UNIVERSITÀ DEGLI STUDI DI MILANO
Scuola di Dottorato in Scienze Biologiche e Molecolari
XXVIII Ciclo

IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF GENES CONTROLLING REPRODUCTIVE MERISTEMS DEVELOPMENT IN THE MODEL SPECIES RICE (*ORYZA SATIVA*)

ISRAR UD DIN

PhD Thesis

Scientific tutor: **Professor Martin M. Kater**

Academic year 2015-2016

SSD: BIO/18-GENETICA

Thesis performed at the Department of Biosciences,
UNIVERSITÀ DEGLI STUDI DI MILANO, via Celoria 26, 20133 Milano

Table of Contents

PART-I

Abstract

Meristems are groups of pluripotent cells that contribute to plant growth and the development of its organs. In rice the inflorescence architecture, which determines the grain yield is established at early stages of reproductive development. Activity of the different types of reproductive meristem and the timing of transition between them shape the branching pattern and the number of spikelets on the inflorescence. A few genes that specify the identity of these meristems and regulate the transitions have been reported but the molecular mechanisms, underlying this process are still not clear. To gain a better understanding we used a precise laser microdissection and RNA-sequencing approach in *Oryza sativa* ssp. *japonica* cv. Nipponbare to elucidate the landscape of gene expression in four reproductive meristem types: the rachis meristem (RM), the primary branch meristem (PBM), the elongating primary branch meristem (including axillary meristems) (ePBM/AM), and the spikelet meristem (SM). We found that genes could be grouped based on specific expression behavior in these meristem types. The bioinformatics analysis of the datasets resulted in the identification of several genes potentially involved in branching. We generated loss of function mutants using a genome editing CRISPR/Cas9 approach, for two promising candidate genes, namely *G1L1* and *G1L2*, both belonging to the ALOG family encoding a conserved domain: "Domain of Unknown Function 640" (DUF640). The knock-outs revealed an interesting branching phenotype that is consistent with the expression profile of these genes. Future research of our datasets combined with mutant analysis are expected to provide important new insights into the molecular mechanism that control rice inflorescence development, which has as an ultimate scope the improvement of grain yield, a trait that has without any doubt top priority for a sustainable agriculture of the future.

1. INTRODUCTION

1.1. THE MODEL SPECIES RICE (*Oryza sativa*)

Rice is one of the most important staple food crops, feeding almost half of the world population. Furthermore, it is considered to be one of the main model plants for monocotyledon research because it offers important advantages as an experimental model. Rice has a small genome size and there is the availability of sequenced genomes for several rice species and varieties and this information is stored in publically available searchable databases. It is a self-fertilizing annual plant, which completes its lifecycle in 6-8 months. Furthermore, rice can efficiently be transformed by means of *Agrobacterium tumefaciens*. Being an important cereal crop, rice belongs to the *Poaceae* family, together with wheat, maize, sorghum and barley. As stated by Clark et al. (1995), all these cereals have a monophyletic origin therefore rice can function as an important source of information for studying other cereal crops.

1.2. MORPHOLOGY OF RICE INFLORESCENCE/PANICLE ARCHITECTURE

The rice inflorescence morphology has been studied during different developmental stages and described in detail by Hoshikawa (1989) and Ikeda et al. (2004), the latter using scanning electron microscopy imaging. The rice inflorescence, like in many other grasses, is a compound raceme, also known as a panicle (Figure1A). The main axis of the panicle is known as the rachis, which starts from the upper node of the highest internode. A rachis, which is around 18 cm long, has 6 to 15, usually 8 to 10 nodes; from each of these nodes arise a primary-rachis branch in the axil of a vestigial leaf, the bract. The length of the internodes greatly varies and sometime it can be observed that two or three primary branches growing almost adjoining. The primary branches follow spiral phyllotaxy, which is different from the alternate phyllotaxy of lateral organs in vegetative phase. The length of the primary branch also varies being the longest in the fourth and fifth primary branches and similar to rachis it also has nodes and internodes and several of these nodes develop secondary branches. In rare cases tertiary branches arise from the basal nodes of secondary branches, otherwise each node of secondary branches and distal nodes of primary branches develop short stalks known as the pedicels or peduncles, each of which terminates into a single spikelet. The terminal flowers develop in primary, secondary and tertiary branches but not in the rachis. At the base of the last developing primary branch a small knot like node can be

seen which is the degenerated tip of the rachis (Figure1 B). The difference between a branch and a peduncle is that the branch bears nodes and long enough to produce more than one spikelet while the peduncle is without nodes, produce a single spikelet. On average a single primary panicle axis (rachis) produces 10 or more primary branches and about 150 spikelets. (Figure1 A).

1.3. DEVELOPMENTAL STAGES OF INFLORESCENCE/PANICLE DIFFERENTIATION

The first step in panicle formation is when the plant goes through the transition from vegetative to reproductive development. This transition is dependent on internal signals, like plant age, phytohormone levels, but also environmental conditions such as temperature and day-length (photoperiod) (Andres and Coupland, 2012). After the flag leaf primordium differentiates, the first bract primordium similar to a leaf primordium develops at the opposite position of the flag leaf at the shoot apex and then the first node of the panicle is formed at the base of this bract. This is the first stage of panicle differentiation (Figure1 C-D). Due to the similarity in primordial structures, the bract is considered a homogenous organ of the leaf. At this stage the cone shaped reproductive apex, the rachis meristem elongates producing bract 2 and primary branch primordia simultaneously develop (Figure1 E-F). Very rapidly higher order bract primordia are formed in a spiral manner and primary rachis branches develop in the axil of each bract (Figure1 G). After formation of all its primary branch primordia the rachis meristem looses its activity and stops growing further. The primary branch primordia at the proximal end, which developed earlier, don't elongate till all the primary branch primordia are established. The growth potential of primary branch primordia 1 and 2 is about half compared to the higher ones, i.e. 7 to 10, and secondary branch primordia developing on these are also different in number and length, being less and smaller than the primary ones.

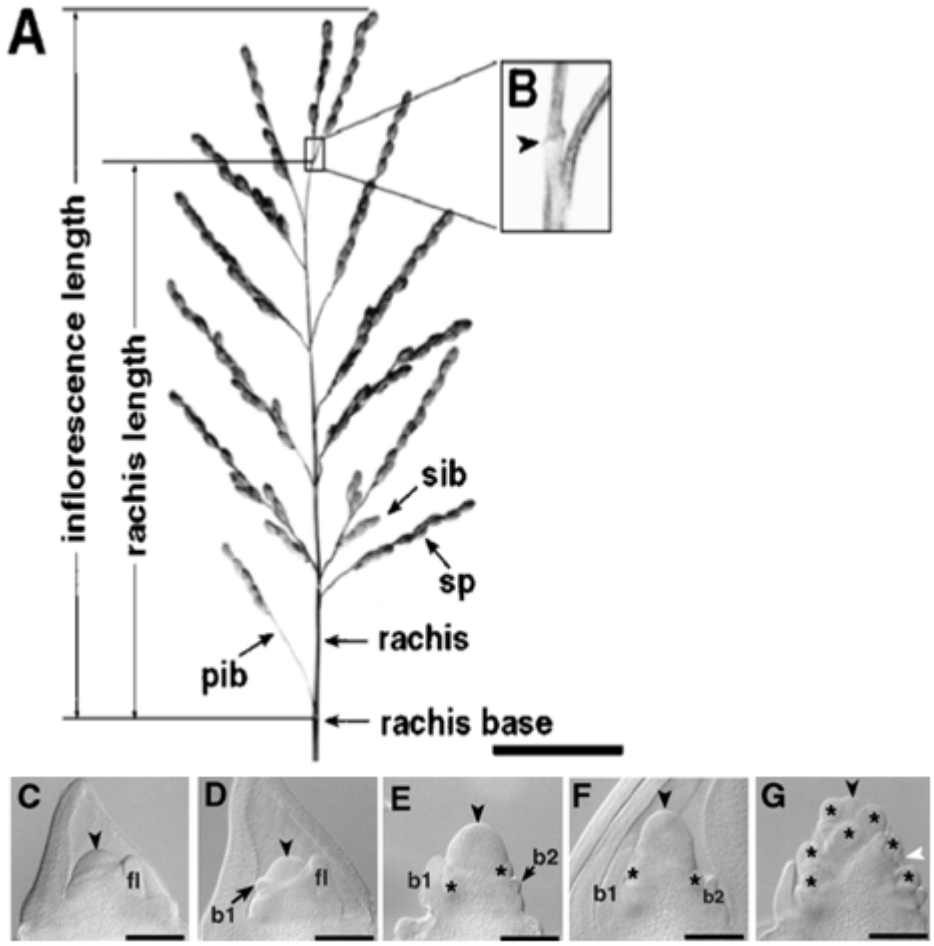

Figure 1.1. Rice panicle morphology and early developmental stages (A). Rice panicle; (B). Degenerative point; (C). First stage of rachis meristem; (D). First bract primordium formation; (E). Formation of two primary branch primordia. Rachis meristem is the largest at this stage; (F). Rachis meristem forming several primary branch primordia; (G). Final stage of primary branch formation. The white arrowhead indicates bract differentiating into hairs. Meristems are indicated by black arrowheads. pib, primary branch; sib, secondary branch; sp, spikelet; fl, flag leaf; b1, first bract; b2, second bract. Bar = 150 μm for (A) to (G). (Ikeda et al., 2004)

1.4. ACTIVITIES OF REPRODUCTIVE MERISTEMS DURING INFLORESCENCE/PANICLE DEVELOPMENT

Meristems are groups of actively dividing cells, present in specific growing regions especially in shoot apices (shoot apical meristem), root apices (root apical meristem), base of nodes and leaf blades in monocot stems (intercalary meristem) that contribute to primary growth of the plant and also produces axillary meristems which form lateral organs (lateral roots, tillars, leaves, branches etc) of the plant. In reproductive phase shoot apical meristem is converted to inflorescence meristem that form reproductive organs, inflorescence and flowers. Variation in

shoot architecture of higher plants is due to several key traits including the nodes number, length of the internodes, activity of meristems present in these nodes and other factors, for example the mechanism of regulation of the stem cells pool at the apices (Bennett and Leyser 2006). The architecture of the rice inflorescence is established during the early stages of development by the activities of reproductive meristems (Rachis meristem, branch meristems and spikelet meristems). The number of primary branches is determined by the timing of rachis meristem abortion, while the transformation of meristem identities, from indeterminate (branch meristems) to determinate (spikelet meristems) defines the pattern of branching. As soon as spikelet differentiation occurs, the branching complexity is fixed and branch primordia rapidly elongate which leads to heading.

1.5. MOLECULAR MECHANISM OF INFLORESCENCE/PANICLE DEVELOPMENT

The branching pattern of the inflorescence not only plays an important role in systematic studies (Classification of plants in groups and subgroups based on morphological homology in plant organs) but it also determines the grain yield of the crop (Kirchoff and Bockhoff, 2013). In higher vascular plants the life cycle is completed by passing through several developmental stages, the two most pronounced are the vegetative phase change, (juvenile to adult transition) and transition from vegetative to reproductive phase (conversion of shoot apical meristem to inflorescence meristem) (Poethig, 2003), the later is critical for reproductive success of the plant species to enter into the next generation.

1.5.1. TRANSITION TO REPRODUCTIVE PHASE AND THE DOUBLE ROLE OF FLOWERING TIME GENES IN REGULATION OF HEADING DATE AND INFLORESCENCE ARCHITECTURE

In ancestral angiosperms the shoot apex is believed to terminate in a flower (Stebbins, 1974), while in most of the descendant angiosperms, a diverse pattern of flowering exists in which the reproductive phase has been prolonged by introducing an intermediate step, the inflorescence meristem, which has the potential to produce axillary meristems that instead of developing flowers produce inflorescence branches.

The main factors, which regulate the phase transition in plants, are those that control shoot apical

meristem (SAM) identity and maintenance, specific miRNAs and flowering signal integrators (Fan et al., 2015). The integrators of the flowering time pathways modulate the switch from vegetative to inflorescence meristem identity by integrating the endogenous signaling pathways and environmental controlled pathways. Some genes involved in the flowering pathways are conserved in both *Arabidopsis*, a long-day plant, and rice (*Oryza sativa*), a short-day plant, but in rice the regulatory pathways seem to be more complex having several heading date genes that regulate flowering under short day and long day conditions which are not found in *A. thaliana* (Brambilla & Fornara, 2013). Genetic module of the photoperiodic flowering pathways as shown in Figure 1.2, in Arabidopsis, a long day plant, flowering is triggered in a light dependent fashion when FLAVIN BINDING KELCH REPEAT F-BOX PROTEIN 1 (FKF1) and GIGANTEA (GI) form a protein complex (Sawa et al., 2007). This complex targets CYCLING DOF FACTORs (CDFs), which encode transcriptional repressors, which regulate the activities of *CONSTANTS* (*CO*), that finally activates the expression of *FLOWERING LOCUS T* (*FT*), which encode a "florigen" protein (Fornara et al., 2009; Putterill et al., 1995; Kardailsky et al., 1999). The florigen moves from the leaves through the phloem to the SAM where it induces flowering (Wigge et al., 2005; Corbesier et al., 2007). Together with linear GI-FKF1-CDFs pathway several other independent mechanisms operate regulating the expression of *CO* at transcriptional and post-transcriptional level (Ito et al., 2012). Some rice varieties are day neutral but tropical varieties flower only when the day length or photoperiod is short. In rice *Heading date 1* (*Hd1*) a homolog of *Arabidopsis CO* is regulated by *OsGI* and activates the expression of *Hd3a*, an orthologue of *Arabidopsis FT*, which triggers flowering in SD. While *Hd3a* is suppressed under LD, which delays flowering (Yano et al., 2000; Hayama et al., 2003). In alternative and independent of the Hd1 pathway, *EARLY HEADING DATE 1* (*Ehd1*), is unique to rice and having no counter part in *Arabidopsis*. It encodes a B-type response regulator and activates the expression of *RICE FLOWERING LOCUS T1* (*RFT1*), and *Hd3a* (Doi et al., 2004). An increase in the day length above 13.5h resulted in low expression of *Hd3a* in "Norin8", a *japonica* rice (*Oryza sativa*), while expression of *RFT1* is less affected (Itoh et al., 2010). Geographic distribution and sequence variation indicates that *RFT1* functional alleles were selected in the process for adaptation to regions with high latitude and short day length (Zhao et al., 2015). Several factors have been reported to regulate the expression of *Ehd1*, such as *Heme Activator Protein like 1* (*OsHAPL1*) (Zhu et al., 2017), *Grain number*, *plant height and heading date 7*

(*Ghd7*) (Xue et al., 2008), *Days To Heading on chromosome 8* (*DTH8/Ghd8/OsHAP3H*) (Xue et al., 2008; Wei et al., 2010), *OsCOL4* (Lee et al., 2010) and *OsCOL10* (Tan et al., 2016) act as repressors of *Ehd1*, while *Ehd1* expression is positively regulated by *Rice Indeterminate1* (*RID1*)*/Early heading date 2* (*Ehd2*) (Matsubara et al., 2008; Wu et al., 2008), *Early heading date 3* (*Ehd3*), *Early heading date 4* (*Ehd4*) (Matsubara et al., 2011; Gao et al., 2013) and *OsMADS51* (Kim et al., 2007).

A MYB-type protein, TaMYB72 from wheat (*Triticum aestivum*), when expressed in rice, triggers flowering in long day conditions, which indicates functional conservation of genes in the control of flowering time in these two diverse crops (Zhang et al., 2016). In several independent flowering pathways and multiple genes are involved in accomplishing heading date in rice through complex interactions and finally targeting *Hd3a* and *RFT1*, which encode for florigen proteins that moves from leaves to the shoot apical meristem (SAM) through the phloem and play a key role in reprogramming of meristem identity and floral transition (Komiya et al., 2009). Hd3a forms a complex with 14-3-3, an intracellular receptor that links OsFD1, a bZIP transcription factor to Hd3a and constitutes the florigen activation complex (FAC) in the SAM, which activates the transcription of the downstream target genes *OsMADS14* and *OsMADS15*, that are members of the AP1/FUL family, a subgroup of the MADS-box gene family (Taoka et al., 2011). Furthermore, AP1-like genes, which specify floral organ identity in *Arabidopsis*, are shown to regulate the switch from SAM to inflorescence meristem (IM) in wheat and rice (Kobayashi et al., 2012). Transcriptome analysis of meristem phase transition from SAM to IM revealed that *PAP2*, a SEPALLATA subfamily MADS-box gene and 3 AP1/FUL family genes, *OsMADS14*, *OsMADS15*, and *OsMADS18* show overlapping expression pattern. The *PAP2* loss of function mutant (*pap2-1*) doesn't show any phenotype at early inflorescence developmental stages, however suppression of *OsMADS14*, *OsMADS15*, and *OsMADS18*, by RNAi resulted in a delay of phase transition from SAM to IM in the *pap2-1* background and instead of an inflorescence multiple shoots were produced. It has been reported that these four MADS-box genes act coordinately in the specification of inflorescence meristem identity downstream of the florigen signal (Kobayashi et al., 2012). Some key regulators of flowering time including *Heading date1* (*Hd1*), *Early heading date* (*Ehd1*) (Endo-Higashi and Izawa, 2011), *Ehd2*, *Rice Indeterminate1* (*RID1*) or *Oryza sativa Indeterminate 1* (*OsID1*) (Matsubara et al., 2008, Park et al., 2008, Wu et al., 2008), *Ehd3* (Matsubara et al 2011), *Ehd4* (Gao et al., 2013), *Grain number,*

Plant height, Heading date 7 (*Ghd7*) (Xue et al., 2008) and *Ghd8/Days to Heading 8* (*DTH8*) (Wei et al., 2010, Yan et al., 2011) play a role in inflorescence development besides their main role in regulating flowering time in rice. Some of these QTLs including *Hd1*, *Ehd1*, *Ghd7*, and *Ghd8*, have already been used in breeding programs because of their functional association with higher grain yield (Endo-Higashi and Izawa 2011). However to explain how these flowering time genes are directly associated with inflorescence development and grain yield is still not understood and further genetic and molecular investigation is needed to elucidate this important developmental process.

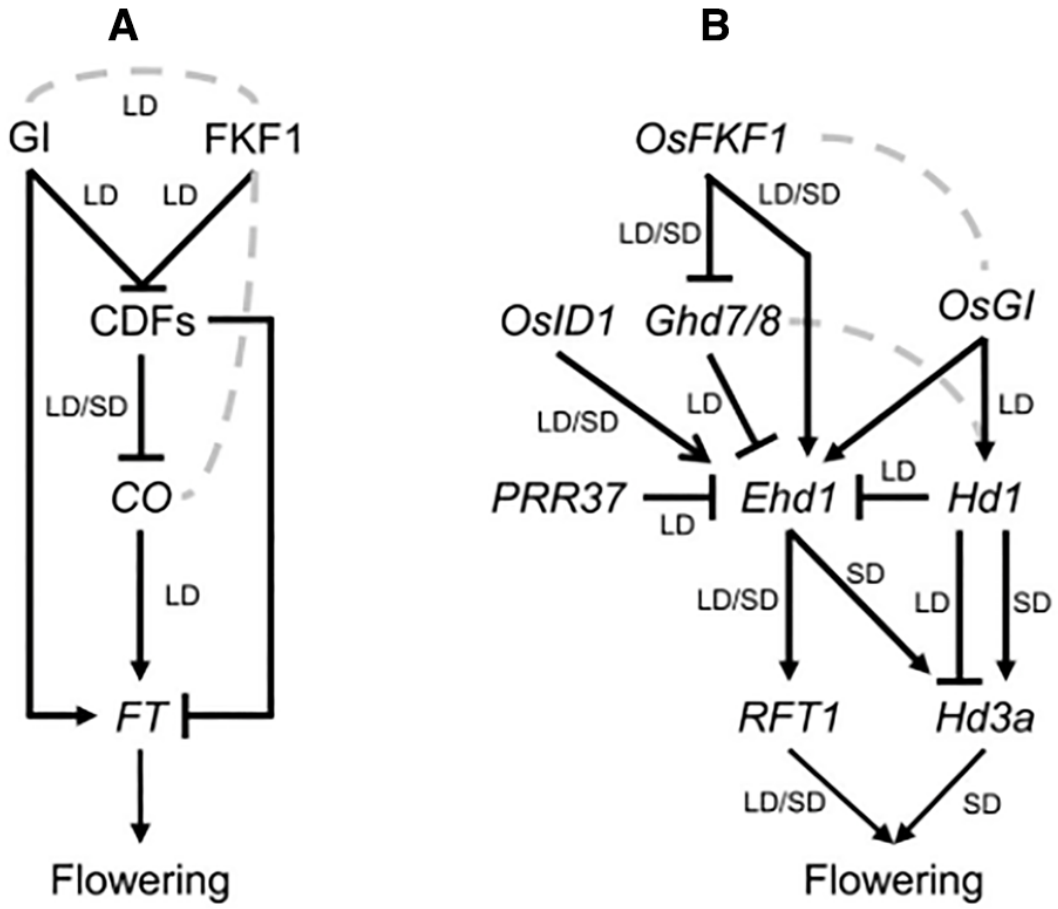

Figure 1.2. A schematic representation of genetic network controlling photoperiodic flowering in Arabidopsis (A) and Rice (B). Figure taken from (Brambilla et al., 2017)

1.5.2. INFLORESCENCE MERISTEM IDENTITY AND MAINTENANCE

Inflorescence form and architecture is mainly determined by the pattern of inflorescence branches created in the due course of development and reflects the number, arrangement, activities and duration of meristems present in the reproductive phase of the plant life cycle (Vollbrecht et al., 2005). Stem cells homeostasis within the reproductive meristems is crucial for normal inflorescence development.

The CLAVATA-WUSCHEL signaling pathway

In Arabidopsis, the CLAVATA (CLV)–WUSCHEL (WUS) feedback pathway is a well-characterized signaling pathway that regulates the coordination of stem cells proliferation with differentiation in the shoot apical meristem (SAM), (Clark et al., 1997, Gaillochet et al., 2015; Pautler et al., 2013). *WUS* is a homeobox gene required for the specification of stem cell identity and mutations in this gene cause premature termination of the SAM and floral meristem after initiation of a few organs (Laux et al., 1996; Xin et al., 2017; Zhou et al., 2015). The role of the CLV genes including *CLV1*, *CLV2* and *CLV3*, major components of the signaling pathway, is to restrict the size of the stem cells compartment. High expression of *WUS* in *clv* mutants implies that the CLV genes act as negative regulators of *WUS* at the transcript level (Schoof et al., 2000). *CLV1* and *CLV2* encode an extracellular leucine-rich-repeat receptor with or without kinase domain respectively (Clark et al., 1997, Jeong et al., 1999; Somssich et al., 2016), while CLV3 is a founding member of the CLE peptide family, which regulates the size of the stem cell niche (Clark et al., 1997; Kayes and Clark 1998; Muller et al., 2008; Ogawa et al., 2008; Ohyama et al., 2009). A compromised expression of any one of these proteins results in over proliferation of the inflorescence and floral meristems and hence formation of extra floral organs. (Clark et al., 1997; Fletcher et al., 1999; Jeong et al., 1999). In both the distantly related species, Arabidopsis and rice, the SAM is organized into a tunica with two clonal layers L1 and L2 and a corpus with one clonal layer the L3. Despite that both species have differences in SAM structure, it is suggested that the CLV signaling pathway is partially conserved in these distant related species. Mutations in the *FLORAL ORGAN NUMBER 1* (*FON1*) an ortholog of *CLV1* (Suzaki et al 2004) and *FON4/2* an ortholog of *CLV3* (Chu et al., 2006; Suzaki et al., 2006) genes also show an enlargement of the floral meristem. Unlike *CLV1*, which is expressed mainly in the corpus of the

SAM in *Arabidopsis* (Clark et al., 1997; Jeong et al., 1999), *FON1* shows a broad expression pattern including all the meristems producing the aerial parts of the rice plant. Compared to the *clv1* mutant, having an enlarged inflorescence meristem, *fon1* has only a phenotype in the floral meristem and produces a normal inflorescence meristem suggesting that there might be an additional FON1-like uncharacterized protein (leucine-rich-repeat receptor kinases), which may putatively regulate inflorescence meristem development in rice (Suzaki et al., 2004).

FON2 SPARE1 (*FOS1*), which encodes a CLE peptide and together with *FON2* maintain FM. Overexpression of *FOS1* led to the termination of SAM, which means it additionally function in Vegetative SAM maintenance. *FOS1* was identified in rice *indica* varieties however its phenotype was suppressed in *japonica* rice due to functional polymorphism (FNP) in the signaling peptide. Based on genetic analysis they have shown that *FOS1* doesn't require *FON1*, a putative receptor of *FON2*, which suggest that *FOS1* and *FON2* works as signaling molecules in regulation of meristem maintenance in independent manner (Suzaki et al., 2009). A third FON2-related gene in rice, *FON2-LIKE CLE PROTEIN1* (*FCP1*), widely expressed in SAM and RAM has a role in vegetative SAM maintenance (Kinoshita et al., 2007; Suzaki et al., 2008). Constitutive expression of *FCP1* results in smaller SAM size and inhibits the initiation of adventitious roots, while *FCP1* RNAi lines, showed similar seedling phenotypes and meristem size to those of the wild types (Suzaki et al., 2008). *FCP2* which is a close paralog of *FCP1* shows similar expression and mutants with double knock down of *FCP1* and *FCP2* transcripts by RNAi, failed to regenerate shoots, which indicates the redundant function of *FCP1* and *FCP2* (Suzaki et al., 2008). *FCP1* negatively regulate *WOX4*, a *WUSCHEL-RELATED HOMEOBOX4*, that functions in the SAM maintenance by regulating the expression of a homeobox gene (*OSH1*) and *FON2* (Ohmori et al., 2013). *TAB1* a WUS-like gene that is expressed in the pre-meristem zone controls the formation of axillary meristem by up-regulating *OSH1* (Tanaka et al., 2015). In short, CLV-WUS pathway in monocots and eudicots is partially conserved but individual components of the signaling pathway have functionally diversified during evolution. Also it is intriguing to understand if there is a functional counter part of the Arabidopsis WUS gene in grasses and if so how it modulates stem cells homeostasis in reproductive meristems.

The Role of Homeobox genes in Meristem Maintenance

In plants SAM formation and maintenance relies on Class 1 KNOTTED 1–like homeobox (KNOX) genes. Knockout mutants of KNOX genes, such as *Oryza sativa homeobox 1* (*OSH1*), maize knotted 1 (KN1), and *Arabidopsis SHOOT MERISTEMLESS* (*STM*), are shown to exhibit abnormal inflorescence development (Long et al., 1996; Tsuda et al., 2011; Vollbrecht et al., 2000). The *osh1* mutant has a smaller inflorescence and a decreased number of spikelets (Tsuda et al., 2011). These reports suggested that KNOX proteins play an important role in shoot apical meristem maintenance. The key role of KNOX proteins has been investigated in Arabidopsis and Rice, which control the shoot apical meristem maintenance by repressing the level of gibberellin (GA) and accumulating high level of cytokinin. GA is shown to promote cell elongation and differentiation while cytokinin regulates cell division and meristem function (Barazesh et al., 2008; Jasinski et al 2005; Sakamoto et al., 2006). The maize *KN1* is shown to link the gene network involved in regulating the specification of plant meristem identity by targeting as many as 643 genes, mostly transcription factors and other proteins involved in hormonal pathways including auxin, cytokinin, GA, and brassinosteroid (Bolduc et al., 2012). The mediation of KNOX protein function by hormones and the mechanism by which KNOX proteins detect the hormone change is still to be investigated.

In rice, it is reported that OSH1 binds to the evolutionarily conserved cis-elements in the promoter region of other KNOX genes and regulate their expression in a direct manner. This regulation between HOX genes is stated to be an indispensable mechanism for self-maintenance of the SAM (Vollbrecht et al., 2000). A negative regulator of class 1 KNOX genes in rice, *OPEN BEAK* (*OPB*), an orthologue of the *JAGGED* gene in *Arabidopsis* is expressed in different meristematic tissues including reproductive meristems. In mutants, *open beak-1* (*opb-1*) and *opb2*, a higher expression of many KNOX genes was observed and plants displayed pleiotropic defects in leaf morphology and reproductive meristems (Horigome et al., 2009). However, the mechanism of interaction of these positive regulator i.e cytokinin and negative regulators like *OPB* for transcriptional regulation of KNOX genes during reproductive stage development is still not clear.

1.5.3. AXILLARY MERISTEM INITIATION DURING INFLORESCENCE / PANICLE DEVELOPMENT

The role of Hormones and TFs

In grasses axillary meristems produce tillers during vegetative development and after phase transition they give rise to branches in the inflorescence and these developmental events are under hormonal, genetic and environmental control. During both vegetative and reproductive phases, the plant hormone auxin plays a key role in axillary meristem initiation (McSteen and Hake 2009). Many genes have been identified through screening for mutants affected in axillary meristem initiation and outgrowth, which results in reduced inflorescence branching. Some of these genes have a key role in hormone biosynthesis, transport and signaling pathways. Genes involved in auxin biosynthesis, transport and signaling have conserved and distinct roles in regulating the axillary meristem initiation and differentiation during inflorescence development in eudicots and monocots. The homologs of the *Arabidopsis* YUCCA family responsible for auxin biosynthesis, the PIN-FORMED 1 (PIN1) and PINOID (PID), auxin transport proteins and the auxin signaling components ARFs (Auxin response factors) in rice and maize are studied and have shown to have similar roles (Benjamens and Scheres 2008; Forestan and Varotto 2012; Zhao 2010).

Two *Arabidopsis* cytochrome P450s *CYP79B2/CYP79B3* and *YUCCA* have been investigated and they have shown to be involved in the conversion of Tryptophan (Trp) into Indole-3-acetaldoxime (IAOx); an important metabolite in the Indole-3-acetic acid (IAA) biosynthetic pathway (Zhao et al., 2001, 2002). In rice the homologues of *CYP79B2/CYP79B3* are still to be identified but seven YUCCA (*OsYUCCA1-7*) genes are identified and constitutive expression of *OsYUCCA-1* resulted in increased level of IAA and plants displayed characteristic phenotypes similar to auxin overproduction. They concluded from these results that *OsYUCCA-1* has a key role in IAA biosynthesis (Yamamoto et al., 2007). The Trp-dependent IAA biosynthetic pathway in rice is still not well elucidated and the enzymes, which catalyze other steps in this pathway, should be investigated.

Auxin transport during inflorescence development is mediated through phosphorylation of PIN1 (auxin efflux carriers) and its homologues by PINOID (PID) that encode Ser/Thr protein kinase including OsPID/OsBIF2 in rice, BIF2 in maize, PID in Arabidopsis and PsPID in Pea

(Gallavotti et al., 2008; Skirpan et al., 2009; Xu et al., 2005). Several rice mutants have been identified that show axillary meristem (AM) patterning and panicle development defects. These mutants include *frizzy panicle* (*fzp*), *lax panicle1* (*lax1*), and *monoculm1/small panicle* (*moc1/spa*) (Komatsu et al., 2001, 2003; Li et al., 2003). They encode mostly transcriptional regulators that constitute the network that operates during AM initiation and maintenance, however the mutual interactions between these factors are still not clear. The rice *MOC1* gene, that encodes a transcriptional regulator belongs to the GRAS (GAI, RGA and SCR) family, is an ortholog of *LATERAL SUPPRESSOR* (*LS*) of tomato and *LAS* of *Arabidopsis* (Li. et al., 2003; Greb et al., 2003). The *FZP* gene in rice, which encodes a protein containing an *APETALA2/ethylene-response* factor (*AP2/ERF*) domain, is an ortholog of the maize *BRANCHED SILKLESS1* (*BD1*). *FZP* and *BD1* determine floral meristem identity in rice and maize respectively, possibly by repressing the AM initiation (Komatsu et al., 2003). Rice *LAX1* (Komatsu et al., 2003) and maize *BARREN STALK1* (*BA1*) (Gallavotti et al., 2004) encoding bHLH protein are described as key regulators of AM initiation. Despite that *BA1* and *LAX1* likely have an important role in auxin-mediated inflorescence morphogenesis, the molecular mechanism by which they act remains unclear. Additionally *LAX2* in rice, physically interacts with *LAX1* and together they regulate AM formation in rice (Tabuchi et al., 2011). Rice *ABERRANT SPIKELET AND PANICLE 1* (*ASP1*) a homolog of *TOPLESS* (*TPL*), which is a transcriptional corepressor in Arabidopsis, plays a role in auxin-related inflorescence development. In the asp1 mutant, fewer inflorescence branches and flowers were observed similar to auxin-related pleiotropic defects (Yoshida et al., 2012). Though the exact function of ASP1 in auxin signaling has not been clearly demonstrated a putative link between ASP1 function and auxin response has been proposed (Yoshida et al., 2012). Cytokinin accumulation in reproductive meristems is also associated with the meristems function and regulation of inflorescence architecture. In rice the *LONELY GUY* (*LOG1*) gene, specifically expressed in the apical and axillary meristems, encodes a cytokinin-activating enzyme that catalyzes the final step in the bioactive cytokinin biosynthetic pathway, is required for continuous function of meristem during inflorescence development (Kurakawa et al., 2007). A *LOG* homolog, *LONELY GUY LIKE PROTEIN6* (*OsLOGL6*) or *An-2* gene, is shown to promote awn length by increasing the level of endogenous cytokinin (Gu et al., 2015). *Gn1a*, which encode an enzyme for cytokinin degradation, CYTOKININOXIDASE/DEHYDROGENASE (OsCKX2), decreases the cytokinin

level in reproductive meristems and reducing transcript levels of *OsCKX2* leads to higher yield of plants by producing more grains (Ashikari et al., 2005). The *UNBRANCHED3* (*UB3*), an ortholog of *OsSPL14* studied in maize and rice respectively, has shown to regulate the expression of *LOG1* and Type-A response regulators (ARRs) which are involved in cytokinin biosynthesis and signaling. Transcriptome analysis of shoot apices and young panicle, showed that many genes associated with cytokinin biosynthesis and signaling were down-regulated in UB3 overexpression lines, including *LOG1* which functions in cytokinin biosynthesis and five Type-A Response Regulator genes (*OsRR1, OsRR4, OsRR6, OsRR9* and *OsRR10*), involved in the cytokinin response pathway, while *OsCKX2* gene from cytokinin oxidase/dehydrogenase family was upregulated which suggest rapid degradation of cytokinin in SAM tissue (Du et al., 2016).

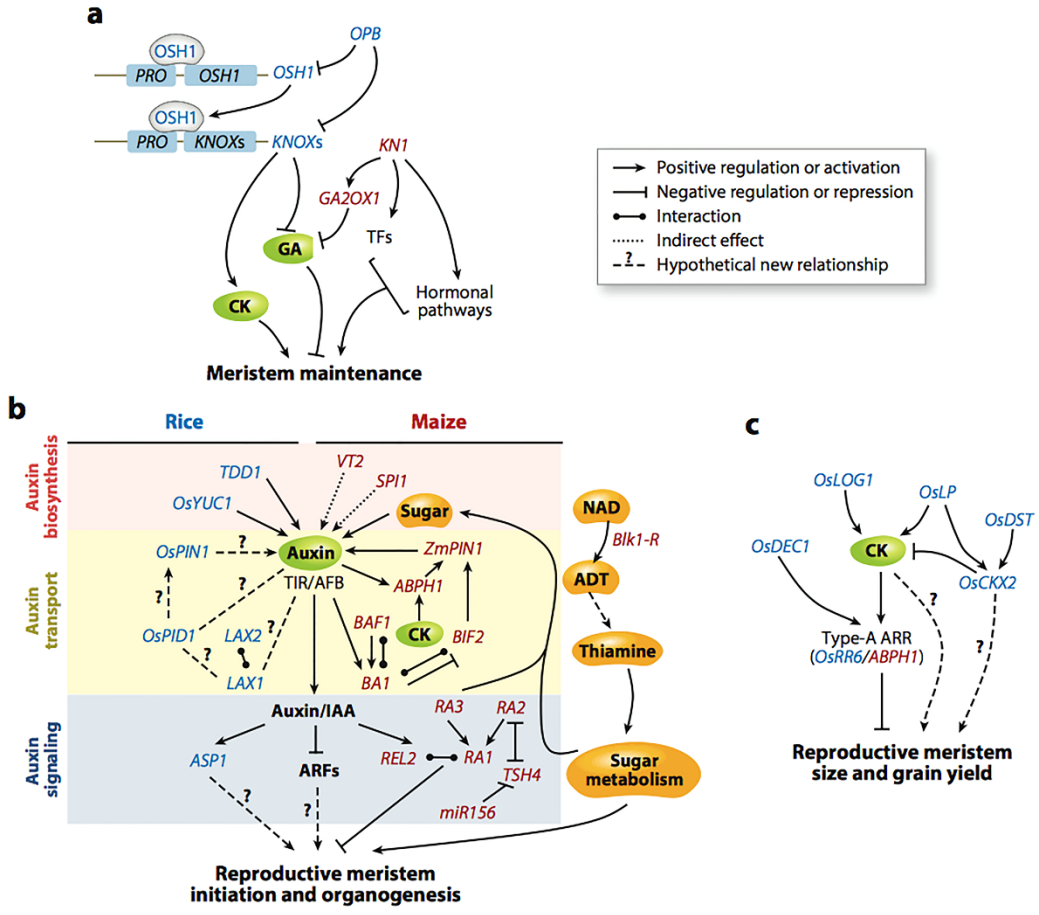

Figure 1.3. Schematic representation of hormones (Auxin, Cytokinin and KNOXs) involved in regulation of meristem activities during inflorescence development in Rice (Shown in blue) and Maize (Shown in red). (a) genetic module for KNOX genes regulating meristem maintenance (b) Mechanism of auxin biosynthesis, transport and

signaling and its role in reproductive meristem activities. (c) Cytokinin ontrolling reproductive meristems size and grain yield. Abbreviations: ARR, auxin-response regulator; ARF, auxin-response factor; IAA, indole-3-acetic acid; CK, cytokinin; GA, gibberellin; ADT, adenosine diphospho-5-(β- ethyl)-4-methylthiazole-2-carboxylic acid; PRO, promoter; TF, transcription factor; TIR/AFB, transport inhibitor response/auxin-binding F-box protein. (From Zheng and Yuan, 2014)

1.5.4. BRANCHING POTENTIAL OF THE INFLORESCENCE

Arabidopsis *TERMINAL FLOWER1* (*TFL1*) and its *Antirrhinum* ortholog *CENTRORADIALIS* (*CEN*), is needed for inflorescence meristem to maintain the indeterminate SAM identity and mutation in these genes led to conversion of inflorescence meristem to terminal flower (Alvarez et al., 1992; Bradely et al., 1996). The function of *TFL1/CEN* homologues in rice *RCN1* and *RCN2* have been investigated to understand if there is a similar mechanism underlying inflorescence development. In plants overexpressing *RCN1* and *RCN2*, resulted in a delay of phase transition and plants produced more inflorescence branches (Nakagawa et al., 2002). In *Arabidopsis TFL1* activity is suppressed in the floral meristem by the interaction of *APETALA1* (*AP1*) and four MADS-domain factors, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), *AGAMOUS-LIKE 24* (*AGL24*), *SHORT VEGETATIVE PHASE* (*SVP*), and *SEPALLATA4* (*SEP4*) when indeterminate IM acquires determinate FM identity. *OsMADS50* and *OsMADS56*, homologous to *Arabidopsis SOC1*; *OsMADS22*, *OsMADS47*, and *OsMADS55*, homologous to *Arabidopsis SVP* and *AGL24*, seem to play a similar role because when these MADS-box genes (*OsMADS50*, *OsMADS56*, *OsMADS22*, *OsMADS47* and *OsMADS55*) were knocked-down in the *pap2-1* background, a significant increase in panicle branching was observed and the phenotype was similar to plants overexpressing *RCN1* and *RCN2*. When expression of four RCN genes were examined in this particular *pap2-1* mutant, *RCN4* expression was found upregulated and knocking down of RCN genes produced even smaller inflorescences, which indicates that *SOC1*, *AGL24*, *SVP* and *SEP4* orthologs in rice regulate panicle branching by suppressing *TFL1* orthologs in rice (Liu et al., 2013). The precise genetic interaction between these MADS-box genes and other inflorescence determinants is still not elucidated. The *SQUAMOSA PROMOTER BINDING PROTEIN*, a transcription factor that binds to the promoter of the MADS-box gene SQUAMOSA was first reported in *Antirrhinum majus* and emerged as an important regulator of a wide range of biological processes (Klein et al., 1996). In rice now 19 SPL (*SQUAMOSA PROMOTER BINDING PROTEIN-LIKE*) genes have been identified by genome wide identification including

OsSPL14 which is encoded by a QTL known as the *WFP* (*WEALTHY FARMER PANICLE*), that regulates panicle branching, and tiller number in rice. The expression of *OsSPL14* is post-transcriptionally regulated by *OsmiR156* and a single point mutation in the *OsmiR156* target site resulted in the production of fewer tillers and more panicle branches that led to ideal plant architecture (Jiao et al., 2010, Miura et al., 2010). However the molecular mechanism by which *OsSPL14* regulate panicle branching is still not elucidated and study of its downstream targets can provide a clue to understand this regulatory mechanism.

1.5.5. SPIKELET MERISTEM IDENTITY AND DETERMINATION

An important event in panicle development, which determines the architecture and final yield of the inflorescence, is transformation of indeterminate meristems (lateral and terminal meristems of inflorescence branches) to determinate meristems (spikelet meristems), which differentiate into functional units of the inflorescence known as the florets. Many key players involved in this transition have been reported including *ABERRANT PANICLE ORGANIZATION-1* (*APO1*) an ortholog of *Arabidopsis UNUSUAL FLORAL ORGAN* (*UFO*) and *APO2*, which is identical to *RFL* an ortholog of *Arabidopsis LEAFY* (*LFY*). *APO1* encodes an F-box protein and suppresses the transition from indeterminate to determinate meristems and hence produces more branch meristems, which ultimately results in higher spikelet numbers (Ikeda et al., 2007). *APO2* controls meristem size and apo2 mutants exhibit smaller meristem and altered phylotaxy of primary branches (Ikeda et al., 2012). Another key regulator of this event is *TAWAWA1* (*TAW1*), which belongs to ALOG family and encodes a nuclear protein containing an uncharacterized DUF640 domain, which is conserved in all land plants. *TAW1* shows high expression in the SAM and reproductive meristems and in the dominant gain-of-function mutant *tawawa1-D*, indeterminate meristem activity is extended by delaying spikelet meristem specification, which resulted in more secondary and tertiary inflorescence branches and hence producing more spikelets. In the *tawawa1-D* mutant three genes, *OsMADS22*, *OsMADS47*, and *OsMADS55* belongs to *SVP* subfamily of MADS-box genes were highly up-regulated while spikelet identity genes showed reduced expression including *OsMADS7* (*SEP3*), *OsMADS8* (*SEP3*), *OsMADS16* (*AP3*), *OsMADS4* (*PI*), *OsMADS3* (*AG*), and *OsMADS58* (*AG*) (Yoshida et al., 2013). Chapter 3 of this thesis is dedicated to the study of other ALOG genes in rice inflorescence development.

1.5.6. FLORAL MERISTEM IDENTITY AND DETERMINATION

In grass species the florets are grouped into a single basic unit of the inflorescence known as the spikelet. In rice each spikelet bears a single floret. *FRIZZY PANICLE* or *BRANCHED FLORETLESS 1* (*FZP/BFL1*) an ortholog of maize *BD1*, encodes an Ethylene Response Factor (ERF), a transcription factor from the AP2/ERF family, which is expressed in reproductive meristems at the time of rudimentary glumes differentiation and hence prevents axillary meristem formation within spikelets and establishes the floret formation. In *fzp* mutants floret formation is arrested and the meristems acquires spikelet identity producing axillary meristems in the axils of rudimentary glumes resulted in a highly branched panicle (Komatsu et al., 2003). *SUPERNUMERARY BRACT* (*SNB*) and *INDETERMINATE SPIKELET-1* (*OsIDS1*) are two AP2 family genes, which are functionally associated and control floral meristem specification. Compared to the *snb* single mutant, the *snb osids1* double mutant further delayed the transition to floral meristem and the mutants exhibit fewer inflorescence branches and spikelet numbers (Lee and An 2012). Another member of the AP2/ERF family, *MULTI-FLORET SPIKELET 1* (*MFS1*), is expressed in the spikelet and floral meristems and regulates spikelet meristem fate and determines floral meristem identity. In mfs1 mutants the transition to floral meristem is suppressed and floral organs identity is altered. *MFS1* positively regulate *SNB* and *OsIDS1*, which control floral meristem specification and *G1/ELONGATED EMPTY GLUME* (*ELE*), a member of the ALOG family, which specifies sterile lemma identity (Ren et al., 2013). Overexpression of the microRNA, *OsmiR172* mostly phenocopies the *snb osids1* double mutant in rice (Zhu et al., 2009; Lee and An 2012). This indicates that *miR172* is involved in regulation of IDS1-like genes in rice, which play a key role in specifying the fate of the spikelet meristem. The AP2 family genes in grasses are functionally diverged and further investigations on these genes will add more insight into the evolution and functional diversification of AP2 members in specifying the fate of reproductive meristems in grasses.

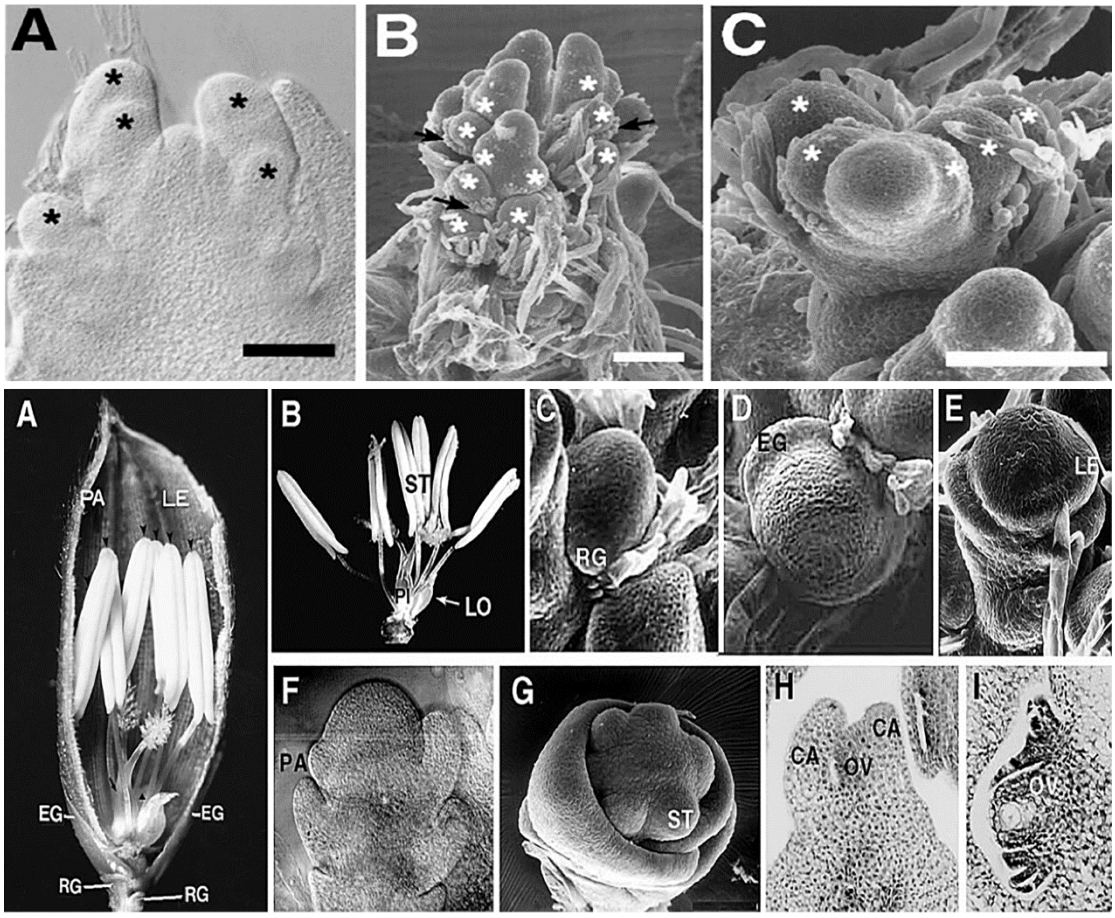

Figure 1.4. Late Stages of rice inflorescence development. (Ikeda et al 2004). (A) Primary branches eleongation, (B) Secondary or axillary branch primordia (SBP) differentiation, (C) Top view of (B) showing distichous phyllotaxy of SBP. Spikelet development (Itoh et al., 2005). (A) Mature spikelet, (B) Mature flower. (C) Early SM producing RG primordium. (D) Formation of EG primordia. (E) Formation of LE primordium. (F) Formation of PA primordium. (G) Formation of ST primordia. (H) Formation of CA and OV primordia. (I) Formation of embryo sac. Abbreviations: PA, palea; LE, lemma; EG, empty glume; RG, rudimentary glume; ST, stamen; PI, pistil; LO, lodicule; CA, carpel; OV, ovule.

1.6. NEW TECHNOLOGIES AND ITS IMPACT ON FUNCTIONAL GENOMICS

The human population is increasing with rapid pace and expected to reach 9.7 billion by 2050 (United Nation, 2015 Revision). In order to produce sufficient food for such a huge population, sustainable crop productivity has posed a difficult challenge to the scientists in the era of limited natural resources, unpredictable global climatic changes and other biotic and abiotic stresses (Fischer et al., 2002). A tremendous progress has been made with rapid development in technologies to bring a revolution in the areas of biology and biotechnology, especially in computational biology, functional genomics, genetic engineering and molecular

breeding. Application of advanced biological and biotechnological approaches has procreated greater impact on agriculture in terms of sustainable production and improving crop yield (Thao and Tran, 2016).

The main aim of functional genomics studies is to explore the information stored in the genes and regulatory elements of the genome, to decipher the gene networks that underlie certain biological pathways and to study the overall functions while looking into the phenotypes or responses of mutant alleles to biotic and abiotic stresses.

According to an updated 7.1 release of rice pseudo-molecules and annotation, from 12 rice chromosomes, 55,986 genes (loci) have been identified, of which 6,457 have 10,352 additional alternative splicing isoforms reaching to a total of 66,338 transcripts in the rice genome (http://orygenesdb.cirad.fr/index.html). To ascribe biological functions to all these predicted genes, a great effort has been done by different groups around the globe in generating mutant libraries through Tos17 tagging, T-DNA insertion, Ds/dSpm tagging, and chemical/ irradiation mutagenesis. From analysis of 246 566 flanking sequence tags (FSTs), among 211470 unique hits, 60.49% of nuclear genes contain at least one insertion and 68.16% were having insertion in the genic region and additionally 57% of non-TE (transposable element) related genes have insertion tags (Wang et al., 2013). Advancement in omics technologies, computational tools and genome editing approaches (see below) have greatly eased and speeded up functional genomic research in rice and other model organisms.

1.6.1. NEXT GENERATION SEQUENCING AND TRANSCRIPTOMICS

Since the discovery of the DNA double helix in 1953, the first effort to sequence 10 out of 20 residues of a linear λDNA strand was made by (Wu et al., 1968) which further continued until modern and reliable methods like chemical method by Maxam-Gilbert (Maxam and Gilbert 1977) and the dideoxy sequencing method by Sanger (Sanger et al., 1977) were developed to obtain longer reads. Initially this technology progressed with slow pace but incremental improvements made complete genome sequencing, ranging from small phages of 5386 bases to the human genome of ~3 billion bases, possible (Hutchison 2007). To address more biological questions and to overcome the limitations of traditional sequencing methods with respect to throughput and costs, second generation or next generation sequencing (NGS) approaches emerged (Margulies et al., 2005) and revolutionized the field of genome sequencing. Application

of NGS technologies are not limited to the static genome analysis but give further insights into the genome organization, gene interaction and epigenetic control in eukaryotic cells with several approaches like ChIP-seq used for studying DNA-protein interaction, Hi-C that investigate the 3D architecture of whole genome and Methylated DNA Immunoprecipitation (MeDIP) to find the methylated cytosine bases distributed in the genome. Furthermore, these NGS methods are also successfully applied for gene expression or transcriptome analysis using RNA-seq approaches resulting in high-resolution RNA expression analysis (Nowrousian, 2010).

Laser microdissection coupled with transcriptome analysis is successfully applied to probe the cell specific gene expression dynamics under certain growth conditions or developmental stages (Ogo et al., 2014; Mantegazza et al 2014). Knowledge of gene expression, i.e the timing, growth or developmental stage and the cell or tissue where the gene is expressed provides a clue to evaluate its biological function. For gene expression analysis initially hybridization based microarray techniques and real-time PCR methods were used to measure the mRNA level, however both lack exquisite sensitivity and are not reliable to deal with alternative splicing isoforms and non-annotated transcripts. To overcome these limitations, tag based sequencing methods like serial analysis of gene expression (SAGE) (Velcules et al., 1995; Herbers and Carninci 2005), cap analysis of gene expression (CAGE) (Kodiziu et al., 2006; Shiraki et al., 2003) and massively parallel signature sequencing (MPSS)(Brenner et al., 2000; Reinartz et al., 2002) were developed which provided a digital readout of mRNA level. These sequencing approaches were best in their ability to provide the accurate digital expression of genes at levels below the sensitivity of hybridization based microarray techniques, but still have some disadvantages like most were based on expensive Sanger sequencing technology, and a significant portion of the short tags cannot be uniquely mapped to the reference genome. Moreover, only a portion of the transcript is analyzed and isoforms are generally indistinguishable from each other. By contrast, the RNA-seq approach by NGS technologies offers rapid and cost effective sequencing capacity, which complemented perfectly the SAGE and other traditional sequencing methods for both mapping and quantifying the transcriptome (Mardis 2008). Wide range of applications and better reliability make RNA-Seq a preferred method for the scientists to evaluate absolute transcript levels, single nucleotide polymorphism (SNPs), to detect novel transcripts and isoforms, map exon/intron boundaries, and many more, working with both sequenced and unsequenced organisms (Mutz et al., 2013).

1.6.2. BRIDGING TRANCRIPTOMICS AND FUNCTIONAL GENOMICS THROUGH GENOME EDITING TOOLS

Functional characterization of genes through genetic mutation is critical for crop genetic improvement, which has been proven from studies on naturally occurring allelic variations and its application in breeding programs. Moreover characterization of thousands of mutants generated artificially by T-DNA/transposon insertion, physical or chemical mutagenesis in model plants species including Arabidopsis (Kuromori et al., 2006) and rice (Wu et al., 2003 and Yang et al., 2013) has paved the way to understand and exploit various biological mechanisms but the fact that random mutagenesis is linked to unwanted background mutations and screening of large scale mutant libraries is costly and time consuming (McCallum et al., 2000). Other approaches like RNA interference (Smith et al., 2000b), antisense RNA (Mol et al., 1990) and virus-induced gene silencing (Baulcombe, 1999) can suppress the expression of corresponding mRNAs and may compromise their functions but silencing is often only partial and unstable which make these tools less successful compared to genetic mutants. An advantage of these systems is the possibility to silence genes tissue specifically.

In the last decade a few exciting new technologies for genome engineering including Zinc Fingers Nucleases (ZFNs), transcription activator-like effector nucleases (TALEN) and recently Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 have emerged, which are capable to target a nuclease to specific DNA sequence creating a double strand break (DSB) at the desired target locus (Gaj et al., 2013). Activation of the cellular response called DNA damage response triggers the DNA repair machinery which follows two possible mechanisms i.e homology-directed repair (HDR) and error-prone non-homologous end joining (NHEJ) to operate and ultimately leads to several types of mutations including point mutations, deletions, insertions, inversions, duplications and translocations at specific sites (Wyman et al., 2006; Joung and sander, 2013). Success of these methods mainly depends on their designing and engineering procedures, selection, specificity, precision and genome editing efficiency.

1.6.3. Zinc Finger Nucleases (ZFNs)

Zinc Finger Nucleases here after (ZFNs) were developed based on the functional principles of the Cys2-His2 zinc finger domains (Miller et al., 1985), which constitutes one of the most common structural motifs in eukaryotes (Rubin et al., 2000). In ZFNs a custom-designed sequence-specific DNA-binding zinc finger domain at the N-terminus is fused with a non-specific nuclease domain of the type IIS restriction enzyme FokI at the C-terminus that cleave DNA (Kim et al., 1996). For the nuclease activity of ZFNs, dimerization of FoKI domains is a crucial step, which is established by the binding of a ZFN-FokI hybrid molecules to the target DNA which is 18-36 bp long including 5-7 bp spacer or cleavage site (Smith et al., 2000).
For gene modification, ZFNs have been successfully applied in living system including human cells (Perez et al., 2008), animals (Doyon et al., 2008; Meng et al., 2008) and plants (Osakabe et al., 2010; Townsend et al., 2009; Shukla et al., 2009). However, carrying some shortcomings such as difficult screening procedures (Hsu et al., 2012), not all sequences can be targeted, off target activity (Cathomen et al., 2008), possible toxicity to the host cell etc. limit its application in functional genomics studies.

1.6.4. TRANSCRIPTION ACTIVATOR-LIKE EFFECTOR NUCLEASES (TALENS)

Success of ZFNs in genetic research inspired the scientists to further improve and explore the customized nuclease technology and its application for therapeutic and other research purposes. TRANSCRIPTION ACTIVATOR-LIKE EFFECTOR NUCLEASES (TALENs), which were relatively easier to design emerged as a better alternative to ZFNs for genome modification. Similar to ZFNs, TALENs also contains a sequence specific customizable DNA-binding domain fused to a nonspecific FokI nuclease domain (Joung and Sander 2013). The DNA-binding domain is derived from a naturally occuring transcription activators like (TAL) effector family secreted by a pathogenic gram negative bacteria of the genus *Xenthomonas* and is made up of a central domain comprised of 1.5 to 33.5 tandem repeats or TAL repeats, each containing 33 to 35 highly conserved amino acids and it recognize one specific DNA base pair (Boch et al., 2009 and Kay et al., 2005). The effector proteins are pumped into the plant cells via a type III secretion system, bind to the promoters of target genes and reprogram the host cell transcription machinery which leads to plant developmental changes such as cell division and enlargement (Boch et al.,

2009). Within each TAL repeat at position 12th and 13th there are two hypervariable residues also known as Repeat Variable diresidues (RVDs), which confer TAL effectors, its DNA binding specificity. The most commonly occuring RVDs are His/Asp (HD) accounting for cytosine (C) binding, Asn/Gly (NG) recognizes thymine (T), Asn/Ile (NI) for Adenine (A) and NN specific for G or NK which is less common but in some contexts appears to have higher efficiency for G than NN (Moscou and Bogdanove 2009; Mak et al., 2012). Due to ease in manipulation and performing the same features as ZFNs, TALENs have been widely applied for genome modification and gene functional analysis in zebrafish (Sander et al., 2011; Huang et al., 2011), rat (Tesson et al 2011), human cells (Miller et al., 2011; Hockemeyer et al., 2011) Arabidopsis (Christian et al 2010; cermak et al., 2011; Li et al., 2012), rice (Li et al., 2012), wheat (Wang et al., 2014), potato (Sawai et al., 2014) and tomato (Lor et al., 2014).

1.6.5. CLUSTERED REGULARLY INTERSPACED SHORT PALINDROMIC REPEATS (CRISPR)-CAS TECHNOLOGY

Prokaryotes have developed various defense strategies to face a wide range of threats posed by viruses and one of these is the RNA-directed genome targeting nuclease system called Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas system, which provide adaptive immunity against mobile genetic elements such as viruses and plasmids (Marraffini, 2015). CRISPRs were reported for the first time in 1987, from sequencing the iap gene in *Escherichia coli* (Ishino et al., 1987). Downstream of iap an unusual repeat cluster comprising of a series of 29 nucleotide repeats were found separated by 32 nucleotides unique spacer sequences. However its biological significance was not found until a family of unique sequences of Short Regularly Spaced Repeats (SRSRs) were reported in prokaryotes (Mojica et al., 2000). Subsequently it was demonstrated in *Archaeoglobus fulgidus* that these SRSRs loci transcribed into snmRNAs (Tang et al., 2002). To avoid the confusion in nomenclature it was decided to use the acronym CRISPR for SRSRs and in the same report investigation of the flanking sequences of CRISP loci in prokaryote genome, leads to the discovery of four CRISPR associated genes cas1 to 4, showing a clear homology in different prokaryotes species (Jansen et al., 2002). The functional role of these repeats was demonstrated based on findings from three independent studies. They showed that the spacers were identical in sequence to plasmid and viral sequences (Bolotin et al., 2005, Mojica et al., 2005, Pourcel 2005), and the more spacers

there were in *Streptococcus thermophiles* strains, the fewer phages could infect them (Pourcel et al., 2005). Based on these observations it was suggested that CRISPR-Cas probably plays a role as immune system in prokaryotes. It was indicated that CRISPRs function via a similar mechanism to that of RNAi in eukaryotes (Mojica et al., 2005, Makarova et al., 2006). In type II CRISPR/Cas from *Streptococcus pyogenes*, a pre-crRNA is processed into a short CRISPR RNA (crRNA) by RNAse III in the prescence of transactivating crRNA (tracRNA) and Cas9. The tracrRNA forms a ribonucleoprotein complex with a nuclease protein Cas9 and crRNA recognizes a specific sequence in the target genome known as protospacer followed by a 3 nucleotide unique sequence called Protospacer Adjacent motif (PAM), and thereby recruite the Cas9 nuclease to the desired target locus at PAM site to generate the DSB in the DNA (Jinek et al., 2012, Mali et al., 2013).

In a very short period of time the CRISPR-Cas9 genome editing technology has been applied successfully in a wide range of organisms including human cells (Mali et al., 2013; Tsai et al., 2014; Guilinger et al., 2014), mice (Shen et al., 2013), zebrafish (Chang et al., 2013, Hwang et al., 2013), wheat (Shan et al., 2013), rice (Miao et al., 2013; Xie et al., 2013), Arabidopsis (Li et al., 2013; Feng et al., 2014), tobacco (Li et al., 2013), sweet orange (Jia et al., 2014). The only limitation with this system is the off-target activity introducing unwanted mutations but several efforts have been made to further enhance the specificity and flexibility of this system to eliminate or reduce the potential off target effects. Using Cas9 from different bacteria which recognize longer target sequence (Hou et al., 2013), using a modified form of Cas9 i.e Cas9 nickase (Cho et al., 2014) and manipulating the length of the protospacer sequence in the guide RNA (Fu et al., 2014) has greatly reduced the off target effects.

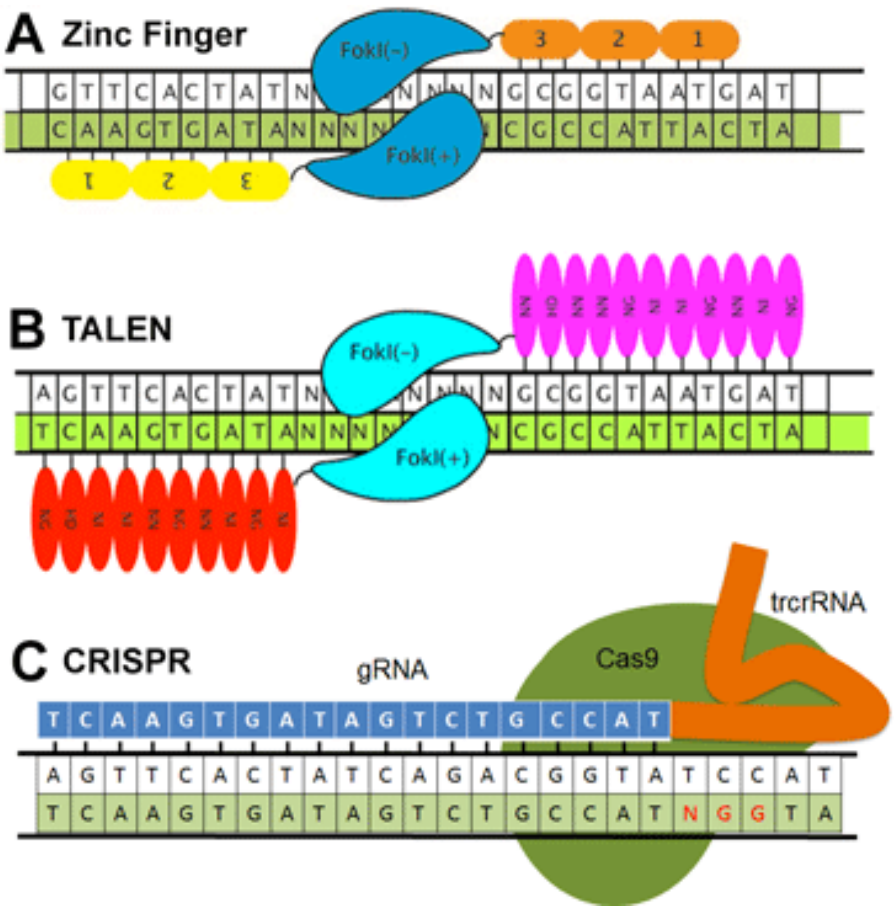

Figure 1.5. Diagrammatic presentation of the Genome Editing Tools taken from. Addgene plasmid repository (https://www.addgene.org/)

1.7. AIM OF THE PROJECT

Inflorescence development in rice is a complex process starting from the conversion of Shoot Apical Meristem (SAM) to Inflorescence Meristem (IM), transition of indeterminate (IM and Branch Meristems) to determinate meristems (Spikelet Meristems), formation of reproductive branches and finally flowers and grain formation. The timing of meristem activities in a particular stage and transition to the next stage determines the architecture and final grain yield of the inflorescence. The molecular mechanism that control these activities and developmental events is however still poorly understood and many genes involved in these processes have for sure not been identified yet, we are in a way just seeing the tip of the iceberg. The aim of this study is to find and characterize new genes potentially involved in regulating these various developmental events during inflorescence development in rice.

1.7.1. TASKS

- Morphology of inflorescence at early stages of development?

As described earlier meristem transitions occur very fast and these early stages of meristem are difficult to obesrve. We decided to perform histological analysis on these early stages to understand the exact timing of the distinct stages of meristem formation. We were able to identify four distinct developmental stages of meristems and the exact timing of their development, including IM, Primary Branch Meristem (PBM), elongated PBM/Axillary Meristems and SM.

- How to get meristem specific and high quality mRNA from the distinct stages?

For this step we planned to isolate the specific meristem types with micrometric precision using laser microdissection methodology and to define the accurate transcriptome for each distinct meristem type. For this purpose we improved the protocol for embedding meristem tissue and membrane slides preparation for laser microdissection, to get good quality RNA for sequencing. We analyzed the transcriptome of the four selected meristem types collected from wild type rice, Nipponbare inflorescences. Chapter 2 of this thesis is dedicated to this transcriptome analysis.

- Picking up candidate genes and generation of loss of function mutants.

Candidate genes were selected from the transcriptome datasets taking into consideration the literature and expression patterns. The expression profiles of the selected candidate gene was further confirmed by RT-PCR and in-situ hybridization and for functional characterization the CRISPR/Cas9 genome editing approach was applied to produce loss of function mutants. Chapter 3 of this thesis is related to this study.

Overall this work is based on modern technologies for cell specific transcriptome analysis using the laser microdissection technique coupled with NGS RNA-Seq technology. We presented a highly valuable resource for gene expression profiling during early reproductive stages, which could be exploited by scientists to get a clue for studying the biological function of putative candidates involved in cereal inflorescence development. Moreover the mutant lines we produced with the CRISPR-Cas9 genome editing approach will be used to get further insight into the molecular mechanism of inflorescence development in rice.

REFERENCES

Andres F, Coupland G. 2012. The genetic basis of flowering responses to seasonal cues. Nature Reviews Genetics 13(9):627-639.

Ashikari M, Sakakibara H, Lin S et al., (2005) Cytokinin oxidase regulates rice grain production. Science, 309, 741–745.

Barazesh S, McSteen P. 2008. Hormonal control of grass inflorescence development.Trends Plant Sci. 13:656–62.

Baulcombe DC. 1999. Fast forward genetics based on virus-induced gene silencing. Curr. Opin. Plant Biol. 2:109–113.

Benjamins R, Scheres B. 2008. Auxin: the looping star in plant development. Annu. Rev. Plant Biol. 59:443–65.

Bennett T and Leyser O. 2006. Something on the side: axillary meristems and plant development Plant Mol. Bio. 60:843–854.

Boch J, Bonas U. 2010. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol. 48: 419–436.

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S et al., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326: 1509–1512.

Bogdanove AJ, Schornack S, Lahaye T. 2010. TAL effectors: finding plant genes for disease and defense. Curr Opin Plant Biol. 13: 394–401.

Bolduc N, Yilmaz A, Mejia-Guerra MK, Morohashi K, O'Connor D et al., 2012. Unraveling the *KNOTTED1* regulatory network in maize meristems. Genes Dev. 26:1685–90.

Bolotin A, Quinquis B, Sorokin A & Ehrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151, 2551–2561.

Bomblies K, Wang R-L, Ambrose BA, Schmidt RJ, Meeley RB, Doebley J. 2003. Duplicate *FLORICAULA/LEAFY* homologs *zfl1* and *zfl2* control inflorescence architecture and flower patterning in maize. Development 130:2385–95.

Bommert P, Je BI, Goldshmidt A, Jackson D. 2013. The maize Gα gene *COMPACT PLANT2* functions in CLAVATA signalling to control shoot meristem size. Nature 502:555–58.

Bommert P, Satoh-Nagasawa N, Jackson D,HiranoH-Y. 2005. Genetics and evolution of inflorescence and flower development in grasses. Plant Cell Physiol. 46:69–78.

Bonas U, Stall RE, Staskawicz B. 1989. Genetic and structural characterization of the avirulence gene avrBs3 from *Xanthomonas campestris* pv. *vesicatoria*. Mol Gen Genet 218: 127–136.

Brambilla V, Fornara F. 2013. Molecular control of flowering in response to day length in rice. J. Integr. Plant Biol. 55(5), 410–418.

Brambilla V, Gomez-Ariza J, Cerise M, Fornara F. 2017.The Importance of Being on Time: Regulatory Networks Controlling Photoperiodic Flowering in Cereals. Front. Plant Sci. 8:665.

Brenner S et al., 2000. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nature Biotechnol. 18, 630–634.

Carraro N, Forestan C, Canova S, Traas J, Varotto S. 2006. *ZmPIN1a* and *ZmPIN1b* encode two novel putative candidates for polar auxin transport and plant architecture determination of maize. Plant Physiol. 142:254–64.

Cathomen T, Joung JK. 2008. Zinc-finger nucleases: the next generation emerges. Mol Ther 16: 1200–1207.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C et al., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res: gkr218.

Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X et al., 2013. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Res 23: 465–472.

Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A et al., 2010. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics, 186: 757–761.

Chu H, Qian Q, LiangW, Yin C, Tan H et al., 2006. The *FLORAL ORGAN NUMBER4* gene encoding a putative ortholog of *Arabidopsis CLAVATA3* regulates apical meristem size in rice. Plant Physiol. 142:1039–52.

Clark LG., Zhang W, Wendel JF. 1995. A Phylogeny of the Grass Family (Poaceae) Based on ndhF Sequence Data. Systematic Bot. 20:436-460.

Clark SE, Williams RW, Meyerowitz EM. 1997. The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. Cell 89:575–85.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q et al., 2007. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science 316 1030–1033.

Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK et al., 2012. Structural basis for sequence-specific recognition of DNA by TAL effectors. Science, 335: 720–723.

Doi K, Izawa T, Fuse T, Yamanouchi U et al., 2004. *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of *Hd1*. Genes Dev. 18,926–936.

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C et al., 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol 26: 702–708.

Du Y, Liu L, Li M, Fang S, Shen X, Chu J, Zhang Z. 2017. *UNBRANCHED3* regulates branching by modulating cytokinin biosynthesis and signaling in maize and rice. New Phytol. doi: 10.1111/nph.14391.

Endo-Higashi N, Izawa T. 2011. Flowering time genes *Heading date 1* and *Early heading date 1* together control panicle development in rice. Plant Cell Physiol. 52:1083–94.

Fan T, Li X, Yang W, Xia K, Ouyang J, Zhang M. 2015. Rice *osa-miR171c* Mediates Phase Change from Vegetative to Reproductive Development and Shoot Apical Meristem Maintenance by Repressing Four *OsHAM* Transcription Factors. PLoS One. 29:10(5).

Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang DL et al., 2014. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proc Natl Acad Sci USA 111: 4632–4637.

Fischer G, Shah M and van Velthuizen H. 2002. "Climate Change and Agricultural Vulnerability". International Institute for Applied Systems Analysis. Report prepared under UN Institutional Contract Agreement 1113 for World Summit on Sustainable Development. Laxenburg, Austria.

Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM. 1999. Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. Science 283:1911–14.

Forestan C, Varotto S. 2012. The role of PIN auxin efflux carriers in polar auxin transport and accumulation and their effect on shaping maize development. Mol. Plant 5:787–98.

Gaillochet C, Daum G, Lohmann JU. 2015. O cell, where art thou? The mechanisms of shoot meristem patterning. Curr. Opin. Plant Biol. 23:91–97.

Gaj T, Gersbach CA, Barbas CF. 2013. ZFN, TALEN, and CRISPR/Cas-based methods for

genome engineering. Trends Biotechnol 31: 397–405.

Gallavotti A, Yang Y, Schmidt RJ, Jackson D. 2008. The relationship between auxin transport and maize branching. Plant Physiol. 147:1913–23.

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK et al., 2004. The role of *barren stalk1* in the architecture of maize. Nature, 432, 630–635.

Gao H, Zheng XM, Fei G et al., 2013. *Ehd4* encodes a novel and *Oryza*-genus-specific regulator of photoperiodic flowering in rice. PLoS Genetics 9, e1003281.

Greb T, Clarenz O, Schafer E, Muller D, Herrero R, Schmitz G, Theres K. 2003. Molecular analysis of the *LATERAL SUPPRESSOR* gene in *Arabidopsis* reveals a conserved control mechanism for axillary meristem formation. Genes Dev 17: 1175–1187.

Gu B, Zhou T, Luo J et al., 2015. *An-2* encodes a cytokinin synthesis enzyme that regulates awn length and grain production in rice. Mol. Plant, 8, 1635–1650.

Guilinger JP, Thompson DB, Liu DR. 2014. Fusion of catalytically inactive Cas9 to Fok I nuclease improves the specificity of genome modification. Nat Biotechnol 32: 577–582.

Harbers, M. & Carninci, P. 2005. Tag-based approaches for transcriptome research and genome annotation. Nature Methods 2, 495–502.

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K. 2003. Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature 422, 719–722.

Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP et al., 2011. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotech. 29: 731–734.

Horigome A, Nagasawa N, Ikeda K et al., 2009. Rice OPEN BEAK is a negative regulator of class 1 knox genes and a positive regulator of class B floral homeotic gene. Plant J. 58:724–36.

Hoshikawa K. 1989. The growing rice plant: An Anatomical Monograph. Nobunkyo, Tokyo107-Japan.

Hsu PD, Zhang F. 2012. Dissecting neural function using targeted genome engineering technologies. ACS Chem Neurosci 3: 603–610.

Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. 2011. Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol; 29: 699–700.

Huq E, Tepperman JM, and Quail PH. 2000. *GIGANTEA* is a nuclear protein involved in phytochrome signaling in *Arabidopsis*. Proc. Natl. Acad. Sci. USA 97 9789–9794.

Hutchison CA, III. 2007. DNA sequencing: bench to bedside and beyond. Nucleic Acids Res. 35:6227–6237.

Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD et al., 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31: 227–229.

Ikeda K, Ito M, Nagasawa N, Kyozuka J, Nagato Y. 2007. Rice *ABERRANT PANICLE ORGANIZATION 1*, encoding an F-box protein, regulates meristem fate. Plant J 51(6): 1030–1040.

Ikeda K, Sunohara H and Nagato Y. 2004. Developmental course of inflorescence and spikelet in rice. Breed. Sci. 54, 147–156.

Ikeda-Kawakatsu K, Maekawa M, Izawa T, Itoh J, Nagato Y. 2012. *ABERRANT PANICLE ORGANIZATION 2/RFL*, the rice ortholog of *Arabidopsis LEAFY*, suppresses the transition from inflorescence meristem to floral meristem through interaction with APO1. Plant J. 69(1):168-80.

Ishino Y, Shinagawa H, Makino K, Amemura M & Nakata A. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. 169, 5429–5433.

Ito S, Song YH, Josephson-Day AR, Miller RJ, Breton G et al., 2012. FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator *CONSTANS* in *Arabidopsis*. Proc. Natl. Acad. Sci. U.S.A. 109, 3582–3587.

Itoh H, Nonoue Y, Yano M, Izawa T. 2010. A pair of floral regulators sets critical day length for Hd3a florigen expression in rice. Nat. Genet. 42, 635–638.

Jansen R, Embden JD, Gaastra W & Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. 43, 1565–1575.

Jasinski S, Piazza P, Craft J, Hay A, Woolley L et al., 2005. KNOX action in *Arabidopsis* is mediated by coordinate regulation of cytokinin and gibberellin activities. Curr. Biol. 15:1560–65.

Jeong S, Trotochaud AE, Clark SE. 1999. The *Arabidopsis CLAVATA2* gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. Plant Cell 11:1925–33.

Jia HG, Wang N. 2014. Targeted genome editing of sweet orange using Cas9/sgRNA. PloS one 9: e93806.

Jiao Y, Wang Y, XueD,Wang J, Yan M et al., 2010. Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. Nat. Genet. 42:541–44.

Joung JK, Sander JD. 2013. TALENs: a widely applicable technology for targeted genome editing. Nature reviews Mol Cell Biol. 14: 49–55.

Kardailsky I, Shukla, VK, Ahn JH, Dagenais N et al., 1999. Activation tagging of the floral inducer *FT*. Science 286 1962–1965.

Kayes JM, Clark SE. 1998. *CLAVATA2*, a regulator of meristem and organ development in Arabidopsis. Development 125:3843–51.

Kim SL, Lee S, Kim HJ, Nam HG, An G. 2007. *OsMADS51* is a short-day flowering promoter that functions upstream of *Ehd1*, *OsMADS14*, and *Hd3a*. Plant Physiology 145, 1484–1494.

Kim YJ, Cha J, Chandrasegaran S. 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain Proc Natl Acad Sci USA, 93:1156–1160.

Kinoshita A, Nakamura Y, Sasaki E, Kyozuka J, Fukuda H, sawa S. 2007. Gain-of-function phenotypes of chemically synthetic CLAVATA3/ESR-related (CLE) peptides in *Arabidopsis thaliana* and *Oryza sativa*. Plant Cell Physiol. 48(12):1821-5.

Kirchoff BK, Bockhoff RC. 2013. Inflorescences: concepts, function, development and evolution. Ann. Bot. 112(8):1471-1476.

Klein J, Saedler H and Huijser P. 1996. A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene *SQUAMOSA*. Mol. Gen. Genet. 250, 7–16.

Kobayashi K, YasunoN, Sato Y, Yoda M, Yamazaki R et al., 2012. Inflorescence meristem identity in rice is specified by overlapping functions of three AP1/FUL-like MADS box genes and *PAP2*, a SEPALLATA MADS box gene. Plant Cell 24:1848–59.

Kodzius R et al., CAGE: cap analysis of gene expression. Nature Methods 3, 211–222 (2006).

Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H et al., 2003a. *LAX* and *SPA*: major regulators of shoot branching in rice, Proc. Natl Acad. Sci. USA. 100:11765-11770.

Komatsu K, 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. Dev Biol. 231(2):364-73.

Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I et al., 2003. LAX and SPA: major

regulators of shoot branching in rice. Proc. Natl. Acad. Sci. USA 100:11765–70.

Komiya R, Yokoi S, Shimamoto K. 2009. A gene network for long-day flowering activates *RFT1* encoding a mobile flowering signal in rice. Development 136:3443–50.

Kurakawa T, Ueda N, Maekawa M, Kobayashi K, Kojima M, Nagato, Y, Sakakibara H and Kyozuka J. 2007. Direct control of shoot meristem activity by a cytokinin-activating enzyme. Nature, 445, 652–655.

Kuromori T, Wada T, Kamiya A, Yuguchi M, Yokouchi T, Imura Y, Takabe H, Sakurai T, Akiyama K et al., 2006. A trial of phenome analysis using 4000 Ds-insertional mutants in gene-coding regions of *Arabidopsis* Plant J., 47:640–651.

Laux T, Mayer KF, Berger J, Jürgens G. 1996. The *WUSCHEL* gene is required for shoot and floral meristem integrity in *Arabidopsis*. Development. 122(1):87-96.

Lee DY, An G. 2012. Two AP2 family genes, *SUPERNUMERARY BRACT* (*SNB*) and *OsINDETERMINATE SPIKELET 1* (*OsIDS1*), synergistically control inflorescence architecture and floral meristem establishment in rice, Plant J. 69:445-461.

Lee DY, Lee J, Moon S, Park SY and An G. 2007. The rice heterochronic gene *SUPERNUMERARY BRACT* regulates the transition from spikelet meristem to floral meristem. Plant J. 49, 64–78.

Lee YS, Jeong DH, Lee DY et al. 2010. OsCOL4 is a constitutive flowering repressor upstream of *Ehd1* and downstream of *OsphyB*. The Plant Journal 63, 18–30.

Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J et al., 2013. Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. Nat Biotechnol 31: 688–691.

Li L, Piatek MJ, Atef A, Piatek A, Wibowo A, Fang X et al., 2012. Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. Plant Mol Biol .78: 407–416.

Li S, Zhao B, Yuan D, Duan M, Qian Q et al., 2013. Rice zinc finger protein DST enhances grain production through controlling *Gn1a/OsCKX2* expression. Proc. Natl. Acad. Sci. USA 110:3167–72.

Li T, Liu B, Spalding MH, Weeks DP, Yang B. 2012. High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotech. 30: 390–392.

Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D et al., 2003. Control of tillering in

rice, Nature. 422:618-621.

Liu C, Teo ZWN, Bi Y, Song S, Xi W et al., 2013. A conserved genetic pathway determines inflorescence architecture in Arabidopsis and rice. Dev. Cell 24:612–22.

Long JA, Moan EI, Medford JI, Barton MK. 1996. A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. Nature 379:66–69.

Lor VS, Starker CG, Voytas DF, Weiss D, Olszewski NE. 2014. Targeted mutagenesis of the tomato *PROCERA* gene using transcription activator-like effector nucleases. Plant Physiol. 166: 1288–1291.

Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science, 335: 716–719.

Makarova KS, Grishin NV, Shabalina SA, Wolf YI & Koonin EV. 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct 1, 7.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE et al., 2013. RNA-guided human genome engineering via Cas9. Science 339: 823–826.

Mantegazza O, Gregis V, Chiara M, Selva C, Leo G, Horner DS, Kater MM. 2014. Gene coexpression patterns during early development of the native Arabidopsis reproductive meristem: novel candidate developmental regulators and patterns of functional redundancy. Plant J. 79(5):861-77.

Mardis ER. 2013. Next-generation sequencing platforms. Annu. Rev. Anal. Chem. 6, 287–303.

Margulies M, Egholm M, Altman WE, Attiya S et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380.

Matsubara K Yamanouchi U Wang ZX Minobe Y Izawa T Yano M . 2008. *Ehd2*, a rice ortholog of the maize *INDETERMINATE1* gene, promotes flowering by up-regulating Ehd1. Plant Physiology 148, 1425–1435.

Matsubara K, Yamanouchi U, Nonoue Y, Sugimoto K, Wang ZX, Minobe Y, Yano M. 2011. *Ehd3*, encoding a plant homeodomain finger-containing protein, is a critical promoter of rice flowering. The Plant Journal 66, 603–612.

Maxam AM, and Gilbert W. 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U. S. A. 74:560–564.

McCallum CM, Comai L, Greene EA, Henikoff S. 2000. Targeted screening for induced mutations Nat. Biotechnol. 18:455–457.

McSteen P, Hake S. 2001. *barren inflorescence2* regulates axillary meristem development in the maize inflorescence. Development 128:2881–91.

McSteen P. 2009. Hormonal regulation of branching in grasses. Plant Physiol. 149:46–55.

Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. 2008. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat Biotechnol 26: 695–701.

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X et al., 2013. Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res 23: 1233–1236.

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF et al., 2011. A TALE nuclease architecture for efficient genome editing. Nat Biotech. 29: 143–148.

Miura K, Ikeda M, Matsubara A, Song X-J, ItoM et al., 2010. *OsSPL14* promotes panicle branching and higher grain productivity in rice. Nat. Genet. 42:545–49.

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J & Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60, 174–182.

Mojica FJ, Diez-Villasenor C, Soria E & Juez G. 2000. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol. Microbiol. 36, 244–246.

Mol JN, Van der Krol AR, Van Tunen AJ, Van Blokland R, De Lange P, Stuitje AR. 1990. Regulation of plant gene expression by antisense RNA. FEBS Lett. 268:427–430.

Moscou MJ, Bogdanove AJ. 2009. A simple cipher governs DNA recognition by TAL effectors. Science 326: 1501.

Muller R, Bleckmann A, Simon R. 2008. The receptor kinase CORYNE of *Arabidopsis* transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. Plant Cell 20:934–46.

Mutz K-O, Heilkenbrinker A, Lönne M, Walter J-G, Stahl F. 2013. Transcriptome analysis using next-generation sequencing. Curr Opin Biotechnol. 24:22–30.

Nakagawa M, Shimamoto K and 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. Plant J. 29, 743–750.

Nardmann J, Werr W. 2006. The shoot stem cell niche in angiosperms: expression patterns of *WUS* orthologues in rice and maize imply major modifications in the course of mono- and dicot evolution. Mol. Biol. Evol. 23:2492–504.

Nowrousian M. 2010. Next-Generation Sequencing Techniques for Eukaryotic Microorganisms: Sequencing-Based Solutions to Biological Problems. EUKARYOTIC CELL 9(9):1300–1310.

Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y. 2008. *Arabidopsis* CLV3 peptide directly binds CLV1 ectodomain. Science 319:294.

Ogo Y, Kakei Y, Itai RN, Kobayashi T, Nakanishi H, Nishizawa NK. 2014. Tissue-specific transcriptional profiling of iron-deficient and cadmium-stressed rice using laser capture microdissection. Plant Signaling & Behavior 9:e29427.

Ohmori Y, Tanaka W, Kojima M, Sakakibara H, Hirano HY. 2013. *WUSCHEL-RELATED HOMEOBOX4* is involved in meristem maintenance and is negatively regulated by the CLE gene *FCP1* in rice. Plant Cell 25: 229–241.

Ohyama K, Shinohara H, Ogawa-Ohnishi M, Matsubayashi Y. 2009. A glycopeptide regulating stem cell fate in Arabidopsis thaliana. Nat. Chem. Biol. 5:578–80.

Osakabe K, Osakabe Y, Toki S. Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases. Proc Natl Acad Sci USA 2010; 107: 12034–12039.

Park DH, Somers DE, Kim YS, Choy YH, Lim HK, Soh MS, Kim HJ, Kay SA, Nam HG. 1999. Control of Circadian Rhythms and Photoperiodic Flowering by the *Arabidopsis GIGANTEA* Gene. Science 285:1579-1582.

Park SJ, Kim SL, Lee S, Je BI, Piao HL et al., 2008. Rice *Indeterminate 1* (*OsId1*) is necessary for the expression of *Ehd1* (*Early heading date 1*) regardless of photoperiod. Plant J. 56:1018–29.

Pautler M, Tanaka W, Hirano H-Y, Jackson D. 2013. Grass meristems I: shoot apical meristem maintenance, axillary meristem determinacy and the floral transition. Plant Cell Physiol. 54:302–12.

Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O et al., 2008. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat. Biotech. 26: 808–816.

Poething RS. 2003. Phase change and the regulation of developmental timing in plants. Science. 301(5631):334-6.

Pourcel C, Salvignol G & Vergnaud G. 2005. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology 151, 653–663.

Putterill J, Robson F, Lee K, Simon R, and Coupland G. 1995. The *CONSTANS* gene of arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80, 847–857.

Reinartz J et al., Massively parallel signature sequencing (MPSS) as a tool for in depth quantitative gene expression profiling in all organisms. Brief. Funct. Genomic Proteomic 1, 95–104 (2002).

Ren D, Li Y, Zhao F, Sang X, Shi J, Wang N, Guo S, Ling Y, Zhang C, Yang Z et al., 2013. *MULTI-FLORET SPIKELET1*, which encodes an AP2/ERF protein, deter- mines spikelet meristem fate and sterile lemma identity in rice. Plant Physiol. 162, 872–884.

Sakamoto T, Sakakibara H, Kojima M, Yamamoto Y, Nagasaki H et al., 2006. Ectopic Expression of KNOTTED1-Like Homeobox Protein Induces Expression of Cytokinin Biosynthesis Genes in Rice. Plant Physiol. 142(1): 54–62.

Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK et al., 2011. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 2011; 29: 697–698.

Sanger F, Nicklen S, and Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. 74:5463–5467.

Sawa M, Nusinow DA, Kay SA, and Imaizumi T. 2007. *FKF1* and *GIGANTEA* complex formation is required for day-length measurement in Arabidopsis. Science 318, 261–265.

Sawai S, Ohyama K, Yasumoto S, Seki H, Sakuma T, Yamamoto T et al., 2014. Sterol side chain reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. Plant cell, 26: 3763–3774.

Schoof H, Lenhard M, Haecker A, Mayer KFX, Jurgens G et al., 2000. The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes. Cell 100:635–44.

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z et al., 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol. 31: 686–688.

Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z et al., 2013. Generation of gene-modified mice via

Cas9/RNA-mediated gene targeting. Cell Res. 23: 720–723.

Shiraki T et al., 2003. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc. Natl Acad. Sci. USA 100, 15776–15781.

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE et al., Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature 2009; 459: 437–441.

Skirpan A, Culler AH, Gallavotti A, Jackson D, Cohen JD, McSteen P. 2009. *BARREN INFLORESCENCE2* interaction with *ZmPIN1a* suggests a role in auxin transport during maize inflorescence development. Plant Cell Physiol. 50:652–57.

Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res. 28(17):3361-9.

Smith NA, Singh SP, Wang MB, Stoutjesdijk PA, Green AG, Waterhouse PM. 2000. Total silencing by intron-spliced hairpin RNAs Nature 407:319–320.

Somssich M, Je BI, Simon R, Jackson D. 2016. CLAVATA-WUSCHEL signaling in the shoot meristem. Development. 143(18):3238-48.

Stebbins, GL. 1974. Flowering plants: Evolution above the species level. Cambridge, Mass, Belknap Press of Harvard University Press.

Suzaki T, Ohneda M, Toriba T, Yoshida A, Hirano HY. 2009. *FON2 SPARE1* redundantly regulates floral meristem maintenance with *FLORAL ORGAN NUMBER2* in rice. PLoS Genet. 5(10):e1000693.

Suzaki T, Sato M, Ashikari M, Miyoshi M, Nagato Y, Hirano H-Y. 2004. The gene *FLORAL ORGAN NUMBER1* regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to *Arabidopsis CLAVATA1*. Development 131:5649–57.

Suzaki T, Toriba T, Fujimoto M, Tsutsumi N, Kitano H, Hirano H-Y. 2006. Conservation and diversification of meristem maintenance mechanism in *Oryza sativa*: function of the *FLORAL ORGAN NUMBER2* gene. Plant Cell Physiol. 47:1591–602.

Suzaki T, Yoshida A, Hirano H-Y. 2008. Functional diversification of CLAVATA3-related CLE proteins in meristem maintenance in rice. Plant Cell 20: 2049–2058.

Tabuchi H, Zhang Y, Hattori S, OmaeM et al., 2011. *LAX PANICLE2* of rice encodes a novel nuclear protein and regulates the formation of axillary meristems. Plant Cell 23:3276–87.

Tan J, Jin M, Wang J et al . 2016. *OsCOL10*, a CONSTANS-like gene, functions as a flowering time repressor downstream of *Ghd7* in rice. Plant & Cell Physiology 57, 798–812.

Tanaka W, Ohomori Y, Ushijima T, Matsusaka H, Matsushita T, Kumamaru T, Kawano S, Hirano HY. 2015. Axillary Meristem Formation in Rice Requires the WUSCHEL Ortholog *TILLERS ABSENT1*. Plant Cell. 27(4):1173-84.

Tang TH et al., 2002. Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. Proc. Natl Acad. Sci. USA 99, 7536–7541.

Taoka K, Ohki I, Tsuji H, Furuita K, Hayashi K et al., 2011. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. Nature 476:332–35.

Tesson L, Usal C, Menoret S, Leung E, Niles BJ, Remy S et al., 2011. Knockout rats generated by embryo microinjection of TALENs. Nat Biotech. 29: 695–696.

Thao NP, Tran LS. 2016. Enhancement of Plant Productivity in the Post-Genomic Era. Current Genomics 17(4):295-6.

Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK et al., High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 2009; 459: 442–445.

Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D et al., 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol 32: 569–576.

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 23:4368–81.

United Nations. 2015. World Population Prospects; The 2015 Revision. New York: United Nations.

Velculescu VE, Zhang L, Vogelstein B & Kinzler KW. 1995. Serial analysis of gene expression. Science 270, 484–487.

Vollbrecht E, Reiser L, Hake S. 2000. Shoot meristem size is dependent on inbred background and presence of the maize homeobox gene, *knotted1*. Development 127:3161–72.

Vollbrecht E, Springer PS, Goh L, Buckler ES, Martienssen R. 2005 Architecture of floral branch systems in maize and related grasses. Nature. 436(7054):1119-26.

Wang N, Long T, Yao W, Xiong L, Zhang Q, Wu C. 2013. Mutant resources for the functional analysis of the rice genome. Mol. Plant. 6(3):596–604.

Wang Y, Cheng X, Shan Q et al., 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotech. 32: 947–951.

Wei X, Xu J, Guo H, Jiang L, Chen S et al., 2010. *DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. Plant Physiol. 153:1747–58.

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M et al., 2005. Integration of spatial and temporal information during floral induction in Arabidopsis. Science 309:1056–1059.

Wu C, Li X, Yuan W, Chen G, Kilian A, Li J, Xu C, Li X, Zhou DX et al., 2003. Development of enhancer trap lines for functional analysis of the rice genome. The Plant Journal 35: 418–427.

Wu CY, You CJ, Li CS, Long T, Chen GX, Byrne ME, Zhang QF. 2008. RID1, encoding a Cys2/His2-type zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. Proc. Natl. Acad. Sci. USA, 105, 12915-20.

Wu R, and AD, Kaiser. 1968. Structure and base sequence in the cohesive ends of bacteriophage lambda DNA. J. Mol. Biol. 35:523–537.

Wyman C, Kanaar R. DNA double-strand break repair: all's well that ends well. Annu Rev Genet 2006; 40: 363–383.

Xie K, Yang Y. 2013. RNA-guided genome editing in plants using a CRISPR-Cas system. Mol Plant 6: 1975–1983.

Xin W, Wang Z, Liang Y, Wang Y, Hu Y. 2017. Dynamic expression reveals a two-step patterning of *WUS* and *CLV3* during axillary shoot meristem formation in Arabidopsis. J Plant Physiol. 214:1-6.

Xu M, Zhu L, ShouH, Wu P. 2005. A PIN1 family gene, *OsPIN1*, involved in auxin-dependent adventitious root emergence and tillering in rice. Plant Cell Physiol. 46:1674–81.

Xue W, Xing Y, Weng X, Zhao Y, Tang W et al., 2008. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. Nat. Genet. 40:761–67.

Yamamoto Y, Kamiya N, Morinaka Y, Matsuoka M, Sazuka T. 2007. Auxin biosynthesis by the YUCCA genes in rice. Plant Physiol. 143:1362–1371.

Yan W-H, Wang P, Chen H-X, Zhou H-J, Li Q-P et al., 2011. A major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. Mol. Plant 4:319–30.

Yang Y, Li Y, Wu C. 2013. Genomic resources for functional analyses of the rice genome. Curr

Opin Plant Biol. 16: 157–163.

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T. 2000. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. Plant Cell 12, 2473–2484.

Yoshida A, Ohmori Y, Kitano H, Taguchi-Shiobara F, Hirano HY. 2012. *ABERRANT SPIKELET AND PANICLE1*, encoding a TOPLESS-related transcriptional co-repressor, is involved in the regulation of meristem fate in rice. Plant J. 70:327–39.

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y et al., 2013. *TAWAWA1*, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. Proc. Natl. Acad. Sci. USA 110:767–72.

Zhang L, Liu G, Jia J et al . 2016. The wheat MYB-related transcription factor *TaMYB72* promotes flowering in rice. Journal of Integrative Plant Biology 58, 701–704.

Zhao J, Chen H, Ren D, Tang H, Qiu R, Feng J et al., 2015. Genetic interactions between diverged alleles of *Early heading date 1* (*Ehd1*) and *Heading date 3a* (*Hd3a*)/ *RICE FLOWERING LOCUS T1* (*RFT1*) control differential heading and contribute to regional adaptation in rice (*Oryza sativa*). New Phytol. 208(3):936-48

Zhao Y, Christensen SK, Fankhauser C, Cashman JR, Cohen JD, Weigel D, Chory J. 2001. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. Science 291:306–309.

Zhao Y, Hull AK, Gupta NR, Goss KA, Alonso J, Ecker JR, Normanly J, Chory J, Celenza JL. 2002. Trp-dependent auxin biosynthesis in Arabidopsis: involvement of cytochrome P450s *CYP79B2* and *CYP79B3*. Genes Dev.16:3100–3112.

Zhao Y. 2010. Auxin biosynthesis and its role in plant development. Annu. Rev. Plant Biol. 61:49–64.

Zhou Y, Liu X, Engstrom EM, Nimchuk ZL, Pruneda-Paz JL, Tarr PT, Yan A, Kay SA, Meyerowitz EM. 2015. Control of plant stem cell function by conserved interacting transcriptional regulators. Nature. 517(7534):377-80.

Zhu S, Wang J, Cai M, Zhang H, Wu F, Xu Y, Li C, Cheng Z et al., 2017. The OsHAPL1-DTH8-Hd1 complex functions as the transcription regulator to repress heading date in rice. J Exp Bot. 68 (3): 553-568.

Zhu, QH, Upadhyaya, NM, Gubler, F and 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.

PART-II

2. PUBLICATIONS

Title: Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection

Thomas W. R. Harrop[1,†], **Israr Ud Din**[2,3,†], Veronica Gregis[2], Michela Osnato[2,‡], Stefan Jouannic[1],Helene Adam[1,*] and Martin M. Kater [2,3,*]

1. Institut de Recherche pour le Developpement, UMR DIADE, 911 Avenue Agropolis, 34394 Montpellier, France,
2. Dipartimento di Bioscienze, Universita degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy,
3. Dipartimento di Scienze Farmacologiche e Biomolecolari, Universita degli Studi di Milano, Via Celoria 26, 20133 Milano,Italy

Received 19 December 2015; revised 11 February 2016; accepted 15 February 2016; published online 1 March 2016.

†These authors contributed equally to this work.

‡Present address: Centre for Research in Agricultural Genomics, Campus UAB, 08193 Cerdanyola, Spain.

*For correspondence (e-mails martin.kater@unimi.it or helene.adam@ird.fr).

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2016) **86,** 75–88 doi: 10.1111/tpj.13147

RESOURCE

Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection

Thomas W. R. Harrop[1,†], Israr Ud Din[2,3,†], Veronica Gregis[2], Michela Osnato[2,‡], Stefan Jouannic[1], Hélène Adam[1,*] and Martin M. Kater[2,3,*]

[1]*Institut de Recherche pour le Développement, UMR DIADE, 911 Avenue Agropolis, 34394 Montpellier, France,*
[2]*Dipartimento di Bioscienze, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy, and*
[3]*Dipartimento di Scienze Farmacologiche e Biomolecolari, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy*

Received 19 December 2015; revised 11 February 2016; accepted 15 February 2016; published online 1 March 2016.
*For correspondence (e-mails martin.kater@unimi.it or helene.adam@ird.fr).
†These authors contributed equally to this work.
‡Present address: Centre for Research in Agricultural Genomics, Campus UAB, 08193 Cerdanyola, Spain.

SUMMARY

In rice, inflorescence architecture is established at early stages of reproductive development and contributes directly to grain yield potential. After induction of flowering, the complexity of branching, and therefore the number of seeds on the panicle, is determined by the activity of different meristem types and the timing of transitions between them. Although some of the genes involved in these transitions have been identified, an understanding of the network of transcriptional regulators controlling this process is lacking. To address this we used a precise laser microdissection and RNA-sequencing approach in *Oryza sativa* ssp. *japonica* cv. Nipponbare to produce quantitative data that describe the landscape of gene expression in four different meristem types: the rachis meristem, the primary branch meristem, the elongating primary branch meristem (including axillary meristems), and the spikelet meristem. A switch in expression profile between apical and axillary meristem types followed by more gradual changes during transitions in axillary meristem identity was observed, and several genes potentially involved in branching were identified. This resource will be vital for a mechanistic understanding of the link between inflorescence development and grain yield.

Keywords: inflorescence development, rice, reproductive meristems, RNA sequencing.

INTRODUCTION

Agricultural development is essential to ensure food production and security for a growing population (Borlaug, 2007). Rice is a staple food for over half the world's population, including many developing countries. A sustainable increase in the production of rice under the constraints of a changing climate and diminishing water and land availability will require plants with improved grain output, making the establishment of high-yield rice varieties a goal of modern breeding programmes (Peng *et al.*, 2008). Rice yield is a complex trait influenced by genetic and epigenetic factors, and is progressively defined during the life cycle of the plant, first during the vegetative phase, in which the number of fertile tillers is established, and then during the reproductive phase and grain-filling phase (Ikeda *et al.*, 2004).

The branched inflorescence of rice is a compound raceme, classified as a panicle, and is composed of a rachis (the main axis), primary branches, higher-order branches, and spikelets. The number of seeds on the panicle is influenced by the complexity and arrangement of branches and spikelets. The establishment and activity of apical and axillary meristems conditions branching in two phases during the early stages of panicle development: the timing of rachis meristem abortion determines the number of primary branches, whilst the transition of indeterminate branch meristems to determinate spikelet meristems specifies the complexity of branching. After spikelet differentiation, branching complexity is fixed and the rachis and branches elongate rapidly and heading and flowering occur (Ikeda *et al.*, 2004). The landscape of gene expres-

sion is a characteristic that differentiates meristem types and is involved in the control of meristem identity transitions. Mutant analyses and mapping of quantitative trait loci have identified a number of genes required for the initiation and development of panicles, as well as genes that control the number and size of grains and panicles (Xing and Zhang, 2010; Wang and Li, 2011). Some of these genes are involved in the patterning of axillary meristems and panicle branching, such as the nuclear regulatory factor-encoding genes *MONOCULM 1*, *LAX PANICLE 1* and *LAX PANICLE 2* (Wang and Li, 2011). A large set of genes related to floral development, which may also affect panicle architecture, has been identified (Yoshida and Nagato, 2011). Some of these genes may have been under selection during domestication or are associated with crop improvement in Asian rice (Xing and Zhang, 2010; He *et al.*, 2011; Wang and Li, 2011; Xu *et al.*, 2012; Ikeda *et al.*, 2013).

Understanding the events during development that determine panicle characteristics such as branching complexity and its plasticity will be vital for sustainable improvement of rice yield potential using targeted breeding programmes. Despite advances in the characterization of individual genes and their interactions, a complete understanding of the control of panicle morphology and grain yield will require mechanistic studies that explain the interactions of gene regulatory network components with each other and with the environment (Azpeitia *et al.*, 2013). One step towards this goal is to describe the differences in gene expression between different meristem types during development. In this article, we describe the measurement and analysis of genome-wide expression in meristematic tissues from the early stages of panicle development in *Oryza sativa* ssp. *japonica* cv. Nipponbare, using a precise laser microdissection and RNA-sequencing approach.

RESULTS

Panicle morphology and sampling for laser microdissection

In rice, the start of the reproductive phase and subsequent initiation of inflorescence development involves fast morphological transformations (described by Ikeda *et al.*, 2004). A detailed histological analysis of rice inflorescences was performed to select four morphologically distinct inflorescence meristem types for laser microdissection (LMD). After differentiation of the flag leaf, the first stage of reproductive development is the conversion of the shoot apical meristem (SAM) to the rachis meristem (RM). The first bract primordium was produced opposite the flag leaf (Figure 1a). After the establishment of the RM, some cells differentiated into primary branch meristem (PBM) in the axils of newly developed bracts (Figure 1b). After the formation of primary branches, bract growth ceased and primary branches elongated (ePBM) (Figure 1c). During primary branch elongation, the PBM can give rise to axillary meristems (AM), which may differentiate into secondary and higher-order branches or be converted directly to spikelet meristem (SM). For this analysis, SM differentiation was considered as the final stage of panicle development (Figure 1d). The PBM and secondary branch meristem (SBM) are both converted to a terminal SM. Each terminal or axillary SM produces one floret meristem (FM), which differentiates into a single floret. The RM, PBM and ePBM/AM stages are the indeterminate stages in which meristematic cells are maintained, whilst the SM has a determinate fate in which the stem cell activity will be lost and from which florets will differentiate (Ikeda *et al.*, 2004).

Genome-wide expression analysis of inflorescence development in rice

To investigate gene expression during the development of the rice inflorescence, LMD was used to collect meristematic tissues from the RM, PBM, ePBM/AM and SM of early panicles (Figure 1e–l). RNA isolated from the meristem samples was amplified and used to produce cDNA libraries for sequencing. RNA input for amplification was between 2.3 and 79.6 ng and the average RNA integrity number (RIN) was 7.2 (Table S1). Initially, two biological replicates were prepared for each stage, but after performing a principal components analysis (PCA) on transformed expression values, two libraries that were produced from degraded RNA ($RIN < 6.5$) were excluded from the analysis (Figure S1). Following this, further samples were prepared to provide a total of three biological replicates with intact RNA ($RIN \geq 7$) for each meristem type. A single sequencing library was produced from each biological replicate.

At least 53 million single-end, 50-base reads were produced from each library, yielding between 16 and 40 million uniquely mapped reads within genes, after excluding reads likely to have originated from rRNA and tRNA. Using strict cut-offs for gene expression, 11 652 unique genes were detected in two or more libraries from at least one stage (Data S1). As a preliminary assessment of the RNA sequencing (RNA-seq) results, the qualitative and quantitative expression of 20 genes were compared with previously reported patterns. 16 of the 20 genes were detected in two or more libraries from at least one meristem type in the LMD dataset (Figure S2 and Table S2). The expression of a further four uncharacterized genes that were detected in specific meristem types by RNA-seq was confirmed by RNA *in situ* hybridization. Each of the four genes was detected in the same meristem types by both methods (Figure 2). These results suggest that the tissues used for RNA sequencing were accurately dissected and represent the intended meristem types.

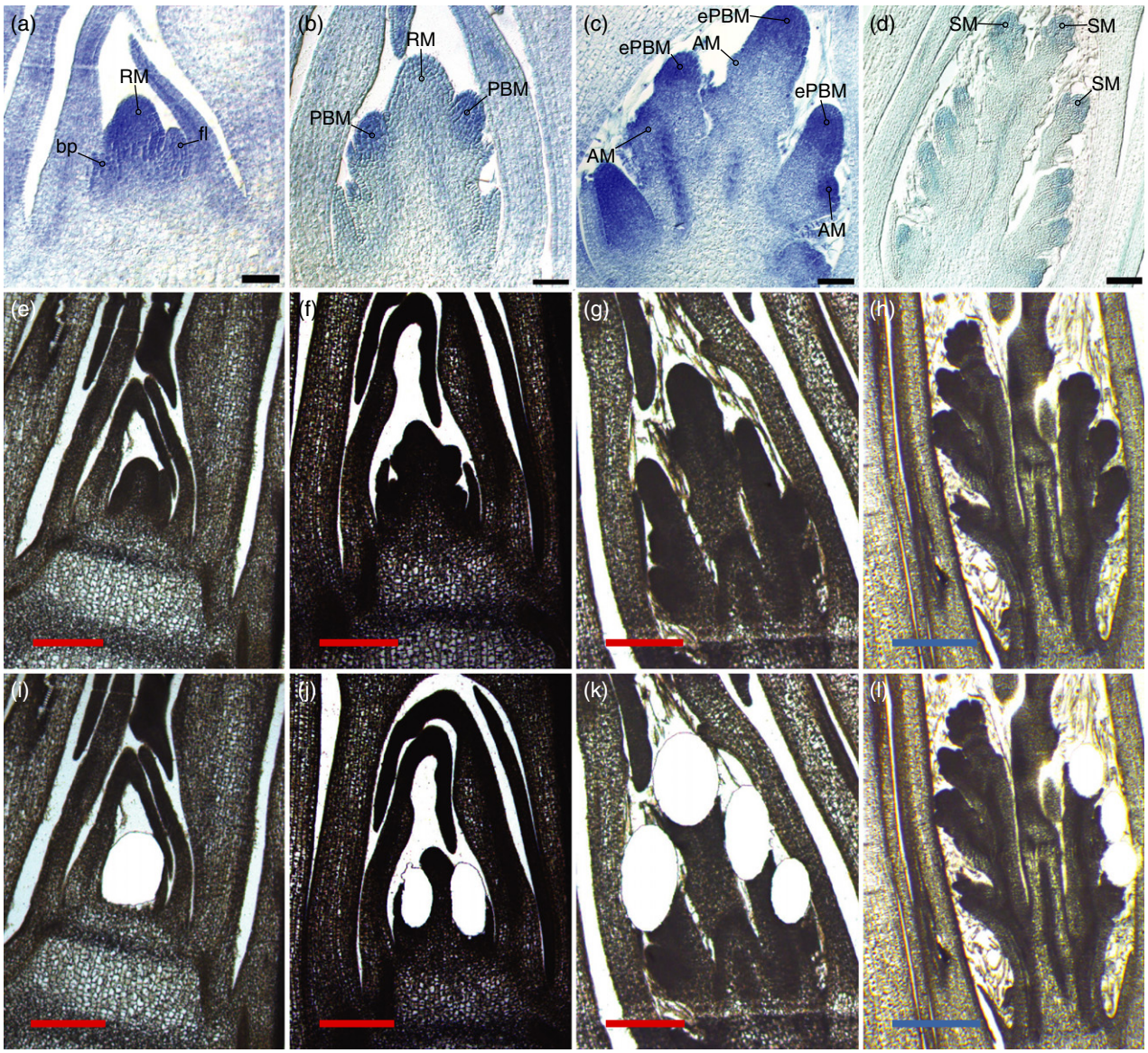

Figure 1. Morphology of the early stages of inflorescence development and laser microdissection of meristem samples.
(a–d) Toluidine blue-stained sections of developing panicles at (a) rachis meristem (RM), (b) primary branch meristem (PBM), (c) elongating primary branch meristem with axillary meristem (ePBM/AM), and (d) spikelet meristem (SM) stages of differentiation. The position of the first bract primordium (bp) and flag leaf (fl) are indicated in (a).
(e–l) Laser microdissection (LMD) samples were collected from RM (e, i), PBM (f, j), ePBM/AM (g, k) and SM (h, l). Images show the samples before (e–h) and after (i–l) dissection. Scale bars represent 50 μm (a–c), 100 μm (d), 200 μm (e–g, i–k), or 320 μm (h, l).

Patterns of gene expression during inflorescence development

To recover common expression patterns, read counts were transformed using the variance-stabilizing transformation (VST) included in the DESeq2 software package (Love *et al.*, 2014). Detected genes were ranked by variance, and standardized, VST-transformed read counts for the 3884 genes with the highest variance (33% of expressed genes) were clustered using the fuzzy *c*-means algorithm implemented in the Mfuzz package (Kumar and Futschik, 2007). To determine the number of cluster cores (*c*), the minimum distance between cluster centroids, the formation of empty clusters, and PCA plots of cluster members were monitored for clusters produced with *c* values between 2 and 25 (Figure S3), leading to the use of eight cluster cores (Figure 3). Several common patterns were recovered by clustering: genes that increase or decrease in expression steadily during development (clusters 1 and 4), genes that change expression between the apical meristem (RM) and axillary meristem (PBM to SM) samples (clusters 7 and 5) and genes that

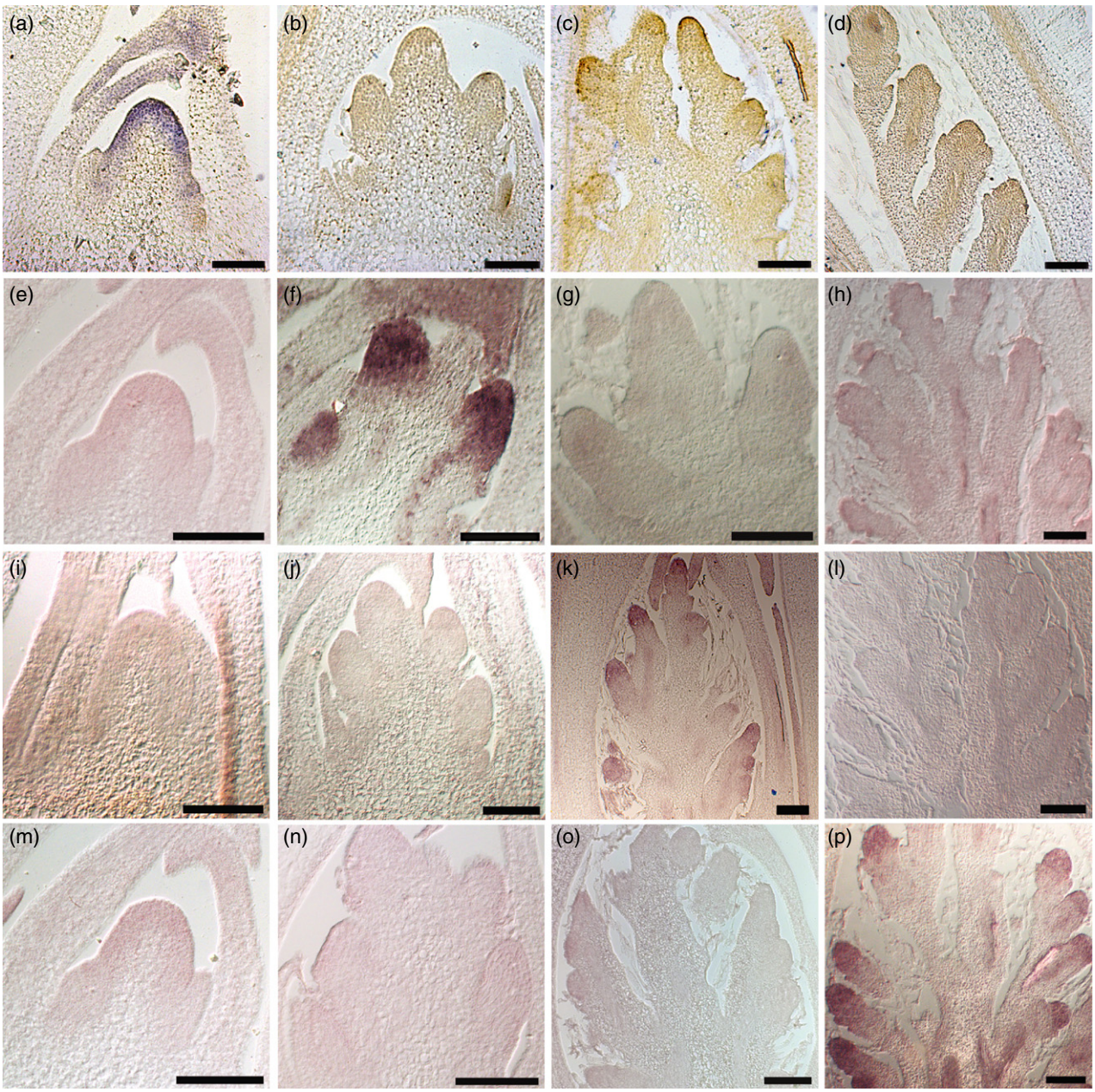

Figure 2. RNA *in situ* hybridization analysis to confirm the specific expression patterns of four genes detected in the RNA-sequencing dataset.
(a–p) Expression of *LOC_Os09g27730* (a–d), *LOC_Os01g04670* (e–h), *LOC_Os10g04270* (i–l) and *LOC_Os10g059908* (m–p) was analysed at the RM (a, e, i, m), PBM (b, f, j, n), ePBM/AM (c, g, k, o) and SM (d, h, l, p) stages. Each gene was detected in the same stage by *in situ* hybridization as by RNA sequencing. The scale bars represent 100 μm.

change expression gradually over the course of changes in axillary meristem identity (clusters 6 and 2). There were also two weaker clusters containing fewer genes, which had complex expression patterns involving changes in expression particular to the PBM (clusters 3 and 8).

Dynamic expression of transcription factor families

Examination of the clustered genes suggested that transcription factor (TF) genes were overrepresented in most clusters (Data S2). Using lists available in the Plant Transcription Factor Database (Pérez-Rodríguez *et al.*, 2010), geneset enrichment analysis (GSEA) was used to provide an overview of expression of TFs and other regulators in the meristem samples (Figure 4). Unlike soft clustering, which was used to group genes based on shared expression dynamics, this analysis was used to investigate the overall strength of expression of the families based on $\log_2$-fold changes (L_2FC) in read count of their members.

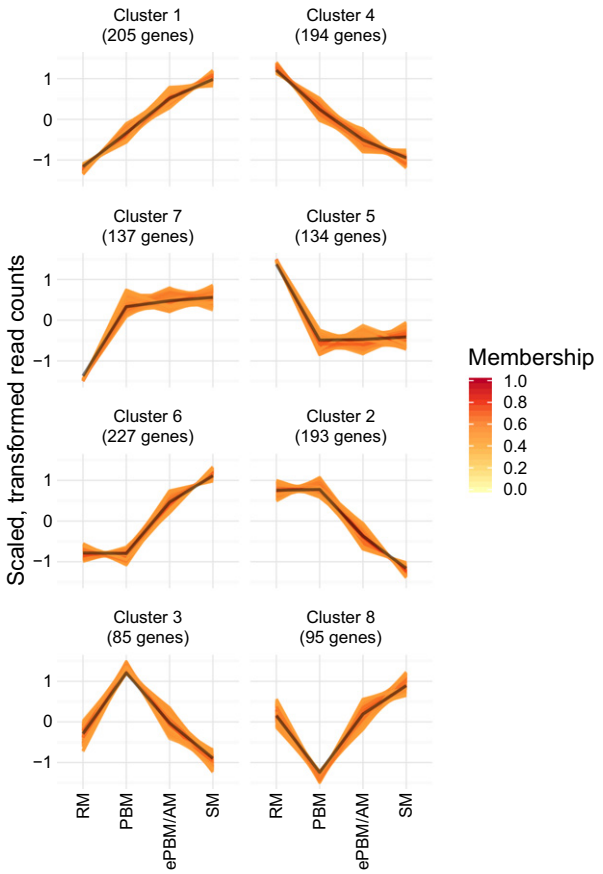

Figure 3. Common patterns of gene expression in developing inflorescences.
Fuzzy *c*-means clustering of normalised, variance-stabilized read counts was used to recover eight common patterns of expression. Each line describes the expression pattern of one gene, with the gene's membership to the cluster represented by the colour of the line. The core values for each cluster are plotted in black. The ordering of the panels in the plot was arranged to enable side-by-side comparison of complementary clusters.

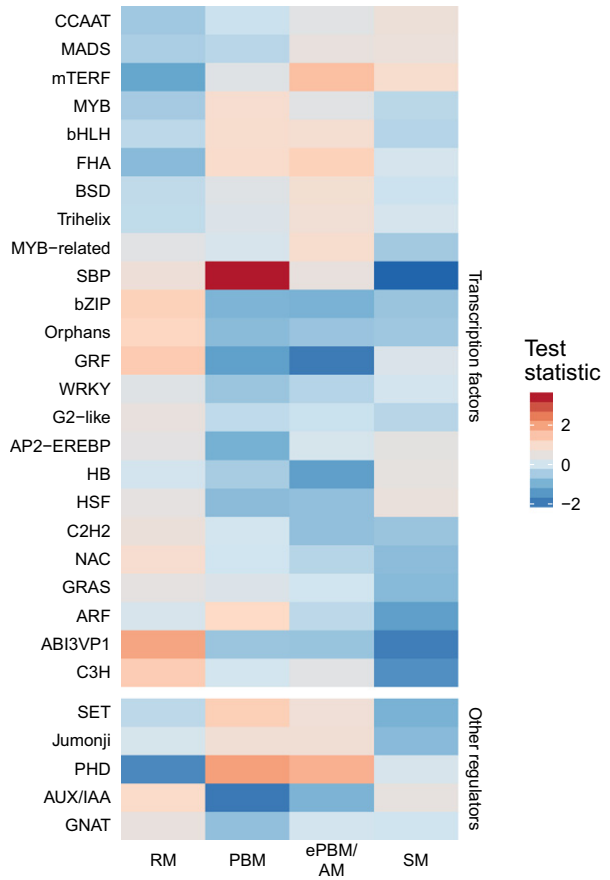

Figure 4. Geneset enrichment analysis (GSEA) of transcription factor and other transcriptional regulator families.
Test statistics were calculated from the $\log_2$-fold change between the stage of interest and all other stages for the expressed genes in each geneset. The absolute value of the test statistic indicates the magnitude of the geneset-level change, and a positive test statistic (red) indicates enrichment and a negative statistic (blue) indicates depletion.

The most pronounced change in expression appears to occur between the RM and branch meristem samples (i.e. the change from apical to axillary meristem). Several families (including ABI3VP1 and GROWTH-REGULATING FACTOR TF genes and Aux/IAA genes) are more highly expressed in RM than in the other meristem types. Another group, including MYB and SBP TF genes and SET and PHD regulator genes, was enriched in one or both of the PBM and ePBM/AM samples but depleted or not enriched in the RM samples.

Homeodomain genes are involved at various stages of plant development and in several hormone response pathways (Chan *et al.*, 1998; Himmelbach *et al.*, 2002; Sawa *et al.*, 2002). The expression of this family in inflorescence meristem types was explored in greater detail using a heatmap of scaled, transformed read counts. Hierarchical clustering of the transformed counts recovered five prominent groups of expression (Figure 5). In general, there was no clear relationship between homeodomain subfamily and expression pattern. However, there was an apparent enrichment of class IV HD-Zip genes with lowest expression in the RM, a peak in indeterminate axillary meristems and then a gradual decrease during axillary meristem transitions.

DISCUSSION

Several previous transcriptomic studies have measured gene expression in whole rice inflorescences at later stages of development than those described here (e.g. Wang *et al.*, 2010; Sato *et al.*, 2011; Sharma *et al.*, 2012; Khanday *et al.*, 2013; Jiang *et al.*, 2014). Another study used microarray experiments at earlier stages to highlight the importance of TFs in panicle development (Furutani *et al.*, 2006). LMD has also been used for precise control of developmental stage during the collection of whole panicle sections for microarrays (Kobayashi *et al.*, 2012). This

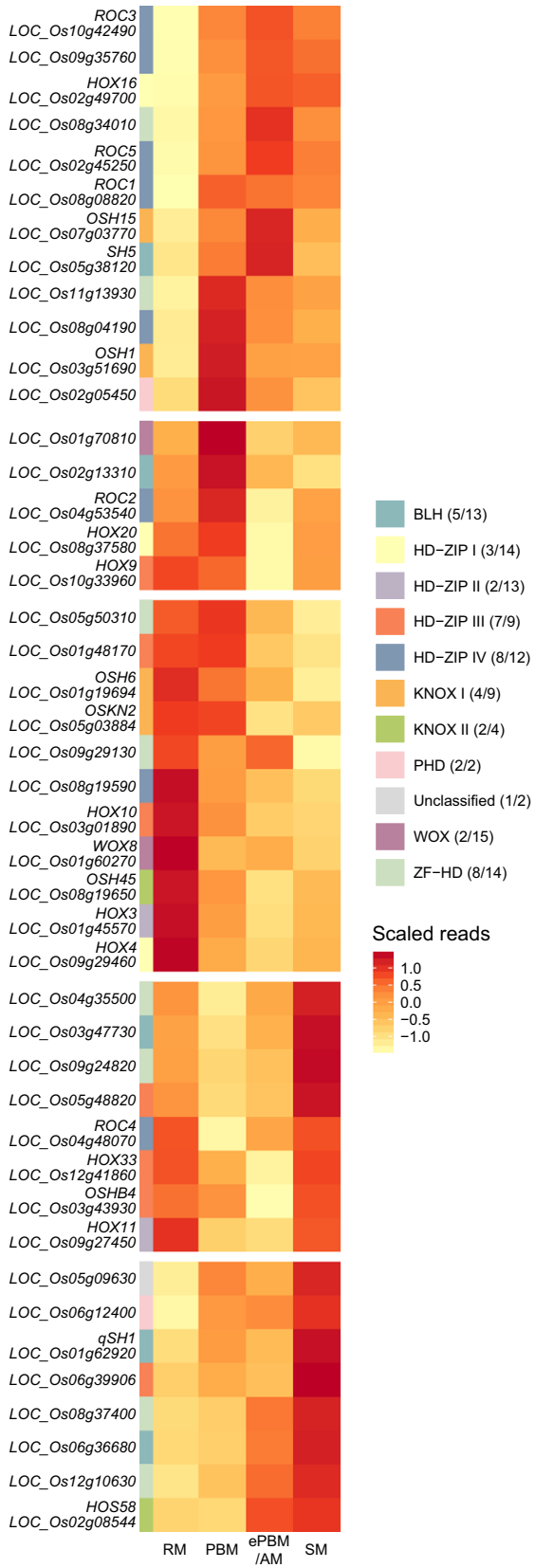

Figure 5. Five prominent patterns of expression of homeobox genes in reproductive meristems.
Scaled, transformed read counts for the homeobox genes that were detected in the dataset are represented by the continuous heat scale. The colour of the *y*-axis denotes the homeobox subfamily for each gene (Jain *et al.*, 2008), and the number of genes detected and the total number of genes for each subfamily is given in brackets after the subfamily name in the legend. The expression patterns were separated into five groups using hierarchical clustering, including a set of homeobox genes with lowest expression in the RM and highest expression in indeterminate axillary meristems, which contains *OSH1* and five of the eight expressed class IV HD-Zip genes.

identified three MADS genes, which are members of the *SQUA*-like (*FUL*-like) clade, that are co-expressed with *PANICLE PHYTOMER2* (*PAP2*/*MADS34*), and a quadruple knockdown of these four genes resulted in defects in inflorescence development (Kobayashi *et al.*, 2012). All four genes were detected in all stages in the RNA-seq dataset presented in this article (Data S1), confirming their expression in reproductive meristems. In contrast with previous transcriptomic datasets, only the meristematic regions of young panicles were collected by LMD for the RNA-seq analysis described here, to restrict the measurement of gene expression to those tissues.

Switch in gene expression between apical and axillary meristems

Clusters 5 and 7 contain genes that change in expression between apical (RM) and axillary meristems (PBM to SM). Axillary meristem initiation requires auxin synthesis and transport in *Arabidopsis thaliana* and maize (reviewed by Gallavotti, 2013), and excess auxin may be inactivated by conjugation to amino acids by GH3 enzymes (Staswick *et al.*, 2005), but these processes are not as well understood in rice. Cluster 7 contains the auxin response factor *ARF6A* and the small auxin-up RNA *SAUR33*, as well as two ABC transporters (*ALS1* and *MDR12*), while the complementary cluster 5 contains four uncharacterized genes with annotations relating to auxin (*IAA10* and *IAA16*, two auxin-responsive transcriptional regulators; *ABCB27*, an ABC transporter; and *LOC_Os01g63770*, an AUX/LAX transporter), and the GH3 gene related to auxin homeostasis, *GH3-8*. Transgenic rice plants presumably expressing *GH3-8* in the endogenous pattern of *FLO-LFY HOMOLOG OF RICE* (*RFL*) (*i.e.* PBM and SBM; Ikeda-Kawakatsu *et al.*, 2012) have smaller panicles with fewer branches (Yadav *et al.*, 2011), and GH3-8 conjugates amino acids to auxin *in vitro* (Ding *et al.*, 2008), suggesting that the panicle phenotype may be related to reduced auxin signalling. Although auxin response factors and Aux/IAA response genes are generally more strongly expressed at earlier stages (Figure 4), the presence of auxin-related genes in complementary clusters suggests that the transcriptional

effects of auxin signalling are dependent on meristem type. Understanding of the relationship between auxin and meristem specification will require functional analysis of the components of the regulatory network involved in the response in these tissues.

Two *STMADS11*-like (*SVP*-like) MADS genes, *MADS22* and *MADS55*, decrease in expression between PBM and SM, whilst the third (*MADS47*) was not detected (Figure 6a). *MADS22* and *MADS55* are activated by the *G1*-Like (G1L/ALOG) transcriptional activator *G1L5* (*TAWAWA1*/*TAW1*), but activation of *MADS47* has not been reported (Yoshida *et al.*, 2013). Overexpression of *TAW1* increases the production of secondary branches and leads to the formation of tertiary branches, resulting in higher grain yield (Yoshida *et al.*, 2013). *TAW1* was recovered from the same cluster as *MADS55*, and the presence and co-expression of genes that are involved in inflorescence development along with other TF genes indicates that the clusters capture biologically meaningful groups of genes, including uncharacterized signalling components. Only three ALOG genes (which are annotated with the Pfam domain PF04852) were detected in the dataset. Two were retrieved from clusters 5 and 4 (*G1L2* and *TAW1* respectively), and the third (*G1L1*) has a similar expression pattern (Figure 6b). *taw1* missense mutants have small inflorescences with reduced grain yield, but this phenotype was stronger when *TAW1* was targeted by RNA interference (Yoshida *et al.*, 2013). This finding may suggest functional redundancy between the three co-expressed ALOG genes, *G1L1*, *G1L2* and *TAW1*, with more than one affected by the RNAi construct used against *TAW1.* Although the increase in secondary branches linked to *TAW1* overexpression suggests that it acts by suppression of SM identity, perhaps via downstream factors that may promote BM activity such as the *STMADS11*-like genes, *taw1* plants also have fewer primary branches (Yoshida *et al.*, 2013), and the peak expression in the RM of *G1L1*, *G1L2*, *TAW1*, *MADS22* and *MADS55* suggests an additional role before the transition from apical to axillary meristem. The difference in expression of annotated TF genes between RM and PBM samples and the recovery of co-expression clusters of primarily uncharacterized genes that appear to be either switched on or off between the RM and the PBM indicate a significant change in gene expression between apical and axillary meristems.

Gradual changes in expression during transitions in axillary meristem identity

Compared with the switch from apical to axillary meristem, more gradual changes in gene expression were apparent during the sequential transition from PBM to SM (Figure 3). This transition corresponds to the acquisition of determinate fate by the axillary meristem. The enrichment of MADS genes in the SM sample and their overrepresentation in cluster 6 is consistent with their involvement in flower development in rice and other plants (Figure 4; e.g. Kyozuka and Shimamoto, 2002; Nagasawa *et al.*, 2003; Yamaguchi *et al.*, 2006; Yao *et al.*, 2008; Dreni *et al.*, 2011; the role of MADS genes in Arabidopsis flower development was reviewed by Prunet and Jack, 2013). Two recent studies in *A. thaliana* and the domesticated tomato, *Solanum lycopersicum*, used an LMD and RNA-seq approach to investigate inflorescence development (Park *et al.*, 2012; Mantegazza *et al.*, 2014). These two datasets were used with the dataset presented in this article to compare the expression of the MIKC^{C}-type class of MADS-box genes (Henschel *et al.*, 2002), suggesting several differences between the species (Figure S4 and Methods S1). *A. thaliana AGL6* is the only gene from the *AGL6*-like clade that is more highly expressed in indeterminate meristems than determinate meristems. *agl6* mutants have fewer axillary buds on the inflorescence stem, suggesting that *AGL6* promotes axillary meristem formation (Huang *et al.*, 2012). The rice genes from the *AGL6*-like clade, *MFO* (*MADS6*) and *MADS17*, redundantly promote floral meristem determinancy (Ohmori *et al.*, 2009), and are correspondingly enriched in determinate meristems. Expression of the putative tomato *AGL6*-like gene *SOLYC01G093960.2* is also enriched in determinate meristems, suggesting that the role of *AGL6* in axillary meristem formation may be specific to the Arabidopsis lineage. In contrast, consistent with their functions in floral organ specification in rice and *A. thaliana*, genes in the *AGL2*-like clade are generally more highly expressed in determinate meristems. *PAP2*, which was strongly expressed at all stages with higher expression in the RM and PBM (Data S1), acts with another *AGAMOUS-LIKE2*-like (*AGL2*-like/*SEPELLATA*) MADS gene, *LEAFY HULL STERILE 1* (*LHS1*/*MADS1*), to specify floral organs (Gao *et al.*, 2010). *pap2* mutants also have altered primary and secondary branch number, suggesting an earlier role is SM specification (Gao *et al.*, 2010; Kobayashi *et al.*, 2010). *LHS1*, which is involved in floral meristem determination and floral organ development (Jeon *et al.*, 2000; Agrawal *et al.*, 2005), was enriched in the SM (Figure S4).

The strong enrichment of SQUAMOSA promoter binding protein-like (SPL/SBP) genes in the PBM and their presence in cluster 2 suggests involvement in the early stages of inflorescence development. *WEALTHY FARMER'S PANICLE* (*WFP*/*SPL14* or *IDEAL PLANT ARCHITECTURE 1*/*IPA1*), *SPL3*, *SPL12* and *SPL17* are all highly expressed in the RM and PBM, gradually decrease in expression and have weakest expression in the SM (Figure 6c), suggesting a redundant function in indeterminate axillary meristems. These genes are all putative targets of *miR156*, a microRNA that causes reduced panicle size and delayed flowering when overexpressed (Xie *et al.*, 2006; Wang *et al.*, 2015), and *SPL* gene expression is controlled by *miR156* and *miR529*

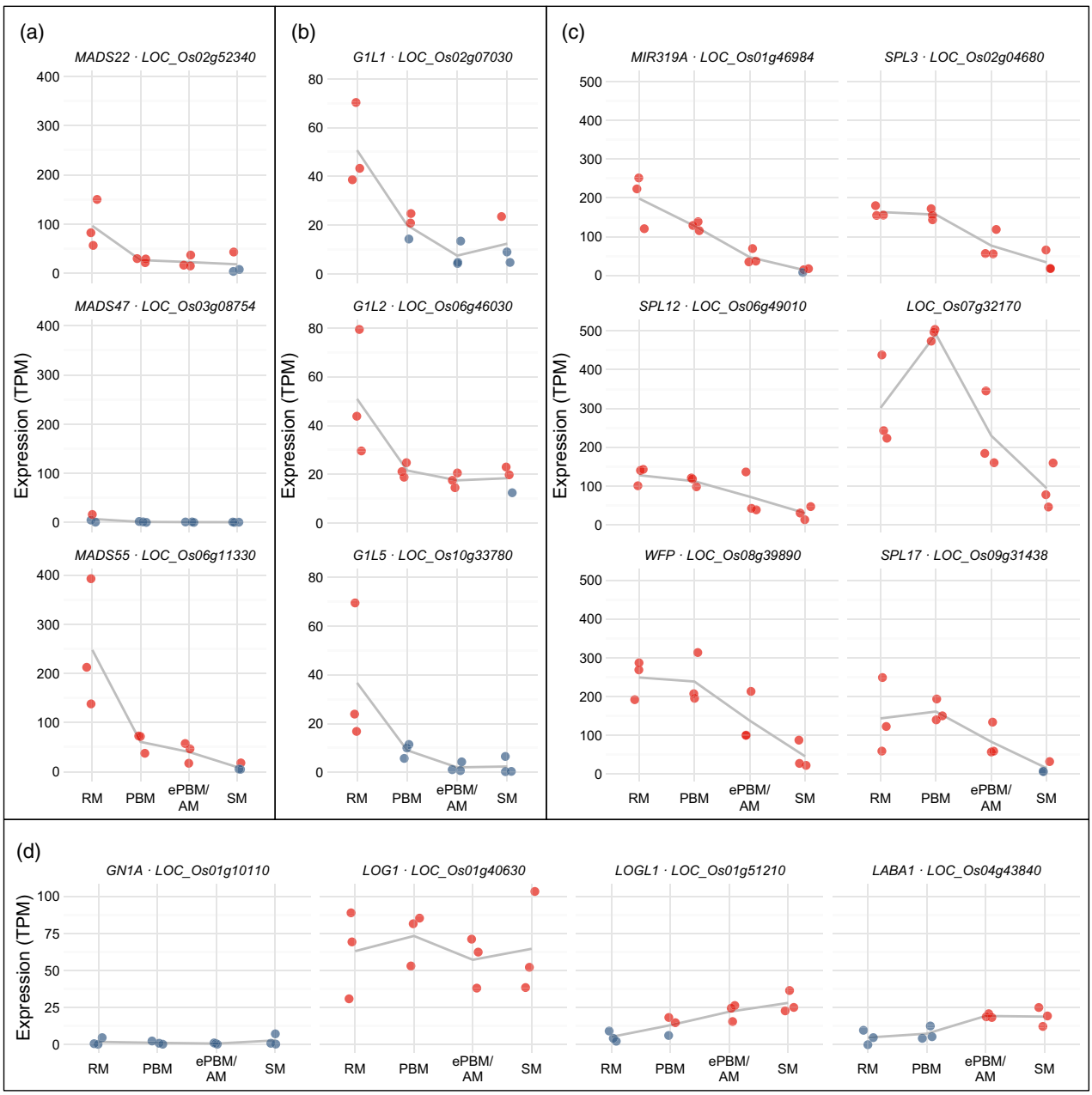

Figure 6. Expression of selected genes in transcripts per million (TPM).
A red point indicates that the gene was above the detection cut-off, and blue indicates that the gene was not detected.
(a) Two reported *STMADS11*-like (*SVP*-like) targets of TAW1, *MADS22* and *MADS55* (Yoshida *et al.*, 2013), are more strongly expressed in the RM than in other meristems, and the other *STMADS11*-like gene, *MADS47*, was only detected in one library.
(b) Only three ALOG genes (*G1L1*, *G1L2* and *G1L5*/*TAW1*) were detected, and they share a similar pattern of expression.
(c) Several *SPL* genes and a co-regulated microRNA precursor, *MIR319A*, which was recovered in cluster 2, are highly expressed in RM and PBM before decreasing in expression in ePBM/AM and SM.
(d) *GN1A* (*CKX2*) was not detected in the LMD dataset, but *LOG1* is expressed in all meristem types, and two other genes possibly related to cytokinin activation, *LABA1* (*LOGL6*) and *LOGL1*, are expressed more highly in ePBM/AM and SM than in RM and PBM.

(Jeong *et al.*, 2011; Wang *et al.*, 2015). The precursor for another microRNA, *MIR319A*, was recovered from the same cluster as *SPL3* and *SPL12* and has a similar expression pattern to these genes, *SPL17* and *WFP* (Figure 6c). In *A. thaliana*, *miR319a* targets *TEOSINTE BRANCHED*/*CYCLOIDEA*/*PROLIFERATING CELL FACTORS* (*TCP*) TFs and is involved in floral organ development (Palatnik *et al.*, 2003; Nag *et al.*, 2009). No direct interaction with *SPL* genes or role in reproductive development has been reported for *miR319a* in rice. Overexpression of *WFP*,

which was detected by *in situ* hybridization experiments in the RM and branch meristems in wild-type plants, is associated with increased branching and grain yield (Jiao *et al.*, 2010; Miura *et al.*, 2010). WFP binds to the promoter region and appears to cause upregulation of the *DENSE AND ERECT PANICLE 1* (*DEP1*) gene (Lu *et al.*, 2013), which is present in cluster 3 and encodes a major QTL for grain production and panicle morphology in high-yield *O. sativa* ssp. *japonica* varieties. Probable loss-of-function *dep1* mutants have larger SAMs with higher cell number, and reduced expression of a cytokinin oxidase gene, *GRAIN NUMBER 1A* (*GN1A*/*CKX2*) (Huang *et al.*, 2009). WFP also binds to the promoters of genes that are involved in the determination of inflorescence architecture via the cytokinin, auxin and gibberellic acid pathways, including *LONELY GUY 1* (*LOG1*), *PIN PROTEIN 1B* and *SLENDER RICE1* respectively, suggesting a role in the coordination of hormone signalling (Lu *et al.*, 2013).

A homologue of *LOG1*, *LONELY GUY LIKE PHOSPHORIBOHYDROLASE 1* (*LOGL1*), is present in cluster 1 with genes that increase in expression between RM and SM. In rice, cytokinins may increase branching complexity by promoting IM or BM activity (reviewed by Han *et al.*, 2014). The LOG1 enzyme converts cytokinin nucleotides to the active form *in vitro*, and *log1* mutants have lower expression of two cytokinin-inducible *RESPONSE REGULATOR* genes, resulting in early termination of BM and IM and small panicles with small SAMs and branching defects (Kurakawa *et al.*, 2007). Similarly, reduced expression of *GN1A*, which encodes an enzyme that inactivates cytokinin, is associated with increased grain production in several high-yield *O. sativa* ssp. *indica* varieties (Ashikari *et al.*, 2005). Although *GN1A* transcripts were not detected in this LMD dataset (Figure 6d), these results suggest that an increase in active cytokinin is associated with grain yield via an effect on meristem activity. However, the function of the *LONG AND BARBED AWN1* (*LABA1*/*LOGL6*/*An-2*) gene provides evidence for an inverse role. The LABA1 protein present in *O. rufipogon*, which shares an identical primary sequence with the protein from *O. sativa* ssp. *japonica* cv. Nipponbare, also activates cytokinin *in vitro* (Gu *et al.*, 2015; Hua *et al.*, 2015). Introduction of the *Laba1* allele from *O. rufipogon* into an *O. sativa* ssp *indica* background, which contains a non-functional *laba1* allele, results in a higher concentration of endogenous cytokinins and higher expression of *RESPONSE REGULATOR* genes (Hua *et al.*, 2015), but reduces the number of grains per panicle and the number of tillers per plant (Gu *et al.*, 2015). *laba1* is a domestication allele found at high frequencies in cultivated accessions, implying that it was affected by artificial selection either for awn phenotype or grain yield (Gu *et al.*, 2015; Hua *et al.*, 2015). *LABA1* was detected in the ePBM/AM and SM samples but not the RM or PBM samples. Although the function of LOGL1 has not been reported, the expression patterns of *LABA1* and *LOGL1* suggest increasing cytokinin activation along the course of axillary meristem determination (Figure 6d). Such a mechanism would support a cytokinin function secondary to the maintenance of IM and BM, but it is not clear how this would influence meristem identity or panicle branching at the molecular level. A type-A *RESPONSE REGULATOR* (*RR3*) is present in cluster 5 (containing genes more highly expressed in RM), highlighting the possible complexity of the roles of cytokinins in inflorescence development.

Homeodomain genes and meristem identity

log1 mutants also appear to have reduced *HOMEOBOX 1* (*OSH1*) expression (Kurakawa *et al.*, 2007). *OSH1* encodes a class I *Knotted1-like homeobox* (*KNOX*) homeodomain TF, and the expression of class I *KNOX* genes in meristems in the developing inflorescence has been detected by *in situ* hybridization (Sentoku *et al.*, 1999). *osh1* mutants have defects in SAM maintenance during vegetative growth and lower expression of other *KNOX* genes, and expression of *KNOX* genes can be induced by cytokinin treatment during shoot regeneration (Tsuda *et al.*, 2011). Another class I *KNOX* gene, *HOMEOBOX 15* (*OSH15*) is expressed in a similar pattern to *OSH1* (Figure 5), as previously reported in *O. sativa* cv. Nipponbare (Sato *et al.*, 1998), and a third, *OSH6*, has a peak in expression in the RM and PBM (Figure 5). OSH1 binds to the upstream region of the *OSH15* locus, and *osh1 osh15* double-mutants have lower induction of *OSH6* after treatment with cytokinin (Tsuda *et al.*, 2011), supporting a role of *KNOX* genes in the response to this phytohormone. The higher expression of *OSH1* and *OSH15* in indeterminate axillary meristems is compatible with a function in the promotion of BM identity. In *A. thaliana*, cytokinin promotes the expression of the homeodomain gene *WUSCHEL*, which itself enhances cytokinin signalling (Leibfried *et al.*, 2005; Lindsay *et al.*, 2006; Gordon *et al.*, 2009). *wox1* plants carrying a likely null mutation in the rice orthologue of *WUSCHEL*, *WUSCHEL-LIKE HOMEOBOX 1* (*WOX1*/*WUS*) (Nardmann and Werr, 2006), have defects in axillary meristem formation and lower expression of *OSH1* (Tanaka *et al.*, 2015). *WOX1* was not detected in the meristem tissues described here, but another *WUSCHEL homeobox* (*WOX*) gene, *WOX8*, has a peak in expression in RM, and a third, *LOC_Os01 g70810*, is predominantly expressed in the RM and PBM (Figure 5). Although these results and previous studies support a role of *KNOX* and *WOX* genes in the maintenance of indeterminate meristem identity, possibly coordinated by phytohormone signalling, more work is required to elucidate this mechanism.

In *A. thaliana*, class III homeodomain-leucine-zipper (HD-Zip) genes are involved in meristem initiation and regulation (Prigge *et al.*, 2005). The class III HD-Zip gene *REVOLUTA* is required for axillary meristem formation (Talbert

et al., 1995; Otsuga *et al.*, 2001). Seven of the nine annotated class III HD-Zip genes in *O. sativa* were detected (Figure 5), including two homologues of *REVOLUTA*, *HOX9* and *HOX10* (Prigge and Clark, 2006), which are both highly expressed in the RM. Three class IV HD-Zip genes, *RICE OUTMOST CELL-SPECIFIC GENE 1* (*ROC1*), *ROC3* and *LOC_Os09g35760*, were recovered in cluster 7, containing genes that had lowest expression in the RM and higher expression in the samples from axillary meristems. Five of the eight detected class IV HD-Zip genes and several genes from other homeobox subclasses, including *OSH1*, also follow this pattern (Figure 5). *Zea mays* class IV HD-Zip genes are expressed in the outer cell layer of the SAM and transgenic overexpression of *OUTER CELL LAYER 1* delays flowering (Depège-Fargeix *et al.*, 2011; Javelle *et al.*, 2011). In rice, several class IV HD-Zip genes have been detected in the vegetative SAM by *in situ* hybridization (Ito *et al.*, 2003), but their role in axillary meristems in the developing inflorescence has not been explained. There are also smaller groups of homeobox genes that peak in expression in the SM or RM samples, suggesting multiple roles and possible redundancy for homeodomain proteins during inflorescence development.

In conclusion, the molecular mechanisms controlling the meristematic activities that guide panicle development in rice remain largely uncharacterized. The dataset presented here addresses this by describing both the transcriptome profiles associated with specific meristem identities and the changes in gene expression that occur during the early stages of panicle development. A marked expression switch was evident between apical and axillary meristems, followed by gradual changes during transitions between PBM and SM. This could be explained by the difference between the fates of the RM, which aborts after PBM differentiation (Ikeda *et al.*, 2004), and the axillary meristems, which undergo transition to determinate SMs. The latter process may be comparable to the gradual meristem maturation observed during tomato inflorescence development (Park *et al.*, 2012), rather than to the continual production of determinate floral meristems characteristic of the Arabidopsis inflorescence meristem. Promising uncharacterized genes were identified, notable examples including TF genes with expression patterns similar to known regulators of panicle architecture, and putative hormone-related genes such as the auxin-responsive genes. Functional studies of these genes will reveal the mechanics of the regulatory networks that establish gene expression patterns in meristems and their control by hormone signalling. Genes with expression specific to certain meristem types were also identified, which may prove valuable as marker genes, for example in developmental studies of mutants affected in determination of reproductive meristem identity. Combined with current technology for the creation of mutants such as CRISPR/Cas9 genome editing, this dataset promises to accelerate research on the molecular mechanisms controlling panicle development, which is a vital for efforts to achieve the sustainable increase in grain yield required to address population growth.

EXPERIMENTAL PROCEDURES

Plant material and sampling

O. sativa ssp. *japonica* cv. Nipponbare plants were grown in a growth chamber with 70% relative humidity at 30°C during the day and 26°C at night. After 7 weeks under long day conditions (14 h light and 10 h darkness) plants were transferred to short day conditions (10 h light and 14 h darkness) to induce floral transition. Inflorescence meristems were harvested from 6 to 14 days after the change of photoperiod.

Morphological analysis

Panicles were harvested in FAA (formaldehyde 10%, acetic acid 5%, ethanol 50%), infiltrated under vacuum for 15 min, stored overnight at 4°C and embedded in paraffin as described by Huijser *et al.* (1992). Tissues were cut into 8-μm sections using an RM2155 microtome (Leica, Wetzlar, Germany), mounted on glass slides and stained with 0.5% w/v toluidine blue for observation with an Axiophot D1 microscope (Zeiss, Oberkochen, Germany) and image capture with an Axiocam MRc 5 camera (Zeiss).

Tissue embedding, laser microdissection and sequencing

Each inflorescence was harvested in 2 ml ice-cold 3:1 ethanol: acetic acid fixative solution, infiltrated twice under a mild vacuum for 15 min and stored for 20 h in fresh fixative at 4°C. Embedding for dissection was performed as described by Mantegazza *et al.* (2014). Tissues were cut into 8-μm sections on an RM2155 or RM2255 microtome (Leica) and dissected on a LMD6000 or LMD7000 laser dissector (Leica). RNA was isolated using the ARCTURUS PicoPure RNA Isolation Kit (ThermoFisher, Waltham, MA, USA) and assayed on a 2100 Bioanalyzer with the RNA 6000 Pico Kit (Agilent, Santa Clara, CA, USA) before amplification and cDNA synthesis with the Ovation RNA-Seq System V2 (NuGEN, San Carlos, CA, USA). Library preparation with the Ovation Ultralow Library Sytem (NuGEN) and 50-base, single-end sequencing on the HiSeq 2000 platform (Illumina, San Diego, CA, USA) were performed by IGA Technology Services (Udine, Italy).

RNA *in situ* hybridization

Samples harvested from the main stem from four developmental stages of panicle development were embedded in Paraplast X-TRA (Sigma-Aldrich, St. Louis, MO, USA) as described by Huijser *et al.* (1992). Digoxigenin-labelled antisense and sense RNA probes were generated with the DIG RNA Labelling Kit SP6/T7 (Roche, Penzberg, Germany) according to the manufacturer's instructions. To generate the probes, cDNA was amplified using the following primers: *LOC_Os01g04670*, 5′-GTGTCAAGGCATCGCCAAC-3′ and 5′-CATCAGCTGGCTGCTTTACC-3′; *LOC_Os10g04270*, 5′-TCCCAGTTGACCGAGAACTG-3′ and 5′-TGAACTTCCGTCACGAACTCC-3′; *LOC_Os10g05990*, 5′-CCTCCGGCAAAGAACTGATG-3′ and 5′-CCGTCATTGGGACTAGTTTGTCAG-3′. For detection of *LOC_Os09g27730*, a digoxigenin-labelled LNA probe with the sequence 5′-TCTG{A}CGACG{T}GCG{A}C{T}GGT-3′ was synthesised (Eurogentec, Liège, Belgium). Hybridization was performed as described by Coen *et al.* (1990) using an NBT/BCIP

(Roche) or VECTOR Blue Alkaline Phosphatase Substrate Kit (Vector Laboratories, Burlingame, CA, USA) for detection.

Data analyses

Analysis of the sequenced libraries was carried out with bash and R scripts (R Core Team, 2015), which were arranged into a pipeline using the ruffus package for python3 (Goodstadt, 2010). All of the code used for the analysis and to generate text, figures and tables for this report, along with the versions and parameters of the software used, are available with revision history in a public GitHub repository at https://github.com/evoreprice/lmdPaper.

Briefly, adaptor trimming was performed with cutadapt (Martin, 2011) and reads were mapped against the MSUv7 genome and annotation (Ouyang *et al.*, 2007; Kawahara *et al.*, 2013) downloaded from Phytozome 10.3 (Goodstein *et al.*, 2012) using STAR (Dobin *et al.*, 2012) in 2-pass mode. The 'quantMode' argument of STAR was used for read counting. rRNA and tRNA contamination was estimated using htseq-count (Anders *et al.*, 2014) to count the number of reads that mapped to regions annotated as tRNA or rRNA on the Rap-DB (Ohyanagi *et al.*, 2006) or to regions where reads simulated with wgsim (https://github.com/lh3/wgsim) from rRNA sequences in the TIGR Plant Repeats Database (Ouyang and Buell, 2004) also mapped. Differential expression analysis to calculate $\log_2$-fold change (L_2FC) values and transformed read counts was performed with the DESeq2 package (Love *et al.*, 2014). To determine a strict cut-off for unexpressed genes, expression values in transcripts per million were calculated for each gene (Li *et al.*, 2010; Wagner *et al.*, 2012). These values were compared with pseudoexpression values calculated from intergenic regions of the genome, using intervals of a similar size distribution to those used for calculating the 'genic' expression values. For each library, the cut-off was placed at the 95th percentile of the distribution of intergenic expression values. Transformed counts and L_2FC values from expressed genes were used for soft clustering with the Mfuzz package (Kumar and Futschik, 2007) and geneset enrichment analysis with the gage package (Luo *et al.*, 2009) respectively. Gene lists were analysed using annotations from several databases, including TIGR (Kawahara *et al.*, 2013), Oryzabase (Kurata and Yamazaki, 2006) and OGRO (Yamamoto *et al.*, 2012). All plots were produced with the ggplot2 package (Wickham, 2009).

Accession numbers

All rice genes are referred to by either their locus identifier (MSUv7; Ouyang *et al.*, 2007) or their official gene symbol provided on Oryzabase (McCouch, 2008).

Data availability

Sequence data from this article have been deposited with the National Centre for Biotechnology Information Sequence Read Archive under accession number SRP067488.

ACKNOWLEDGEMENTS

We thank James Tregear for feedback and discussion, and Jean-Luc Verdeil and Geneviève Conejero for assistance with laser microdissection at the PHIV imaging platform (CIRAD, Montpellier, France). This project was jointly supported by Fondation Agropolis through the *Investissements d'avenir* programme (ANR-10-LABX-0001-01) and Fondazione Cariplo under the EVOREPRICE 1201-004 project. Israr Ud Din's Ph.D. fellowship was funded by IBGE, The University of Agriculture Peshawar, KPK, Pakistan.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Principal components analysis of transformed read counts for each library.

Figure S2. Comparison of the LMD RNA-sequencing dataset with published *in situ* results.

Figure S3. Assessment of clustering results.

Figure S4. Differential expression of genes encoding putative MIKC proteins in *A. thaliana*, *S. lycopersicum* and *O. sativa* ssp. *japonica.*

Table S1 Library and sequencing statistics.

Table S2 References for Figure S2.

Methods S1. Comparative analysis of the MADS family.

Data S1. Expression values for all genes in transcripts per million with annotations.

Data S2. List of clustered genes with annotations.

REFERENCES

Agrawal, G.K., Abe, K., Yamazaki, M., Miyao, A. and Hirochika, H. (2005) Conservation of the E-function for floral organ identity in rice revealed by the analysis of tissue culture-induced loss-of-function mutants of the *OsMADS1* gene. *Plant Mol. Biol.* **59**, 125–135. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16217607&retmode=ref&cmd=prlinks.

Anders, S., Pyl, P.T. and Huber, W. (2014) HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England).* Available at: http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btu638.

Ashikari, M., Sakakibara, H., Lin, S. *et al.* (2005) Cytokinin oxidase regulates rice grain production. *Science*, **309**, 741–745. Available at: http://www.sciencemag.org/content/309/5735/741.full.

Azpeitia, E., Dávila-Velderrain, J., Villarreal, C. and Álvarez-Buylla, E.R. (2013) Gene regulatory network models for floral organ determination. In *Flower Development*. New York, NY: Springer New York, pp. 441–469. Available at: http://link.springer.com/10.1007/978–1-4614–9408–9_26.

Borlaug, N. (2007) Feeding a hungry world. *Science*, **318**, 359–359. Available at: http://www.sciencemag.org/content/318/5849/359.full.

Chan, R.L., Gago, G.M., Palena, C.M. and Gonzalez, D.H. (1998) Homeoboxes in plant development. *Biochim. Biophys. Acta*, **1442**, 1–19. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9767075&retmode=ref&cmd=prlinks.

Coen, E.S., Romero, J.M., Doyle, S., Elliott, R., Murphy, G. and Carpenter, R. (1990) *floricaula:* a homeotic gene required for flower development in Antirrhinum majus. *Cell*, **63**, 1311–1322. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=1702033&retmode=ref&cmd=prlinks.

Depège-Fargeix, N., Javelle, M., Chambrier, P., Frangne, N., Gerentes, D., Perez, P., Rogowsky, P.M. and Vernoud, V. (2011) Functional characterization of the HD-ZIP IV transcription factor OCL1 from maize. *J. Exp. Bot.* **62**, 293–305. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20819789&retmode=ref&cmd=prlinks.

Ding, X., Cao, Y., Huang, L., Zhao, J., Xu, C., Li, X. and Wang, S. (2008) Activation of the indole-3-acetic acid-amido synthetase GH3–8 suppresses expansin expression and promotes salicylate- and jasmonate-independent basal immunity in rice. *Plant Cell*, **20**, 228–240. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=18192436&retmode=ref&cmd=prlinks.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2012) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, **29**, 15–21. Available at: http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/bts635.

Dreni, L., Pilatone, A., Yun, D., Erreni, S., Pajoro, A., Caporali, E., Zhang, D. and Kater, M.M. (2011) Functional analysis of all AGAMOUS subfamily

members in rice reveals their roles in reproductive organ identity determination and meristem determinacy. *Plant Cell*, **23**, 2850–2863. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=21810995&retmode=ref&cmd=prlinks.

Furutani, I., Sukegawa, S. and Kyozuka, J. (2006) Genome-wide analysis of spatial and temporal gene expression in rice panicle development. *Plant J.* **46**, 503–511. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16623909&retmode=ref&cmd=prlinks.

Gallavotti, A. (2013) The role of auxin in shaping shoot architecture. *J. Exp. Bot.* **64**, 2593–2608. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=23709672&retmode=ref&cmd=prlinks.

Gao, X., Liang, W., Yin, C. *et al.* (2010) The *SEPALLATA*-like gene *OsMADS34* is required for rice inflorescence and spikelet development. *Plant Physiol.* **153**, 728–740. Available at: http://www.plantphysiol.org/content/153/2/728.full.

Goodstadt, L. (2010) Ruffus: a lightweight Python library for computational pipelines. *Bioinformatics (Oxford, England)*, **26**, 2778–2779. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20847218&retmode=ref&cmd=prlinks.

Goodstein, D.M., Shu, S., Howson, R. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22110026&retmode=ref&cmd=prlinks.

Gordon, S.P., Chickarmane, V.S., Ohno, C. and Meyerowitz, E.M. (2009) Multiple feedback loops through cytokinin signaling control stem cell number within the *Arabidopsis* shoot meristem. *Proc. Natl Acad. Sci. USA*, **106**, 16529–16534. Available at: http://www.pnas.org/content/106/38/16529.full.

Gu, B., Zhou, T., Luo, J. *et al.* (2015) *An-2* encodes a cytokinin synthesis enzyme that regulates awn length and grain production in rice. *Mol. Plant*, **8**, 1635–1650. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1674205215003238.

Han, Y., Yang, H. and Jiao, Y. (2014) Regulation of inflorescence architecture by cytokinins. *Front. Plant Sci.* **5**, 669. Available at: http://journal.frontiersin.org/article/10.3389/fpls.2014.00669/abstract.

He, Z., Zhai, W., Wen, H. *et al.* (2011) Two evolutionary histories in the genome of rice: the roles of domestication genes. *PLoS Genet.* **7**, e1002100. Available at: http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1002100.

Henschel, K., Kofuji, R., Hasebe, M., Saedler, H., Münster, T. and Theißen, G. (2002) Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patens*. *Mol. Biol. Evol.* **19**, 801–814. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=12032236&retmode=ref&cmd=prlinks.

Himmelbach, A., Hoffmann, T., Leube, M., Höhener, B. and Grill, E. (2002) Homeodomain protein ATHB6 is a target of the protein phosphatase ABI1 and regulates hormone responses in *Arabidopsis*. *EMBO J.* **21**, 3029–3038. Available at: http://emboj.embopress.org/cgi/doi/10.1093/emboj/cdf316.

Hua, L., Wang, D.R., Tan, L. *et al.* (2015) *LABA1,* a domestication gene associated with long, barbed awns in wild rice. *Plant Cell*, **27**, 1875–1888. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=26082172&retmode=ref&cmd=prlinks.

Huang, X., Qian, Q., Liu, Z. *et al.* (2009) Natural variation at the DEP1 locus enhances grain yield in rice. *Nat. Genet.* **41**, 494–497. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=19305410&retmode=ref&cmd=prlinks.

Huang, X., Effgen, S., Meyer, R.C., Theres, K. and Koornneef, M. (2012) Epistatic natural allelic variation reveals a function of AGAMOUS-LIKE6 in axillary bud formation in *Arabidopsis*. *Plant Cell*, **24**, 2364–2379. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22730404&retmode=ref&cmd=prlinks.

Huijser, P., Klein, J., Lönnig, W.E., Meijer, H., Saedler, H. and Sommer, H. (1992) Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene *squamosa* in *Antirrhinum majus*. *EMBO J.* **11**, 1239–1249. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=1563342&retmode=ref&cmd=prlinks.

Ikeda, K., Sunohara, H. and Nagato, Y. (2004) Developmental course of inflorescence and spikelet in rice. *Breed. Sci.* **54**, 147–156. Available at: http://doi.org/10.1270/jsbbs.54.147.

Ikeda, M., Miura, K., Aya, K., Kitano, H. and Matsuoka, M. (2013) Genes offering the potential for designing yield-related traits in rice. *Curr. Opin. Plant Biol.* **16**, 213–220. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=23466256&retmode=ref&cmd=prlinks.

Ikeda-Kawakatsu, K., Maekawa, M., Izawa, T., Itoh, J.-I. and Nagato, Y. (2012) *ABERRANT PANICLE ORGANIZATION 2/RFL,* the rice ortholog of Arabidopsis *LEAFY,* suppresses the transition from inflorescence meristem to floral meristem through interaction with *APO1*. *Plant J.* **69**, 168–180. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1365–313X.2011.04781.x/full.

Ito, M., Sentoku, N., Nishimura, A., Hong, S.K. and Sato, Y. (2003) Roles of rice *GL2*-type homeobox genes in epidermis differentiation. *Breed. Sci.* **53**, 245–253. Available at: http://doi.org/10.1270/jsbbs.53.245.

Jain, M., Tyagi, A.K. and Khurana, J.P. (2008) Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. *FEBS J.* **275**, 2845–2861. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=18430022&retmode=ref&cmd=prlinks.

Javelle, M., Klein-Cosson, C., Vernoud, V., Boltz, V., Maher, C., Timmermans, M., Depège-Fargeix, N. and Rogowsky, P.M. (2011) Genome-wide characterization of the HD-ZIP IV transcription factor family in maize: preferential expression in the epidermis. *Plant Physiol.* **157**, 790–803. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=21825105&retmode=ref&cmd=prlinks.

Jeon, J.S., Jang, S., Lee, S. *et al.* (2000) *leafy hull sterile1* is a homeotic mutation in a rice MADS box gene affecting rice flower development. *Plant Cell*, **12**, 871–884. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=10852934&retmode=ref&cmd=prlinks.

Jeong, D.-H., Park, S., Zhai, J., Gurazada, S.G.R., De Paoli, E., Meyers, B.C. and Green, P.J. (2011) Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant Cell*, **23**, 4185–4207. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22158467&retmode=ref&cmd=prlinks.

Jiang, G., Xiang, Y., Zhao, J., Yin, D., Zhao, X., Zhu, L. and Zhai, W. (2014) Regulation of inflorescence branch development in rice through a novel pathway involving the pentatricopeptide repeat protein sped1-D. *Genetics*, **197**, 1395–1407. Available at: http://www.genetics.org/cgi/doi/10.1534/genetics.114.163931.

Jiao, Y., Wang, Y., Xue, D. *et al.* (2010) Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* **42**, 541–544. Available at: http://www.nature.com/doifinder/10.1038/ng.591.

Kawahara, Y., de la Bastide, M., Hamilton, J.P. *et al.* (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice*, **6**, 1–10. Available at: http://link.springer.com/article/10.1186/1939–8433–6-4.

Khanday, I., Yadav, S.R. and 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. Available at: http://www.plantphysiol.org/content/161/4/1970.full.

Kobayashi, K., Maekawa, M., Miyao, A., Hirochika, H. and 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. Available at: http://pcp.oxfordjournals.org/cgi/doi/10.1093/pcp/pcp166.

Kobayashi, K., Yasuno, N., Sato, Y., Yoda, M., Yamazaki, R., Kimizu, M., Yoshida, H., Nagamura, Y. and Kyozuka, J. (2012) Inflorescence meristem identity in rice is specified by overlapping functions of three *AP1/FUL*-like MADS box genes and *PAP2,* a *SEPALLATA* MADS box gene. *Plant Cell*, **24**, 1848–1859. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22570445&retmode=ref&cmd=prlinks.

Kumar, L. and Futschik, M.E. (2007) Mfuzz: a software package for soft clustering of microarray data. *Bioinformation*, **2**, 5–7. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=18084642&retmode=ref&cmd=prlinks.

Kurakawa, T., Ueda, N., Maekawa, M., Kobayashi, K., Kojima, M., Nagato, Y., Sakakibara, H. and Kyozuka, J. (2007) Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature*, **445**, 652–655.

Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=17287810&retmode=ref&cmd=prlinks.

Kurata, N. and Yamazaki, Y. (2006) Oryzabase. An integrated biological and genome information database for rice. *Plant Physiol.* **140**, 12–17. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16403737&retmode=ref&cmd=prlinks.

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 Physiol.* **43**, 130–135. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=11828031&retmode=ref&cmd=prlinks.

Leibfried, A., To, J.P.C., Busch, W., Stehling, S., Kehle, A., Demar, M., Kieber, J.J. and Lohmann, J.U. (2005) WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature*, **438**, 1172–1175. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16372013&retmode=ref&cmd=prlinks.

Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A. and Dewey, C.N. (2010) RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics (Oxford, England)*, **26**, 493–500. Available at: http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btp692.

Lindsay, D.L., Sawhney, V.K. and Bonham-Smith, P.C. (2006) Cytokinin-induced changes in *CLAVATA1* and *WUSCHEL* expression temporally coincide with altered floral development in *Arabidopsis*. *Plant Sci.*. Available at: http://www.sciencedirect.com/science/article/pii/S016894520600032X.

Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=25516281&retmode=ref&cmd=prlinks.

Lu, Z., Yu, H., Xiong, G. *et al.* (2013) Genome-wide binding analysis of the transcription activator ideal plant architecture1 reveals a complex network regulating rice plant architecture. *Plant Cell*, **25**, 3743–3759. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=24170127&retmode=ref&cmd=prlinks.

Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D. and Woolf, P.J. (2009) GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*, **10**, 161. Available at: http://www.biomedcentral.com/1471–2105/10/161.

Mantegazza, O., Gregis, V., Chiara, M., Selva, C., Leo, G., Horner, D.S. and Kater, M.M. (2014) Gene coexpression patterns during early development of the native Arabidopsis reproductive meristem: novel candidate developmental regulators and patterns of functional redundancy. *Plant J.* **79**, 861–877. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=24923650&retmode=ref&cmd=prlinks.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12. Available at: http://journal.embnet.org/index.php/embnetjournal/article/view/200/479.

McCouch, S.R. (2008) Gene nomenclature system for rice. *Rice*. Available at: http://link.springer.com/article/10.1007/s12284–008–9004–9.

Miura, K., Ikeda, M., Matsubara, A., Song, X.-J., Ito, M., Asano, K., Matsuoka, M., Kitano, H. and Ashikari, M. (2010) *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* **42**, 545–549. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20495564&retmode=ref&cmd=prlinks.

Nag, A., King, S. and Jack, T. (2009) miR319a targeting of *TCP4* is critical for petal growth and development in *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **106**, 22534–22539. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0908718106.

Nagasawa, N., Miyoshi, M., Sano, Y., Satoh, H., Hirano, H., Sakai, H. and Nagato, Y. (2003) *SUPERWOMAN1* and *DROOPING LEAF* genes control floral organ identity in rice. *Development*, **130**, 705–718. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=12506001&retmode=ref&cmd=prlinks.

Nardmann, J. and Werr, W. (2006) The shoot stem cell niche in angiosperms: expression patterns of *WUS* orthologues in rice and maize imply major modifications in the course of mono- and dicot evolution. *Mol. Biol. Evol.* **23**, 2492–2504. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16987950&retmode=ref&cmd=prlinks.

Ohmori, S., Kimizu, M., Sugita, M., Miyao, A., Hirochika, H., Uchida, E., Nagato, Y. and Yoshida, H. (2009) *MOSAIC FLORAL ORGANS1,* an *AGL6*-like MADS box gene, regulates floral organ identity and meristem fate in rice. *Plant Cell*, **21**, 3008–3025. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=19820190&retmode=ref&cmd=prlinks.

Ohyanagi, H., Tanaka, T., Sakai, H. *et al.* (2006) The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic Acids Res.* **34**, D741–D744. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkj094.

Otsuga, D., DeGuzman, B., Prigge, M.J., Drews, G.N. and Clark, S.E. (2001) *REVOLUTA* regulates meristem initiation at lateral positions. *Plant J.* **25**, 223–236. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=11169198&retmode=ref&cmd=prlinks.

Ouyang, S. and Buell, C.R. (2004) The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res.* **32**, D360–D363. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkh099.

Ouyang, S., Zhu, W., Hamilton, J. *et al.* (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res.* **35**, D883–D887. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=17145706&retmode=ref&cmd=prlinks.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C. and Weigel, D. (2003) Control of leaf morphogenesis by microRNAs. *Nature*, **425**, 257–263. Available at: http://www.nature.com/doifinder/10.1038/nature01958.

Park, S.J., Jiang, K., Schatz, M.C. and Lippman, Z.B. (2012) Rate of meristem maturation determines inflorescence architecture in tomato. *Proc. Natl Acad. Sci. USA*, **109**, 639–644. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1114963109.

Peng, S., Khush, G.S., Virk, P., Tang, Q. and Zou, Y. (2008) Progress in ideotype breeding to increase rice yield potential. *Field. Crop. Res.* **108**, 32–38. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0378429008000749.

Pérez-Rodríguez, P., Riaño-Pachón, D.M., Correa, L.G.G., Rensing, S.A., Kersten, B. and Mueller-Roeber, B. (2010) PlnTFDB: updated content and new features of the Plant Transcription Factor Database. *Nucleic Acids Res.* **38**, D822–D827. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=19858103&retmode=ref&cmd=prlinks.

Prigge, M.J. and Clark, S.E. (2006) Evolution of the class III HD-Zip gene family in land plants. *Evol. Dev.* **8**, 350–361. Available at: http://doi.wiley.com/10.1111/j.1525–142X.2006.00107.x.

Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N. and Clark, S.E. (2005) Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. *Plant Cell*, **17**, 61–76. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15598805&retmode=ref&cmd=prlinks.

Prunet, N. and Jack, T.P. (2013) Flower development in *Arabidopsis:* there is more to it than learning your ABCs. In *Flower Development*. New York, NY: Springer New York, pp. 3–33. Available at: http://link.springer.com/10.1007/978–1-4614–9408–9_1.

R Core Team. (2015) *R: A Language and Environment for Statistical Computing*. Available at: http://www.R-project.org/.

Sato, Y., Sentoku, N., Nagato, Y. and Matsuoka, M. (1998) Isolation and characterization of a rice homeobox gene, *OSH15*. *Plant Mol. Biol.* **38**, 983–998. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9869405&retmode=ref&cmd=prlinks.

Sato, Y., Antonio, B., Namiki, N. *et al.* (2011) Field transcriptome revealed critical developmental and physiological transitions involved in the expression of growth potential in *japonica* rice. *BMC Plant Biol.* **11**, 10. Available at: http://www.biomedcentral.com/1471–2229/11/10.

Sawa, S., Ohgishi, M., Goda, H., Higuchi, K., Shimada, Y., Yoshida, S. and Koshiba, T. (2002) The *HAT2* gene, a member of the HD-Zip gene family, isolated as an auxin inducible gene by DNA microarray screening, affects auxin response in *Arabidopsis*. *Plant J.* **32**, 1011–1022. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=12492842&retmode=ref&cmd=prlinks.

Sentoku, N., Sato, Y., Kurata, N., Ito, Y., Kitano, H. and Matsuoka, M. (1999) Regional expression of the rice *KN1*-type homeobox gene family during embryo, shoot, and flower development. *Plant Cell*, **11**, 1651–1664. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=10488233&retmode=ref&cmd=prlinks.

Sharma, R., Agarwal, P., Ray, S. *et al.* (2012) Expression dynamics of metabolic and regulatory components across stages of panicle and seed development in *indica* rice. *Funct. Integr. Genomics*, **12**, 229–248. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22466020&retmode=ref&cmd=prlinks.

Staswick, P.E., Serban, B., Rowe, M., Tiryaki, I., Maldonado, M.T., Maldonado, M.C. and Suza, W. (2005) Characterization of an Arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. *Plant Cell*, **17**, 616–627. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15659623&retmode=ref&cmd=prlinks.

Talbert, P.B., Adler, H.T., Parks, D.W. and Comai, L. (1995) The *REVOLUTA* gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of *Arabidopsis thaliana*. *Development*, **121**, 2723–2735. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=7555701&retmode=ref&cmd=prlinks.

Tanaka, W., Ohmori, Y., Ushijima, T., Matsusaka, H., Matsushita, T., Kumamaru, T., Kawano, S. and Hirano, H.-Y. (2015) Axillary meristem formation in rice requires the *WUSCHEL* Ortholog *TILLERS ABSENT1*. *Plant Cell*, **27**, 1173–1184. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=25841039&retmode=ref&cmd=prlinks.

Tsuda, K., Ito, Y., Sato, Y. and Kurata, N. (2011) Positive autoregulation of a *KNOX* gene is essential for shoot apical meristem maintenance in rice. *Plant Cell*, **23**, 4368–4381. Available at: http://www.plantcell.org/content/23/12/4368.full.

Wagner, G.P., Kin, K. and Lynch, V.J. (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22872506&retmode=ref&cmd=prlinks.

Wang, Y. and Li, J. (2011) Branching in rice. *Curr. Opin. Plant Biol.* **14**, 94–99. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=21144796&retmode=ref&cmd=prlinks.

Wang, L., Xie, W., Chen, Y. *et al.* (2010) A dynamic gene expression atlas covering the entire life cycle of rice. *Plant J.* **61**, 752–766. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20003165&retmode=ref&cmd=prlinks.

Wang, L., Sun, S., Jin, J. *et al.* (2015) Coordinated regulation of vegetative and reproductive branching in rice. *Proc. Natl Acad. Sci. USA*. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=26631749&retmode=ref&cmd=prlinks.

Wickham, H. (2009) *ggplot2: Elegant Graphics for Data Analysis*. Springer New York. Available at: http://had.co.nz/ggplot2/book.

Xie, K., Wu, C. and Xiong, L. (2006) Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiol.* **142**, 280–293. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16861571&retmode=ref&cmd=prlinks.

Xing, Y. and Zhang, Q. (2010) Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol.* **61**, 421–442. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20192739&retmode=ref&cmd=prlinks.

Xu, X., Liu, X., Ge, S. *et al.* (2012) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* **30**, 105–111. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22158310&retmode=ref&cmd=prlinks.

Yadav, S.R., Khanday, I., Majhi, B.B., Veluthambi, K. and Vijayraghavan, U. (2011) Auxin-responsive *OsMGH3,* a common downstream target of *OsMADS1* and *OsMADS6,* controls rice floret fertility. *Plant Cell Physiol.* **52**, 2123–2135. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22016342&retmode=ref&cmd=prlinks.

Yamaguchi, T., Lee, D.-Y., Miyao, A., Hirochika, H., An, G. and Hirano, H.-Y. (2006) Functional diversification of the two C-class MADS box genes *OSMADS3* and *OSMADS58* in *Oryza sativa*. *Plant Cell*, **18**, 15–28. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16326928&retmode=ref&cmd=prlinks.

Yamamoto, E., Yonemaru, J., Yamamoto, T. and Yano, M. (2012) OGRO: the overview of functionally characterized genes in rice online database. *Rice*. Available at: http://link.springer.com/article/10.1186/1939–8433–5-26.

Yao, S.-G., Ohmori, S., Kimizu, M. and Yoshida, H. (2008) Unequal genetic redundancy of rice *PISTILLATA* orthologs, *OsMADS2* and *OsMADS4,* in lodicule and stamen development. *Plant Cell Physiol.* **49**, 853–857. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=18378529&retmode=ref&cmd=prlinks.

Yoshida, H. and Nagato, Y. (2011) Flower development in rice. *J. Exp. Bot.* **62**, 4719–4730. Available at: http://jxb.oxfordjournals.org/content/62/14/4719.full.

Yoshida, A., Sasao, M., Yasuno, N. *et al.* (2013) *TAWAWA1,* a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proc. Natl Acad. Sci. USA*, **110**, 767–772. Available at: http://www.pnas.org/content/110/2/767.full.

PART-III

3. FUNCTIONAL CHARACTERISATION OF ALOG FAMILY GENES, *G1L1* AND *G1L2* IN RICE (*Oryza sativa*)

3.1. SUMMARY

Meristems are groups of pluripotent cells that contribute to plant growth and the development of its organs. Based on meristems fate, the rice inflorescence or panicle development has two main stages; the indeterminate meristems stage in which the length and number of branches of the inflorescence are established by the activities of the inflorescence meristem (IM), branch meristems (BMs), and the determinate spikelet meristems (SMs). The timing of determinate SMs specification plays an important role in determining the inflorescence architecture and is associated with final grain yield. A good example is *TAWAWA1* (*TAW1*) of rice, a member of the plant specific ALOG family, which encodes an uncharacterized nuclear protein containing a DUF640 domain. The dominant *taw1-D* mutant delayed the specification of SMs identity resulting in more panicle branches. The important role of *ALOG* genes related to meristem function has been reported in other species including Arabidopsis and Tomato. Very little is known about other *ALOG* genes in rice, therefore we investigated the expression dynamics of ALOG family members in detail and our results show that two other genes, *G1L1* and *G1L2*, showed a similar expression pattern to *G1L5*/*TAW1*. We used the CRISPR/Cas9 genome editing tool to generate single and double knock out mutant lines for *G1L1* and *G1L2* to study their functional role in rice panicle development. Preliminary results suggest a role for *G1L1* and *G1L2* in controlling the branch numbers and final grain yield of the panicle.

3.2. INTRODUCTION

Environmental cues and endogenous flowering pathways induce the switch from vegetative to reproductive growth leading to the conversion of the vegetative Shoot Apical Meristem (SAM) into the Inflorescence Meristem (IM) (Kobayashi & Weigel, 2007), which activity is important for inflorescence development. Crop productivity is largely dependent on inflorescence architecture. Variation in inflorescence architecture (length of the main axis and branches, number of branches and number of flowers) among species and cultivars determines the final yield (MacAlister et al., 2012; Weberling, 1989).

Rice has a complex determinate inflorescence known as panicle; composed of a main axis called rachis with 8-10 nodes each 2-4cm apart, which produce primary, secondary and rarely tertiary branches (Hoshikawa, 1989). These branches bear spikelets, which ultimately develop florets, which are the rice flowers (Ikeda et al., 2004). The architecture of the rice inflorescence is determined by the activities of meristems. The fate of a meristem is specified after a series of transitions, starting from shoot apical meristem (SAM) to rachis meristem (RM) or IM that produces BMs, which convert to determinate SMs that each finally form a single fully fertile flower or floret (Itoh et al., 2005).

Several genes in rice have been described to understand the molecular mechanism governing meristem phase transition especially from BMs to SMs, which is considered a critical step in determining the architecture of the inflorescence. A conserved role of rice *TERMINAL FLOWER*/*CENTRORADIALIS* (*RCN*) genes, homologs of *Arabidopsis TFL1/CEN*, has been proposed, which repress floral/spikelet meristem identity specification in both distant species and its overexpression results in a delay in flowering time and in highly branched inflorescences (Nakagawa et al., 2002). However, several studies on *UNUSUAL FLORAL ORGANS (UFO)* and *LEAFY (LFY)*, and its rice orthologs *ABERRANT PANICLE ORGANIZATION1 (APO1)* and *APO2,* respectively, showed a functional divergence in monocot and eudicot species with respect to inflorescence meristem fate. Loss of function mutations in both *APO1* and *APO2* resulted in precocious SMs identity specification and small inflorescences with fewer branches (Wilkinson & Haughn, 1995; Ikeda et al., 2005, 2009, 2012).

The SQUAMOSA PROMOTER BINDING PROTEIN, a transcription factor that binds to the promoter of *SQUAMOSA*, which is a MADS-box gene, was first reported in *Antirrhinum majus* and emerged as an important regulator of a wide range of biological processes. *Oryza sativa*

SQUAMOSA PROMOTER BINDING PROTEIN-LIKE14 (OsSPL14) is encoded by a QTL known as the *WFP (WEALTHY FARMER PANICLE)*, that regulates panicle branching, and tiller number in rice. The expression of *OsSPL14* is post-transcriptionally regulated by *OsmiR156* and a single point mutation in the *OsmiR156* target site resulted in the production of fewer tillers and more panicle branches, a highly desired inflorescence phenotype by breeders (Jiao et al., 2010, Miura et al., 2010). However, the molecular mechanism by which *OsSPL14* regulates panicle branching is still not understood and study of its downstream targets may provide a clue to understand this regulatory mechanism. *PANICLE PHYTOMER2 (PAP2)* encodes a MADS-box protein named OsMADS34, which belongs to a grass specific subclade of *SEPALLATA (SEP)* subfamily proteins and it controls meristem phase transition from BMs to SMs (Gaoet al., 2010;Kobayashi et al., 2010). Quadruple knockdown of the three AP1-like genes (*OsMADS14, OsMADS15* and *OsMADS18)* and *PAP2* block the transition from vegetative to reproductive phase and IM identity failed to establish which indicate that these four genes act redundantly to determine IM identity (Kobayashi et al., 2012). The function of *TERMINAL FLOWER* 1(*TFL1*)*/CENTRORADIALIS* (*CEN*) homologues in rice *RCN1* and *RCN2* have been investigated to understand if there is a similar mechanism underlies inflorescence development. In plants overexpressing *RCN1* and *RCN2* transcripts, were delayed in phase transition and plants produced more inflorescence branches (Nakagawa et al., 2002). Other MADS-box genes *SOC1*, *AGL24*, *SVP*, and *SEP4* homologs in rice, *OsMADS50*, *OsMADS56*, *OsMADS22*, *OsMADS47*, and *OsMADS55* respectively, control inflorescence branching by suppressing *RCN* genes in lateral meristems (Liu et al., 2013), however the putative genetic interactions between these genes and other inflorescence determinants are still not elucidated.

ALOG (*Arabidopsis* LSH1 & *Oryza* G1) (Zhao et al., 2004; Yoshida et al., 2009), is a family of eukaryotic transcription factors containing a single DUF640 domain that is considered to be derived from DNA Binding Domains (DBDs) of mobile and selfish elements (Lyer & Aravind, 2012), however its DNA-binding specificity and mechanism of action is still not clear. Only few members of this family have been studied, the first one identified was Arabidopsis *LSH1*, which showed hypersensitivity to continuous light and regulates hypocotyl length (Zhao et al., 2004). Furthermore, in Arabidopsis, two direct transcriptional targets of *CUP SHAPED COTYLEDON1 (CUC1)*, including *LSH3* and *LSH4,* which are specifically expressed in the boundary region between the SAM and lateral organs. Constitutive expression of *LSH3* and *LSH4* resulted in

ectopic meristem formation, which suggests that they are putatively involved in regulation of meristem function by modulating meristem cells proliferation and differentiation (Cho and Zambryski, 2011; Takeda et al., 2011), however it will be interesting to study the loss-of-function phenotypes of these LSH genes. In a dominant gain of function mutant of *TAW1*, an enhanced branching of the inflorescence was observed with a higher spikelet number due to a delayed switch from BMs to SMs identity. Terminating Flower (*TMF*), the closest homolog of *TAW1* in tomato, has been shown to have a similar function. In the *tmf* loss-of-function mutant, the primary SAMs instead of producing a sympodial inflorescence terminates in a single flower (MacAlister et al., 2012). Another *TAW1* homolog in maize was identified by transcriptome analysis of SAMs during embryogenesis, named *ALOG1,* and was shown to be expressed in the boundaries between SAM and the lateral meristems (Takacs et al., 2012).

3.3. MATERIALS AND METHODS

3.3.1. PLANT MATERIALS AND GROWTH CONDITIONS

Oryza sativa L. ssp. japonica cv. Nipponbare plants were grown in a growth chamber with 70% relative humidity at 30°C during the day and 26°C at night. After 8 weeks under long day conditions (14 h light and 10 h darkness) plants were transferred to short day conditions (10h light and 14h darkness) to induce floral transition. For mRNA expression analysis reproductive meristems were collected from young inflorescence harvested from 6 to 14 days after the change of photoperiod. Transgenic plants were grown in the greenhouse under short day conditions (10h light and 14h darkness) at 28°C day temperature and 24°C night temperature with 80% relative humidity.

3.3.2. Vector Construction and Rice Transformation

For the generation of single knock-out mutants, for each gene *G1L1*, *G1L2*, 20-bp target specific protospacers (Table 3.1) were selected using the CRISPR-P database (*http://cbi.hzau.edu.cn/crispr/)* and cloned into the BsaI site of pOs-sgRNA entry vectors shown in Figure 3.1 (B) under U3 promoter and then combined into the destination vector containing the Cas9 under maize Ubiquitin promoter Figure 3.1 (A) using the Gateway LR Clonase II Enzyme mix following the procedure reported by Miao et al. (2013). For double (*g1l1*, *g1l2*) knockout lines, protospacers were designed containing BsaI sites according to Xie et al. (2015), amplified with PCR using the pGTR plasmid as template and then ligated with Golden Gate (GG) and assembled with PCR. To make the Polycistronic tRNA-gRNA (PTGs) the PCR reaction was purified and digested with FokI and cloned into the destination vector pRGEB32 shown inFigure1 (C) containing the Cas9 under the control of the UBI promoter. The destination vectors were introduced into *Agrobacterium tumefaciens* strain EH105 and embryogenic calli from *Oryza sativa* L. ssp. *Japonica* cv. Nipponbare seeds were transformed according to the methods described by Hiei et al. (1994) and Toki (1997).

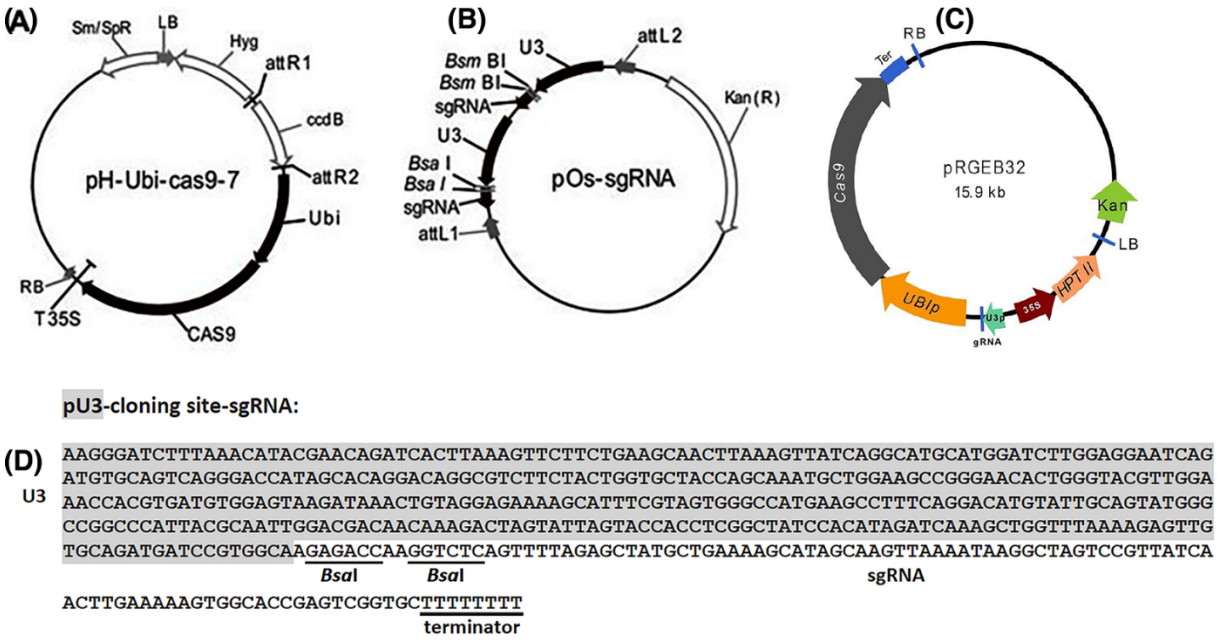

Figure 3.1. Schematic representation of the vectors. Destination vector containing cas9, Miao et al. (2013). (A) sgRNAplasmid for cloning protospacer (B) and pOS-sgRNA map showing U3 sequence, BsaI site and sgRNA scaffold (D). Destination vector for Multiplex CRISPR/Cas9 system (C) by Xie et al. (2015).

3.3.3. Mutant screening in transgenic plants

Genomic DNA was extracted from T_o-hygromycin-resistant plants and genotyped by PCR using the Cas9 and gRNA specific primers (Table 3.1). Subsequently, from the positive plants, DNA fragments across the target sites were amplified with PCR using the gene-specific primer pairs (Table 3.1). The PCR amplicons were purified with isopropanol and sequenced. The sequencing chromatograms were analysed carefully for mutations.

3.3.4. RNA isolation and RT-PCR analysis

To analyse gene expression, total RNA from different tissues was extracted with the NucleoSpin® RNA Plant kit (http://www.mn-net.com) and DNA contamination was removed using the TURBO DNA-free™ Kit according to the manufacturer's instructions (*https://www.thermofisher.com*). The RNA was reverse transcribed using the ImProm-II™ Reverse Transcription System (*https://ita.promega.com*) and the cDNA was used as template in

RT-PCR reactions included three different biological. As a control, we amplified rice *Elongation Factor1* (*EF1*) fragment for all the samples. Primers used for RT-PCR are given in (Table 3.1).

3.3.5. *In-situ* hybridization

Reproductive meristems harvested from the main stem at different stages of early panicle development were embedded in Paraplast X-TRA (*http://www.sigmaaldrich.com*) as described by Huijseret al. (1992). To generate the sense and antisense probes, gene fragments were amplified from cDNA using gene specific primers (Table 3.1), cloned into pGEM®-T Easy Vector and confirmed by sequencing. Digoxigenin-labeled antisense and sense RNA probes were generated with the DIG RNA Labeling Kit SP6/T7 from promega (*https://ita.promega.com*) according to the manufacturer's instructions. Hybridization was performed as described by Coenet et al. (1990) using an NBT/BCIP from Promega for detection.

Table 3.1:List of Primers

PROTOSPACERS FOR SINGLE CRISPR/Cas9		
GENE	OLIGOS-ID	SEQUENCE
G1L1	Osp1012 (+ve)	GGCACATCCGCGACACGCAGTCCA
	Osp1013 (-ve)	AAACTGGACTGCGTGTCGCGGATG
G1L2	Osp1016 (+ve)	GGCACTGGAGCTGTCGCGGTGCAG
	Osp1017 (-ve)	AAACCTGCACCGCGACAGCTCCAG
G1L5	Osp1024 (+ve)	GGCAGTGCCCCTTCTTCGGCCACC
	Osp1025 (-ve)	AAACGGTGGCCGAAGAAGGGGCAC
PROTOSPACERS FOR MULTIPLEX CRISPR/Cas9		
GENE	OLIGOS-ID	SEQUENCE
G1L1	Osp1271 (+ve)	TAGGTCTCACGACACGCAGTCCAGTTTTAGAGCTAGAA
	Osp1272 (-ve)	CGGGTCTCAGTCGCGGATGTGCACCAGCCGGG
G1L2	Osp1273 (+ve)	TAGGTCTCAGTCGCGGTGCAGGTTTTAGAGCTAGAA
	Osp1274 (-ve)	CGGGTCTCACGACAGCTCCAGTGCACCAGCCGGG

G1L5	Osp1275 (+ve)	TAGGTCTCACTTCTTCGGCCACCGTTTTAGAGCTAGAA
	Osp1276 (-ve)	CGGGTCTCAGAAGAAGGGGCACTGCACCAGCCGGG
PRIMERS FOR PRODUCING PTGs FOR *p*RGEB32 VECTOR		
GENE	OLIGOS-ID	SEQUENCE
L5AD5F	Osp1198	CGGGTCTCAGGCAGGATGGGCAGTCTGGGCAACAAAGCAC CAGTGG
L3AD5R	Osp1199	TAGGTCTCCAAACGGATGAGCGACAGCAAACAAAAAAAAAAGC ACCGACTCG
S5AD5F	Osp1200	CGGGTCTCAGGCAGGATGGGCAGTCTGGGCA
S3AD5R	Osp1201	TAGGTCTCCAAACGGATGAGCGACAGCAAAC
PRIMERS FOR GENOTYPING TRANSGENIC PLANTS FOR Cas9		
	OLIGOS-ID	SEQUENCE
	Atp5706 (Fw)	GTGAAGCTCAATAGAGAGGACC
	Atp5718 (Rev)	CTTGATAATCTTGAGGAGGTCGTGG
PRIMERS FOR AMPLIFYING TARGET SITE		
GENE	OLIGOS-ID	SEQUENCE
G1L1	Osp1840	GGAGATGGACATGATCGGCATGG
	Osp1841	GAAGTGCGCCGGGAACAAGAAGTG
G1L2	Osp1362	AGGTTTGCTGCTGCTTGTGC
	Osp1363	TGAGACGAAGACGAGGAGGTG
G1L5	Osp1837	GCAGATCGACGATGGAGTTCGTG
	Osp1838	GCTTCTTGCGCTTCTTCTTCTCG
PRIMERS FOR SEQUENCING TARGET SITE		
GENE	OLIGOS-ID	SEQUENCE
G1L1	Osp1842	GCAGGTACGAGTCGCAGAAGC
G1L2	Osp1362	AGGTTTGCTGCTGCTTGTGC
G1L5	Osp1837	GCAGATCGACGATGGAGTTCGTG

PRIMERS FOR qRT-PCR		
GENE	OLIGOS-ID	SEQUENCE
EF1	RT1212	TGGTATGGTGGTGACCTTTG
	RT1213	GTACCCACGCTTCACATCCT
G1L1	RT2541 (Fw)	GCACACCACACCTACCATGA
	RT2542 (Rev)	GGCTGCAGAGATCGAAGTGT
G1L2	RT1387 (Fw)	TTGCAGTGGTCTTCTTCGCA
	RT1389 (Rev)	AGAGTTTGAGGTGCAGATGTGA
G1L5	RT2543 (Fw)	GAGCTGCTAGCCTCCTACG
	RT2544 (Rev)	GCTAGTAGCAAGAGCAGCCTA
PRIMERS FOR *in-situ* HYBRIDIZATION		
GENE	OLIGOS-ID	SEQUENCE
G1L1	Osp1384 (Fw)	ACACCAAGCAGAAGCAGCAG
	Osp1385(Rev)	ATGCAAATCACCACGCATCC
G1L2	Osp1387 (Fw)	CACACTTCATGCACGGACAC
	Osp1388(Rev)	TGCTATATGCTGCTGATCTCTG
G1L5	Osp1390 (Fw)	GCGTCAGCTACGAGAAGAAG
	Osp1391(Rev)	ATTAGATGCAGTAGCAGCAGC

3.4. RESULTS

3.4.1. THE CHROMOSOMAL LOCATION, GENE STRUCTURE, AND CONSERVED DOMAIN OF ALOG FAMILY GENES

The ALOG family in rice is composed of 10 genes. The locations of *ALOG* genes on the 12 rice chromosomes are shown in Figure 3.2. *G1L7*, *G1L4*, *G1L2*, *G1* and *G1L5/TAW1* are located on chromosome 1, 4, 6, 7 and 10 respectively, whereas *G1L8* and *G1L9* are both located on chromosome 5 and the remaining three genes, *G1L1*, *G1L3* and *G1L6/TH1* are located on chromosome 2. There is no *ALOG* gene located on chromosome 3, 8, 9, 11 and 12. ALOG protein size range from 202 to 284 amino acids while the molecular weight ranges from 21.673 to 29.165 kDa as shown in Table 3.2. The conserved DUF640 domain is shown in Figure 3.3.

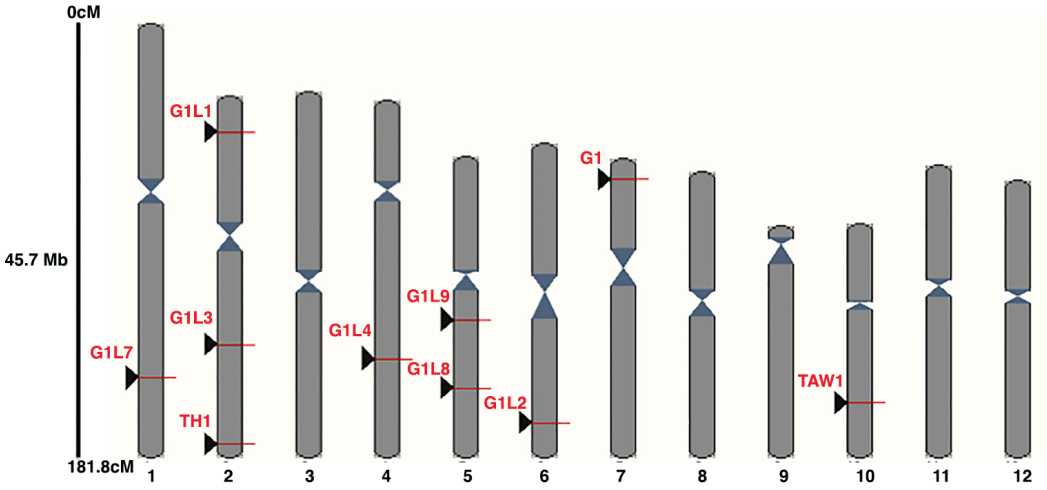

Figure 3.2. Distribution of *ALOG* genes on 12 rice chromosomes. Scale correspondent to Chr1. is 45.7 Mb

Table 3.2:ALOG family, gene Locus IDs, CDS, Protein length (AA) and Mol. Weight (kDa)

Gene Name	**Locus ID**	**CDS (nucleotides)**	**Protein Length (Amino Acids)**	**Mol. Weight (kDa)**
G1	*LOC_Os07g04670*	831	276	28.496
G1L1	*LOC_Os02g07030*	813	270	28.628
G1L2	*LOC_Os06g46030*	834	277	29.165
G1L3	*LOC_Os02g41460*	630	209	21.808
G1L4	*LOC_Os04g43580*	609	202	21.673
G1L5 (TAW1)	*LOC_Os10g33780*	615	204	21.835
G1L6 (TH1)	*LOC_Os02g56610*	747	248	25.885
G1L7	*LOC_Os01g61310*	639	212	21.743
G1L8	*LOC_Os05g39500*	717	238	24.725
G1L9	*LOC_Os05g28040*	855	284	29.440

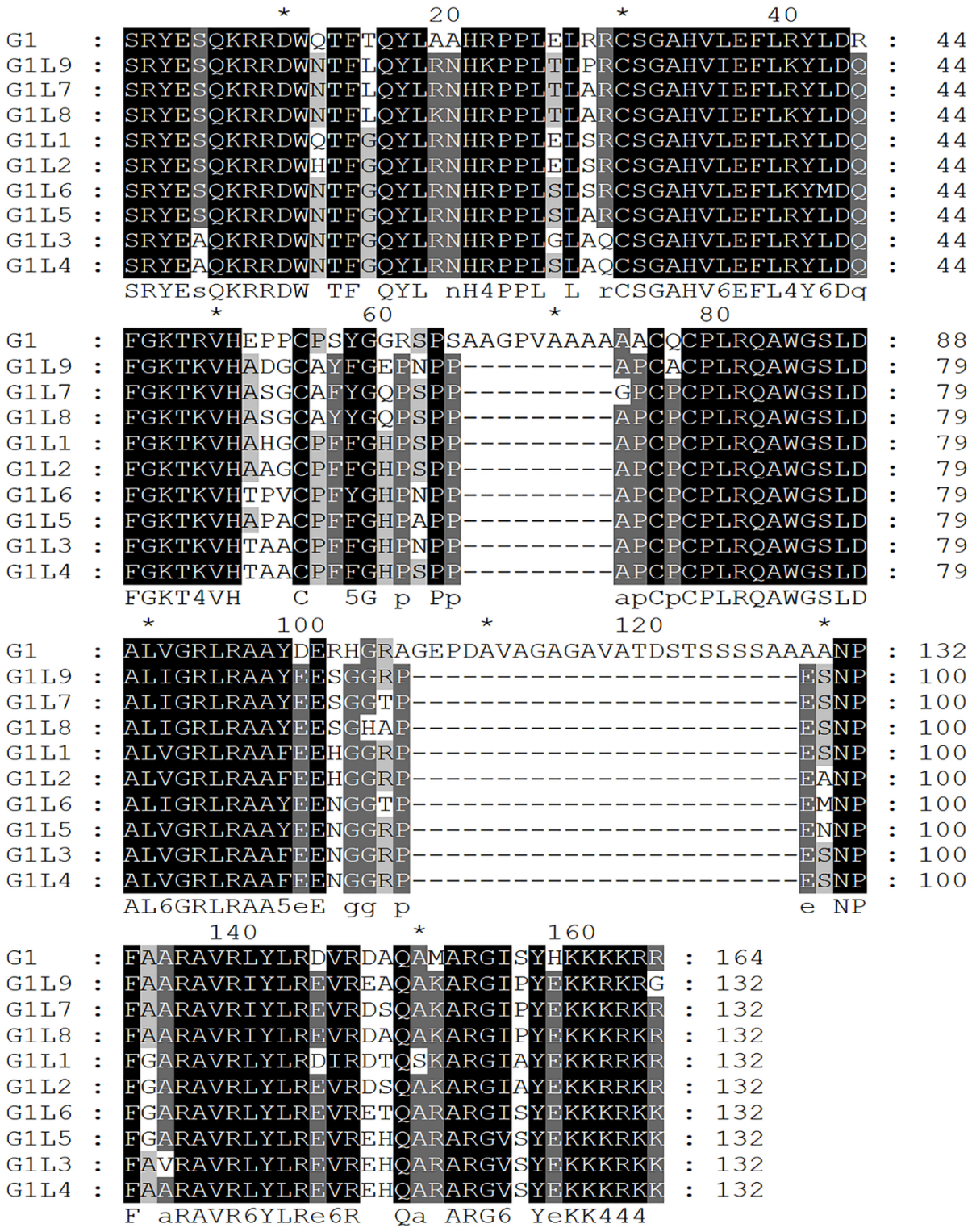

Figure 3.3. Sequence alignment of ALOG/DUF640 proteins in rice.

3.4.2. EXPRESSION PATTERN OF THE ALOG GENES DURING RICE PANICLE DEVELOPMENT AND OTHER TISSUES

Rice inflorescence architecture is determined by the activities of reproductive meristems. Reproductive meristems go through quick transformation as soon shoot apical meristem convert to IM/RM, which then form branch meristems and finally transform into determinate spikelet meristems that differentiate into floral meristems.

We investigated the expression dynamics of all members of the ALOG family using RNA seq data obtained from transcriptome analysis of four different types of reproductive meristems that were specifically dissected by laser micro-dissection microscopy (Harrop et al., 2016; see chapter 2 of this thesis). We dissected Rachis meristem (RM), Primary branch meristem (PBM), Elongated primary branch meristem when it produces axillary meristems (ePBM) and spikelet meristem (SM) shown in (Figure 3.4) Our results showed that among the 10 ALOG family genes, *G1L1*, *G1L2* and *G1L5* (*TAW1*) are highly expressed in all the four meristem types and that the expression pattern of *G1L1* and *G1L2* was similar to *G1L5* (*TAW1*) and therefore they might have a similar function in inflorescence development.

To investigate the expression patterns of these genes in more detail, we analysed vegetative (root tips, whole roots, young leaves, mature leaves, shoot apical meristem), reproductive (i.e Rachis meristem or inflorescence meristem, branch meristems and spikelet plus early floret meristems) and fruit tissues (i.e milk seeds and mature seeds) by real-time quantitative PCR (RT-qPCR). A high expression level for all three genes was observed starting from SAM to PBM and SM. *G1L1* was low expressed in vegetative tissues except the SAM and showed high expression in reproductive meristems, milk seeds and mature seeds. *G1L2* was more widely expressed and showed high expression in reproductive meristems but was also expressed in vegetative tissues including root tips, whole roots, young leaves, SAM and fruit tissue i.e mature seeds. *G1L5* was expressed in SAM and reproductive meristems and root tips. These results indicate that these *ALOG* genes might play an important role in meristem activities during growth and development of the plant organs.

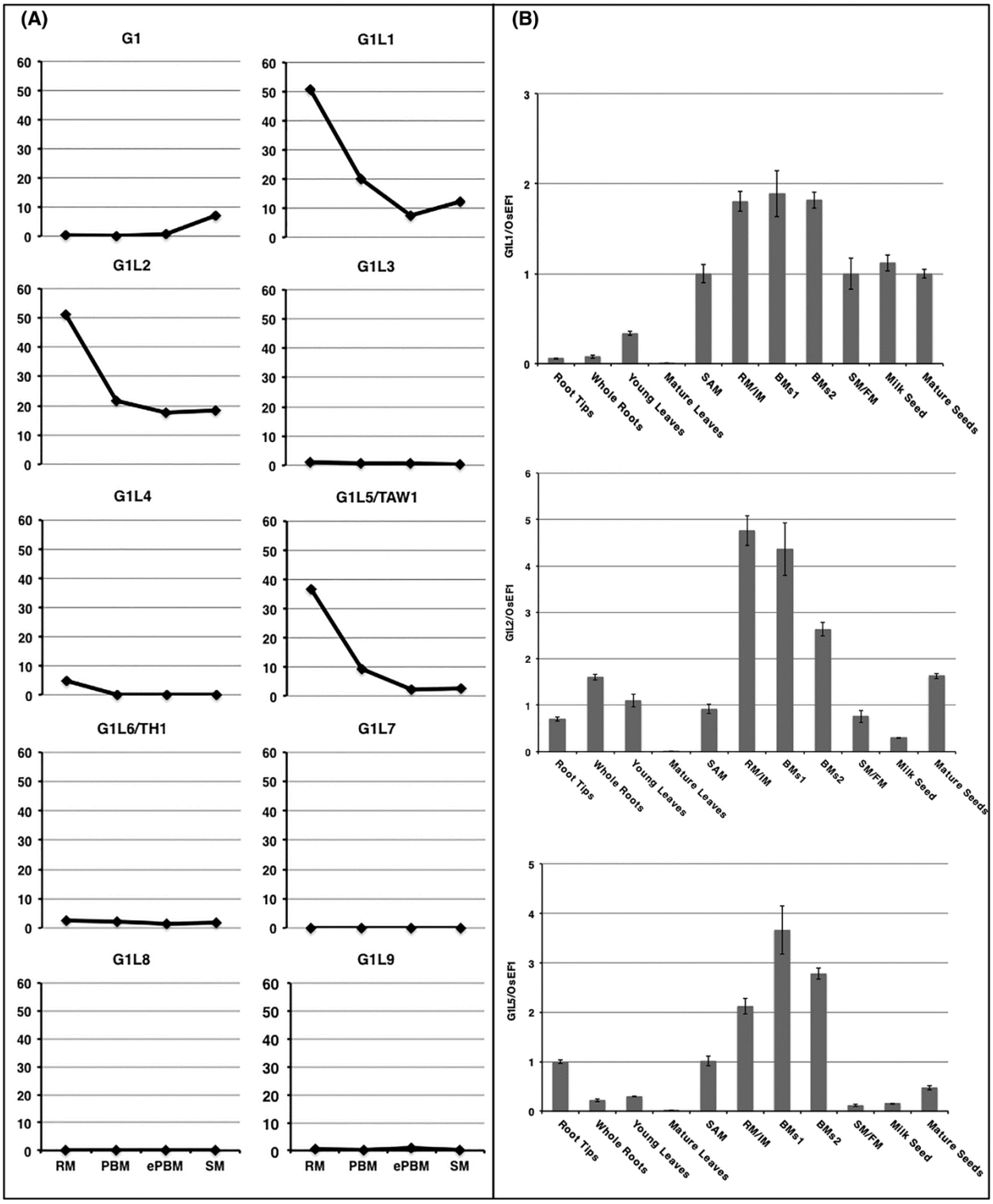

Figure 3.4. Gene Expression in Transcript Per Million (TPM) of Rice ALOG family members in four meristem types including RM, PBM, ePBM and SM (A). Data retrieved from our study on reproductive meristem transcriptome analysis (Harrop et al., 2016; see chapter 2 of this thesis). Expression of the selected genes analysed by RT-PCR across different tissues including vegetative, reproductive and fruit tissues (B).

3.4.3.EXPRESSION ANALYSIS OF *G1L1* AND *G1L2* BY IN SITU HYBRIDIZATION ACROSS REPRODUCTIVE MERISTEMS DURING EARLY STAGES OF PANICLE DEVELOPMENT

To study the spatial expression pattern of *ALOG* genes in different meristem types during inflorescence development, we selected 2 out of 10 genes i.e *G1L1* and *G1L2* from the ALOG family based on expression data obtained from the transcriptome analysis of 4 types of reproductive meristems. We applied the mRNA in situ hybridization technique using highly specific antisense digoxigenin-labeled RNA probes. For control experiment, sense digoxigenin-labeled probes were used as negative control while *G1L5* (*TAW1*) antisense probe was used as positive control. Similar to *TAW1*, both *G1L1* and *G1L2* genes showed expression in all the mersitem types including IM, BMs and SM as shown in Figure 3.5.

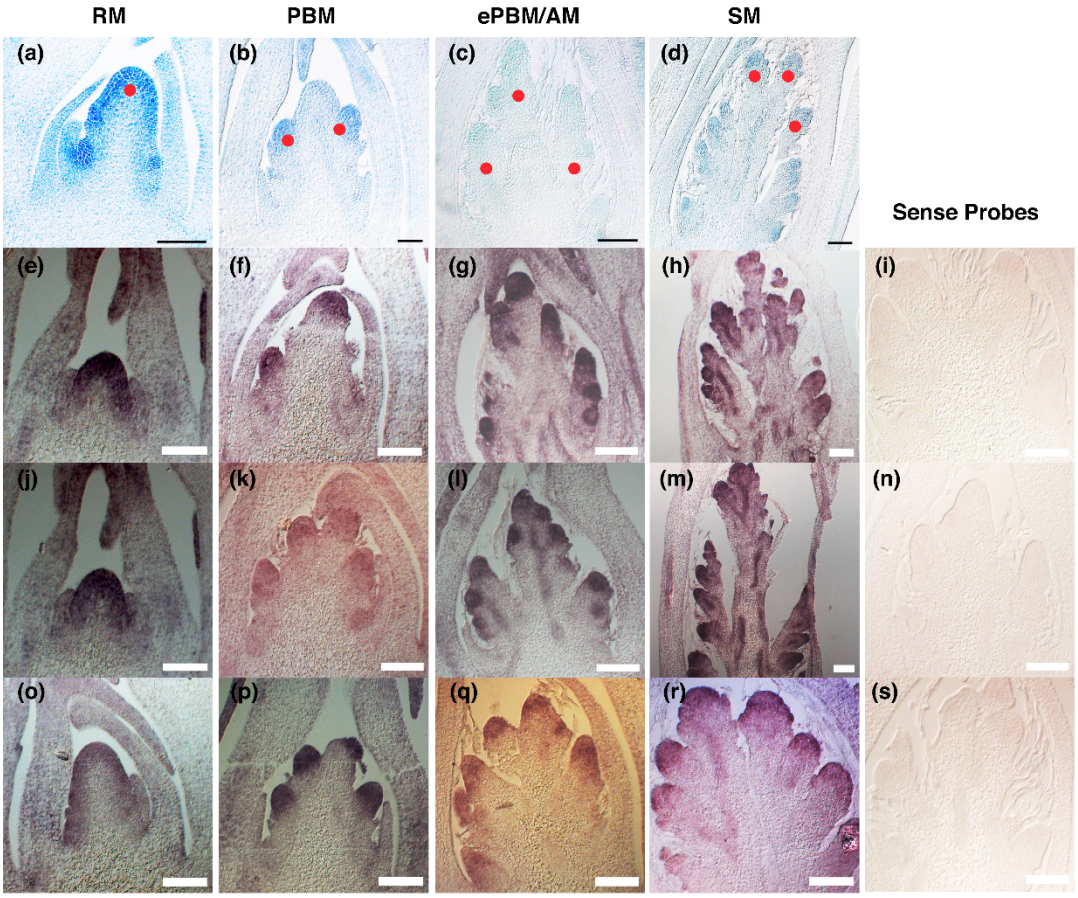

Figure 3.5. Morphological stages of early panicle development (a-d) showing Rachis Meristem (RM), Primary Branch meristem (PBM), Elongated Primary Branch Meristem or Axillary Meristem (ePBM/AM) and Spikelet

Meristem (SM). Selected ALOG genes expression pattern analysed by *in situ* hybridization in four stages. *G1L1* antisense probe (e-h), *G1L1* sense probe as negative control (i), *G1L2* antisense probe (j-m), *G1L2* sense probe (n), *G1L5* antisense probe (o-r), *G1l5* sense probe (s). Scale bars represent 50 μm (a-c) and 100 μm (d-s).

CRIPR-Cas9 technology

In order to generate mutants for the *ALOG* genes we applied the CRISPR-Cas9 technology (Miao et al., 2013; Xie et al., 2015) in rice to specifically induce mutation in two ALOG family genes, *G1L1* and *G1L2*. For *G1L1* (*OS02G0166800*), which has two exons, sgRNA was designed to target the coding sequence in the first exon and was expected to induce mutation in this region (397-417 bp from the ATG), as shown in Figure 3.6. Another sgRNA was designed to target *G1L2* (*OS06G0672400*), which has a single exon, at 131 bp downstream from the ATG start codon (Figure 3.7) and we were expecting mutations 131-151 bp downstream of the start codon in the coding sequence of *G1L2*. To generate double mutants, the sgRNAs for multiplex CRISPR/Cas9 were designed targeting the same region as the one used in the single construct. Using *Agrobacterium* mediated transformation, rice embryogenic calli were transformed with the single and multiplex CRISPR/Cas9 constructs. Positive plants were selected from regeneration on hygromycin containing media and screened for mutations.

3.4.4. *g1l1* LOSS OF FUNCTION MUTANT AND RELATED PHENOTYPE

From the first round of transformation only two plants were generated. One plant was wild-type at the target locus whereas the other line was heteroallelic. One allele showed a 6 bp (GCAGCT) deletion while the other showed 1 bp (T) deletion. The plant having no mutation in *G1L1* was used as control and both vegetative and reproductive growth was carefully observed. Preliminary phenotypic analysis of the T0 line showed that the *g1l1* mutant flowered very late and plant height was short compared to the control. For agronomic traits analysis three panicles were taken from three tillers coming from the same plant for both the control and *g1l1* mutant. The mutant exhibited differences in panicle shape and architecture producing fewer branches and total number of seeds compared to control plant. The panicle length observed was 13.83 cm in control plant while 11.1 cm in mutant. Control plant produced 8 primary and 5.5 secondary branches while mutant plant produced 2.67 and 1.5 primary and secondary branches respectively. We also found a significant difference in total number of seeds per panicle, which were 40.67 in control compared to 20 seeds per panicle in mutant. *g1l1* spikelets were

completely sterile and for this reason we were able to analyse only the T0 generation. The data presented in Figure 3.6, is obtained from T0 generation. In the second round of transformation we have generated several transgenic plants for the single *G1L1* CRISPR/Cas9 construct but they are still to be analysed.

3.4.5. *g1l2* LOSS OF FUNCTION MUTANT AND RELATED PHENOTYPE

For the sgRNA targeting *G1L2* we obtained 6 independent lines in T0 for *G1L2*. By sequencing we found that 90% of the regenerants were bearing a single bp insertion of either T, A, C or a heteroallelic mutation of T insertion on one allele and A insertion on the other allele, 3bp upstream of the PAM sequence on the target site. In the T1 generation, two mutants that were homozygous for the C insertion (Figure 3.7, c), and which have lost the Cas9 were selected for further characterization. As given in Figure 3.7, wild type and g1l2 plants showed a slight difference in panicle length which was 20 cm in wild type plants compared to 17 cm in the mutants and number of Primary Branches (PBs) per panicle were 9 in the wild type plants while 5.5 in the mutants. A difference was observed in Secondary Branches (SBs) of the wild type (11 SBs per panicle) compared to the mutants (5 SBs per panicle). Wild type plants produced on average 77 seeds per panicle while mutants produced 36 seeds per panicle. These observations were consistent with the T0 generation. No obvious phenotypic affect was observed in vegetative organs i.e tiller numbers, number of leaves, plant height and roots. The data presented in Figure 3.7, is obtained from the T1 generation.

3.4.6. *g1l1 g1l2* DOUBLE MUTANT AND RELATED PHENOTYPE

For double mutants, we obtained 9 independent transformants, which were genotyped for the presence of Cas9 to check the transformation efficiency and it was 100%. All the positive lines were sequenced and screened for mutations in both genes, *G1L1* and *G1L2*. We found 7 lines carrying mutations on both loci and two lines with only mutations at the *G1L2* locus. One line shown in Figure 3.8, is biallelic for *G1L1* with a 1bp (T) insertion and a 2bp (AG) deletion. The same line is also biallelic for *G1L2* with a 1bp (T) insertion and a 1bp (A) insertion. Detail for all the mutations in the 9 independent lines is shown in Table 3.3. All the independent lines showed similar phenotype in the panicle but compared to the single mutants the phenotype was more drastic. Control plants from the same transformation but wild-type for both the *G1L1* and *G1L2*

loci have a panicle length of 13.83 cm and produced 8 primary and 2.67 secondary branches per panicle with average of 40.67 total numbers of seeds per panicle. The double mutants shown average panicle length of 13 cm with 5.4 primary and 0.857 secondary branches respectively which are less when compared to control plants and single mutants *g1l1* or *g1l2* as well. The mutants produced fewer panicle branches with less seeds compared to the wild type control. Average total numbers of seeds per panicle produced by double mutants were 20.74 compared to 40.67 in control plants, 20 seeds in *g1l1* and 36 in *g1l2*. Similar to *g1l1* single mutant, also in the double mutant spikelets were completely sterile.

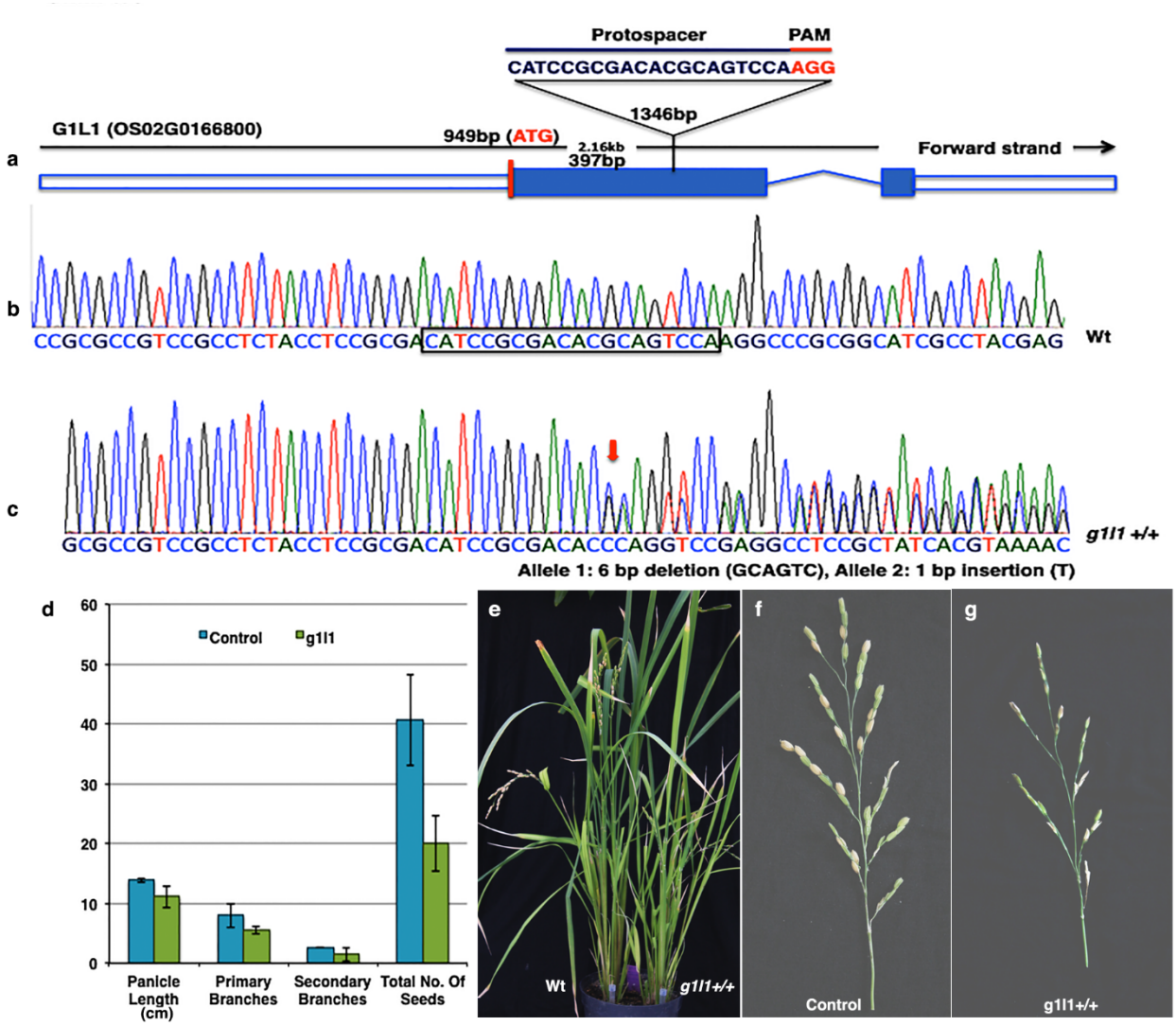

Figure 3.6. *g1l1* loss of function mutant and related phenotype. *G1L1* transcript structure representing ATG site (949 bp downstream of 5' end), protospacer sequence and target site (397 bp upstream of ATG) for gRNA. Chromatograms showing Wt. *G1L1* genomic sequence (b) and *g1l1* mutant with a 6bp (GCAGTC) deletion and a

1bp (T) insertion. Phenotypic characterization of Wt. and *g1l1* panicle for panicle length, number of primary and secondary branches and total number of seeds (d). Mature plants Wt/Control and g1l1 (e). Mature panicle of Wt/Control and *g1l1* (g).

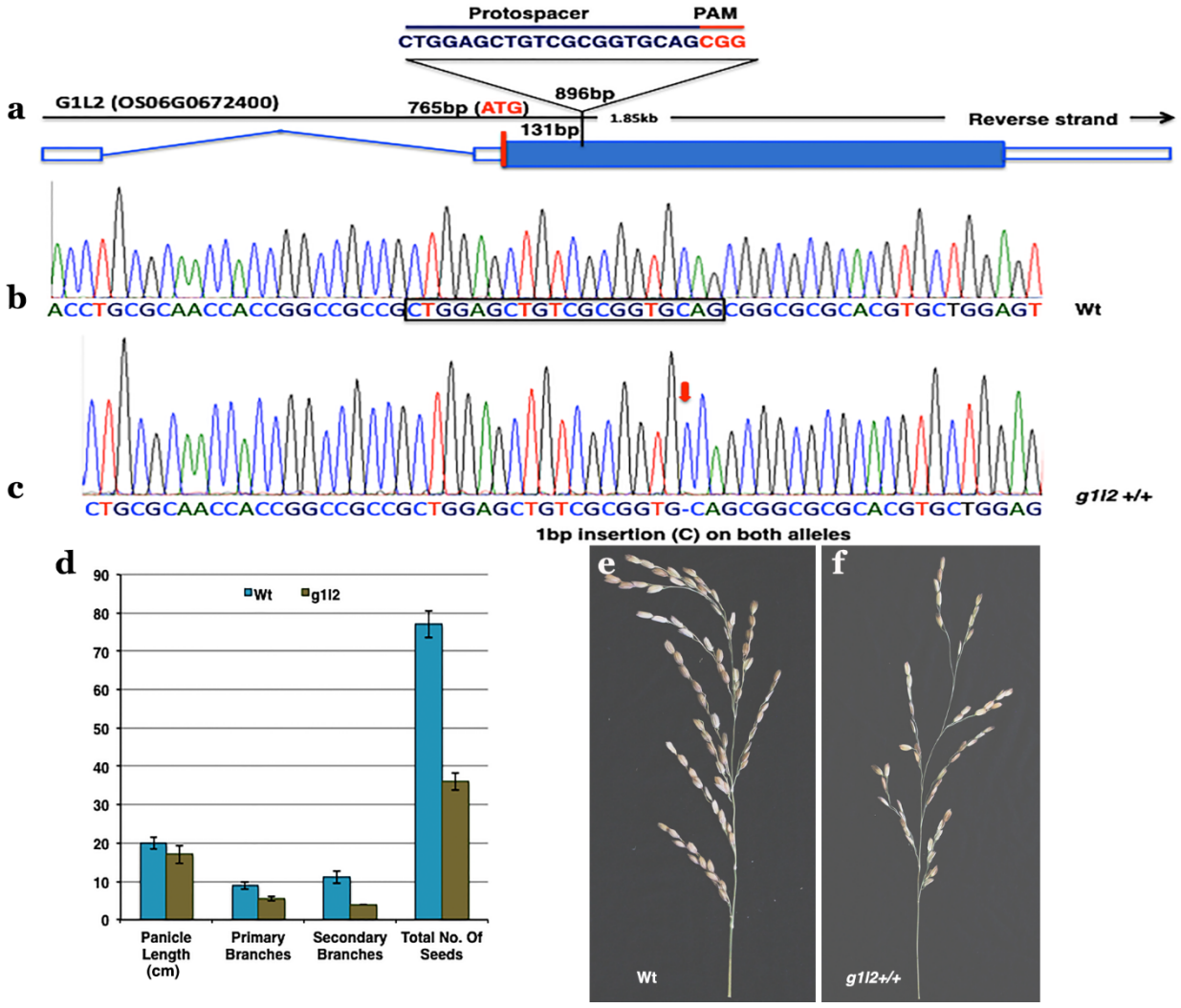

Figure 3.7. *g1l2* loss of function mutant and related phenotype. *G1L2* transcript structure representing ATG site (765 bp downstream of 5' end), protospacer sequence and target site (131 bp upstream of ATG) for gRNA. Chromatograms showing Wt *G1L2* genomic sequence (b) and *g1l2* mutant with 1bp C insertion 3bp upstream of PAM site. Phenotypic characterization of Wt and *g1l2* panicle for panicle length, number of primary and secondary branches and total number of seeds (d). Wt (left) and *g1l2* (right). Mature panicle of Wt (e) and *g1l2* (f).

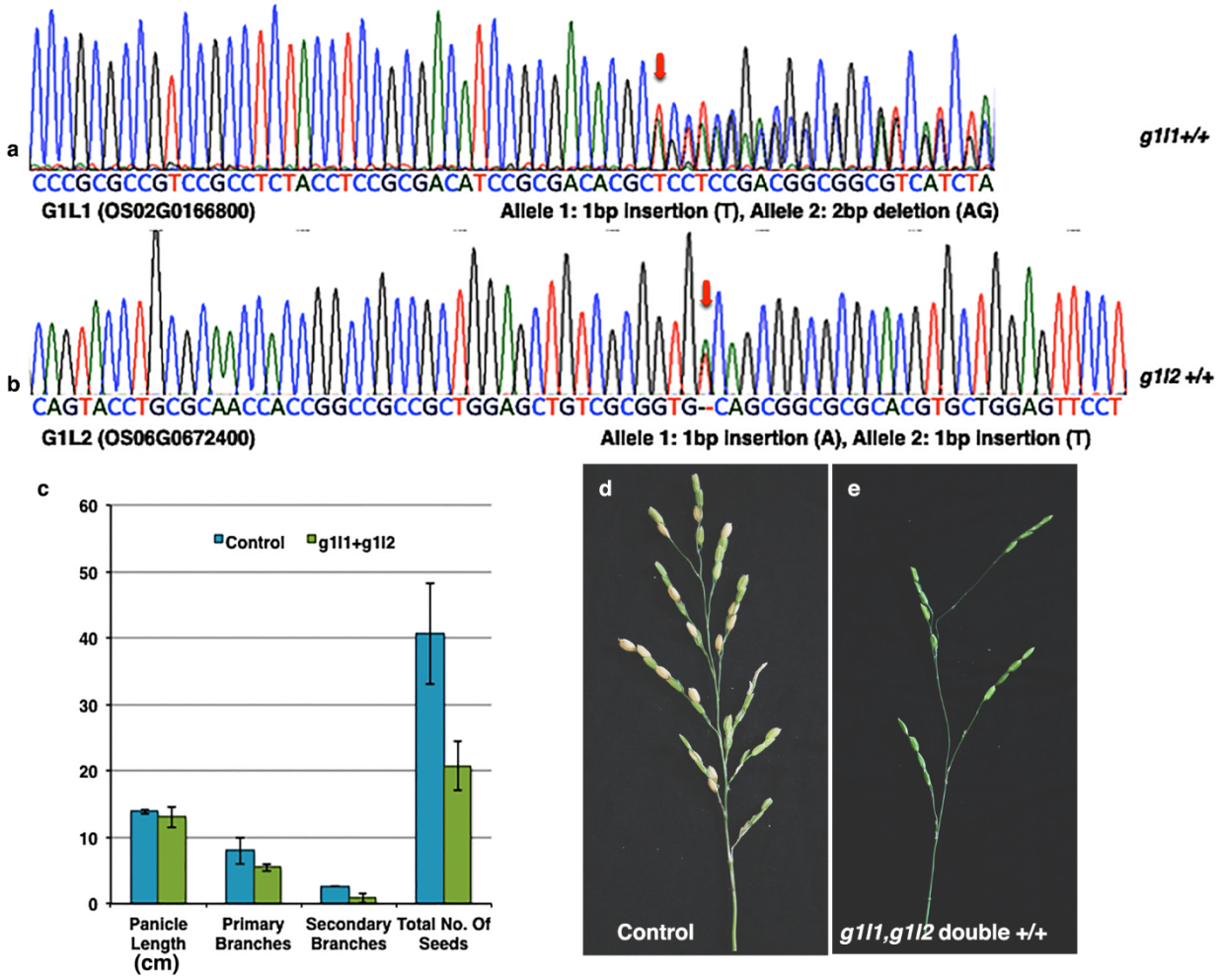

Figure 3.8. *g1l1 g1l2* double mutant and related phenotype. Chromatograms representing *g1l1 g1l2* double mutant showing biallelic mutation for both genes. Single T insertion on one allele and 2bp (AG) deletion on second allele for *g1l1* (a), 1bp A insertion on one allele and 1bp T insertion on second allele for *g1l2*. Phenotypic characterization of Wt/Control and *g1l1 g1l2* panicle for panicle length, number of primary and secondary branches and toatal number of seeds (c).Mature panicle of Control (d) and *g1l1 g1l2* (e).

Table 3.3: Detail of mutations at *G1L1* and *G1L2* loci for the double mutants *g1l1+g1l2*

T0 Plants	Cas9	Mutations	
		g1l1	***g1l2***
07	*+ve*	Allele 1: 1bp insertion (T) Allele 2: 2bp deletion (AG)	Allele 1: 1bp insertion (A) Allele 2: 1bp insertion (T)
13	*+ve*	Allele 1: 1bp insertion (T) Allele 2: 2bp deletion (AG)	Allele 1: 1bp insertion (A) Allele 2: 1bp insertion (T)
15	*+ve*	Allele 1: Wt Allele 2: Wt	Allele 1: 1bp insertion (T) Allele 2: 31 bp deletion
16	*+ve*	Allele 1: 1bp insertion (T) Allele 2: 1bp insertion (T)	Allele 1: 1bp insertion (T) Allele 2: 1bp insertion (T)
17	*+ve*	Allele 1: 30 bp deletion Allele 2: 21 bp deletion	Allele 1: 1bp insertion (G) Allele 2: 1bp insertion (T)
18	*+ve*	Allele 1: 2bp deletion (AG) Allele 2: 1bp insertion (T)	Allele 1: 1bp insertion (T) Allele 2: 1bp insertion (A)
19	*+ve*	Allele 1: 2bp deletion (CT) Allele 2:3bp deletion (ACT)	Allele 1: 1bp insertion (T) Allele 2: 1bp insertion (T)
20	*+ve*	Allele 1: Wt Allele 2: Wt	Allele 1: 1bp insertion (G) Allele 2: 1bp insertion (T)
21	*+ve*	Allele 1: 1bp insertion (G) Allele 2: 1bp insertion (T)	Allele 1: 1bp insertion (G) Allele 2: 1bp insertion (T)

3.5. DISCUSSION

The ALOG family in rice consists of 10 genes distributed on 7 out of 12 chromosomes as shown in Figure 3.2. All these genes encode protein with a conserved domain named DUF640 (domain of unknown function 640) in PFAM database (http://pfam.xfam.org). Phytozome Blast (https://phytozome.jgi.doe.gov/pz/portal.html) and NCBI Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi) revealed the presence of ALOG members in all the land plants including *Physcomitrella patens* (bryophyte), *Selaginella moellendorffii* (lycophyte), *Amborella trichopoda* (angiosperm) but not in algae, fungi and animals (data not shown), which indicates that ALOG family is specific to land plants and evolved during early evolution. Arabidopsis and rice have 10 members each while tomato encodes 11 ALOG family members. In rice members of ALOG family are present in all the species including wild type, *Oryza rufipogon*, *Oryza barthii* and cultivated species, *Oryza glaberrima* and *Oryza sativa* (*japonica* and *indica*), however the number of these genes in each species is not the same.

Only three members of ALOG family have been studied in rice. They showed to play roles in inflorescence and spikelet development. *LONG STERILE LEMMA1 (G1)/ELONGATED EMPTY GLUME (ELE)* was the first member of the ALOG family to be characterized in rice by Yoshida et al. (2009). Its expression was investigated in floral organs where it was only expressed in sterile lemma primordia through out development while there was no expression in lemma. Its function in specification of the sterile lemma by repressing lemma identity and demonstrates that continuous expression of *G1* in sterile lemma primordia throughout development is consistent with the *g1* phenotype in which spikelets display a long sterile lemma. The phenotype of this *g1* mutant is caused by a missense mutation in the conserved domain, which suggests that the conserved ALOG domain critical is for protein function. *G1* function was further confirmed by the fact that the *Oryza grandiglumis* species, which has a natural allelic variation at the *G1* locus, also produces spikelets with long sterile lemmas. *TRIANGULAR HULL1 (TH1)* transcript was detected in vascular tissues of lemmas, paleas, inflorescence branches and leaves using *in situ* hybridization. Stronger signal was observed in distal part of the lemma but not in the basal part. The *th1* mutant reduced the size of tubercles on the surface of the lemma and palea and the phenotype was stronger in the distal part of the lemma compared to the basal part which resulted in beak like spikelet. However, the number and spatial distribution of tubercles was not affected. These findings show that *TH1* functions in lemma and palea morphogenesis (Li et al., 2012).

Another member of the ALOG family, *TAW1* which is expressed in Shoot Apical Meristem (SAM), Inflorescence Meristem (IM), Branch Meristems (BMs) and Spikelet Meristems (SMs), plays an important role in determining rice inflorescence architecture by delaying SMs identity and enhancing BMs activities producing more reproductive branches and higher spikelet numbers in the inflorescence. ALOG genes in other species for example *LSH3* and *LSH4* of Arabidopsis are shown to be involved in the regulation of meristem function by modulating meristem cells proliferation and differentiation (Cho and Zambryski, 2011; Takeda et al., 2011) and *Terminating Flower* (*TMF*), a closest homolog of *TAW1* in tomato has been shown to have a similar and conserved function. In the *tmf* mutant the primary SAM instead of producing a sympodial inflorescence, terminates in a single flower (MacAlister et al., 2012). All the aforementioned reports suggest that ALOG family play important role in inflorescence development by putatively regulating meristem activities and function. Here we study two other members of the ALOG family, namely *G1L1* and *G1L2*, located on Chr2 and Chr6 and which share similar expression patterns with *TAW1*.

Our analysis of the RNA-Seq Data from reproductive rice meristems (Harrop et al., 2016) revealed that only *G1L1, G1L2* and *G1L5*/*TAW1* were expressed in the reproductive meristems. The *taw1* missense mutation resulted in small inflorescences with fewer branches and grain number and this phenotype was even stronger in RNAi lines. Our hypothesis is that the increased phenotype of the RNAi lines might be explained by the fact that *TAW1* is co-expressed with *G1L1* and *G1L2* and that the RNAi silencing constructs may also be affecting the activity of these genes. This suggests that these genes have a redundant function. To investigate the role of *G1L1* and *G1L2* in inflorescence development we performed a functional analysis of these genes. The expression pattern of the selected genes was further confirmed by RT-PCR in both vegetative and reproductive tissues and also by in-situ hybridization in reproductive meristems. Our results demonstrate that *G1L1*, *G1L2* and *TAW1* are expressed in all the reproductive meristems analysed by RNA-seq, RT-PCR and in-situ hybridization.

By loss of function mutant analysis, generated with CRISPR/Cas9 genome editing technology, we demonstrated that both single mutants *g1l1* and *g1l2* produced less inflorescence branches and reduced grain yield and the phenotype was stronger in the double mutant in respect to the single mutants however panicle length was not affected significantly. These results suggest that *G1L1* and *G1L2* control lateral meristem initiation/formation in the inflorescence but unlike

TAW1 there was no significant difference observed in flowering time in single and double *g1l1*, *g1l2* mutants. It will be interesting to study the triple mutant of *G1L1*, *G1L2* and *TAW1* to check if there is any functional redundancy between these ALOG genes.

Meristem size is regulated by the maintenance of the stem cells pool, which replenished after it consumes in lateral organs formation (Barton 2010). Mutation in CLV pathway genes enlarge the size of meristems, which lead to increase in shoot and inflorescence branches, more flowers and ectopic organs formation in flowers and fruits (Laufs et al., 1998). Inflorescence branches and grain numbers are determined by reproductive meristems formation, maintenance and differentiation. Branching pattern of the inflorescence can also be explained by rate of initiation of BMs and transition to SMs (Prusinkiewicz et al., 2007; Park et al., 2014). Studying the activities of early reproductive shoot apex while analysing 8 cultivars of *japonica* rice revealed that the initial size of the IM control the number of PBs but has no association with the number of spikelets per PB. Spikelet numbers on each primary branch are determined by the cell division activity in early reproductive apex (Mu et al 2005). It is also reported that altered inflorescence architecture in *apo1* and *apo2* mutants, having a small panicle with little PBs and SBs compared to wild type, is putatively associated with smaller IM displaying reduced cell proliferation rate. Overexpression of *APO1* resulted in larger IM and highly branched inflorescence (Ikeda et al., 2009; 2012). Similarly overexpression of *SPL14,* a regulator of *miR156*, leads to smaller IM and produce smaller inflorescence with fewer branches (Wang et al., 2015).

Our results indicated that *G1L1* and *G1L2* might be involved in regulating meristem activities and function during inflorescence development. It will be interesting to analyse the meristem size in *g1l1* and *g1l2* mutants to see if the branching phenotype in these mutants is related to meristem size. Also by checking the expression of meristem marker gene like *OSH1* in the *g1l1* and *g1l2* mutant background can further explain the role of these genes in meristem maintenance. By investigating co-expression and identifying putative target genes in the mutant back ground will further add to our understanding how these genes interact with each other in order to regulate panicle branching and grain numbers in rice.

REFERENCES

Barton MK. 2010. Twenty years on: the inner workings of the shoot apical meristem, a developmental dynamo. *Dev. Biol.* 341, 95–113.

Cho E, Zambryski PC. 2011. Organ boundary1 defines a gene expressed at the junction between the shoot apical meristem and lateral organs. Proc Natl Acad Sci U S A. 108: 2154-2159.

Coen ES, Romero JM, Doyle S, Elliott R, Murphy G and Carpenter R. 1990. floricaula: a homeotic gene required for flower development in Antirrhinum majus. Cell, 63, 1311–1322.

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.

Harrop, TW, Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, et al., 2016. Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. Plant J. 86, 75–88.

Hiei Y, Ohta S, Komari T and Kumashiro T. 1994. Efficient transformation of rice (Oryza sativa L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. Plant J. 6(2):271-282.

Hoshikawa K. 1989. The growing rice plant: An Anatomical Monograph. Nobunkyo, Tokyo107-Japan.

Huijser P, Klein J, Lönnig WE, Meijer H, Saedler H. and Sommer H. 1992. Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene *squamosa* in *Antirrhinum majus*. EMBO J. 11, 1239–1249.

Ikeda K, Nagasawa N, Nagato Y. 2005. *ABERRANT PANICLE ORGANIZATION1* temporally regulates meristem identity in rice. Dev Biol, 282:349-360.

Ikeda K, Sunohara H and Nagato Y. (2004) Developmental course of inflorescence and spikelet in rice. Breed. Sci. 54, 147–156.

Ikeda-Kawakatsu K, Maekawa M, Izawa T, Itoh J, Nagato Y. 2012. *ABERRANT PANICLE ORGANIZATION 2/RFL*, the rice ortholog of *Arabidopsis LEAFY*, suppresses the transition from inflorescence meristem to floral meristem through interaction with APO1. Plant J. 69:168-180.

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.

Itoh J, Nonomura K, Ikeda K, Yamaki S, Inukai Y, Yamagishi H, Kitano H, Nagato Y. 2005. Rice plant development: from zygote to spikelet. Plant Cell Physiol. 46:23-47.

Iyer LM, Aravind L. 2012. ALOG domains: provenance of plant homeotic and developmental regulators from the DNA-binding domain of a novel class of DIRS1-type retroposons. Biol Direct 7: 39.

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.

Kellogg EA, Camara PE, Rudall PJ, Ladd P, Malcomber ST, Whipple CJ, Doust AN. 2013. Early inflorescence development in the grasses (Poaceae). Front Plant Sci. 4:250.

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.

Kobayashi K, Yasuno N, Sato Y, Yoda M, Yamazaki R, Kimizu M, Yoshida H, Nagamura Y, Kyozuka J. 2012. Inflorescence meristem identity in rice is specified by overlapping functions of three AP1/FUL-like MADS box genes and *PAP2*, a SEPALLATA MADS box gene. Plant Cell. 24:1848-1859.

Kobayashi Y, Weigel D. 2007. Move on up, it's time for change—mobile signals controlling photoperiodic-dependent flowering. Genes and Development 21: 2371–2384.

Laufs P, Olivier G, Claudia J, Kiên K, Traas J. 1998. Cellular parameters of the shoot apical meristem in Arabidopsis. Plant Cell. 10:1375–1389.

Liu C, 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. Dev. Cell 24:612-622.

MacAlister CA, Park SJ, Jiang K, Marcel F, Bendahmane A, Izkovich Y, Eshed Y, Lippman ZB. 2012: Synchronization of the flowering transition by the tomato *TERMINATING FLOWER* gene. Nat. Genet. 44:1393-1398.

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X et al., 2013. Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res 23: 1233–1236.

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. Nat Genet. 42:545-549.

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. Plant J. 29:743-750.

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

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

Takacs EM, Li J, Du C, Ponnala L, Janick-Buckner D, Yu J, Muehlbauer GJ, Schnable PS, Timmermans MC, Sun Q et al.: Ontogeny of the maize shoot apical meristem. 2012. Plant Cell, 24:3219-3234.

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

Toki S. 1997. Rapid and efficient Agrobacterium-mediated transformation in rice. Pl. Mol. Biol. Reporter. 15(1):16-21.

Wang L, Sun S, Jin J, Fu D, Yang X, Weng X, Xu C, Li X, Xiao J, Zhang Q. 2015. Coordinated regulation of vegetative and reproductive branching in rice. Proc Natl Acad Sci USA. 112:15504–15509.

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

Wilkinson MD, Haughn GW. 1995. *UNUSUAL FLORAL ORGANS* controls meristem identity and organ primordia fate in *Arabidopsis*. Plant Cell. 7:1485–99.

Xie K, Minkenberg B, Yang Y. 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci U S A. 112(11):3570-5.

Yoshida A, Suzaki T, Tanaka W, Hirano HY. 2009. The homeotic gene *Long Sterile Lemma* (*G1*) specifies sterile lemma identity in the rice spikelet. Proc Natl. Acad. Sci. USA 106:20103-20108.

Zhao L, Nakazawa M, Takase T, Manabe K, Kobayashi M, Seki M, Shinozaki K, Matsui M. 2004. Overexpression of *LSH1*, a member of an uncharacterised gene family, causes enhanced light regulation of seedling development. Plant J. 37:694-706.

4. CONCLUSION AND FUTURE PROSPECTS

Nowadays functional genomics studies apply sophisticated technologies to gain deep understanding of the molecular genetic mechanisms that are underlying important biological processes. In rice, which feeds almost half of the world population, the functional characterization of genes can help to understand their function and this knowledge can be used for classical and modern breeding approaches to improve the yield performance of the crop. This thesis focuses on the molecular mechanisms controlling inflorescence development by transcriptome analysis using laser microdissected inflorescence meristems. These datasets were subsequently used to select genes that putatively control meristem functions and these were targeted by a genome editing approach to create mutations in these genes.

The global population is expected to reach 9.7 billion people by 2050 (United Nation, 2015 Revision). Population burst of this enormity is placing unprecedented demand on natural resources and agriculture production to meet the challenge of growing society's food need (Foley et al., 2011). This combined with effects of global climate change renders meeting with the required food production even more challenging. For improving crop productivity, at one side it is important to exploit modern technologies and to ensure proper crop and soil management practices while limiting degradation of natural resources and environment. On the other side it is pertinent to produce new crop varieties that can perform better under different environmental constrains like biotic and abiotic stresses and give higher yield with less input (Fan et al., 2012). Genetic improvement offers great potential to develop new crop varieties with higher yield (Miflin, 2000). Functional genomic studies, which rely on molecular genetics approaches (forward and reverse genetic screens) are followed to identify new genes that influence yield related traits (McCouch et al., 1995; Ashikari & Matsuoka 2006). Due to the availability of sequenced genomes and well established transformation techniques in *Arabidopsis* and rice, molecular genetic approaches are relatively straight forward to identify the genes that regulate certain function, compared to conventional breeding approaches which are laborious and time consuming (Takeda & Matsuoka 2008). Molecular functional genomic studies help to explore the information stored in the genes and regulatory elements of the genome, to decipher the gene networks that underlie certain biological pathways and to study the overall functions while looking into the phenotypes or responses of mutant alleles to biotic and abiotic stresses. As

described in **Chapter 1,** from 12 rice chromosomes, 55,986 genes (loci) have been identified, of which 6,457 have 10,352 additional alternative splicing isoforms reaching to a total of 66,338 transcripts in the rice genome (http://orygenesdb.cirad.fr/index.html). In an attempt to ascribe biological functions to all these predicted genes, a great effort has been done by different groups around the globe in generating mutant libraries through Tos17 tagging, T-DNA insertion, Ds/dSpm tagging, and chemical/ irradiation mutagenesis.

Advancement in omics technologies, computational tools and innovation of new genome editing methods have greatly eased and speeded up functional genomic research in rice and other model organisms. As mentioned in **Chapter 1**, over the last decade NGS technology is becoming faster and cheaper, which can easily be applied to study genes expression pattern in different cells, tissues and organs or under different growth conditions. The emerging CRISPR/Cas technology can be used to induce specific mutation in one gene, multiple genes, nearly everywhere in the genome to elucidate the functional elements in the genome. Genome editing using CRISPR/Cas9 method can also be applied to induce mutations in QTLs that negatively regulate the agronomic traits. The aforementioned approaches have contributed to our understanding and provided information on genetic components that can be exploited effectively for crop improvement. In rice CRISPR/Cas technology is successfully applied to modify several traits, including herbicide tolerance, biotic and abiotic stress response, fertility, architecture and yield by inducing targeted mutations in single or multiple genes (Ikeda et al., 2016; Li et al., 2016a,b,c; Osakabe et al., 2016; Sun et al., 2016; Wang et al., 2016; Xu et al., 2014, 2016; Zhou et al., 2015).

The molecular mechanisms controlling the meristematic activities that determine panicle architecture in rice, which is the focus of this thesis, remain largely uncharacterized. We recently published transcriptome datasets for specific inflorescence meristem types providing insight into the spatiotemporal expression of genes occuring during early stages of panicle development. These data will yield useful information and will set a baseline for reverse genetics approaches to understand the molecular mechanisms underlying rice panicle development. We believe that functional characterization of genes selected based on expression patterns as reported in this thesis will provide in the future deep insight into molecular network involved in inflorescence meristems identity specification.

4.1 MERISTEMS SPECIFIC TRANSCRIPTOME ANALYSIS

A complete overview of developmental stages and morphology of the rice inflorescence is described in detail in **Chapter 1.** To decipher the transcriptional changes at the early dynamic phase of inflorescence development, starting from the switch to the reproductive phase and subsequent initiation of morphologically distinct meristems, we started to analyze the morphology of these reproductive meristems to understand better their differences at different developmental stages. Based on this analysis we were able to select 3 distinct stages of indeterminate meristems including Inflorescence Meristems (IMs), Primary Branch Meristems (PBMs), elongated PBMs/Axillary Meristems (ePBMs/AM) and 1 of the determinate meristems Spikelet Meristems (SMs), described in **Chapter 2**. We used for these studies the japonica rice cv., Nipponbare, because of its photoperiod sensitive nature, which allowed us to synchronize flowering since the floral transition can be induced by changing the photoperiod after 7 weeks from long day conditions (14h light and 10h dark), to short day conditions (10h light and 14h dark). The four meristem stages can be harvested starting from 6 to 14 days (Green house) and 10 to 18 (Phytotron) under short day.

In last decade several transcriptome studies have been reported in rice, measuring gene expression profiles of various developmental stages, i.e transcriptome analysis performed on 48 organs and tissues through out the life cycle of the rice plant under natural conditions including later stages of inflorescence development (Sato *et al.*, 2011); spatiotemporal expression profiling of 19 stages covering both vegetative and reproductive phases by microarray analysis (Sharma et al., 2012); a global expression profile of young panicles in the *OsMADS1*-RNAi background (Khanday et al., 2013), and genome wide expression profiling of a wild type japonica breeding line, 933 and a Clustered-Spikelet mutant *sped1-D* (Guanghuai et al., 2014). Laser microdissection and RNA-seq has also been used for studying the role of MADS-box genes controlling inflorescence meristem identity (Kobayashi *et al.*, 2012). In contrast to previous transcriptomic datasets, we have defined the precise transcriptome of only the early reproductive meristems develop in sequential and time dependent manner following laser dissection combined with RNA-seq method in order to restrict the measurement of gene expression to specific meristem types providing highly specific and accurate datasets.

Genes that are coordinately regulated and share similar expression patterns cluster together and can yield clues for predicting function of uncharacterized genes if any gene in the cluster has a

known function (Mantegazza et al., 2014). By clustering genes, we recovered 8 clusters (1-8), each containing genes that exhibit a common pattern of expression demonstrated in **Chapter 2.** A marked change in transcript profile was observed between RM and axillary meristems, followed by gradual changes during transitions between PBM and SM. This could be explained by the difference between the fates of the rachis meristem, which aborts after PBM differentiation (Ikeda et al., 2004), and the axillary meristems, which undergo transition to determinate SMs. The latter process may be comparable to the gradual meristem maturation observed during tomato inflorescence development (Park *et al.*, 2012), rather than to the continual production of determinate floral meristems characteristic of the Arabidopsis inflorescence meristem. Numerous new genes were identified, including transcription factor and putative hormones related genes potentially controlling meristem activities during panicle development based on sharing common expression patterns with known regulators. Functional characterization of these genes will contribute to the generation of a genetic model describing the regulatory networks that determine inflorescence architecture by controlling reproductive meristems identities and specification.

When performing cell or tissue specific gene expression, there is always the concern that the cells or tissue are contaminated even if dissected precisely with a laser microdissector. We checked by *in-situ* hybridization the localization of the expression of four uncharacterized genes that are according to the RNA-seq data expressed only in one of the 4 selected stages. These experiments confirmed that the expression was really specific for each meristem type. These genes with meristem specific expression may be used as marker genes in the future, especially for analyzing mutants affected in determinacy of reproductive meristem identity. However, it might also interesting to analyze the function of these genes.

To further confirm the reliability of our data we picked the genes already reported in literature and compared the expression to our RNA-seq data and we found the same expression for most of the genes we analyzed, however there were a few exceptions but this might be justified by the fact that RNA-seq is a sensitive procedure and the results might vary with genetic background of the plant as well as the growth conditions.

Furthermore, the analysis of the ALOG genes as described in **chapter 3** showed that also these genes are expressed as predicted by our transcriptome datasets.

4.2 FROM RNA-SEQ TO GENOME EDITING

Transcriptome analysis helps to elucidate the quantitative expression dynamic of the complete set of transcripts in a specific cell types for a certain developmental stage or physiological condition. (Wang et al., 2009). RNA-Seq method, which is rapidly developing has the advantage to offer a global view of the transcriptome and its organization for a number of cell types in different organisms (Cloonan et al., 2008). Analysing the transcriptome is helpful in functional genomic studies, which can help to interpret the biological function of genetic elements to understand for example development and diseases (wang et al., 2009). In plants many studies on transcriptome analysis have been reported (e.g. Jiang et al., 2014; Kobayashi *et al.*, 2012; Mantegazza et al., 2014; Sato *et al.*, 2011; Sharma et al., 2012), in which they have presented large gene expression data sets. It is convenient to pick candidate genes from these datasets, however to study the function of these genes there is need to generate loss of function alleles for these genes. As mentioned in **Chapter 1,** the limited availability of T-DNA insertion lines and limitations associated with random mutagenesis, RNA interference, antisense RNA and virus-induced gene silencing make genome editing technology a preferred tool for researchers to generate loss of function mutants. Among the current generation of nucleases with capabilities to precisely target loci in the genome, the CRISPR/Cas system emerged more rapidly because of several advantages, including versatility, simplicity, high efficiency, specificity and acquiescence to multiplexing, and for these particular reasons it seems to be immensely promising in plants (Sergiu et al., 2017). Moreover by implementing a multiplex approach to target multiple genes simultaneously we will be able to elucidate pathways and help engineering of complex multigenic yield related traits in crop plants, even when these genes are closely linked in the genome (Kumar and Jain 2015). The CRISPR/Cas9 type-II system consists of a Cas9, an endonuclease, which is more versatile in terms of genome engineering in many organisms (Sontheimer et al., 2015). We successfully applied the CRISPR/Cas9 system for targeted mutagenesis in a single gene (Miao et al., 2013) and another multiplex system for generating double and triple mutants (Xie et al., 2015).

Based on our transcriptome data the selected 2 ALOG family genes, *G1L1* and *G1L2*, that were co-expression with *TAW1*, another member of the same family shown to be involved in inflorescence development (Yoshida et al., 2013) (**Chapter 2**). The fact that having high sequence similarity plus the mutant phenotype of the single *taw1* mutant is enhanced when using

an RNAi approach suggests that there could be a functional redundancy between these genes. Based on this hypothesis we generated single and multiple knockout mutants (**Chapter 3**). Based on preliminary observations both *g1l1* and *g1l2* single mutants show a phenotype in panicle branching and the phenotype is enhanced in the *g1l1 g1l2* double mutant. From this we can conclude that they function independently in regulating panicle development. However the *g1l1* single mutant and double mutant were sterile in the T0 generation complicating deep analysis of these genes.. We try to avoid the sterility of these lines by crossing with wild type, as male or female and this work is still in progress. Additionally *taw1*, *g1l1* and *g1l2* triple mutant lines are growing in our greenhouse but at the time of writing this thesis (June 2017) they are not flowering yet. At this point we are not yet able to demonstrate if there is a genetic relationship between *TAW1* and *G1L1* and/or *G1L2*. Moreover further analysis is needed on meristem size of the mutant plants and expression analysis of meristem marker genes in these mutants to understand better the function of these ALOG genes.

4.3 FINAL CONCLUSION

We have shown that transcriptome analysis of morphologically distinct meristems using laser microdissection and RNA-Seq, is a reliable approach to study the transcriptional dynamic at early stages of rice panicle development. We also combined new technologies like laser microdissection and genome editing through CRISPR/Cas9 to evaluate the function of genes putatively involved in inflorescence development. Based on the data obtained so far this strategy seems to be very promising.

I believe that the data that we presented here will accelerate research on the molecular mechanisms controlling panicle development and will pave the way toward achieving a sustainable increase in grain yield required to meet with the challenge of producing enough food for the rapidly growing population.

REFERENCES

Ashikari M, Matsuoka M. 2006. Identification, isolation and pyramiding of quantitative trait loci for rice breeding. Trends Plant Sci. 11(7):344-50.

Cloonan N, Forrest ARR, Kolle G, Gardiner BBA, Faulkner GJ et al., Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat Meth. 2008;5(7):613–9.

Fan M, Shen J, Yuan L, Jiang R, Chen X, Davies WJ, Zhang F. 2011. Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China. J. Exp. Bot. 63(1):13-24.

Foley JA, Ramankutty N, et al., 2011. Solutions for a cultivated planet. Nature. 478(7369):337-42.

Guanghuai Jiang, Yanghai Xiang, Jiying Zhao, Dedong Y et al., 2014. Regulation of Inflorescence Branch Development in Rice Through a Novel Pathway Involving the Pentatricopeptide Repeat Protein sped1-D. Genetics. 197(4): 1395–1407.

Ikeda K, Sunohara H and Nagato Y. 2004. Developmental course of inflorescence and spikelet in rice. *Breed. Sci.* 54, 147–156.

Ikeda T, Tanaka W, Mikami M, Endo M and Hirano HY. 2016. Generation of artificial drooping leaf mutants by CRISPR-Cas9 technology in rice. Genes Genet. Syst. 90, 231–235.

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, Yasuno N, Sato Y, Yoda M, Yamazaki R et al., 2012. Inflorescence meristem identity in rice is specified by overlapping functions of three AP1/FUL-like MADS box genes and PAP2, a SEPALLATA MADS box gene. Plant Cell 24:1848–59.

Kumar V and Jain M. 2015. The CRISPR-Cas system for plant genome editing: advances and opportunities. J Exp Bot. 66(1):47-57.

Li J, Meng X, Zong Y, Chen K, Zhang H, Liu J, Li J et al., 2016a. Gene replacements and insertions in rice by intron targeting using CRISPR–Cas9. Nature Plants, 2, 16139.

Li M, Li X, Zhou Z, Wu P, Fang M, Pan X, Lin Q. et al., 2016b. Reassessment of the four yield-related genes Gn1a, DEP1, GS3, and IPA1 in rice using a CRISPR/Cas9 system. Front. Plant Sci. 7, 377.

Li Q, Zhang D, Chen M, Liang W, Wei J, Qi Y and Yuan Z. 2016c. Development of japonica photo-sensitive genic male sterile rice lines by editing carbon starved anther using CRISPR/Cas9. J. Genet. Genomics, 43, 415–419.

Mantegazza O, Gregis V, Chiara M, Selva C, Leo G, Horner DS, Kater MM. 2014. Gene coexpression patterns during early development of the native Arabidopsis reproductive meristem: novel candidate developmental regulators and patterns of functional redundancy. Plant J. 79(5):861-77.

McCouch SR, Doerge RW. 1995. QTL mapping in rice. Trends Genet. 11(12):482-7.

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X et al., 2013. Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res 23: 1233–1236.

Miflin B. 2000. Crop improvement in the 21st century. J Exp. Bot. 51(342):1-8.

Osakabe Y, Watanabe T, Sugano SS, Ueta R, Ishihara R, Shinozaki K and Osakabe K. 2016. Optimization of CRISPR/Cas9 genome editing to modify abiotic stress responses in plants. Sci. Rep. 6, 26685.

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.

Sato Y, Antonio B, Namiki N, Motoyama R, Sugimoto K et al. 2011. Field transcriptome revealed critical developmental and physiological transitions involved in the expression of growth potential in japonica rice. BMC Plant Biol. 11:10.

Sergiu C, Diana G, Amin H, Alina-Andreea Z, Pierre C, and Ioana BN. 2017. CRISPR/Cas9: Transcending the Reality of Genome Editing. Mol Ther Nucleic Acids. 16; 7: 211–222.

Sharma R, Agarwal P, Ray S, Deveshwar P, Sharma P, Sharma N, et al., 2012. Expression dynamics of metabolic and regulatory component across stages of panicle and seed development in *indica* rice. Funct. Integr. Genomics 12, 229–248.

Sontheimer EJ, Barrangou R. 2015. The bacterial origins of the CRISPR genome-editing revolution. Hum Gene Ther. 26(7):413–24.

Sun Y, Zhang X, Wu C, He Y, Ma Y, Hou H, Guo X. et al., 2016. Engineering herbicide-resistant rice plants through CRISPR/Cas9-mediated homologous recombination of acetolactate synthase. Mol. Plant, 9, 628–631.

Takeda S, Matsuoka M. 2008. Genetic approaches to crop improvement: responding to environmental and population changes. Nat Rev Genet. 9(6):444-57.

Wang F, Wang C, Liu P, Lei C, Hao W, Gao Y, Liu YG. et al. 2016. Enhanced rice blast resistance by CRISPR/Cas9-targeted mutagenesis of the ERF transcription factor gene OsERF922. PLoS ONE, 11, e0154027.

Wang Z, Gerstein M, Snyder M (2009) RNA-seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57–63.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Review Genetics. 10: 57-63.

Xie K, Minkenberg B, Yang Y. 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci U S A. 112(11):3570-5.

Xu R, Li H, Qin R, Wang L, Li L, Wei P and Yang J. 2014. Gene targeting using the Agrobacterium tumefaciens-mediated CRISPR-Cas system in rice. Rice, 7, 1.

Xu R, Yang Y, Qin R, Li H, Qiu C, Li L, Wei P et al., 2016. Rapid improvement of grain weight via highly efficient CRISPR/Cas9-mediated multiplex genome editing in rice. J. Genet. Genomics, 43, 529–532.

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, et al., 2013. TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. Proc. Natl. Acad. Sci. USA 110:767–72.

Zhou J, Peng Z, Long J, Sosso D, Liu B, Eom JS, Huang S.et al., 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. Plant J. 82, 632–643.

5 ACKNOWLEDGMENTS

Thanks to Almighty Allah for the countless mercy He took on me; the help, the power, the courage and whatever He bestowed upon me for the completion of my PhD and making it easier for me. I acknowledge my deepest sense of gratitude for whom this world has been made, the Holy Prophet (PBUH) who had spread the light of learning in the world and forever city of knowledge for the whole humanity.

No words could ever adequately express my feelings to my affectionate parents, my loving wife and my sweet kids and all other family members, who provided me all what I needed and remembered me in their prayers.

I feel honor and obliged to express my deep gratitude to my respected and honorable supervisor Professor **Martin M. Kater**, whose wholehearted advice, valuable suggestions and support embodies the inculcation of spirit of hard work and their sincere guidance and encouragement will always serve as beacon of light through the course of my life. I would like to extend my heartiest thanks to **Veronica Gregis**, **Michela Osnato** and **Hélène Adam**, for their lively supervision, continuous guidance and every possible cooperation throughout the PhD pursuit.

With deep sense of gratitude and benevolence, I offer words of thanks to Prof. **Simona Masiero**, **Tom Harrop**, **Edward Kiegle**, PhD students (**Elia Lacchini**, **Emanuela Franchini** and **Muneeb Kiani**) and my master student **Cecilia Bertone** for providing all privileges and help during my research and thesis work. I specially acknowledge all the MADS group members and colleagues who cooperated and accompanied me during this period. Thanks dears for filling my life with everlasting and happiest moments and being with me so far.

I would like to thank the **Agriculture University Peshawar**, in particular the rector, **Prof. Zahoor Ahmad Swati** for giving me the opportunity and providing all financial and social support to pursue my PhD at abroad. I would also like to thank **Prof. Safdar Hussain Shah**, **Prof. Iqbal Munir** and **Nasir Jamal** for helping to promote funding through out my PhD.

Special heartiest thanks to everyone who stay with me as a support. May Allah bless them all. (Ameen)

ISRAR UD DIN