, a broad comparative study of expression patterns of *LHS1* orthologs across the grasses has revealed a high degree of variability, both within and between florets (Malcomber and Kellogg 2004).

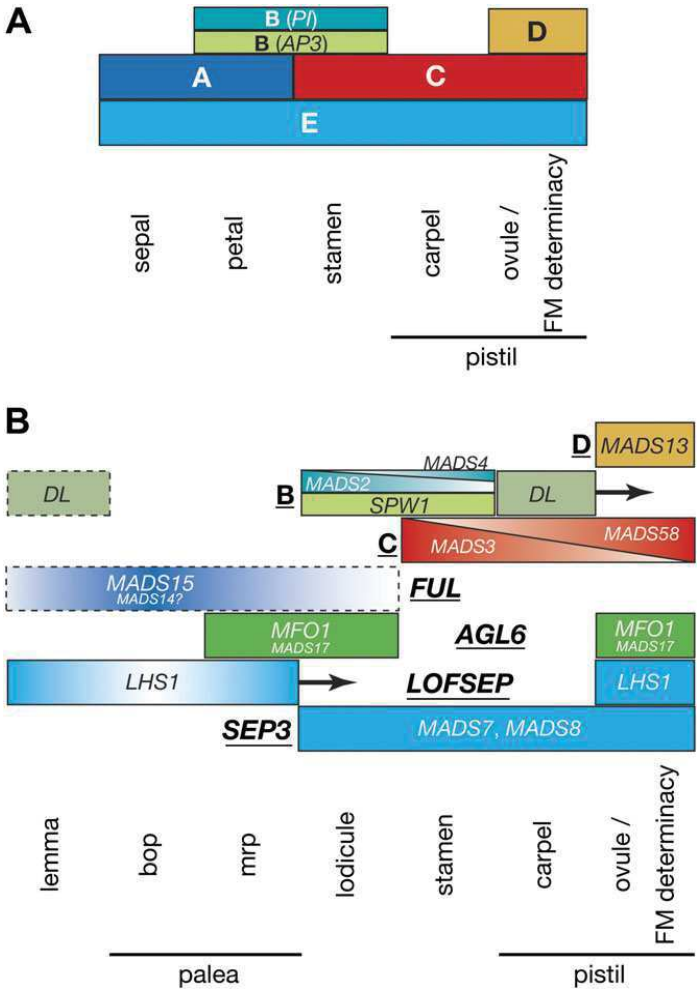


Figure 1.21: Genetic models for flower development in *Arabidopsis* (A) and rice (B). (A) Combinatorial actions of A-, B-, C-, D-, and E-function specify sepal, petal, stamen, carpel, and ovule identities in *Arabidopsis*. Heterodimerization of *AP3* and *PI* is necessary for B-function. (B) Underlined 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 *MADS15* 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).

1.4 Objectives

The diversity of panicle architecture is notable between the Oryza species from the AA genome group. This group includes 8 diploid species with two crop species, namely *O. sativa* and *O. glaberrima*, *O. rufipogon*, *O. meridionalis*, *O. barthii*, *O. longistaminata* and *O. glumeapatula* (Vaughan et al. 2003). The species complex presents the particularity to possess two cultivated species: *O. sativa* deriving from *O. rufipogon* in Asia, and *O. glaberrima* deriving from *O. barthii* in Africa. Moreover, an inter and intra-specific natural variations of morphological traits is observable in the different species both wild and crop species on the different continents. The panicle structure (or complexity) changed in domesticated species from a panicle with few grains, to a domesticated highly branches panicle carrying larger numbers of seeds. The difference in panicle structure is also significant between the two domesticated species, *O. glaberrima* showing an intermediate panicle complexity between *O.sativa* and the two wild-relatives *O.rufipogon* and *O.barthii*.

As detailed in the previous sections**,** a large set of genes related to panicle development and architecture were characterized in *O. sativa*, the Asian rice crop. These genes were identified in the context of mutant analyses and intra-specific diversity through QTL characterizations. However, only few panicle-associated QTLs were identified in the context of rice domestication (crop vs. wild relative) (Furuta et al. 2014) . The main point is that none of these genes were studied in the context of the evolution of panicle structure and rice domestication. Moreover, none of the orthologous of the panicle-related genes were reported up to date in African species.

In this context, several questions still needed to be addressed:

1. What are the morphological factors contributing to both inter- and intra-specific diversity of panicle architecture in Asian and African rice species?
2. What are the molecular bases of the inter-specific diversity of panicle structure notably in the context of domestication (crop vs. wild relative)?
3. What is the molecular basis of the phenotypic convergence observed for both Asian and African domestications (i.e. higher panicle complexity)? To which extend, the genes targeted during the two independent domestication processes were similar or not?
4. What are the molecular factors associated to intra-specific diversity of panicle structure in Asian and/or African rice? To which extend are they similar?

In the context of my PhD, I aimed to address some of these questions or at least to get results to partly

address them. I decided to focus on two points/approaches:

1. Comparative analyses of panicle morphology and genome expression in the context of inter-specific diversity (i.e. domestication).
2. Molecular and phenotypic bases of the intra-specific diversity in a crop species.

My work mainly focused on African rice species in IRD-Montpellier. In the context of a partnership between IRD-Montpellier and my previous lab in the Agricultural genetics Institute (AGI) in Ha Noi (Viet Nam), an International Joint Laboratory (LMI RICE) was initiated three years ago. This lab involved my two supervisors Pr. Pascal Gantet (as head of the LMI) and Dr. Stefan Jouannic (as PI of the panicle project developed in there). Consequently, my PhD was done spending 9 months in France and 3 months in Vietnam per year.

In this context, **the first objective** of my thesis project was to analyze the main **morphological events,** which govern panicle structure diversity between the two African rice species (*O. glaberrima* vs. *O. barthii*). This work was performed at early stages of panicle development through histological analysis; The **second objective,** corresponding to the main objective of my PhD project, was to carry out a comparative analysis of **genome expression** between the two African species with the aim getting a better understanding of the molecular basis of the panicle diversity between the two African species and how it could be related to the domestication. For this purpose, I developed a comparative analysis of expression pattern of orthologs of *O. sativa* landmark genes related to panicle development and more especially to branching process and meristem fate control. The idea was to determine to which extent an alteration of their expression patterns (spatial, timing and/or quantitative) might be related to panicle structure diversity between the two species (section 2.2.3). This approach was completed by genome sequence analysis and also by a transcriptomic analysis on the small RNA population through Next Generation Sequencing (NGS) in order to determine to which extent this class of RNAs may be associated to panicle structure diversity (section 2.2.2). The **third objective** was developed in the context of the partnership between IRD-Montpellier and LMI RICE in Viet Nam. A collection of **Vietnamese Landraces** from *O. sativa* was developed and I initiated the **phenotypic diversity** of the panicles in this collection with the aim to apply genome-wide association studies (GWAS) in order to evidence genomic regions potentially of interest to explain the intra-specific diversity of panicle structure in Asian rice and for selection-assisted markers related to yield improvement breeding programs in Viet Nam (section 2.3).

2.

RESULTS

2. RESULTS

2.1 Introduction

Rice panicle shows both inter-specific (crop vs. wild relative) and intra-specific phenotypic diversity (Asian or/and African rice). The observation rise a question about the nature of the morphological and molecular factors contributing to the variety of panicle architecture in Asian and African rice species, as well as about molecular mechanism related to panicle modification during rice domestication. Firstly, in order to understand which early morphological events may underlie the differences of panicle architecture at mature stage, I complete a description of the morphological events occurring during panicle establishment and development between the 2 of African rice species *O. glaberrima* and *O. barthii*. Secondly, deep sequencing of small RNA transcriptomes of the two species was performed to characterize the expression of small RNAs during rice panicle development between the domesticated species and its wild-relative. Moreover, a detailed analysis of the spatial and temporal expression patterns of a set of orthologs of *O. sativa* landmark genes related to meristem activity and meristem fate control in African rice was conducted.

Moreover, taking advantage from a parallel project which supplies information about genetic diversity of Asian rice collection (including 188 Vietnamese *O. sativa* landrace accessions and 40 reference accessions), we carried out an analysis of the diversity of panicle architecture of these landraces. Rather than understanding phenotypic bases of the intra-specific diversity in a Vietnamese rice collection, the project purpose is to identify genomic regions of possible significance in the determination of panicle diversity within *O. sativa* Vietnamese landraces, which may be suitable for marker-assisted selection for breeding programs related to yield potential.

My results are presented in two parts in the following section. In the first part, will be presented two manuscripts (as first author) related to the morphological and genome expression analyses in *O. glaberrima* and *O. barthii*. In the first manuscript we report on the heterochronic alteration of *miR2118*-trigerred phasiRNAs expression during panicle development related to the time-shift of panicle meristem state. The bio-informatic part of this work was mainly conducted by F. Sabot and I carried out the wet-lab work. This manuscript has been submitted to Plant Physiology journal and is still under reviewing process at this time. In the second one, we report on the conservation of spatial expression patterns of African rice orthologs of *O. sativa* panicle-related landmark genes. However an alteration of both timing and level of expression were observed between the two species, supporting the hypothesis that panicle branching diversity may be related to differential meristem fate acquisition during the developmental time-course of the panicle. These landmark genes were selected from the literature on *O. sativa* (see Table 1 in Annexes). In the context of my PhD, I selected 9 of them as the

key-acting genes in the two phases of panicle development: axillary meristem establishment and meristem fate transition (i.e. indeterminate to determinate). This manuscript is still in preparation and not yet submitted for publication.

In the second part of the Results, I present you the initial data set on the panicle phenotyping of the Vietnamese collection of *O. sativa* landraces. Some of the data we obtained were included in a manuscript by Phung et al. (2014) submitted to BMC Plant Biology journal. I was co-author of this manuscript, as participating to the collect of the first phenotypic data from this field assay. The manuscript reports on the first characterization of genotypic structure of this collection using DArT and GBS-derived markers, aiming that this collection would be suitable for GWAS. This manuscript is presented in Annexes.

2.2 Morphological and genome expression analyses in the African rice species

2.2.1 Main results described in the two manuscripts

- Using *O. sativa* as reference, we completed the description of the staging of inflorescence development in African rice species. The observation suggests that the overall panicle morphology at early stages is quite similar between the 3 species (*O. glaberrima, O. barthii* and *O. sativa*). However, it seems that spikelet meristem and floret differentiation occur later in *O. glaberrima* than in *O. barthii* (Fig. 1- Manuscript 1).
- Through small RNA transcriptomics, it was shown that 21-nt small RNAs mainly produced from un-annotated regions of the genomes exhibited a large fraction with higher expression in *O. barthii* than in *O. glaberrima.* This fraction of un-annotated 21-nt small RNAs derived mainly from phased siRNAs (phasiRNAs) generated from about a thousand of loci producing non-coding polyA tailed RNAs (Fig. 2 and Supplemental Fig. S2, S3 and S6 – Manuscript 1).
- These 21-nt phasiRNAs and their regulators *miR2118* and *MEL1* (a gamete-specific AGO protein encoding gene) not only express at different levels between *O. barthii* and *O. glaberrima,* but a shift in their timing of expression (i.e. heterochrony), was observed during panicle development, with a later expression in *O. glaberrima* than in *O. barthii* (Fig. 3 and Supplemental Fig. S7 – Manuscript 1).
- The associated long non-coding RNA precursors were detected in the differentiating spikelet meristems whereas *miR2118, MEL1* and 21-nt phasiRNAs were observed only at the early floret differentiation stage, in the outer cell layer of stamens (i.e. epidermis) and extended to pollen sac in later stages, consider as male gametogenesis in African rice (Fig. 3 - Manuscript 1).

- The expression patterns of orthologs of *O. sativa* landmark genes were highly conserved in *O. glaberrima* and *O. barthii* (Fig. 2 – Manuscript 2)
- The timing and level of landmark gene expressions are different between the two African species. The expression levels of the genes promoting the establishment of lateral meristems in domesticated rice are higher and the timing of the expression longer than in their wild relatives. (Fig. 1 – Manuscript 2)
- The coding sequences and promoter regions for these genes are highly conserved between the two of African rice species. Only few SNPs and INDELS were evidenced between the two species with some of them related to putative transcription factor binding sites (Figure 3 – Manuscript 2)

2.2.2 Manuscript 1: Time-shift of panicle meristem states in African rice species.

Ta KN, Sabot F et al.

Time-shift of panicle meristem states in African rice species

Ta KN[1,3]*, Sabot F[1]*, Adam H[1], Vigouroux Y[1], De Mita S[1], Ghesquière A[1], Do NV[3], Gantet P[2,3], Jouannic S[1,3].

1. IRD, UMR DIADE, 911, avenue Agropolis, BP64501, F-34394 Montpellier Cedex 5, France
2. UM2, UMR DIADE, 911, avenue Agropolis, BP64501, F-34394 Montpellier Cedex 5, France
3. IRD, AGI LMI RICE, University of Science and Technology of Hanoi, Agronomical Genetics Institute, Pham Van Dong road, Hanoi, Vietnam

*: These two authors contributed equally to this work

Current address:

S. De Mita INRA, Université de Lorraine, UMR 1136 Interactions Arbres/Microorganismes, F-54280 Champenoux, France

Corresponding author: Stefan Jouannic - E-mail: stephane.jouannic@ird.fr

Abstract

Rice species exhibit a wide range of panicle structure. The branching complexity of this inflorescence depends on the activity of lateral and branch apical meristems during its development. One hypothesis is that the timing of floral fate acquisition of these meristems modulates branching complexity. To explain these variations, much emphasis has been placed on changes in transcriptional regulation, but no large-scale study has reported yet changes in small RNA regulation related to this process. To evaluate this aspect, we performed deep sequencing of small RNA transcriptomes of two closely related species with striking panicle structure: *Oryza glaberrima* and its wild ancestor, *Oryza barthii*. Our study revealed a drastic change in a significant fraction of the 21-nucleotide small RNA population, corresponding mainly to *miR2118*-triggerred phased siRNAs (or phasiRNAs), with under-expression in *O. glaberrima.* These changes were affected by a heterochronic alteration of phasiRNA expression during panicle development, as well as their polyA-tailed mRNA precursors and their regulators, *miR2118* and the gamete-specific *Argonaute* gene *MEL1*, with delayed expression in *O. glaberrima*. Moreover, we show that *miR2118*-triggerred 21-nt phasiRNAs are specifically expressed in differentiating male gametes, even if their precursors were already expressed in spikelet meristems. Our study reveals a major reshaping of the regulation network from a specific class of small RNAs related to stamen development as well as meristem state-specific genes like the E-function MADS-box gene *LHS1*, related to differential panicle structure. The latter suggests that branching complexity of panicle might partly rely on differential timing of determinate state acquisition.

Introduction

Panicle structure acts as a major component of rice yield potential by determining the number of seeds produced, and was consequently one of the main morphological traits selected during rice domestication, with higher branching complexity (Sweeney and McCouch, 2007). Panicle architecture relies on the establishment and the activity of apical and axillary meristems deriving from the vegetative shoot apical meristem (SAM). After the SAM enters the reproductive phase, it transforms to a rachis meristem (RM) and primary branch meristems (PBMs) are generated from the periphery of the RM. Secondary and eventually tertiary branch meristems (SBMs, TBMs) are established in following steps on extending branches. These different branch levels generate the singled-flowered spikelets initiated from axillary and apical branch meristems (Ikeda et al., 2004; Itoh et al., 2005). Computer-based modeling suggests that the complexity level of panicle and more broadly the evolution of inflorescence architecture may be related to the timing of acquisition of floral fate (i.e. determinate fate) of apical and lateral reproductive meristems (Prusinkiewicz et al., 2007; Koes, 2008).

A genome-wide analysis of gene expression during Asian rice (*Oryza sativa*) panicle development has shown that only a small set of genes (357 out 22000) were expressed differentially in the early stages of panicle development, with a high proportion of transcription factors (Furutani et al., 2006). This study revealed that the initiation of the reproductive phase triggers the onset of regulatory networks leading to panicle development, and various transcription factors might play key roles in these networks. Mutant analyses and map-based cloning of QTLs have identified a number of genes required for the initiation and development of panicles, as well as genes controlling numbers and sizes of grains and panicles (Xing and Zhang, 2010; Wang and Li, 2011). Among these genes, some are involved in the patterning of axillary meristem and the branching of panicle, such as the nuclear regulatory factor-encoding genes *MONOCULM1*/*SMALL PANICLE* (*MOC1*/*SPA*), *LAX PANICLE1* (*LAX1*) and *LAX2* (Wang and Li, 2011). Analyses of the *aberrant panicle organization1* (*apo1*) and *apo2* mutants indicated that APO2 (the rice ortholog of *Arabidopsis* LEAFY) interacts with APO1 (the rice ortholog of *Arabidopsis* UFO) to suppress precocious conversion of branch meristem to spikelet meristem (Ikeda-Kawakatsu et al., 2012). Moreover, a large set of genes related to floral development has been identified in rice, notably MADS-box genes related to the ABC model which may also affect the panicle architecture (Yoshida and Nagato, 2011). Some of these panicle-associated genes have been identified as potential domestication (or crop improvement) genes in Asian rice, through QTL characterization (Xing and Zhang, 2010; Wang and Li, 2011; Ikeda et al., 2013) and detection of genomic regions under selection (He et al., 2011; Xu et al., 2011; Wang et al., 2012). Among these genes is the *miR156*- and *miR529*-targeted *OsSPL14* gene, identified through the characterization of *Wealthy Farmer's Panicle* (*WFP*) and *Ideal Plant Architecture* (*IPA*) QTLs (Jiao et al., 2010; Miura et al., 2010; Jeong et al., 2011).

A prevailing view in evolutionary developmental biology (Evo-Devo) is that morphological traits evolved mostly by changes in expression patterns of functionally conserved genes rather than through the emergence of new genes (Doebley and Lukens, 1998; Caroll, 2008). Although much emphasis has been placed on changes in transcriptional regulation, gene expression is regulated at many levels. In this sense, regulation of genome expression by small RNAs appeared to be an important mechanism in the control of plant development (morphogenesis and phase transition), through post-transcriptional regulation of mRNAs (*miRNAs* and *ta-siRNAs*) and silencing of gene expression (*siRNAs*) *via* site-specific DNA methylation (Jones-Rhoades et al., 2006; Arikit et al., 2013). Several small RNAs, required for initiation and development of panicles, have been identified in Asian rice *O. sativa* (Wang and Li, 2011; Ikeda et al., 2013). The microRNAs *miR156* and *miR172* have been reported to have a major impact on panicle architecture through the post-transcriptional regulation of their respective mRNA targets (Zhu et al., 2009; Jiao et al., 2010; Miura et al., 2010; Lee and An, 2012). Finally, a specific class of siRNAs corresponding to secondary siRNAs, namely phased siRNAs (or phasiRNAs), including trans-acting siRNAs (or ta-siRNAs) associated with *TAS3* loci, were shown to be associated with panicle development and more specifically with male gametogenesis (Johnson et al., 2009; Song et al., 2012a,b; Arikit et al., 2013; Komiya et al., 2014). Some phasiRNAs were shown to be panicle-specific 21- or 24-nucleotide (nt) small RNAs, produced from numerous polyA-tailed long non-coding RNA (lncRNA) generating loci through an RDR6-dependent pathway triggered by the microRNAs *miR2118* and *miR2275* respectively (Song et al., 2012a,b; Komiya et al., 2014), in conjunction with the gamete-specific Argonaute (AGO) protein, MEIOSIS ARRESTED AT LEPTOTENE1 (MEL1) for the 21-nt phasiRNAs (Komiya et al., 2014). However, their function during male gametogenesis is still unclear. In contrast, the *TAS3*-associated ta-siRNAs, triggered by the microRNA *miR390*, target mRNAs of *Auxin Response Factor* (*ARFs*) genes involved in various developmental processes (Allen et al., 2005; Liu et al., 2007; Nogueira et al., 2007; Song et al., 2012b).

In rice, two related crop species arose independently in Asia and Africa from two different wild species that had been isolated for over a million years. Asian rice, *O. sativa*, was domesticated as long as 10 000 years ago from *Oryza rufipogon*, while African rice, *Oryza glaberrima*, was domesticated about 3 000 years ago from *Oryza barthii* (Second, 1982; Linares, 2002; Caicedo et al., 2007; Vaughan et al., 2008; Huang et al., 2012; Orjuela et al. 2014; Wang et al. 2014). African rice domestication is linked with a single domestication origin in West Africa (Linares, 2002; Wang et al. 2014), associated with a severe genetic bottleneck (Li et al., 2011; Nabholz et al., 2014; Orjuela et al., 2014). It was suggested that several small RNA loci have experienced direct selection during Asian rice domestication, such as the microRNA loci *MIR164e*, *MIR390*, *MIR395a/b* and the ta-siRNA locus *TAS3a2* (Wang et al., 2010; Wang et al., 2012). Some of these microRNAs showed differential

expression between the two Asian species, suggesting that expression of miRNAs could also be a target of domestication (Wang et al., 2012). However, a large-scale study of changes in small RNA expression related to panicle structure diversity and evolution is still lacking.

The two African species exhibit striking panicle architecture, from a low branching complexity and few seeds in *O. barthii* to a more complex panicle and more seeds in *O. glaberrima* (Linares, 2002). In order to study the relationship that may exist between the variation of panicle architecture in African rice and small RNA expression, we conducted a comparative analysis of panicle morphology and genome expression during early events of panicle development, between the domesticated African species *O. glaberrima* and its wild-relative *O. barthii,* For this purpose, RNA-seq analysis of the small RNA population and various landmark genes related to reproductive meristem fate control were performed. Our comparative study showed, significant expression changes between the two species of the gamete-specific 21-nt phasiRNAs as well as their regulators *miR2118* and *MEL1*, and floral fate controlling genes, such as the E-function gene *LEAFY HULL STERILE1/OsMADS1*. The latter suggests that branching complexity of panicle might partly rely on differential rate of determinate state acquisition of the meristems.

Results

Structure and early development of African rice panicles

Rice panicle consists of a series of different order of branches: rachis (first axis), higher order of axis (i.e. primary branches or PBs, secondary branches or SBs and eventually tertiary branches or TBs). The single-flowered spikelets (SPs) are established on each panicle branch from apical and lateral meristems. A quantification of the main morphological traits of the African rice panicles was conducted using *P-TRAP* software (Al Tam et al., 2013) and *O. sativa* spp. *japonica* (cv Nipponbare) as a reference (Fig. 1, A-B). The panicle structure changed in domesticated species *O. glaberrima* from a panicle with few PBs and SBs (even without SB) bearing relatively few SPs (or grains) in the wild ancestor (*O. barthii*), to a highly branched structure (i.e. higher number of PBs and SBs) carrying larger numbers of SPs and seeds, as illustrated by the accessions CG14 and B88 respectively (Fig. 1, A-B). The *O. glaberrima* CG14 panicle size and complexity are even higher than *O. sativa* Nipponbare ones in our conditions. For the three species, the rachis meristem is aborted and a vestige remains at the distal end of the rachis (Fig. 1A). No tertiary branches were observed for these three accessions.

In order to understand which early morphological events may underlie the differences of panicle architecture at mature stage, a description of early panicle development in *O. glaberrima* (accession

CG14) and in its wild relative *O. barthii* (B88) was carried out, using the Asian species *O. sativa* spp. *japonica* (cv Nipponbare) as a reference (Fig. 1C). According to the description of panicle development in *O. sativa* (Ikeda et al., 2004), we further divided this period into 4 stages (i.e. stage 1 to 4) for the African species, with stage 0 corresponding to vegetative shoot apical meristem (SAM) shortly before phase transition (Fig 1Ca-c). Rachis meristem (RM) formation is considered as beginning of stage 1, with a RM taller and slightly wider than vegetative SAM. In turn, the rachis meristem produces primary branch meristems (PBMs) leading to the formation of PBs (Fig. 1Cd-f). At stage 2, PBs elongate and contribute to the higher order branches through the establishment of secondary branch meristems (SBMs) (Fig. 1Cg-l). At stage 3, the spikelets and floret meristems are differentiated from all branch and axillary meristems in the panicle (Fig. 1Cm-o). At stage 4, floret organs are initiated (data not shown).

The overall panicle morphology at early stages is quite similar between the 3 species (Fig. 1C). The RM is still present at the beginning of stage 2 (Fig. 1Cg-i) but not detectable at late stage 2 (Fig. 1Cj-l), indicating an abortion of the RM at this stage. However, abortion timing during this period cannot be determined. More PBs seem to be established in *O. glaberrima* than in the two other species at this stage (Fig. 1Cj-l). However, at stage 3, more SPs and florets seems to be differentiated in *O. barthii* in comparison to the other two species for a similar morphological stage (Fig. 1Cm-o). These observations would suggest that in *O. barthii* panicle, SPs and floret differentiation may occur earlier than in *O. glaberrima*. To better understand what mechanisms are involved at the molecular level, and obtain an insight into factors that may govern the diversity of panicle structure in African rice, we analyzed the expression of panicle-associated small RNAs and of various landmark genes related to reproductive meristem fate control.

Panicle-associated small RNAs in African rice species

To obtain a comprehensive survey of panicle derived-small RNAs in African rice and the qualitative and quantitative differences between the two species, we compared small RNA populations in panicles of *O. glaberrima* and *O. barthii*, using genome-wide small RNA-seq analysis. In order to focus on differences in expression resulting from inter-specific variations and to buffer genotypic variations, we used two RNA bulks extracted from developing panicles of 10 genotypes for each species (Supplemental Table S1). Over 33.1×10^6 and 33.9×10^6 high quality reads were generated from *O. glaberrima* and *O. barthii* libraries respectively (Supplemental Table S2). A total of 64% and 64.5% of small RNAs clusters (distinct small RNAs) ranging from 18 to 28 nucleotides from *O. barthii* and *O. glaberrima*, respectively, were mapped to the reference genome *O. sativa* ssp *japonica* cv Nipponbare MSU v7.0, and were similarly distributed over the *O. sativa* genome (Supplemental Figs. S1 and S2). As expected, the 21- and 24-nucleotide (nt) small RNAs were the predominant

populations of small RNAs in the two species, with a higher number of distinct 24-nt small RNAs (Supplemental Fig. S1).

The 21-nt small RNA population from the two species, corresponding to 23 798 distinct sequences, exhibited a surprisingly large fraction of small RNAs with a higher expression in *O. barthii* than in *O. glaberrima* (Fig. 2A). This subpopulation corresponds to 29% of the 21-nt small RNAs mapped, considering small RNAs that were at least five times more expressed in *O. barthii* than in *O. glaberrima* (threshold expression ratio ≥ 5.0). Other size classes (from 18- to 28-nt except 21-nt) were not distinguished by such a pattern of distribution between the two species (Supplemental Fig. S3). The overall 21-nt small RNAs were then categorized into five distinct classes corresponding to the different functional compartments of the rice genome: miRNAs (11.3%), ncRNAs (13.1%), repeats (12.4%), genes (CDS, intron, UTR: 25.5%) and unannotated regions (37.7%) of *O. sativa* reference genome. However, the distribution of the over-expressed fraction in the wild species showed a strong bias towards unannotated regions (81.4%; *p*-value=0.0) with the remaining classes all severely under-represented: miRNAs (2.5%; *p*-value=8.0×10^{-139}), ncRNAs (0.6%; *p*-value=3.5×10^{-301}), repeats (0.8%; *p*-value=2.3×10^{-261}) and genes (14.8%; *p*-value=1.2×10^{-83}) (fig. 3A). This indicated that the over-expressed 21-nt small RNAs detected in *O. barthii* panicles were mainly produced from regions corresponding to unannotated regions of *O. sativa* genome (Fig. 2A and Supplemental Fig. S3).

Previous studies on *O. sativa* also showed that the *miR2118*-triggered 21-nt phasiRNAs specifically expressed in the rice panicle originated from unannotated regions of the genome (Johnson et al., 2009; Song et al., 2012a; Komiya et al., 2014). To determine whether the over-represented 21-nt small RNA fraction in *O. barthii* corresponds to phasiRNAs, we used a dedicated program (see Materials and Methods section). We detected 4 100 distinct phasiRNAs from *O. barthii* and *O. glaberrima*, distributed across 892 loci (denoted "phased loci"), using *O. sativa* genome as reference (Supplemental Fig. S4, A-B). In addition, 952 distinct 21-nt small RNAs mapped to these 892 loci, but were not detected as phased siRNAs by our procedure. In total, 5 052 distinct 21-nt small RNA sequences associated with these 892 phased loci were found, corresponding to 21.2% of the panicle-derived 21-nt small RNA population from the two species. Out of these 5 052 small RNAs, 3 694 were detected in both species, 1 352 in *O. barthii* only and six in *O. glaberrima* only (see Supplemental Fig. S4 for details). As previously reported, the detected phased loci were present on the 12 chromosomes but not at the same density, and were organized in clusters or super-clusters, as on chromosome 12 (Supplemental Fig. S2). As expected, 86.5% of these 21-nt phasiRNAs mapped to unannotated regions of the *O. sativa* genome. This fraction of 21-nt phasiRNAs represents 49% of the unannotated 21-nt small RNA population (4 418 over 8 977 sequences). The 892 phased loci were scanned for the presence of the 22-nt miRNA *miR2118* recognition site in their vicinity (Johnson et al., 2009; Song et

al., 2012a; Komiya et al., 2014). About 60% of the detected phased loci shared the conserved motif in the vicinity of one end of the detected phased loci (from two to 452 bp from the end).

For 96% of the phased loci, there was a significant difference in read count between the two species, and about 71% of them with an abundance ratio ≥5.0 (see Supplemental Fig. S4 for details). These 21-nt siRNAs were mainly over-accumulated in *O. barthii* and contributed to 52% of the *O. barthii* five fold over-accumulated 21-nt small RNAs (Fig. 2A). The abundance of phasiRNAs from a single locus was not equivalent, with the predominance of one or two phasiRNAs (Fig. 2B and Supplemental Fig. S5). For the differentially expressed loci, this unequal distribution over the locus was conserved between the two species, with over-accumulation of all the detected phasiRNAs from a single locus (Figure 2B). Interestingly, the fraction of differentially expressed phased loci (with a higher expression in *O. barthii*) sharing the *miR2118* recognition site was significantly higher than in the not-differentially expressed phased loci between the two African species (see Supplemental Fig. S4A for details). Therefore, phased locus over-expression in *O. barthii* might result from differences in *miR2118* regulation between the cultivated species and its ancestor.

Comprehensive analysis of conserved miRNA families lead to the identification of 146 miRNA families expressed in African rice panicles. Interestingly, *miR159/319* families contributed 64% of the mature miRNA expression in young panicles of both African species (see Supplemental Fig. S4 for details), contrasting to previous studies of panicle-derived miRNAs in *O. sativa* (Jeong et al., 2011; Peng et al., 2011). This may be related to the specific developmental stages used in our study or may reflect differences between Asian and African rice species. Most miRNA families were expressed at similar levels in *O. glaberrima* and *O. barthii*. Differential expression affected only a few families with a trend toward higher expression in *O. barthii* (Fig. 2C). These families included the 22-nt *miR2118* and five other ones, namely *miR2275*, *miR5495*, *miR5497*, *miR5516* and *miR5519* (Fig. 2C). The 22-nt miRNA *miR2275* was previously reported to trigger 24-nt phasiRNAs (Johnson et al., 2009; Song et al., 2012a). However, despite over-expression of *miR2275* in *O. barthii,* no over-expression of any subset of the 24-nt small RNAs was detected (Supplemental Fig. S3). The miRNAs *miR5495* and *miR5497* were previously shown to be *O. sativa*-specific miRNAs expressed in pollen (Wei et al., 2011). Finally, our study provided evidence that the osa-*MIR5516* and osa-*MIR5519* loci were panicle-expressed 21-nt phasiRNA phased loci, namely phased_592, phased_594 and phased_327 respectively (Supplemental Fig. S4). In the same way, other annotated *MIR* loci were evidenced in our analysis of 21-nt phased loci as overlapping loci, suggesting that these loci are phasiRNA generating loci rather than true miRNA loci: *osa-MIR5486, osa-MIR5488*, *osa-MIR5506*, *osa-MIR5514*, *osa-MIR5517*, *osa-MIR5527*, *osa-MIR5530*, *osa-MIR5791*, *osa-MIR5796*, *osa-MIR5800* and *osa-MIR5822* (Supplemental Fig. S4).

Heterochronic variation in *miR2118* and 21-nt phasiRNAs during panicle development

The differential expression of these microRNAs and 21-nt phasiRNAs observed between *O. barthii* and *O. glaberrima* may reflect a difference in their expression level, but also heterochrony, i.e. a shift in the timing of their expression during panicle development. To test these hypotheses, expression analysis of *miR2118,* 21-nt phasiRNAs and their RNA precursors was performed on panicles collected from *O. glaberrima* (accession CG14) and *O. barthii* (accession B88) from morphological stage 1 to stage 4 of panicle development (i.e. unbranched rachis meristem to floret differentiation) (Figs. 1C and 3A). Four distinct phased loci from distinct clusters and chromosomes were considered for the RNA precursors as well as the phasiRNA with the highest level of accumulation for each locus (Fig. 3 and Supplemental Figs. S5 and S7).

While neither mature *miR2118* nor 21-nt phasiRNA *phasiPH12-1* were detected in early stage of panicle development (i.e. unbranched rachis meristem, stage 1), their expression was initiated later in *O. glaberrima* than in *O. barthii*, (i.e. early branching activity, stage 2)., The expression level of the miRNA and the phasiRNA reached a peak at stage 4 (i.e. floret differentiation) in *O. glaberrima* but was already reached at stage 3 (i.e. branching and spikelet differentiation) in *O. barthii* and remained at the same level in stage 4, leading to a higher expression level in *O. glaberrima* than in *O. barthii* at stage 4 (Fig. 3A). Similar patterns of expression were observed during panicle development for 21-nt phasiRNAs from the four phased loci (Supplemental Fig. S7). These findings indicate that the later expression of 21-nt phasiRNAs during panicle development in *O. glaberrima* is associated with a later expression of *miR2118*.

Similarly to Komiya et al. (2014), we have shown that the 21-nt phasiRNAs from the African species were generated from polyA-tailed long non-coding RNAs (Supplemental Fig. S6). The accumulation patterns of these polyA-tailed lncRNAs were similar to those of the phasiRNAs with later expression in *O. glaberrima* than in *O. barthii* (Fig. 3A and Supplemental Fig. S7). Furthermore, we investigated the accumulation levels of others factors involved in the biogenesis of the *miR2118*-triggered 21-nt phased RNAs, such as the OsDCL4 DICER-like protein, and the gamete-specific *Argonaute* protein MEL1 (Song et al., 2012a,b; Komiya et al., 2014). During early panicle development, the accumulation level of *OsDCL4* mRNAs slightly decreases over the 4 stages, while *MEL1* mRNA accumulation increases from stage 1 to stage 4 in a similar pattern to the small RNAs and lncRNAs (i.e. with later expression in *O. glaberrima* than *O. barthii*) (Fig. 3A and Supplemental Fig S7). Taken together, these data suggest that the accumulation level of phasiRNAs may depend on *miR2118,* phased loci-associated lncRNAs and *MEL1* accumulation levels rather than *OsDCL4* one.

To determine the spatial expression pattern of *miR2118, phasiPH12-1* and *PH12* lncRNA during African rice panicle development, *in situ* hybridization analysis was performed in young panicles of the two species (Fig. 3B). Similar patterns were observed for the three species with signal associated with mature *miR2118* observed only at the early floret differentiation stage (Fig. 3B). The signal was limited to differentiating stamens, with an expression pattern restricted to the outer cell layer of stamens (i.e. epidermis) and extended to pollen sac in later stages. Interestingly, both signals of *phasiPH12-1* and *PH12* lncRNA were observed in the same location in later stages (i.e. pollen sac). However, while *PH12* lncRNAs were detected from stage 3 (i.e spikelet meristem differentiation) to stage 4 (floret differentiation), *phasiPH12-1* was only detected in the pollen sac of florets posterior to *miR2118* detection. The temporal and spatial expression analysis of *miR2118*, *phasiPH12-1* and *PH12* lncRNA, as well as the higher expression level of pollen-specific *miR5495* and *miR5497* observed in the small RNA sequencing data, indicates that gamete-associated small RNAs are expressed earlier in *O. barthii* than in *O. glaberrima*. This suggests that stamen differentiation, and so floret differentiation, might occur later in *O. glaberrima* than *in O. barthii* during panicle development for a similar morphological differentiation.

Time-shift in landmark genes expression patterns in African rice

In order to support this finding, we analyzed the expression patterns of landmark genes related to meristem activity and meristem fate transition, relying on *O. sativa* panicle-related genes. Firstly, genes such as *Oryza sativa homeobox1 (OSH1), LAX PANICLE1 (LAX1), SQUAMOSA Promoter Binding Protein-Like14* (*OsSPL14*), controlling the initiation/maintenance of lateral meristems, were used as molecular markers for branching activity (Jiao et al., 2010; Miura et al., 2010; Tsuda et al., 2011; Woods et al., 2011). Secondly, genes related to meristem fate control were analyzed: while *ABERRANT PANICLE ORGANIZATION1* (*APO1*) and *APO2* gene products interact to suppress the transition from branch meristem to spikelet meristem (Rao et al., 2008), the E-function MADS-box *LHS1* gene promotes the transition to determinate meristems (SMs and FMs) (Cui et al., 2010; Khanday et al., 2013). In our conditions, qRT-PCR analysis at branching stage (stage 2) shows that the orthologs of branch-promoting genes *OSH1, LAX1* and *OsSPL14* are over-accumulated in *O. glaberrima* compared to *O. barthii* (Fig. 4A), suggesting a higher meristematic activity in the domesticated species in comparison to its wild-relative. Similarly, the mRNA accumulation of orthologs of the meristem-fate genes *APO1* and *APO2* was higher in *O. glaberrima* than in *O. barthii*, suggesting a delay of determination state in the meristem of *O. glaberrima*. Accordingly, we observed a lower accumulation of the *LHS1* mRNAs in *O. glaberrima* in comparison to *O. barthii* (Fig. 4A). This finding indicates that, for a similar morphological state, higher activity of the branch-promoting and spikelet-repressing genes occurs in *O. glaberrima* in comparison to *O. barthii* at the branching

stage, in contrast to the spikelet-promoting genes.

Using *in situ* hybridization, we investigated the expression pattern of *LHS1* ortholog in *O. glaberrima* and *O. barthii* panicles at late branching stage (i.e. at the transition between stage 2 and 3). In the two species, *LHS1* mRNAs were detected specifically in SMs (Fig. 4B). However, while the *LHS1* mRNA signal was observed in some lateral meristems in *O. glaberrima*, it was detected in all lateral meristems in *O. barthii* (Fig. 4B). This indicates that at a similar morphological stage all the terminal and lateral meristems had acquired the spikelet fate in *O. barthii*, whereas in *O. glaberrima* few had done so. This is coherent with the gamete-associated small RNAs profiling we observed.

Discussion

Conservation of miR2118-triggered 21-nt phasiRNAs in African rice

Using panicle-derived small RNA transcriptome sequencing in *O. glaberrima* and its wild ancestor, *O. barthii,* we show that 29% of the 21-nt small RNA population is drastically repressed (or non activated) in cultivated genotypes during the branching stage of panicle development. The large alteration of the 21-nt small RNA population in rice suggests change in their key regulator(s). This fraction of small RNAs corresponds mainly to 21-nt phasiRNAs generated from nearly a thousand non-coding RNA loci, associated with altered expression of the 22-nt microRNA *miR2118*. This is in agreement with what was observed in Asian rice *O. sativa* in which panicle-specific 21-nt phasiRNAs were produced from multiple non-coding RNA loci through an RDR6/DCL4-dependent pathway triggered by the microRNA *miR2118* (Johnson et al., 2009; Song et al., 2012a,b). The synthesis of secondary siRNAs (or phasiRNAs) in reproductive organs from non-coding loci has only been reported in rice, maize and *Brachypodium*, suggesting a recent origin of these secondary siRNAs from a common ancestor of grasses (Johnson et al., 2009; Vogel et al., 2010; Song et al., 2012a,b). However, the low conservation of these loci across these species suggests that species-specific functions of these 21-nt phasiRNAs may occur (Komiya et al., 2014). Interestingly, the *miR2118*-triggered secondary siRNA synthesis is also conserved across distantly related species, such as *Medicago truncatula*, tobacco and tomato. However, in these eudicot species, the *miR2118* family members (including *miR482* family members) were recruited to initiate secondary siRNAs on mRNAs encoding NBS-LRR proteins involved in disease resistance (Zhai et al., 2011; Li et al., 2012; Shivaprasad et al., 2012).

This study revealed conserved features of 21-nt phasiRNAs and associated loci between Asian and African rice species. However, this analysis was limited to the fraction of small RNAs African rice phasiRNAs conserved in *O. sativa* (ie. mapped to the *O. sativa* genome), and does not provide a

complete overview of the complexity of phased loci in African rice species and their evolution during the African domestication process. The genome sequence of *O. glaberrima* CG14 accession was recently published (Wang et al., 2014). However, using the same bio-informatic workflow, we got a better mapping results of our small RNA sequence data using the reference genome *O. sativa* ssp *japonica* cv Nipponbare MSU v7.0 than the *O. glaberrima* CG14 released sequence: 72.1% of small RNAs reads from both African species using Nipponbare MSU v7.0 sequence in contrast to 61.6 % and 56.2 % from *O. barthii* and *O. glaberrima* respectively using *O. glaberrima* reference genome AGl1.1. For this reason, we kept our small RNA sequence data analysis using the *O. sativa* Nipponbare reference genome. Nevertheless, it might be expected that higher diversification occurred between Asian and African rice since divergence about 1 million years ago than between *O. glaberrima* and *O. barthii* since divergence in Africa about 3000 years ago. Consequently, our quantitative analysis in African rice, which is limited to small RNAs that are conserved in the *O. sativa* reference, will be only marginally biased by evolutionary differences at the sequence level between the two African species. However, the number of overlapping phased loci between African rice species and *O. sativa* is quite low. About 1136 phased loci were identified from 4-cm long panicles of *O. sativa* ssp *indica* var 93-11. However, only 416 out of the 892 detected phased loci in African species overlapped phased loci detected in *O. sativa* spp *indica* (Song et al., 2012a). It may be because they were identified based on different stages of panicle development, rather than a quantitative effect between the two studies, as more small RNA sequences were characterized in our study. Moreover, in our comparative analysis, 48% of the over-expressed small RNAs in *O. barthii* mapping to *O. sativa* genome were not associated with the detected phased loci, and also originated mainly from unannotated regions of the genome. It is possible that these remaining over-expressed small RNAs belong to phased loci specific to African rice genomes that were not detectable in our analysis because of sequence divergence from the *O. sativa* reference, or that they are associated with non-coding loci unrelated to phased small RNA generating loci. In addition, 40% of the detected loci associated with differentially expressed 21-nt phasiRNAs between the two African species had no *miR2118* recognition site motif, suggesting a distinct mechanism of production. It cannot be ruled out that a cascade effect with *miR2118*-triggered phased loci occurred (as supported by shared sequences between phased loci). Another possible explanation is that *miR2118* recognition sites are present in the African rice genome at these loci but not in the *O. sativa* genome.

The miR2118-triggered 21-nt phasiRNAs as markers of male-gametogenesis in African rice

The function of the 21-nt phasiRNAs during panicle development is still unclear. MicroRNA *miR2118* has been reported to be preferentially expressed in rice and maize stamens, suggesting a major role in male gametogenesis (Song et al., 2012a). Recently, it was shown that the function of this class of

phased small RNAs was dependent on the germ line-specific Argonaute (AGO) protein MEL1 through a direct interaction between MEL1 and phasiRNAs (Komiya et al., 2014). Moreover, as mature *miR2118* was detected in the MEL1-binding small RNA fraction, it was argued that MEL1 may play a role in the first steps of the 21-nt phasiRNA biogenesis pathway, as being involved as the AGO protein in the *miR2118*-driven RISC triggering to the first cleavage of lncRNA precursors (Komiya et al., 2014). This is in agreement with the expression pattern observed for *MEL1* and *miR2118*. MEL1 mRNAs were first detectable by *in situ* hybridization in the hypodermis of developing stamen primordia in a spotty pattern, similarly to what we observed for *miR2118* microRNA, and became restricted to microsporangia and pollen sac (Nonomura et al., 2007). In our study, we have shown that while *PH12* precursor lncRNAs were detected in the spikelet meristem differentiation stage (stage 3), *phasiPH12-1* was only detected in the stamens of differentiating florets and co-localized with *PH12* precursor lncRNAs, but posterior to *miR2118* detection, which was first limited to the hypodermis of differentiating stamens and extended to pollen sac in later stages. Together, these data suggest that the 21-nt phasiRNA regulatory network is initiated early during panicle development from the spikelet meristem establishment, before *miR2118* and *MEL1* expression, through the induction of lncRNA precursor expression. The co-expression of *miR2118* and *MEL1* in hypodermis of differentiating stamen would initiate the 21-nt phasiRNA biogenesis pathway leading to the accumulation of the phasiRNAs. However, the *trans*-acting factors involved in the regulation of the lncRNA precursors are still unknown (Komiya et al., 2014).

Although direct evidence for the involvement of *miR2118* and the associated phasiRNAs and long ncRNAs in male gametogenesis is still lacking, their involvement is supported by the phenotypes of *rdr6* and *dcl4* mutants affected in flower and stamen development (Liu et al., 2007; Song et al., 2012b). Moreover the *mel1* mutant phenotype indicates that the MEL1 AGO protein mediates the regulation of germ-line mother cell development and meiosis in both male and female organs, suggesting that the MEL1-phasiRNAs complexes may be involved in these processes (Nonomura et al., 2007; Komiya et al., 2014). The question that remains is whether these phasiRNAs can be considered as *ta*-siRNA (secondary siRNAs targeting other mRNAs) or as products of silencing of long non-coding RNAs of still unknown function. The co-expression levels of the phasiRNAs and their precursors suggest that these phasiRNAs act as trans-acting siRNAs. However, among the 1565 predicted putative targets (targeted by 915 phasiRNAs from 544 distinct phased loci), none were significantly differentially expressed from the same panicle stages of the two African species, using mRNA-seq data from Nabholz and colleagues (2014) (data not shown), similarly to what was observed for analysis of *O. sativa indica* panicle RNA degradome data (Song et al., 2012a). It still cannot be excluded that the phasiRNAs do not trigger mRNA cleavage but act through inhibition of translation or DNA methylation, as suggested by Song et al. (2012a). However, the main accumulation

of MEL1 in the cytoplasm would favor for *trans*-acting function of MEL1-phasiRNA complex on mRNAs (Komoiya et al, 2014).

Panicle structure complexity variation in African rice

Theoretical modeling of inflorescence architecture, based on variations of a parameter called *vegetativeness* (*veg*), suggests that panicle branching complexity and its evolution, as well as for other inflorescence types (racemes and cymes), depends on difference of timing of floral fate acquisition of apical and/or lateral meristems (Prusinkiewicz et al., 2007). This is related to the fact that as soon as a meristem acquires the floral fate (determinate state), it is no longer able to establish new lateral meristems contributing to the elaboration of the branch structure. This biological basis is supported by the analysis of various mutants affected in the floral meristem identity in different species, notably in *Arabidopsis thaliana*, *Antirrhinum majus*, petunia and tomato (Koes, 2008; Moyroud et al., 2010; Park et al., 2014). Moreover, a comparative study of the diversity of inflorescence architecture in tomato based on meristem-specific transcriptome analysis also confirmed this model (Park et al., 2012). This study revealed that the gene-regulatory network related to inflorescence branching is initiated early during meristem maturation and that evolutionary diversity in inflorescence architecture in Solanaceae is modulated by heterochronic shifts in the acquisition of floral fate and that abrupt morphogenetic changes may be related to progressive molecular changes associated with maturation of the tomato SAM (Park et al., 2012). In the case of grass inflorescence and more specifically rice panicle, it was also reported that the variation of panicle architecture depends on the activity or the expression levels of meristem fate controlling genes but also branch promoting genes (Kyozuka et al., 2014; Zhang and Yuan, 2014). Among these genes, *APO1* and *APO2* genes (the orthologs of Arabidopsis *UFO* and *LFY* genes respectively) are reported as branch meristem-promoting genes, through the regulation of cell proliferation at least in apical inflorescence meristem (Ikeda-Kawakatsu et al., 2009, Ikeda-Kawakatsu et al., 2012), contrasting with what is observed in Eudicot species in which the *UFO*- and *LFY*-like genes act as flower-promoting genes. Another branch meristem-related gene was identified through the characterization of a rice QTL, *Ideal Plant Architecture1* (*IPA1*)/*Wealthy Farmer's Panicle* (*WFP*). *IPA1*/*WFP* encodes a SQUAMOSA promoter binding protein–like (SBP) box protein named OsSPL14. This gene is post-transcriptionally regulated by two closely related microRNAs, *miR156* and *miR529*, the last one acting as the main regulator of *OsSPL14* during reproductive phase (Jiao et al., 2010; Miura et al., 2010; Jeong et al., 2011). A point mutation in the common region of *miR156*- and *miR529*-targeted sites within *OsSPL14* causes increased accumulation of *OsSPL14*, resulting in rice plants with fewer tillers, and larger inflorescences with more branches and spikelets (Jiao et al., 2010; Miura et al., 2010). The expression of these protein-coding genes is initiated at a very early stage of panicle development, even before the transition from vegetative stage, with expression in all

branch meristems (Ikeda-Kawakatsu et al., 2009; Ikeda-Kawakatsu et al., 2012; Jiao et al., 2010; Miura et al., 2010).

In our comparative expression analysis of African rice panicle development, all the genes related to branching activity were over-accumulated at the branching stage in the domesticated species. This result is in agreement with the higher branch phenotype observed in *O. glaberrima* compared to *O. barthii*. At the same time, histological analysis of early stages of panicle development in the two African rice species suggests that spikelet meristem differentiation occurs later in *O. glaberrima* than in *O. barthii*. This is supported by later initiation of expression of the 21-nt phasiRNA pathway members (i.e. *miR2118*, *MEL1*, *lncRNAs* and *phasiRNAs*), especially the *lncRNA* precursors, which are expressed in spikelet meristems, as well as the spikelet-associated MADS-box gene *LHS1/OsMADS1*. This gene contributes to meristem termination by specifying a single floret on the spikelet meristem and floret organ development, through the regulation of transcription factors and the balance between auxin and cytokinin signaling pathways (Khanday et al., 2013). Its expression is clearly initiated at the spikelet differentiation state before floret establishment (our study; Gao et al., 2010; Kobayashi et al., 2010). Therefore, our findings indicate that not only spikelet fate acquisition is different between the two species, but also branch meristem activity through differential expression of related genes. In fact, these two developmental processes are connected, as higher branching activity delays spikelet specification, as observed with the *apo1* and *apo2* mutants (see above) but also with the *tawawa* (*taw*) gain-of-function mutant characterized by enhanced branch meristem activity and delayed spikelet meristem specification, resulting in a prolonged branch formation stage and an inflorescence with more branches (Yoshida et al., 2013). *TAW* gene is a member of the small gene family ALOG and acts partly through the regulation of flowering time-related *SVP*-like genes to control the rice panicle architecture (Yoshida et al., 2013). A parsimonious hypothesis would be that the expression of very early acting factors of panicle development might be differentially affected between the two African rice species to lead to the final panicle structure differences observed between the two species.

Conclusion

Here, we provide evidence that the male-gametogenesis–specific 21-nt phasiRNA pathway triggered by *miR2118* is conserved in both wild-relative and domesticated African rice species. This pathway is in fact initiated when spikelet meristems are established, through the initiation of lncRNA precursor expression, and the gamete-specificity of the 21-nt phasiRNAs may be driven by *MEL1* and *miR2118* stamen-specific expression. Our study also provides evidence that the differential expression of some spikelet-related genes (both protein-coding and non-coding genes) as well as branch-promoting genes affecting the panicle structure may have been impacted by the domestication of African rice *O.*

glaberrima. However, the functional link between the major regulatory changes of branch and spikelet-related gene expression and the panicle phenotypic variations associated with domestication still remains to be determined. It will be of great interest to determine whether a similar situation is observed in the Asian rice species (*O. sativa* vs. *O. rufipogon*), supporting the phenotypic convergence observed between the two domestication processes (higher branching complexity in domesticated vs. wild species; Sweeney and McCouch, 2007).

Materials and Methods

Plant materials, sampling and panicle phenotyping

The set of rice (Oryza *sativa*, *glaberrima* and *barthii*) used for the mature panicle phenotyping were grown in summer 2011 in fields of CIAT (Centro Internacional para la Agricultura Tropical, Cali, Colombia) (n= 3 plants per variety, 2 replicates). Three panicles per plant were harvested at mature stage for each replicate (n=18 panicles per variety). Genotypes include Nipponbare, CG14 and B88 for *O. sativa*, O. *glaberrima* and O. *barthii* respectively. *P-TRAP* software has been used to quantify the panicle traits (Al Tam et al., 2013).

For histological analysis and expression analysis, Nipponbare, CG14 and B88 plants were grown in growth chamber at IRD, Montpellier (France). Chamber settings were as follows: on 14-10h day/night cycle at 32°C/28°C and humidity at 60%. Flowering was induced by short day conditions (10-h day/night cycle). 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.

For Illumina sequencing, 10 accessions of *O. glaberrima* and 10 accessions of *O. barthii* (see supplementary table S1) were grown in the greenhouse at IRD, Montpellier. Around 15 panicles from each accession were collected from 4 to 15 days after induction, corresponding to stage 1 (rachis and primary branch meristem) to stage 3 (spikelet differentiation) of panicle development. Total RNAs (mRNAs and small RNAs) were extracted using an RNeasy Plant Mini Kit with RLT and RWT buffers (Qiagen, France). DNAs treatments were performed using the RNAeasy-free DNase set (Qiagen). Two bulks of total RNAs corresponding to a mix of total RNAs from the 10 accessions of the two species were used for sequencing. For expression analysis, the RNA samples were used individually on another set of sampling.

Histology studies

Samples were fixed overnight at 4°C in fixation buffer (4% paraformaldehyde, 0.1 M phosphate buffer, pH7). Samples were dehydrated through a graded ethanol series (30, 50, 70, 90, 100%, v/v) for 2 hours and stored at 4°C. and embedded in Technovit resin (Heraeus Kulzer, Germany). Blocks were sectioned at 4-5µm thickness using a HM650 microtome (Thermo Scientific Microm, Walldorf, Germany). Slides were double-stained with PAS stain (periodic acid–Schiff reagent) for the detection of carbohydrate compounds and naphthol blueblack (NBB) for the detection of proteins. Slides were observed with a Leica DMRB microscope and photographed by Evolution MP5.0 color Media Cybernetics camera.

Illumina sequencing and data processing

Purified small RNA sequencing was performed by Eurofins/MWG Operon (Germany) on an Illumina *Hi-seq 2000* using the TrueSeq™ SBS v5 sequencing kit. The raw data (accession number GSE48346 in NCBI Gene Expression Omnibus) were trimmed by removing adapter sequences and low quality sequences using *CutAdapt* (Martin 2011). All the trimmed reads ranging from 18 to 28 nucleotides were clustered and mapped to *O. sativa ssp japonica* cv Nipponbare genome (MSU release version 7; http://rice.plantbiology.msu.edu/) using *BLAST* (Altschul et al., 1990). The 18-28 nucleotide reads were annotated using successive hierarchical *BLAST* versus (in order) miRBase v17.0 (Kozomara et al., 2011), Rfam v7, home-made repeat database (successive curated concatenation of *RetrOryza*, RepBase, TREP and TIGRRepeats), CDS then gene features from *Oryza sativa ssp japonica var* Nipponbare MSU v7.0 annotation, and finally the MSU v7.0 rice genome. The *BLAST* and post-filters parameters used were probability of 85%, e-value of 10^{-3}, on a size of 85% of the reads (minimum size of 16). The same *BLAST* parameters were used throughout the analysis. Mapping from *O. glaberrima* and *O. barthii* were then compared and filtered using a series of homemade *Perl* scripts (available on demand). The 21-mers were used in phasing analysis with the *ta-si Prediction* tool from the *UEA sRNA workbench* facilities (http://srna-workbench.cmp.uea.ac.uk/; Stocks et al. 2012). Once the loci were identified, we used the *EMBOSS* software suite v6.5.7.0 (Rice et al., 2000) to extract -500/+500 bases around each locus, and treated them using *MEME* v4.8.1 (Bailey and Elkan, 1994).

Statistical tests of all the processed data were performed using *g-test* and a fixed *p*-value of 10^{-3}. Depending on the experiment, the degree of freedom was adjusted but was generally 1. All the calculations were performed using homemade *Perl* scripts and CPAN statistical modules.

Expression analysis

Quantitative RT-PCRs: First-stand cDNA synthesis was done using SuperScript III cDNA First-strand synthesis system (Invitrogen). Quantitative stem-loop RT-PCR analyses on small RNAs were performed using 100 ng of total RNA according to Varkonyi-Gasic et al. (2007) with RTs run at 42 °C in conjunction with small RNA-specific primers and PCRs at 60 °C for annealing. Quantitative RT-PCR analyses on mRNAs were performed using 1 µg of total RNA in conjunction with polydT or random hexamer primers according to the manufacturer's instructions.

Quantitative RT-PCRs were performed using LightCycler 480 thermocycler (Roche, France) in conjunction with SYBR Green I master mix (Roche, France) in 8 µL reaction mix containing 2 µL of diluted RTs and 0.8 µL of forward and reverse primers at 10 µM. The Q-PCR amplification conditions include 3 stages: pre-incubator (95°C in 10 minutes); amplification with 45 cyclers (95°C 15s and 60°C 30s); melting curve (95°C 5s and 70°C 1 min). In stem-loop Q-RT-PCR, the levels of miRNA were normalized by mature *miR159* expression level. In classic Q-RT-PCR, mRNAs were normalized

using the rice *Actin* genes (LOC_Os03g50885.1) transcripts expression level. Each set of experiments was repeated three times, and the relative quantification method with efficiency corrected calculation model (Souaze et al., 1996) was used to evaluate quantitative variations. The primers used are listed in Supplemental Table S3. Statistical significance of the quantitative values differences between *O. glaberrima* and *O. barthii* for each stage was calculated using two-tails t-test.

In situ hybridization: To obtain DNA templates for the RNA probe synthesis, PCR amplifications were performed with gene-specific antisense primers tailed with a T7 RNA polymerase binding site (see supplementary Table S3 for primer sequences). The resulting DNA fragments were used directly as templates for synthesizing antisense ribo-probes incorporating UTP–digoxigenin (Roche) as the label in conjunction with a T7 Maxi Script kit (Ambion). For *miR2118*, *PH12* precursor and *phasiPH12-1* detection, 0.02 μM of a 5' digoxigenin–labeled LNA probe complementary to the target (see Supplemental Table S3 for primer sequences) was used. *In situ* hybridization experiments were carried out as described by Adam et al., (2007). Detection was performed using the Vector Blue Alkaline Phosphatase Substrate Kit III (Vector Laboratories). Slides were observed and photographed by Evolution MP5.0 color Media Cybernetics camera in conjunction with a Leica DMRB microscope. These images processed using Photoshop CS6.

Acknowledgments

We thank members of the lab for their help with plant care and RNA extractions. We thank members of IRD-Montpellier bioinformatics platform for their technical support. We also thank O. Leblanc, C. Brugidou and T. Harrop for critical reading of the manuscript.

Financial sources

This project was jointly supported by Agropolis Fondation (through the « Investissements d'avenir » programme (ANR-10-LABX-0001-01)) and Fondazione Cariplo under the reference ID EVOREPRICE 1201-004. This work was also supported by the Institut de Recherche pour le Développement (IRD) institutional funding, Agropolis foundation (ARCAD project for post-doctoral fellowship of SDM), Vietnamese government (program 322 to KNT) and the Global Rice Science Partnership (GRiSP) scholarship program (KNT).

Literature cited

Adam H, Marguerettaz M, Qadri R, Adroher B, Richaud F, Collin M, Thuillet AC, Vigouroux Y, Laufs P, Tregear JW, et al (2011) Divergent Expression Patterns of *miR164* and *CUP-SHAPED COTYLEDON* Genes in Palms and Other Monocots: Implication for the Evolution of Meristem Function in Angiosperms. Mol Biol Evol **28**: 1439–1454

Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-Directed Phasing during Trans-Acting siRNA Biogenesis in Plants. Cell **121**: 207–221

Al-Tam FM, Adam H, Anjos AD, Lorieux M, Larmande P, Ghesquière A, Jouannic S, Shahbazkia HR (2013) P-TRAP: a Panicle Trait Phenotyping Tool. BMC Plant Biol **13**: 122

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal Mol Biol **215**: 403–410

Arikit S, Zhai J, Meyers BC (2013) Biogenesis and function of rice small RNAs from non-coding RNA precursors. Curr Opin Plant Biol **16**: 170–179

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In Altman R, Brutlag D, Karp P, Lathrop R, Searls D, eds, Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, AAAI Press, Menlo Park, California, pp 28-36

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, et al (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS Genet **3**: 1745–1756

Carroll S (2008) Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell **134**: 25–36

Cui R, Han J, Zhao S, Su K, Wu F, Du X, Xu Q, Chong K, Theissen G, Meng Z (2010) Functional conservation and diversification of class E floral homeotic genes in rice (*Oryza sativa*). Plant J **61**: 767–781

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell **127**: 1309–1321

Doebley J, Lukens L (1998) Transcriptional regulators and the evolution of plant form. Plant Cell **10**: 1075–1082

Furutani I, Sukegawa S, Kyozuka J (2006) Genome-wide analysis of spatial and temporal gene expression in rice panicle development. Plant J **46**: 503–511

Gao X, Liang W, Yin C, Ji S, Wang H, Su X, Guo C, Kong H, Xue H, Zhang D (2010) The SEPALLATA-like gene *OsMADS34* is required for rice inflorescence and spikelet development. Plant Physiol **153**: 728-740

He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X, Greenberg AJ, Hudson RR, Wu C-I, Shi S (2011) Two evolutionary histories in the genome of rice: the roles of domestication genes. PLoS Genet **7**: e1002100

Huang X, Kurata N, Wei X, Wang Z-X, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, et al (2012) A map of rice genome variation reveals the origin of cultivated rice. Nature **490**: 497–501

Ikeda K, Sunohara H, Nagato Y (2004) Developmental course of inflorescence and spikelet in rice. Breeding sci **54**: 147–156

Ikeda M, Miura K, Aya K, Kitano H, Matsuoka M (2013) Genes offering the potential for designing yield-related traits in rice. Curr Opin Plant Biol **16**: 213–220

Ikeda-Kawakatsu K, Yasuno N, Oikawa T, Iida S, Nagato Y, Maekawa M, Kyozuka J (2009) Expression level of *ABERRANT PANICLE ORGANIZATION1* determines rice inflorescence form through control of cell proliferation in the meristem. Plant Physiol **150**: 736–747

Ikeda-Kawakatsu K, Maekawa M, Izawa T, Itoh J-I, 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

Itoh J-I, Sato Y, Nagato Y, Matsuoka M (2006) Formation, maintenance and function of the shoot apical meristem in rice. Plant Mol Biol **60**: 827–842

Jeong D-H, Park S, Zhai J, Gurazada SGR, De Paoli E, Meyers BC, Green PJ (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, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, et al (2010) Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. Nat Genet **42**: 541–544

Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan G-L, Walbot V, Sundaresan V, Vance V, Bowman LH (2009) Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. Genome Res **19**: 1429–1440

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. Ann Rev Plant Biol **57**: 19–53

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 Physiol **161**: 1970–1983

Kobayashi K, Maekawa M, Miyao A, Hirochika H, Kyozuka J (2010) *PANICLE PHYTOMER2* (*PAP2*), encoding a SEPALLATA subfamily MADS-box protein, positively controls spikelet meristem identity in rice. Plant Cell Physiol **51**: 47-57

Komiya R, Ohyanagi H, Niihama M, Watanabe T, Nakano M, Kurata N, Nonomura K-I (2014) Rice germline-specific Argonaute MEL1 protein binds to phasiRNAs generated from more than 700 lincRNAs. Plant J **78**: 385–397

Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. Nucl Acids Res **39**: D152–D157

Lee D-Y, 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

Li Z-M, Zheng X-M, 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

Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B (2012) MicroRNA regulation of plant innate immune receptors. Proc Natl Acad Sci USA **109**: 1790–1795

Linares OF (2002) African rice (*Oryza glaberrima*): history and future potential. Proc Natl Acad Sci USA **99**: 16360–16365

Liu B, Chen Z, Song X, Liu C, Cui X, Zhao X, Fang J, Xu W, Zhang H, Wang X, et al (2007) *Oryza sativa dicer-like4* reveals a key role for small interfering RNA silencing in plant development. Plant Cell **19**: 2705–2718

Liu Y, Wang Y, Zhu QH, Fan L (2013) Identification of phasiRNAs in wild rice (*Oryza rufipogon*). Plant Sign Behav **8**: 25079

Martin M (2011) *Cutadapt* removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **18**: 10-12

Mattick JS (2007) A new paradigm for developmental biology. J Exp Biol **210**: 1526-1547

Miura K, Ikeda M, Matsubara A, Song X-J, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M (2010) *OsSPL14* promotes panicle branching and higher grain productivity in rice. Nature Genet **42**: 545–549

Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F (2010) LEAFY blossoms. Trends Plant Sci **15**: 346–352

Nabholz B, Sarah G, Sabot F, Ruiz M, Adam H, Nidelet S, Ghesquière A, Santoni S, David J, Glémin S (2014) Transcriptome population genomics reveals severe bottleneck and domestication cost in the African rice (*Oryza glaberrima*). Mol Ecol **23**: 2210-2227

Nogueira FTS, Madi S, Chitwood DH, Juarez MT, Timmermans MCP (2007) Two small regulatory RNAs establish opposing fates of a developmental axis. Genes Dev **21**: 750–755

Nonomura K, Morohoshi A, Nakano M, Eiguchi M, Miyao A, Hirochika H, Kurata N (2007) A germ cell specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. Plant Cell **19**: 2583-2594

Orjuela J, Sabot F, Chéron S, Vigouroux Y, Adam H, Chrestin H, Sanni K, Lorieux M, Ghesquière A (2014) An extensive analysis of the African rices genetic diversity through a global genotyping. Theor Appl Genet DOI: 10.1007/s00122-014-2374-z

Park SJ, Jiang K, Schatz MC, Lippman ZB (2012) Rate of meristem maturation determines inflorescence architecture in tomato. Proc Natl Acad Sci USA **109**: 639–644

Park SJ, Eshed Y, Lippman ZB (2014) Meristem maturation and inflorescence architecture-lessons from the Solanaceae. Curr Opin Plant Biol **17**:70-77

Peng T, Lv Q, Zhang J, Li J, Du Y, Zhao Q (2011) Differential expression of the microRNAs in superior and inferior spikelets in rice (*Oryza sativa*). J Exp Bot **62**: 4943–4954

Poethig RS (2009) Small RNAs and developmental timing in plants. Curr Opin Genet Dev **19**: 374–378

Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E (2007) Evolution and development of inflorescence architectures. Science **316**: 1452–1456

Rao NN, Prasad K, Kumar PR, Vijayraghavan U (2008) Distinct regulatory role for RFL, the rice LFY homolog, in determining flowering time and plant architecture. Proc Natl Acad Sci USA **105**: 3646–3651

Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet **16**: 276-277

Second G (1982) Origin of the genic diversity of cultivated rice (Oryza spp.): study of the polymorphism scored at 40 isozyme loci. Jap J Genet **57**: 25–57

Shivaprasad PV, Chen H-M, Patel K, Bond DM, Santos BACM, Baulcombe DC (2012) A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. Plant Cell **24**: 859–874

Song X, Li P, Zhai J, Zhou M, Ma L, Liu B, Jeong D-H, Nakano M, Cao S, Liu C, et al (2012) Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. Plant J **69**: 462–474

Song X, Wang D, Ma L, Chen Z, Li P, Cui X, Liu C, Cao S, Chu C, Tao Y, et al (2012) Rice RNA-dependent RNA polymerase 6 acts in small RNA biogenesis and spikelet development. Plant J **71**: 378–389

Souaze F, Ntodou-Thome A, Tran CY, Rostene W, Forgez P (1996) Quantitative RT-PCR: limits and accuracy. Biotechniques **21**: 280-285

Stocks MB, Moxon S, Mapleson D, Woolfenden HC, Mohorianu I, Folkes L, Schwach F, Dalmay T, Moulton V (2012) The UEA sRNA workbench: a suite of tools for analyzing and visualizing next generation sequencing microRNA and small RNA datasets. Bioinformatics **28**: 2059-2061

Sweeney M, McCouch S (2007) The complex history of the domestication of rice. Ann Bot **100**: 951–957

Tsuda K, Ito Y, Sato Y, Kurata N (2011) Positive autoregulation of a *KNOX* gene is essential for shoot apical meristem maintenance in rice. Plant Cell **23**: 4368-4381

Vaughan DA, Lu BR, Tomooka N (2008) The evolving story of rice evolution. Plant Sci **174**: 394-408

Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. Plant Meth **3**: 12

Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon-Smith M, Lail K, et al (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature **463**: 763–768

Wang Y, Shen D, Bo S, Chen H, Zheng J, Zhu Q-H, Cai D, Helliwell C, Fan L (2010) Sequence variation and selection of small RNAs in domesticated rice. BMC Evol Biol **10**: 119

Wang Y, Li J (2011) Branching in rice. Curr Opin Plant Biol **14**: 94–99

Wang Y, Bai X, Yan C, Gui Y, Wei X, Zhu Q-H, Guo L, Fan L (2012) Genomic dissection of small RNAs in wild rice (*Oryza rufipogon*): lessons for rice domestication. New Phytol **196**: 914–925

Wang M, Yu Y, Haberer G, Marri PR, Fan C, Goicoechea JL, Zuccolo A, Song X, Kudrna D, Ammiraju JS et al. (2014) The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. Nat Genet. doi: 10.1038/ng.3044

Wei L, Yan L, Wang T (2011) Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of *Oryza sativa*. Genome Biol **12**: R53

Woods DP, Hope CL, Malcomber ST (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

Xing Y, Zhang Q (2010) Genetic and molecular bases of rice yield. Ann Rev Plant Biol **61**: 421–442

Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L, et al (2012) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. Nat Biotechnol **30**: 105–111

Yoshida H, Nagato Y (2011) Flower development in rice. J Exp Bot **62**: 4719–4730

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, Chen R, Yamazaki R, Tokunaga H, Kitaguchi Y, Sato Y, et al (2013) TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. Proc Natl Acad Sci USA **110**: 767–772

Zhang D, Yuan Z (2014) Molecular control of grass inflorescence development. Annu Rev Plant Biol **65**:553-578

Zhai J, Jeong D-H, De Paoli E, Park S, Rosen BD, Li Y, González AJ, Yan Z, Kitto SL, Grusak MA, et al (2011) MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. Genes Dev **25**: 2540–2553

Zhao K, Wright M, Kimball J, Eizenga G, McClung A, Kovach M, Tyagi W, Ali ML, Tung C-W, Reynolds A, et al (2010) Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. PLoS ONE **5**: e10780

Zhu Q-H, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C (2008) A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome Res **18**: 1456–1465

Zhu Q-H, Upadhyaya NM, Gubler F, Helliwell CA (2009) Over-expression of *miR172* causes loss of spikelet determinacy and floral organ abnormalities in rice (*Oryza sativa*). BMC Plant Biol **9**: 149

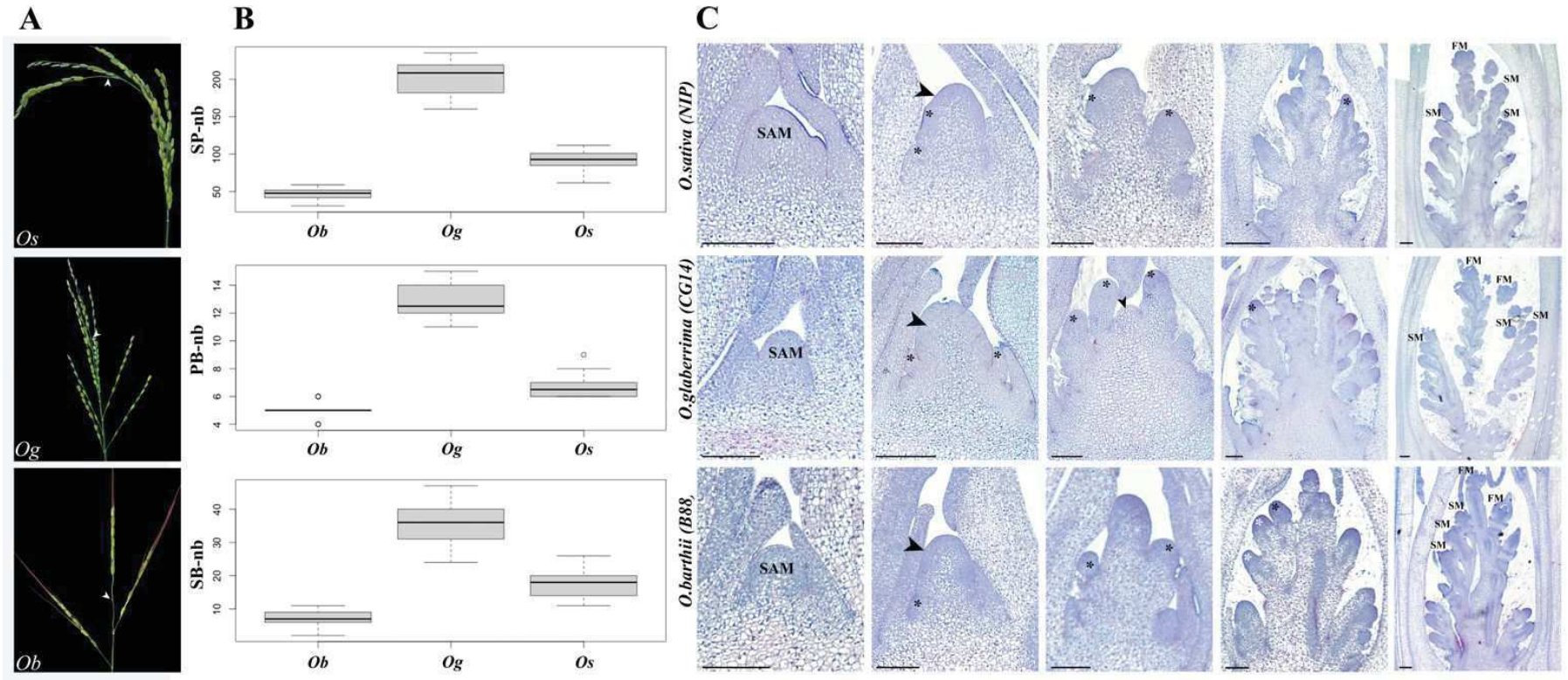


Figure 1. Panicle development in Asian and African rice species. A, Mature panicle of *O. barthii var.* B88, *O. glaberrima* cv CG14 and *O. sativa* cv Nipponbare; Ob, *O. barthii*; Og, *O. glaberrima*; Os, *O. sativa*; white arrowhead: vestige of aborted rachis meristem. B, Comparison of panicle traits at mature stage between *O. barthii var.* B88, *O. glaberrima* cv CG14 and *O. sativa* cv Nipponbare. Pb_nb, primary branches number; Sb_nb, secondary branches number; Sp_nb, spikelet number; Ob, *O. barthii*; Og, *O. glaberrima*; Os, *O. sativa*. These morphological traits were quantified using P-TRAP software (Al Tam et al., 2013); n= 18 panicles for each accession. C, Histological organization of panicle architecture at early developmental stages of *O. barthii var.* B88, *O. glaberrima* cv CG14 and *O. sativa* cv Nipponbare, respectively; (a-c) Stage 0: mature vegetative shoot apices; (d-f) Stage 1: elongation of rachis meristem (arrowheads) and formation of primary branch meristems (PBM) (*); (g-l) Stage 2: rachis meristem (arrowheads) with elongating primary branches (PBs). At the end of this stage (i.e. j-l), secondary branches (SBs) (white *) are initiated from PBs (*); (m-o) Stage 3: spikelet meristem (SM) and floret meristem (FM) differentiation. Scale bar. 100µm.

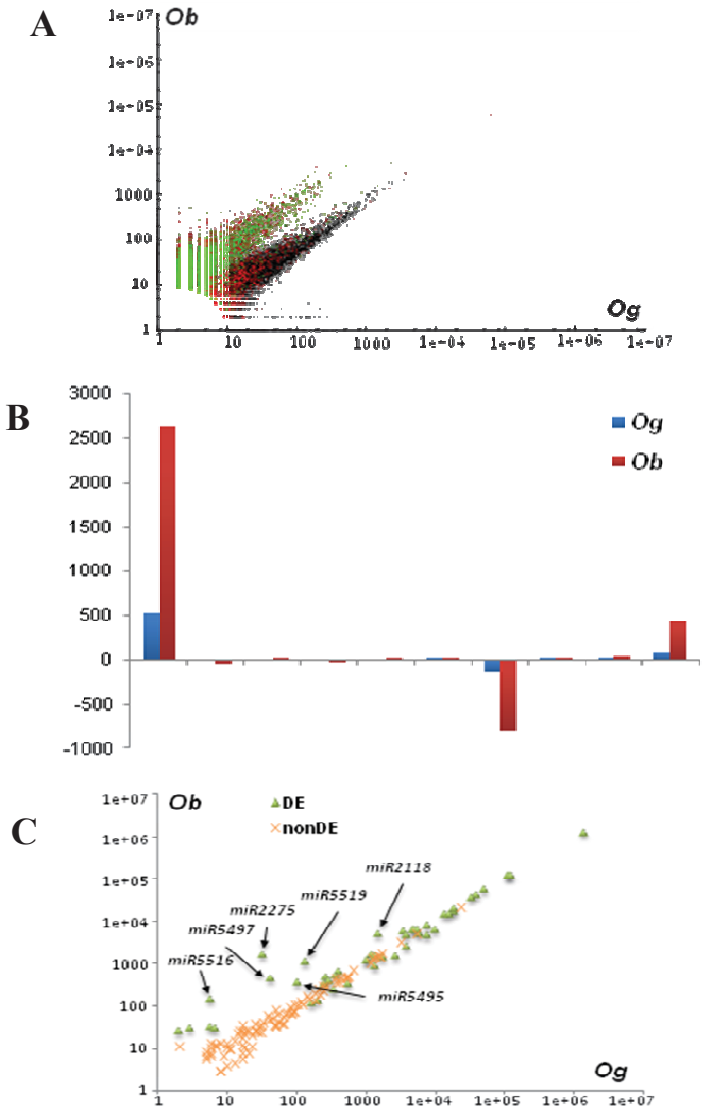


Figure 2. Relative abundance of panicle-derived 21-nt small RNAs from *O. barthii* and *O. glaberrima*. A, Relative abundance of 21-nt small RNAs between *O. barthii* (*Ob*) and *O. glaberrima* (*Og*). LogPlot of normalized abundance of distinct small RNA sequences. Black dots represent global 21-nt small RNAs, red dots unannotated 21-nt small RNAs, and green dots 21-nt small RNAs associated with detected phased loci. **B,** Relative abundance of detected phasiRNAs generated from a single locus (*PH-12*) between *O. barthii* (*Ob*) and *O. glaberrima* (*Og*). The detected phasiRNAs are ordered from the 5' end to the 3' end and named *phasiPH12-1* to *phasiPH12-10* (note that they may not be consecutive on the locus sequence). The absolute number on the y-axis represents the number of reads and their origin from strand (+) or strand (-) is indicated. **C,** Relative abundance of miRNA families between *O. barthii* (*Ob*) and *O. glaberrima (Og)*. LogPlot of normalized abundance of miRNA reads. Ob: *O. barthii*; Og: *O. glaberrima*. Green triangles represent the differentially expressed miRNA family, and orange crosses non-differentially expressed miRNA families between *O. barthii* and *O. glaberrima*.

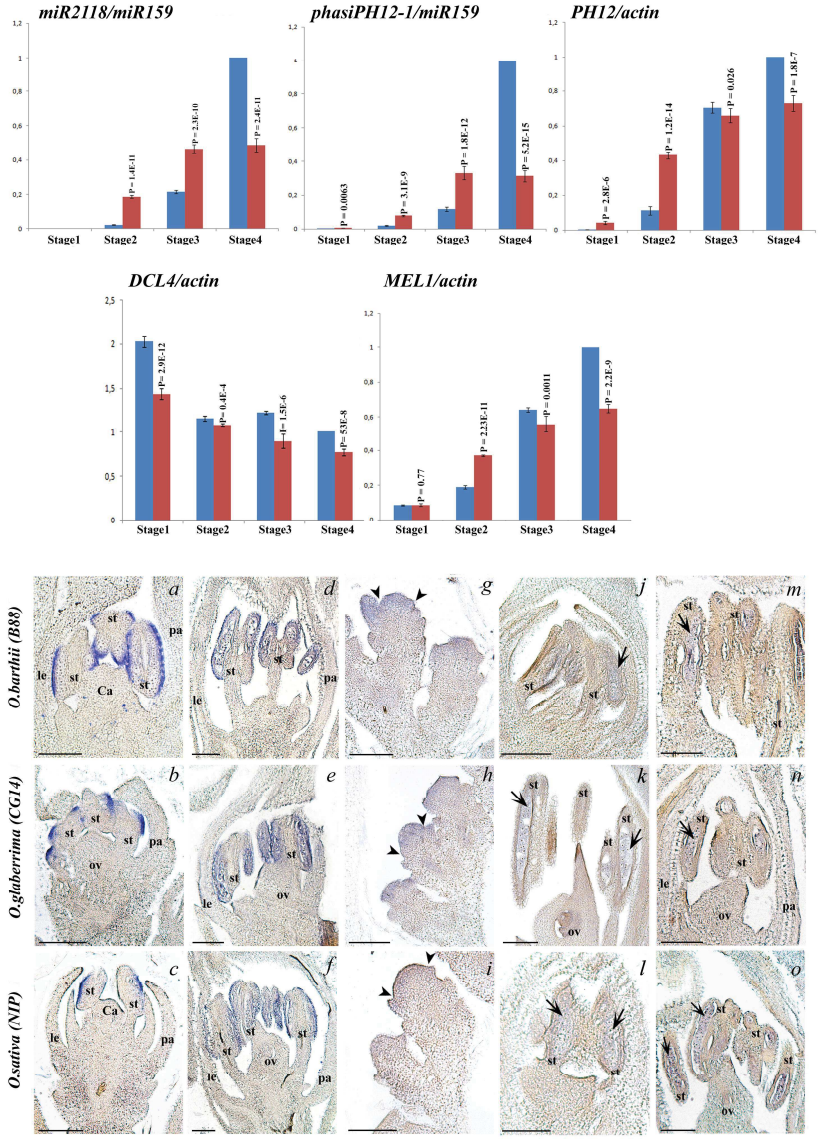


Figure 3. Expression analysis of small RNAs and related-regulators during panicle development. A, qRT-PCR analysis of *miR2118, PH12* precursor*, phasiPH12-1*, *OsDCL4* and *MEL1* expression levels during panicle development (from stage 1 to stage 4) in *O. barthii* (red) and *O. glaberrima* (blue). Relative expression levels are normalized using ratio values in *O. glaberrima* stage 4 as calibrator. B, *In situ* hybridization of mature *miR2118* (a-f), *PH12* mRNA precursor (g-l) and *phasiPH12-1* (m-o) in spikelet meristem (SMs), floret meristem (FMs) and differentiated floret in *O. barthii* (B88), *O. glaberrima* (CG14) and *O. sativa* (Nipponbare). *miR2118* transcripts were detected in the outer cell layer of stamens (*i.e.* epidermis) (a-c) and then extent to pollen sac (d-f). Precursor *PH12* mRNAs were detected in the spikelet meristem (g-i) and in pollen sac of the anthers (arrow) (j-l). The expression of *phasiPH12-1* was detected in the pollen sac of the anthers in florets (m-o). le, lemma; pa, palea; st, stamen; ca, carpel; ov, ovule. Arrowheads, lemma and palea primordium formation in SMs; arrow, anthers. Scale bar: 100μm.

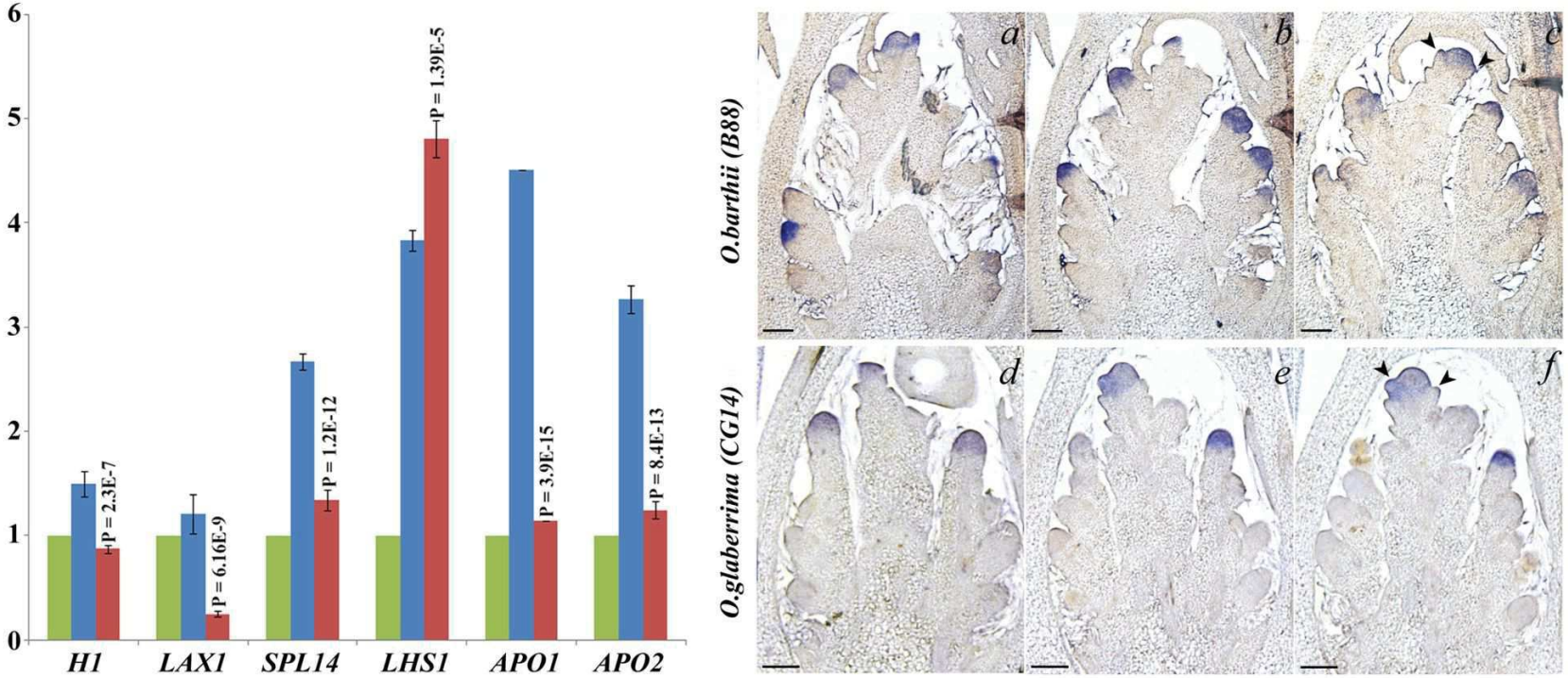


Figure 4. Expression analysis of landmark genes in young panicles from Asian and African rice. A, qRT-PCR analysis of *OSH1, LAX1, OsSPL14, LHS1, APO1* and *APO2* expression levels in panicle at stage 2 in *O. glaberrima (blue), O. barthii (red)* and *O. sativa (green).* Relative expression levels are normalized using ratio values in *O. sativa* as calibrator. B, *In situ* hybridization of *LHS1* in spikelet meristems (SMs) of *O. glaberrima* (a-c) and *O. barthii* (d-f); arrowheads, lemma and palea primordium formation in SMs. Scale bar: 100μm

Supplemental data

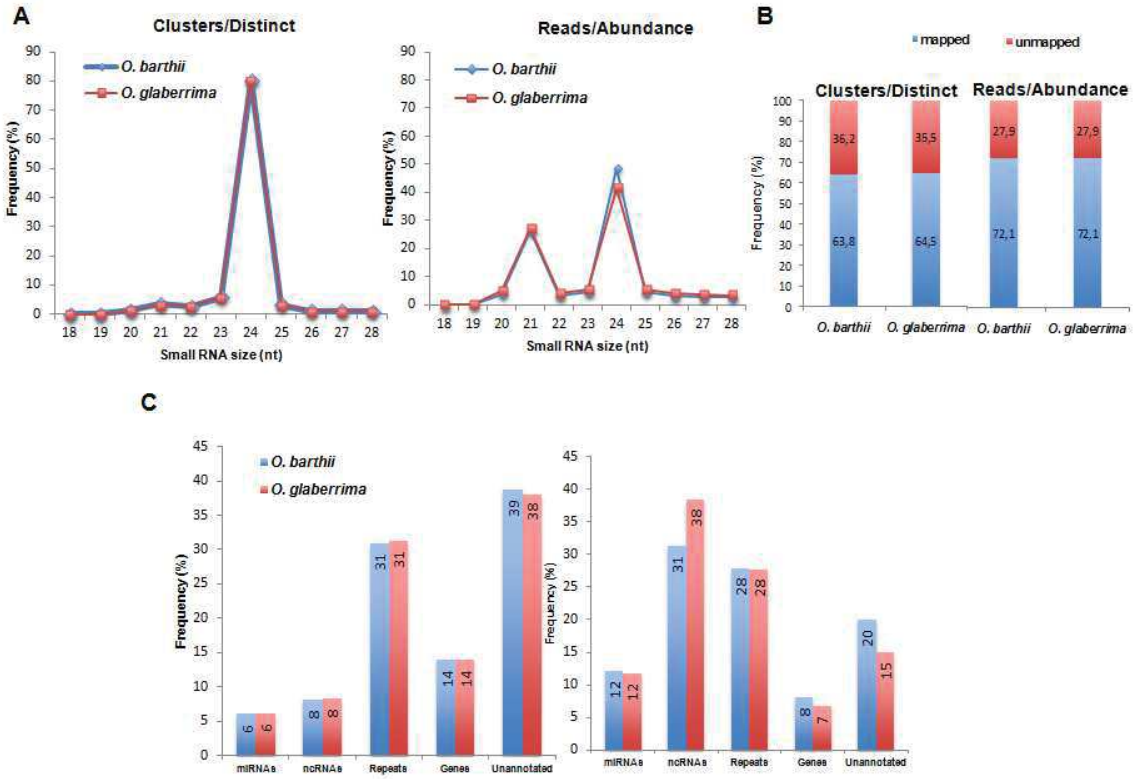


Figure S1. Size distribution, mapping and annotation of *O. barthii* and *O. glaberrima* panicle-derived small RNAs on *O. sativa Nipponbare* genome. A, Small RNA size distribution of clusters (distinct) and reads (abundance). The size of small RNAs was plotted *versus* frequency (percentage relative to total abundance). **B,** Mapping rate of clusters (ie. distinct sequences) and total reads on *O. sativa nipponbare* MSU7.0. **C,** Small RNA annotation distribution of clusters (distinct) and individual reads (abundance).

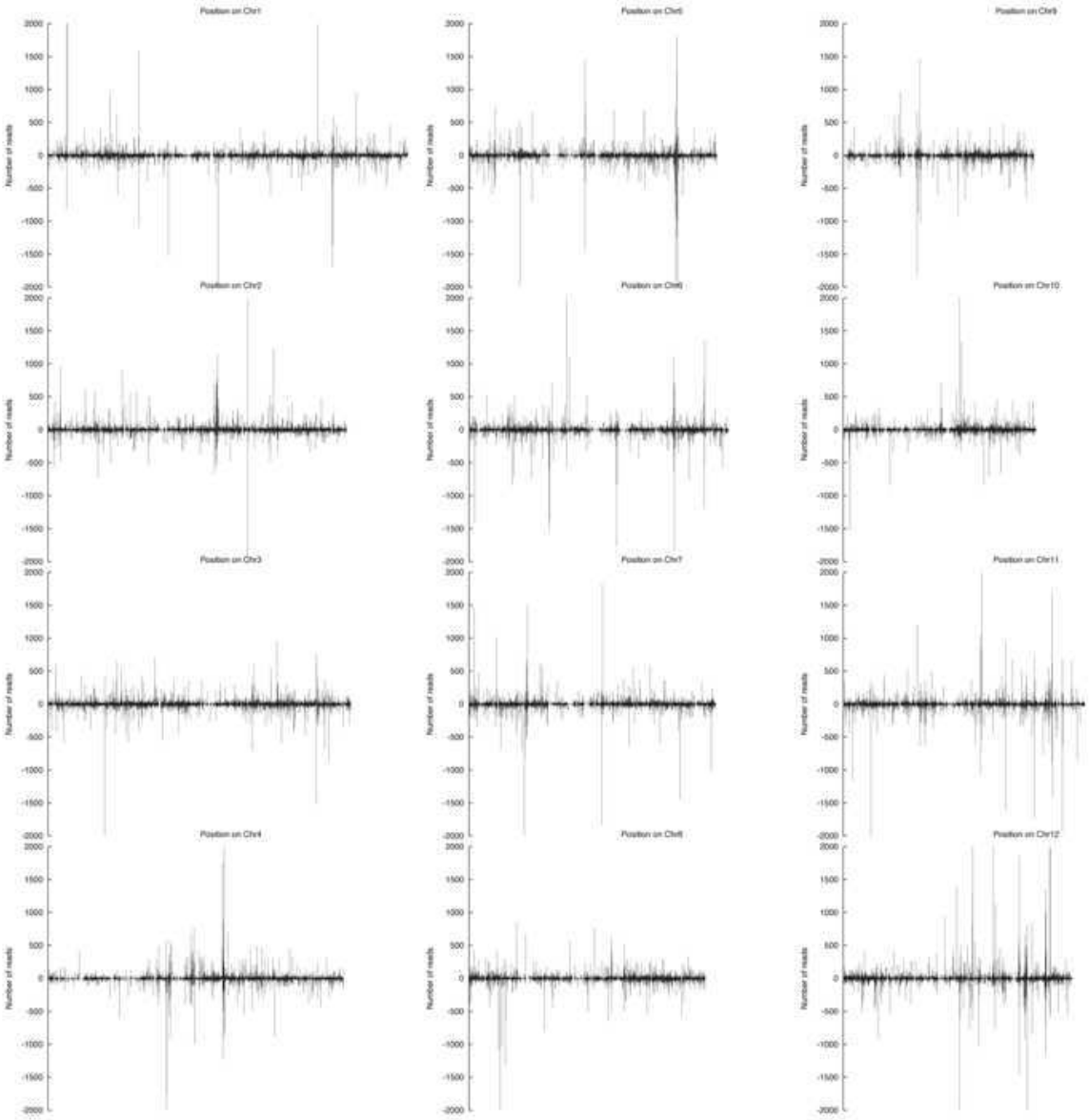


Figure S2. Genomic distribution and abundance of *O. barthii* and *O. glaberrima* panicle-derived small RNAs on *O. sativa Nipponbare* genome.

A, Distribution and abundance of *O. barthii* reads vs. *O. sativa Nipponbare* genome MSU7.0.

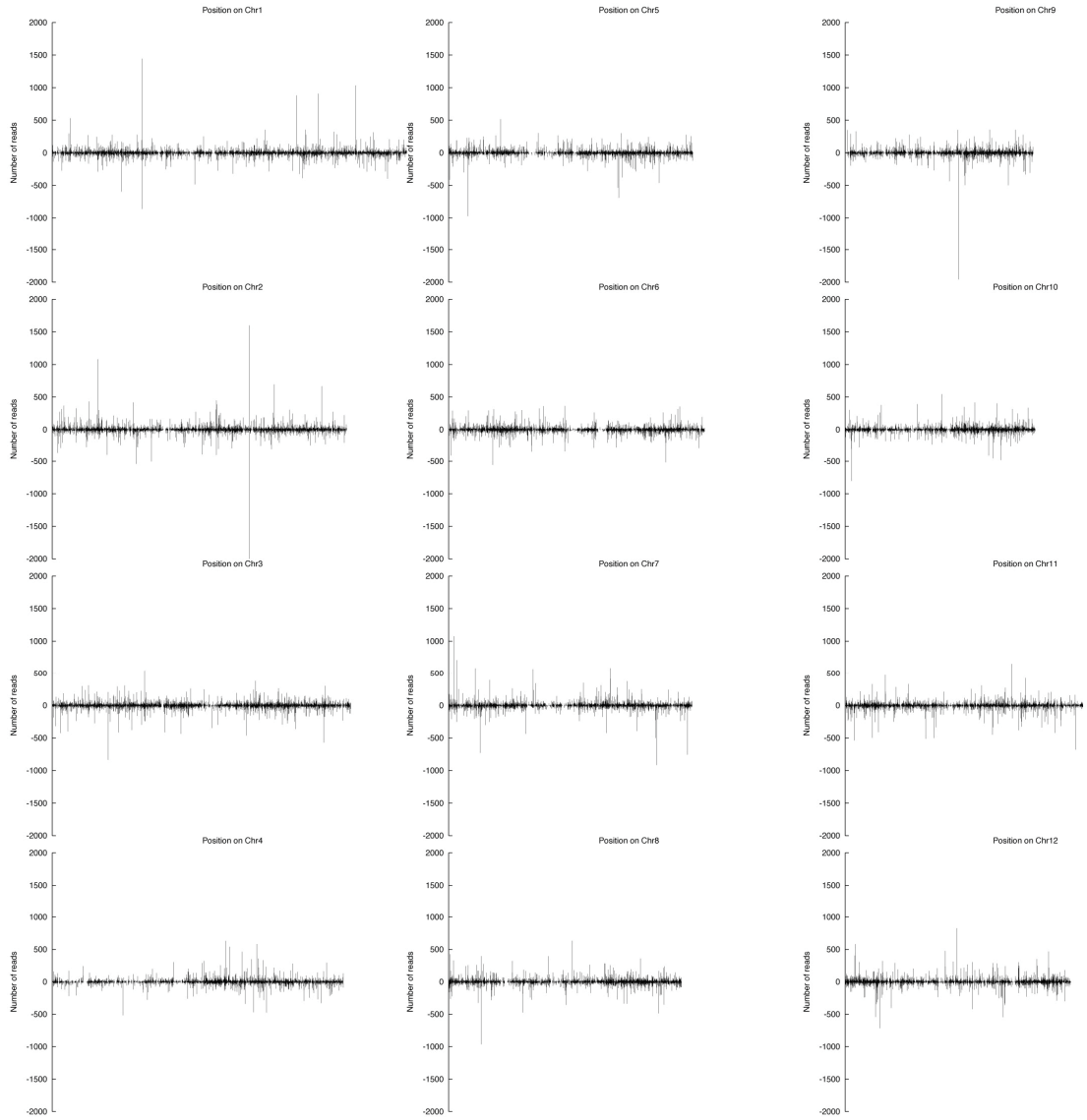


Figure S2 B, Distribution and abundance of *O. glaberrima* reads vs. *O. sativa Nipponbare* genome MSU7.0.

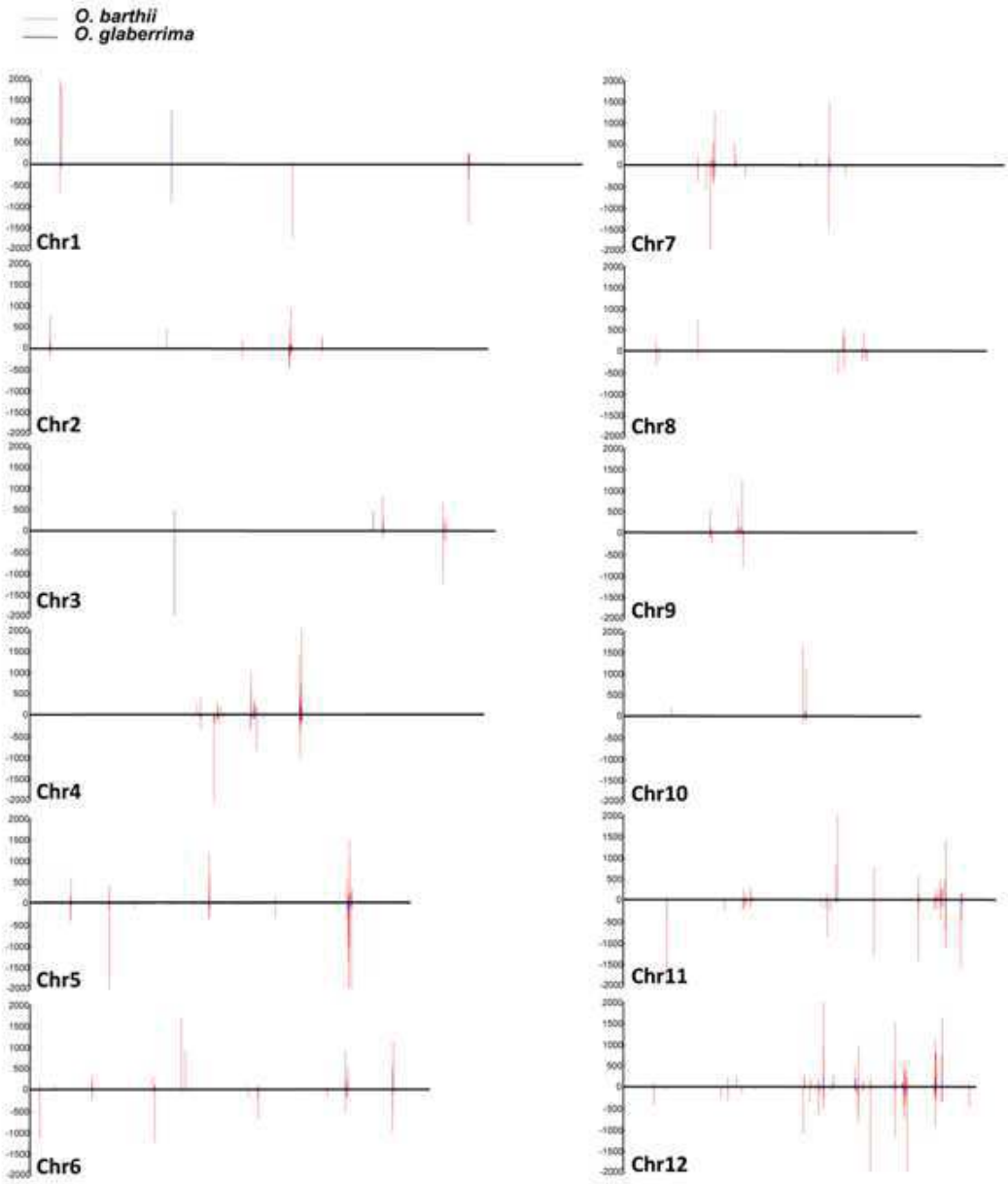
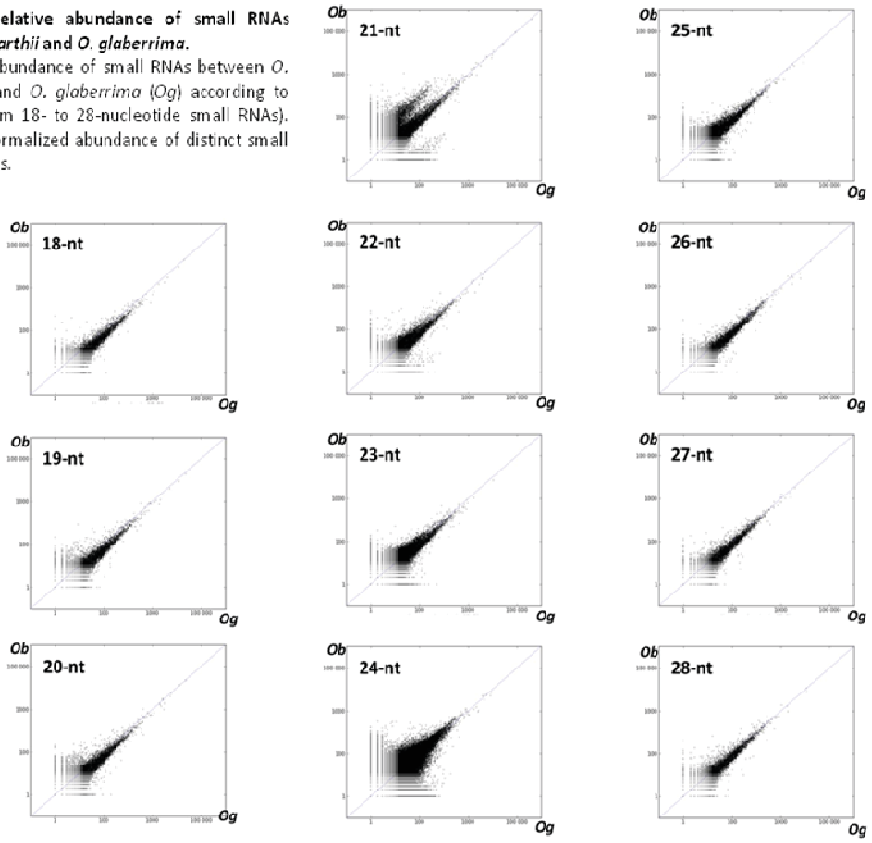
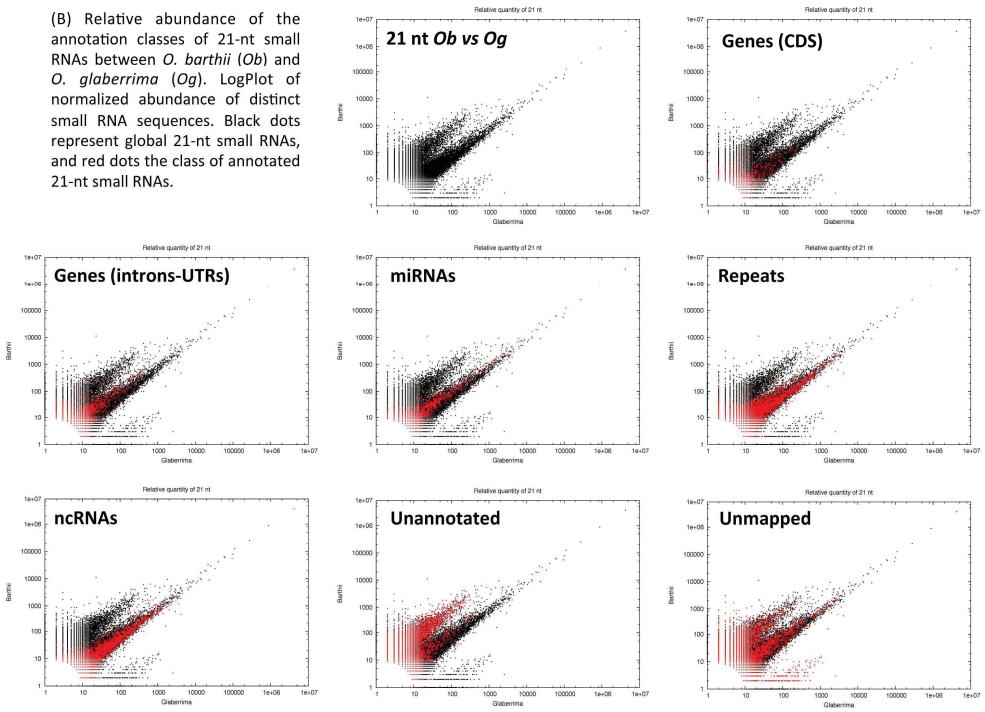


Figure S2. C, *O. barthii*- and *O. glaberrima*-derived phased 21-nt small RNAs on *O. sativa nipponbare* genome (MSU 7.0).

Figure S3: Relative abundance of small RNAs between *O. barthii* and *O. glaberrima*.
(A) Relative abundance of small RNAs between *O. barthii* (*Ob*) and *O. glaberrima* (*Og*) according to their size (from 18- to 28-nucleotide small RNAs). LogPlots of normalized abundance of distinct small RNA sequences.



(B) Relative abundance of the annotation classes of 21-nt small RNAs between *O. barthii* (*Ob*) and *O. glaberrima* (*Og*). LogPlot of normalized abundance of distinct small RNA sequences. Black dots represent global 21-nt small RNAs, and red dots the class of annotated 21-nt small RNAs.



(C) Relative abundance of the annotation classes of 24-nucleotide small RNAs between *O. barthii* (*Ob*) and *O. glaberrima* (*Og*). LogPlot of normalized abundance of distinct small RNA sequences. Black dots represent whole 24-nucleotide small RNAs, red dots represent the class of annotated 24-nucleotide small RNAs.

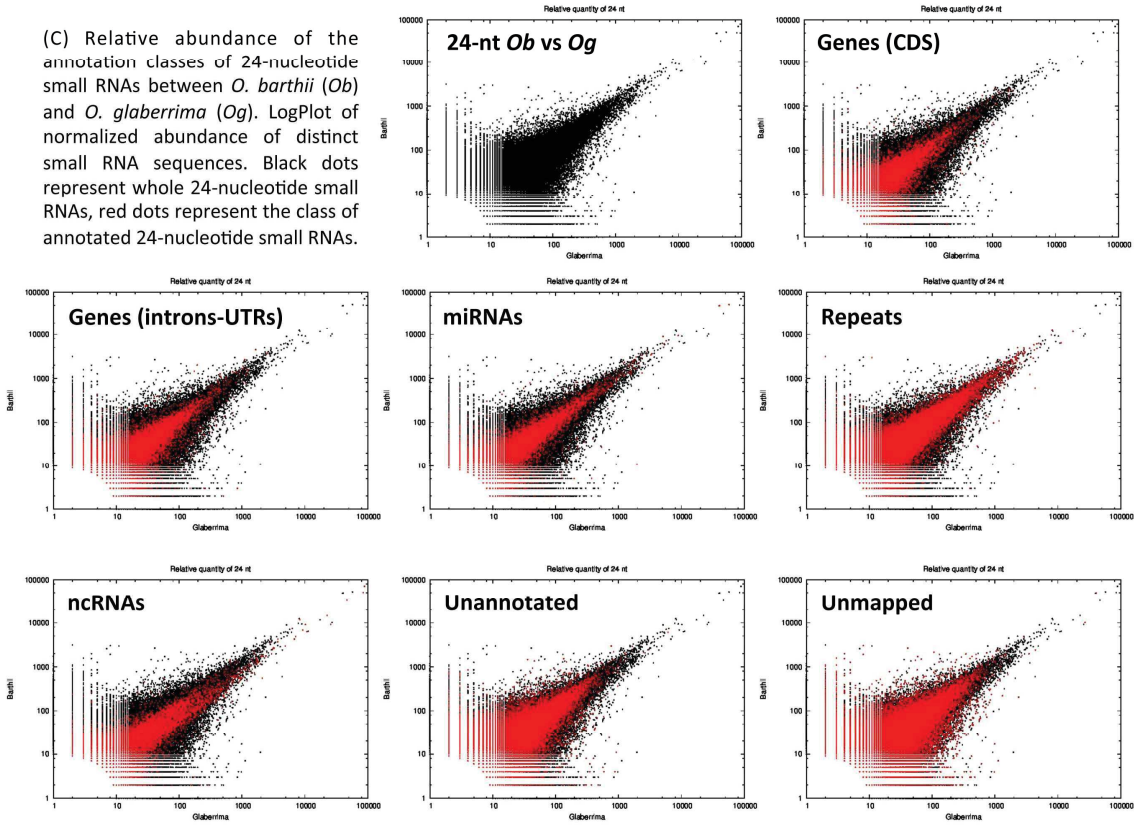
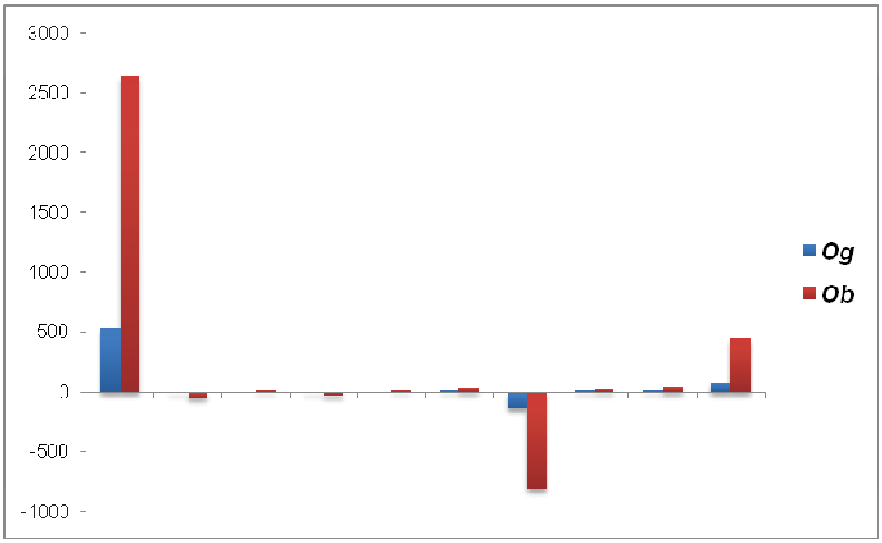


Figure S5. Sequences and phasiRNA abundances of the 4 phased loci used for validation. The 4 phased loci reported here were used for stem-loop RT-PCR (phasiRNAs) and classical RT-PCR (lncRNAs) validation. The positions of the phased loci on MSU7.0 chromosomes are indicated. The alignment between *miR2118f* and the phased locus is indicated. The *miR2118* recognition site is in blue. The phasiRNAs detected in the plus strand are highlighted alternately in green and in red. The phasiRNAs detected in the minus strand are underlined alternately in green and in red. The phasiRNA boxed in yellow is the most abundant one and the one used for stem-loop RT-PCR validation and qRT-PCR analysis. Histograms illustrate the relative abundance of the detected phasiRNAs per locus in the two species.

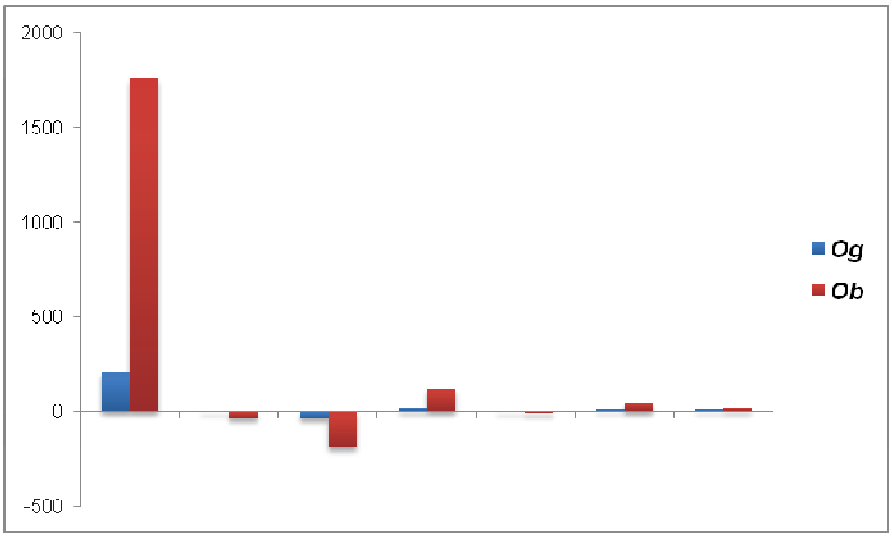
>phased_12 Chr1-2244143..2244489

5'AAUAGGCUUUGGAGACACCGGGAAUGGAAAACAUAGUUAGAUAUCAAUCUCUCGAUCGUUGAUGCCAUAAGCUACUUGGCU
•••••..•.••••.••.•.••••
3' AUCCUUACCCUCCGUAGUCCUU 5' *miR2118f*
UGCAAGUUUCCAUCCUGCGAGCUGUUAACCAGAAAGGGGAUGGUUGACAGUAGCCUGGAGUUUGGAUCGUUAUUGUGGAACUUG
AAUUAUGAAUCUGUCAUACAUUCCAAAAUUUACGUUGAGAAAUUGAGCUGAAAUCUGCUCGUUGAAGGAACCAACGACUCCUGC
CACGUUCAAUCUUUUUGCCAGAGAAUGAAGAGUAUGCAUGGAGGUGAAUUCUGAGCUUAUCGGAUCAGUUACUCAAUUGUUUGA
AGAUUAUACCUUGAAGA 3'



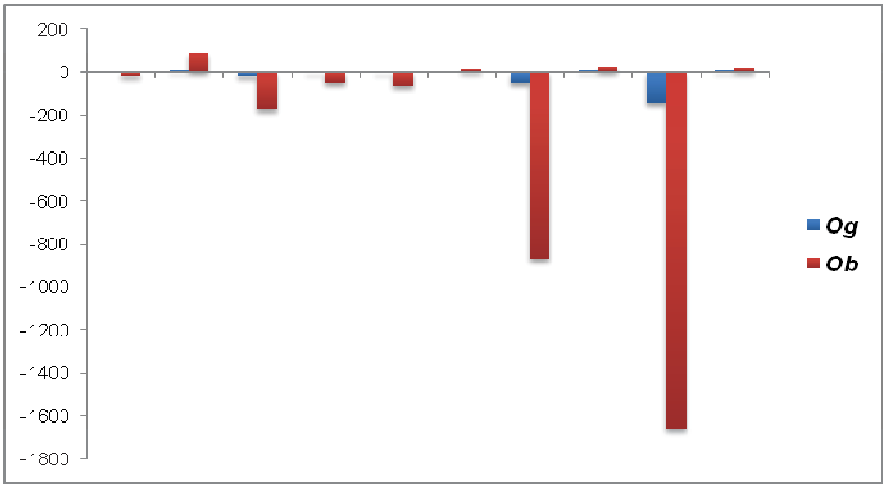
>phased_557 Chr4-21018743-21018938

5'AUAGGGCAUGGGAUGCGUC-GGAAAUUCAGGCAGCUCGUUCUGAUGAUUUGCGGCAUAGGAAACGUUGGGGAGCAGCGGGU
..••.•.••••.••.••.••••
3' AUCCUUACCCUCCGUAGUCCUU 5' *miR2118f*
ACGCUUCUGCUGUUUGUUAAUUCCUCGACCGUCGACUGGUGGUUUGUCGGUGUGACCAGAUGUGUCCUGUUCUAUCCUUUGGUA
CGCGACAAAUUUCUGCAACCAUAAAACCCCAA 3'



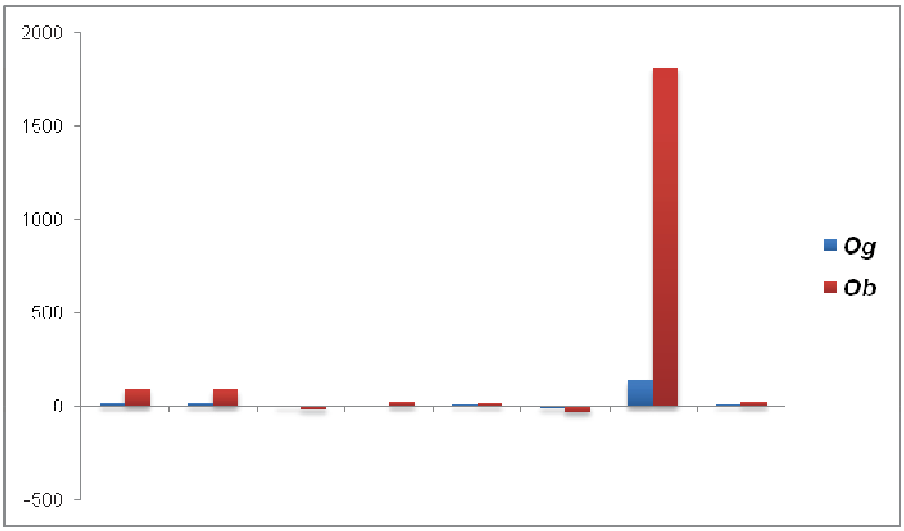
>phased_612 Chr5-24872805..24873056

5′ GUUGCGAAUGUGAUGCAUUGGGAAUACGAGUGGAACGAACCAGCCAUCAGAAAGCUCAUGCCAACCAGCGAGGUGUACUA
•..•••••.••.••••..••••
3′ AUCCUUACCCUCCGUAGUCCUU 5' *miR2118f*
AUAAGCUAGCUACUAUGAGCGACGCUCUUCAGGGAUGCCAACCAACGUGUAGGUUUCUCGUCAAUCCAAUCCUUGGUCACCGUU
CCCCAAGUCCUGCUGGACACUUUUGGUAAAGAAAAAGAUAAUGAAAUUCUUCUGAUCCCGUUAUGUCUCUUCUGCAAAUGUUAA
AUUUUUCAUUAUAAGCUGCAAUUCAGAUUCUUUGUAUAG 3′



>phased_779 Chr7-16046295..16046641

5′UUUGGGUUUGGGAGGCUUGGGGAAGUGAGGUGCUACUAUCUUGCAUUUCUUGCUUAGCUGCUGCGAAGAUUUCAGCUGCA
•.••..•••••••••.•..••••
3' AUCCUUACCCUCCGUAGUCCUU 5' *miR2118f*
UCAUUUGCUCAGCUGGUGUGAAGAUUCCUGCUCACUCGCUUUGUCCACAAUAGGGCAUCACUAUCGGCAAUGACAAGAGUAGUG
GAAAAAAUAGAAGCAGUACAUGAAAGCUUCUACAGUGACCGUAUCGAACGCCAUAGAUGAUCAUGAAUCAGCAGUAGAUGAUAU
CAGAUGAUCUAUCUCCUGGCUUUUAGUUGUAUGAACGCUGGAUUUGAGGUGAAUUACUUGUUUGUGUUGCUGAUCGAGUUUUGA
GCGUGCGCCUAACUAGUU 3′


2000
1500
1000
500
0
-500
Og
Ob

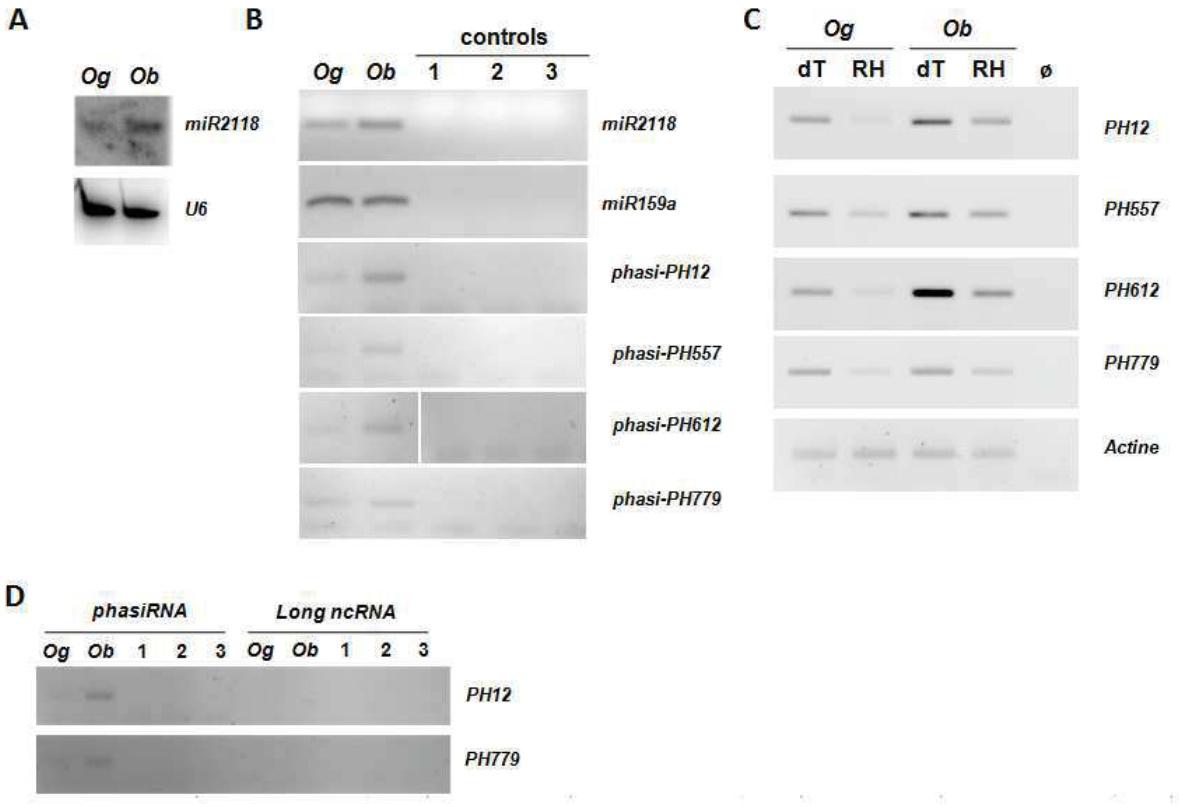


Figure S6. RNA-seq data validation by northern-blotting and semi-quantitative RT-PCR for miRNAs, phasiRNAs and ncRNAs associated with phased loci in panicle-derived small RNA bulks from *O. barthii* (*Ob*) and *O. glaberrima* (*Og*). A, Northern blot hybridization using *miR2118* as probe. U6 probe was used as control. **B,** Stem-loop RT- PCR analysis of various miRNAs and phasiRNAs. *miR159a* probe was used as a loading control. **C,** Classic RT-PCR analysis of ncRNAs associated with 21-nt phased small RNA loci using from polydT primer (dT) or random hexamer primers (RH) for the RTs. The *Actin* gene (*Os03g50885*) was used as a loading control. **D,** Test of specificity of stem-loop RT-PCRs against phasiRNAs on *PH12* and *PH779* loci, in conjunction with phasiRNA specific forward primer (PH12-F and PH779-F primers; phasiRNA label) and long ncRNA forward primer (PH12-F2 and PH779-F2 primers; long ncRNA label). Controls: 1. RT reaction using *O. glaberrima* RNA bulk without stem-loop RT primer; 2. RT reaction using *O. barthii* RNA bulk without stem-loop RT primer; 3. RT-PCR without stem-loop RT matrix; **Ø:** RT-PCR without RT. See Supplemental Table S3 for primer and probe sequences.

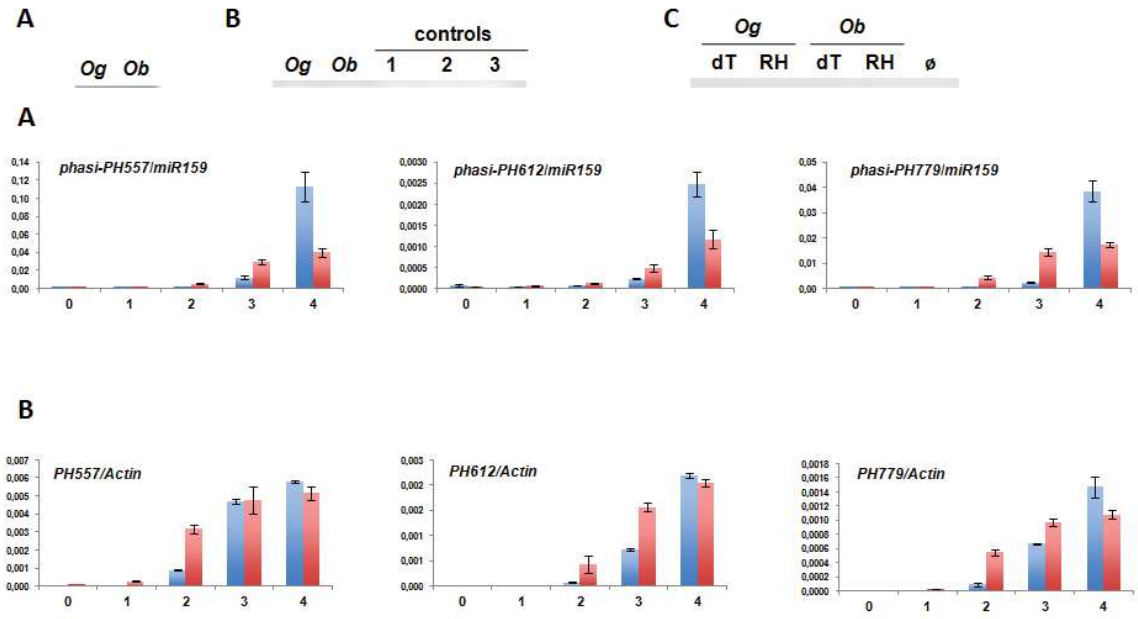


Figure S7. Expression patterns of phasiRNAs and ncRNAs associated with 21-nucleotide phased loci in *O. glaberrima* and *O. barthii*. (*A*) Histograms of relative expression levels of 21-nucletiode phasiRNAs from 3 distinct phased loci (*PH557, PH612* and *PH779*) during panicle development in *O. glaberrima* CG14 (blue bars) and *O. barthii* B88 (blue bars) using real-time stem-loop RT-PCR amplifications. *miR159a* mature miRNA was used as reference small RNA. 0: stage 0, shoot apical meristem before flowering induction; 1 to 4: stages 1 to 4, young panicle at different developmental stages (see Materials and Methods section). See Supplemental Table S3 for primer sequences. (*B*) Histograms of relative expression levels of lncRNAs associated with 21-nt phased loci *PH557*, *PH612* and *PH779* during panicle development in *O. glaberrima* CG14 and *O. barthii* B88 using real time polydT RT-PCR amplifications. The *Actin* gene (*Os03g50885*) was used as reference gene. See (A) for the legend. See Supplemental Table S3 for primer sequences.

Supplemental Tables

Table S1: lists of genotypes used in this study

Species	Accession nb	Origin
O. barthii	B197	Lake Tchad
O. barthii	B117	Botswana
O. barthii	B88	Cameroun
O. barthii	B74	Lake Tchad
O. barthii	B70	Lake Tchad
O. barthii	B68	Lake Tchad
O. barthii	B64	Nigeria
O. barthii	B48	Mali
O. barthii	B5	Lake Tchad
O. barthii	B49	Lake Tchad
O. glaberrima	Tog 6208	Guinea
O. glaberrima	Tog6211	Nigeria
O. glaberrima	AC104589	Burkina Faso
O. glaberrima	MG12	Mali
O. glaberrima	CG14	Senegal
O. glaberrima	Tog 5681	Nigeria
O. glaberrima	Tog 7020	Sierra Leon
O. glaberrima	Tog 6221	Burkina Faso
O. glaberrima	Tog 5887	Liberia
O. glaberrima	Tog 5500	Nigeria

Table S2: summary statistics of small RNA libraries

	O. barthii	***O. glaberrima***
Raw data	71.4 $x10^6$ reads	73.3 $x10^6$ reads
High quality clustered data	33.9 $x10^6$ reads	33.1 $x10^6$ reads
	9.5 $x10^5$ clusters	9.0 $x10^5$ clusters
Annotated data	26.3 $x10^6$ reads	26.0 $x10^6$ reads
	6.4 $x10^5$ clusters	6.2 $x10^5$ clusters

Table S3: list of primers used in this study. The bases in brackets represent the ones that were modified LNAs. The sequences in italic correspond to stem-loop region for stem-loop RT-PCRs. The underline sequences correspond to T7 RNA polymerase binding site used for RNA probe synthesis.

Name	Sequence
Northern-blotting	
miR2118f	TAGGAATGGGAGGCATCAGGAA
U6	GCAGGGGCCATGCTAATCTTCTCTGTATCGT
Stem-loop qRT-PCR	
miR2118-RT	*GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC*TAGGAA
miR2118f-F	CGGCGGTTCCTAATGCCTCCCA
miR159b-RT	*GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC*CAGAGC
miR159-F	CGGCGGTTTGGATTGAAGGGA
Univ-RT	GTGCAGGGTCCGAGGT
Classic qRT-PCR	
PH12-RT	*GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC*CCCTTT
PH12-F	CGGCGGCGAGCTGTTAACCAG
PH557-RT	*GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC*ATGCCG
PH557-F	CGGCGGCGTTCTGATGATTTG
PH612-RT	*GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC*CCGTTC
PH612-F	CGGCGGCCAGCAGGACTTGGG
PH779-RT	*GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC*TCATCT
PH779-F	CGGCGGTCATGAATCAGCAGT
PH12-F2	TCTCGATCGTTGATGCCATA
PH12-R1	GGAGTCGTTGGTTCCTTCAA
PH557-F2	TCTGATGATTTGCGGCATAG
PH557-R1	GGGGTTTTATGGTTGCAGAA
PH612-F2	GAACCAGCCATCAGAAAGCT
PH612-R1	AAGTGTCCAGCAGGACTTGG
PH779-F2	TTGTCCACAATAGGGCATCA
PH779-R1	ACCTCAAATCCAGCGTTCAT
ACT-F	CATTCCAGCAGATGTGGATTG

ACT-R	TCTTGGCTTAGCATTCTTGG
OSH1_F	CAGTTCGTGATGATGGAC
OSH1_R	CTAAAACCGACCCCTGCATTA
APO1-F	GTTCTACTGCATGAGCTCGTC
APO1-R	TGCACCTTGCTCCATACGTTC
APO2-F	AGGTGCAATCCATGGCTAAG
APO2-R	GCATCTTGGGCTTGTTGATG
LHS1-F	GTGACCATTCCCTGCAGATT
LHS1-R	GTCTGCTGCTTCATTGCTCA
SPL14-F	CTGCCTGAATTTGACCAAGG
SPL14-R	AAGCTTCTGAACCTGCGATG
LAX1-F	ATTACCGGTTGGTCATGGTC
LAX1-R	AAGCGATCGAGCAAACAAGT
OsDCL4-F	TCAGAAGAAGGCTGCACAAG
OsDCL4-R	CGAACGTCCTCTTCTTTTGG
OsMEL1-F	TCCCAAGATCAAGGAGAACG
OsMEL1-R	ACAAGCAACCAGCTCCAAAC

***In situ* hybridizations**

LNA-miR2118f	TAGG(A)ATGGG(A)GGC(A)TC(A)GGAA
LNA-PH12	CCC(T)TTC(T)GGT(T)AAC(A)GCTCG
LHS1-HIS-F	GAAGAGCAAGGAGCAACAGC
LHS1-HIS-R	AATCTGCAGGGAATGGTCAC
LHS1-HIS-T7R	GCGAAATTAATACGACTCACTATAGGGCGAAAATCTGCAGGGAATGGTCAC
PH12_T7R	GCGAAATTAATACGACTCACTATAGGGCGAAGGAGTCGTTGGTTCCTTCAA

2.2.3 Manuscript 2: Differential expression of panicle-related landmark genes between *Oryza glaberrima* and its wild-relative *Oryza barthii*.

Ta KN et al.

Differential expression of panicle-related landmark genes between *Oryza glaberrima* and its wild-relative *Oryza barthii*.

Ta KN[1,2], Gantet P[2,3], Do NV[2], Ghesquière A[1], Jouannic S[1,2], Adam H[1]*

1. IRD, UMR DIADE, 911, avenue Agropolis, BP64501, F-34394 Montpellier Cedex 5, France
2. IRD, LMI RICE, University of Science and Technology of Hanoi, Agronomical Genetics Institute, Pham Van Dong road, Hanoi, Vietnam
3. UM2, UMR DIADE, 911, avenue Agropolis, BP64501, F-34394 Montpellier Cedex 5, France

Key words: panicle, branching, meristem fate, African rice

Abstract

Panicle architecture has a direct impact on grain yield potential and thus has been a major trait selected during rice domestication. The panicle consists of a series of branching orders: rachis, primary branch, secondary branch, potentially tertiary branch and finally spikelets. Spikelets are grass-specific short branches bearing single flowers. All branches are generated from axillary meristems and their identities are subsequently defined according to their position and the timing of initiation. Rice panicle development is controlled by a set of genes that regulate axillary meristem initiation, identity meristem, meristem fate, cell division and hormone signaling that has been well described in *Oryza sativa.* The African rice *Oryza glaberrima* has been domesticated from its wild relative *Oryza barthii* about 3000 years ago. During domestication process, the panicle complexity changed from a panicle with few primary and secondary branches bearing relatively few grains in *O. barthii*, to a domesticated highly branched panicle carrying larger numbers of seeds. In the present study, we tested whether an alteration of spatial and temporal expression of panicle-related landmarks genes could be related to the panicle diversity observed between the two African species. Our results have shown a conservation of the spatial expression pattern of the landmarks genes studied but have highlighted a differential timing and level of the expression of these genes during the panicle development in the two species. These results suggest that variation of panicle complexity in African rice may be related to a heterochronic change in acquisition of meristem determinacy.

Introduction

The African rice *Oryza glaberrima* and the Asian rice *Oryza sativa* are the only two species of cultivated rice in the world. While *O. sativa* was domesticated about 10 000 years ago, *O. glaberrima* has a shorter history as it derived from its wild ancestor *Oryza barthii* along the Niger river in Mali about 3 000 years ago (Vaughan et al. 2008; Huang et al. 2012). Recently, different studies brought evidences that African rice domestication was linked with a single domestication origin in West African, associated with a severe genetic bottleneck (Li et al. 2011; Nabholz et al. 2014; Orjuela et al. 2014; Wang et al. 2014). However, in contrast to Asian rice domestication that has been the topic of extensive works; African rice domestication has been less studied in terms of molecular genetics. Although African rice maintained very low genetic diversity compared to Asian rice (Li et al. 2011; Nabholz et al. 2014; Orjuela et al. 2014; Wang et al. 2014), its close relationship with Asian rice species and its more simple domestication history, make it also a good model for the study of the evolution of morphological traits and the associated molecular gene network along rice domestication.

Several morphological traits were selected during the domestication, such as tillering, seed color, seed shattering, and many other traits was known as the "domestication syndrome" (Doebley et al. 2006). Among them, panicle architecture is one of the main morphological trait selected during the two domestication processes in rice (Sweeney and McCouch 2007). Rice panicle architecture results from the establishment and the activity of apical and axillary meristems deriving from the vegetative shoot apical meristem (SAM) (Ikeda et al. 2004). In the reproductive phase, the SAM converts into rachis meristem (RM), which will produce primary branch meristems (PBMs) until its abortion. These PBMs will contribute to the establishment of the primary branches as well as the secondary branch meristems (SBMs), with eventually tertiary branch meristems (TBMs) from the developing secondary branches. Finally, all the lateral and terminal meristems convert to spikelet meristems (SMs) then florets. Therefore, rice panicle architecture is determined by two phases: meristem establishment and activity (branching process) and meristem fates (transition from branch meristems to SM). For decades, quantitative trait locus (QTL) mapping and mutant characterization in *O. sativa* revealed a large number of genes required for the initiation and the development of the panicle, as well as genes controlling numbers and size of grains (Xing and Zhang 2010; Wang and Li 2011). Some mutations such as *monoculm1/small panicle (moc/spa)*, *lax panicle1 (lax1), lax panicle2 (lax2)* have been shown to affect the patterning of axillary meristems (AMs) and the panicle development (Komatsu et al. 2001; Komatsu et al. 2003). In the *tawawa1*-D gain of function mutant, the rachis meristem (RM) activity is extended compared to wild type plants and spikelet specification is delayed, resulting in prolonged branch formation and increased numbers of spikelets (Yoshida et al., 2012). The *aberrant panicle organization1* (*apo1*) and *aberrant panicle organization2* (*apo2*) mutants form small panicles

with reduced numbers of branches and spikelets while the overexpression of *APO1* and *APO2* genes causes large panicles with an increased number of spikelets (Ikeda et al. 2007; Rao et al. 2008). Moreover, the *APO1* gene was also identified as a QTL related to number of primary rachis branches and vascular bundle formation in panicles (Terao et al. 2009). Two other genes were characterized as QTLs related to grain yield, namely *DENSE AND ERECT PANICLE* (*DEP1*) and *IDEAL PLANT ARCHITECTURE 1* (*IPA1*)*/WEALTHY FARMER'S PANICLE* (*WFP*). *DEP1* encodes for a previously unknown protein and shows effect on spikelet number by enhancing meristematic activity and promoting cell proliferation (Huang et al. 2009). *IPA1/WFP* corresponds to *OsSPL14*, a SBP-box (SQUAMOSA promoter binding protein-like) protein-encoding gene that is targeted by the microRNAs *miR156* and *miR529* to promote panicle branching (Miura et al. 2010; Jiao et al. 2010; Jeong et al. 2011).

The two African species *O. glaberrima* and *O. barthii* are characterized by differential panicle complexity in terms of primary branch number and branching order level (*i.e.* establishment of secondary branching). In a previous study, we have shown that the differential panicle complexity in the two African species was associated to differential timing of expression of the spikelet-specific gene *LEAFY HULL STERILE1/OsMADS1* (*LHS1*) and of the *miR2118*-triggered phasiRNA pathway related to male gametogenesis (Ta et al submitted). To investigate whether the differential panicle complexity between the two African rice species was associated to differential expression of other genes related to panicle development, we carried out a detailed analysis of the spatial and temporal expression patterns of a set of orthologs of *O. sativa* landmark genes related to meristem activity and meristem fate control. Our work has shown a high conservation of the spatial expression pattern of these genes between the two species, but differences in their expression levels and timing as well as polymorphisms in their promoter regions were evidenced.

Results

Expression of panicle landmark genes during rice panicle development

To obtain further insights into mechanism regulating panicle development of African rice, we further divided the early panicle development into 4 stages as described in Ta et al (submitted). Briefly, stage 1 corresponds to the elongation of rachis meristem and the formation of primary branch. In stage 2, higher order branch is determined. The transition from branch meristems (BMs) to spikelet meristems (SMs) and the differentiation of floral organs/flower development occurred at stage 3 and stage 4 respectively (see Supplemental Figure S1). In this study, we focused on the branching process (*i.e.* meristem establishment and activity) and the transition phase to SMs (*i.e.* meristem fate), as the basic complexity of the rice panicle architecture is determined at this stage.

Landmarks genes related to panicle development in *O. sativa* have been used to achieve gene-expression diversity between the two African species. Firstly, genes controlling the initiation and/or the maintenance of lateral meristems were used as molecular markers of branch meristem activity, such as *Oryza sativa homeobox1* (*OSH1*)*, LAX PANICLE1* (*LAX1*) (Tsuda et al. 2011; Woods et al. 2011) and the *SQUAMOSA promoter binding protein-like14* (*SPL14*) as well as its microRNA regulators *miR156* and *miR529* (Jiao et al. 2010; Miura et al. 2010; Jeong et al. 2012). Quantitative RT-PCR analyses showed differences of transcript levels during the early panicle development between the two African species *O. glaberrima* and *O. barthii*. *OSH1* and *LAX1* mRNAs were detected at higher levels in *O. glaberrima* than in its wild-relatives during both branching phase spikelet/floret phase corresponding from stage 2 to stage 4 (Fig. 1). This may be related to a higher number of branch meristems in *O. glaberrima* than in *O. barthii.* However, in terms of kinetics, the expression levels of both *OSH1* and *LAX1* genes reached a peak level at stage 2 or 3 in *O. glaberrima* then decrease later on, in contrast to *O. barthii* in which both genes reached a peak of expression level at stage 1, decreasing gradually later on (Fig 1). This finding indicates that the timing of accumulation of these genes is maintained a longer time in the domesticated rice in comparison with its wild-relative. The accumulation patterns of *SPL14* orthologous genes in the two African species were quite similar with a peak of accumulation at stage 1, decreasing gradually in later stages (Fig.1). Similarly to the two previous genes, the accumulation level of *SPL14* was higher in *O. glaberrima* than in *O. barthii*, with an exception at stage 4 with similar relative ratio values for the two species (Fig. 1). In parallel, the *miR529* accumulation pattern mirrored the *SPL14* one pattern with a sharp increase of accumulation level reaching a peak at stage 3, with a 40- and 70-fold change in *O. glaberrima* and *O. barthii* respectively related to *O. glaberrima* stage 1 values (Fig. 1). Overall, the *miR529* microRNA showed a higher accumulation level in *O. barthii* than in *O. glaberrima*, in agreement with the opposite situation for *SPL14* genes in these two species. We also investigated *miR156* accumulation levels during panicle development in relation to *SPL14* genes and *miR529*. The overall accumulation levels of *miR156* during panicle development do not change at the same range that *miR529* ones (*i.e*. two-fold increase), in parallel with lower Cp values for *miR156* (Fig. 1). The accumulation level of *miR156* reached a peak at stage 2 in *O. glaberrima,* while the peak was observed at stage 3 and 4 in *O. barthii* (Fig. 1). Based on opposite accumulation patterns of *SPL14* and *miR529* as well as the high induction of *miR529* in panicle in contrast to *miR156,* these would suggest that *SPL14* genes were mainly regulated by *miR529* rather than *miR156* in early stage of rice panicle development. Moreover, these results show the lower level of accumulation of *SPL14* genes in *O. glaberrima* was related to a higher accumulation of *miR529* in this species. Overall, the two African species were characterized by difference in their accumulation level and not the timing of accumulation of *miR529* and its putative mRNA target.

The second sub-set of genes used as molecular markers was related to meristem fate control. This sub-set of landmark genes included *LEAFY HULL STERILE1/OsMADS1* (*LHS1*) gene, which promotes the formation of spikelet/floret meristems (Khanday et al. 2013), as well as *ABERRANT PANICLE ORGANIZATION1* (*APO1*), *ABERRANT PANICLE ORGANIZATION2* (*APO2*) and *TAWAWA1 (TAW1)* genes, reported as suppressor of the transition from branch meristems to spikelet meristems (Rao et al. 2008; Ikeda et al. 2007; Yoshida et al. 2012; Ikeda-Kawakatsu et al. 2012). The qRT-PCR results have shown that the accumulation of *LHS1* transcripts increased sharply from stage 1 to stage 4 in both species, with a higher accumulation in *O. barthii* than in *O. glaberrima* at all stages. In contrast, the accumulation levels of *APO1, APO2* and *TAW1* genes decreased from stage 1 to stage 4, in parallel to a lower accumulation in *O. barthii* than in *O. glaberrima*, especially during branching phase (*i.e*. stages 1 and 2). However, the decreasing accumulation level for the *APO1* and *APO2* over the developmental kinetics is more rapid in *O. barthii* than in *O. glaberrima*, indicating that the accumulation of theses transcripts was maintained over a longer time in *O. glaberrima* than in its wild-relative. As a result, the higher expression of *APO1*, *APO2* and *TAW1* and the lower and later expression of *LHS1* in *O. glaberrima,* as well as the peak of *APO1* and *APO2* expression maintained over a longer period in *O. glaberrima,* support the fact of a panicle with a higher and longer branching phase in *O. glaberrima* than in *O. barthii*.

Spatial expression pattern of landmark genes regulating rice panicle development.

In order to determine whether the differences of accumulation levels of these landmark genes may be related to alteration of their spatial expression pattern or not, we performed *in situ* hybridization analyses for these genes (Fig2 and Supplemental Fig. S2). Overall, the spatial patterning of the accumulation of the corresponding transcripts was conserved between the two African species (Fig. 2). However a diversity of patterns was observed between the genes. In our study, a strong accumulation of *OSH1* transcripts was observed in the vascular bundles but also in the entire of branch meristems (but not in the epidermis) until initiation of spikelet/floret meristems (Fig 2 and Supplemental Fig. S2) as previously reported in *O. sativa* (Sentoku et al 1999). In contrast to this gene, *LAX1* transcripts were restricted specifically in the adaxial boundary region of new BMs and remains in the young SMs (Fig. 2 and Supplemental Fig. S2) as reported by Komatsu et al (2003) in *O. sativa*. Similarly, *SPL14* transcripts were detected in the boundary region of new BMs in the two species (Fig. 2 and Supplemental Fig. S2) as reported by Luo et al (2012) in *O. sativa*. In contrast, the two microRNA regulators, *miR529* and *miR156,* were both detected in the entire new branch meristems (including epidermis) and were maintained in the florets with differentiating organs (Fig. 2 and Supplemental Fig. S2). This finding would suggest that the microRNA-mediated regulation of *SPL14* transcripts is not relying on a dampening system but on exclusion/restriction one.

In agreement to what was reported by Kobayashi et al. (2010) in *O. sativa*, a strong signal corresponding to *LHS1* transcripts was detected in the spikelet meristem and weakly in palea and lemma of FMs but not in differentiating floret inner organs neither in branch meristems (Fig 2 and Supplemental Fig. S2). However, the other landmark genes acting as regulator of meristem fates, such as *APO2* and *TAW1,* were not only expressed in the SMs as *LHS1*, but also in all lateral meristems from early stage of rice panicle development (Fig 2 and Supplemental Fig. S2). These findings are in agreement with the qRT-PCR developmental kinetics, in which both *APO2* and *TAW* genes were expressed at higher level in earlier stage, namely stages 1 and 2. Unfortunately, we were unable to get reproducible *in situ* hybridization signals using several probes corresponding to *APO1* transcripts for the two African species as well as for the Asian one *O. sativa*.

Divergence of promoter structure between *O. glaberrima* and *O. barthii*

According to the above results, the spatial expression of landmark genes tested in here is highly conserved among *Oryza sativa* species complex. However, differences of level and timing of expressions were observed for those genes between the two African species. This would suggest that during domestication in African rice, those genes might have experienced differential *cis*- or/and *trans*-regulations.

To determine whether polymorphisms/mutations could affect protein sequences, we annotated the corresponding genomic sequence between *O. glaberrima* and *O. barthii,* taking advantage of the recent release of *O. glaberrima* genome and associated *O. barthii* data (Wang et al. 2014). The results demonstrated a high conservation of coding sequence between the two African species (data not shown), with the exception of *OSH1* gene having two SNPs in the coding region of *O. barthii* compare to *O. glaberrima,* leading to one amino acid change. However, the sequence comparison of the two African species with the Asian rice crop species *O. sativa* evidenced few SNPs and INDELs in coding regions of those genes leading to amino acids changes of *APO1, APO2, LAX1, OSH1* and *SPL14* protein sequences in the two African species in comparison to *O. sativa* (Supplemental Table 2). Finally, the recognition site of both *miR156* and *miR529* microRNAs in *SPL14* genes were compared between the two African species and *O. sativa* (supplemental Fig. S3). These sequences were identical between the three species, indicating that the differential expression pattern of *SPL14* between the two African species did not rely on alteration of the recognition sites of the microRNA regulators.

To obtain further insight into factors involved in the differential gene expression between the two African species, we carried out a comparative analysis of the genomic sequences corresponding to the 2 Kb-long regions upstream the ATG codons (considered in here after as promoter region), in order to evidence promoter sequence polymorphisms which may be related to putative *cis*-regulating domains. The result revealed some SNPs and INDELs (*i.e.* insertions, deletions) between *O. barthii* and *O.*

glaberrima LHS1 (5 SNPs and 3 INDELS), *APO1* (6 SNPs and 1 INDEL), *APO2* (2 SNPs and 1 INDEL) and *SPL14* (3 SNPs) genes (Fig. 3). Some of these polymorphic sites share similarity between *O. barthii* and *O. sativa*, others between *O. glaberrima* and *O. sativa* and finally others were polymorphic in the 3 species (Fig.3). Interestingly some of these sites were related to putative transcription factor binding sites (TFBSs) in promoter region of *LHS1* and *APO1* (Fig 3). The putative TFBSs related to polymorphic sites found in *LHS1* promoters were related to hormone responses (ABA and Jasmonate) (Busk and Pagès 1997; Simpson et al. 2003; Figueroa and Browse 2012), and XCPE1 (X core promoter element1) that drives RNA polymerase II transcription (Ohler and Wassarman 2010).

Those found in *APO1* promoter sequences were related to the regulation of seed maturation (MAT) domain *via* abscissic acid signal identified in *Arabidopsis*, the binding sites of MADS-box proteins *APETALA1 (AP1), APETALA3 (AP3)* specifying the floral organ identity and finally one related to the binding site of WUSCHEL (WUS) protein, which regulates the maintenance of stem cell populations in shoot meristem (Chern et al, 1996a; Chern t al, 1996b; Ikeda, et al. 2009; Laux et al. 1996). All these putative TFBSs derived from studies in *A. thaliana* but were not functionally evidenced in rice.

Consequently, the annotation of the coding and promoter regions of the landmark genes showed highly conserved features between the domesticated species and its wild relative. Although we could determine few SNPs and INDELs in their promoter regions, with some of them related to putative TFBSs, it couldn't be ruled out that the differential expression in terms of quantity and timing between the two species, may rely on differential trans-acting factors.

Discussion

Landmark gene expression patterns are conserved in African rice species

Domestication process in African rice (about 3000 years) was associated with several morphological and physiological changes selected by human and natural (Li et al. 2011). One of the main morphological traits selected during African rice domestication is the panicle branching complexity. The panicle structure changed from a panicle with few primary and secondary branches bearing relatively few grains in *O. barthii*, to a highly branched panicle carrying larger numbers of seeds in *O. glaberrima*. The prevailing view in Evolutionary developmental biology (Evo-Devo) is that forms evolved mostly by changes in expression patterns of functionally conserved genes rather than through the emergence of new genes (Doebley and Lukens 1998; Carroll 2008). Our comparative study of the domesticated African rice species and its wild relative provides another example in favor of this statement. This study was based on the assumption that orthologous genes in terms of speciation relationship have a conserved function in African rice species compared to *O. sativa*. This is supported

by the fact that orthologous genes between the two species are conserved in term of sequences (*i.e*. no polymorphism affecting the amino acid sequences), suggesting that at least the biochemical properties of the proteins did not diverged. It can be suspected that higher diversification occurred during Asian and African rice diversification (about 1 million years ago) rather than during *O. glaberrima* and *O. barthii* diversification in Africa (about 3000 years ago). However, we reported that the spatial expression patterns of the genes tested in here during rice panicle development were similar between *O. glaberrima* and *O. barthii*, but also in Asian rice species *O. sativa* and in other monocot species. This is illustrated by the genes promoting meristem activity such as *LAX1* and *SPL14*. The *O. sativa LAX1* gene and maize *BARREN STALK1* (*BA1*) ortholog expressed in the boundary region of initiating branch meristem and regulate the initiation of all reproductive axillary meristem (Komatsu et al. 2001; Gallavotti et al. 2004). This function seems to be conserved in other Poaceae according to the conservation of the expression pattern of the orthologous genes (Woods et al., 2011). Expression patterns of *LAX1*-like orthologous genes in Eudicots indicate that this function is partially conserved. However, differences in the temporal distribution of *LAX1*-like ortholog mRNAs in sampled eudicots suggest that these genes are required for the initiation and early maintenance of axillary meristems but not during later stages of outgrowth, as in Poaceae (Woods et al., 2011). The *O. sativa OsSPL14* and maize *TASSELSHEATH4* (*TSH4*) ortholog mRNAs accumulate at the boundary adjacent to all branch meristems (Chuck et al. 2010; Miura et al., 2010; Luo et al., 2012). The involvement of *O. sativa OsSPL14* gene in the control of panicle complexity has been evidenced through the independent characterization of *Wealthy Farmer's Panicle* (*WFP*) / *Ideal Plant Architecture* (*IPA*) QTLs related to grain yield and plant architecture (Jiao et al. 2010; Miura et al. 2010). A higher accumulation of this gene was related to a higher branching complexity of the panicle (Jiao et al. 2010; Miura et al. 2010). These QTLs are related to polymorphisms in the promoter region of this gene but also in the overlapping region of the recognition sites of the two closely related microRNAs *miR156* and *miR529*. The *miR156* microRNA was stated as the regulator of *OsSPL14* during panicle development (Jiao et al. 2010; Miura et al. 2010). The mutation of *OsSPL14* gene, reported in the two *japonica* cultivars Aikawa1 and Shaoniejing (SNJ) by Miura et al. (2010) and Jiao et al. (2010) respectively, is in fact shared by the recognition sites of the two microRNAs (see supplemental Fig. S3). *In fine*, *miR529* appears to be the main regulator of *SPL14-like* genes in both Asian and African species during the reproductive phase based on its expression level and the detected truncated *OsSPL14* mRNA in developing panicles (Jeong et al., 2011; our study). *In situ* hybridization analysis of *SPL14-like* gene and miRNA expression patterns in African rice species revealed that their expression patterns did not overlap or at least partially: *miR529* was detected in the entire but not in the flank of branch meristems where *SPL14* mRNAs were accumulated. These spatially separated expression domains would suggest a regulatory mechanism based on spatial restriction or mutual exclusion but not relying on dampening

regulation (Voinnet 2009). Similar pattern was observed for *miR156* and *SPL14*-like genes in Arabidopsis (*i.e*. *SPL9* gene) and rice during vegetative phase with an expression of *SPL14*-like gene in leaf anlagen and in developing leaf primordia in conjunction to *miR156* accumulation in both the shoot apical meristem proper and leaf primordia (Wang et al. 2008, 2009; Xie et al. 2012). This suggests a similar SPL regulatory mechanism independently of the type of microRNA and the developmental context. However, the *A. thaliana SPL9* expression pattern in *se1* and *ago1-27* mutant backgrounds has shown that *SPL9* spatial expression pattern was not affected, suggesting that *miR156* is not the main regulator of its spatial accumulation but is more to dampen overall levels of *SPL* RNA in leaf primordia (Wang et al. 2008). The *miR156*-*SPL* regulatory module is involved in the vegetative to reproductive phase transition as well as in vegetative branching in both eudicots and monocots (Poethig 2009). Interestingly, the microRNA *miR529* is found in grasses and the moss *Physcomitrella patens* but not in Eudicot species (Cuperus et al. 2011). This indicates that this microRNA emerged quite early during plant evolution but was lost in Eudicots after the Angiosperm radiation. The dual regulation of *SPL14*-like genes by *miR156* and *miR529* suggests a developmental or cellular context-dependent regulation with a specific reproductive regulation of this class of genes by *miR529* in grasses, and that *miR156* and *miR529* may contribute to complex regulation of plant architecture in crops.

APO1 and *APO2/RFL* genes are orthologs of *UNUSUAL FLORAL ORGANS* (*UFO*) and *LEAFY (LFY)* in *A. thaliana* respectively. These two genes are stated as negatively regulators of the shift to SM identity (*i.e*. determinate fate) (Ikeda-Kawakatsu et al. 2009, 2012). These two genes appear to play opposite roles to those of *UFO* and *LFY,* which promote floral fate (*i.e*. determinate fate) in *A. thaliana* and other eudicot species (Moyroud et al. 2010) Moreover, mutants of the two duplicate *LFY* homologs in maize (*i.e*. *zfl1, zfl2*) display the disruption of floret meristem, as well as the reproductive phase transition (Bomblies et al. 2003). However, based on their expression pattern and mutant phenotypes, it would be more accurate to consider *APO1* and *APO2* as promoting factors of meristematic (indeterminate) growth in grass inflorescences, through the regulation of cell proliferation at least in apical inflorescence meristem (Ikeda-Kawakatsu et al., 2009, 2012; Bomblies et al, 2003). The delay of SM specification may be considered as a consequence of meristem functioning alteration in both mutants. Similarly, *TAW1* gene belonging to the small ALOG (for *Arabidopsis LSH1* and *Oryza G1*) gene family was stated as a suppressor of SM transition. However, based on its expression pattern and dominant mutant phenotype, this gene should be considered as a meristematic (indeterminate) growth-promoting gene in inflorescence, partly through the regulation of flowering time-related *SVP*-like genes (Yoshida et al., 2012). This is in agreement with our expression patterns observed in African species, in which *TAW1* and *APO2* mRNAs accumulated in both branch meristem and spikelet meristem, with highest accumulation in early stages of panicle development.

The closest relatives of *TAW1* in *A. thaliana* are *LSH3/OBO1* and *LSH4* genes, which are expressed in shoot organ boundary cells in which they may suppress organ differentiation (Takeda et al. 2011). Consequently, the expression patterns of this class of genes differed between Eudicots and grasses, leading to divergent functions in meristems. In addition, it was shown that APO1 and APO2 proteins interact at the molecular level to control inflorescence and flower development (Ikeda-Kawakatsu et al. 2012). The temporal and spatial expression of *APO1*, *APO2* and *TAW1* are quite similar during Asian rice panicle development. However their genetic/functional relationship still remains unknown.

In contrast to the previous genes, the *O. sativa* spikelet-promoting *LHS1/OsMADS1* gene, as well as the African rice orthologs, only expressed in the spikelet meristems (Kobayashi et al. 2010; our study). The maize *LHS1* orthologs *ZMM8* and *ZMM14* genes expressed only in the upper floret, and within floral organs of certain sampled taxa, indicating that these genes are involved in the determinacy of the spikelet meristem and in the distinction of the upper from the lower florets in maize inflorescence (Cacharron et al 1999). In addition, wheat *LHS1* ortholog (*i.e*. *WLHS1*) also expressed slightly difference from the Oryza *LHS1*, the transcript of *WLHS1* accumulated at high levels in the floret organ (*i.e*. lemma, palea, pistil, glume) (Shitsukawa et al. 2007). The *LHS1*-like *SEPALLATA* (SEP) genes have been linked with the origin and diversification of the grass spikelet (Malcomber et al. 2006). The plasticity of their expression in the grasses suggests that they could be associated with differences in the structures of their respective inflorescence (Christensen and Malcomber 2012).

Differential level and timing of landmark gene expressions in African rice

A model of inflorescence (notably panicles) evolution was proposed on the basis of the difference in the time required for apical and lateral meristems to acquire floral fate (Prusinkiewicz et al. 2007). This model is supported by various mutant analysis and detailed transcriptomic time course analysis of the diversity of inflorescence architecture in Solanaceae (Park et al. 2011). These authors revealed that the program for inflorescence branching is initiated early during meristem maturation and that evolutionary diversity in inflorescence architecture is modulated by heterochronic shifts in the acquisition of floral fate. In our previous study, based notably on deep sequencing of panicle-derived small RNA transcriptomes in *O. glaberrima* and *O. barthii*, we observed a later initiation of expression in *O. glaberrima* compared to *O. barthii* of the *miR2118*-triggered 21-nt phasiRNA pathways (*i.e*. *miR2118*, *MEL1*, *lncRNAs* and *phasiRNAs*) and the spikelet-associated MADS-box gene *LHS1/OsMADS1,* both being specifically initiated in spikelet meristems (Ta et al., submitted). This finding suggested that spikelet meristem fate acquisition was set up later in *O. glaberrima* than in *O. barthii*. Moreover, our comparative expression analyses of landmark genes associated to branch meristem functioning in African rice panicle development, provided evidences that modification of expression level and/or timing of these genes during the panicle development occurred between the

two species. All these genes were over-accumulated during panicle development in the crop species, with the exception of *miR529*. This finding is certainly related to the higher numbers of established meristems in *O. glaberrima* compared to *O. barthii*. Moreover, a heterochronic shift of the spikelet-specific marker gene *LHS1*/*OsMADS1* and the microRNA *miR529* was observed with a later expression in *O. glaberrima* than in *O. barthii* (Fig 4*)*. In parallel, the branching-related genes *LAX1, APO1* and *APO2* were expressed at high level over a longer period in early panicle development in *O. glaberrima* than in *O. barthii* (Fig 4). This heterochronic shift in term of gene expression indicates a prolonged branching phase and a delayed spikelet meristem fate acquisition in *O. glaberrima,* which can be related to its higher branch complexity compared to *O. barthii* (Fig 4).

This differential expression during panicle development in the two species may be considered as a consequence of genomic evolution affecting *cis*- and/or *trans*-regulatory mechanisms. This is often the case for traits associated with dynamic processes which are more readily through regulator (*i.e*. *cis*- and/or *trans*-elements) rather than coding mutations (Wray 2007). Obviously, gene expression changes underlie many evolution changes in phenotype, thus identifying the genetic variants that regulated gene expression is significant. However in contrast to identifying variation in coding regions of the genome, characterizing the extent of *cis*-acting or/and *trans*-acting regulatory variation presents a much greater challenge. Numerous studies from other species (*i.e*. maize, wheat and tomato) identified regulator mutations with functionally significant consequences for phenotype and domestication (Doebley et al. 2006). In rice, *GS5* and *qSH1* are major QTLs controlling seed size and seed shattering respectively, two domestication-related traits. These QTLs were related to nucleotide changes in *cis*-regulatory elements leading to expression modification (Konishi et al. 2006; Li et al. 2010). The *qSH1* locus was recently analyzed in the context of African rice domestication (Wang et al. 2014). The mutation in the promoter region preventing the expression of this gene in the abscission zone in *O. sativa* was not found in *O. glaberrima* either in *O. barthii* and no difference in term of expression level was detected between the two species as well as polymorphism within the coding region (Wang et al. 2014). This finding would suggest that African rice domestication did not affect this specific gene. In contrast, another seed shattering-related gene, namely *Sh4*, shows several polymorphisms between *O. glaberrima* and *O. sativa* but the causative mutation in its coding region in Asian rice is not present in African rice (Li et al. 2006; Wang et al. 2014). But in that case, it was shown that the promoter region of *OgSh4* might have been selected during domestication, leading to a lower expression level of *OgSh4* in comparison to the one of its ortholog in *O. barthii* (Wang et al. 2014).

In order to determine to which extent *cis*-modifications may have occurred for the panicle-related genes studied in here during African rice domestication/evolution, promoter sequence comparison of these genes was carried out between the two species. Despite the low sequence divergence between

the two African rice species (Nabholz et al. 2014; Orueja et al., 2014; Wang et al. 2014), few SNPs and INDELs were evidenced in the promoter regions of *O. barthii LHS1*, *APO1*, *APO2* and *SPL14* genes in comparison to *O. glaberrima*, with some of them related to putative TFBSs. Notably, in *APO1* promoter region, SNPs were found into two TFBSs belonging to MADS box family: a TFBS related to ABA response (Vicente-Carbajosa and Carbonero 2005) and an short sequence insertion in *O. barthii* related to a WUSCHEL (WUS) protein-binding site which regulates the stem cell identity in shoot meristem and floral meristem integrity (Laux et al. 1996; Ikeda et al. 2009). Similarly, for *LHS1* gene, 8 polymorphic sites were evidenced in *O. barthii vs. O. glaberrima* (5 SNPs and 8 INDELs), with 3 of them related to putative TFBSs. However, these TFBSs were identified in *A. thaliana* and no equivalent data has been described in rice. Then, we cannot conclude to which extent the polymorphisms observed may affect the binding activity on these sites. However, these polymorphic sites could be of interest in the context of genome-wide association studies using the recently evaluated African rice collection (Orueja et al. 2014) or in bi-parental populations, in order to decipher if they could be associated to panicle trait variation. However, none of the genes studied in here co-localized with genomic regions under a selective sweep in *O. glaberrima* in comparison to *O. barthii* (Wang et al. 2014), leading to the hypothesis that these genes were not directly under human selection pressure during African rice domestication, as for Asian rice species. The global alteration of expression of these genes, independently of their function during panicle development, in *O. glaberrima* regarding its wild-relative would suggest that the expression of (a) very early acting factor(s) in panicle development might be differentially affected between the two African rice species. The early acting genes such as *APO1*, *APO2* and *TAW1* may be favored in the understanding of their differential expression between the two species. An interesting point is that in our comparative analysis of African rice species, we observed a differential timing of expression between *APO1* and *APO2* genes on one hand and *TAW1* gene on the other hand. This would suggest two distinct pathways for initial steps of panicle development at least in term of regulatory pathways.

Conclusion

In conclusion, a global modification of gene expression and not a specific gene alteration was observed during early panicle development between the two African species (Fig. 4). The expression patterns of these genes were modified at both quantitative and timing levels (i.e. heterochronic shift) but a conservation of their spatial patterns was observed. This heterochronic shift of gene expression suggests a prolonged branching phase and a delayed spikelet meristem fate acquisition in *O. glaberrima,* which can be related to its higher branch complexity compared to *O. barthii* (Fig 4). Moreover it can be speculated that the alteration of spikelet-specific genes and small RNAs such as *LHS1*/*OsMADS1* and *miR2118*-triggered 21-nt phasiRNAs (Ta et al., submitted; this study), as well as axillary meristem initiation genes such as *LAX1* and *SPL14*, may be considered as a consequence of an

alteration of reproductive meristem-promoting gene expressions such as *APO1*, *APO2* and *TAW1* from the early stages of panicle development. It will be of great importance to understand the regulatory processes related to these genes in rice species in order to understand the initial steps of panicle architecture control and its evolution.

Acknowledgments

We would to thanks F. Morcillo and S. Dussert for access to Genomatix software, M. Colin for technical advices for histological work as well as S. Chéron and H. Chrestin for plant care in greenhouse. F. Sabot and H. Pidon are also thanks for their help for their bio-informatic support. This project was jointly supported by Agropolis Foundation (through the « Investissements d'avenir » programme (ANR-10-LABX-0001-01)) and Fondazione Cariplo under the reference ID EVOREPRICE 1201-004. This work was also supported by the Institut de Recherche pour le Développement (IRD) institutional funding, Vietnamese government (program 322 to KNT) and the Global Rice Science Partnership (GRiSP) scholarship program (KNT).

Materials and Methods

Plant materials

Plants of *Oryza glaberrima* CG14 accession and *Oryza barthii* B88 accession were grown in growth chamber (14-h day/night cycle at 32°C/28°C, and humidity at 60%) at IRD, Montpellier (France). Flowering was induced by short day conditions (10-h day/night cycle). Panicles were collected at 4 different stages: stage 1, rachis and primary branch meristem; stage 2, elongated primary branch and secondary meristem; stage 3, spikelet differentiation; stage 4, young flowers with differentiates organs were used for RNA isolation and *in situ* hybridization.

Gene expression analysis

DNA samples were extracted from leave following CTAB method (*Doyle and Doyle, 1987*). Total RNAs (mRNAs and small RNAs) from different stages (stage 1 to stage 4) during rice panicle development were extracted using an RNeasy Plant Mini Kit with RLT and RWT buffers (Qiagen, France). DNAs treatments were performed using the RNAeasy-free DNase set (Qiagen). The concentrations of DNA and RNA samples were determined using a NanoDrop ND-1000 Spectro apparatus. The integrity and size distribution of total DNA were checked by agarose-gel electrophoresis.

The mRNA RT-PCRs (qRT-PCRs) analyses were perform using 1µg of total RNA in conjunction with 1µL of 50mM oligo$(dT)_{20}$ and the SuperScript III cDNA First-strand synthesis system (Invitrogen) according to manufacturer's instructions. Quantitative stem-loop RT-PCR analyses on small RNAs (*miR156, miR529*) were performed using 100 ng of total RNA according to Varkonyi-Gasic et al. (2007). A mix of total RNA in conjunction with 1µL small RNA-specific RT primers (1µM) in final volume of 9.5µL was incubated at 70°C for 5 min and on ice for 5 min to denature RNA. Then the mix was added 10.5µL of the reaction mix (4µL $MgCl_2$ (25mM), 4µL Improm-II 5X buffer, 1µL dNTP (10mM), 1µL RNAsin (40U/L) and 1 µL Improm-II RT). Pulsed RT reaction was performed following the conditions: 16°C for 30 min, (30°C 30 sec ; 42°C 30 sec ; 50°C 1sec) for 60 cycles, 70°C for 15 min.

Quantitative RT-PCRs for both mRNA- and small RNA-derived RTs were performed in an optical 384-well plate using LightCycler 480 thermocycler (Roche, France). Triplet reactions for each sample contained 4 µL SYBR Green Master Mix (Roche), 2 µL of diluted RTs and 0.8 µL of forward and reverse primers (10 µM each) in a final volume of 10 µL. The Q-PCR amplification were performed following conditions: 10 min 95°C, (15s 95°C and 30s 60°C) for 45 cyclers, 5s 95°C and 1 min 70°C. Target cDNAs and small RNAs were normalized using the rice *Actin* gene (*LOC_Os03g50885*) and mature *miR159* accumulation level, respectively. Each set of experiments

was repeated three times, and the relative quantification method with efficiency corrected calculation model (Souaze et al., 1996) was used to evaluate quantitative variation. Statistical tests were performed using t-test with two-tail test and a fixed *p*-value of 0.01. The primers used are listed in Supplementary Table S1.

In situ hybridization

In situ hybridization experiments were carried out as described by Adam et al., (2007). To obtain DNA templates for the RNA probe synthesis, PCR amplifications were performed with gene-specific antisense primers tailed with a T7 RNA polymerase binding site (see supplementary Table S1 for primer sequences). The resulting DNA fragments were used directly as templates for synthesizing antisense ribo-probes incorporating UTP–digoxigenin (Roche) as the label in conjunction with a T7 Maxi Script kit (Ambion). Detection was performed using the Vector Blue Alkaline Phosphatase Substrate Kit III (Vector Laboratories). Slides were observed and photographed by Evolution MP5.0 color Media Cybernetics camera in conjunction with a Leica DMRB microscope microscope and photographs were taken with a Q-capture pro 7 imaging system. These images processed using Photoshop CS6 (Adobe, France).

Genes sequencing and data processing

The sequence of candidate genes and their promoter from *O. glaberrima* and *O. barthii* were obtained using sequence data reference of *O.sativa* by BLASTn program of Gramene database (http://blast.gramene.org/Multi/blastview). Identification of putative transcription factor binding sites (TFBSs) in promoter regions was done using Genomatix software (http://www.genomatix.de/). To obtain complete sequence of some genes (*i.e*. *APO1, APO2* and *SPL14*) from *O. barthii*, PCR amplifications were done using total DNA from *O. barthii* B88 accession. The primers used are listed in Table S1. PCR products were in pGEM®-T Easy vector (Promega) and sequenced by Beckman Coulter Genomics (www.cogenicsonline.com). The sequences was annotated using clustalW alignment program from Genious and Mobile Pasteur facilities (http://mobyle.pasteur.fr/cgi-bin/portal.py

References

Bomblies K, Wang RL, Ambrose BA, Schmid, RJ, Meeley RB, Doebley J (2003) Duplicate *FLORICAULA/LEAFY* homologs *zfl1* and *zfl2* control inflorescence architecture and flower patterning in Maize. Development **130**: 2385–2395.

Busk K and Pagès. M (1997) Protein Binding to the Abscisic Acid-Responsive Element Is Independent of VIVIPAROUS1 in Vivo. The Plant Cell **9**: 2261–70.

Cacharron NJ, Saedler H, Theissen G (1999) Expression of MADS box genes *ZMM8* and *ZMM14* during inflorescence development of *Zea Mays* discriminates between the upper and the lower floret of each spikelet. Development Genes and Evolution **209**: 411–20.

Carroll SB (2008) Evo-Devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. Cell **134**: 25–36.

Chern MS, Bobb AJ, Busto MM (1996) The regulator of MAT2 (ROM2) protein binds to early maturation promoters and represses PvALF-activated transcription. The Plant Cell 8: 305-321.

Chern MS, Eiben HG, Bustos MM (1996) The Developmentally Regulated bZIP Factor ROM1 Modulates Transcription from Lectin and Storage Protein Genes in Bean Embryos. The Plant Journal*:* For Cell and Molecular Biology **10**: 135–48.

Christensen AR and Malcomber ST (2012) Duplication and diversification of the *LEAFY HULL STERILE1* and *Oryza sativa MADS5 SEPALLATA* lineages in graminoid Poales. EvoDevo **3**: 4.

Chuck G, Whipple C, Jackson D, Hake S (2010) The maize SBP-Box transcription factor encoded by tasselsheath4 regulates bract development and the establishment of meristem boundaries. Development **137**: 1243–50.

Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and functional diversification of MIRNA genes. The Plant Cell **23**: 431–442.

Doebley John, and Lewis Lukens (1998) Transcriptional regulators and the evolution of plant form. The Plant Cell **10**: 1075–1082.

Doebley, John F, Brandon SG, Bruce DS (2006) The molecular genetics of crop domestication. Cell **127**: 1309–21.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin **19**: 11-15.

Figueroa P and Browse J (2012) The Arabidopsis JAZ2 Promoter Contains a G-Box and Thymidine-Rich Module That Are Necessary and Sufficient for Jasmonate-Dependent Activation by MYC Transcription Factors and Repression by JAZ Proteins. Plant & Cell Physiology **53**: 330–43. doi:10.1093/pcp/pcr178.

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Pè ME, Schmidt RJ (2004) The role of *barren stalk1* in the architecture of maize. Nature **432**: 630–635.

Huang X, Qian Q, Liu Z, Sun S, He S, Luo D, Xia G, Chu C, Li J, Fu X (2009) Natural variation at the *DEP1* locus enhances grain yield in rice. Nature Genetics **41**: 494–97.

Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, Guo Y, Lu Y, Zhou C, Huang T, Zhang L, Wang Y, Feng L, Furuumi H, Kubo T, Miyabayashi T, Yuan X, Xu A, Dong G,

Zhan Q, Li C, Fujiyam A, Toyada A, Lu A, Feng Q, Aian Q, Li J, Han B (2012) A map of rice genome variation reveals the origin of cultivated rice. Nature **490**: 497–501.

Ikeda K, Sunohara H, Nagato Y (2004) Developmental course of inflorescence and spikelet in rice. Breeding Science **54**: 147–156.

Ikeda K, Ito M, Nagasawa N, Kyozuka J, Nagato Y (2007) Rice *ABERRANT PANICLE ORGANIZATION1*, encoding an F-box protein, regulates meristem fate. The Plant Journal **51**: 1030–1040.

Ikeda M, Mitsuda N, Takagi OM (2009) Arabidopsis WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. The Plant Cell **21**: 3493–3505.

Ikeda-Kawakatsu K, Yasuno N, Oikawa T, Iida S, Nagato Y, Maekawa M, Kyozuka J (2009) Expression level of *ABERRANT PANICLE ORGANIZATION1* determines rice inflorescence form through control of cell proliferation in the meristem. Plant Physiology **150**: 736–747.

Ikeda-Kawakatsu K, Maekawa M, Takeshi I, Tioh JI, 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.

Jeong DH, Park S, Zhai J, Gurazada SGR, De Paoli E, Meyers BC, Green PJ (2011) Massive analysis of rice small RNAs: Mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. The Plant Cell **23**: 4185 –4207.

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, Qian Q, Li J (2010) Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. Nature Genetics **42**: 541–544.

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.

Kobayashi K, Maekawa M, Miyao A, Hirochika H, Kyozuka 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.

Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H, Shimamoto K, Kyozuka J (2003) *LAX* and *SPA*: major regulators of shoot branching in rice. Proceedings of the National Academy of Sciences USA **100**: 11765–11770.

Komatsu M, Maekawa M, Shimamoto K, 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.

Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyozuka 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.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science **312**: 1392–1396.

Laux T, Mayer KF, Berger J, Jürgens G (1996) The *WUSCHEL* gene is required for shoot and floral meristem integrity in Arabidopsis. Development **122**: 87–96.

Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J, et al (2011) Natural variation in GS5 plays an important role in regulating grain size and yield in rice. Nature Genetics **43**: 1266–1269.

Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. Science **311**: 1936–1939.

Li ZM, Zheng XM, Ge S (2011) Genetic diversity and domestication history of African rice (*Oryza glaberrima*) as inferred from multiple gene sequences. Theoretical and Applied Genetics **123**: 21–31.

Luo L, Li W, Miura K, Ashikari M, Kyozuka J (2012) Control of tiller growth of rice by *OsSPL14* and strigolactones, which work in two independent pathways. Plant & Cell Physiology **53**: 1793–1801.

Malcomber ST, Preston JC, Reinheimer R, Kossuth J, Kellogg EA (2006) Developmental gene evolution and the origin of grass inflorescence diversity. Advances in Botanical Research **44**: 425–481.

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.

Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F (2010) LEAFY blossoms. Trends in Plant Science **15**: 346–352.

Nabholz B, Sarah G, Sabot F, Ruiz M, Adam H, Nidelet S, Ghesquière A, Santoni S, David J, Glémin S (2014) Transcriptome population genomics reveals severe bottleneck and domestication cost in the African Rice (*Oryza glaberrima*). Molecular Ecology **23**: 2210–2227.

Ohler U, Wassarman DA (2010) Promoting developmental transcription. Development **137**: 15–26.

Orjuela J, Sabot F, Chéron S, Vigouroux Y, Adam H, Chrestin H, Sanni K, Lorieux M, Ghesquière A (2014) An extensive analysis of the African rice genetic diversity through a global genotyping. Theoretical and Applied Genetics, doi 10.1007/s00122-014-2374-z.

Park SJ, Jiang K, Schatz MC, Lippman ZB (2011) Rate of meristem maturation determines inflorescence architecture in tomato. Proceedings of the National Academy of Sciences USA **109**: 639-644.

Poethig RS (2009) Small RNAs and developmental timing in plants. Current Opinion in Genetic & Development **19**: 374–378.

Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E (2007) Evolution and development of inflorescence architectures. Science **316**: 1452–1456.

Rao NN, Prasad K, Kumar PR, Vijayraghavan U (2008) Distinct regulatory role for *RFL*, the rice LFY homolog, in determining flowering time and plant architecture. Proceedings of the National Academy of Sciences USA **105**: 3646–3651.

Shitsukawa N, Tahira C, Kassai KI, Hirabayashi C, Shimizu T, Takumi S, Mochida K, Kawaura K, Ogihara Y, Murai K (2007) Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. The Plant Cell **19**: 1723–1737.

Souaze F, Ntodou-Thome A, Tran CY, Rostene W, Forgez P (1996) Quantitative RT-PCR: limits and accuracy. Biotechniques **21**: 280-285.

Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K, Shinozaki KY (2003) Two Different Novel Cis-Acting Elements of erd1, a clpA Homologous Arabidopsis Gene Function in Induction by Dehydration Stress and Dark-Induced Senescence. The Plant Journal: For Cell and Molecular Biology (**33**): 259–70.

Sweeney M, McCouch S (2007) The complex history of the domestication of rice. Annals of Botany **100**: 951–957.

Ta KN, Sabot F, Adam H, Vigouroux Y, De Mita S, Ghesquière A, Do, NV, Gantet P, Jouannic S Time-shift of panicle meristem states in African rice species. **Submitted.**

Takeda S, Hanano K, Kariya A, Shimizu S, Zhao L, Matsui M, Tasaka M, Aida M (2011) *CUP-SHAPED COTYLEDON1* transcription factor activates the expression of *LSH4* and *LSH3*, two members of the ALOG gene family, in shoot organ boundary cells. The Plant Journal **66**: 1066–1077.

Terao T, Nagata K, Morino K, Hirose T (2009) A gene controlling the number of primary rachis branches also controls the vascular bundle formation and hence is responsible to increase the harvest index and grain yield in rice. Theoric and Applied Genetics **120**: 875–893.

Tsuda K, Ito Y, Yato Y, Kurata N (2011) Positive autoregulation of a *KNOX* gene is essential for shoot apical meristem maintenance in rice. The Plant Cell **23**: 4368–4381.

Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. Plant Methods **3**: 12.

Vaughan DA, Lu BR, Tomooka N (2008) The evolving story of rice evolution. Plant Science **174**: 394–408.

Vicente-Carbajosa J, Carbonero P (2005) Seed maturation: developing an intrusive phase to accomplish a quiescent state. International Journal of Developmental Biology **49**: 645-651.

Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell **136**: 669–687.

Wang M, Yu Y, Haberer G, Marri PR, Fan C, Goicoechea JL, Zuccolo A, Song X, Kudrna D, Ammiraju JSS, Cossu RM, Maldonado C, Chen J, Lee S, Sisneros N, de Baynast K, Golser W, Wissotski M, Kim W, Sanchez P, Ndjiondjop MN, Sanni K, Long M, Carney J, Panaud O, Wicker T, Machado CA, Chen M, Mayer KFX, Rounsley S, Wing RA (2014) The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. Nature Genetics **46**: 982-988.

Wang Y, Li J (2011) Branching in rice. Current Opinion in Plant Biology **14**: 94–99.

Wang J-W, Schwab R, Czech B, Mica E, Weigel D (2008) Dual effects of *miR156*-targeted *SPL* genes and CYP78A5/KLUH on plastochron length and organ size in *Arabidopsis thaliana*. The Plant Cell **20**: 1231–1243.

Wang J-W, Czech B, Weigel D (2009) *miR156*-regulated *SPL* transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. Cell **138**: 738–749.

Woods DP, Hope CL, Malcomber ST (2011) Phylogenomic analyses of the *BARREN STALK1/LAX PANICLE1 (BA1/LAX1)* genes and evidence for their roles during axillary meristem development. Molecular Biology and Evolution **28**: 2147 –2159.

Wray GA (2007) The evolutionary significance of *cis*-regulatory mutations. Nature Reviews Genetics **8**: 206–216.

Xie K, Shen J, Hou X, Yao J, Li X, Xiao J, Xiong L (2012) Gradual increase of *miR156* regulates temporal expression changes of numerous genes during leaf development in rice. Plant Physiology **158**: 1382–1394.

Xing Y and Zhang Q (2010) Genetic and molecular bases of rice yield. *A*nnual Review of Plant Biology **61**: 421–42.

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, Chen R, Yamazaki R, Tokunaga H, Kitaguchi Y, Sato Y, Nagamura Y, Ushijima T, Kumamaru T, Iida S, Maekawa M, Kyozuka J (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.

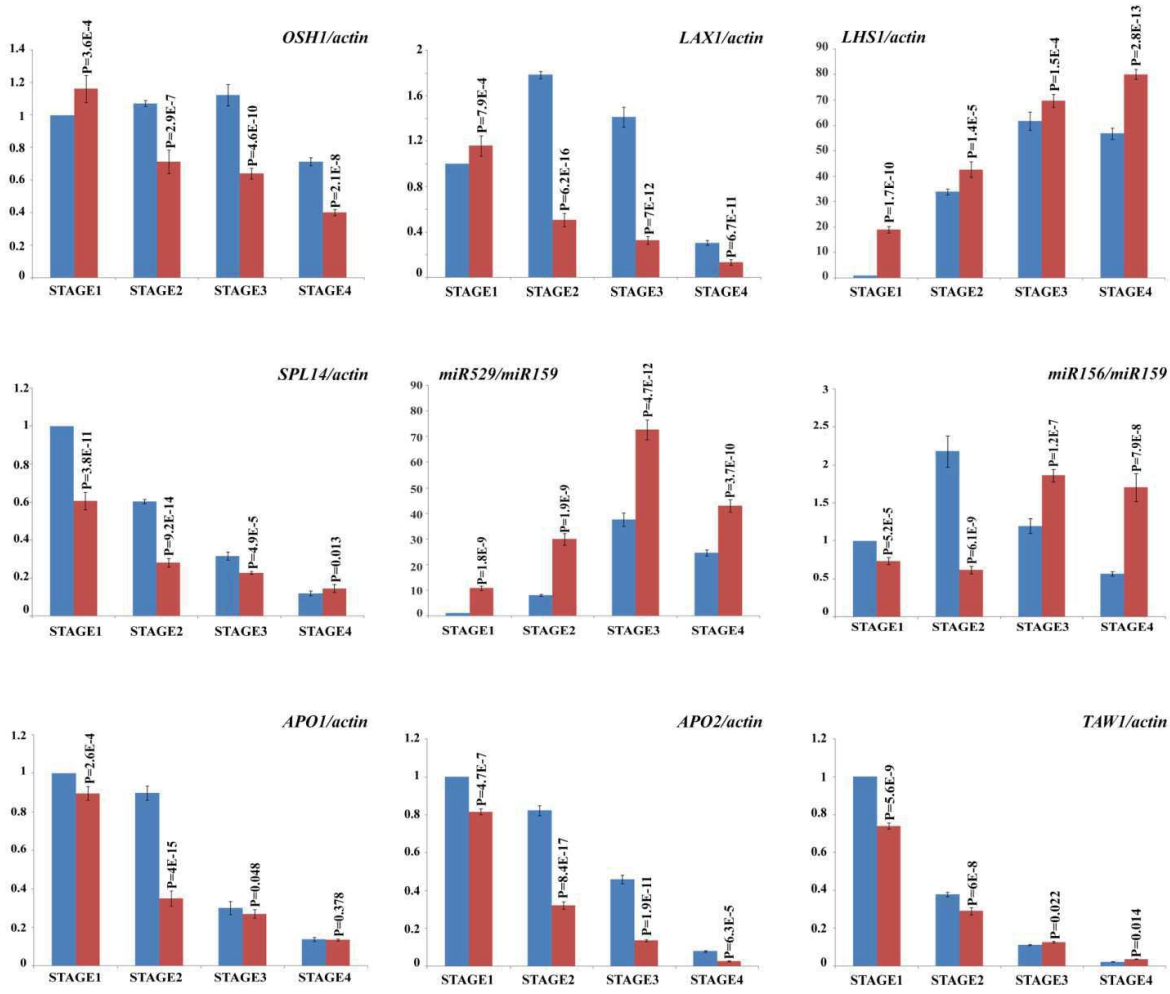


Figure 1. Expression analysis of landmark genes during panicle development in African rices. qRT-PCR analysis of *OSH1*, *LAX1, LHS1*, *SPL14*, *miR529, miR156, APO1, APO2* and *TAW1* accumulation levels during panicle development (from stage 1 to stage 4) in *O. glaberrima* (blue) and in *O. barthii* (red). Target cDNAs and small RNAs accumulation levels were normalized using the rice *Actin* gene (*LOC_Os03g50885*) and mature *miR159* microRNA accumulation level, respectively. Relative expression level ratios are normalized using *O. glaberrima* stage 1 ratio values as calibrator (y-axis). The p-values from t-test analyses (*O. glaberrima vs. O barthii* per stage) are indicated.

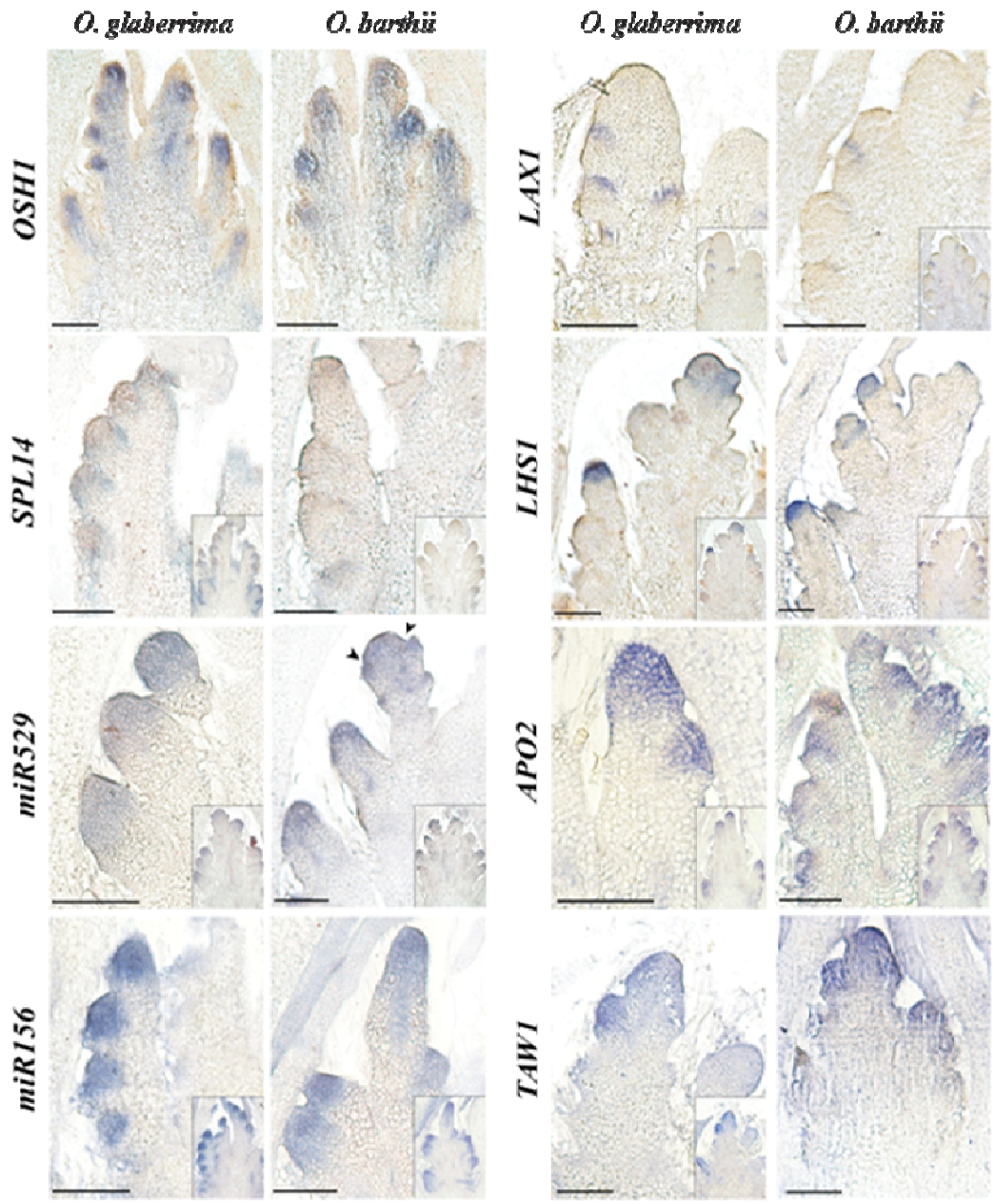


Figure 2. *In situ* expression patterns of panicle-related landmark genes. *In situ* hybridization analysis of *OSH1*, *LAX1, SPL14*, *miR529*, *miR156*, *LHS1, APO2* and *TAW1* gene expression in branch meristems (stages 2-3) in *O. glaberrima* and *O. barthii. OSH1* transcripts were detected in the inner of new branch meristems (BMs) while *LAX1* expresses in boundary regions of new BMs. The microRNAs *miR156* and *miR529* and their putative *SPL14* mRNA target are accumulated in different domains in branch meristems: *SPL14* mRNA accumulate in the boundary region while *miR156* and *miR529* were detected in the entire new branch meristems. *LHS1* mRNAs accumulate specially in SMs while *APO2* and *TAW1* mRNAs accumulate in BMs. The inset photos show the complete panicle section corresponding to the close views of the branch. Arrowheads, lemma and palea primordium formation in SMs; Scale bars: 100μm.

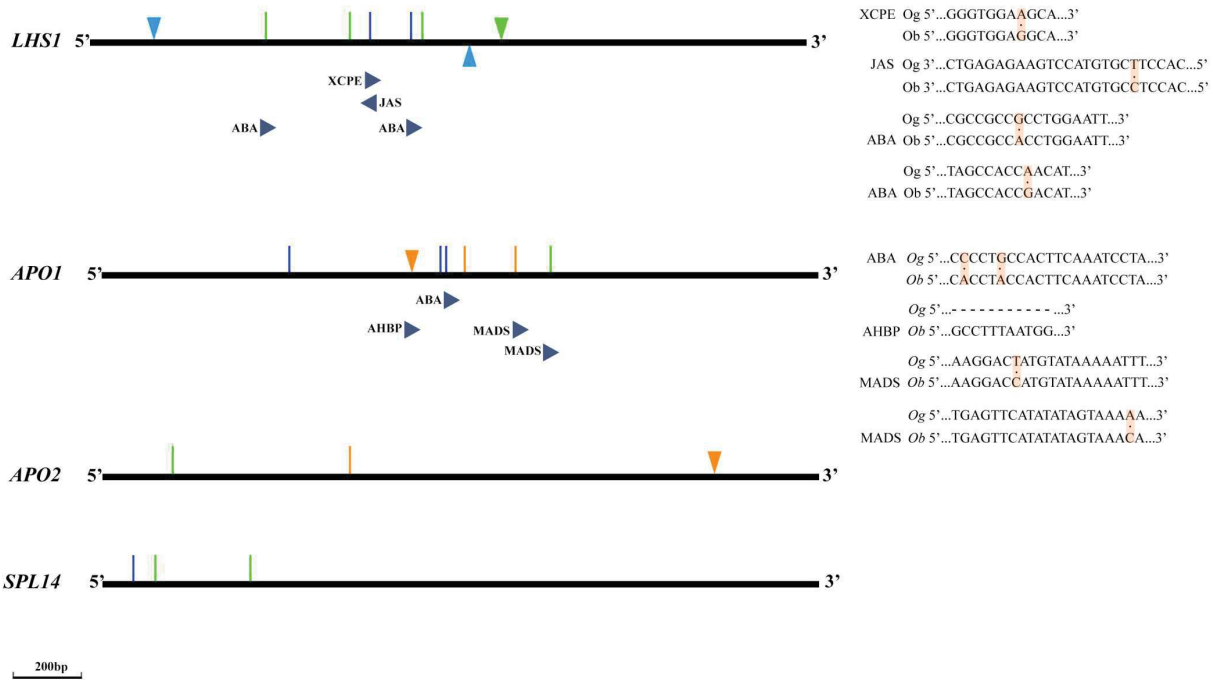


Figure 3. Promoter sequence comparisons of *LHS1, APO1, APO2* and *SPL14* orthologous genes in *O. glaberrima* and *O. barthii.* INDELs were defined according to *O. glaberrima* cv CG14 sequence. INDELs are represented by triangles, up side for insertion and down side for deletion in *O. barthii*. Vertical bars represent SNPs. Green color indicates polymorphic sites in *O. barthii* vs. *O. glaberrima* but identical to *O. sativa.* Blue color indicates polymorphic sites in *O. barthii* vs. *O. glaberrima* and *O. sativa*. Orange color indicates polymorphic sites for the 3 species. Putative transcription factor binding sites related to INDELs or SNPs are indicated by arrowheads (indicating site orientation). The sequence alignments of these sites between *O. barthii* (Ob) and *O. glaberrima* (Og) are indicated on the right and the corresponding polymorphic site is highlighted.

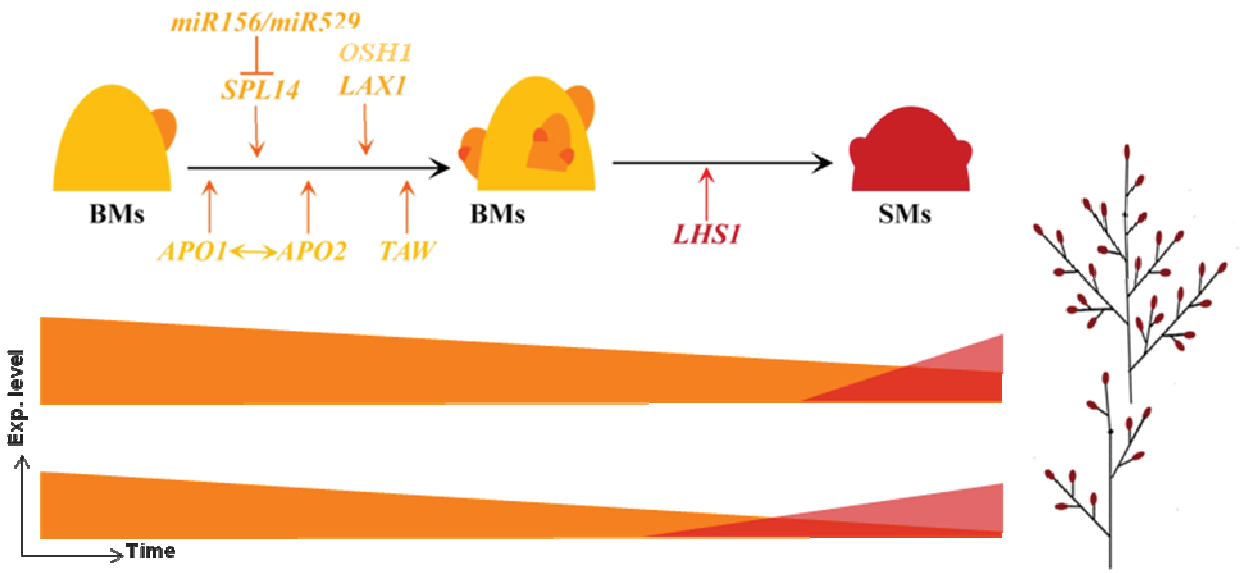
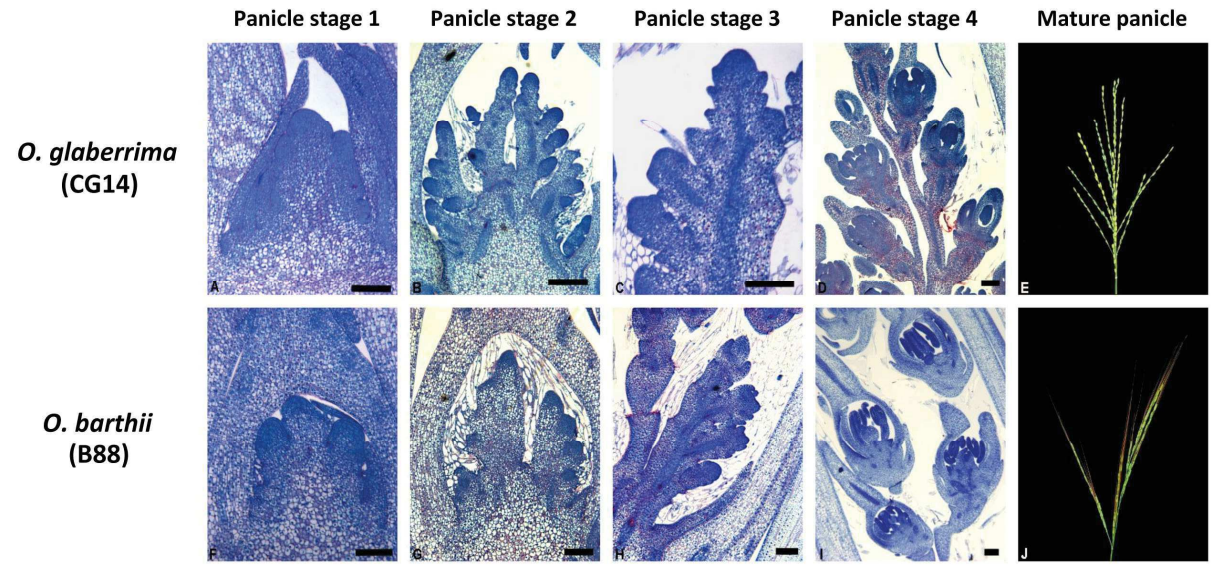


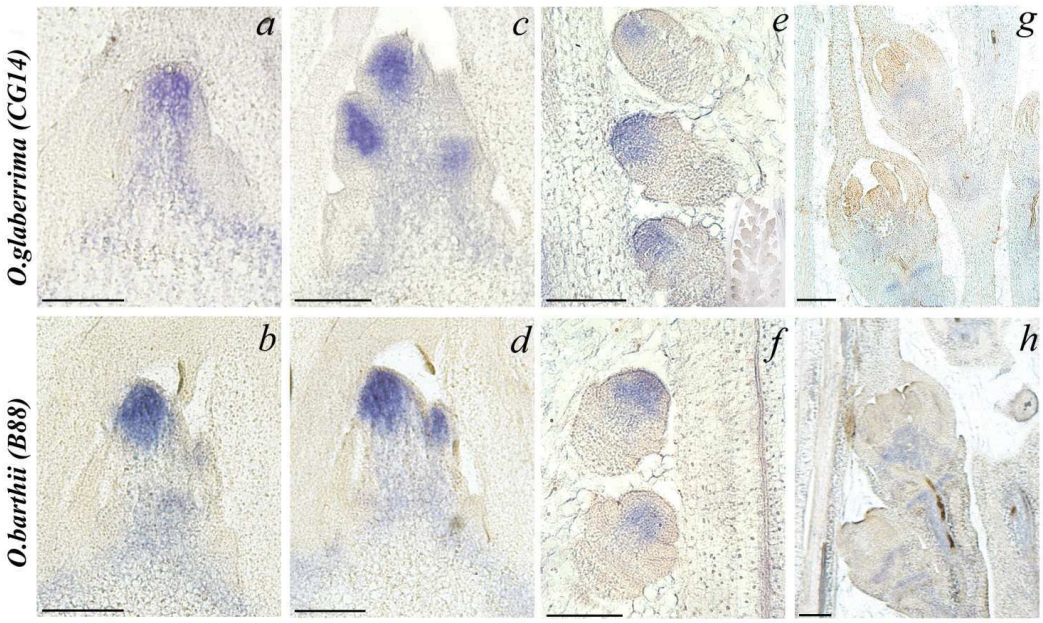
Figure 4. Model of variation in complexity of African rice panicle architecture. In the upper panel, the two main phases of panicle development are indicated: branching activity (primary and other order branching) and spikelet/floret meristem development. Yellow, branch meristem (BMs) - indeterminate phase; Red, spikelet/floret meristem (SMs/FMs) - determinate phase. The function of the selected landmark genes are indicated (i.e. branching activity and BM to FM transition) as well as their promoting or repressing activity (arrows and flat head respectively). The lower panel illustrates that the variation between *O. glaberrima* and *O. barthii* panicle architecture complexity may be related to a differential expression timing and level of genes controlling branching process and transition from indeterminate to determinate fate during the rice panicle development, with a delayed expression of flower fate-related genes in *O. glaberrima* consequently to a higher/longer branching activity.

Supplemental data

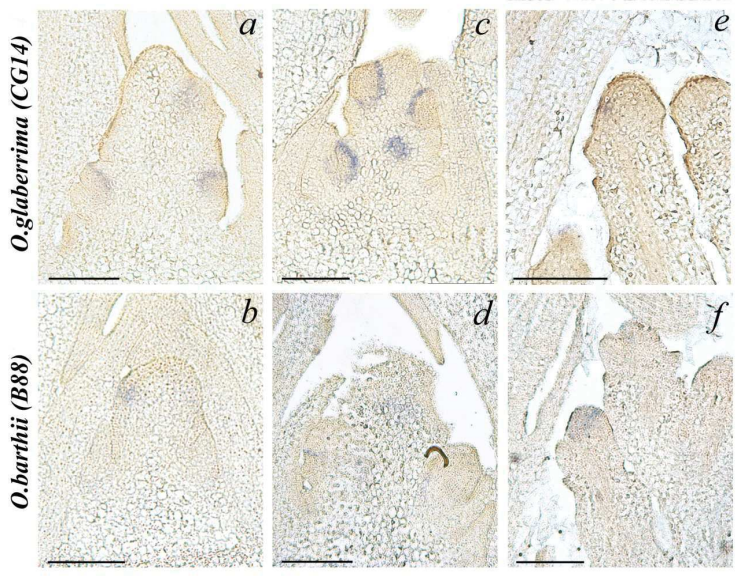


Supplemental Figure S1. Histological description of selected developmental stages of African rice panicles used for qRT-PCR analyses. *O. glaberrima*: A-E; *O. barthii*: F-J; stage 1: A and F; stage 2: B and G; stage 3: C and H; stage 4: D and I; mature stage: E and J. Scale bars: 100 µm

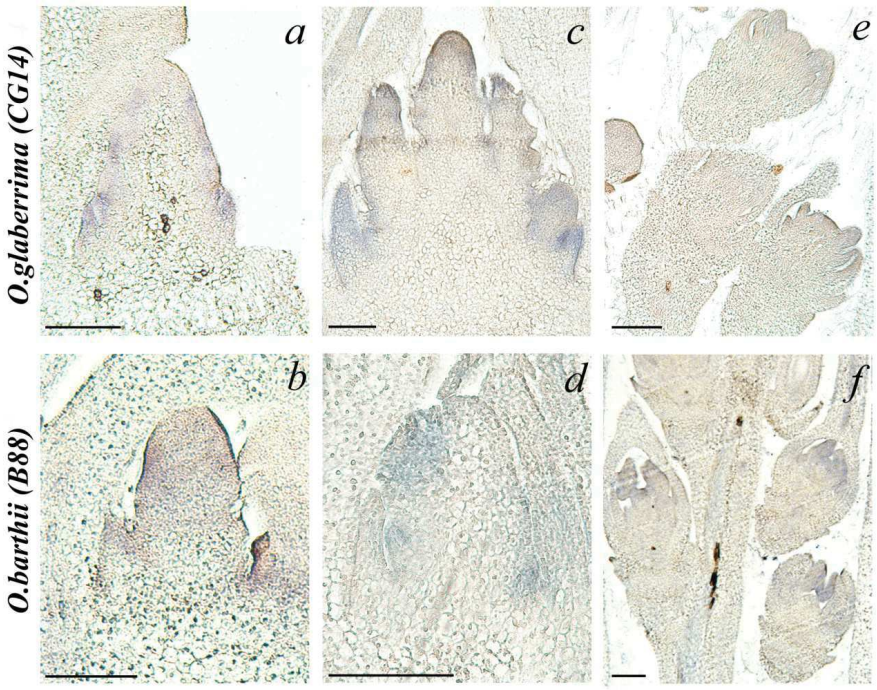
Supplemental Figure S2. *In situ* hybridization analysis of *OSH1, LAX1, SPL14, miR529, miR156, LHS1, TAW1* and *APO2* gene expression during panicle development of *O. glaberrima* and *O. barthii*. This figure illustrates the *in situ* hybridization data during panicle development not presented in the Fig. 2 of the main text.



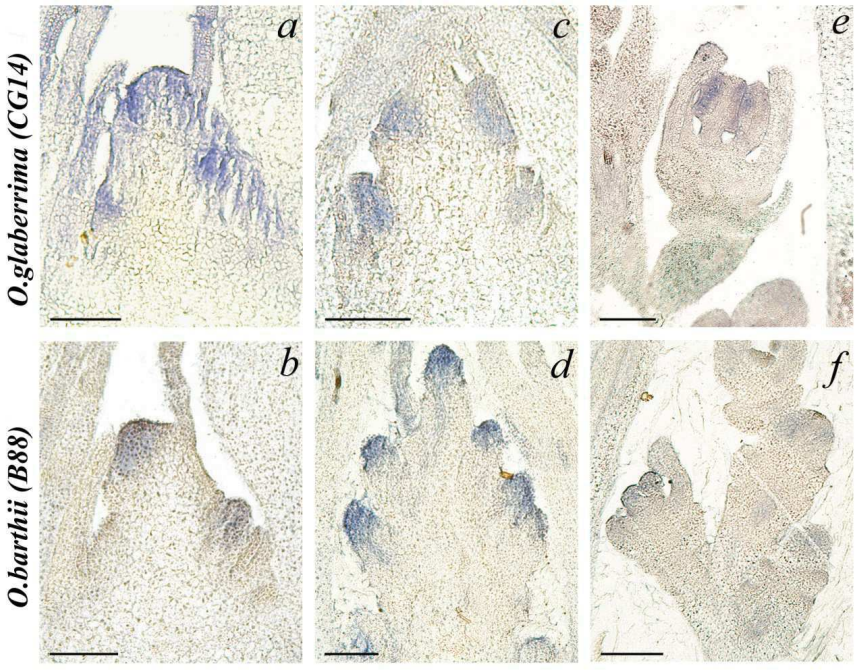
OSH1. In situ mRNA patterns of *OSH1* gene at stages 1, 2 and 4 in *O. glaberrima* cv CG14 (a, c, e, g) and *O. barthii* var. B88 (b, d, f, h) respectively . Stage 1: a, b; stage 2: c, d, e, f; stage 4: g, h. Scale bars: 100µm.



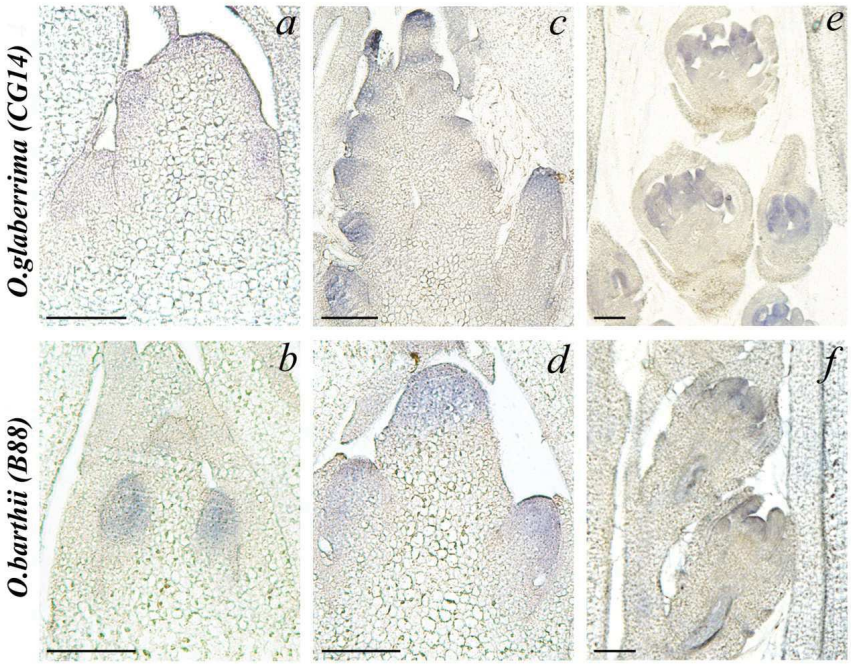
LAX1. In situ mRNA patterns of *LAX1* gene at stages 1, 2 and 3 in *O. glaberrima* cv CG14 (a, c, e) and *O. barthii* var. B88 (b, d, f) respectively. Stage 1: a, b; stage 2: c, d; stage 3: e, f. Scale bars: 100µm.



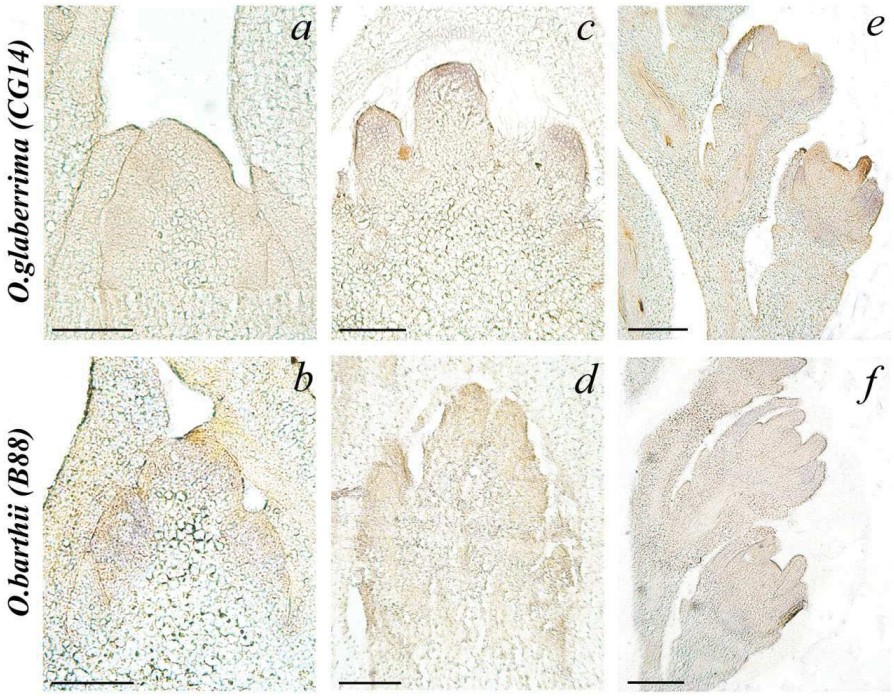
SPL14. In situ mRNA patterns of *SPL14* gene at stages 1, 2 and 4 in *O. glaberrima* cv CG14 (a, c, e) and *O. barthii* var. B88 (b, d, f) respectively. Stage 1: a, b; stage 2: c, d; stage 4: e, f. Scale bars: 100µm.



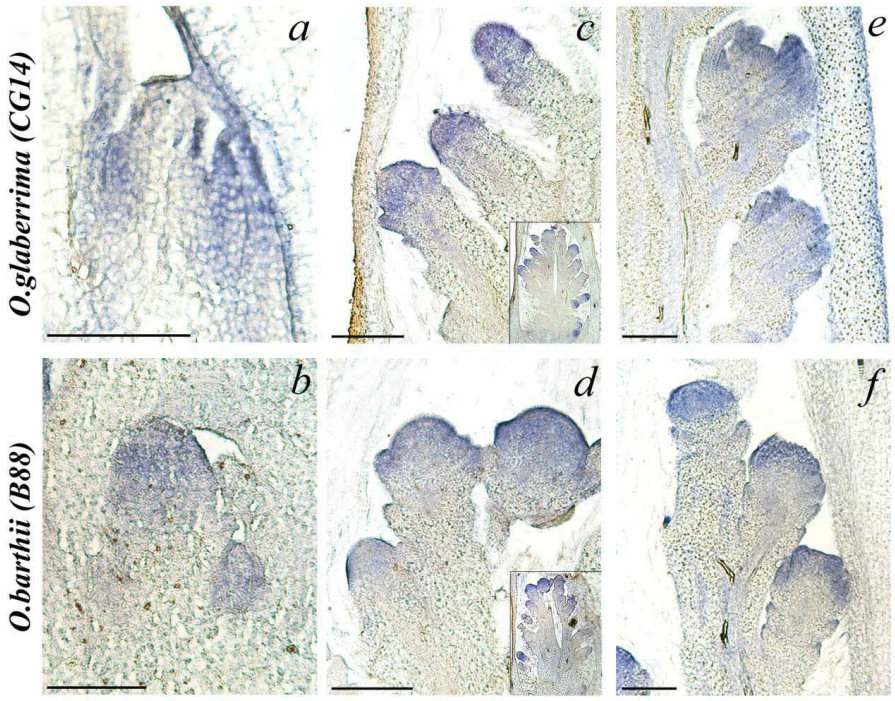
miR156. In situ mRNA patterns of *miR156* microRNA at stages 1, 2 and 4 in *O. glaberrima* cv CG14 (a, c, e) and *O. barthii* var. B88 (b, d, f) respectively. Stage 1: a, b; stage 2: c, d; stage 4: e, f. Scale bars: 100µm.



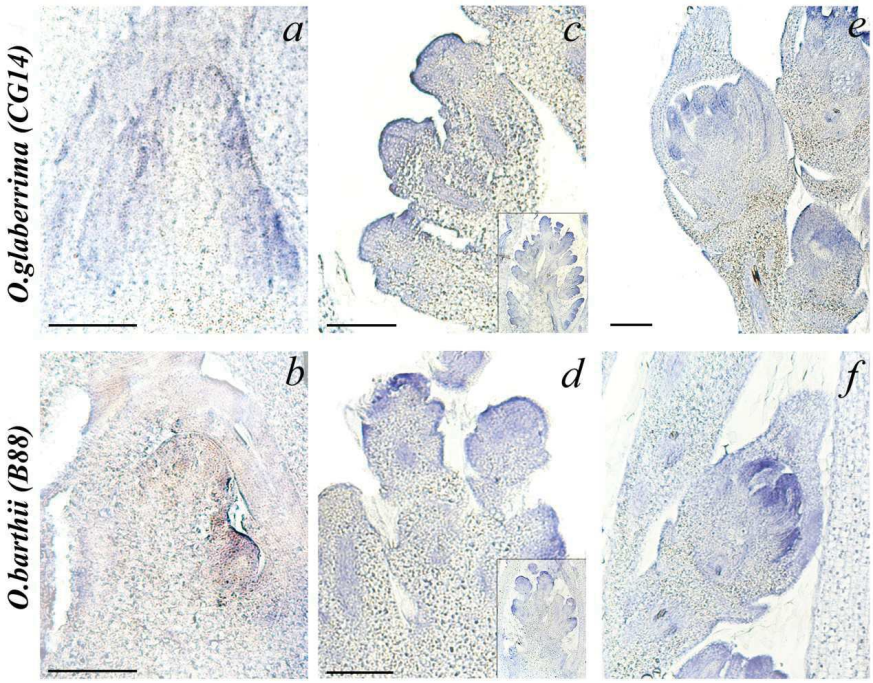
miR529. In situ mRNA patterns of *miR529* microRNA at stages 1, 2 and 4 in *O. glaberrima* cv CG14 (a, c, e) and *O. barthii* var. B88 (b, d, f) respectively. Stage 1: a, b; stage 2: c, d; stage 4: e, f. Scale bars: 100µm.



LHS1. In situ mRNA patterns of *LHS1* gene at stages 1, 2 and 4 in *O. glaberrima* cv CG14 (a, c, e) and *O. barthii* var. B88 (b, d, f) respectively. Stage 1: a, b; stage 2: c, d; stage 4: e, f. Scale bars: 100µm.



APO2. In situ mRNA patterns of *APO2* gene at stages 1, 3 and 4 in in *O. glaberrima* cv CG14 (a, c, e) and *O. barthii* var. B88 (b, d, f) respectively. Stage 1: a, b; stage 3: c, d; stage 4: e, f. Scale bars: 100µm.



TAW1. In situ mRNA patterns of *TAW1* gene at stages 1, 3 and 4 in in *O. glaberrima* cv CG14 (a, c, e) and *O. barthii* var. B88 (b, d, f) respectively. Stage 1: a, b; stage 3: c, d; stage 4: e, f. Scale bars: 100µm.

Supplemental Figure S3. *miR156* and *miR529* recognition sites in *SPL14* mRNA target in *O. sativa* (*OsSPL14*), *O. glaberrima* (*OgSPL14*) and *O. barthii* (*ObSPL14*). The single nucleotide change in *O. sativa* from C to A at the *Osa-miR156* targeted site in *OsSPL14* as reported in the *japonica* cultivars Aikawa1 and Shaoniejing (SNJ) by Miura et al. (2010) and Jiao et al. (2010) respectively is highlighted in red. The numbers above the sequence indicate the location of the nucleotide in the *OsSPL14* coding sequence. Dots indicate identical nucleotide sequences in the region corresponding to the recognition sites of *Osa-miR156* and *Osa-miR529*.

Supplemental Table 1: List of primers used in this study.

The underlined sequences correspond to the T7 promoter used for RNA probe synthesis. The bold sequences in the sequence of stem-loop RT primers correspond to the stem-loop part.

Primer name	Sequence
RT-PCR	
ACT-F	CATTCCAGCAGATGTGGATTG
ACT-R	TCTTGGCTTAGCATTCTTGG
APO1-F	GTTCTACTGCATGAGCTCGTC
APO1-R	TGCACCTTGCTCCATACGTTC
APO2-F	AGGTGCAATCCATGGCTAAG
APO2-R	GCATCTTGGGCTTGTTGATG
LHS1-F	GTGACCATTCCCTGCAGATT
LHS1-R	GTCTGCTGCTTCATTGCTCA
SPL14-F	CTGCCTGAATTTGACCAAGG
SPL14-R	AAGCTTCTGAACCTGCGATG
OsMADS3-F	TGAGGAGCAAGGTTGTTGAG
OsMADS3-R	AGGCTGCTGCATGATGTTC
OsMADS58-F	AAACGGAGCTGCAGAATGAC
OsMADS58-R	GCTGAACCCATCATGTTCAC
TAW-F	CTAGTTACTCCACTCCACTC
TAW-R	GTAGTTTTGCTAGTAGCAAG
Stem-loop RT-PCR	
miR159b-RT	**GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGA**CCAGAGC
miR159-F	CGGCGGTTTGGATTGAAGGGA
Univ-RT	GTGCAGGGTCCGAGGT
miR156-RT	**GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGA**CGTGCTC
miR156-F	GCGGCGGTGACAGAAGAGAGT
miR529-RT	**GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGA**CAGGCTG
miR529-F	CGGCGCAGAAGAGAGAGAGTA
***in situ* hybridization**	
OsH1_F	CAGTTCGTGATGATGGAC
OsH1_R	CTAAAACCGACCCCTGCATTA
OsH1_T7_F	GCGAAATTAATACGACTCACTATAGGGCGAACAGTTCGTGATGATGGAC
OsH1_T7_F	GCGAAATTAATACGACTCACTATAGGGCGAACTAAAACCGACCCCTGCATTA
OsSPL14_HIS_F	AGTGGCACAGGAACGTAGCTCCT
OsSPL14_HIS_R	GCACAGCTCGAGTCGGTGGCGGCAC
OsSPL14_T7_HIS_F	GCGAAATTAATACGACTCACTATAGGGCGAAAGTGGCACAGGAACGTAGCTCCT

OsSPL14_T7_HIS_R	GCGAAATTAATACGACTCACTATAGGGCGAAGCACAGCTCGAGTCGGTGGCGGCAC
APO2_F	ATCTCGGAGCTCGGGTTCACG
APO2_F1	GCCGACCGCAAGGACAGCAA
APO2_R	CGCAAACATGGGTACACGACG
APO2_T7F1	GCGAAATTAATACGACTCACTATAGGGCGAA GCCGACCGCAAGGACAGCAA
APO2_T7R	GCGAAATTAATACGACTCACTATAGGGCGAA CGCAAACATGGGTACACGACG
LHS1-HIS-F	GAAGAGCAAGGAGCAACAGC
LHS1-HIS-R	AATCTGCAGGGAATGGTCAC
LHS1-HIS-T7F	GCGAAATTAATACGACTCACTATAGGGCGAAGAAGAGCAAGGAGCAACAGC
LHS1-HIS-T7R	GCGAAATTAATACGACTCACTATAGGGCGAAAATCTGCAGGGAATGGTCAC
TAW1_insitu_F	GCGTCAGCTACGAGAAGAAG
TAW1_insitu_R	GTAGTTTTGCTAGTAGCAAG
TAW1_insitu_T7+F	GCGAAATTAATACGACTCACTATAGGGCGAAGCGTCAGCTACGAGAAGAAG
TAW1_insitu_T7+R	GCGAAATTAATACGACTCACTATAGGGCGAAGTAGTTTTGCTAGTAGCAAG
T7_insitu	GCGAAATTAATACGACTCACTATAGGGCGAA
LNA-miR156 (Exiqon)	GTGCTCACTCTCTTCTGTCA
LNA-miR529 (Exiqon)	AGGC[G]TAC[C]TCTC[C]T[C]T

Sequencing

APO1_SEQ_OB_F1	GGTTCTGGTGTTCGACGTGGC
APO1_SEQ_OB_R1	CCGGGGCTCGTACGCGAACAC
APO1_SEQ_OB_R	ACCGGATAACGGGGTAGAAG
APO2_SEQ_OB_F2	ATCCCAACGATGCCTTCTCGG
APO2_SEQ_F1	CGCAAGGACAGCAAGCTAGTA
APO2_SEQ_R1	GCTCCCCGCCATGTCATGCTC
SPL14_SEQ_OB_F1	GGAGAGAAAGGAGGCTCGTCGG
SPL14_SEQ_OB_F2	CACTGTGGGTGCAGTGTCTT
SPL14_SEQ_OB_R1	AGCAAAGCAAAAGCAGTGGT

Supplemental Table S2. Summary of annotated polymorphisms in nucleic coding and protein sequences in *O. sativa* in comparison to *O. glaberrima* and *O. barthii.* CDS: gene coding sequence. The position and the type of non-synonymous mutations in the amino acid sequences are indicated. The polymorphic changes are indicated in *O. sativa* in comparison to *O. glaberrima*.

Gene	CDS	Protein	Note
OSH1	6	5	(aa22-23: del HQ), (aa38: T➔A), (aa48: T➔L), (aa94-95: del SA), (aa58: G➔A)
LAX1	2	1	(aa125: A➔S)
APO1	12	3	(aa17: V➔I), (aa260: del A), (aa372: del AG)
APO2	7	1	(aa139: E➔D)
LHS1	0	0	
SPL14	7	4	(aa95: del E), (aa237: G➔R), (aa239: L➔D), (aa317: T➔A)

2.2.4 Comments on genome expression analysis

The inflorescence architecture in the nature could be divided into three main classes (i.e. raceme, cyme and panicle), which diverged during the evolution in distinct plant families. However, the mechanisms that control inflorescence architecture are still largely unknown. In 2007, Prusinkiewicz et al proposed a single model of inflorescence evolution based on variations of a parameter called *vegetativeness* (*veg*) – a measure of the "state" of a meristem. This meristematic "state" can be resumed by its ability to establish new branching structure (i.e. axillary meristem) and to change to a determinate meristem. The indeterminate meristems (apical meristem and/or lateral meristem(s)) convert to determinate meristem (flower/floret meristem), which are no longer able to establish new lateral meristem. Thus, this process will have a big impact on the inflorescence type according to the position/number of the meristems contributing to flowers and the timing of their determinate acquisition. The modeling was improved by several mutants affected in the floral meristem identity in different species, notably in *Arabidopsis thaliana*, *Antirrhinum majus*, petunia and tomato (Koes, 2008; Moyroud et al., 2010; Park et al., 2014), indicating that very early acting genes such as *LFY*, *UFO* and *TFL1* from *A. thaliana* and their orthologs in other species may have contributed significantly to the evolution of inflorescence structure.

In the context of the evolution of rice panicle architecture, no study reports a comparative analysis of inter-specific diversity at genome expression level especially for the small RNA population, which is known to have a big impact on plant development. Several genes were reported as involved in the initiation and functioning of reproductive meristems (including lateral meristem, spikelet meristem and floret meristem) during rice panicle development. However none of them were studied in the context of panicle evolution. In the frame of my PhD work, two approaches were developed regarding comparative genome expression analysis in the two African rice species, *O. glaberrima* and *O. barthii*. The first one was a genome-wide analysis of the small RNA population expressed during panicle development. Only few microRNAs were affected in term of accumulation level in this transcriptomic analysis, including the 22-nt *miR2118* known to trigger the panicle-specific 21-nt phasiRNA pathway (Song et al. 2012a; Komiya et al. 2014). In parallel, we observed a significant difference on timing of expression of this class of siRNAs during panicle development in *O. glaberrima* vs. *O. barthii*. It was shown that this phasiRNA pathway is related to male gametogenesis and constitutes *in fine* a marker of determinate fate of the inflorescence meristems in rice panicle. This work suggested that the on-set of spikelet/floret differentiation was earlier in *O. barthii* than in *O. glaberrima,* as supported by the expression pattern of *LHS1* gene, a marker of floret differentiation. To complete this work, a second approach was developed through the expression analysis of a set of orthologs of *O. sativa* landmark genes related to panicle development. This set of 9

genes was related to the formation/maintenance of axillary meristems on one hand and to spikelet/floret meristems differentiation on the other hand. Although spatial expression of those genes was highly conserved between domesticate species and its wild-relative, their expression pattern were modified at both quantitative and timing levels (i.e. heterochronic shift). While the genes related to branching activity or meristem functioning (i.e. *OSH1, LAX1, SPL14, APO1* and *APO2*) were upper-accumulated and sustained on a longer period during the panicle development in the crop species, the gene promoting spikelet/floret meristem fate activity (i.e. *LHS1*) behaved in opposite way. Both quantitative and heterochronic shift of gene expression suggests a prolonged branching phase and a delayed spikelet meristem fate acquisition in *O. glaberrima* in regard to its wild-relative *O. barthii* (Figure 2.1). This finding is supported by the histological analysis of early stages of panicle development in the two species, suggesting that spikelet meristem differentiation occurs later in *O. glaberrima* than in *O. barthii*.

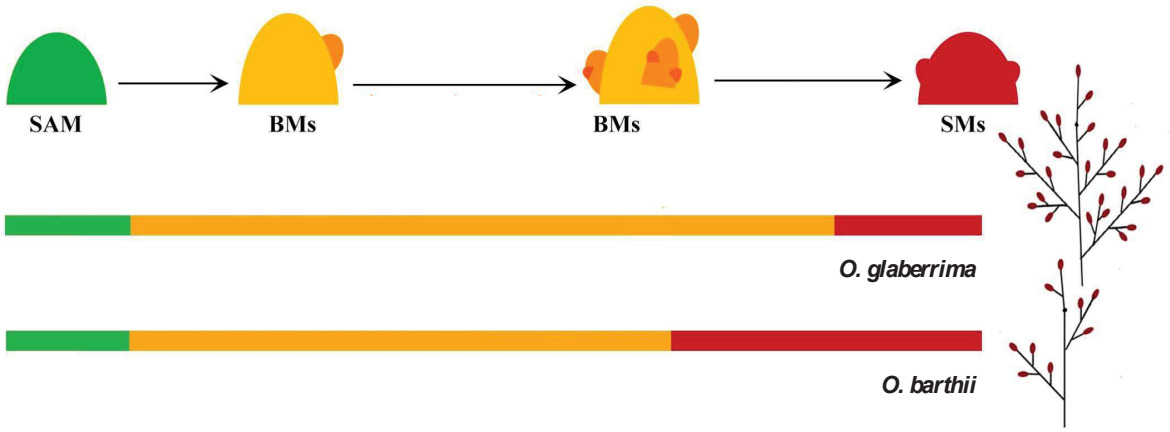


Figure 2.1: Model of differential timing of determinate meristematic state acquisition between the two African species *O. glaberrima* and *O. barthii*. In the upper panel are illustrated the two main phases of panicle development (orange: branching phase; red: spikelet/floret phase). The vegetative phase is illustrated in green. In the lower panel are represented the different phases in term of duration until complete differentiation of the panicle for *O. glaberrima* and *O. barthii*. The differential panicle complexity between the two species is illustrated on the right.

All together, these findings indicate that the variation of rice panicle architecture in the two African species is not only related to the activity of meristem fate controlling genes, but also branch-promoting genes. It should be considered that the delay of acquisition of determinate fate in *O. glaberrima* might be rather a consequence of the higher and longer branching activity in this species. Moreover, considering the differential expression of these genes during panicle development between the two species, raise a question that whether the existence of some mutations related to *cis*- and/or *trans*-regulatory factors may affect the expression of these genes. Although characterizing these regulators presents a much more challenge in comparison to the identification of variations in coding

region of genome, several studies from rice and other species (*i.e*. maize, wheat and tomato) identified regulator mutations with functionally significant consequences for phenotype and domestication (Doebley et al. 2006; Tang et al. 2010). However, in contrast to the variation of expression level and timing of the panicle-related genes studied here, their coding sequences were highly conserved and none of them co-localized with genomic regions under selective sweep in *O. glaberrima* in comparison to *O. barthii* (Wang et al. 2014). This result suggests that these genes were not directly under human selection pressure during domestication. Otherwise, the global alteration of expression of these genes would be independent during panicle development between *O. glaberrima* and O. *barthii.* The finding suggests that the expression of (a) very early acting factor(s) in panicle development might be differentially affected between the two African rice species. In this context, it will be of great interest to understand the regulatory processes related to very early acting genes such as *APO1*, *APO2* and *TAW1* in rice species in order to understand the initial steps of panicle architecture control and its evolution.

2.3 Panicle phenotyping of *O. sativa* from a Vietnamese landrace collection.

Natural intra-specific diversity is a resource of valuable genetic characters or alleles related to plant development, biotic and abiotic resistance. The use of this resource for breeding programs is of great interest to development high yield potential cultivars able to adapt to changing environmental conditions. Thanks to the development of Genome-Wide Association Studies (GWAS) in crops (Han and Huang 2013) in conjunction with the development of NGS technology, there is a renewed interested in genetic resources with the objective of identify valuable alleles in genetic resources. GWAS are used to characterize genetic bases of the variation of complex quantitative traits by establishing statistical links between phenotypes and genotypes (Nordborg and Weigel 2008). The first advantage of GWAS on classical QTL detection in mapping populations is that the analysis can be conducted on a panel of varieties without having to develop specific mapping populations. The second advantage is that GWAS enable to explore a larger diversity of alleles existing in genetic resources.

Recent publications illustrated the application of GWAS in rice for main agronomic traits (Huang et al. 2010; Zhao et al. 2011; Huang, et al. 2012). In these studies, large panel of worldwide *O. sativa* varieties were used. However, a very small part of the rice genetic diversity from specific countries (such as Viet Nam) was exploited in these studies. This may constitute a limiting step in the aim to provide new alleles in local breeding programs. For this purpose, the establishment of a core-collection of Vietnamese rice traditional varieties was carried out. Traditional varieties (or landraces) from Vietnam may constitute a valuable resource of alleles of interest of morphological traits as well as for biotic and abiotic tolerances for local and world-wide breeding programs.

In the initial step, a panel of 223 accessions of *O. sativa* Vietnamese landraces was established three years ago in LMI RICE (Hanoi, Vietnam) in collaboration with the Plant Resource Centre (PRC, Hanoi, Vietnam), which provided the seeds. The selection of this panel was done in collaboration with B. Courtois (UMR AGAP, CIRAD, France) on the basis of the geographic distribution of the accessions all over the Vietnam and for their representativeness of the different eco-systems of culture in Vietnam. Moreover, a panel of 40 accessions representing the word-wide diversity of *O. sativa* was provided by CIRAD (France) as well as 2 accessions of *O. glaberrima* as out-group, provided by IRD-Montpellier (France).

The genetic characterization of these two panels was carried out first with DArT markers then with GBS-derived SNP markers. Finally, 185 Vietnamese accessions were genotyped by this way. Population structure and linkage disequilibrium decay were finely analyzed in order to assess up to which extent the Vietnamese panel was adapted for GWAS. This work was done in the frame of

Phung Phuong Nhung's PhD in LMI RICE and in collaboration with B. Courtois (CIRAD, France). A manuscript reporting on this work has been submitted to BMC Plant Biology:

Characterization of a panel of Vietnamese rice varieties using DArT and SNP markers in view of association mapping. NTP Phung, CD Mai, P. Mournet, J. Frouin, G. Droc, NK Ta, S. Jouannic, LT Thi, VN Do, P Gantet, B Courtois.

I am co-author of this manuscript by participating in the phenotypic data collection from the field assay conducted in 2011. Regarding the Vietnamese University policy on PhD report and defense, Phung Phuong Nhung's manuscript is not included to this PhD report.

For phenotypic data collection, the accessions from the two panels were grown under field conditions in the Plant Resource Center located at An-Khan-Hoai Duc, near Hanoi during the 2011 wet season. The same plots were used to collect DNA from single plant for genotyping, to measure few key parameters (flowering dates, plant height, tiller number, seed shape, glutinous/non glutinous seed trait) and to collect panicles. The experimental design was based on a randomization of the accessions with 3 replications. For each accession, plants were grown in a 1.0 m^2 plot with an effective of 30 plants. The initial design was to collect the first panicle from 5 plants per accession per repeat in order to get in fine 15 panicles per accessions.

Finally, 1581 panicles were collected from 158 Vietnamese accessions and 33 from the reference panel, due to pest diseases, weather conditions and the absence of flowering induction for some accessions. The detailed are reported in Table 2 in Annexes. Some of the sampled Vietnamese accessions were not genotyped (21 accessions), leading to 137 Vietnamese accessions suitable for further studies from this field assay. In parallel, for some of the genotyped accessions, we were unable to collect panicles (46 accessions). Series of photos was done in order to provide images for the passport information for each accession: in the field, in the lab for panicle shape/bearing, in the lab for spread panicles (Figure 2.2). An illustration of the diversity of panicle structure for this panel of accessions is shown in Figure 2.3.

Various panicle traits were measured by hand: rachis length, node number, primary branch number, primary branch length, secondary branch number and spikelet number. This work was mainly done in LMI RICE during my visits there with the help of I. Bourrier (UM2). The data acquisition was completed just this year due to the large set of panicles to analyze. Some preliminary analyses were performed on this set of data but more time will be needed for complete the statistical analysis of this set. The Figure 2.4 illustrates the distribution and frequency of the measured traits in the accessions from the two panels.

To elucidate the main components of variation in panicle architecture in this panel, 8 panicle traits were subjected to principal component analysis (Table 2.1). The contributions of PC1 and PC2 were 35.2% and 19.7% respectively. The sum of the top three components accounted for more than 66% of the total variance. For the top three principal components, substantial loadings occurred with the following: for PC1, increase in spikelet number (sp_nb), in secondary branches (TA_nb), in length of primary branches (SA_average) and in the total length of the axis of the panicle (TotalPanicleLength); for PC2, increase in internode length along the rachis (PA_length.SA_nb), the length of the rachis (PA_length) and the number of primary branches (SA_nb); and for PC3, increase in tertiary branches (QA_nb) (Figure 2.5). From this analysis, it can be concluded that several variables (i.e. measured traits) can explain the observed diversity and not only one. In summary, the main factor of the first component corresponds to the axillary meristems produced on the primary branches, which give secondary, tertiary and spikelet meristems. In the second component, the main component is more related to primary branches traits (numbers and length). In a following step, I will analyze the projection of the accessions according to their genotype (i.e. their affiliation to one of the sub-groups of the population as reported by Phung Phuong Nhung's analysis), in order to evidence if the same components contribute to variations in the different sub-groups and to evidence any structuration of the traits according the genotype, which would indicate to which extent we will have to consider the whole panel or sub-sets for GWAS.

In conclusion from this preliminary analysis, the observed panicle diversity on the Vietnamese panel mainly relies on two groups of variables, one related to the primary branching process and on the secondary branching process. This is of great interest in the frame of GWAS applied to this panel, as it might be possible to putatively identify genomic regions related to the two aspects of panicle development.

Figure 2.2: Examples of panicle images collected (genotype G9, Lốc trắng sớm plei cầu). A, panicle in the field; B, panicle shape (in lab); C, spread panicle.

Fig 2.3: Panicle architecture diversity in the Vietnamese collection panel. Illustrations of spread panicles

A second experiment has been set up this year, based only on the genotyped Vietnamese accessions without the late flowering ones (159 accessions *in fine*). This field assay was based on a randomization of the accessions with 2 replications. The sampling of the panicles is under going and the analysis will be carried out in 2015. For this purpose, the 2D-image analysis software developed by IRD-Montpellier in collaboration with F. Al Tam and HR Shahbazkia (Universidade do Algarve, Faro, Portugal), named P-TRAP (Faroq et al. 2013) will be used in order to collect more data from spread panicles and to avoid error-prone manual counting.

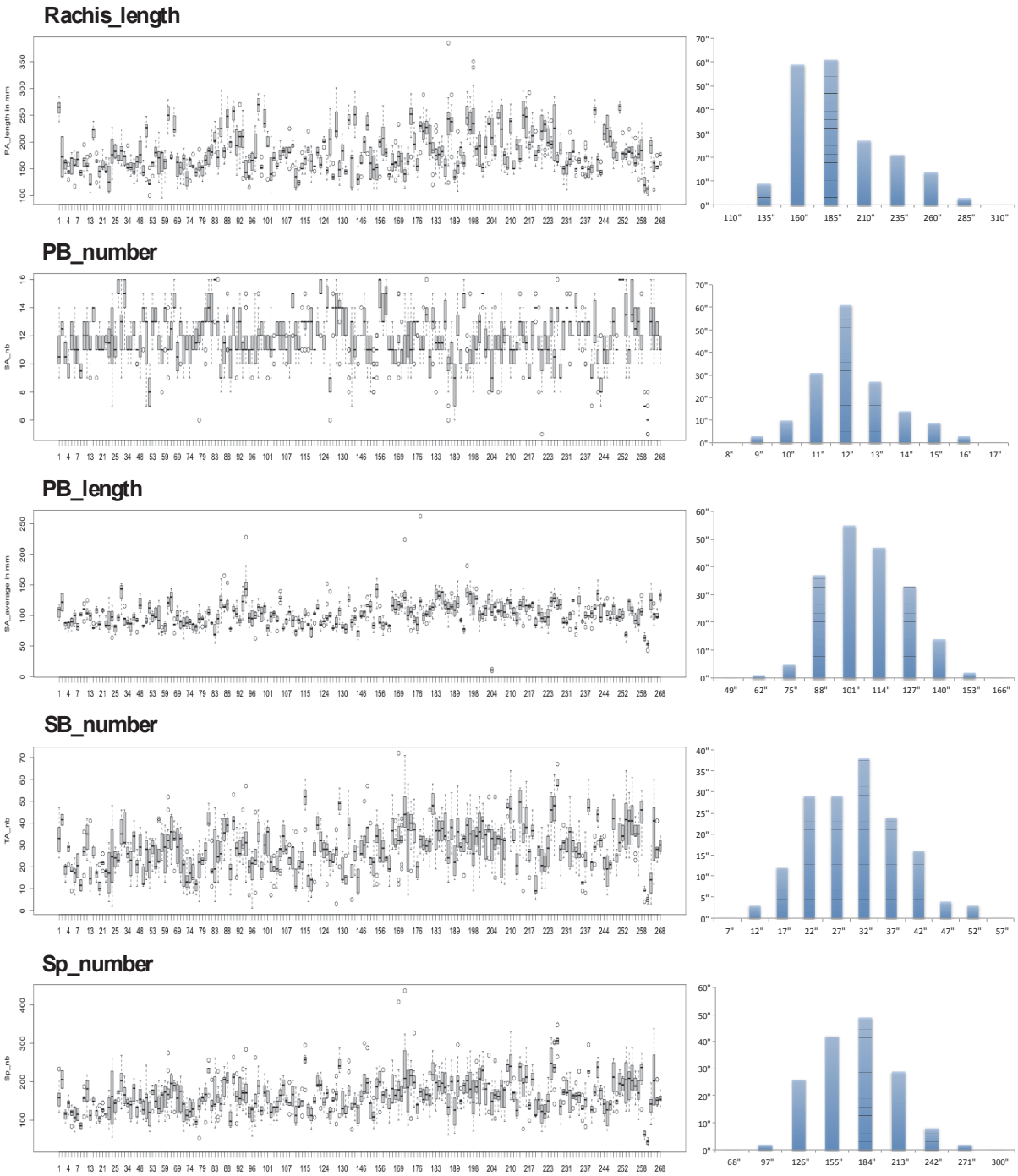


Figure 2.4 Distribution and frequency plots of panicle traits measured across the panel of Vietnamese land races and reference accessions. Left panel: distribution plots; number on x-axis indicates accession ID; y-axis indicates traits values (number or length in mm); the stand deviation of the values is indicated by grey boxes with the average by black dash. The frame of min and max values is indicated by dotted line. Right panel: Frequency plots for each measured traits. The y-axis represents number of accessions for each class. The x-axis represents trait values (number or length in mm).

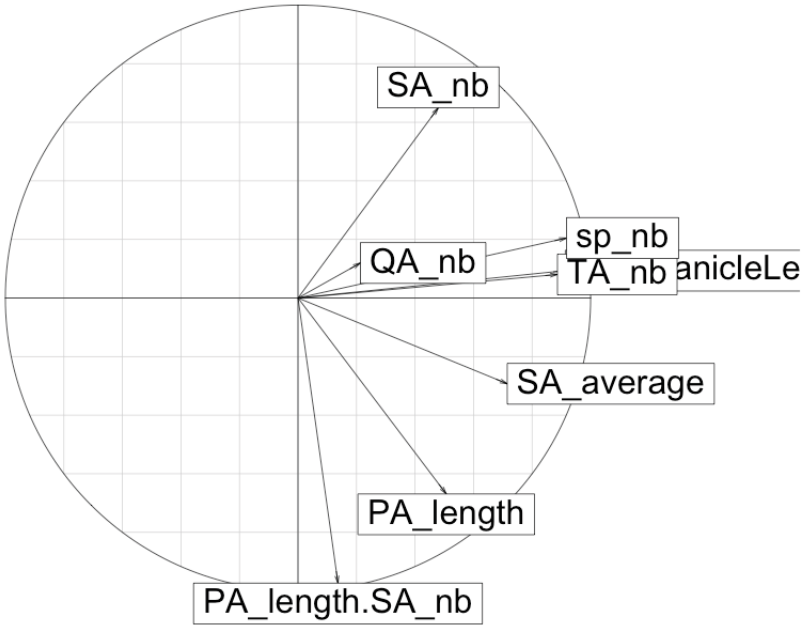


Figure 2.5: Principal Component analysis (PCA) plot for panicle traits from the Vietnamese and reference panels. PA_length: rachis length (PA= primary axis); SA_average: primary branch length average (SA= secondary axis); SA_nb: primary branch number; TA_nb: secondary branch number (TA= tertiary axis); QA_nb: tertiary branch number (QA= quaternay axis); Sp_nb: spikelet number; Total panicle length: sum of rachis, primary branch and secondary branch length; PA_length.SA_nb: ratio of PA_length and SA_nb values (i.e. average internode length).

Principal Component	1st	2nd	3rd	4th	5th	6th	7th	8th
Eigen value	3.52	1.97	1.11	0.83	0.49	0.039	0.014	0.01
Percentage contribution	35.2	19.7	11.1	8.3	4.9	3.9	1.4	0.1
Cumulative contribution	44.05	68.73	82.62	93.00	99.20	99.70	99.88	100

Table 2.1: Eigenvectors for 8 panicle traits in principal component analysis in Vietnamese and reference panels.

3.

CONCLUSIONS AND PERSPECTIVES

3. CONCLUTIONS AND PERSPECTIVES

This PhD work mainly focused on the inter-specific comparison between two African rice species, the domesticated *O. glaberrima* and its wild-relative *O. barthii*. This work provided the first inter-specific genome-wild analysis of the small RNA population expression during panicle development. By using deep sequencing of small RNA transcriptomes, it revealed a drastic change in the 21-nt small RNA population expression between the domesticated species and its wild-relatives, corresponding mainly to a delay of expression timing of the male gametogenesis-associated *miR2118*-triggered 21-nt phasiRNAs pathway in *O. glaberrima*. In addition, this work provided evidences on a heterochronic shift of panicle-related gene expression in conjunction with differential accumulation between the two species. This finding indicated that both spikelet-related genes and branch-promoting gene expression were affected, suggesting a prolonged branching phase and a delayed spikelet meristem fate acquisition in domesticated rice. Moreover, this would suggest that (a) very early acting factor(s) involved in panicle development might be at the origin of the differential expression of those genes between the two species. However, there are still several questions remaining to be addressed. The first point is that our study is a descriptive analysis that needs to be complete by functional and genetic analyses, in order to evidence whether these alteration of gene expressions lead to the difference of panicle architecture observed between the two species. Moreover, beside the landmark genes and small RNAs studied here, it will be important to determine whether new/other factor related to panicle architecture (*i.e.* hormones, new genes, etc.) may be involved in the rice domestication/evolution in African rice, but also in Asian rice. I propose some works in short-term and long-term perspective to answer those questions.

The short/mid-term perspectives

Firstly, how to describe at cellular level the heterochonic alteration we proposed during panicle development? It must be reminded that panicle architecture is determined at early stage of panicle development, which is a very short-time process (about 10-15 days), difficult to analyze. We used the classic optical microscopy in conjunction with specific histological staining to observe the early stages of inflorescences. However, the method used is destructive and do not allow *in vivo* imaging to follow developmental time course of the panicle. Thus, we suggest using the scanning electron microscopy (SEM) to get better views of panicle surface and branching structure (3D image). Recently, Dhondt et al. (2010) presented a new development in high-resolution X-ray computed tomography (HRXCT) that is a minimally invasive structural imaging method that allows 3D reconstruction of scanned

objects. The advantage of HRXCT is the non-destructive nature of this technology for *in vivo* imaging and will permit to follow the daily *in vivo* development of panicle. The combination of these techniques may allow getting a complete description of the staging of inflorescence development in different species. These comparisons will help us to understand in detail what are the cellular events and at which early staging of development difference during panicle development could be observed between the species. In the frame of my PhD work, we tried to apply this technology to African rice panicles using facilities at UM2 (Montpellier). However, the resolution of the X-ray tomograph was not high enough to get good quality images of the inner structures of the panicle within the plants.

Secondly, how to get evidences that (a) very early acting factor(s) in panicle development may explain the difference of panicle architecture observed? Based on the landmark gene expression analysis, we have shown differential expression of very early-acting genes such as *APO1, APO2* and *TAW1*. This last gene is a member of the small gene family ALOG and acts partly through the regulation of flowering time-related *SVP*-like genes to control the rice panicle architecture (Yoshida et al., 2013) where as *APO1* and *APO2* encode for F-protein and a plant-specific transcription factor, respectively. Interestingly, the orthologs of *APO1* and *APO2* in other species such as Arabidopsis and tomato were shown to be molecular factors of inflorescence evolution. However it has to be kept in mind that difference of in the function of these genes were significant between eudicots and at least grasses. The *apo1, apo2* and *taw1* lost function mutants were characterized by lower panicle complexity (Ikeda et al. 2007; Ikeda-Kawakatsu et al. 2012; Yoshida et al. 2013). The gain of function *apo1-D1* and *taw1-D2* mutants, as well as the overexpressing transgenic plants, were characterized on the inverse by higher branching panicles by enhanced branch meristem activity and delayed spikelet meristem specification (Ikeda-Kawakatsu et al. 2009; Ikeda-Kawakatsu et al. 2012; Yoshida et al. 2013). Initial histological analysis provided evidences that these genes may act through the control of cell proliferation within the inflorescence meristem (Ikeda et al. 2005; Ikeda-Kawakatsu et al. 2009; Ikeda-Kawakatsu et al. 2012; Yoshida et al. 2013). However, clear evidences of the relationship between inflorescence meristem size alteration and differential timing of the main phases of panicle development are still lacking. Thus, it will be interesting to complete histology analysis of these mutants in comparison with the *O. sativa* wild type but also with African species, by using classical but also new imaging tools. The observation could give us insights about cellular evens occurring in very early stages of panicle development that determine the panicle complexity in the mutations and the different wild-type species. In a next step, the expression analysis of panicle-related genes in the mutant backgrounds in comparison with the wild-type might not only explain relation among *APO1, APO2* and *TAW1* as well as the relation of those genes with other panicle-relative genes.

The long-term perspectives

Because the time limit of the PhD project, it was not possible to analyze other factors that may be possibly involved in the rice panicle development. For instance phytohormones were known to play a critical role in regulating branching of inflorescence (Barazesh and Mcsteen 2008). Among them, auxin is required for axillary meristem initiation during both vegetative and inflorescence and was shown to involved in inflorescence branching in different species such as Arabidopsis and maize (Benková et al. 2003; Morita and Kyozuka 2007; Barazesh and Mcsteen 2008). Although few genes were established that regulated auxin transport (*i.e. OsPIN1, LAX*) in *O. sativa*, the function of auxin and its activity during early stage of panicle development in rice is still not evidenced (Morita and Kyozuka 2007; McSteen 2009). It will be of great interest to analyze auxin marker lines in *O. sativa* such as DR5-GUS (*i.e.* a synthetic promoter for auxin response marker) in both wild-type and mutant backgrounds, in order to evidence relationship between landmark genes and auxin pathway. 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. Finally, preliminary results from mRNA transcriptomic analysis in *O. sativa* and African rice species in our lab indicate that Brassinosteroid pathway-related genes are differentially expressed between the different meristematic state and between the different species, opening a field related to this phyto-hormone.

This type of analysis is currently limited to the Asian species *O. sativa*. Indeed, no transformation protocol was reported for the African species, as well as for the wild species in *Oryza* genus. It will be of great importance for future comparative analysis to develop an efficient protocol of genetic transformation for these species. This is currently undergoing in LMI RICE-AGI lab, having an expertise in rice transformation (both *japonica* and *indica*), but will need sustained efforts to get an optimized protocol for African rice species.

In order to identify key regulators and evidence new factors controlling rice panicle development related to meristem state control, two mains approaches will be developed. The first one relies on genome-wide expression analysis through NGS technology. By using the advance technique such as LASER Microdissection Microscopy (LMM) to sampling the correct meristems in conjunction with illumine mRNA-seq/smallRNA-seq and bioinformatics analysis, the study not only will allow to identify genes/smallRNA related to differentiation during panicle development but also genes/smallRNA highly correlate to specific meristem states (*i.e.* rachis, branch and spikelet meristems) among Asian and African species. The findings will provide evidences of

genes/smallRNAs showing convergent evolution in expression during the two domestication events and also new candidate genes regulated panicle development. As a result, the study contributes to the understanding about evolution/domestication as well as the molecular mechanism controlling the panicle development.

The second approach, complementary to the first one, relies on the identification of the genetic factors contributing to the differential panicle architecture between and within rice species. As already mentioned, a genome-wide association study is currently developed on a collection of *O. sativa* Vietnamese landraces, a project in which I will continue to work after my PhD in order to identify genomic regions or/and DNA elements (*i.e.* single nucleotide polymorphisms - SNPs) related to panicle diversity within this Vietnamese panel of *O. sativa* species, which may be suitable for marker-assisted selection related to yield improvement breeding programs. This approach will have to be completed through other approach as bi-parental mapping population, detailed bio-informatic analysis in order to get a short list of putative candidate genes related to panicle traits in the identified regions. This will lead to functional analysis (mutant or transgenic lines characterization) for validation. As a mirror, a similar approach on an African rice collection (both *O. glaberrima* and *O. barthii*) is under development at IRD-Montpellier in collaboration with INERA (Burkina-Faso) for the phenotyping. These parallel works will allow to whether similar factors are related to intra-specific diversity of panicle structure in both Asian and African species.

In the context of genetic factors contributing to the inter-specific diversity of panicle structure, an analysis is undergoing using a specific inter-specific *O. sativa* x *O. glaberrima* population (Chromosome Segment Substitution Lines or CSSLs) (Gutiérrez et al. 2010). A population of 63 lines representing the complete genome of *O. glaberrima* in *O. sativa* background has been already phenotyped for panicle traits. Those lines were already genotyped at high density through GBS method, allowing in the near future the identification of *O. glaberrima* genomic regions conferring a specific panicle trait in contrast to *O. sativa*. Moreover an inter-specific mapping population between *O. glaberrima* and *O. barthii* is under development in IRD-Montpellier to complete this analysis in order to evidence genetic factors in the context of African rice domestication.

All together, these multi-scale approaches integrating both cellular, genetic, genomics and molecular analyses will led to a complete view of the events controlling early reproductive development in rice and the factors involved in the diversity of panicle architecture in the context of domestication and between Asian and African rice.

4.

MATERIALS AND METHODS

4. MATERIALS & METHODS

4.1 Materials

4.1.1 Chemicals and kits

Unless otherwise indicated, 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 (Carlsbad CA, USA), and Promega (Madison WI, USA). SYBR Green I kit was provided by Roche (France). Taq DNA polymerase was purchased from Promega (USA). All the enzymes were supplied and used with their buffers.

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.

4.1.2 Plant materials

The set of rice (*Oryza sativa*, *Oryza glaberrima* and *Oryza barthii*) used for the mature panicle phenotyping were grown in summer 2011 in fields of CIAT (Cali, Colombia) (n= 3 plants per variety, 2 replicates). Three panicles per plant were harvested at mature stage for each replicate (n=18 panicles per variety). Genotypes include Nipponbare, CG14 and B88 for *O. sativa*, *O. glaberrima* and *O. barthii* respectively. P-TRAP software has been used to measure the panicle traits (Al Tam et al., 2013).

For histological study and expression analysis of landmarks genes, Nipponbare, CG14 and B88 plants were grown in growth chamber at IRD, Montpellier (France). Chamber settings were as follows: on 14-h day/night cycle at 32°C/28°C, and humidity at 60%. Flowering was induced by short day conditions (10-h day/night cycle). 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.

For Illumina sequencing, 10 accessions of *O. glaberrima* and 10 accessions of *O. barthii* (Table 2) were grown in the greenhouse at IRD, Montpellier. Around 15 panicles from each accession were collected from 4 to 15 days after induction, corresponding to stage 1 (rachis and primary branch meristem) to stage 3 (spikelet differentiation) of panicle development.

For intra-specific analysis of Vietnamese rice, 218 accessions (Table 2) were grown in summer 2011 in fields of PRC (Plant Resource Center, Hanoi, Vietnam). There were 188 traditional accessions (provided by the PRC), originating from different districts of Vietnam and diverse rice ecosystems, and 30 reference accessions (provide by the CIRAD, France) from a core collection representing the varietal group diversity of *Oryza sativa.* Five mature panicles from 5 individuals per accession in 3 repeats were collected for analysis. Structural traits of the panicles were measured by hand: rachis length, the number of spikelet per panicle, the number of primary branch per panicle, the number of secondary and tertiary branch per primary branch, and the number of nodes associated. Other traits were collected during the development of the plants, including plant high, dry weight of plant, the number of tiller and efficient tiller per plant. A dedicated Excel macro program was used to manage data of all accessions.

4.2 Methods

4.2.1 Isolation of plant nucleic acids

DNA isolation

DNA samples were extracted from leave following CTAB method (*Winnepenninckx et al. 1993*). About 100 mg leave were ground into powder in liquid nitrogen and placed in a 1.5 ml microfuge tube. The leaf tissue was homogenized in 700µL CTAB buffer (1% w/v CTAB, 700mM NaCl, 10mM Tris-HCl pH8 and 50 mM EDTA pH8.0), vortexed very well and incubated at 65°C during 60 min for cell lyses. A volume of 700µL of chloroform/isoamyl alcohol (24:1) was added, mixed by hand and then centrifuged at 14,000 rpm for 5 min at 4°C. The upper aqueous layer was transferred to a new 1.5mL microfuge tube which contained of 700µL cold iso-propanol, well mixed and kept at -20°C for 60 min. The tubes were centrifuged at 14,000 rpm for 20 minutes at 4°C. Supernatant was discarded and washed by adding 1mL 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 re-suspended in 50µL sterile dH_2O. The concentrations of DNA samples were determined using a NanoDrop ND-1000 Spectro apparatus. The integrity and size distribution of total DNA were checked by agarose-gel electrophoresis.

RNA isolation

Total RNAs (mRNAs and small RNAs) from different stage (stage 1 to stage 4) during rice panicle development were extracted using an RNeasy Plant Mini Kit (Qiagen) with a modified protocol. Samples were ground into powder in liquid nitrogen and transferred into a 15 mL tube

containing 3 mL mixture (2.5 mL RLT buffer and 250 µL of β-mercaptoethanol) followed by vortexing for 20 seconds. The lysate was transferred to 2 QIAshredder Spin Columns per sample followed by centrifugation at 13,000 rpm for 2 min at room temperature. The filtrate was carefully collected in 15 mL tube and this step was repeated until the complete loading of the 3 mL lysate on the QIAshredder Spin Columns. 1,5 volumes of Ethanol 100% (v/v) was added to the filtrate and mixed by pipetting. The solution was added to 2 RNeasy Mini Spin Columns prior to centrifugation at 13000 rpm for 30 seconds at room temperature. The filtrate was then discarded. This step was repeated until complete addition of the recovered volume to the RNeasy Mini Spin Columns. Then, 350 µL of RWT buffer was added to RNeasy Mini Spin Columns followed by centrifugation for 30 seconds at 13,000 rpm. DNAse treatments were performed using an RNase-free DNase kit (Qiagen, France) according to the manufacturer's instructions, incubating for 15 min at room temperature. Then, 350 µL of RWT buffer was added to the column and centrifugation was carried out for 30 seconds at 13,000 rpm to discard the flow. A volume of 500 µL of RPE buffer was added before centrifugation for a further 30 seconds. This step was repeated twice. Then, the column was placed in a new 1.5 mL tube and 35 µL of RNAse-free water was added to the column and allowed to impregnate for 10 min at room temperature. Centrifugation was then performed for 1 min at 13000 rpm. The eluate from the first column was recovered and added to the 2nd column followed by centrifugation. The concentration of the RNA samples was determined using a NanoDrop ND-1000 Spectro apparatus. The collected RNAs were stored at –80°C.

4.2.2 Illumina sequencing and data processing

Purified small RNA sequencing was performed by Eurofins/MWG Operon (Germany) on an Illumina *Hi-seq 2000* using the TrueSeq^{TM} SBS v5 sequencing kit. The raw data (accession number GSE48346 in NCBI Gene Expression Omnibus) were trimmed by removing adapter sequences and low quality sequences using *CutAdapt* (Martin 2011). All the trimmed reads ranging from 18 to 28 nucleotides were clustered and mapped to *O. sativa ssp japonica cv Nipponbare* genome (MSU release version 7; http://rice.plantbiology.msu.edu/) using *BLAST* (Altschul et al. 1990). The 18-28 nucleotide reads were annotated using successive hierarchical *BLAST* versus (in order) miRBase v17.0 (Kozomara et al. 2011), Rfam v7, home-made repeat database (successive curated concatenation of *RetrOryza*, RepBase, TREP and TIGRRepeats), CDS then gene features from *Oryza sativa ssp japonica cv Nipponbare* MSU v7.0 annotation, and finally the MSU v7.0 rice genome. The *BLAST* and post-filters parameters used were probability of 85%, e-value of 10^{-3}, on a size of 85% of the reads (minimum size of 16). The same BLAST parameters were used throughout the analysis. Mapping from *O. glaberrima* and *O. barthii* were then compared and filtered using a series of homemade *Perl*

scripts (available on demand). The 21-mers were used in phasing analysis with the *ta-si Prediction* tool from the *UEA sRNA workbench* facilities (http://srna-workbench.cmp.uea.ac.uk/; Stocks et al. 2012). Once the loci were identified, we used the *EMBOSS* software suite v6.5.7.0 (Rice et al. 2000) to extract -500/+500 bases around each locus, and treated them using *MEME* v4.8.1 (Bailey and Elkan 1994). A search for putative initial targets of phased small RNAs was performed using *psRNATarget* server facilities (Dai and Zhao 2011; http://plantgrn.noble.org/psRNATarget/) with default parameters except the maximum expectation value of 1 (i.e. the threshold of the scoring schema of miRU according to Zhang (2005).

Statistical tests of all the processed data were performed using *g-test* and a fixed *p*-value of 10^{-3}. Depending on the experiment, the degree of freedom was adjusted but was generally 1. All the calculations were performed using homemade *Perl* scripts and CPAN statistical modules.The sequence of candidate genes and their promoter were obtained using BLASTn program (ref) using *O. glaberrima* (AGl1.1) and *O. barthii* genome sequences from database of Gramene database (http://blast.gramene.org/Multi/blastview). Identification of putative transcription factor binding sites (TFBSs) in promoter regions was done using Genomatix software (http://www.genomatix.de/). To obtain complete sequence of some genes (i.e. *APO1, APO2* and *SPL14*) from *O. barthii*, PCR amplifications were done using total DNA from O. barthii B88 accesion. The primer used are listed in Table 1. PCR products were in pGEM®-T Easy vector (Promega) and sequenced by Beckman Coulter Genomics (www.cogenicsonline.com). The sequences was annotated using Mobile Pasteur facilities (http://mobyle.pasteur.fr/cgi-bin/portal.py)

4.2.3 Genes expression analysis

First-stand cDNA synthesis was done using SuperScript III cDNA First-strand synthesis system (Invitrogen) using 1µg of total RNA in conjunction with 1µL of 50mM oligo$(dT)_{20}$. Then cDNA were obtained using the SuperScript^{TM}III Reverse Transcription System (Invitrogen, USA) according to the manufacturer's instructions.

Semi-quantitative RT-PCRs were performed using GoTaq polymerase (Promega) as follows: 4 µL of 5X buffer GoTaq buffer; 0.5 µL of both sense and antisense gene specific primers (10 µM each) (see primer list on Table1); 0.5 µL dNTP mix (10mM); 0.05 µL Taq polymerase (5U/µL) and 13.35 µL of H_20. PCR amplifications were performed with an Applied Biosystems GeneAmp PCR System 9700, using the following conditions: 2 min 94°C; (30 sec 94°C; 30 sec 55°C; 30 sec 72°C) for 25 to 30 cycles; 10 min 72°C. The integrity and size distribution of the genes were checked by agarose-gel electrophoresis.

Quantitative RT-PCRs were performed in an optical 384-well plate using LightCycler 480 thermocycler (Roche, France). Triplet reactions for each sample contained 4 µL SYBR Green Master Mix (Roche), 2 µL of diluted RTs and 0.8 µL of forward and reverse primers (10 µM each) in a final volume of 10 µL (see primer list on Table 1). The Q-PCR amplification conditions include 3 stages: pre-incubator (10 min 95°C); amplification with 45 cyclers (15s 95°C and 30s 60°C); melting curve (5s 95°C and 1 min 70°C). Target cDNAs were normalized using transcripts accumulation level from the rice *Actin* gene (*LOC_Os03g50885*) (Table1). Each set of experiments was repeated three times, and the relative quantification method with efficiency corrected calculation model (Souaze et al., 1996) was used to evaluate quantitative variation. The primers used are listed in Table 1. Statistical tests were performed using t-test with two-tail test and a fixed *p*-value of 0.01

4.2.4 miRNA and phasiRNA expression analysis

Stem-loop RT technique followed by end-point PCR was performed to synthesis RT of small RNAs (Figure 4.1). Base on the spatial constraint of the structure, the stem-loop RT primers was used to provide better specificity and sensitivity. First, the stem-loop RT primer was hybridized to the miRNA molecule and then reverse transcribed in a pulsed RT reaction.

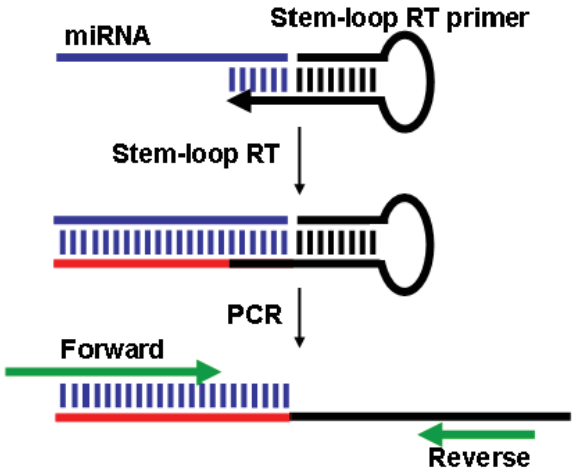


Figure 4.1 Schema showing stem-loop RT-PCR miRNA assays (Varkonyi-Gasic et al. 2007). Stem-loop RT primers bind to the 3' portion of miRNA molecules, initiating reverse transcription of the miRNA. Then, the RT product is amplified using a miRNA specific forward primer and the universal reverse primer

Small RNAs were performed using 100 ng of total RNA according to Varkonyi-Gasic et al. (2007). A mix of total RNA in conjunction with 1µL small RNA-specific RT primers (1µM) (Table1) in final volume of 9.5µL was incubated at 70°C for 5 min then on ice for 5 min to denature RNA. The mix was then centrifuged 10 sec at maximum speed and 10.5µL of the reaction mix (4µL $MgCl_2$

(25mM), 4µL Improm-II 5X buffer, 1µL dNTP (10mM), 1µL RNAsin (40U/L) and 1 µL Improm-II RT) was added. Pulsed RT reaction was performed following the conditions: 16°C for 30 min, (30°C 30 sec ; 42°C 30 sec ; 50°C 1sec) for 60 cycles, 70°C for 15 min.

To analyze the expression of miRNA and phasiRNAs, quantitative RT-PCRs were performed using a small RNA-specific forward primer and the universal reverse primer (Table1). The levels of small RNAs were normalized by using mature *miR159* accumulation level. Statistical analysis using T-Test with two-tail test a fixed *p*-value of 0.01

4.2.5 Northern blot hybridizations

Small RNA northern blot hybridizations were performed with 15 µg of the two RNA bulks separated in a 15% (w/v) poly-acrylamide gel and electro-transferred to a Zeta-Probe GT nylon membrane. Blots were cross linked under UV and hybridized overnight at 40 °C in PerfectHyb™ Plus hybridization buffer (Sigma) with ^{32}P-radiolabeled oligonucleotide probes complementary to the miRNA sequences. As a loading control, each blot was hybridized with an oligonucleotide labeled probe complementary to U6 snRNA. After two washes at 50 °C in 2xSSC-0.1% (v/v) SDS solution, blots were scanned using a Typhoon 8600 imager system (Amersham). The sequences of the oligonucleotides used as probes are listed in Table 1

4.2.6 Histology analysis

Rice panicle samples (from stage 1 to stage 4) were fixed overnight at 4°C in fixation buffer (4% (v/v) paraformaldehyde, 0.1 M phosphate buffer, pH7). Samples were dehydrated through a graded EtOH series (30, 50, 70, 90, 100% (v/v)), 30 min for each concentration of EtOH and stored at 4°C overnight. The sample then was embedded in Technovit resin (Heraeus Kulzer, Germany). Blocks were sectioned at 4-5µm thickness using a HM650 microtome (Thermo Scientific Microm, Walldorf, Germany). Slides were double-stained with PAS stain (periodic acid–Schiff reagent) for the detection of carbohydrate compounds and naphthol blueblack (NBB) for the detection of proteins. Slides were observed with a Leica DMRB microscope and photographed by Evolution MP5.0 color Media Cybernetics camera.

4.2.7 *In situ* hybridization

4.2.7.1 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, according to the principle detailed in figure 2.2. Transcription using the RNA polymerase corresponding to promoter 1 leads to sense RNA

(*i.e.* the same sequence as the mRNA). If the RNA polymerase for promoter 2 is used, antisense RNA is transcribed.

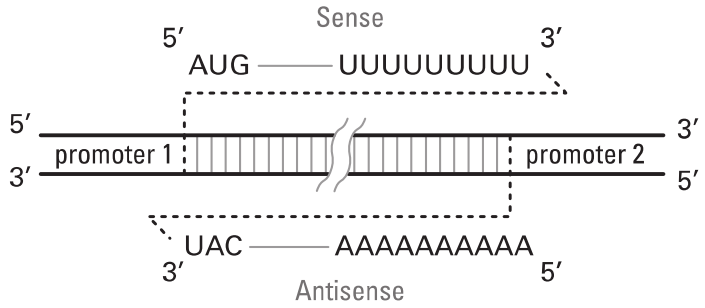


Figure 4.2 Principle of *in situ* hybridizations

cDNA synthesized from total RNAs of difference stages during panicle development was used to prepare probe. PCR_0 amplifications were performed using specific primers. Each reaction contains 4 µL of 5X buffer GoTaq buffer; 0.5 µL of both sense and antisense gene specific primers (10 µM each); 0.5 µL dNTP mix (10mM); 0.05 µL Taq polymerase (5U/µL) and 13.35 µL of H_20. PCR amplifications were performed following conditions: 2 min 94°C; (30 sec 94°C; 30 sec 55°C; 30 sec 72°C) for 25 to 30 cycles; 10 min 72°C. The PCR_0 products were cloned into pGEM®-T Easy and JM109 competent cells (Promega) according to the manufacturer's instructions. After transformation, colony screening was performed by direct PCR on colonies using universal primers (forward and reverse M13). Plasmid DNA from positive clones was prepared using a QIAprep Spin Miniprep Kit (Qiagen) and sequenced by Beckman Coulter Genomics (www.cogenicsonline.com) for sequence confirmation (i.e. specific to candidate genes, size from 150bp to 400bp; 30% Adenin). Plasmid DNAs were then used as matrice for PCR amplification of T7 RNA promoter-containing DNA fragments used for sense or antisense RNA transcription. Firstly, PCR_1 used PCR_0 products (1/200 dilution) as template with 2 reactions (PrimerT7 sense + Primer Antisense; PrimerT7 antisense + Primer Sense) (Fig 2.2). Then, PCR_2 used PCR_1 products (1/200 dilution) as template with 2 reactions (T7 primer + Primer Antisense; Primer Sense + T7 primer). PCR_1 and PCR_2 conditions are similar to PCR_0 condition (see above). PCR_2 product was purified by EtOH precipitation, and eluted in 50µL RNAse-free H_2O. The concentrations of DNA samples were determined using a NanoDrop ND-1000 Spectro apparatus. PCR_2 produce (1µg/µl) was 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. DNA fragments from PCR reaction Sense+T7 provided antisense RNA probes while the DNA fragments from PCR reaction T7+Antisense provided sense RNA probes. 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. For *miR2118*, *miR529* and *phasiPH12* detection, 0.02 µM of a 5' digoxigenin–labeled LNA probe complementary to the target was used.

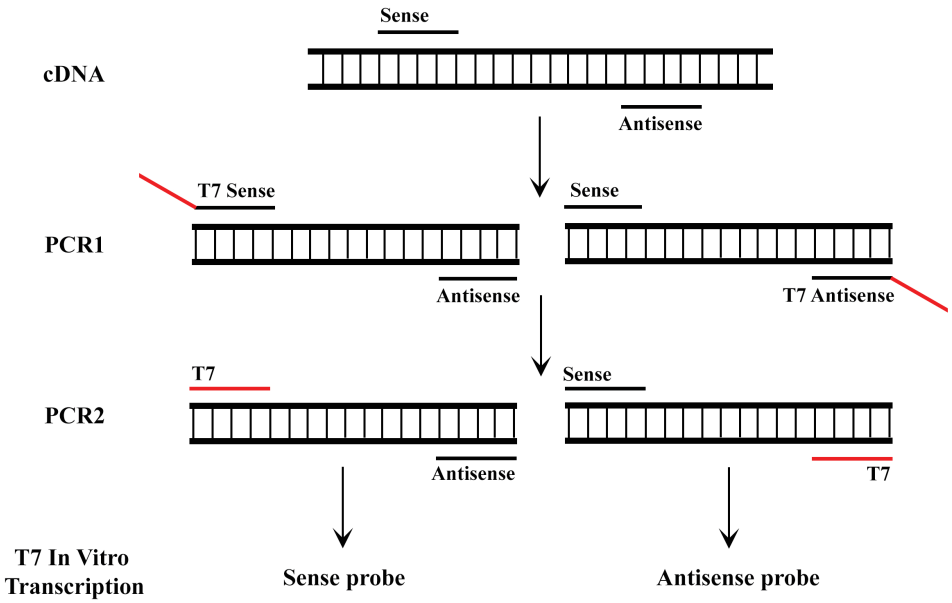


Figure 4.3 Schematic showing sense and antisense probe assays.

4.2.7.2 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.

4.2.7.3 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 histoclear (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 this 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 that 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 slide of sample was prepared one day before *in situ* hybridization experiment. The plastic block of paraffin containing the samples was removed and histological sections of 8 µm of thickness were made using a microtome (Leica GmbH D6907 instrument Nussiloch, 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 in 3 months

4.2.7.4 In situ hybridization

The experiments were carried out as described by Adam et al., (2007) contain 6 steps. In the first step the tissues were dewaxed by histoclear (3 times for 10 min) and hydrated through a graded EtOH series, (100° - 2 times for 10 min, 70° and 50° for 5 min, DEPC water – 2 times for 10 min). Then, the 2nd step, the slides were treated by proteinase K in 200ml 1X Proteinase K buffer (added 134µl 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 (2 times) at 2 min. In step 3, the slide was dehydrated by cleaned through a graded EtOH series (50°, 70°, 100° - 2 times) for 1 min per solution. Hybridization mix was prepared at step 4 include 50µ formamide 100%, 10 µl SSC 20X buffer, 20 µl Sulfate Dextran 50%, 4 µl Denhardt 50X, 1 µl ARNt (11ng/ml), 1.5 µl probe (200ng/µl) and DEPC water to have final volume 100 µl per slide. Mix was warmed up at 65°C for 5 min and kept in ice. Hybridation 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, plaque Omnislide thermo cycler) at 42°C overnight. In order to accomplish step 5,the slide was washed by several buffers to remove non hybridized single-stranded RNA probe containing SSC 2X buffer (one time for 5 min in room temperature and one time for 45 min 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, SSC 1X buffer for 15 min and PB 1X (2 times for 10 min). In the last step, the slide was incubated in 700 µl of 1% blocking solution in PBS1X buffer for 1 hour in humid chamber. Then, this solution was replaced by 500 µl of 1% blocking solution with 1/500 antibody anti-deoxygenize 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

4.3 Medias, solutions and buffers

Medium LB

Bacto®-tryptone	10g.L^{-1}
Bacto®-yeast extract	5g.L^{-1}
NaCl	10g.L^{-1}
In deionised water	

Medium LB agar

Bacto®-tryptone	10g.L^{-1}
Bacto®-yeast extract	5g.L^{-1}
NaCl	10g.L^{-1}
Bacto-Agar	1,5% (w/v)

SOC medium (100ml)

Bacto®-tryptone	2.0g
Bacto®-yeast extract	0.5g
1M NaCl	1mL
1M KCl	0.25mL
2M Mg^{2+} stock, filter-sterilized	1mL
2M glucose, filter-sterilized	1mL
pH	7.0

CTAB 2X buffer (1l)

CTAB	20g
0.5M EDTA pH 8.1	40ml
1M Tris HCL pH8.0	100ml
5M NaCl	280ml
β-mercaptoethanol [added just before use]	0.2%

Solution for *insitu* hybridization

All the solution are *RNAse* free and traited with DEPC water

DEPC water (1l)

DEPC	100µl
Water (miliQ)	900 µl

Mixed well and incubated 4 hour before autoclave

Proteinase K 10X buffer (500ml)

Tris-HCL (1M)	60.57g
EDTA (0.5M)	84.05g
pH	8

TRIS 10X buffer (500ml)

Tris-HCL (1M)	60.57g
pH	8.2

PBS 10X buffer (1l)

Na_2HPO_4 (70mM)	9.94g
NaH_2PO_4 (30mM)	3.6g
NaCl (1.2M)	70.128g
KCl (27mM)	2g
pH	7.4

SSC 20X buffer (1l)

NaCl (3M)	175.5g
$Na_3C_6H_5O_7$ (300mM)	88g
pH	7

NTE 10X buffer (1l)

Tris-HCL (100mM)	12.114g
NaCl (5M)	292.2g
EDTA (10mM)	3.671g
pH	7.5

Revelation 10X buffer

Tris-HCL (1M)	121,14g
pH	8.2

RNAse A (10g/)

RNAse power	0.04g

NTE1X	4ml

Boiled for 5 min at 100°C, then aliquoted in 1.5ml microfuge tube .

Blocking solution 10 % (100ml)

Blocking reagent	10g
Maleic acid ($C_4H_4O_4$)	1.161g

Heated by microwave to solubilise, mixed well, autoclaved and aliquoted. Solution was stored at -20°C.

Denhardt 50X (500mL)

NaCl	0.8766g
pH	7.5

5.

BIBLIOGRAPHY

5. BILIOGRAPHY

Abdullah AB, Ito S, 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, **2**:28–43.

Adam H, Marguerettaz M, Qadri R, Adroher B, Richaud F, Collin M, Thuillet AC, Vigouroux Y, Laufs O, Tregear JW, Jouannic S (2010) Divergent Expression Patterns of *miR164* and *CUP-SHAPED COTYLEDON* Genes in Palms and Other Monocots: Implication for the Evolution of Meristem Function in Angiosperms. Molecular Biology and Evolution **28**: 1439–54.

Ahearn KP, Johnson HA, Weigel D, Wagner DR (2001) *NFL1*, a *Nicotiana Tabacum LEAFY*-Like Gene, Controls Meristem Initiation and Floral Structure. Plant and Cell Physiology **42**: 1130–39.

Ainsworth, Charles (2008) Flowering and Its Manipulation. John Wiley & Sons.

Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-Directed Phasing during Trans-Acting *siRNA* Biogenesis in Plants. Cell **121**: 207–21.

Arikit S, Zhai J, Meyers BC (2013) Biogenesis and Function of Rice Small RNAs from Non-Coding RNA Precursors. Current Opinion in Plant Biology **16**: 170-179.

Arite T, Umehara M, Ishikawa S, Hanada A, Maekawa M, Yamaguchi S, Kyozuka J (2009) *d14*, a Strigolactone-Insensitive Mutant of Rice, Shows an Accelerated Outgrowth of Tillers. Plant and Cell Physiology **50**: 1416–24.

Arthur W (2002) The Emerging Conceptual Framework of Evolutionary Developmental Biology. Nature **415**: 757–64.

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin Oxidase Regulates Rice Grain Production. Science **309**: 741–45.

Axtell MJ (2013) Classification and Comparison of Small RNAs from Plants. Annual Review of Plant Biology **64**: 137–59.

Axtell MJ, and Bowman JM (2008) Evolution of Plant microRNAs and Their Targets. Trends in Plant Science **13**: 343–49.

Axtell MJ, Jan C, Rajagopalan R, Bartel DP (2006) A Two-Hit Trigger for siRNA Biogenesis in Plants. Cell **127**: 565–77.

Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, Pamphilis C, Albert VA, Aono N, Aoyama T et al (2011) The Selaginella Genome Identifies Genetic Changes Associated with the Evolution of Vascular Plants. Science (New York, N.Y.) **332**: 960–63.

Barazesh S and McSteen P (2008) Hormonal Control of Grass Inflorescence Development. Trends in Plant Science **13**: 656–62.

Benková E, Michniewicz M, Sauer M, Teichmann T, Seifertová D, Jürgens G, Friml J (2003) Local, Efflux-Dependent Auxin Gradients as a Common Module for Plant Organ Formation. Cell **5**: 591–602

Benlloch R, Berbel A, Mislata AS, Madueno F (2007) Floral Initiation and Inflorescence Architecture: A Comparative View. Annals of Botany **100**: 659–76.

Bennett MJ, Marchant A, Green HG, May ST, Ward SP, Millner PA, Walker AR, Schulz B, Feldmann KA (1996) *Arabidopsis AUX1* Gene: A Permease-like Regulator of Root Gravitropism. Science (New York, N.Y.) **273**: 948–50.

Bergthorsson U, Adams KL, Thomason B, Palmer JD (2003) Widespread Horizontal Transfer of Mitochondrial Genes in Flowering Plants. Nature **424**: 197–201.

Besnard F, Vernoux T, Hamant O (2011) Organogenesis from Stem Cells in Planta: Multiple Feedback Loops Integrating Molecular and Mechanical Signals. Cellular and Molecular Life Sciences **68**: 2885–2906.

Betrán E, Emerson J, Kaessmann H, Long M (2004) Sex Chromosomes and Male Functions: Where Do New Genes Go? Cell Cycle (Georgetown, Tex.) **3**: 873–75.

Betrán E, and Long M (2002) Expansion of Genome Coding Regions by Acquisition of New Genes. Genetica **115**: 65–80.

Bonnet E , Peer YV, Rouzé P (2006) The Small RNA World of Plants. New Phytologist **171**: 451–68.

Booker J, Sieberer T, Wright W, Williamson L, Willett B, Stirnberg P, Turnbull C, Srinivasan M, Goddard P, Leyser O (2005) *MAX1* Encodes a Cytochrome P450 Family Member That Acts Downstream of *MAX3/4* to Produce a Carotenoid-Derived Branch-Inhibiting Hormone. Developmental Cell **8**: 443–49.

Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK (2005) Endogenous siRNAs Derived from a Pair of Natural *Cis*-Antisense Transcripts Regulate Salt Tolerance in *Arabidopsis*. Cell **123**: 1279–91.

Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S (2006) *ramosa2* Encodes a LATERAL ORGAN BOUNDARY Domain Protein That Determines the Fate of Stem Cells in Branch Meristems of Maize. The Plant Cell Online **18**: 574–85.

Bowman JL, and Eshed Y (2000) Formation and Maintenance of the Shoot Apical Meristem. Trends in Plant Science **5**: 110–15.

Bradley D (1997) Inflorescence Commitment and Architecture in *Arabidopsis*. Science **275**: 80–83.

Brewer PB, Dun EA, Ferguson BJ, Rameau C, Beveridge CA (2009) Strigolactone Acts Downstream of Auxin to Regulate Bud Outgrowth

in Pea and *Arabidopsis*. Plant Physiology **150**: 482–93.

Bridges CB (1936) The bar 'gene' a duplication Science (New York, N.Y.) **83**: 210–11.

Caetano-Anollés G, Yafremava LS, Gee H, Caetano-Anollés D, Kim HS, Mittenthal JE (2009) The Origin and Evolution of Modern Metabolism. The International Journal of Biochemistry & Cell Biology **41**: 285–97.

Carpenter R, and Coen ES (1990) Floral Homeotic Mutations Produced by Transposon-Mutagenesis in Antirrhinum Majus. Genes & Development **4**: 1483–93.

Carroll SB. (2008) Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell **134**: 25–36.

Causier B, ZSchwarz-Sommer Z, Davies B (2010) Floral Organ Identity: 20 Years of ABCs. Seminars in Cell & Developmental Biology **21**: 73–79.

Chae E, Tan QK, Hill TA, Irish VF (2008) An *Arabidopsis* F-Box Protein Acts as a Transcriptional Co-Factor to Regulate Floral Development. Development (Cambridge, England) **135**: 1235–45.

Chapman MA, Pashley CH, Wenzler J, Hvala J, Tang S, Knapp SJ, Burke JM (2008) A Genomic Scan for Selection Reveals Candidates for Genes Involved in the Evolution of Cultivated Sunflower (*Helianthus Annuus*). The Plant Cell **20**: 2931–45.

Charlesworth D, Liu FL, Zhang L (1998) The Evolution of the Alcohol Dehydrogenase Gene Family by Loss of Introns in Plants of the Genus Leavenworthia (*Brassicaceae*). Molecular Biology and Evolution **15**: 552–59.

Chen X (2004) A microRNA as a Translational Repressor of *APETALA2* in *Arabidopsis* Flower Development." Science (New York, N.Y.) **303**: 2022–25.

Chen X (2009) Small RNAs and Their Roles in Plant Development. Annual Review of Cell and Developmental Biology **25**: 21–44.

Chuck G, Cigan AM, Saeteurn K, Hake S (2007) The Heterochronic Maize Mutant *Corngrass1* Results from Overexpression of a Tandem microRNA. Nature Genetics **39**: 544–49.

Coen ES, and Meyerowitz EM (1991) The War of the Whorls: Genetic Interactions Controlling Flower Development. Nature **353**: 31–37.

Cong B, Liu J, Tanksley SD (2002) Natural Alleles at a Tomato Fruit Size Quantitative Trait Locus Differ by Heterochronic Regulatory Mutations. Proceedings of the National Academy of Sciences **99**: 13606–11.

Conti L, and Bradley D (2007) *TERMINAL FLOWER1* Is a Mobile Signal Controlling *Arabidopsis* Architecture. The Plant Cell Online **19**: 767–78.

Crawford S, Shinohara N, Sieberer T, Williamson L, George G, Hepworth J, Müller D, Domagalska MA, Leyser O (2010) Strigolactones Enhance Competition between Shoot Branches by Dampening Auxin Transport. Development (Cambridge, England) **137**: 2905–13.

Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and Functional Diversification of MIRNA Genes. The Plant Cell Online **23**: 431–42.

De Robertis EM (2008) Evo-Devo: Variations on Ancestral Themes. Cell **132**: 185–95.

Dhondt S, Hannes V, Denis VL, Veerle C, and Dirk I (2010) Plant Structure Visualization by High-Resolution X-Ray Computed Tomography. Trends in Plant Science **15**: 419–22.

Doebley J (1993) Genetics, Development and Plant Evolution. Current Opinion in Genetics & Development **3**: 865–72.

Doebley JF, Gaut BS, Smith BD (2006) The Molecular Genetics of Crop Domestication. Cell **127**: 1309–21.

Doebley J, Stec A, Hubbard L (1997) The Evolution of Apical Dominance in Maize. Nature **386**: 485–88.

Doerner P (2003) Plant Meristems: A Merry-Go-Round of Signals Review. Current Biology **13**: R368–74.

Dolgosheina EV, Morin RD, Aksay G, Sahinalp SC, Magrini V, Mardis ER, Mattsson J, Unrau PJ (2008) Conifers Have a Unique Small RNA Silencing Signature. RNA **14**: 1508–15.

Domagalska MA, and Leyser O (2011) Signal Integration in the Control of Shoot Branching. Nat Rev Mol Cell Biol **12**: 211–21.

Dorweiler J and Doebley J (1997) Developmental Analysis of *Teosinte Glume architecture1*: A Key Locus in the Evolution of Maize (Poaceae). American Journal of Botany **84**: 1313.

Drouin G, and Dover GA (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. Journal of Molecular Evolution **31**: 132–50.

Faroq AT, Adam H, Anjos A, Lorieux M, Larmande P, Ghesquière A, Jouannic S, Shahbazkia HR (2013) P-TRAP: A Panicle Trait Phenotyping Tool. BMC Plant Biology **13**: 1–14.

Finnegan EJ and Matzke MA (2003) The Small RNA World. Journal of Cell Science **116**: 4689–93.

Fornara F, Parenicová L, Falasca G, Pelucchi N, Masiero S, Ciannamea S, Lopez-Dee Z, Altamura MM, Colombo L, Kater MM (2004) Functional Characterization of *OsMADS18*, a Member of the AP1/SQUA Subfamily of MADS Box Genes. Plant Physiology **135**: 2207–19.

Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD (2000) *fw2.2:* A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size. Science (New York, N.Y.) **289**: 85–88.

Freeling M (2009) Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome, Segmental, or by Transposition. Annual Review of Plant Biology **60**: 433–53.

Friml J, Yang X, Michniewicz M, Weijers D, Quint A, Tietz O, Benjamins R, Ouwerkerd PBF, Ljung K, Sandberg G, Hooykaas PJJ, Palne K, Offringa R (2004) A PINOID-Dependent Binary Switch in Apical-Basal PIN Polar Targeting Directs Auxin Efflux RID A-2076-2010. Science **306**: 862–65.

Furner J, Pumfrey JE (1992) Cell Fate in the Shoot Apical Meristem of *Arabidopsis thalian.* Development **115**: 755–64.

Furuta T, Uehara K, Angeles-Shim RB, Shim J, Ashikari M, Takashi T (2014) Development and Evaluation of Chromosome Segment Substitution Lines (CSSLs) Carrying Chromosome Segments Derived from *Oryza rufipogon* in the Genetic Background of *Oryza sativa L*. Breeding Science **63**: 468–75.

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Pè ME, Schmidt RJ (2004) The Role of *Barren stalk1* in the Architecture of Maize. Nature **432**: 630–35.

Gao X, Liang W, Yin C, Ji S, Wang H, Su S, Guo C, Kong H, Xue H, Zhang D (2010) The *SEPALLATA*-like Gene *OsMADS34* Is Required for Rice Inflorescence and Spikelet Development. Plant Physiology **153**: 728–40.

Ghildiyal M and Zamore PD (2009) Small Silencing RNAs: An Expanding Universe. Nature Reviews Genetics **10**: 94–108.

Gilad Y, Rifkin SA, Pritchard JK (2008) Revealing the Architecture of Gene Regulation: The Promise of eQTL Studies. Trends in Genetics **24**: 408–15.

Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pagès V, Dun EA, Pillot JP, Letisse F (2008) Strigolactone Inhibition of Shoot Branching. Nature **455**: 189–94.

Gutiérrez AG, Carabalí SJ, Giraldo OX, Martínez CP, Correa F, Prado G, Tohme J, Lorieux M (2010) Identification of a Rice Stripe Necrosis Virus Resistance Locus and Yield Component QTLs Using *Oryza sativa* x *O. glaberrima* Introgression Lines. BMC Plant Biology **10:** 6.

Hake S (2008) Inflorescence Architecture: The Transition from Branches to Flowers. Current Biology **18**: R1106–8.

Hammer K (1984) Das Domestikationssyndrom. Die Kulturpflanze **32**: 11–34.

Han B and Huang X (2013) Sequencing-Based Genome-Wide Association Study in Rice. Current Opinion in Plant Biology **16**: 133–38.

Hannon GJ (2002) RNA Interference. Nature **418**: 244–51.

Harlan JR (1992) Crops and Man, 2nd Edition.

Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE (2006) Dissecting *Arabidopsis thaliana* DICER Function in Small RNA Processing, Gene Silencing and DNA Methylation Patterning. Nature Genetics **38**: 721–25.

Henz SR, Cumbie JS, Kasschau KD, Lohmann JU, Carrington JC, Weigel D, Schmid M (2007) Distinct Expression Patterns of Natural Antisense Transcripts in *Arabidopsis*. Plant Physiology **144**: 1247–55.

He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X, Greenberg AJ, Hudson RR, Wu CI, Shi S (2011) Two Evolutionary Histories in the Genome of Rice: The Roles of Domestication Genes. PLoS Genetics **7**:100-120.

Huang P, Molina J, Flowers JM, Rubinstein S, Jackson SA, Purugganan MD, Schaal BA (2012) Phylogeography of Asian Wild Rice, Oryza Rufipogon: A Genome-Wide View. Molecular Ecology **21**: 4593–4604.

Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, Xia G, Chu G, Li J, Fu X (2009) Natural Variation at the *DEP1* Locus Enhances Grain Yield in Rice. Nat Genet **41**: 494–97.

Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y et al (2012) A Map of Rice Genome Variation Reveals the Origin of Cultivated Rice. Nature **490**: 497–501.

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C et al (2010) Genome-Wide Association Studies of 14 Agronomic Traits in Rice Landraces. Nature Genetics **42**: 961–67.

Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W et al (2012) Genome-Wide Association Study of Flowering Time and Grain Yield Traits in a Worldwide Collection of Rice Germplasm. Nature Genetics **44**: 32–39.

Hwang I, Sheen J, Müller B (2012) Cytokinin Signaling Networks. Annual Review of Plant Biology **63**: 353–80.

Ikeda-Kawakatsu K, Maekawa M, Izawa T, Itoh JI, 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**: 168–80.

Ikeda K, Ito M, Nagasawa N, Kyozuka J, Nagato Y (2007) Rice *ABERRANT PANICLE ORGANIZATION 1*, Encoding an F-box Protein, Regulates Meristem Fate. The Plant Journal **51**: 1030–40.

Ikeda K, Sunohara H, Nagato Y (2004) Developmental Course of Inflorescence and Spikelet in Rice. Breeding Science **54**: 147–56.

Irish VF (2010) The Flowering of *Arabidopsis* Flower Development. The Plant Journal **61**: 1014–28.

Ishii T, Xu Y, McCouch SR (2001) Nuclear- and Chloroplast-Microsatellite Variation in A-Genome Species of Rice. Genome / National Research Council Canada **44**: 658–66.

Ishikawa S, Maekawa M, Arite T, Onishi K, Takamure I, Kyozuka J (2005) Suppression of Tiller Bud Activity in Tillering Dwarf Mutants of Rice. Plant and Cell Physiology **46**: 79–86.

Itoh JI, Nonomura KI, Ikeda K, Yamaki S, Inukai Y, Yamagishi H, Kitano H, Nagato Y (2005) Rice

Plant Development: From Zygote to Spikelet. Plant and Cell Physiology **46**: 23–47.

Izawa T, Konishi S, Shomura A, Yano M (2009) DNA Changes Tell Us about Rice Domestication. Current Opinion in Plant Biology **12**: 185–92.

Jack T (2004) Molecular and Genetic Mechanisms of Floral Control. The Plant Cell Online **16**:1–17.

Jacobs SWL, and Everett J (2000) *Grasses: systematics and evolution*. http://www.publish.csiro.au/pid/2422.htm.

Jasinski S, Piazza P, Craft J, Hay A, Woolley L, Rieu L, Phillips A, Hedden P, Tsiantis M (2005) *KNOX* Action in *Arabidopsis* Is Mediated by Coordinate Regulation of Cytokinin and Gibberellin Activities. Current Biology **15**: 1560–65.

Jeong DH, Park S, Zhai J, Gurazada SGR, De Paoli E, Meyers BC, Green PJ (2011) Massive Analysis of Rice Small RNAs: Mechanistic Implications of Regulated MicroRNAs and Variants for Differential Target RNA Cleavage. The Plant Cell Online **23**: 4185–4207.

Jeon JS, Lee S, An G (2008) Intragenic Control of Expression of a Rice MADS Box Gene *OsMADS1*. Molecules and Cells **26**: 474–80.

Jeon JS, Lee S, Ki-hong J (2000) Production of Transgenic Rice Plants Showing Reduced Heading Date and Plant Height by Ectopic Expression of rice MADS Box Gene. Molecular Breeding **6**:518-592.

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G Zeng D, Lu Z, Zhu X, Qian Q, Li J (2010) Regulation of *OsSPL14* by *OsmiR156* Defines Ideal Plant Architecture in Rice. Nature Genetics **42**: 541–44.

Jin, Hailing, Vladimir Vacic, Thomas Girke, Stefano Lonardi, and Jian-Kang Zhu. 2008. "Small RNAs and the Regulation of Cis-Natural Antisense Transcripts in *Arabidopsis*." *BMC Molecular Biology* 9: 6.

Jin YK and Bennetzen JL (1994) Integration and Nonrandom Mutation of a Plasma Membrane Proton ATPase Gene Fragment within the Bs1 Retroelement of Maize. The Plant Cell **6**: 1177–86.

Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan GL, Walbot V, Sundaresan V, Vance V, Bowman LH (2009) Clusters and Superclusters of Phased Small RNAs in the Developing Inflorescence of Rice. Genome Research **19**: 1429–40.

Kanno A, Nakada M, Akita Y, Hirai M (2007) Class B Gene Expression and the Modified ABC Model in Nongrass Monocots. The Scientific World Journal **7**: 268–79.

Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC (2007) Genome-Wide Profiling and Analysis of *Arabidopsis* siRNAs. PLoS Biol **5**: e57.

Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC (2003) P1/HC-Pro, a Viral Suppressor of RNA Silencing, Interferes with *Arabidopsis* Development and miRNA Unction. Developmental Cell **4**: 205–17.

Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Zhu JK, Staskawicz BJ, Jin H (2006) A Pathogen-Inducible Endogenous siRNA in Plant Immunity. Proceedings of the National Academy of Sciences of the United States of America **103**: 18002–7.

Kellogg EA (2007) Floral Displays: Genetic Control of Grass Inflorescences. Current Opinion in Plant Biology **10**: 26–31.

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**: 1970–83.

Khush GS (2001) Green Revolution: The Way Forward. Nature Reviews Genetics **2**: 815–22.

Kidner CA, Martienssen RA (2005) The Developmental Role of microRNA in Plants. Current Opinion in Plant Biology, Growth and development, **8**: 38–44.

Kimura M (1991) The Neutral Theory of Molecular Evolution: A Review of Recent Evidence. Idengaku Zasshi **66**: 367–86.

Kim VN (2005) Small RNAs: Classification, Biogenesis, and Function. Mol Cells **19**: 1–15.

Kirchoff BK (2003) Shape Matters: Hofmeister's Rule, Primordium Shape, and Flower Orientation. International Journal of Plant Sciences **164:** 505–17.

Kobayashi K, Maekawa M, Miyao A, Hirochika H, Kyozuka 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.

Koes R (2008) Evolution and Development of Virtual Inflorescences. Trends in Plant Science **13**: 1–3.

Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H, Shimamoto K, Kyozuka J (2003) *LAX* and *SPA*: Major Regulators of Shoot Branching in Rice. *P*roceedings of the National Academy of Sciences of the United States of America **100**: 11765–70.

Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyozuka J (2003) *FRIZZY PANICLE* Is Required to Prevent the Formation of Axillary Meristems and to Establish Floral Meristem Identity in Rice Spikelets. Development **130**: 3841–50.

Komatsu M, Maekawa M, Shimamoto K, 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**: 364–73.

Komiya R, Ohyanagi H, Niihama M, Watanabe T, Nakano M, Kurata N, Nonomura KI (2014) Rice Germline-Specific Argonaute MEL1 Protein

Binds to phasiRNAs Generated from More than 700 lincRNAs. The Plant Journal **78**:385-397

Kondo T, Sawa S, Kinoshita A, Mizuno S, Kakimoto T, Fukuda H, Sakagami Y (2006) A Plant Peptide Encoded by *CLV3* Identified by in Situ MALDI-TOF MS Analysis. Science (New York, N.Y.) **313**: 845–48.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP Caused Loss of Seed Shattering during Rice Domestication. Science (New York, N.Y.) **312**: 1392–96.

Kramer EM, Jaramillo MA, Di Stilio VS (2004) Patterns of Gene Duplication and Functional Evolution during the Diversification of the AGAMOUS Subfamily of MADS Box Genes in Angiosperms. Genetics **166**: 1011–23.

Kurakawa T, Ueda N, Maekawa M, Kobayashi K, Kojima M, Nagato Y, Sakakibara H, Kyozuka J (2007) Direct Control of Shoot Meristem Activity by a Cytokinin-Activating Enzyme. Nature **445**: 652–55.

Kyozuka J, Konishi S, Nemoto K, Izawa T, Shimamoto K (1998) Down-Regulation of *RFL*, the *FLO/LFY* Homolog of Rice, Accompanied with Panicle Branch Initiation. Proceedings of the National Academy of Sciences of the United States of America **95**: 1979–82.

Kyozuka 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 Physiology **43**: 130–35.

Laufs P, Peaucelle A, Morin H, Traas J (2004) MicroRNA Regulation of the *CUC* Genes Is Required for Boundary Size Control in *Arabidopsis* Meristems. Development (Cambridge, England) **131**: 4311–22.

Laux T, Mayer KF, Berger J, Jürgens G (1996) The *WUSCHEL* Gene Is Required for Shoot and Floral Meristem Integrity in *Arabidopsis*. Development (Cambridge, England) **122**: 87–96.

Law JA, and Jacobsen SE (2010) Establishing, Maintaining and Modifying DNA Methylation Patterns in Plants and Animals. Nature Reviews Genetics **11**: 204–20.

Lee DY 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. The Plant Journal **69**: 445–61.

Leibfried A, To JPC, Busch W, Stehling S, Kehle A, Demar M, Kieber JJ, Lohmann JU (2005) *WUSCHEL* Controls Meristem Function by Direct Regulation of Cytokinin-Inducible Response Regulators. Nature **438**: 1172–75.

Lenhard M, and Laux T (1999) Shoot Meristem Formation and Maintenance. Current Opinion in Plant Biology **2**: 44–50.

Li C, Zhou A, Sang T (2006) Rice Domestication by Reducing Shattering. Science **311**: 1936–39.

Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B (2012) MicroRNA Regulation of Plant Innate Immune Receptors. Proceedings of the National Academy of Sciences **109**: 1790–95.

Li H, Liang W, Yin C, Zhu L, Zhang D (2011) Genetic Interaction of *OsMADS3, DROOPING LEAF,* and *OsMADS13* in Specifying Rice Floral Organ Identities and Meristem Determinacy. Plant Physiology **156**: 263–74.

Linares OF (2002) African Rice (*Oryza glaberrima*): History and Future Potential. Proceedings of the National Academy of Sciences **99**: 16360–65.

Lin H, Wang R, Qian Q, Yan M, Meng X, Fu Z, Yan C, jiang B, Su Z, Li J, Wang Y (2009) *DWARF27*, an Iron-Containing Protein Required for the Biosynthesis of Strigolactones, Regulates Rice Tiller Bud Outgrowth. The Plant Cell **21**: 1512–25.

Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, Paran I, Eshed Y, Zamir D (2008) The Making of a Compound Inflorescence in Tomato and Related Nightshades. PLoS Biology **6**: e288.

Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X, Teng S, Hiroshi F, Yuan M, Luo D, Han B Li J (2003) Control of Tillering in Rice. Nature **422**: 618–21.

Li ZM, Zheng XM, Ge S (2011) Genetic Diversity and Domestication History of African Rice (*Oryza glaberrima*) as Inferred from Multiple Gene Sequences. Theoretical and Applied Genetics **123**: 21–31.

Liu C, Norman Teo ZW, Bi Y, Song S, Xi W, Yang X, Yin Z, Yu H (2013) A Conserved Genetic Pathway Determines Inflorescence Architecture in Arabidopsis and Rice. Developmental Cell **24**: 612–22.

Londo JP, Chiang YC, Hung KH, Chiang TU, Schaal BA (2006) Phylogeography of Asian Wild Rice, *Oryza rufipogon*, Reveals Multiple Independent Domestications of Cultivated Rice, *Oryza sativa.* Proceedings of the National Academy of Sciences of the United States of America **103**: 9578–83.

Long M, Betrán E, Thornton K, Wang W (2003) The Origin of New Genes: Glimpses from the Young and Old. Nature Reviews Genetics **4**: 865–75.

Malcomber ST and Kellogg EA (2004) Heterogeneous Expression Patterns and Separate Roles of the *SEPALLATA* Gene *LEAFY HULL STERILE1* in Grasses. The Plant Cell **16**: 1692–1706.

Mallo M, Wellik DM, Deschamps J (2010) *Hox* Genes and Regional Patterning of the Vertebrate Body Plan. Developmental Biology **344**: 7–15.

Mallory AC, Dugas DV, Bartel DP, Bartel B (2004) MicroRNA Regulation of NAC-Domain Targets Is Required for Proper Formation and Separation of Adjacent Embryonic, Vegetative, and Floral Organs. Current Biology **14**: 1035–46.

Mallory AC, Reinhart BJ, Bartel D, Vance VB, Bowman LH (2002) A Viral Suppressor of RNA Silencing Differentially Regulates the

Accumulation of Short Interfering RNAs and Micro-RNAs in tobacco. Proceedings of the National Academy of Sciences of the United States of America **99**: 15228–33.

Margis R, Adriana F. Fusaro, Neil A. Smith, Shaun J. Curtin, John M. Watson, E. Jean Finnegan, and Peter M. Waterhouse. 2006. "The Evolution and Diversification of Dicers in Plants." *FEBS Letters* 580 (10): 2442–50.

Martin C, Ellis N, Rook F (2010)VDo Transcription Factors Play Special Roles in Adaptive Variation? Plant Physiology **154**: 506–11.

Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJM (2009) RNA-Mediated Chromatin-Based Silencing in Plants. Current Opinion in Cell Biology **21**: 367–76.

Ma Z, Coruh C, Axtell MJ (2010) *Arabidopsis lyrata* Small RNAs: Transient MIRNA and Small Interfering RNA Loci within the *Arabidopsis* Genus. The Plant Cell **22**: 1090–1103.

McKim S and Hay A (2010) Patterning and Evolution of Floral Structures - Marking Time. Current Opinion in Genetics & Development **20**: 448–53.

McSteen P (2006) Branching Out: The Ramosa Pathway and the Evolution of Grass Inflorescence Morphology. The Plant Cell Online **18**: 518–22.

McSteen P and Leyser O (2005) Shoot Branching. Annu. Rev. Plant Biol **56**: 353–74.

McSteen P (2009) Hormonal Regulation of Branching in Grasses. Plant Physiology **149**: 46–55.

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–U102.

Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, Huang P, Jackson s, Schaal B, Bustamante CD, Boyko AR, Purugganan M (2011) Molecular Evidence for a Single Evolutionary Origin of Domesticated Rice. Proceedings of the National Academy of Sciences **108**: 8351–56.

Montes RAC, de Fátima Rosas-Cárdenas F, De Paoli E, Accerbi M, Rymarquis LA, Mahalingam G, Marsch-Martínez N, Meyers BC, Green PJ, de Folter S (2014) Sample Sequencing of Vascular Plants Demonstrates Widespread Conservation and Divergence of microRNAs. Nature Communications 5:3722.

Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of *ARGONAUTE7-miR390* Interaction and Dual Functionality in *TAS3* Trans-Acting siRNA Formation. Cell **133**: 128–41.

Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A (2005) Gene Duplication and Exon Shuffling by Helitron-like Transposons Generate Intraspecies Diversity in Maize. Nature Genetics **37**: 997–1002.

Morita Y and Kyozuka J (2007) Characterization of *OsPID*, the Rice Ortholog of *PINOID*, and Its Possible Involvement in the Control of Polar Auxin Transport. Plant and Cell Physiology **48**: 540–49.

Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, Baulcombe DC (2009) Uniparental Expression of PolIV-Dependent siRNAs in Developing Endosperm of *Arabidopsis*. Nature **460**: 283–86.

Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F (2010) LEAFY Blossoms. Trends in Plant Science **15**: 346–52.

Müller R, Bleckmann A, Simon R (2008) The Receptor Kinase CORYNE of *Arabidopsis* Transmits the Stem Cell-Limiting Signal *CLAVATA3* Independently of *CLAVATA1*. The Plant Cell **20**: 934–46.

Murray JAH, Jones A, Godin C, Traas J (2012) Systems Analysis of Shoot Apical Meristem Growth and Development: Integrating Hormonal and Mechanical Signaling. The Plant Cell Online **24**: 3907–19.

Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H, Nagato Y (2003) *SUPERWOMAN1* and *DROOPING LEAF* Genes Control Floral Organ Identity in Rice. Development **130**: 705–18.

Nakagawa M, Shimamoto K, Kyozuka J (2002) Overexpression of *RCN1* and *RCN2*, Rice *TERMINAL FLOWER 1/CENTRORADIALIS* Homologs, Confers Delay of Phase Transition and Altered Panicle Morphology in Rice. The Plant Journal **29**: 743–50.

Nei M, Rogozin IB, Piontkivska H (2000) Purifying Selection and Birth-and-Death Evolution in the Ubiquitin Gene Family. Proceedings of the National Academy of Sciences **97**: 10866–71.

Nesbitt TC and Tanksley SD (2002) Comparative Sequencing in the Genus Lycopersicon: Implications for the Evolution of Fruit Size in the Domestication of Cultivated Tomatoes. Genetics **162**: 365–79.

Molinero-Rosales N and Jamilena M (2000) *FALSIFLORA*, the Tomato Orthologue of *FLORICAULA* and *LEAFY,* Controls Flowering Time and Floral Meristem Identity. The Plant Journal **20**: 685–93.

Nobuta K, Lu C, Shrivastava R, Pillay M, De Paoli E, Accerbi M, Arteaga-Vazquez M, Sidorenko L, Jeong D, Yen Y, Green P, Chandler VL, Meyers BC (2008) Distinct Size Distribution of Endogeneous siRNAs in Maize: Evidence from Deep Sequencing in the *mop1-1* Mutant. Proceedings of the National Academy of Sciences of the United States of America **105**: 14958–63.

Nordborg M and Weigel D (2008) Next-Generation Genetics in Plants. Nature **456**: 720–23.

Nurminsky DL, Nurminskaya MV, Aguiar D, Hartl DL (1998) Selective Sweep of a Newly Evolved Sperm-Specific Gene in Drosophila. Nature **396**: 572–75.

Ochman H (2001) Lateral and Oblique Gene Transfer. Current Opinion in Genetics & Development **11**: 616–19.

Oikawa T and Kyozuka J (2009) Two-Step Regulation of *LAX PANICLE1* Protein Accumulation in Axillary Meristem Formation in Rice. The Plant Cell Online **21**: 1095–1108.

Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD (2006) Selection Under Domestication: Evidence for a Sweep in the Rice Waxy Genomic Region. Genetics **173**: 975–83.

Palatnik JF, Wollmann H, Schommer C, Schwab R, Boisbouvier J, Rodriguez R, Warthmann N? Allen E, Dezulian T, Huson D, Carrington JC, Weigel D (2007) Sequence and Expression Differences Underlie Functional Specialization of *Arabidopsis* microRNAs *miR159* and *miR319*. Developmental Cell **13**: 115–25.

Paterson AH (2002) What Has QTL Mapping Taught Us about Plant Domestication? New Phytologist **154**: 591–608.

Patthy L (1996) Exon Shuffling and Other Ways of Module Exchange. Matrix Biology: Journal of the International Society for Matrix Biology **15**: 301–10

Patthy L (1999) Genome Evolution and the Evolution of Exon-Shuffling : a Review. Gene **238**: 103–14.

Pautler M, Tanaka W, Hirano HY, Jackson D (2013) Grass Meristems I: Shoot Apical Meristem Maintenance, Axillary Meristem Determinacy and the Floral Transition. Plant and Cell Physiology **54**: 302–12.

Pearl SA, Bowers JE, Reyes-Chin-Wo S, Michelmore RW, Burke JM (2014) Genetic Analysis of Safflower Domestication. BMC Plant Biology **14**: 43.

Peng S, Khush GS, Virk P, Tang Q, Zou Y (2008) Progress in Ideotype Breeding to Increase Rice Yield Potential. Field Crops Research **108**: 32–38.

Pina SD, Souer E, Koes R (2014) Arguments in the Evo-Devo Debate: Say It with Flowers! Journal of Experimental Botany **65**: 2231-42

Prasad K, Parameswaran S, 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. The Plant Journal **43**: 915–28.

Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E (2007) Evolution and Development of Inflorescence Architectures. Science **316**: 1452–56.

Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ (2006) Distinct Catalytic and Non-Catalytic Roles of *ARGONAUTE4* in RNA-Directed DNA Methylation. Nature **443**: 1008–12.

Rao RR, Prasad K, Kumar PR, Vijayraghavan U (2008) Distinct Regulatory Role for *RFL*, the Rice *LFY* Homolog, in Determining Flowering Time and Plant Architecture. Proceedings of the National Academy of Sciences **105**: 3646–51.

Rastogi S and Liberles DA (2005) Subfunctionalization of Duplicated Genes as a Transition State to Neofunctionalization. BMC Evolutionary Biology **5**: 28.

Rebocho AB, Bliek M, Kusters E, Castel R, Procissi A, Roobeek I, Souer E, Koes R (2008) Role of *EVERGREEN* in the Development of the Cymose Petunia Inflorescence. Developmental Cell **15**: 437–47.

Reinhardt D, Pesce ER, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C (2003) Regulation of Phyllotaxis by Polar Auxin Transport. Nature **426**: 255–60.

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in Plants. Genes & Development **16**: 1616–26.

Remizowa MV, Rudall PJ, Choob VV, Sokoloff DD (2013) Racemose Inflorescences of Monocots: Structural and Morphogenetic Interaction at the Flower/inflorescence Level. Annals of Botany **112**: 1553–66.

Rijpkema AS, Vandenbussche M, Koes R, Heijmans K, Gerats T (2010) Variations on a Theme: Changes in the Floral ABCs in Angiosperms. Seminars in Cell & Developmental Biology **21**: 100–107.

Ritter MK, Padilla CM, Schmidt RJ (2002) The Maize Mutant *Barren stalk1* Is Defective in Axillary Meristem Development. American Journal of Botany **89**: 203–10.

Ron M, Saez MA, Williams LE, Fletcher JC, McCormick S (2010) Proper Regulation of a Sperm-Specific Cis-Nat-siRNA Is Essential for Double Fertilization in *Arabidopsis*. Genes & Development **24**: 1010–21.

Samach A, Klenz JE, Kohalmi SE, Risseeuw E, Haughn GW, Crosby WL (1999) The *UNUSUAL FLORAL ORGANS* Gene of *Arabidopsis thaliana* Is an F-Box Protein Required for Normal Patterning and Growth in the Floral Meristem. The Plant Journal: For Cell and Molecular Biology **20**: 433–45.

Sarla N and Swamy BPM (2005) *Oryza glaberrima*: A Source for the Improvement of *Oryza sativa*. Current Science **89**: 955–63.

Satoh-Nagasawa N, Nagasawa N, Malcomber S, Sakai H, Jackson D (2006) A Trehalose Metabolic Enzyme Controls Inflorescence Architecture in Maize. Nature **441**: 227–30.

Schoof H, Lenhard M, Haecker A, Mayer KF, Jürgens G, Laux T (2000) The Stem Cell Population of *Arabidopsis* Shoot Meristems in Maintained by a Regulatory Loop between the *CLAVATA* and *WUSCHEL* Genes. Cell **100**: 635–44.

Second G (1982) Origin of the Genic Diversity of Cultivated Rice (*Oryza* Spp.): Study of the Polymorphism Scored at 40 Isozyme Loci. Jpn J Genet **57**: 25–57.

Shannon, S, and DR Meeks-Wagner. 1991. "A Mutation in the *Arabidopsis* TFL1 Gene Affects Inflorescence Meristem Development." *The Plant Cell* **3** (9): 877–92.

Shigyo M, Hasebe M, Ito M (2006) Molecular Evolution of the *AP2* Subfamily. Gene **366**: 256–65.

Shivaprasad PV, Chen HM, Patel K, Bond DM, Santos BACM, Baulcombe DC (2012) A MicroRNA Superfamily Regulates Nucleotide Binding Site–Leucine-Rich Repeats and Other mRNAs. The Plant Cell Online **24**: 859–74.

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**: 547–55.

Slack F, and Ruvkun G (1998) Heterochronic Genes in Development and Evolution. Biological bulletin rarine biological laboratory **195**: 375–76.

Smartt J and Simmonds N (1995) Evolution of Crop Plants. 2 edition. Wiley-Blackwell.

Soltis DE, Ma H, Frohlich MW, Soltis PS, Albert VA, Oppenheimer DG, Altman NS, de Pamphilis C, Leebens-Mack J (2007) The Floral Genome: An Evolutionary History of Gene Duplication and Shifting Patterns of Gene Expression. Trends in Plant Science **12**: 358–67.

Song X, Li P, Zhai J, Zhou M, Ma L, Liu B, Jeong DH, Nakano M, Cao S, Liu C, Chu C, Wang XJ, Green PJ, Meyers BC, Cao X (2012) Roles of *DCL4* and *DCL3b* in Rice Phased Small RNA Biogenesis: Phased Small RNA Clusters Biogenesis in Rice. The Plant Journal **69**: 462–74.

Song X, Wang D, Ma L, Chen Z, Li P, Cui X, Liu C, Cao S, Chu C, Tao Y, Cao X (2012) Rice RNA-Dependent RNA Polymerase 6 Acts in Small RNA Biogenesis and Spikelet Development. The Plant Journal **71**: 378–89.

Souer E, Rebocho AB, Bliek M, Kusters E, de Bruin RAM, 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. The Plant Cell **20**: 2033–48.

Souer E, Krol A, Kloos D, Spelt C, Bliek M, Mol J, Koes R (1998) Genetic Control of Branching Pattern and Floral Identity during Petunia Inflorescence Development. Development **125**: 733–42.

Souer E, Houwelingen A, Kloos D, Mol J, Koes R (1996) The No Apical Meristem Gene of Petunia Is Required for Pattern Formation in Embryos and Flowers and Is Expressed at Meristem and Primordia Boundaries." Cell **85**: 159–70.

Sunkar R and Zhu JK (2004) Novel and Stress-Regulated microRNAs and Other Small RNAs from *Arabidopsis*. The Plant Cell **16**: 2001–19.

Sweeney MT, Thomson MJ, Pfeil BE, McCouch S (2006) Caught Red-Handed: *Rc* Encodes a Basic Helix-Loop-Helix Protein Conditioning Red Pericarp in Rice. The Plant Cell Online **18**: 283–94.

Tabuchi H, Zhang Y, Hattori S, Omae M, Shimizu-Sato S, Oikawa T, Qian Q, nishimura M, Kitano H, Xie H, Fang X, Yoshida H, Kyozuka J, Chen F, Sato Y (2011) *LAX PANICLE2* of Rice Encodes a Novel Nuclear Protein and Regulates the Formation of Axillary Meristems. The Plant Cell Online **23**: 3276–87.

Takeda T, Suwa Y, Suzuki M, Kitano H, Tanaka MU, Ashikari M, Matsuoka M, Ueguchi C (2003) The *OsTB1* Gene Negatively Regulates Lateral Branching in Rice. The Plant Journal **33**: 513–20.

Talmor-Neiman M, Stav R, Klipcan L, Buxdorf K, Baulcombe DC, Arazi T (2006) Identification of Trans-Acting siRNAs in Moss and an RNA-Dependent RNA Polymerase Required for Their Biogenesis. The Plant Journal: For Cell and Molecular Biology **48**: 511–21.

Tang G, Reinhart BJ, Bartel DP, Zamore PD (2003) A Biochemical Framework for RNA Silencing in Plants. Genes & Development **17**: 49–63.

Tang H, Sezen U, Paterson AH (2010) Domestication and Plant Genomes. Current Opinion in Plant Biology **13**: 160–66.

Teo ZWN, Song S, Wang YQ, Liu J, Yu H (2014) New Insights into the Regulation of Inflorescence Architecture. Trends in Plant Science **19**: 158–65.

Theissen G and Saedler H (2001) Plant Biology. Floral Quartets. Nature **409**: 469–71.

Umehara M, Hanada A, Yoshida S, Akiyama K, Arite T, Takeda-Kamiya N, Magome H, Kamiya Y, Shirasu K, Yoneyama K, Kyozuka J, Yamaguchi S (2008) Inhibition of Shoot Branching by New Terpenoid Plant Hormones. Nature **455**: 195–200.

Valencia S, Antonio M, Liu J, Hannon GJ, Parker R (2006) Control of Translation and mRNA Degradation by miRNAs and siRNAs. Genes & Development **20**: 515–24.

Vaucheret H (2008) Plant ARGONAUTES. Trends in Plant Science **13**: 350–58.

Vaughan DA, Lu BR, Tomooka N (2008) The Evolving Story of Rice Evolution RID C-3130-2009. Plant Science **174**: 394–408.

Vaughan DA, Morishima H, Kadowaki K (2003) Diversity in the Oryza Genus. Current Opinion in Plant Biology **6**: 139–46.

Veron AS, Kaufmann K, Bornberg-Bauer E (2007) Evidence of Interaction Network Evolution by Whole-Genome Duplications: A Case Study in MADS-Box Proteins. Molecular Biology and Evolution **24**: 670–78.

Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K et al (2010) Genome Sequencing and Analysis of the Model Grass Brachypodium Distachyon. Nature **463**: 763–68.

Voinnet O (2009) Origin, Biogenesis, and Activity of Plant MicroRNAs. Cell **136**: 669–87.

Vollbrecht E, Springer PS, Goh L, Iv ESB, Martienssen R (2005) Architecture of Floral Branch Systems in Maize and Related Grasses. Nature **436**: 1119–26.

Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, Bomblies-Yant K, Lukens L, Doebley J (2005) The Origin of the Naked Grains of Maize. Nature **436**: 714–19.

Wang RL, Stec A, Hey J, Lukens L, Doebley J (1999) The Limits of Selection during Maize Domestication. Nature **398**: 236–39.

Wang W, Brunet FG, Nevo E, Long M (2002) Origin of Sphinx, a Young Chimeric RNA Gene in Drosophila Melanogaster. Proceedings of the National Academy of Sciences of the United States of America **99**: 4448–53.

Wang Y and Li J (2008) Molecular Basis of Plant Architecture. Annual Review of Plant Biology **59**: 253–79.

Wang Y and Li J (2011) Branching in Rice. Current Opinion in Plant Biology **14**: 94–99.

Wang ZY, Zheng FQ, Shen GZ, Gao JP, Snustad DP, Li MG, Zhang JL, Hong MM (1995) The Amylose Content in Rice Endosperm Is Related to the Post-Transcriptional Regulation of the Waxy Gene. The Plant Journal: For Cell and Molecular Biology **7**: 613–22.

Waters MT, Brewer PB, Bussell JD, Smith SM, Beveridge CA (2012) The *Arabidopsis* Ortholog of Rice *DWARF27* Acts Upstream of *MAX1* in the Control of Plant Development by Strigolactones. Plant Physiology **159**: 1073–85.

Weberling F (1989) Morphology of Flowers and Inflorescences. Cambridge University Press.

Weigel D and Meyerowitz EM (1994) The ABCs of Floral Homeotic Genes. Cell **78**: 203–9.

Weir I, Lu J, Cook H, Causier B, Schwarz-Sommer Z, Davies B (2004) *CUPULIFORMIS* Establishes Lateral Organ Boundaries in Antirrhinum. Development (Cambridge, England) **131**: 915–22.

Woods DP, Hope CL, Malcomber ST (2011) Phylogenomic Analyses of the *BARREN STALK1/LAX PANICLE1 (BA1/LAX1)* Genes and Evidence for Their Roles During Axillary Meristem Development. Molecular Biology and Evolution **28**: 2147–59.

Wray GA (2007) The Evolutionary Significance of *Cis*-Regulatory Mutations. Nature Reviews Genetics **8**: 206–16.

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA (2003) The Evolution of Transcriptional Regulation in Eukaryotes. Molecular Biology and Evolution **20**: 1377–1419.

Winnepenninckx B, Backeljau T, De Wachter R (1993) Extraction of High Molecular Weight DNA from *Molluscs*. Trends in Genetics **9**: 407.

Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y (2010) DNA Methylation Mediated by a microRNA Pathway. Molecular Cell **38**: 465–75.

Xiao, Han, Ning Jiang, Erin Schaffner, Eric J. Stockinger, and Esther van der Knaap (2008) A Retrotransposon-Mediated Gene Duplication Underlies Morphological Variation of Tomato Fruit. Science (New York, N.Y.) **319**: 1527–30.

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and Functional Diversification of Small RNA Pathways in Plants. PLoS Biol **2**: e104.

Xing Y and Zhang Q (2010) Genetic and Molecular Bases of Rice Yield. Annual Review of Plant Biology **61**: 421–42.

Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X, Zhang Q (2008) Natural Variation in *Ghd7* Is an Important Regulator of Heading Date and Yield Potential in Rice. Nat Genet **40**: 761–67.

Xu M, Zhu L, Shou H, Wu P (2005) A *PIN1* Family Gene, *OsPIN1*, Involved in Auxin-Dependent Adventitious Root Emergence and Tillering in Rice. Plant and Cell Physiology **46**: 1674–81.

Yamaki S, Miyabayashi T, Eiguchi M, Kitano H, Nonomura KI, Kurata N (2010) Diversity of Panicle Branching Patterns in Wild Relatives of Rice. Breeding Science **60**: 586–96.

Yanai O, Shani E, Doleza K, Tarkowski P, Sablowski R, Sandberg G, Samach A, Ori N (2005) *Arabidopsis* KNOXI Proteins Activate Cytokinin Biosynthesis. Current Biology: CB **15**: 1566–71.

Yang T, Xue L, An L (2007) Functional Diversity of miRNA in Plants. Plant Science **172**: 423–32.

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, Chen R, Yamazaki R, Tokunaga H, Kitaguchi Y, Sato, Nagamura Y, Ushijima T, Kumamaru T, Iida S, Maekawa M, Kyozuka J (2012) *TAWAWA1*, a Regulator of Rice Inflorescence Architecture, Functions through the Suppression of Meristem Phase Transition. Proceedings of the National Academy of Sciences **110**: 767–72.

Yoshida H and Nagato Y (2011) Flower Development in Rice. Journal of Experimental Botany **62**: 4719-4730.

Zalewski W, Galuszka P, Gasparis S, Orczyk W, Nadolska-Orczyk A (2010) Silencing of the *HvCKX1* Gene Decreases the Cytokinin Oxidase/dehydrogenase Level in Barley and Leads to Higher Plant Productivity. Journal of Experimental Botany **61**: 1839–51.

Zazimalova E, Krecek P, Skupa P, Hoyerova K, Petrasek J (2007) Polar Transport of the Plant Hormone Auxin - the Role of PIN-FORMED (PIN) Proteins RID C-8240-2009. Cellular and Molecular Life Sciences **64**: 1621–37.

Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA (2006) Conservation and Divergence of Plant microRNA Genes. The Plant Journal **46**: 243–59.

Zhang D and Zheng Y (2014) Molecular Control of Grass Inflorescence Development. Annual Review of Plant Biology **65**: 553–78.

Zhang L, Zhao YL, Gao LF, Zhao GY, Zhou RH, Zhang BS, Jia JZ (2012) *TaCKX6-D1*, the Ortholog of Rice *OsCKX2*, Is Associated with Grain Weight in Hexaploid Wheat. The New Phytologist **195**: 574–84.

Zhang LB, Zhu Q, Wu ZQ, Ross-Ibarra J, Gaut BS, Ge S, Sang T (2009) Selection on Grain Shattering Genes and Rates of Rice Domestication." *The New Phytologist* **184** (3): 708–20.

Zhang X, Xia J, Lii YE, Barrera-Figueroa BE, Zhou X, Gao S, Lu L et al (2012) Genome-Wide Analysis of Plant Nat-siRNAs Reveals Insights into Their Distribution, Biogenesis and Function. Genome Biology **13** (3): R20.

Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung A, Bustamante CD, McCouch S (2011) Genome-Wide Association Mapping Reveals a Rich Genetic Architecture of Complex Traits in *Oryza sativa*. Nature Communications **2**: 467.

Zou J, Chen Z, Zhang S, Zhang W, Jiang G, Zhao X, Zhai X, Pan X, Zhu L (2005) Characterizations and Fine Mapping of a Mutant Gene for High Tillering and Dwarf in Rice *(Oryza sativa L.*). Planta **222**: 604–12.

ANNEXES

Table 1: List of landmark genes controling rice panicle architecture and their orthologs in other species. TF: Transcription factor; AM: Axillary meristem; SM: Spikelet meristem; FM: Floret meristem; IM: Inflorescence meristem

Gene name	Function	Gene product	Arabidopsis ortholog	Maize ortholog	Petunia ortholog	Tomato ortholog	References
Tillering							
MOC1 (MONOCULM1)	Tiller initiation/maintenance	GRAS TF	*LATERAL SUPPRESSOR (*LAS*)*			*LATERAL SUPPRESSOR (*LS*)*	McSteen and Leyser 2005; Ashikari et al. 2005; Li et al. 2003
OsTB1	Downstream *MOC1*	TCP domain	*BRANCHED1 (BRC1)*	*Teosinte branched1* (*tb1*)			Takeda et al. 2003. Doebley et al. 1997
DWARF (D)	AMs maintenance		*MORE AXILARY BRANCHING (MAX)*				Zou et al. 2005; Ishikawa et al. 2005 ; Wang and Li 2011; Waters et al. 2012
Axillary meristem initiation							
OSH1	Meristem maintenance	Homeodomain TF	*SHOOT MERISTEMLESS (STM)*	*KNOTTED-1* (*KN1*)			Tsuda et al. 2011
LAX1 (LAX PANICLE1)	AM initiation	Basic helix-loop-helix (bHLH) TF	*LARY MERISTEM FORMATION (ROX)*	*BARREN STALK1 (BA1)*			Komatsu et al. 2003 ; Komatsu et al. 2001 ; Gallavotti et al. 2004
LAX2 (LAX PANICLE2)	AM formation, maintenance	Nuclear protein with a plant-specific conserved domain					Tabuchi et al. (2011)
OsPIN1	Auxin transport	Auxin efflux carrier	*PIN1*	*ZmPIN1*			Xu et al. 2005
OsPINOID	Regulates auxin transport (regulate the localization of PIN1 protein)	Serine/threonine protein kinase	*PINOID*	barren inflorescence2 (bif2)			Morita and Kyozuka (2007)
FRIZZY PANICLE (FZP)	AM formation	ethylene-responsive element binding factor (ERF)	PUCHI	*BRANCHED SILKLESS1*			Komatsu et al. 2001
Axillary meristem outgrowth							
ABERRANT PANICLE ORGANIZATION 1 (APO1)	Transition from IM to SM; negative regulator of SM fate	F-box protein	*UNUSUAL FLORAL ORGAN* (*UFO*)		*DOUBLE-TOP (DOT)*		Ikeda et al. 2007
ABERRANT PANICLE ORGANIZATION 2 (APO2)	Transition from IM to SM	*LEAFY TF*	*LEAFY (LFY)*	zfl1, zfl2	*ABERRANT LEAF AND FLOWER (ALF)*	*FALSIFLORA (FA)*	J Kyozuka et al. 1998; Rao et al. 2008

TAWAWA1 (TAW1)	Suppressing the transition from IM to SM	nuclear protein belonging to the ALOG family				Yoshida et al 2012
ABERRANT SPIKELET AND PANICLE1 (ASP1)	Transition from IM to SM	TOPLESS-related transcriptional co-repressor	TOPLESS (TPL)	ramosa enhancer locus2 (rel2)		Yoshida et al. 2012
LEAFY HULL STERILE (LHS1)/OsMADS1	Transition from IM to FM; identity of the palea and lemma	subgroup of *LOF-SEP* genes	SEP-like			Khanday et al. 2013
GRAIN NUMBER1 (Gn1a)	cytokinin accumulation	cytokinin oxidase/dehydrogenase (*OsCKX2*), an enzyme that degrades cytokinin	CYTOKININ OXIDASE (CKX)			Ashikari et al. 2005
LONELY GUY (LOG)	cytokinin accumulation	enzyme that catalysis the final step of cytokinin biosynthesis within meristem				Kurakawa et al. 2007
DENSE AND ERECT PANICLE (DEP1)	Regulation branching	PEBP domain protein				Huang et al. 2009
OsSPL14 (IPA1/WFP)	Regulation braching	SBP-box (SQUAMOSA promoter binding protein-like) protein TF	*SPL9*	*TSH4 (Tasselseed4)*	*SISPL9*	Jiao et al. 2010 Miura et al. 2010
Osa-miR156	Target OsSPL14	MicroRNA miR156	*miR156*	*miR156*		Jiao et al. 2010 Miura et al. 2010
Osa-miR529	Target OsSPL14	MicroRNA miR529	*miR529*	*miR529*		Jeong et al. 2012
Spikelet differentiation						
Reduced Culm Number1 and 2 (RCN1, RCN2)	Regulate flowering time		*TERMINAL FLOWER1 (TFL1)/CENTRORADIALIS (CEN)*			Nakagawa et al. 2002
Ghd7	Regulate flowering time	a CCT-domain protein				Xue et al. (2008)
INDETERMINATE SPIKELET 1 (OsIDS1)	FM determinacy	APETALA2 TF (A class)	TARGET OF EAT1 (TOE1), TOE2, TOE3	indeterminate spikelet1 (ids1)/ Tasselseed6 (Ts6)		Chuck et al. 2007
SUPERNUMERARY BRACT (SNB)	FM determinacy	APETALA2 TF (A class)	TARGET OF EAT1 (TOE1), TOE2, TOE3	sister of indeterminate spikelet1 (sid1)		Chuck et al. 2007
Osa-miR172	Regulate SNB and OsIDS	MicroRNA miR172	miR172	tasselseed4 (ts4)		Chuck et al. 2007; Lee et al. 2010
PANICLE PHYTOMER (PAP2/OsMADS34)	SM and FM determinacy	SEPALLATA (SEP) family	SFP			Gao et al. 2010
OsMADS3	FM determinacy	MADS box TF (C class)	AGAMOUS (AG)	zmm2, zmm23		Kyozuka and Shimamoto 2002
OsMADS58	FM determinacy	MADS box TF (C class)	AGAMOUS (AG)	zag1		Yamaguchi et al. 2006
MFO1/ OsMADS6	FM determinacy	MADS-box TF	AGL6-like	bearded-ear		Ohmori et al. 2009

				(bde)/zag3	
OsMADS17	FM determinacy	MADS-box TF	AGL6-like		Ohmori et al. 2009
OsMADS14	FM determinacy	APETALA1 MADS box TF (A class)	APETALA1 (A class)		Fornara et al. 2004; Jeon et al. 2000
OsMADS18	FM determinacy	APETALA1 MADS box TF (A class)	APETALA1		Fornara et al. 2004; Jeon et al. 2000

Floral organ patterning

SUPERWOMAN1 (SPW1)	establishment of lodicule and stamen	*APETALA3* (*AP3*) (B class)	*APETALA3* (*AP3*)		Nagasawa et al. 2003
OsMADS2, OsMADS4	establishment of lodicule and stamen	*PISTILLATA* (*PI*) (B class)	*PISTILLATA* (*PI*)		Yao et al. 2008
DROOPING LEAF (DL)	lemma specification	YABBY protein family	*CRABS CLAW* (*CRC*),	ZmDL	Nagasawa et al. 2003
OsMADS13	Specification of ovule	MADS box TF (D class)	SEEDSTICK (STK)	zag2	Li et al. 2011

Accession Name	Local ID	nb of panicles	nb of repeat(s)	GBS data	Origin-Ecosystem
Tép hải phòng	G1	15	3	✓	VIETNAM-0
Tà cô lào cai	G2	1	1	✓	VIETNAM-0
An tu đỏ vỏ	G3	4	1	✓	VIETNAM-0
Nhông đỏ hải dương	G4	5	1	✓	VIETNAM-0
Nhông trắng hải phòng	G5	5	1	✓	VIETNAM-0
Sớm giai hưng yên	G6	5	1	✓	VIETNAM-0
Tẻ trắng hòa bình	G7	2	1	✓	VIETNAM-0
Chọn từ 502 học viện	G8	2	1	✓	VIETNAM-0
Lốc trắng sớm plei cầu	G9	5	1	✓	VIETNAM-0
Tám tròn hải dương	G11	10	2	✓	VIETNAM-0
Tám đen hải phòng	G13	10	2	✓	VIETNAM-0
Nếp thanh tâm hải phòng	G15	5	1		VIETNAM-0
Nếp gà gáy hải dương	G17	5	1	✓	VIETNAM-0
Nếp quýt hải dương	G18	5	1	✓	VIETNAM-0
Gié trắng hòa bình	G21	5	1	✓	VIETNAM-0
Trứng trắng tuyên quang	G22	3	1	✓	VIETNAM-0
Nếp đo	G23	4	1		VIETNAM-0
Tám xoan hải hậu	G24	14	3	✓	VIETNAM-0
Nếp vàng ong lạc sơn HB	G25	4	1	✓	VIETNAM-0
Đoàn kết	G27	5	1	✓	VIETNAM-0
Khẩu pe lạnh	G28	15	3		VIETNAM-0
Nếp con	G29	5	1	✓	VIETNAM-4
Nàng đùm	G34	5	1		VIETNAM-0
Nếp cẩm	G37	9	2	✓	VIETNAM-2
Nếp cẩm	G39	9	3	✓	VIETNAM-2
Nếp bà lão	G46	8	2	✓	VIETNAM-6
Lúa ngoi	G48	15	3	✓	VIETNAM-6
DT10	G49	5	1	✓	VIETNAM-0
Lúa nếp 3 tháng dạng 1	G50	10	2	✓	VIETNAM-4
Ba trăng hướng	G51	5	1	✓	VIETNAM-4
Lúa can đỏ	G53	6	2	✓	VIETNAM-0
Lúa hẻo	G55	5	1		VIETNAM-4
Nếp ghim hương	G57	10	2	✓	VIETNAM-2
Nếp hương lăng	G58	5	1	✓	VIETNAM-6
Nếp mậm	G59	10	2	✓	VIETNAM-0
Chiêm đỏ	G63	15	3	✓	VIETNAM-0
Nước mặn dạng 1	G65	6	2	✓	VIETNAM-3

Accession Name	Local ID	nb of panicles	nb of repeat(s)	GBS data	Origin-Ecosystem
Nếp 3 tháng	G68	5	1	✓	VIETNAM-1
Cốc mọi dạng 1	G69	4	1	✓	VIETNAM-2
Cốc mọi dạng 2	G70	5	1	✓	VIETNAM-2
Lúa cang dạng 1	G72	6	2	✓	VIETNAM-2
Lúa cang dạng 2	G73	11	3	✓	VIETNAM-2
Nếp quạ có râu dạng 2	G74	5	1	✓	VIETNAM-2
Nếp quạ	G75	5	1		VIETNAM-2
Cang kiến dạng 1	G77	4	1	✓	VIETNAM-2
Cang kiến dạng 2	G78	11	2	✓	VIETNAM-2
Lúa đá dạng 2	G79	5	1	✓	VIETNAM-2
Ba ktong	G80	10	2	✓	VIETNAM-4
Ba đỏ dạng 1	G81	5	1	✓	VIETNAM-4
Ba dú dạng 2	G82	5	1		VIETNAM-4
Nếp vàng	G83	5	1	✓	VIETNAM-4
Ba chơ k'tê	G84	10	2	✓	VIETNAM-0
Chành trụi	G85	15	3	✓	VIETNAM-6
Khẩu pan pua	G87	9	2	✓	VIETNAM-0
Ble mạ mùa	G88	10	2	✓	VIETNAM-0
Khẩu bò khá	G89	5	1	✓	VIETNAM-0
Blào clía	G90	5	1	✓	VIETNAM-4
Blào cố kén	G91	10	2	✓	VIETNAM-4
Blào cô cẩm	G92	5	1	✓	VIETNAM-4
Pờ lề pờ lầu xá	G93	15	3	✓	VIETNAM-4
Lúa đỏ	G94	13	3	✓	VIETNAM-4
Lúa chăm	G95	9	2	✓	VIETNAM-2
Chiêm rong	G96	10	2	✓	VIETNAM-1
Tám thơm	G97	5	1	✓	VIETNAM-1
Ngoi tía	G98	15	3	✓	VIETNAM-3
Lúa chăm biển	G99	5	1	✓	VIETNAM-2
Khẩu quại dạng 2	G100	10	2	✓	VIETNAM-4
Dieo kbin	G101	5	1	✓	VIETNAM-0
Tzo koh dạng 2	G102	15	3	✓	VIETNAM-0
Cu pủa dạng 1	G103	5	1	✓	VIETNAM-0
Cu pủa dạng 2	G104	5	1	✓	VIETNAM-0
Nếp thái lan	G105	14	3	✓	VIETNAM-0
Nếp hải hậu	G106	5	1	✓	VIETNAM-1
Nếp thái bình lùn	G107	5	1	✓	VIETNAM-1

Accession Name	Local ID	nb of panicles	nb of repeat(s)	GBS data	Origin-Ecosystem
Tám ấp bẹ	G108	5	1	✓	VIETNAM-1
Mành gié	G109	5	1	✓	VIETNAM-4
Rằn trắng	G110	5	1	✓	VIETNAM-4
Nếp rẫy	G111	5	1	✓	VIETNAM-4
Nàng thiệt	G113	5	1	✓	VIETNAM-1
Koi lòi	G115	7	2	✓	VIETNAM-0
Koi pu	G116	5	1		VIETNAM-0
Khảo sang	G117	9	2	✓	VIETNAM-4
L26	G119	5	1	✓	VIETNAM-0
Nếp lào trắng	G122	5	1		VIETNAM-4
Bao thai lùn	G123	14	3		VIETNAM-1
Nếp đen	G124	5	1	✓	VIETNAM-1
Nếp nương	G125	15	3	✓	VIETNAM-1,4
Khâu đắm đòi	G126	5	1	✓	VIETNAM-4
Khẩu bao thai	G127	5	1		VIETNAM-1
Khâu đăm	G128	9	2	✓	VIETNAM-4
Lc 93-2	G129	10	2	✓	VIETNAM-4
Lúa da bò	G130	5	1	✓	VIETNAM-4
Padai long khánh	G131	5	1	✓	VIETNAM-4
A 330	G133	15	3	✓	VIETNAM-1
Padai caloc	G134	5	1	✓	VIETNAM-4
Padai dlùc	G135	15	3	✓	VIETNAM-4
Lúa thanh trà	G142	5	1	✓	VIETNAM-2
Nàng loan hạt tròn	G146	5	1	✓	VIETNAM-2
Nếp địa phương	G150	5	1	✓	VIETNAM-1
Lốc sớm	G152	5	1	✓	VIETNAM-1,4
Tẻ nương	G153	5	1	✓	VIETNAM-4
Nếp thơm	G154	14	3	✓	VIETNAM-1
Khẩu pe lạnh	G155	14	3	✓	VIETNAM-4
Lúa k	G156	5	1	✓	VIETNAM-0
Sợ crioong	G157	10	2	✓	VIETNAM-0
Va tai ana acu	G158	3	1	✓	VIETNAM-4
Bn 1	G161	5	1	✓	VIETNAM-1
Giống 90 ngày	G165	10	2	✓	VIETNAM-2
Chín tèo	G166	10	2	✓	VIETNAM-0
Jasmine 95	G169	15	3	✓	VIETNAM-1
Om 504 japan	G170	13	3	✓	VIETNAM-1

Accession Name	Local ID	nb of panicles	nb of repeat(s)	GBS data	Origin-Ecosystem
Vnd 95-20	G172	13	3	✓	VIETNAM-1
Tám thơm trung quốc	G173	15	3	✓	VIETNAM-1
Blẩu đu	G174	15	3		VIETNAM-0
Ngọ doi	G175	15	3		VIETNAM-4
Khẩu boong đương	G176	9	2		VIETNAM-5
Chăm hơm	G177	10	2	✓	VIETNAM-0
Kháu chính phủ	G178	14	3	✓	VIETNAM-0
Blao pu lau	G179	10	2	✓	VIETNAM-0
Blau plan pieng	G181	10	3	✓	VIETNAM-4
Khẩu mổ	G182	14	3	✓	VIETNAM-4
Khẩu pe lạnh	G183	10	2	✓	VIETNAM-4
Pe lạnh	G184	10	2		VIETNAM-4
Khẩu pe lạnh	G185	10	2		VIETNAM-4
Khẩu nỏ	G186	15	3	✓	VIETNAM-4
Khẩu đường phường	G187	15	3	✓	VIETNAM-4
Bèo cú	G188	15	3		VIETNAM-4
Khẩu năm rinh	G189	4	1	✓	VIETNAM-4
Plề phmả chua	G190	15	3	✓	VIETNAM-4
Khẩu tan	G191	5	1	✓	VIETNAM-2
Khẩu bao thai	G192	5	1	✓	VIETNAM-2
Blề blậu lia	G194	15	3	✓	VIETNAM-4
Blề có pòn	G196	14	3	✓	VIETNAM-4
Blề blậu đa	G198	15	3	✓	VIETNAM-4
Chà ni xa	G199	10	2		VIETNAM-4
Chà fu nu	G200	11	2	✓	VIETNAM-4
Chà xư phu lu	G201	5	1	✓	VIETNAM-4
Nống to	G202	8	2	✓	VIETNAM-4
Plàu cà bành	G203	5	1	✓	VIETNAM-4
Plê đờ	G204	10	2	✓	VIETNAM-4
Blề blẩu chớ	G205	9	2	✓	VIETNAM-4
Blê blẩu đơ	G206	5	1	✓	VIETNAM-4
Khẩu lứa	G207	10	2	✓	VIETNAM-2
Khẩu boong lăm	G208	8	2	✓	VIETNAM-2
Blề chớ	G209	5	1	✓	VIETNAM-4
Khẩu lếch	G221	10	2		VIETNAM-4
Plầu ngoàng plặc	G210	5	1	✓	VIETNAM-4
Plầu bulặt	G211	4	1	✓	VIETNAM-4

Accession Name	Local ID	nb of panicles	nb of repeat(s)	GBS data	Origin-Ecosystem
Plầu cà chắt	G212	10	2	✓	VIETNAM-4
Blè blậu đơ	G213	14	3		VIETNAM-5
Blè blậu xá	G214	10	2	✓	VIETNAM-4
Blè blậu soa	G216	9	2	✓	VIETNAM-5
Plầu của pào	G217	10	3	✓	VIETNAM-4
Khẩu la lạnh	G218	5	1		VIETNAM-5
Plề la	G219	5	1	✓	VIETNAM-4
Khẩu mắc có	G220	8	2	✓	VIETNAM-4
Plề mà mủ	G222	10	2	✓	VIETNAM-4
Blè blậu tan	G223	14	3	✓	VIETNAM-4
APO	G224	12	3		PHILIPPINES-4
ASD 1	G225	10	2		INDIA-1
FANDRAPOTSY 104	G226	5	1		MADAGASCAR-1
IR64	G228	15	3		PHILIPPINES-1
MAKALIOKA 34	G230	5	1		MADAGASCAR-1
NONA BOKRA	G231	5	1		INDIA-1
PETA	G232	5	1		INDONESIA-1
POKKALI	G233	5	1		SRI LANKA-1
SWARNA	G234	5	1		INDIA-3
SHAI KUH	G235	5	1		CHINA-1
SINTANE DIOFOR	G236	5	1		BURKINA FASO-1
TAICHUNG NATIVE 1	G237	15	3		TAIWAN-1
TE QING	G238	5	1		CHINA-1
BARAN BORO	G239	15	3		BANGLADESH-1
BLACK GORA	G240	7	1		INDIA-4
KASALATH	G242	10	2		INDIA-4
N 22	G243	5	1		INDIA-4
BASMATI 1	G244	5	1		PAKISTAN-1
BASMATI 370	G245	5	1		INDIA-1
DOM SOFID	G246	9	2		IRAN-1
KAUKKYI ANI	G247	5	1		MYANMAR-3
PANKHARI 203	G248	5	1		INDIA-1
AZUCENA	G249	5	1		PHILIPPINES-4
IRAT 216	G252	5	1		COTE D'IVOIRE-4
GIZA 171	G253	10	2		EGYPT-1
GOTAK GATIK	G254	5	1		INDONESIA-?
IAC 165	G255	12	2		BRAZIL-4
IGUAPE CATETO	G256	10	2		BRAZIL-4
IRAT 13	G257	10	2		CÔTE D'IVOIRE-4
KHAO DAM	G258	9	2		LAO-4
M 202	G260	5	1		USA-1
NIPPONBARE	G262	9	2	✓	JAPAN-1
CG14	G265	9	2		SENEGAL-1

Table 2: List of Vietnamese landrace accessions phenotyped in 2011. The effective of sampled panicles per accession are indicated as well as the number of repeats considered. The availability of GBS data is indicated. The origin and the ecosystem are indicated as follow: country of origin-1=irrigation (lowland); 2=rainfed low land (sunken field); 3= rainfed low land (deep field); 4=rainfed upland; 5=terraced field; 6=marsh, tide; 0=no information