2.0 LITERATURE REVIEW

2.1 Rice

2.1.1 Cultivated Rice

Rice (*Oryza sativa* L.) is one of the most important staple foods in the world as over half of the world's population depends on rice in their diet. The genus *Oryza* L. belongs under the tribe *Oryzeae*, subfamily *Oryzoideae* in the family *Poaceae* (*Gramineae*) (Lu, 1999). This genus is divided into four species complexes, including the *Oryza sativa*, *Oryza officialis*, *Oryza ridelyi* and *Oryza granulata* species complexes (Sweeney and McCouch, 2007).

The *Oryza sativa* complex is divided into cultivated species and wild species. *Oryza sativa* L. and *Oryza glaberrima* Steud. are two cultivated species, whereas *Oryza rufipogon*, *Oryza nivara*, *Oryza barthii*, *Oryza longistaminata*, *Oryza meridionalis* and *Oryza glumaepatula* are wild species (Lu, 1999; Vaughan *et al*., 2003; Sweeney and McCouch, 2007). Both cultivated species and wild species are diploid and share the AA genome (2n=24) (Ge *et al*., 1999).

Cultivated rice *Oryza glaberrima* Steud. is cultivated in West African countries (Sweeney and McCouch, 2007). According to isozyme studies, simple sequence repeats (SSR) and single nucleotide polymorphism (SNP) data, African rice *Oryza glaberrima* Steud. was domesticated from *Oryza barthii* (Second, 1982; Semon

et al., 2005; Sweeney and McCouch, 2007). Recently, NERICA (New Rice for Africa) a new variety, which is progeny from an *Oryza sativa* and *Oryza glaberrima* cross, has become very popular among West African farmers (Jones *et al*., 1997; Gridley *et al*., 2002; Sweeney and McCouch, 2007).

Cultivated rice *Oryza sativa* L. is cultivated in Asia (Sweeney and McCouch, 2007). So far, six different groups or subpopulations have been identified among *Oryza sativa* varieties, including *indica*, *japonica*, *aus*, *aromatic*, *rayada* and *ashina*. Now, only the *indica* and *japonica* subspecies are distinguished as confirmed by restriction fragment length polymorphism (RFLP) studies (Wang and Tanksley, 1989; Sweeney and McCouch, 2007). The genetic divergences between the *indica* and *japonica* subspecies have been shown to be the products of separate domestication events from the wild type species (Cheng *et al*., 2003; Sweeney and McCouch, 2007).

2.1.2 Rice Domestication

Oryza rufipogon Griff. and *Oryza nivara* are two different wild species under the *Oryza sativa* complex (Sweeney and McCouch, 2007). *Oryza rufipogon* is a perennial species, photoperiod sensitive, largely cross-fertilized, and adapted to persistently wet habitats (Sharma *et al*., 2000; Sang and Ge, 2007). Whereas *Oryza nivara* is an annual species, photoperiod insensitive, predominantly self fertilized, and adapted to seasonally dry habitats (Sharma *et al*., 2000; Sang and Ge, 2007). Before *Oryza nivara* was recognized, *Oryza rufipogon* was identified as the wild ancestor of *Oryza*

sativa and widely used to accommodate the wild Asian A-genome taxonomical group (Sharma and Shastry 1965; Sang and Ge, 2007). After *Oryza nivara* was recognized, it may represent the most closely related ancestor of *Oryza sativa* compared with *Oryza rufipogon*. This is because *Oryza rufipogon* is a perennial species but *Oryza sativa* and *Oryza nivara* are annual species (Khush, 1997; Yamanaka *et al*., 2003; Sang and Ge, 2007). However, controversy has persisted on whether *Oryza nivara* is an ecotype or subspecies of *Oryza rufipogon* (Morishima 2001; Cheng *et al*., 2003; Vaughan *et al*., 2003; Sang and Ge, 2007; Zhu *et al*., 2007; Gao and Innan, 2008), and on whether *Oryza sativa* was domesticated once or more times from divergent wild populations (Sang and Ge, 2007; Gao and Innan, 2008).

Oryza sativa ssp. *indica* and *Oryza sativa* ssp. *japonica* are two major ssp. in Asia. *Oryza sativa* ssp. *indica* is usually found in the lowlands of tropical Asia; whereas, *Oryza sativa* ssp. *japonica* is typically found in the upland hills of southern China, Southeast Asia, and Indonesia, as well outside of Asia (Khush, 1997; Sang and Ge, 2007). Rice domestication is suggested to have begun since 9000 years ago within eastern India, Indochina and portions of southern China, where *Oryza rufipogon* grows across this region (Khush, 1997; Sang and Ge, 2007). So far, two major hypotheses have been proposed to explain rice domestication, suggesting either a single or multiple origin of cultivated rice and geographically independent or nonindependent rice domestications from its wild ancestor (Sang and Ge, 2007; Gao

and Innan, 2008). However there is still a lack of strong evidence to support any one hypotheses of rice domestication.

2.1.3 Rice Genome Sequencing Projects

Rice was selected as a model monocotyledon plant because it has the smallest genome of the cereals at 389 Mb, vast germplasm collection, enormous repertoire of molecular genetic resources, and an efficient transformation system (Paterson *et al*., 2005; Sasaki and Sederof, 2003; IRGSP, 2005). High quality *Oryza sativa* ssp. *japonica* Nipponbare and *Oryza sativa* ssp. *indica* variety 93–11 genome sequences have been completed by four independent research teams at the Beijing Institute of Genomics, Syngenta, International Rice Genome Sequencing Project (IRGSP), and Monsanto (Sasaki *et al*., 2002; Yu *et al*., 2002; IRGSP, 2005; Yu *et al*., 2005). After the completion of the rice genome sequencing, the Michigan State University Rice Genome Annotation Project (formerly hosted by The Institute for Genomic Research), the Rice Information System (Rise), Rice Annotation Project Database (RAP-DB), and the Rice Genome Annotation Project of The Institute for Genomic Research (TIGR) are the main rice genome annotation resources to support rice biology research (Zhao *et al*., 2004; Ouyang *et al*., 2007; Rice Annotation Project, 2008; Tanaka *et al*., 2008).

Oryza sativa ssp. *indica* and *japonica* were chosen for full rice genome sequencing because they are major Asian cultivated varieties but they show distinct

divergence from sequence variations to phenotypic changes (Oka and Chang, 1962; Cheng and Lu, 1984; Han and Zhang, 2008). The comparative genomics between these two subspecies were carried out through the grass database Gramene, which detects polymorphisms between them and assists understanding of the origin, domestication, genetic map, marker, and quantitative trait loci (QTL; Han and Zhang, 2008; Liang *et al*., 2008). Besides comparative genomic analysis within the *Oryza* genus, genome sequences of *Oryza* can be used to compare with other closely related grass species, including sorghum and maize, in order to identify important conserved functional units and regulatory elements (Han and Zhang, 2008; Liang *et al*., 2008).

The identification of the function of genes and the understanding of the interactions between them has become another big challenge after completing the full genome sequencing of two major Asian cultivated varieties (Han and Zhang, 2008). The generation of a large mutant library and the isolation of full-length cDNAs are methods used for studying the functional genomics of rice (Han and Zhang, 2008). Insertion mutation, and chemical and physical mutagenesis are major mutational analysis techniques that can be performed to study gene function. These have been applied to the construction of mutant libraries (Jeon *et al*., 2000; Agrawal *et al*., 2001; Wu *et al*., 2003; Hirochika *et al*., 2004; Sakamoto *et al*., 2004; Wu *et al*., 2005; Till *et al*., 2007; Zhang *et al*., 2007).

Insertional mutagenesis and the Tos17 disruption system are two major insertion mutation methods (Jeon *et al*., 2000; Hirochika *et al*., 2004; Zhang *et al*., 2007). For insertional mutagenesis, T-DNA is used to disrupt the gene function (Zhang *et al*., 2007). About 13,804 T-DNA flanking sequence tag lines were isolated from a T-DNA insertion library (Jeon *et al*., 2000; Zhang *et al*., 2007). However, T-DNA inserts were non-randomly distributed on each chromosome and the frequency of T-DNA integration was limited based on the proportion to chromosome size (Zhang *et al*., 2007). Meanwhile, the Tos17 disruption system is another more efficient method, whereby more than 50,000 disruption lines of *japonica* were identified (Miyao *et al*., 2007; Han and Zhang, 2008). Recently, the combination of the Tos17 insertion method in the rice T-DNA mutant library was conducted in order to enhance functional analysis of rice genes (Piffanelli *et al*., 2007; Han and Zhang, 2008).

Chemical and physical mutagenesis was another mutational analysis method to study gene function (Wu *et al*., 2005; Till *et al*., 2007). Normally, chemical agents such as ethyl methanesulfonate (EMS), methyl nitrosourea and diepoxybutane, or physical methods like fast-neutrons, gamma ray, and ion beam irradiation, cause high density mutations (Wu *et al*., 2005; Till *et al*., 2007). The targeting induced local lesions in genomes (TILLING) method was used to screen mutation populations with point mutations, whereas, PCR-based screening methods were used to detect large deletions (Wu *et al*., 2005; Till *et al*., 2007). A total of 66,891 mutant lines were generated by chemical and physical mutagenesis in the IR64 mutant collection of M_4

generation (Krishnan *et al*., 2009). According to Hirochika *et al*. (2004), 180,698 to 460,000 insertions flanking sequence tags estimated were needed to tag every gene in rice. However, only around 200,000 insertion flanking sequence tags were reported by using different mutational analyses (Krishnan *et al*., 2009). Thus, more cost effective methods are needed to generate more insertion flanking sequence tags (Krishnan *et al*., 2009).

Full-length cDNAs provide important information for understanding the function of genes, especially at the transcriptional level (Nishiyama *et al*., 2003; Han and Zhang, 2008). More than 32,000 *japonica* full-length cDNAs were isolated and identified (The Rice Full-Length cDNA Consortium, 2003; Han and Zhang, 2008). Besides that, more than 20,000 full-length cDNAs and 40,000 5' EST were isolated from two *indica* cDNA libraries (Xie *et al*., 2005; Liu *et al*., 2007; Han and Zhang, 2008). Beside two major Asian cultivated varieties, 1,888 full-length cDNAs of wild rice *Oryza rufipogon* W1943 have been collected (Lu *et al*., 2008). Full-length cDNAs of *Oryza* genus not only provided important resources for comparative analysis among rice varieties, moreover it can be used for comparison between two model monocotyledon and dicotyledon species, *Oryza sativa* and *Arabidopsis thaliana* (Liu *et al*., 2007; Lu *et al*., 2008; Tanaka *et al*., 2009). Comparisons between full-length cDNAs of *Oryza sativa* and *Arabidopsis thaliana* were used to study the molecular evolution of transcription start sites, which play important roles in

tissue-specific gene expression and functional variation (Landry *et al*., 2003; Iida and Go, 2006; Tanaka *et al*., 2009).

2.2 Quantitative Trait Loci (QTLs)

The quantitative trait locus (QTL) is a region within a genome that contains genes associated with a particular quantitative trait, including yield, quality and some forms of disease resistance (Collard *et al*., 2005). Conventional phenotypic evaluation for identification of QTLs is complicated. Therefore, the development of molecular markers such as Restriction Fragment Length Polymorphisms was applied to analyze QTL regions during the 1980s (Collard *et al*., 2005; Swamy and Sarla, 2008).

DNA markers play important roles in agriculture for the construction of linkage maps (Collard *et al*., 2005). Examination of chromosomal regions based on linkage maps can help in the identification of genes controlling simple traits and quantitative traits (Mohan *et al*., 1997; Collard *et al*., 2005). The process of constructing linkage maps and conducting QTL analysis is known as QTL mapping (McCouch and Doerge, 1995; Mohan *et al*., 1997; Collard *et al*., 2005). The development of marker-assisted selection (MAS) was used as a molecular tool for plant breeding (Ribaut and Hoisington, 1998; Collard *et al*., 2005). MAS was used to identify the agronomically desirable and undesirable loci in each species to assist in phenotypic selection (Collard *et al*., 2005; Swamy and Sarla, 2008). Once desired QTLs were identified, MAS was used in developing isogenic lines (Collard *et al*., 2005; Swamy and Sarla,

2008). This gives a better understanding of specific effects of QTL alleles on each species. DNA marker technology is widely used for crop improvement in wheat, barley, maize, bean, capsicum and soybean (Collard *et al*., 2005; Ulukan, 2009).

2.3.1 Yield-Related QTL in Rice

In general, wild species have smaller fruit, produce fewer seed and have many undesirable traits compared to cultivars (Sang and Ge, 2007; Swamy and Sarla, 2008). Basically wild species face many challenges to obtain interspecies hybrids and thus may take a longer time for introgression of useful traits in to many crops (Swamy and Sarla, 2008). However, wild species are still selected as donors for enhancing yield in conventional yield improvement programs (Swamy and Sarla, 2008). This is because wild species have large proportion of genetic variations to address the problem of gene pool narrowing in modern rice cultivars (Xiao *et al*., 1998). Besides that, molecular mapping studies have shown that wild species can contribute genes to improving yield in yield-enhancing QTL regions (Xiao *et al*., 1998; Swamy and Sarla, 2008).

Rice yield is a complex trait and shows a continuous phenotypic variation governed by QTL (Swamy and Sarla, 2008). It is determined by a combination of traits such as yield per plot, yield per plant, yield per panicle, tiller per plant, panicles per plant, spikelet per plant, spikelet per panicle, grains per plant, grains per panicle, etc. So far, nine yield-related traits have been identified in three wild species, *Oryza*

rufipogon (2n=24,AA), *Oryza glumaepatula* (2n=24, AA) and *Oryza grandiglumis* (2n=22, CCDD). Table 2.1 shows the chromosome-wise distribution of yield-related QTL from wild species of rice. Distribution of yield-related QTL on rice chromosomes is not random (Swamy and Sarla, 2008). Many yield-related QTLs have been identified on chromosomes 1, 2, 3 and 4.

Trait	**QTLs on chromosome**												**Total QTLs**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Yield per plot	1	1	-	-	-	-	-	5	-	-	-	-	7
Yield per plant	6	7	4	4	1	3	2	-	2	-	3	-	32
Yield per panicle	3	3	1	6	4	1	2	-	-	-	3	-	23
Tiller per plant	1	1	-	1	1	-	-	-	-	-	-	-	4
Panicles per plant	4	4	1	-	-	-	1	-	-	-	2	-	12
Spikelet per plant	1	-	-	-	-	-	-	-	-	-	-	-	1
Spikelet per panicle	7	4	2	2	-	1	1	2	2	-	-	1	21
Grains per plant	5	4	-	1	3	1	-	-	-	-	1	1	20
Grains per panicle	3	3	2	2	1	1	-	1	1	-	1	3	17

Table 2.1: Chromosome-wise distributions of yield-related traits of QTL from *Oryza rufipogon*, *Oryza glumaepatula* and *Oryza grandiglumis*. (Data from Xiao *et al*., 1998; Moncada *et al*., 2001; Brondani *et a*l., 2002; Thomson *et al*., 2003; Septiningsih *et al*., 2003a; Marri *et al*., 2005; Yoon *et al*., 2006; Swamy and Sarla, 2008).

Oryza rufipogon is one of the wild species frequently used as a donor for enhancing yield in conventional yield improvement programs (Xiao *et al*., 1998; Sabu *et al*., 2006; McCouch *et al*., 2007). Although its phenotypes are agronomically undesirable compared to cultivar types, it has a large proportion of genetic variations for the *Oryza* genus which could be used to solve the problem of gene pool narrowing in modern rice cultivars (Xiao *et al*., 1998; Sabu *et al*., 2006; McCouch *et al*., 2007). So far, twenty-nine yield QTLs and twenty-one yield-enhancing QTLs have been mapped and identified on *Oryza rufipogon* (Xiao *et al*., 1998; Xiong *et al*., 1999; Moncada *et al*., 2001; Septiningsih *et al*., 2003a; Thomson *et al*., 2003; Marri *et al*., 2005; Swamy and Sarla, 2008). It was thought that wild species may have a single allele or more for yield improvement, which maybe masked (Sweeney and McCouch, 2007; Swamy and Sarla, 2008). Therefore, a 172 kb genomic DNA region of wild rice *Oryza rufipogon* (annual type, accession number IRGC 105491) around the RM5 locus associated with the yield-enhancing QTL *yld1.1* was sequenced and compared with sequence of *Oryza sativa* ssp. *japonica* and *Oryza sativa* ssp. *indica* (Song *et al*., 2009). Twelve out of fourteen predicted genes were matched with putative genes of known function in *Oryza sativa* ssp. *japonica and Oryza sativa* ssp. *indica* sequence (Song *et al*., 2009). Although the function of the predicted genes of *Oryza rufipogon* has yet to be determined, it was believed that the RM5 region plays an important role in evolution selection because it is highly conserved in coding, intronic and intergenic region when compared with *Oryza sativa* ssp. *japonica and Oryza sativa* ssp. *indica* sequences (Song *et al*., 2009).

2.3.2 Advance Backcross QTL

The advance backcross (AB) QTL strategy was developed to overcome the problem of introduction of undesirable traits from wild species into the elite background (Swamy and Sarla, 2008). This strategy maps yield-related QTLs for introgression into elite breeding material (Tanksley and Nelson, 1996; Swamy and Sarla, 2008). This strategy has many successful examples, especially in rice, regarding transfer of positively contributing alleles from inferior wild species to elite species cultivars (Xiao *et al*., 1998; McCouch *et al*., 2007).

In a Malaysian National Rice Breeding program, Malaysian rice breeders have also applied the AB-QTL strategy towards producing a high quality new variety. A Malaysian accession of wild rice *Oryza rufipogon*, IRGC105491 has been selected as the donor parent with *Oryza sativa* ssp. *indica* cv. MR219, a Malaysian elite rice variety as the recurrent parent (Sabu *el al*., 2006). Twenty-six top-performing BC_2F_2 families were identified according to grain yield per plant, culm length, tillers per plant, filled grains per panicle and 1000-grain weight (Sabu *et al*., 2006). They showed around 27 % higher yield per plant compared to the recurrent parent. The twenty-six top-performing BC_2F_2 families were each self crossed until BC_2F_7 in order to evaluate and validate yield-related QTLs in various environments. AB-QTL technology helps in enhancing productivity and quality of backcross families (Wickneswari pers. comm. 2010). Moreover, it permits efficient use of genetic diversity for breeding program.

2.3.3 Yield-Related Genes in Rice QTL

Yield is a complex trait and controlled by multiple genes (Swamy and Sarla, 2008). Although many QTLs responsible for regulating rice yield components have been identified, only a few functions of yield-related genes in rice have been reported to date (Swamy and Sarla, 2008; Zha *et al*., 2009). Eight *leucine-rich repeat receptor-like kinase* (*LRK*) genes have been identified from the QTL *qGY2-1* of Dongxiang wild rice (*Oryza rufipogon* Griff.), which was found to increase grain yield by 16 % (Zha *et al*., 2009). The *LRK1* gene is one of the eight *LRK* genes that regulate rice branch number by enhancing cellular proliferation (Zha *et al*., 2009). This gene was contained in the donor parent (Dongxiang wild rice) but absent from the recurrent parent genome (Guichao2, *Oryza sativa* ssp. *indica*; Zha *et al*., 2009). The over-expression of *LRK1* in *Oryza sativa* ssp. *indica* variety 9311 has resulted in a 27.09 % increase in total grain yield per plant (Zha *et al*., 2009).

The *grain number 1a* (*Gn1a*) gene was cloned and identified from yield-related QTL in backcross inbred line from a *indica* rice variety, Habataki, and a *japonica* variety, Koshihikari, which contributes directly to grain productivity (Ashikari *et al*., 2005). The gene encodes cytokinin oxidase or dehydrogenase 2 (OsCKX2), which is an enzyme that degrades the phytohormone cytokinin but it promotes cell division (Ashikari *et al*., 2005; Sakamoto, 2006). The reduced expression of OsCKX2 in rice enhanced grain yield (Ashikari *et al*., 2005). This is because more cytokinin accumulated in inflorescence meristems were helped to increase the number of

reproductive organs (Ashikari *et al*., 2005).

The *grain weight 2* (*GW2*) gene was cloned and characterized from a yield-related QTL in backcross inbred line from a *indica* rice variety, Fengaizhan-1, and a *japonica* variety, WY3, which controls rice grain width and weight (Song *et al*., 2007). The gene encodes a RING-type protein with E3 ubiquitin ligase activity (Song *et al*., 2007). This protein plays an important role in the ubiquitin-proteosome pathway and is a negative regulator of cell division in rice. Loss of *GW2* function will result in greater grain width, weight and yield (Song *et al*., 2007). This is because the spikelet hull has become wider and the grain milk filling rate was accelerated (Song *et al*., 2007). Understanding the mechanisms of these yield-related genes involved in cereal crop yield will help improve agricultural traits in future crop breeding programmes (Swamy and Sarla, 2008; Zha *et al*., 2009).

2.3 Receptor-Like Kinase (RLK)/ *Pelle* Protein Kinase Family

The Receptor protein kinase was first reported in animals and it plays important roles in cellular signaling processes (Zhang, 1998). Meanwhile, Maize (*Zea mays*) was the first plant reported to possess receptor protein kinases, better known as plant receptor-like kinases (RLK), which have sequence homology and structures similar to that of animal receptor kinases (Walker and Zhang, 1990; Zhang, 1998). Plant RLK's are grouped into the RLK/*Pelle* family because the plant RLK kinase domains belong to the *Pelle* kinase of RLK/*Pelle* family in *Drosophila melanogaster* (Shiu and

Bleecker, 2001a).

The plant RLK/*Pelle* family can be categorised into three major groups, namely transmembrane receptor kinase, receptor-like cytoplasmic kinase (RLCK) and receptor-like protein (RLP) as shown in Figure 2.1 and 2.2 (Diévart and Clark, 2004; Afzal *et al*., 2008). Transmembrane receptor kinase contains an extracellular domain connected to a cytoplasmic serine/theronine (Ser/Ther) protein kinase via a single pass transmembrane helix (Shiu and Bleecker, 2001a; Diévart and Clark, 2004). However, RLCK only lack an extracellular domain, whereas, RLP only lack an intracelluar cytoplasmic kinase domain (Shiu and Bleecker, 2001a; Diévart and Clark, 2004; Shiu *et al*., 2004; Morillo and Tax, 2006). Transmembrane receptor kinase represents the largest group of plant RLK/*Pelle* family members (Shiu and Bleecker, 2003). According to conserved residues in the extracellular kinase domain, another three groups can be classified in the transmembrane receptor kinase group. Characterization of the rice kinome was based on the presence or absence of a conserved lysine (K) kinase subdomain II, a conserved arginine (R) in kinase subdomain VIb, and aspartic acid (D) in kinase subdomain VII (Dardick and Ronald, 2006). Normally, the activation loop of kinase subdomain VIb contains a conserved arginine (R) and an invariant aspartate (D), and is then referred to as an RD-kinase class (Dardick and Ronald, 2006; Castells and Casacuberta, 2007). However, if the activation loop of the kinase subdomain VII lacks only a conserved arginine (R), it is then referred to as a non-RD-kinase class. If the kinase subdomain lacks both the conserved arginine (R)

and invariant aspartate (D), it is referred to as a RD-minus kinase (Dardick and Ronald, 2006). According to Dardick and Ronald (2006), the RD-kinase class RLK is involved in development whereas the non-RD class RLK is involved in innate immunity. However, the general function of the RD-minus kinase has not yet been understood (Dardick and Ronald, 2006; Afzal *et al*., 2008).

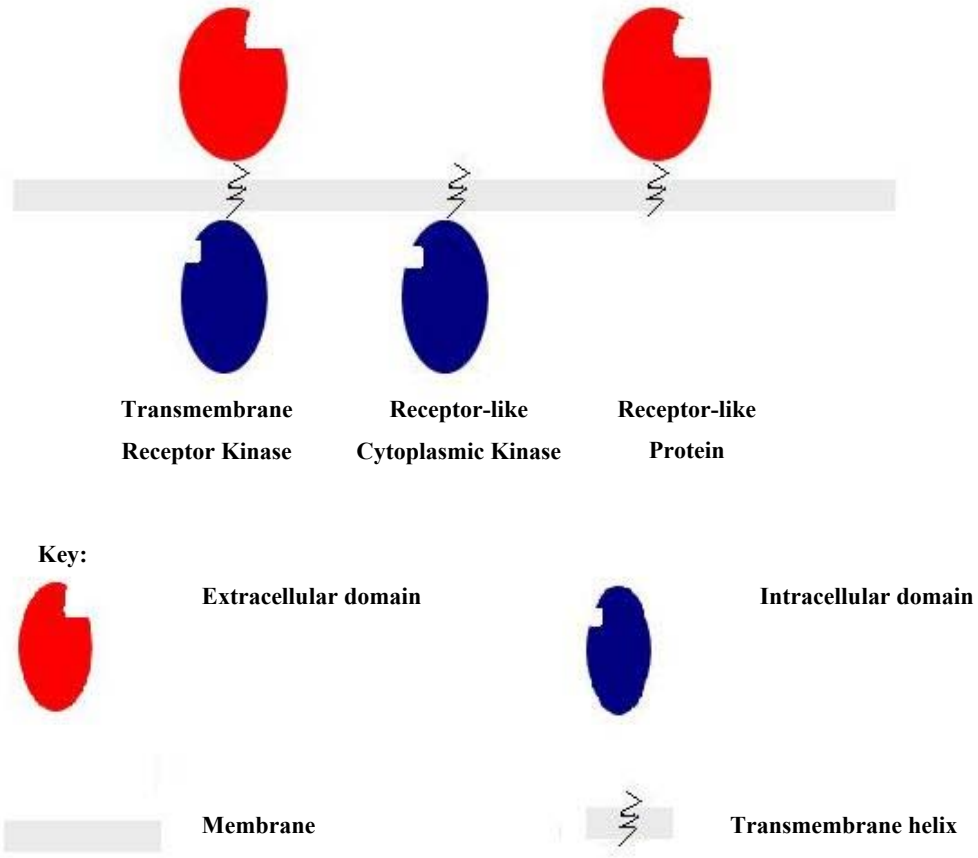

Figure 2.1: The three general RLK class structures in the plant RLK/*Pelle* family, the transmembrane receptor kinase, the receptor-like cytoplasmic kinase (RLCK) and the receptor-like protein (RLP).

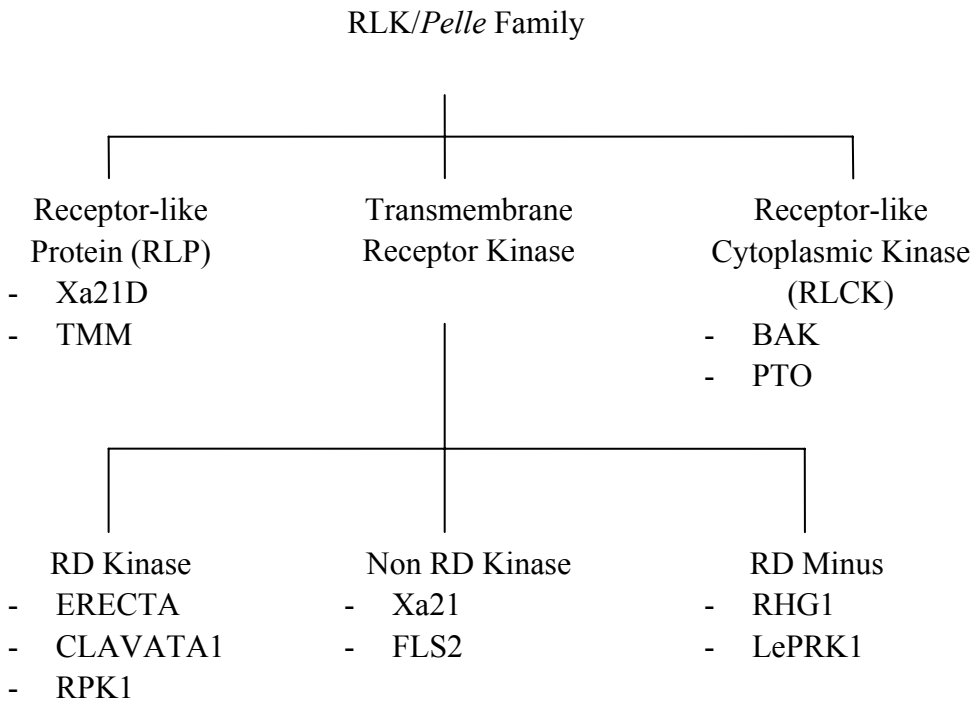

Figure 2.2: The general classification scheme for plant RLK/*Pelle* family (Modified from Afzal *et al*., 2008).

The plants RLKs possess various types of extracellular domain. Therefore, about 15 types of plant RLKs were identified based on extracellular domain only, i.e. C-type lectin-like, CrRLK1-like, domain of unknown function 26 (DUF26)-like, extensin-like, legume (L)-lectin-like, LRK10-like, leucine-rich repeat (LRR)-like, lysine motif (LysM)-like, PERK-like, RKF3-like, thaumatin-like, URK1-like, self-incompatibility locus (S)-domain-like, tumor-necrosis factor receptor (TNFR)-like, and cell wall associated receptor kinase (WAK)-like (Shiu and Bleecker, 2001a; Diévart and Clark, 2004; Shiu *et al*., 2004). The LRR-like type extracellular domains represent the largest group of plant RLK/*Pelle* family members (Shiu and

Bleecker, 2002; Shiu *et al*., 2004). The LRR-like type can be further classified into 13 subfamilies (LRR-I to LRR-XIII) based on their LRR tandem repeat (Shiu and Bleecker, 2002; Shiu *et al*., 2004). Animal LRR-like receptor kinases do not contain cytoplasmic kinase unlike plant RLK (Diévart and Clark, 2004). Animal LRR receptor kinases similar to plant LRRs are involved in growth and development (Diévart and Clark, 2004). LRR-like type of plant RLK and animal receptor kinases may share similar mechanisms of signal perception through cell surface receptors because their LRR motifs are alike (Diévart and Clark, 2004).

The functions of plant RLK/*Pelle* family members can be categorized into two broad categories. The first category of plant RLK is involved in pathogen response and defence systems (Shiu *et al*., 2004; Morillo and Tax, 2006). Disease resistance of plants is a complex response. Plant RLK*s* are involved in recognition of non-self signals from pathogens and activation of defence systems (Shiu *et al*., 2004; Morillo and Tax, 2006). For example, the member of Flagellin Sensitive2 (FLS2) is involved in recognition of flagellin in the *Arabidopsis thaliana* response to pathogen (Gomez-Gomez *et al*., 2000), while Xa21 of rice bestows resistance to bacterial blight pathogen *Xanthomonas oryzae pv oryzae* (Song *et al*., 1995), and Systemin Receptor SR160 is involved in response to the polypeptide hormone systemin in tomato (Scheer *et al*., 2002; Shiu *et al*., 2004; Morillo and Tax, 2006).

The second functional category of plant RLK is involved in cell growth and developmental systems. For example, the CLAVATA1 Receptor Kinase (CLV1) protein acts as a regulator of cell fate, which includes in the fate of stems cells in *Arabidopsis thaliana* meristems (Clark *et al*., 1997: Jeong *et al*., 1999; Fletcher *et al*., 1999; Morillo and Tax, 2006). S-locus receptor kinase (SRK) are involved in recognition during development, which recognises pollen signal in self-incompatibility mechanism in *Brassica*, *Arabidopsis thaliana* and rice (Kachroo *et al*., 2002; Morillo and Tax, 2006). Receptor-like Protein Kinase 1 (RPK1) are involved in regulation of growth process in response of hormone, which plays important role in transcription induction on abscisic acid (ABA) pathway in *Arabidopsis thaliana* (Osakabe *et al*., 2005; Morillo and Tax, 2006). Although more than 1,027 and 1,429 RLK/*Pelle* family members were identified in *Arabidopsis thaliana* and rice respectively, the understanding of their roles are still limited (Ding *et al*., 2009).

2.4 RNA Silencing

RNA silencing, also known as RNA interference (RNAi), is a phenomenon in which small double-stranded RNA (dsRNA) can induce efficient sequence-specific silencing of gene expression (Fire *et al*., 1998). About 20-30 nucleotide (nt) small RNA molecules are generated from longer dsRNA precursors by RNaseIII-like Dicer enzymes (Hannon, 2002; Vaucheret, 2006). Small RNAs are key regulators of gene activities, genetic elements and antiviral defence (Baulcombe, 2004; Brodersen and

Voinnet, 2006; Vaucheret, 2006; Mlotshwa *et al*., 2008b). The small RNA silencing molecules can be categorized into two major groups, including microRNA (miRNA) and small interfering RNA (siRNA; Vaucheret, 2006).

miRNAs were first reported to play an important function in regulating development timing of *C. elegans* (Lee *et al*., 1993; Reinhart *et al*., 2000). In plants, mature miRNAs have sizes ranging from 20-30 nt and play important roles as key posttranscriptional regulators, particularly during development (Groβhans and Filipowicz, 2008). Meanwhile, siRNAs were first reported in transgenic plants, expressing a transgene homologous to an endogenous gene (Napoli *et al*., 1990; van der Krol *et al*., 1990). The expression of the endogenous gene was inhibited, and the phenomenon was called post-transcriptional gene silencing (PTGS) (Napoli *et al*., 1990; van der Krol *et al*., 1990). siRNAs also were discovered during virus infection, where dsRNA was generated through plant DNA viruses and plant ssRNA viruses (Mlotshwa *et al*., 2008b). In addition to siRNA's from transgenic or viral origin the plant genome also encodes endogenous siRNA. Endogenous siRNAs can be categorized into three classes, including RDR2-dependent siRNAs, trans-acting siRNAs (ta-siRNAs), and natural antisense siRNAs (nat-siRNAs).

2.4.1 The Biogenesis and Mechanism of siRNA and miRNA

Several key conserved protein families have been identified and are involved in plant core RNA silencing machinery, including DICER (*DCR*) or DICER-LIKE

(*DCL*), ARGONAUTE (*AGO*) and RNA-DEPENDENT RNA POLYMERASE (*RDR*; Baulcombe, 2004; Chapman and Carrington, 2007). About four *DCL*s, ten *AGO*s and six *RDR*s genes were identified in *Arabidopsis thaliana*, while rice was found to contain eight *DCL*s, nineteen *AGO*s and five *RDR*s genes (Schauer *et al*., 2002; Morel *et al*., 2002; Yu *et al*., 2003; Kapoor *et al*., 2008). During RNAi, a selected strand of a small interfering RNA molecule is incorporated into an effector complex, known as an RNA-induced silencing complex (RISC) (Martinez *et al*., 2002). RISC contains an AGO protein which provides small RNA binding domains for small RNA to bind into an effector complex for cleavage (Tomari and Zamore, 2005; Tolia and Joshua-Tor, 2007). The cleaved target RNAs are produced by the ribonuclease-H activity of *AGO* (Tomari and Zamore, 2005; Tolia and Joshua-Tor, 2007).

RNA polymerase II usually transcribes the biogenesis of plant miRNA independently and then pri-miRNAs are generated (Ramachandran and Chen, 2008). Small 21 nt RNAs are generated from the pri-miRNA by the RNase III *DCL1* to liberate miRNA/miRNA* duplexes (miRNA star is denoted as the opposite strand to the miRNA; Jones-Rhoades *et al*., 2006). The miRNA is then selectively incorporated into the RISC complex to initiate degradation or repression (Baumberger and Baulcombe, 2005; Qi *et al*., 2005).

RDR2-dependent siRNAs, ta-siRNAs and nat-siRNAs are the best understood plant endogenous siRNA. In plants, RDR2-dependent siRNAs represents the largest

siRNA group (Ramachandran and Chen, 2008). *RDR2*-dependent siRNAs are generated by dsRNA depend *RDR2* (Ramachandran and Chen, 2008). DCL3 protein is used to process dsRNA into 21-22 nt siRNA (Zheng *et al*., 2006; Vaucheret, 2008). RISC complex containing AGO4 or AGO6 proteins are required for the biogenesis of RDR2-dependent siRNAs for DNA methylation and histone methylation (Zilberman *et al*., 2003; Xie *et al*., 2004; Qi *et al*., 2005; Zheng *et al*., 2006; Vaucheret, 2008).

Ta-siRNA is involved in regulation of target genes originating from different loci (Vaucheret, 2006). The biogenesis of ta-siRNAs is dependent on the miRNA pathway. The cleavage of ta-siRNA precursor RNAs are generated from specific miRNAs, i.e. *TAS* family genes from *Arabidopsis thaliana* (Allen *et al*., 2005; Yoshikawa, *et al*., 2005). The 21-nt long ta-siRNAs are generated by *DCL4* depending on *RDR6* (Mlotshwa *et al*., 2008). Whilst, 24 nt ta-siRNAs are generated by *DCL3*. In *DCL3* or *DCL4* absence, 22-nt ta-siRNAs can be generated by *DCL2* depending on *RDR2*- and *RDR6*-derived siRNAs (Allen *et al*., 2005; Yoshikawa, *et al*., 2005).

Nat-siRNAs are generated from overlapping bidirectional transcripts from heterogeneous populations of plant small RNAs (Borsani *et al*., 2005; Katiyar-Agarwal *et al*., 2006). The 22-nt nat-siRNAs are produced by *DCL1* and 24-nt nat-siRNAs also can be generated *DCL2* (Borsani *et al*., 2005; Katiyar-Agarwal *et al*., 2006).

2.4.2 RNAi Technology in Plant Systems

RNAi technology has been widely applied to elucidate plant gene function. The dsRNA can be generated exogenously by a transgene encoding a short hairpin RNA. A plant expression vector is used to induce the RNA-mediated gene silencing pathway (Baulcombe, 2004; Chapman and Carrington, 2007). Transient RNAi systems and stable RNAi systems are well established and widely used to study function of plant genes. Transient RNAi systems are easy and rapid but are only suitable for analyzing gene function at cellular level; whereas, stable RNAi systems have a broad range of suitable hosts and are heritable to the next generation, so can be used for characterization of gene function in whole plants (Sato, 2005; Karimi *et al*., 2007).

There are three methods based on transient RNAi systems reported, i.e. virus-amplicon via *Agrobacterium*, introduction of artificially synthesized dsRNA and virus-induced gene silencing (VIGS; Waterhouse and Helliwell, 2003; Sato, 2005). Among them, the VIGS vector system is the most efficient system but is only limited to a few host plant species (Sato, 2005). Therefore, the method of introduction of artificially synthesized dsRNA or introduction of a dsRNA-expression vector via *Agrobacterium* infection was developed (Sato, 2005). This method would be more suitable for application to wider ranges of plant species (Karimi *et al*., 2007). The pANDA vector (Miki and Shimamoto, 2004) and pIPKb007 vector (Himmelbach *et al*., 2007) are well-established examples of stable RNAi systems using Gateway

technology, which have been used for *Agrobacterium*-mediated plant transformation to study gene function (Karimi *et al*., 2007; Isshiki and Kodama, 2010). The stable RNAi system is heritable and able to give additional amplification of silencing via transitivity (Baulcombe, 2007; Moissiard *et al*., 2007). RNA silencing via transitivity occurs when the initial pool of small RNAs (or termed as primary siRNA) are directed against one region of a transcript depending on *RDR* protein (Baulcombe, 2007; Moissiard *et al*., 2007). The propagation of dsRNAs that target the mRNA sequence in transitive silencing depends on *RDR*6 activity in *Arabidopsis thaliana* (Moissiard *et al*., 2007). The siRNAs generated from primary siRNA are termed as secondary siRNAs (Baulcombe, 2007; Moissiard *et al*., 2007). So far, stable RNAi based technology has been used for the improvement of crop traits, viral resistance and the control of plant-feeding pests (Isshiki and Kodama, 2010).