# COMPUTATIONAL IDENTIFICATION, PHYLOGENETIC AND SYNTENY ANALYSIS OF RECEPTOR-LIKE KINASES “RLK” AND RECEPTOR-LIKE PROTEINS “RLP” IN LEGUMES

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Daniel Restrepo Montoya

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Program:
Genomics and Bioinformatics

October 2018

Fargo, North Dakota

North Dakota State University
Graduate School

**Title**

# COMPUTATIONAL IDENTIFICATION, PHYLOGENETIC AND SYNTENY ANALYSIS OF RECEPTOR-LIKE KINASES “RLK” AND RECEPTOR-LIKE PROTEINS “RLP” IN LEGUMES

**By**

Daniel Restrepo Montoya

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University’s regulations and meets the accepted standards for the degree of

**DOCTOR OF PHILOSOPHY**

SUPERVISORY COMMITTEE:

Dr. Juan M. Osorno
Co-Chair

Dr. Phillip E. McClean
Co-Chair

Dr. Julia Bowsher

Dr. Robert Brueggemann

Approved:

11/16/2018
Date

Dr. Phillip E. McClean
Department Chair

# ABSTRACT

Legumes are considered the second most important family of crop plants after the grass family based on economic relevance. In recent years, the field of legume genomics has expanded due to advancements in high-throughput sequencing and genotyping technologies. To date, no published comparative genomic analysis explores receptor-like kinases “RLK” and receptor-like proteins “RLP” among legume genomes. Evaluating these RLK and RLP should provide a source of new information because extensive genetic and phenotypic studies have already discovered the diverse roles of RLK and RLP in cell development, disease resistance, and stress responses among other functions. This study demonstrates that a computational logical approach for classifying the RLK/RLP in legumes/non-legumes is statistically well supported and can be used in other plant species. The analysis of RLK/RLP of 7 legumes and 3 non-legume species evaluated suggests that about 2% are RLK and less than 1% of the proteins are RLP. The results suggest a dynamic evolution of RLK and RLP in the legume family. In fact, between 66% to 85% of RLK and 83% to 88% of RLP belong to orthologous clusters among the species evaluated. The remaining RLK and RLP proteins are classified as singletons. The ratio of the pairwise synteny blocks of RLK/RLP among legumes shows a 1:1 relationship. The exception is *G. max,* which shows an approximately 2:1 ratio due to its recent whole genome duplication (*G. max* vs. the other six legumes). The other legumes show evidence of a similar proportion of plasma membrane proteins among the legume pairwise synteny blocks.

# ACKNOWLEDGMENTS

I would like to thank Dr. Juan M. Osorno and Dr. Phil McClean for their patience and for teaching me about agriculture, breeding, genetics, genomics, bioinformatics, and much more. Dr. Robert Brueggeman for its invaluable tips, insights, and recommendations about plant pathology. Dr Julia Bowsher for its recommendations and support. The Fulbright scholarship. The Francisco Jose de Caldas – Colciencias scholarship, the North Dakota State University, the ILLiad interlibrary loan service system, and the NDSU dry breeding program. The Google Cloud platform service and the free extended trials. Dr. Mark Miller and Dr. Wayne Pfeiffer and the Cipres-Xsede platform. Dr. Luis F. Niño and Ing. Jonathan Narvaez and the Scientific Cluster at Universidad Nacional de Colombia. Dr. Henrik Nielsen at DTU. The NCBI staff - Dr. Peter S. Cooper, Dr. Wayne Matten. The OrthoMCL support team. The MCScanX community. Rebecca Kutty and Michael Stein for their editing suggestions, and finally the *NIX, python, and bash communities.

# DEDICATION

Simon, Maria, the everlasting gardener, and the regretful fisherman.

## TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

# LIST OF APPENDIX TABLES

# LIST OF APPENDIX FIGURES

# CHAPTER 1. INTRODUCTION

In recent years, the field of legume genomics has expanded due to advancements in high-throughput sequencing and genotyping technologies. Over the last eight years, reference genome sequences for many legume crops have been reported [1-7], allowing practitioners to identify structural variations and enhance the efficiency and resolution of genetic mapping and marker-trait association studies on a large scale [8]. To date, no published comparative genomic analysis explores receptor-like kinases "RLK" and receptor-like proteins "RLP" among the genomes of the following legumes: soybean (*Glycine max* (L.) Merrill) [1], common bean (*Phaseolus vulgaris* L.) [2], barrel medic (*Medicago truncatula* L.) [3], mungbean (*Vigna radiata* (L.) R. Wilczek) [4], cowpea (*Vigna unguiculata* L. Walp) [7], adzuki bean (*Vigna angularis*) [5], and pigeonpea (*Cajanus cajan* L.) [6].

Based on economic relevance, legumes are considered the second most important crop plant family after the grass family. Approximately 27% of world crop production is composed of grain legumes, providing 33% of human dietary protein, while pasture and forage legumes are fundamental for animal feed [9]. Recent advances in plant genomics have moved beyond model systems to focus on plant species of economic importance, such as *P. vulgaris* [2], and neglected species, such as *V. unguiculata* [10]. Recently, bioinformatics algorithm packages for sequence, assembly, functional annotation, and high-throughput gene expression and other comprehensive tools have been developed for a range of species [11], including legume-specific computational solutions, such as LegumeIP2.0 [12] and LegumeInfo.org [13]. A major technical challenge is to integrate these various tools to better understand genome structure and function [11]. This study attempted to do so using a comparative genomic approach to identify and computationally evaluate RLK and RLP among legumes/non-legumes. The results provide a source of new information for breeders and geneticist considering that extensive genetic and phenotypic studies have reported diverse functional roles of RLK and RLP (plasma membrane receptors) extending from control of cell development to stress responses [14]. These receptors play a crucial role in plant disease resistance [15], and their identification could lead to the development of high-density receptor candidates for genetic maps and crop improvement [16]. Likewise, associated gene arrangements among taxa, such as shared ancestry of genes and integrated findings from model organisms like *M. truncatula* or *G. max*, provide a valuable background to study less-understood systems [17].

The RLK/RLP legume analysis used in this study followed a comparative genomic approach based on the three objectives discussed in Chapters 2, 3, and 4. The first objective was the computational identification of RLK/RLP in legumes. A computational strategy integrating multiple tools and including a set of logical conditions was proposed to classify RLK/RLP in legumes/non-legumes. The strategy was statistically evaluated using an independent dataset of resistance RLK/RLP experimentally validated by other authors [18-23] to evaluate predictive performance [24]. The computational strategy allows practitioners to identify the quantity of RLK/RLP genes present in genomes and report a catalog of plasma membrane receptors and functional domains per species. The quality of the strategy performance was measured by calculating the specificity, sensitivity, and Matthews coefficient correlation (Figure 1).

The second objective was the phylogenetic analysis of RLK/RLP in legumes. To evaluate the catalog of RLK/RLP among 10 species, a similarity analysis using orthology and phylogenetic tree approaches was used. The predicted RLK/RLP were classified in three categories: ortholog clusters, singletons, and single-copy gene clusters. The distribution of the functional domains present in the predicted RLK/RLP among the legumes/non-legumes was explored using a hierarchical clustering approach of the functional domains among the plasma membrane proteins (Figure 1).

The third objective was the synteny analysis of RLK/RLP in legumes. This analysis followed a pairwise comparison spanning a number of highly similar genes shared within the same or similar genomic region. The analysis focused only on the regions of the species with RLK/RLP genes for the purpose of identifying patterns of evolutionary conservation and divergence across genomes among synteny blocks that have plasma membrane receptors (Figure 1). In total, 10-dicot genomes were included: seven legume species previously mentioned and three non-legume species: *Arabidopsis thaliana* (L.) Heynh [25] and tomato (*Solanum lycopersicum* (L.) H. Karst) [26] as model plants that could allow us to evaluate conservation and divergence, and common grape (*Vitis vinifera* L.) [27], because it represents the basal rosid lineage and has and ancestral karyotypes that facilitate comparisons across major eurosids [27, 28].

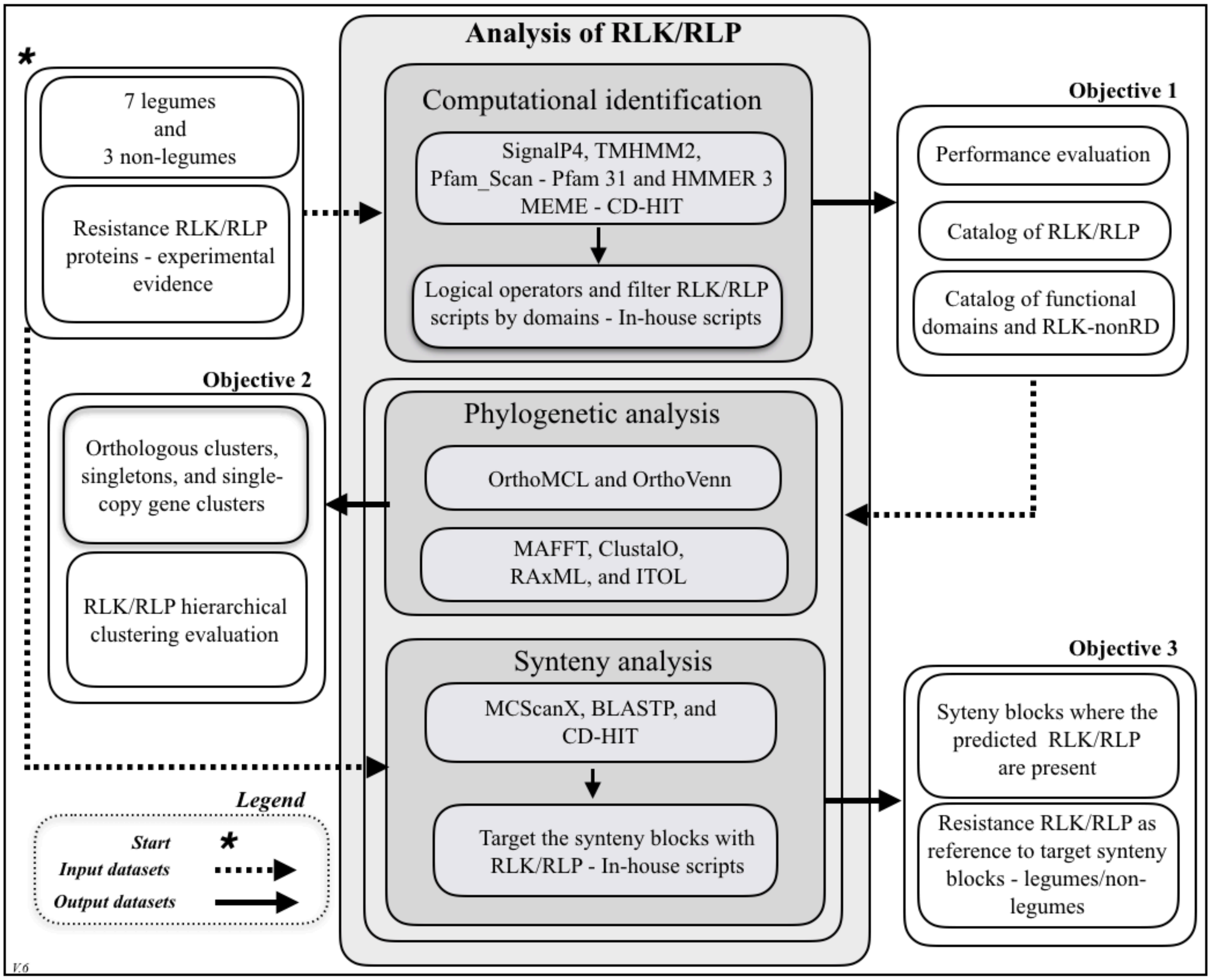


**Figure 1. Flow diagram of the methodological approach for RLK/RLP analysis.**
The figure describes the three objectives: the computational identification, phylogenetic analysis, and synteny analysis of RLK/RLP among species.

# CHAPTER 2. COMPUTATIONAL IDENTIFICATION OF RLK AND RLP IN LEGUMES

## Abstract

### *Background*

The plant cell wall is an active structure that mechanically connects cell tissues and controls the shape of the plant cell by sensing external stimuli and transmitting signals to the cytoplasm. The plasma membrane is enclosed by the cell wall and contains RLK and RLP proteins, which play a fundamental roll in cell to cell interaction and are crucial in plant growth, development, and immunity. These plasma membrane receptors belong to large gene/protein families that are not well computationally classified. They are characterized with automatic annotation platforms, but require manual expertise in computational identification to produce reliable results. Due to the economic relevance of legumes for human and animal consumption, this detailed analysis of these plasma membrane proteins in seven legume and three model (non-legume) species brings a new source of information to legume breeding and genetic research communities.

### *Results*

This computational research presents a logical classification approach based on a computational strategy for RLK and RLP protein identification. The strategy integrates the identification of these classes of plant receptors based on prediction of critical protein features, such as the presence or absence of signal peptide, the occurrence of transmembrane helices, the presence or absence of specific intracellular Pkinase domains, the absence of the NB-ARC domains, and the presence or absence of extra domains (LRR, L/C/G-Lectin, LysM, malectin, thaumatin, WAK/EGF, stress-antifung, and additional domains), which are normally located extracellularly and named as “target domains” for this study. The strategy was evaluated using

independent and experimentally-validated RLK and RLP proteins reporting a sensitivity of over 0.84, a specificity of 1.00, and a Matthews correlation coefficient of 0.90. The computational approach allowed the development of a detailed catalog of plasma membrane receptors (target and non-domains) for seven legumes, classified by Pfam clans/families' combinations per protein and the functional domain prevalence per species. The exclusive domains identified in legumes for RLK are WaaY, APH Pkinase_C, LRR_2, and EGF and for RLP are L-lectin LPRY and PAN_4. *G. max* has a larger set of RLK (1,867 proteins) and RLP (466 proteins) in its repertoire compared with the species evaluated, most likely due to its recent duplication. The RLK-nonRD and RLCK subclasses are reported for the species evaluated as well; in both classes, less than 20% of the total RLK predicted for each species belong to this class. In the domain combination comparison, up to four possible combinations of functional extra domains for RLK proteins were identified apart from the Pkinase domain. Among the 10-species evaluated about 40% of the proteins in the kinome are RLK. The exclusive legume domain combinations identified are B-Lectin/PR5K domains in *G. max*, *M. truncatula*, *V. angularis*, and *V. unguiculata* and a three-domain combination B-lectin/S-locus/WAK in *C. cajan*, *M. trucatula*, *P. vulgaris*, *V. angularis*. and *V. unguiculata*.

***Conclusions***

The analysis of both legumes/non-legumes species suggests that about 2% of the proteins of each genome belong to the RLK family and less than 1% belong to RLP family. Almost all of the extracellular domains evaluated were present in most of the species evaluated. More domain diversity combinations existed in the RLK compared with the RLP proteins and LRR domains, and the dual domain combination LRR/Malectin was the most frequent domain for both groups of plasma membrane receptors among legumes/non-legumes. The C-lectin domain in RLK proteins is usually reported as present only once per genome, a result that was confirmed among

the species evaluated; *G. max*, however, showed two RLK-C-lectin proteins, which would be logical considering it recent history of a whole genome duplication. Legumes exclusively show Pkinase, extracellular domains, and atypical domain combinations in RLK and RLP compared with the non-legumes evaluated. The computational logic approach used for the classification of RLK and RLP allowed the integration of all possible RLK and RLP structural conformations. This shows that the computational logic approach is statistically well supported based on the performance evaluation using an independent set of RLK/RLP experimentally validated, and can be used in other plants.

**Keywords:** Dicots, model plants, resistance genes/proteins, legumes/non-legumes.

## Background of RLK and RLP

Plants have evolved a surveillance system that is continuously monitoring a broad range of stimuli, including tissue damage or altered developmental processes, or establishing a symbiotic interaction. They commonly use pattern recognition receptors (PRR) to perceive 1) microbe-, pathogen-, or damage-associated molecular patterns (MAMP/PAMP/DAMP); 2) virulence factors; 3) secreted proteins; and 4) processed peptides directly or indirectly with specific molecular signatures [29]. These membrane-bound PRR are receptor-like kinases (RLK) or receptor-like proteins (RLP). The two receptor classes are located on the plant plasma membrane and are known as modular transmembrane proteins [18]. In contrast, the intracellular resistance proteins such as the nucleotide binding site-leucine-rich repeat proteins (NB-LRR or NBS-LRR) are encoded by the so-called resistance genes (R genes) and have been targeted to elicit a resistance response to crop pathogens in the field [20]. These intracellular resistance genes are out of the scope of this study.

According to the innate immunity plant system described by the zigzag model [30], the first line of plant cell defense, which can be a constituent of both non-host and host resistance, is based on PRR that activate the immune response upon detection of PAMP. This is called PAMP-triggered immunity (PTI) and confers broad-spectrum resistance against pathogens sharing the same PAMP; successful pathogens, however, suppress PTI-secreting effectors and produce effector-triggered susceptibility. The process of evolution responded to this phenomenon with the development of cytoplasmic R proteins that activate the effector-triggered immunity (ETI). The ETI is pathogen race-specific and is associated with programmed cell death (hypersensitive response or HR), and systemic-acquired resistance (SAR) in the host. These host-pathogen interactions lead to continuous selection pressure for novel pathogen isolates that overcome ETI and, correspondingly, new plant genotypes that resurrect ETI [30-32]. However, recent evidence suggests that a rigorous division cannot be considered among PAMP/effectors, PRR/R proteins, and PTI/ETI [32].

R genes are broadly categorized into eight classes based on their motif organization and membrane domains [33]. Following this classification and depending on their protein structure, three belong to the NBS class, with the remaining five classes related to RLK and RLP, such as Cf-9, Cf-4, and Cf-2 (class III); Xa21 (class IV); Ve1 and Ve2 (class V); RPW8 (class VI); and Pto, and RPg1 (class VIII), the latter of which confers high specificity interaction [33]. Proteins such as PGIP also play an important role for certain defense proteins even though they are not directly involved in pathogen recognition or activation of any defense genes [33]. On the other hand, the known PRR that confer a broad-spectrum resistance are modular transmembrane and they are either RLK or RLP proteins, and their recognition is based on a set of conserved molecules [31]. Most of the characterized RLK/RLP have been shown to be involved in defense/resistance processes in plants (Table A1). However, these plasma membrane receptors are also actively involved in cell growth and development, such as floral organ abscission (*A. thaliana* – HAESA) [34], meristem development (*A. thaliana* – CLAVATA) [35], self-incompatibility (MPLK) [36], abscission (CST) [37], stomatal patterning (TMM) [38], and embryonic patterning (SSP) [39]. In some cases, they are also simultaneously involved in immunity, growth, and development, such as botrytis-induced kinase 1 (BIK1) [40, 41] and brassinosteroid signaling kinases (BSK) [22].

RLK and RLP are structurally identified by the presence of motifs involved in the protein transport system, such as signal peptide. The extracellular domains or ectodomains, which are functional regions located outside of the cell, initiate contact with other molecules or surfaces and lead to signal transduction [18-23, 31, 42]. The presence of transmembrane helices allows the RLK/RLP to be anchored in the plasma membrane [43]. In the case of ectodomains, the LRR are part of the large family of N-glycosylated plant proteins, and many N-glycosylation acceptor sequences are present in all ectodomains [44]. The C (Carbohydrate-binding protein domain)/G (S-receptor-like or S-locus)/L (L-like lectin domain), LysM (Lysin Motif), and malectin classes of lectins are key players in plant immunity [45]. The C/G/L lectins are omnipresent in plants [46]. LysM are the most studied lectins of which 15 RLK-LysM and five RLP-LysM have been functionally characterized [47]. These proteins are known to play an essential role in plant defense signaling and inducing symbiosis. Among these proteins are NFR1 (Nod factor receptor 1) [48], NFR5 (Nod factor receptor 5) [48], LYK3 (putative Medicago ortholog of NFR1) [49], and NFP (LysM protein controlling Nod factor perception) [50], and their functional ligands are lipochitooligosaccharide nod factors [51]. Malectin-like domain-containing and FERONIA protein (FER or protein Sirene) receptors are recognized as critical regulators of cell growth and appear to function as surveyors of cell-wall status [52].

Other ectodomain families include the PR-5 family (Pathogenesis-related protein 5), composed of thaumatin-like proteins (TLPs) that are responsive to biotic and abiotic stress and are widely studied in plants [53]. Cell-wall-associated kinases (the "WAK" family) and their roles in signal transduction and pathogen stress responses arose from studies of the plant model species *A. thaliana* [54, 55]. The hallmark of a WAK is the presence of epidermal growth factor-like repeats ("EGF") in the extracellular domain [2, 3]. In contrast to the WAK, the evolution of the tumor necrosis factor/tumor necrosis factor receptor superfamily ("TNF/TNFR") is complicated and not well understood [56], and even though the TNFR domain of dicots and monocots is conserved, this domain family has some distinguishing characteristics among taxonomic families [57]. The stress-antifung domain family (known as DUF26 – Domain of Unknown Function) belongs to the cysteine-rich receptor-like protein kinases that form one of the largest groups of RLK in plants [58]. The structural details of RLK and RLP are presented in Table 1 [19-21, 59, 60].

**Table 1. Structural organization of RLK and RLP in plants.**

| Plasma membrane protein class | Signal Peptide | Transmembrane helix | Extracellular domain examples (RLK and RLP) | Intracellular domain Ser/Thr Kinase or Tyrosine Kinases |
|---|---|---|---|---|
| RLP | P* | P | All domains reported below | A |
| RLK | P* | P | LRR | P |
| | P* | P | L/C/G-Lectin | P |
| | P* | P | LysM | P |
| | P* | P | TNFR | P |
| | P* | P | Thaumatin | P |
| | P* | P | WAK | P |
| | P* | P | Malectin | P |
| | P* | P | EGF | P |
| | P* | P | Stress-Antifung | P |
| RLCK | A | P or A | Could have extra domains | P |

P: present, A: absent. *Proteins could have other transport system (non-classical export mechanism). RLP: Receptor-like proteins; RLK: Receptor-like kinase; LRR: Leucine Rich Repetitive; L/C/G-Lectin: Lectin Receptor Kinases [61]; LysM: Lysine Motifs; PR5K: Pathogenesis-related 5-like receptor kinases [62]; RLCK: Receptor-like cytoplasmic kinases; S-domain: Self-incompatibility domain [63]; TNFR: Tumor necrosis factor receptors; WAK: Wall-associated kinases; EGF: Epidermal growth-factor-like; Stress-Antifungal: Salt stress response/antifungal. The extra domains reported in this table are the most common domains functionally explored [18-23].

The RLK and RLP typically display high target specificity and selectivity [20, 64]. This condition provided an opportunity to understand how plants differentiate and distinguish favorable and harmful stimuli, as well as how various receptors coordinate their roles under variable environmental conditions [20]. The RLK family belongs to the protein kinase superfamily that has expanded in the flowering plant lineage, in part through recent duplications. Particularly, the flowering plant protein kinase repertoire known as "kinome," (a term coined by Manning et al., 2002 [65]), describes the catalog of protein kinases in a genome and is significantly larger (600 to 2,500 members) than the kinome in other eukaryotes. This large variation among organisms is principally due to the expansion and contraction of a few families; more than 60% of the kinome belongs to the receptor-like kinase/Pelle flowering plants family [66, 67]. The kinase domains can be divided into RD and non-RD families based on the presence or absence of an arginine located before a catalytic aspartate residue [68]. Non-RD kinases lack the strong autophosphorylation activities of RD kinases and display lower enzymatic activities [69]. Non-arginine-aspartate (non-RD) kinases are associated with innate immune receptors that recognize conserved microbial signatures [68]. Computational and comprehensive tools related to the prediction and analysis of resistance genes, such as RLK or RLP, could potentially support plant breeders/geneticists in detecting and understanding new plant disease resistance sources, which may be useful for crop improvement [70].

The receptor-like proteins (RLP) function with RLK to regulate development and defense responses. Also, the similarities between RLP and RLK and their functional relationships are consistent with the possibility that RLK with novel domain configurations may have been created through fusions between existing RLP and RLK [64, 71]. In addition, most RLP, which are secreted or membrane-spanning proteins, are likely integral components of extracellular signaling networks. Fusions between ancestral RLP and RLK/Pelle kinases could, therefore, have led to novel signal transduction pathways by linking ligand perception to different downstream kinase targets. Alternatively, fusions may simply have occurred between RLP and RLK/Pelle that were already components of the same signaling networks [64].

In recent years, more than 20 studies involving computational identification of cytoplasmic resistant proteins (mostly NBS-LRR) from different plants species have been published [16, 72], while analyses to identify RLK/RLP have been fewer (see review by Sekhwal and colleagues [16]), probably due to their functional diversity, which makes them harder to characterize compare to cytoplasmic resistant proteins. These genomic studies targeted many plant species [73], including *Arabidopsis* [74], *Arabidopsis* and rice (*Oryza sativa* L.) [75], grape (*Vitis vinifera* L.) [76], and tomato (*Solanum lycopersicum* (L.) H. Karst) [26], among others. To date, the strategies used similar approaches, but no standardized computational tools or consistent annotation criteria were followed. Thus, the results from different studies are not necessarily comparable [16]. Furthermore, the establishment of robust, independent, and highly diverse data with multiple examples is required to evaluate the performance of the strategies and tools published [24, 77].

Over the last eight years, legume genomics tools have expanded because of advancements in high-throughput sequencing and genotyping technologies resulting in reference genome sequences for many legume crops. This in turn allowed the identification of structural variations and enhanced the efficiency and resolution of large-scale genetic mapping and marker-trait association studies for legumes [8, 78]. Legumes are considered the second most important family of crop plants after the grass family based on their economic relevance. Approximately 27% of world crop production is composed of grain legumes, providing 33% of human dietary protein, while pasture and forage legumes are fundamental for animal feed [9]. To date, no RLK and RLP comparative genomic analyses were published that explores the genomes of soybean

(*Glycine max* (L.) Merrill) GM [1], common bean (*Phaseolus vulgaris* L.) PV [2], barrel medic (*Medicago truncatula* L.) MT [3], mungbean (*Vigna radiata* (L.) R. Wilczek) VR [4], cowpea (*Vigna unguiculata* L. Walp) VU [7], Adzuki bean (*Vigna angularis* var. *angularis*) VA [5], and pigeonpea (*Cajanus cajan* L.) CC [6].

This study reports on the computational identification of receptor-like proteins and receptor-like kinase proteins and probable resistance RLK-nonRD proteins in legumes using probabilistic methods [79-81]. The identification of these plasma membrane receptors is based on the prediction of absence/presence of a signal peptide, transmembrane helix motifs, and extracellular and intracellular domains. The performance of the proposed strategy was evaluated with experimentally-validated RLK (n: 64) and RLP (n: 26) proteins, and the RLK/RLP identification was applied on protein datasets that belong to the seven legume genomes mentioned above. Also, three non-legume model plant species were included to enrich the analysis. These species are *Arabidopsis thaliana* (L.) Heynh, AT [25]; tomato (*S. lycopersicum*) SL [26]; and common grape (*V. vinifera*) VV [27].

## Methods

### *Datasets*

To evaluate the proposed RLK/RLP identification strategy, three datasets were used (RLK, RLP and cytoplasmic resistance genes). All the datasets contain experimentally-validated proteins from 34 plant species (Table A1) and were extracted from the UniProt Consortium [82]. The RLK set contained 66 proteins, the RLP set contained 28 proteins, and the set of cytoplasmic resistance genes (non-RLK/RLP), contained 97 proteins (Table A1) [16, 20, 83-85]. To identify probable RLK and RLP, the analysis focused on seven legumes and also three non-legumes (outgroup set), including *V. vinifera* because it represents the basal rosid lineage and has a close-to-ancestral karyotypes that facilitate comparisons across major eurosids [27, 28]. Also, non-legumes *Arabidopsis* and *S. lycopersicu*m were included because they are model plants that could allow us to evaluate conservation and divergence.

The protein information of the legumes/non-legumes was obtained from the NCBI database: mungbean “VR” (GCF_000741045.1_Vradiata_ver6 ), pigeon pea “CC” (GCA_000340665.1_C.cajan_V1.0), and adzuki bean “VA”(*V. angularis* annotation release 100 - VA); the Phytozome-11 repository: soybean “GM” (gmax_275_wm82.a2.v1), barrel medic “MT” (Mtruncatula_285_Mt4.0v1), and common bean “PV” (Pvulgaris_442_v2.1); and the Phytozome-12 repository: *Arabidopsis* “AT” (Athaliana_167_TAIR10), tomato “SL” (Slycopersicum_390_ITAG2.4), cowpea “VU” (Vunguiculata_469_v1.1), and grape “VV” (Vvinifera_145_Genoscope.12X).

### ***Computational identification of RLK and RLP***

The computational strategy for RLK and RLP discovery is described in Figure 2. The identification of the presence/absence of signal peptide and transmembrane helices was predicted with SignalP 4.0 [79] and TMHMM 2 [80], respectively. The cut-offs used were Eukaryotes (euk): euk SignalP-noTM networks: 0.45 and euk SignalP-TM networks: 0.50 [79]. The selection criteria for TMHMM2 were based on the identification of one or more transmembrane helices, which must exceed the expected number of amino acids (ExpAA) threshold; if this value is larger than 18, it is very likely to be a transmembrane protein or have a signal peptide [80]. In both prediction processes, cut-off values are reported by default.

The PfamScan (pfamscan.pl) script [86] was used to annotate the protein sequences against the Pfam 31.0 library using HMMER 3.1b1 [81]. The selection criteria to assign a protein to each modular organization classes were defined by PfamScan, which states if overlapping matches within a clan are detected, it will then only report the most significant, which will be the lowest E-value match within the clan [87]. In some cases, proteins belonged to two domain classes, but the redundant information was extracted in the counting process. To establish a domain cutoff for Pfam-A searches, the parameter used by default was based on the diverse set of domains to reach these trusted cut-offs, which were defined by Pfam curators and their variable for each domain or family [81].

The PfamScan output was filtered using in-house scripts (https://github.com/drestmont/plant_rlk_rlp/) for the identification of RLK/RLP and their structural domains. The identification of the modular organization domains (Table 1) is defined in the Pfam database [88] as profiles and clans (labelled: CL); the clans are profiles grouped together with a common evolutionary ancestor [86]. The in-house script includes 134 Pfam domains representing the extra domains and the Pkinase reported in Table 1. They are considered "target domains" for this research and are reported in Table 2 (Table A2).

**Table 2. Target domains for the classification of RLK/RLP.**

| Functional family* | Clan or Domain | Number of domains Reported in Pfam 31 |
|---|---|---|
| LRR | CL0022 | 11 |
|  | LRRNT_2 | 1 |
| Pkinase | CL0016 | 35 |
|  | Pkinase_C | 1 |
| L-Lectin | CL0004 | 43 |
| C-Lectin | Lectin_C | 1 |
| G-Lectin | B_lectin | 1 |
|  | S_locus_glycop | 1 |
| LysM | LysM | 3 |
| PR5K | Thaumatin | 1 |
| TNFR | TNFR | 6 |
| PAN | CL0168 | 6 |
| WAK | WAK | 1 |
|  | GUB_WAK | 1 |
|  | WAK_assoc | 1 |
| Malectin | CL0468 | 2 |
| EGF | CL0001 | 18 |
| Stress-antifungal | Stress-antifungal | 1 |
| NB-ARC | NB-ARC | 217 |

*Source: Pfam 31.0 [89].The domains reported in Table 2 are not exclusively present on RLK and RLP. The NB-ARC belong to R genes, which belong to cytoplasmic proteins and were used to exclude false positive proteins.

The identification approach follows this logic (logical operators: and, or, and not) (Figure 2) for RLK: “presence/absence Signal peptide” and “transmembrane helix (at least one)” and “Pkinase domain/s” and “Extracellular domain/s: LRR or L-Lectin or C-Lectin or G- Lectin or LysM or PR5K or TNFR or WAK or Malectin or EGF or Stress-Antifung” not “NB-ARC” domains and, for RLP: “presence/absence Signal peptide” and “transmembrane helix (at least one)” and “Extracellular domain/s: LRR or L-Lectin or C-Lectin or G-Lectin or LysM or PR5K or TNFR or WAK or Malectin or EGF or Stress-Antifung” not “Pkinase domain/s” and not “NB-ARC domains”. Finally, a summary of the domain and family prevalence among species was obtained based on the RLK/RLP identified in the evaluation set and the species explored. The frequency analysis was based on the evaluation of “experimentally-validated protein datasets” (Table A1), and also for the identified proteins, which belong to the species evaluated. After the RLK proteins per species were classified to identify potential non-RD proteins, the entire set of Pkinase sequence domains was broken into subsets using the start and end domain coordinates reported by PfamScan. The MEME command line tool version [90] was used to identify the RD and non-RD motif sites, and the MEME parameters used were as follows: -mod oops -maxw 10 -nmotifs 4 -maxsize 6,000,000. After the motif sites were reported, they were classified as RD ([H][R][D]) and non-RD ([H][^R][D]) motif (regex notation). The kinome was identified annotating the whole set of proteins per species using pfamscan.pl. The proteins with the presence of Pkinase domains were filtered (Figure 4 - footnote).

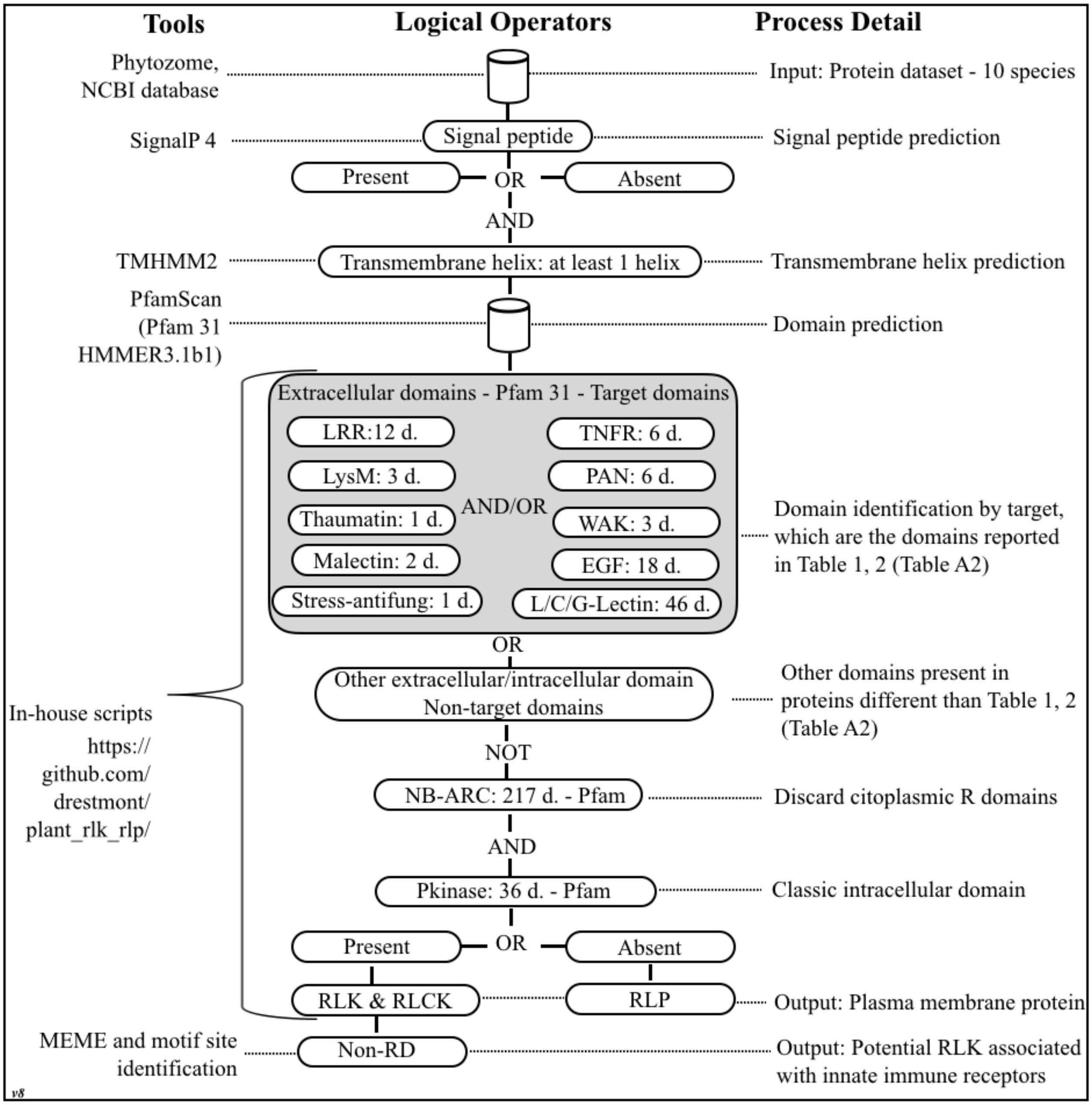


**Figure 2. Computational strategy followed to identify RLK and RLP.**

### *Independent evaluation of predictive performance*

To evaluate the RLK an RLP prediction strategy we calculated if a protein does or does not belong to an RLK or RLP class. Also, to test the prediction performance, three evaluation sets with known outcomes supported by experimental evidence were built (Table A1). For the performance evaluation measurement, sensitivity (range: 0 to 1), specificity (range: 0 to 1), and

Matthews correlation coefficient “MCC” (range: -1 to 1) were selected [24]. In the evaluation datasets, the identification of experimentally-validated proteins for each class became the true positive (RLP and RLK) and true negative data (cytoplasmic resistance genes) [24]. The cytoplasmic resistance genes could have similar ectodomains to RLK/RLP but shown exclusively the NB-ARC domains [91]. The datasets obtained were independently processed using CD-HIT [92] to obtain a non-redundant version using a 90% identity to avoid similar or highly similar overlapping entries [24]. The predictive analysis of RLK/RLP were applied to the non-redundant sets; for the RLK evaluation, the RLP and “cytoplasmic resistance genes” sets were used as true negative proteins; for the RLP evaluation, the RLK and “cytoplasmic resistance genes” sets were used as true negative proteins.

## Results

### *Performance prediction of RLK and RLP*

The independent performance evaluation of the computational strategy demonstrated 56 out of 64 RLK proteins were identified as true RLK, and the remaining proteins were non-detected and considered as false negatives. In contrast, 22 out of 26 RLP proteins were classified as true RLP, and the remaining proteins were non-detected and classified as false negatives. Lastly, none of the 97 proteins belonging to the cytoplasmic R genes set were classified as RLK or RLP classes (Table A3). Based on these results, the performance predictive measures were calculated (Table 3).

**Table 3. Performance evaluation**

| Measure | RLK | RLP |
|---|---|---|
| Sensitivity | 0.87 | 0.84 |
| Specificity | 1 | 1 |
| Matthews correlation coefficient | 0.90 | 0.90 |

Non-redundant datasets used for the performance evaluation are RLK, n:64; RLP, n:26; and Other R genes, n:97. Table A1: list of experimentally-validated proteins used for this evaluation (RLK, RLP, and cytoplasmic resistance proteins), and Table A3: evaluation performance summary.

The evaluation allowed the establishment of a minimum set of conditions to classify the RLK and RLP protein classes. RLK- and RLP-predicted proteins must have at least one transmembrane helix and the presence of at least one extracellular domain (LRR, L/C/G-Lectin, LysM, PR5K, thaumatin, WAK, malectin, EGF, or stress-Antifung). Additionally, for RLK, the presence of the intracellular Pkinase domains is required, and for RLP, the absence of Pkinase and NB-ARC domains is required; these logic conditions are stated in Figure 1.

### *Summary of predicted RLK and RLP*

Based on the number of RLK and RLP identified among the species evaluated, about 3% or less of the total proteins per species belong to these classes of plasma membrane proteins. Specifically, for legumes, the percentage range for RLKs is 0.93% to 2.34%, and for non-legumes is 1.3% to 1.68%. The RLP percentage range for legumes is 0.29% to 0.69%, and for non-legumes is 0.46% to 0.64%. The species analysis evaluated 447,948 proteins, 351,491 from legumes, and 96,457 from non-legumes. Almost 9.35% of the legumes and 9.75% of the non-legumes had a predicted signal peptide, and 4.32% of legumes and 4.41% of non-legumes had at least one transmembrane helix above the threshold. For the subset of proteins without the presence of a signal peptide, 16.56% of legumes and 17.90% of non-legumes reached the TMHMM cut-off. Among the total number of proteins evaluated, 1.94% of legumes and 1.53 of non-legumes belong to the RLK class of proteins, and 0.54% of legumes and 0.51% of non-legumes belong to the RLP class (Table 4). Also, the number of RLK proteins identified as non-

RD, which are potentially kinases associated with 22 innate immune receptors, are reported in Table 4 footnote, and the differentiated protein identified by species are in the supplementary file (Table A4, Table A5).

**Table 4. Summary of total number of RLK and RLP identified across legumes/non-legumes.**

| | | Signal peptide | | | Transmembrane helices | | RLK/RLP proteins identified per species | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Species | Total proteins reported | Pre/ Abs | Number of proteins | % | Number of proteins | % | RLK* | % RLK | RLP* | % RLP |
| *C. cajan* | 48,331 | P | 2,679 | 5.54 | 1,031 | 2.13 | 197 | | 62 | |
| | | A | 45,652 | 94.46 | 5,760 | 11.92 | 253 | | 79 | |
| | | | | | | total | 450 | 0.93 | 141 | 0.29 |
| *G. max* | 88,647 | P | 8,125 | 9.17 | 3,934 | 4.44 | 1,182 | | 282 | |
| | | A | 80,522 | 90.83 | 15,459 | 17.44 | 685 | | 184 | |
| | | | | | | total | 1,867 | 2.11 | 466 | 0.53 |
| *M. truncatula* | 62,319 | P | 6,251 | 10.03 | 2,961 | 4.75 | 647 | | 196 | |
| | | A | 56,068 | 89.97 | 10,383 | 16.66 | 415 | | 167 | |
| | | | | | | total | 1,062 | 1.70 | 363 | 0.58 |
| *P. vulgaris* | 36,995 | P | 4,120 | 11.14 | 1,895 | 5.12 | 571 | | 138 | |
| | | A | 32,875 | 88.86 | 6,349 | 17.16 | 272 | | 77 | |
| | | | | | | total | 843 | 2.28 | 215 | 0.58 |
| *V. angularis* | 37,769 | P | 3,570 | 9.45 | 1,681 | 4.45 | 557 | | 124 | |
| | | A | 34,199 | 90.55 | 6,364 | 16.85 | 279 | | 90 | |
| | | | | | | total | 836 | 2.21 | 214 | 0.57 |
| *V. radiata* | 35,143 | P | 3,450 | 9.82 | 1,584 | 4.51 | 505 | | 141 | |
| | | A | 31,693 | 90.18 | 5,934 | 16.89 | 266 | | 99 | |
| | | | | | | total | 771 | 2.19 | 240 | 0.68 |
| *V. unguiculata* | 42,287 | P | 4,698 | 11.11 | 2,105 | 4.98 | 660 | | 190 | |
| | | A | 37,589 | 88.89 | 7,962 | 18.83 | 333 | | 103 | |
| | | | | | | total | 993 | 2.35 | 293 | 0.69 |

**Table 4. Summary of total number of RLK and RLP identified across legumes/non-legumes. (Continued)**

| | | Signal peptide | | | Transmembrane helices | | RLK/RLP proteins identified per species | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Species | Total proteins reported | Pre/ Abs | Number of proteins | % | Number of proteins | % | RLK* | % RLK | RLP* | %RLP |
| *V. vinifera* | 26,346 | P | 2,043 | 7.75 | 842 | 3.20 | 270 | | 96 | |
| | | A | 24,303 | 92.25 | 4,980 | 18.90 | 174 | | 73 | |
| | | | | | | total | 444 | 1.69 | 169 | 0.64 |
| *A. thaliana* | 35,386 | P | 4,088 | 11.55 | 1,935 | 5.47 | 408 | | 121 | |
| | | A | 31,298 | 88.45 | 5,784 | 16.35 | 148 | | 49 | |
| | | | | | | total | 556 | 1.57 | 170 | 0.48 |
| *S. lycopersicum* | 34,725 | P | 3,258 | 9.38 | 1,480 | 4.26 | 316 | | 107 | |
| | | A | 1,467 | 90.62 | 5,727 | 16.49 | 161 | | 53 | |
| | | | | | | total | 477 | 1.37 | 160 | 0.46 |

For each species, the results were distinguished by the present "P" and absent "A" of signal peptide and follow the logic flow presented in Figure 1. *Non-redundant data reported. For the RLK-nonRD, the results per species are: *A. thaliana*: 48 proteins (8.6%), *C. cajan*: 61 proteins (13.6%), *G. max*: 223 proteins (11.9%), *M. truncatula*: 194 proteins (18.3%), *P. vulgaris*: 124 proteins (14.7%), *S. lycopersicum*: 83 proteins (17.4%), *V. angularis*: 122 proteins (14.6%), *V. radiata*: 113 proteins (14.7%), *V. unguiculata*: 158 proteins (15.9%), and *V. vinifera*: 59 proteins (13.3%). RLK-nonRD IDs are reported in Table A6. The kinome (total set of proteins with a kinase in a genome) per species was calculated and the results for the species are CC: 1,268 p. (35.5% - RLK), GM: 4,497 p. (41.4% - RLK), MT: 2,281 p. (46.6% - RLK), PV: 1,888 p. (44.7% - RLK), VA: 1,898 p. (44% - RLK), VR: 1,772 p. (43.5% - RLK), VU: 2,090 p. (47.5% - RLK), VV: 1,064 p. (41.7% - RLK), AT: 1,431 p. (38.9% - RLK), and SL: 1,194 p. (39.9% - RLK).

Based on the Pfam clans and families of functional domains used to filter the RLK and RLP (Table A2), the computational strategy allowed a determination of whether one or more extra domains were present per protein. For the RLK proteins reported in Table 5, the approach identified up to four combinations of functional extra domains (located extra or intracellularly), besides a Pkinase domain. All species evaluated had proteins with at least one extra domain (Table 1). The exception was the TNFR domain in which the in-house scripts (https://github.com/drestmont/plant_rlk_rlp/) did not identify its presence in any of the datasets; however, when reviewing the approach, it was found that the TNFR domains predicted by Pfam 31, HMMER, and PfamScan did not reach the minimum cut-offs in the prediction process followed.

The G-lectin class of proteins reported in Table 5 is typically composed of three domains (B-lectin/S-locus/PAN); however, different combinations of those three domains were identified. C-lectin is a rare domain, and only soybean species showed more than one such protein. The WAK is typically composed of two domain classes (WAK/EFG), and such proteins possessed one or the other domain. The dual domain combination LRR/Malectin is the most frequent among the atypical dual combinations. Also, atypical domain combinations with a low frequency among the species were identified. Among the legumes, these were the B-Lectin/PR5K combination in GM, MT, VA, and VU and a three-domain combination of B-lectin/S-locus/WAK only in CC, MT, PV, VA, and VU. Among non-legumes, the uncommon dual combinations PAN/WAK and PAN/S-locus/WAK were only found in VV. The only uncommon domain combination found in both legumes/non-legumes was S-locus/WAK in VV and VR.

A four-domain combination, consisting B-lectin/S-locus/PAN/WAK domains was present GM, MT, PV, SL, VA, VR, VU, and VV species. Across all legume/non-legume species, the LRR ectodomain class was the most frequent domain per species. The computational classification strategy also discovered RLK proteins with no other domains and some proteins with the additional domains beyond the signal peptide, a transmembrane helix, and a Pkinase domains. In the case of the RLCK, the proteins that belong to this class are the kinases without signal peptide, but with a transmembrane helix. The RLCK without another plasma membrane attachment domain was not predicted (Table 5).

**Table 5. Receptor-like kinases identified by extracellular domains across the species.**

| | | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Domain class | Domain combinations | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| LRR | lrr | 134 | 579 | 324 | 239 | 254 | 249 | 301 | 136 | 180 | 198 |
| G-lectin: combination of ectodomains | s-locus | 2 | 1 | 1 | 1 | 1 | 1 | 3 | 2 | 0 | 0 |
| | b-lectin | 7 | 20 | 25 | 12 | 12 | 14 | 15 | 7 | 2 | 7 |
| | b-lectin/pan | 2 | 9 | 7 | 12 | 2 | 5 | 17 | 2 | 1 | 5 |
| | s-locus/pan | 5 | 10 | 4 | 5 | 1 | 0 | 7 | 18 | 2 | 0 |
| | b-lectin/s-locus | 11 | 24 | 14 | 15 | 14 | 18 | 15 | 7 | 2 | 10 |
| | b-lectin/s-locus/pan | 31 | 146 | 131 | 41 | 53 | 44 | 96 | 12 | 33 | 42 |
| L-Lectin | l-lectin | 24 | 66 | 46 | 38 | 35 | 36 | 42 | 20 | 44 | 22 |
| C-lectin | c-lectin | 1 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Lectin | lysM | 7 | 27 | 16 | 14 | 11 | 12 | 13 | 5 | 5 | 8 |
| Lectin (Feronia) | malectin | 29 | 99 | 54 | 82 | 58 | 50 | 60 | 29 | 36 | 22 |
| Thaumatin (Osmotin) | pr5k | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| WAK | wak | 11 | 66 | 33 | 41 | 45 | 39 | 46 | 14 | 27 | 17 |
| | egf | 1 | 4 | 0 | 1 | 0 | 0 | 1 | 2 | 0 | 2 |
| | wak/egf | 5 | 10 | 16 | 6 | 3 | 7 | 8 | 7 | 4 | 7 |
| DUF26 recently renamed | stress_antifung | 28 | 173 | 66 | 70 | 57 | 58 | 90 | 22 | 45 | 15 |
| Classically related to G-lectin | pan | 5 | 10 | 1 | 2 | 2 | 0 | 1 | 3 | 0 | 0 |
| Combination of different domain ectodomains identified | lrr/malectin | 12 | 63 | 66 | 32 | 30 | 19 | 28 | 25 | 47 | 7 |
| | pan/wak | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | s-locus/wak | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 7 | 0 | 0 |
| | b-lectin/pr5k | 0 | 1 | 3 | 0 | 1 | 0 | 1 | 2 | 0 | 0 |
| | b-lectin/s-locus/wak | 2 | 0 | 1 | 2 | 3 | 0 | 3 | 3 | 0 | 0 |
| | b-lectin/s-locus/pan/wak | 0 | 8 | 2 | 2 | 2 | 2 | 5 | 8 | 0 | 1 |
| | pan/s-locus/wak | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Table 5. Receptor-like kinases identified by extracellular domains across the species. (Continued)**

| | | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Domain class | Domain combinations | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| RLK - pkinase | rlk – non-target ectodomain | 4 | 11 | 6 | 5 | 3 | 5 | 6 | 0 | 8 | 5 |
| Combination of ectodomains | rlk - not ectodomains | 30 | 180 | 74 | 87 | 93 | 72 | 86 | 0 | 25 | 28 |
| Identified RLCK with/without ectodomains | rlck extra domain | 8 | 7 | 6 | 6 | 10 | 3 | 5 | 5 | 5 | 9 |
| | rlck only pkinase | 91 | 346 | 163 | 128 | 144 | 133 | 142 | 16 | 86 | 70 |

For each species, the results were merge by present “P” and absent “A” of signal peptide. All possible domain combinations were explored and are reported in the “Domain combinations” column (proteins reported are non-redundant). *A. thaliana*: AT, *C. cajan*: CC, *G. max*: GM, *M. truncatula*: MT, *P. vulgaris*: PV, *S. lycopersicum*: SL, *V. angularis*: VA, *V. radiata*: VR, *V. unguiculata*: VU, and *V. vinifera*: VI (Table A4). RLCK: Only kinase domain identified. All proteins reported in this table have at least one transmembrane helix. Extra: proteins that have the presence/absence of signal peptide, at least one transmembrane helix, a Pkinase and other extracellular/intracellular domains different than LRR, L/C/G-Lectin, LysM, Pr5k-Thaumatin, WAK, Malectin, EGF or Stress-Antifung were only considered for the combination identification analysis, but other domains reported in Table A7 named as “non-target” domains could be present.

For the RLP extracellular domain identification and domain combinations reported in Table 6, the computational approach allowed the identification of up to three possible combinations of functional extra domains (which could be located extra or intracellularly) in the proteins evaluated; however, all combinations correspond to the typical combinations reported in Table 1, such as the G-lectin (B-lectin/S-lectin/PAN) present in legumes/non-legumes, the classic WAK/EGF only present in CC and VV (legume/non-legume), and the LRR/Malectin present in all species evaluated. However, the three cases mentioned were of a low frequency compared with other domains, such as LRR or Stress-antifung. As in RLK, for RLP, the most abundant ectodomain was the LRR per species, and no RLP proteins were identified with C-lectin or TNFR domains present; recall that the TNFR domains did not reach the thresholds in the prediction process using Pfam 31, HMMER, and PfamScan.

**Table 6. Receptor-like proteins identified by extracellular domains across the species.**

| | | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Domain details | Domain combinations | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| LRR | lrr | 69 | 247 | 225 | 107 | 104 | 138 | 171 | 78 | 71 | 67 |
| G-lectin: combination of ectodomains identified | s-locus | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| | b-lectin | 1 | 5 | 2 | 5 | 3 | 2 | 5 | 8 | 4 | 1 |
| | s-locus/pan | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| | b-lectin/s-locus | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 1 |
| | b-lectin/s-locus/pan | 2 | 3 | 3 | 3 | 1 | 2 | 5 | 4 | 2 | 1 |
| L-lectin | l-lectin | 31 | 88 | 54 | 35 | 39 | 33 | 34 | 27 | 34 | 36 |
| Lectin | lysM | 3 | 7 | 8 | 5 | 3 | 5 | 4 | 3 | 2 | 4 |
| Lectin (Feronia) | malectin | 5 | 12 | 7 | 3 | 5 | 3 | 5 | 8 | 7 | 3 |
| Thaumatin (Osmotin) | pr5k | 7 | 28 | 19 | 13 | 16 | 17 | 20 | 7 | 15 | 16 |
| WAK | wak | 5 | 14 | 16 | 8 | 9 | 12 | 11 | 10 | 8 | 12 |
| | egf | 4 | 17 | 5 | 6 | 8 | 5 | 8 | 3 | 5 | 3 |
| | wak/egf | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| DUF26 recently renamed | stress_antifung | 12 | 34 | 15 | 22 | 22 | 19 | 24 | 9 | 23 | 14 |
| Classically related to G-lectin | pan | 1 | 3 | 3 | 1 | 1 | 1 | 3 | 3 | 0 | 0 |
| Combination of target ectodomains | lrr/malectin | 1 | 9 | 5 | 9 | 4 | 4 | 4 | 5 | 1 | 3 |

For each species, the results were distinguished by present “P” and absent “A” of signal peptide, all possible domain combinations were explored and are reported in the “Domain combinations” column. Proteins reported are non-redundant. *A. thaliana*: AT, *C. cajan*: CC, *G. max*: GM, *M. truncatula*: MT, *P. vulgaris*: PV, *S. lycopersicum*: SL, *V. angularis*: VA, *V. radiata*: VR, *V. unguiculata*: VU, and *V. vinifera*: VI. All proteins reported in this table have at least one transmembrane helix. Other domains reported in Table A8 named as “non-target” domains could be present.

### *Summary of the presence and prevalence of functional domains*

As a result of the identification process for RLK and RLP, the specific domains that belong to the clans and families in Tables 1 and 2 (Table A2) are reported in Tables 7, 8, and 9 and summarized in Figure 3.

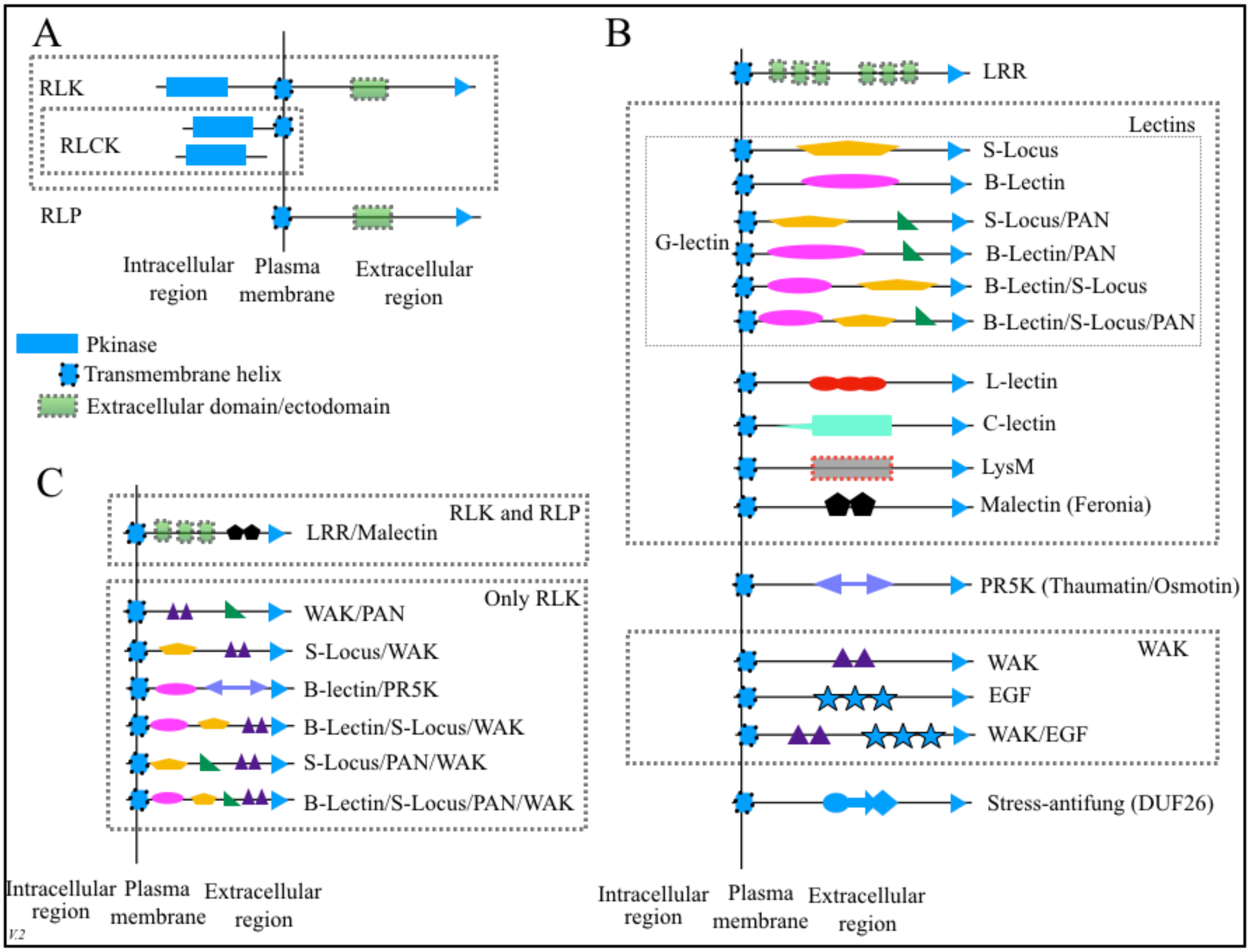


**Figure 3. Summary of the extracellular domains identified in RLK/RLP.**
The domains in this figure resume the domains and the combinations identified. A. Classical RLK/RLP protein structure. B. Ectodomains identified that are also reported by the scientific community (Table 1 and 2). C. Ectodomain combinations identified in RLK/RLP. In B and C, the ectodomains are only represented, in the RLK cases all proteins must have an intracellular Pkinase.

Table 7 shows the domains identified in the RLK and RLP proteins (Table A1) used for the performance evaluation of the plasma membrane identification process. The next two tables are related to the predicted RLK and RLP. Table 8 shows the domains identified in the predicted

RLK, and Table 9 shows the domains identified in the predicted RLP. In the target domains identified on the experimentally-validated RLK and RLP proteins (S1), almost all of the domains were identified for the RLK with the exception of the C-Lectin and TNFR domains; also, two additional domains were found on the proteins evaluated (DUF3403 and CL0384). For the RLP evaluated, only domains belonging to LRR and LysM were identified. Regarding the ectodomain classes reported for RLK and RLP (Table A1), the expected domains were identified using the strategy implemented in this study (Table 7).

**Table 7. Summary of domains identified in the validation dataset.**

| Clan or Domain | Domain or Family | RLK | RLP |
|---|---|---|---|
| PKinase | Pkinase_Tyr | X* | |
| | Pkinase | X | |
| LRR | LRR_8 | X | X |
| | LRRNT_2 | X | X |
| | LRR_1 | X | X |
| | LRR_4 | X | X |
| | LRR_6 | X | X |
| L-Lectin | Lectin_legB | X | |
| G-Lectin | B_lectin | X | |
| | PAN_2 | X | |
| | S_locus_glycop | X | |
| LysM | LysM | X | X |
| PR5K | Thaumatin | X | |
| WAK | GUB_WAK_bind | X | |
| | WAK | X | |
| Malectin | Malectin_like | X | |
| EGF | EGF_CA | X | |
| Stress-Antifung | Stress-antifung | X | |
| DUF3403 | DUF3403 | X | |
| CL0384 | GDPD | X | |

*X: Present. Source: Table A1: the list of experimentally-validated proteins used for this evaluation were RLK, n:64 and RLP, n:26.

As a result of the entire set of domains identified in the predicted RLK, 125 Pfam domains (Table 8 and Table A7) were classified, with 35 domains (Table 8) belonging to the “target domains” (Tables 1 and A2). The remaining domains are included in Table A7. Independent of the Pkinase domains, which are cytoplasmically located, the other domains could be present either extra- or intracellularly. Comparing the domains identified in the predicted RLK and RLP against the target Pfam domains (Table A2) for the identification of extra/intracellular domains, 10 out of 35 Pkinase domains, 7 out of 12 LRR domains, 1 out of 43 L-Lectin domains, 1 out of 1 C-Lectin domains, 5 out of 8 G-Lectin domains, 1 out of 3 LysM domains, 1 out of 1 PR5K domain, 3 out of 3 WAK domains, 2 out of 2 Malectin domains, 3 out of 18 EGF domains, and 1 out of 1 Stress-antifung domain were identified. Also, with the exception of the TNFR, all families and domains reported in Table 1 were identified in all 10 species. Of the non-target domains (90 in total, Table A7), the most prevalent were RCC1_2, DUF3403, Ribonuc_2-5A, NAF, DUF3660, and Glyco_hydro_18, all of which were present in at least eight species (legumes/non-legumes); the remaining domains (84 in total) were present in two or fewer species.

**Table 8. Summary of domains present on the RLK proteins predicted.**

| | | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Clan or domain | Pfam Domain name ID | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| Pkinase | Ins_P5_2-kin | | | | | | x | | | x | |
| | RIO1 | | | x | | | | | | x | |
| | Pkinase | x | x | x | x | x | x | x | x | x | x |
| | PI3_PI4_kinase | x | | | | | | | x | x | |
| | Pkinase_Tyr | x | x | x | x | x | x | x | x | x | x |
| | Choline_kinase | | x | x | | | | | x | x | x |
| | ABC1 | x | x | x | x | x | x | x | x | x | x |
| | Pkinase_C | x | x | | | | | | | | |
| | PIP5K | | x | | | | | | | | x |
| | WaaY | | | x | | | | | | | |
| | APH | | | | x | x | x | x | | | |
| LRR | LRRNT_2 | x | x | x | x | x | x | x | x | x | x |
| | LRR_8 | x | x | x | x | x | x | x | x | x | x |

**Table 8. Summary of domains present on the RLK proteins predicted. (Continued)**

| | | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Clan or domain | Pfam Domain name ID | C C | G M | M T | P V | V A | V R | V U | V V | A T | S L |
| | LRR_1 | x | x | x | x | x | x | x | x | x | x |
| | LRR_4 | x | x | x | x | x | x | x | x | x | x |
| | LRR_6 | x | x | x | x | x | x | x | x | x | x |
| | LRR_2 | | x | | | x | | x | | | |
| | LRR_5 | | | | x | x | x | x | | | x |
| L-Lectin | Lectin_legB | x | x | x | x | x | x | x | x | x | x |
| C-Lectin | Lectin_C | x | x | x | x | x | x | x | x | x | x |
| G-Lectin | B_lectin | x | x | x | x | x | x | x | x | x | x |
| | S_locus_glycop | x | x | x | x | x | x | x | x | x | x |
| PAN | PAN_2 | x | x | x | x | x | x | x | x | x | x |
| | PAN_1 | | x | | | | | x | | | x |
| LysM | LysM | x | x | x | x | x | x | x | x | x | x |
| PR5K | Thaumatin | | x | x | | x | | x | | x | |
| WAK | WAK_assoc | x | x | x | x | x | x | x | x | x | x |
| | WAK | x | x | x | | | | | x | x | x |
| | GUB_WAK_bind | x | x | x | x | x | x | x | x | x | x |
| Malectin | Malectin_like | x | x | x | x | x | x | x | x | x | x |
| | Malectin | x | x | x | x | x | x | x | x | x | x |
| EGF | EGF_CA | x | x | x | x | x | x | x | x | x | x |
| | EGF | | x | x | | | | | | | |
| | EGF_3 | | | x | x | x | x | x | | | x |
| Stress-antifung (DUF26) | Stress-antifung | x | x | x | x | x | x | x | x | x | x |

Present: X.

For the entire set of domains identified in the RLP, 71 domains (Table 9 and Table A8) were identified, 33 (Table 9) belong to the “target domains” (Tables 1 and A2), and the remaining domains are reported in Table A8. All domains present in this dataset are extracellularly located. Comparing the domains identified with the Pfam clans and families from Table A2 used to identify extra/intracellular domains, 8 out of 12 LRR domains, 8 out of 43 L-Lectin domains, 5 out of 8 G-Lectin domains, 1 out of 3 LysM domains, 1 out of 1 PR5K domain, 3 out of 3 WAK domains, 2 out of 2 Malectin domains, 4 out of 18 EGF domains, and 1 out of 1 Stress-antifung domain were identified. Also, with the exception of C-Lectin and the TNFR family, all families and domains are reported in Table 1. Of the non-target domains (38 in total Table A8), the most prevalent were DUF2854, Glyco_hydro_32N, DUF3357, Alliinase_C, Galactosyl_T, zf-RING_2, PA, Peptidase_M8, and Exostosin, all of which were present in at least six species; the remaining domains (29 in total) were present in three or fewer legumes/non-legumes species.

**Table 9. Summary of domains present on the RLP proteins predicted.**

| | | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Clan or Domain | Domain name | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| LRR | LRR_8 | x | x | x | x | x | x | x | x | x | x |
| | LRR_1 | x | x | x | x | x | x | x | x | x | x |
| | LRRNT_2 | x | x | x | x | x | x | x | x | x | x |
| | LRR_2 | | | x | | | | | | x | |
| | LRR_4 | x | x | x | x | x | x | x | x | x | x |
| | LRR_6 | x | x | x | x | x | x | x | x | x | x |
| | LRR_9 | x | | | | | | | | | |
| | LRR_5 | | | | | | | | x | | |
| L-Lectin | Gal-bind_lectin | x | x | x | x | x | x | x | x | x | x |
| | Glyco_hydro_32C | x | x | x | x | x | x | x | | x | x |
| | XET_C | x | x | x | x | x | x | x | x | x | x |
| | Lectin_legB | x | x | x | x | x | x | x | x | x | x |
| | Glyco_hydro_16 | x | x | x | x | x | x | x | x | x | x |
| | Calreticulin | x | x | x | x | x | x | x | x | x | x |
| | SPRY | x | x | x | x | x | x | x | | | |
| | Alginate_lyase2 | | x | | | | | | x | | |
| G-Lectin | B_lectin | x | x | x | x | x | x | x | x | x | x |
| | S_locus_glycop | x | x | x | x | x | x | x | x | x | x |
| PAN | PAN_2 | x | x | x | x | | x | x | x | x | x |
| | PAN_1 | x | x | x | x | x | x | | | x | |
| | PAN_4 | x | x | x | x | x | x | x | | | |
| LysM | LysM | x | x | x | x | x | x | x | x | x | x |
| Thaumatin (PR5K) | Thaumatin | x | x | x | x | x | x | x | x | x | x |

**Table 9. Summary of domains present on the RLP proteins predicted. (Continued)**

| | | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Clan or Domain | Domain name | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| WAK | WAK_assoc | x | x | x | x | x | x | x | x | x | x |
| | WAK | | | | | | | | x | x | |
| | GUB_WAK_bind | x | x | x | x | x | x | x | x | x | x |
| Malectin | Malectin_like | x | x | x | x | x | x | x | x | x | x |
| | Malectin | | x | x | x | | | | x | | x |
| EGF | EGF_alliinase | x | x | | | | | x | x | x | x |
| | cEGF | x | x | x | x | x | x | x | x | x | x |
| | EGF_CA | x | | x | | | | | x | | |
| | EGF_2 | | x | x | x | x | x | x | | | x |
| Stress-antifung | Stress-antifung | x | x | x | x | x | x | x | x | x | x |

X: Present.

## Discussion

The process to evaluate the performance of the computational prediction approach used RLK and RLP proteins directly or indirectly related to resistance. The quality of the validation dataset (Table A1) is ideal because the data come from diverse species and are independent, experimentally-validated, and non-redundant. Based on the legume/non-legume results, the RLK proteins are more diverse in terms of domains compared with the RLP proteins (Table 7). With respect to sensitivity and specificity, the first measure evaluated how well the approach correctly classified a protein as RLK/RLP, with the results suggesting it has the ability to designate a true RLK/RLP as positive, with few false negatives. The second measure evaluated the ability of the approach to correctly classify a protein as non-RLK/RLP. Those results indicate a greater ability to identify few false positive proteins. Based on the Matthews correlation coefficient, the performance evaluation reports a very strong positive value (0.90), which suggests the approach is ideal for RLK/RLP identification [24].

As for the RLK/RLP prediction requirements described in Figure 1, the prediction and identification of RLK using the logic sum of conditions seem to be simple. The Pkinase domain is a requirement for the RLK proteins, in contrast with the logic sum of conditions that a protein needs to be classified as RLP. Interestingly, for the last plasma membrane class mentioned, apart from the conditions that proteins must meet to belong to the RLP class (Table1), one factor that improves the confidence of the prediction and reduces false positive protein is the exclusion of cytoplasmic resistance genes which could be confounded with RLP. This identification is reached excluding the proteins with NB-ARC domains present.

Of the total plasma membrane proteins reported in Table 4, the results for *G. max* indicate a larger set of RLK and RLP compared with the species overall, which is probably due to its recent whole genome duplication [93]. In fact, soybean, had a very recent genome duplication 13 Mya [94], and would be expected to have more plasma membrane receptors than most other plant species [95]. These duplication mechanisms are the main proximal cause for the higher degree of expansion of the protein kinase superfamily in plants [95]. Regarding the RLK-non-RD class (Table A6), with the exception of the non-legume AT (8.6%), the other legumes/non-legumes species (CC (13.6%), GM (12.0%), MT (18.3%), PV (14.7%), SL (17.4%), VA (14.6%), VR (14.7%), VU (15.9%), and VV (13.3%)), have more than 12% RLK with this modification in the kinase domain. This subset of RLK is interesting because it has been previously found that most PRR kinases or PRR-associated kinases have a change in a conserved arginine (R) located adjacent to the key catalytic aspartate (D) (the so-called RD motif) that facilitates phosphotransfer [68, 96], and further comparative analysis must be done.

Compared with RLK, the majority of RLCK reported in Table 5 only contain a Ser/Thr-specific cytoplasmic kinase domain, corresponding to results previously reported results [97]. However, no “target domains” were identified, contrary to the additional domains previously reported, which mention that apart from the Pkinase, the RLCK could have similar domains intracellularly to the ectodomains present in the RLK, such as leucine rich repeat (LRR), lectin, epidermal growth factor (EGF), a domain of unknown function (DUF), U-BOX, and WD40 [97]. With the exception of the non/legume VV (4.7%), all other legumes/non-legumes (AT (16.4%), CC (22%), GM (18.9%), MT (15.9%), PV (15.9%), SL (16.6%), VA (18.4%), VR (17.7%), and VU (14.82%)) had more than 15% of the RLK classified as RLCK. This is important because a number of RLCK have emerged as central components linking PRR to downstream defenses. These PRR are involved in transducing signals from extracellular ligand perception into downstream signaling by phospho-relay [85]; several *Arabidopsis* RLCK are associated with PRR and play important roles in PTI [84].

The number of RLK per species reported in this study (Table 4) is proportionally similar to the 1% to 2% of total gene models per species reported in previous studies, where RLK normally represented about 60% or more of protein kinases [66, 67]. The range of RLK proteins identified in this study was 450-1867 for legume proteins and 444-556 for non-legume proteins. The legumes GM (1,867 proteins) and MT (1,062 proteins) showed the highest number of RLK (Table 3). In contrast, the range of values for legume RLP proteins was 141-466 proteins and 160-170 for non-legume proteins. As with RLK, the legumes GM (466 proteins) and MT (363 proteins) showed the highest number of RLP.

Given that the RLK receptor configuration arises from a fusion between an RLP and an RLCK [98], it could be expected that RLP have similar ectodomains, excluding the LRR and LysM domains experimentally reported in RLP (Table A1). The presence of other extracellular domains, which are mainly associated with RLK (Table 1), was explored to identify probable RLP with the presence of L/C/G-lectin, TNFR, thaumatin, WAK, malectin, EGF, or stress-antifung domain. This approach was based on the similarities reported among two-plasma membrane receptors and suggests a consistent functional relationship and the possibility of novel domain configurations created by their fusion [64]. This approach discovered that for legumes

(0.29% to 0.69%)/non-legumes (0.46% to 0.64%), less than 1% of the proteins present in the evaluated genomes belonged to the RLP class (Tables 4 and 6).

Even though the TNFR domains belonging to both plasma membrane classes in the species evaluated (Tables 5 and 6) were not identified, a detailed evaluation showed that in the prediction process step (Pfam31, HMMER3.1, and PfamScan.pl), the domain match was considered insignificant because the bit score fell below the software threshold. However, RLK proteins have been predicted as RLK with a TNFR extracellular domain and reported in the SMART database in an earlier study [99] for *A. thaliana* (2 proteins)*, G. max* (4 proteins)*, S. lycopersicum* (2 proteins), and *V. vinifera* (3 proteins). Interestingly, with the exception of the *V. vinifera* proteins*,* the eight other proteins were identified as RLK either with non-target domains or only the Pkinase domain (Table 5). All details about these proteins are included in Table A9. Other domains missed could include L-Lectin and TNFR for RLP (Table 6) and the missed TNFR identification in the summary of domains reported in Table 7. This exploration of missing domains suggests that including tools such as SMART could add precision to the predictions in some instances.

Regarding the diverse domain combinations identified in the RLK and RLP listed in Tables 5 and 6, RLK, in particular, vary greatly in their extracellular domain organization. A variety of extracellular domains are present in RLK [23]; compared with RLP, the RLK identified in this study shown interesting domain combinations, such as LRR/Malectin; the S-locus/WAK present only in the legume VA and the non-legume VV; the B-lectin/PR5K present only in the legumes GM, MT, VA, and VU; the B-lectin/S-locus/WAK present only in the legumes CC, MT, PV, VA, and VU; and the B-Lectin/S-locus/Pan/WAK shared among the legumes GM, MT, PV, VA, and VU, and non-legumes VV and SL. The unique non-common ectodomain combination identified in RLP was LRR/Malectin, which was present in all species evaluated (Table 6). This suggests a higher diverse combination of domains in RLK compared with RLP. Some RLK domain combinations were only reported for legumes, while RLP combinations were present among legumes and non-legumes.

Regarding the diversity of the Pfam domains reported in Tables 7, 8, and 9, the use of all possible clans and family domains summarized in Table 1, allowed the classification of all possible domains present in RLK and RLP. This approach suggests an advantage rather than using only target specific domains using motifs [100]. Diversity of the Pfam domains was most evident in the RLK class for the Pkinase, with 10 domains/families, LRR: 7 d/f identified; for the RLP class, diversity was most evident for the LRR: 8 d/f, L-Lectin: 8 d/f, and EGF: 4 d/f. Among the 10 Pkinase domains/families, WaaY in MT; APH in PV, VA, VR and VU; and Pkinase_C in CC and GM were exclusively present in the legumes. For the 7 RLK-LRR, the LRR_2 was exclusively present in the legumes GM, VA, and VU. For other family domains, the EGF domain was only present in the legumes GM and MT. In contrast, for the ectodomains present in RLP, the LRR_9 from the LRR clan was only present in CC; the L-lectin clan with the LPRY domain and the PAN clan with the PAN_4 domain were exclusive to all the legumes. Interestingly, those clans are collections judged likely to be homologous and are valuable because they are built manually and integrate a diverse variety of information sources, allowing a transfer of structural and functional information between families and improving the prediction of structure and function of unknown families [101]. The classification of non-target domains present for RLK (Table A7) and RLP (Table A8) among the species demonstrated that none of the most prevalent domains identified (present in 10-species) in both plasma membrane classes was common, suggesting a bias related to the kind of plasma membrane relation. This suggests that further analysis could be done to explore probable correlations among the domains evaluated.

## Conclusions

The identification of RLK and RLP based on the use of different machine-learning tools publicly available for the prediction of different biological aspects, such as export mechanisms, transmembrane helices, and specific domains that structurally characterize them, allowed this study to propose a simple, logical, and effective set of conditions. These conditions in turn allow for the identification of multiple combinations of extracellular and intracellular domains that could be present and used to classify plasma membrane receptors in plants. The approach used in this study was validated using an independent set of RLK/RLP proteins involved in resistance, which have been experimentally evaluated and reported by the scientific community.

The validation demonstrated that the approach is highly effective in identifying RLK/RLP proteins in terms of sensitivity, specificity, and MCC. The output obtained after the exploration of the set of proteins belonging to 10 species of dicots led to the conclusion that, related to the proteins present in a species, less than 2% and 1% of the proteins identified belong to RLK and RLP, respectively. The domains organization of RLK was more diverse compared with the domain organization of RLP domains. More L-lectin domain diversity exists in RLP (8 domains) compared with RLK (1 domain). Specifically, for the RLK, the non-RD represented 8% to 18%, and the RLCK represented about 15% of this class of plasma membrane proteins per species evaluated. For RLK (exception TNFR) and RLP (C-Lectin and TNFR), almost all of the target proteins reported in Table 1 were present in the legumes/non-legumes evaluated.

Regarding the legume/non-legume comparison, *G. max* contains a larger set of RLK (1,867 proteins) and RLP (466 proteins) compared with the legume/non-legume species. The RLK-nonRD proteins among legumes/non-legumes, with the exception of the non-legume AT (8.6%), had more than 12% RLK with this condition in the kinase domain (CC (13.6%), GM (12.0%), MT (18.3%), PV (14.7%), SL (17.4%), VA (14.6%), VR (14.7%), VU (15.9%), and

VV (13.3%). The RLCK among legumes/non-legumes, with the exception of the non/legume VV (4.7%), had more than 15% of the RLK classified as RLCK (AT (16.4%), CC (22%), GM (18.9%), MT (15.9%), PV (15.9%), SL (16.6%), VA (18.4%), VR (17.7%), and VU (14.82%)). Across all species, the LRR ectodomain class was the most frequent domain per species. C-lectin is a rare domain commonly reported only once per genome, and only the GM species showed more than one such protein, which could be related to the recent whole genome duplication. The approach allowed the identification of up to four possible combinations for RLK with functional extra domains (that could be located extra- or intracellularly) in a protein, besides a Pkinase domain.

For RLK/RLP among legumes/non-legumes, the LRR/Malectin domain combination is the most frequent among the dual combinations. Also, the atypical domain combinations in RLK with low frequency among the species were identified among the legumes as the B-Lectin/PR5K in GM, MT, VA, and VU, and a three-domain combination B-lectin/S-locus/WAK only in CC, MT, PV, VA, and VU. Among non-legumes, the dual PAN/WAK and PAN/S-locus/WAK were found only in VV. The only combination found in both legumes/non-legumes was S-locus/WAK in VV and VR. The unique case of a four-domain combination in legume/non-legume species was B-lectin/S-locus/PAN/WAK present in GM, MT, PV, SL, VA, VR, VU, and VV. In contrast, in legumes/non-legumes, the atypical domain combinations in RLP G-lectin B-lectin/S-lectin/PAN were present, but the typical WAK/EGF domain combination was only present in CC and VV.

Among the 10-Pkinase domains identified, WaaY in MT; APH in PV, VA, VR and VU; and Pkinase_C in CC and GM were exclusively present in legumes. As for the 7-RLK-LRR, the LRR_2 was exclusively present in the GM, VA, and VU legumes. For the RLK-EGF clan, the

EGF domain was only present in the legumes GM and MT. And for the RLP-L-lectin clan and PAN clan, the LPRY domain and the PAN_4 domain were exclusively present in all legumes.

# CHAPTER 3. PHYLOGENETIC ANALYSIS OF RLK AND RLP IN LEGUMES

## Abstract

### *Background*

The adaptability and productivity of legumes are limited by major biotic and abiotic stresses. These stresses directly involve the group of plasma membrane receptor proteins known as RLK and RLP. Evaluating the similarity and homology relations among RLK and RLP and exploring their phylogenetic relationships can improve the understanding of evolutionary relations among various legume species.

### *Results*

This phylogenetic analysis of RLK and RLP proteins in legumes used two approaches: an orthologous analysis, including the identification of orthologous clusters, singletons, and single-copy gene clusters, and a hierarchical clustering evaluation of the plasma membrane receptors and their ectodomains (LRR, L/C/G-Lectin, LysM, malectin, thaumatin, WAK/EGF, stress-antifung, and additional domains). The RLK-nonRD, which are plasma membrane proteins potentially associated with innate immune receptors, were grouped into three main clusters in the hierarchical tree. Typically, a high proportion of RLK and RLP legume proteins belong to orthologous clusters. However, in this study among legumes, between 66% to 90% of the proteins per species were classified in orthologous clusters and the remaining as singletons, and only eight single-copy gene clusters were identified. The ectodomains identified on the RLK and RLP showed clustering patterns among the respective hierarchical trees based on the ectodomain annotation. The clustering analysis of RLK found RLK-nonRD to be present in all classes, with two blocks tightly grouped, the first among RLKC and proteins with WAK/EGF and G-lectin ectodomains and the second among LRR ectodomains. The RLCK were distributed along the

RLK tree. As for the RLP tree, two main clusters were identified, with the first block including the LRR ectodomains, and the second block the remaining ectodomains.

### *Conclusions*

The analysis of the RLK and RLP suggest a dynamic evolution in the legume family, with between 66% to 85% of RLK and 83% to 88% of RLP belonging to orthologous clusters among the species evaluated; the remaining proteins were classified as singletons. In fact, for the 10-species, a lower number of singleton proteins were reported among RLP compared to RLK, suggesting that RLP are more conserved compared to RLK. The RLK and RLP phylogenetic trees are partially resolved and clustered based LRR, malectin, LysM, L/G/C-lectin, thaumatin, WAK/EGF, and stress-antifung and their combinations. The phylogenetic trees suggested that the RLP protein conformation is less diverse compared to the RLK, which is supported by the results reported in chapter 2. The results about the RLK-nonRD clusters suggests that future analysis could reveal new motifs associated to this kinase modification.

**Keywords:** Dicots, orthology, extracellular domains, plasma membrane receptors, legumes/non-legumes.

## Background of RLK and RLP in legumes

Legumes are derived from a common ancestor 60 million years ago (Mya) [93]. Based on morphological characters, three major legume subfamilies exist: mimosoids (Mimosoideae), caesalpiniods (Caesalpinioideae), and papilionoids (Papilionoideae). The latter subfamily contains the cultivated grain legumes or pulses and can be subdivided into four clades: 1) Phaseoloids: *Glycine* spp. Willd., *Phaseolus* spp. L., *Cajanus* spp. L., and *Vigna* spp. Savi; 2) Galeogoids: *Pisum* L., *Lens* Mill., *Lathyrus* L., *Vicia* L., *Medicago* L., and *Cicer* L.; 3) Genistoids: *Lupinus* L.; and 4) Dalbergoids: *Arachis* L. [102]. The domestication of the Fabaceae family as grain legumes has been reported in conjunction with cereals [103]. However, more legumes have been domesticated overall, which makes the Fabaceae family the taxon with the greater number of domesticates [9, 103]. Of the clades, the Phaseolid group of warm-season legumes was domesticated later than the Galeogoids group of cool-season legumes [9].

The Papilionoideae subfamily, the largest clade among the legumes, is monophyletic. It shares a common ancestor, and its chloroplast experienced a 50kb inversion 50 Mya [93]. Research shows that the timing of polyploidy (whole genome duplication, or WGD), which affects most lineages in this clade, occurred after the divergence of the mimosoid and papilionoid clades, but the precise timing is still unknown [104]. Among the most recognized legumes are *Medicago truncatula* L., which is considered a model species for legumes [105], and the cultivated legumes pigeon pea (*Cajanus cajan* L.) [6], soybean (*Glycine max* (L.) Merrill), mungbean (*Vigna radiata* (L.) R. Wilczek) [4], cowpea (*Vigna unguiculata* L. Walp) [106], adzuki bean (*Vigna angularis* var. *angularis*) [5], and common bean (*Phaseolus vulgaris* L.) [2]. In 2005, WGD events were reported that established the legume phylogenetic relationship [93]. Interestingly, during the last 135 to 250 million years of evolution, the protein-coding gene families have been affected by different biological events, such as various gene duplication mechanisms, including WGDs (or polyploidization) and segmental and tandem duplications among other factors [17, 107, 108].

In legumes, several WGD and triplication events occurred soon after the split between monocots and eudicots [109]. Common grape (*Vitis vinifera* L.) divergence is known to have occurred early in eudicot evolution; due to this event, grape is considered ideal for phylogenetic discrimination among legumes [109]. The legume species diverged from one another about 37.6 Mya [110]. Based on the fossil records, the divergence of Fabales from the closest group Rosales and Cucurbitales was estimated at 59.9 Mya. The divergence of Rosids and Asterids occurred around 89.3 Mya, and a Papilionoideae-specific WGD was observed among legumes. In addition, soybean recent duplication occurred about 13 Mya [1]. Soybean, pigeonpea, mungbean, and common bean evolved from a common ancestor about 23.9 Mya (Figure 4).

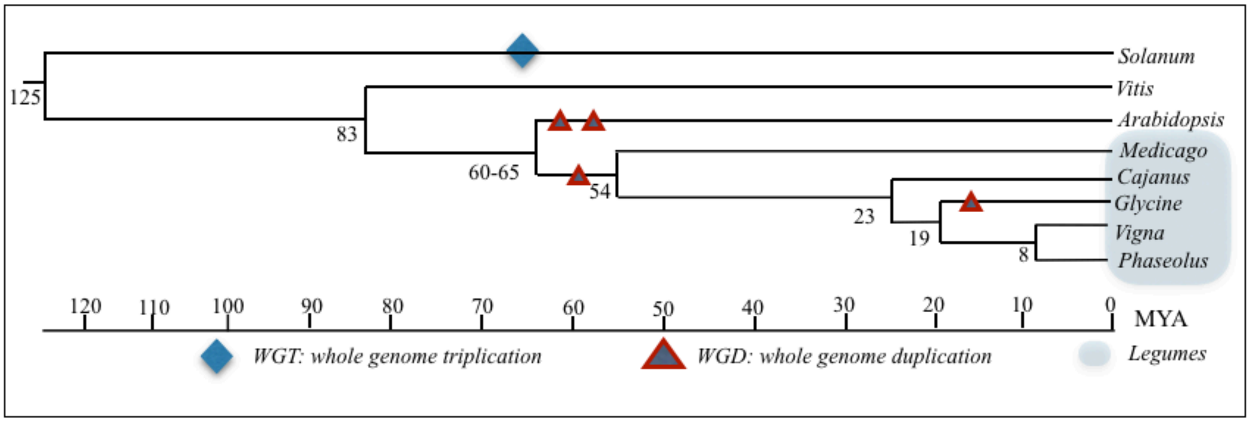


**Figure 4. Taxonomic relationships among legumes/non-legumes.**
Taxonomic relationships within the Papilionoideae subfamily and the species included as outgroups for the RLK and RLP phylogenetic analysis are shown**.** The topology and distances reported were adapted [93, 111, 112]. *Solanum lycopersicum*, *Vitis vinifera* and *Arabidopsis thaliana* included as outgroup species for this study.

Recently, whole-genome sequencing of legumes [111] has allowed for comparative genomic analyses to be conducted. Such research requires a complex genome annotation process that depends on the identification of homologous sequences evaluated at intervals and, preferably, orthologous to sequences of known identity and function in related genomes. Orthologous genes (orthologs) are the result of speciation events that are derived from a common ancestor [113]. Orthologs are interesting because they can be predicted to have conserved all or part of the ancestral biological function [114]. Comparative analysis also allows the identification of orthologous clusters of single-copy genes and singletons. The single-copy status is conserved through evolutionary time [115]. The singletons are genes/proteins that are not present in any orthologous group or remain ungrouped based on the datasets evaluated [116]. This sort of analysis is ideal for RLK, RLP, and RLCKs (cytoplasmic RLK) because of their evolutionary relationships and important roles in plant signaling and because their gene subfamilies are large with complex histories of gene duplication and loss [117]. The evaluation of RLK/RLP among *Aradidopsis*, *Lotus japonica,* and *Medicago truncatula* reported such gene duplication and loss, with a high frequency of reciprocal gene loss in the LRR-RLK/RLP, and RLCK subfamilies. Furthermore, pairwise comparisons showed lineage-specific duplications associated with reciprocal gene loss [117].

RLK/RLP are involved in diverse biological functions [84, 118], and in recent years, evidence is emerging that upon ligand binding, RLK also form homodimers or heterodimers with other kinases and RLP, functioning in multiprotein complexes to initiate plant immunity [119, 120]. While the possible role of RLK and RLP in plant immunity has been functionally characterized, matching these proteins to their ligands is still challenging [47]. Interestingly, an example of crosstalk between RLK pathways controlling immune signaling and growth/development is exemplified by the study of Feronia, which plays multiple roles in the regulation of plant growth and development [84, 121]. At the same time, multiple RLCK have been identified as fundamental components linked to defense strategies [122]. Among these, the botrytis-induced kinase (BIK1) is a central component that integrates signals from pattern recognition receptors (PRR) interacting with CERK1, FLS2, EFR, and PEPR [40, 84, 123, 124].

The ectodomain or extracellular domains are highly variable among RLK and RLP, providing a means to recognize a wide range of ligands. Some RLK and RLP act as PRR that perceive thread signals [84], while, others are involved in the regulation of plant development and growth, symbiosis, reproduction, and tolerance to abiotic stresses [125]. Interestingly, plants prioritize immunity over other functions if pathogen-imposed menaces are present [126]. Among the RLK, the kinase domains can be divided into RD and non-RD families depending on the presence or absence, respectively, of an arginine positioned before a catalytic aspartate residue [68]. Non-RD kinases lack the robust autophosphorylation activities of RD kinases, display lower enzymatic activities [69], and have been associated with innate immune receptors [68].

Also, PRR often belong to the RLK and RLP subfamilies that display lineage-specific expansion (LSE) [84], defined as the proliferation of a protein family in a particular lineage relative to the sister lineage [127]. Plants have higher rates of gene duplication than other organisms, and, together with the functional bias in gene retention, contribute to differences in the degree of LSE [128]. The LSE feature helps in the identification of PRR and allows deployment of disease resistance by cross-species/genus transfer of PRR-coding genes [84].

Among the most common ectodomains present in RLK and RLP are the LRR, lectins, PR5k “thaumatin,” WAK/EGF, TNFR-CR4 family, and stress-antifungal. Phylogenetically, the LRR ectodomain containing RLP seems more similar to the LRR containing RLK. A significant number of RLP can be located within RLK clusters, suggesting that RLP were either derived from or gave rise to RLK [129]. Lectins are essential to plant life and are involved in cell-to-cell communication, development, and resistance. At the membrane surface, lectins are present both as soluble and chimeric RLK and RLP [130] and are classified as G-Lectin, L-Lectin, C-Lectin, LysM, and malectin. Proteins of the PR5k family have high-sequence identity with thaumatins, which are proteins isolated from the West African shrub *Thaumatococcus daniellii* and also known as thaumatin-like proteins [131]. For decades, thaumatin-like proteins have been studied in plants for their antifungal properties [132]. Wall-associated kinases (WAK) comprise a subfamily of proteins within the RLK superfamily, though not all WAK and WAK-like genes encode transmembrane proteins [133]. Also, numerous domains (GubWAK, WAK, WAKassoc, EGF-Ca2+, cEGF, and EGF3) may be found in the extracellular part of the WAK proteins, contributing to its variability [134]. The TNFR or CRINKLY4 (CR4) family of receptor-like kinases is involved in a wide range of developmental processes in plants [57]. And stress-antifung (known as DUF26) belonging to the cysteine-rich receptor kinases forms one of the largest groups of receptor-like protein kinases in plants, but its biological functions remain unclearly defined [58].

Based on the computational identification of RLK/RLP in legume species presented in Chapter 2, an orthology analysis and a functional phylogenetic or hierarchical clustering evaluation of the plasma membrane receptors were undertaken. Both the orthology analysis and functional phylogenetic evaluation were based on amino acid sequences. The purpose of the hierarchical clustering evaluation was to identify the RLK- and RLP-similar relationship among legumes/non-legumes. The seven legumes involved in this evaluation were *G. max*, *P. vulgaris*, *M. truncatula*, *V. angularis*, *V. radiata*, *V. unguiculata*, and *C. cajan*. It allowed the RLK/RLP relationship to be evaluated and the extracellular domains (LRR, L/C/G-lectin, LysM, malectin, thaumatin, WAK/EGF, stress-antifung, and others) to be tracked as a reference. Also, three non-legume species: *Arabidopsis thaliana* (L.) Heynh [25], tomato (*Solanum lycopersicum* (L.) H. Karst) [26], and common grape (*Vitis vinifera* L.) [27], were used as the outgroup species in the hierarchical clustering evaluation; the first two species were included because most of the experimentally-validated RLK/RLP belong to this species (Table A1) and are biological models in plants. Grape also represents the basal rosid lineage and has close-to-ancestral karyotypes that facilitate comparisons across major eurosids [27, 28] (Figure 4).

## Methods

### *Datasets*

The RLK (Table A4), RLP (Table A5), and RLK-nonRD (Table A6) classified for the seven-legume species, and grape (*V. vinifera*) (as the closest legume “outgroup”), tomato (*S. lycopersicum*), and *Arabidopsis* were included as plant models in dicots (Figure 4). The RLK-nonRD dataset represented about 10% of the total RLK, but it was extracted to evaluate its relationship since it is potentially associated with innate immune receptors that recognize conserved microbial signatures [68]. The experimentally-validated RLK and RLP were also included in the evaluations (Table A1). In the RLK dataset, nine proteins were not included because of the presence of string regions showing 4 ≥ more undefined amino acids labeled as "X" in a continuous position in 50% or more of its whole sequence.

### *Orthology analysis*

The homology inference to RLK, RLK-nonRD, and RLP among all species was calculated with the OrthoMCL [135] tool that reports orthologous clusters, Blastp threshold: E-value 1e -5 and MCL inflation parameter of 1.5 (default parameters). The results were visualized with OrthoVenn [115]. The criteria to be included in the orthology analysis followed this order based on the genomes evaluated: A) among the legumes selected, included one species per genera, B) prioritized the genome species included from the Phytozome repository due to its quality standards, and C) included the closest outgroup. These criteria allowed the identification of *M. truncatula:* MT, *C. cajan* (not present in the Phytozome repository)*:* CC, *G. max:* GM, *P vulgaris:* PV, *V. unguiculata:* VU. and the outgroup *V. vinifera:* VV. The orthology analysis for *V. radiata*, *V. angularis*, *A. thaliana*, and *S. lycopersicum* was reported as supplementary. This process allowed functional orthologous clusters (orthologous and paralogous), single-copy gene clusters, and singletons to be identified.

### *Hierarchical clustering analysis of RLK and RLP*

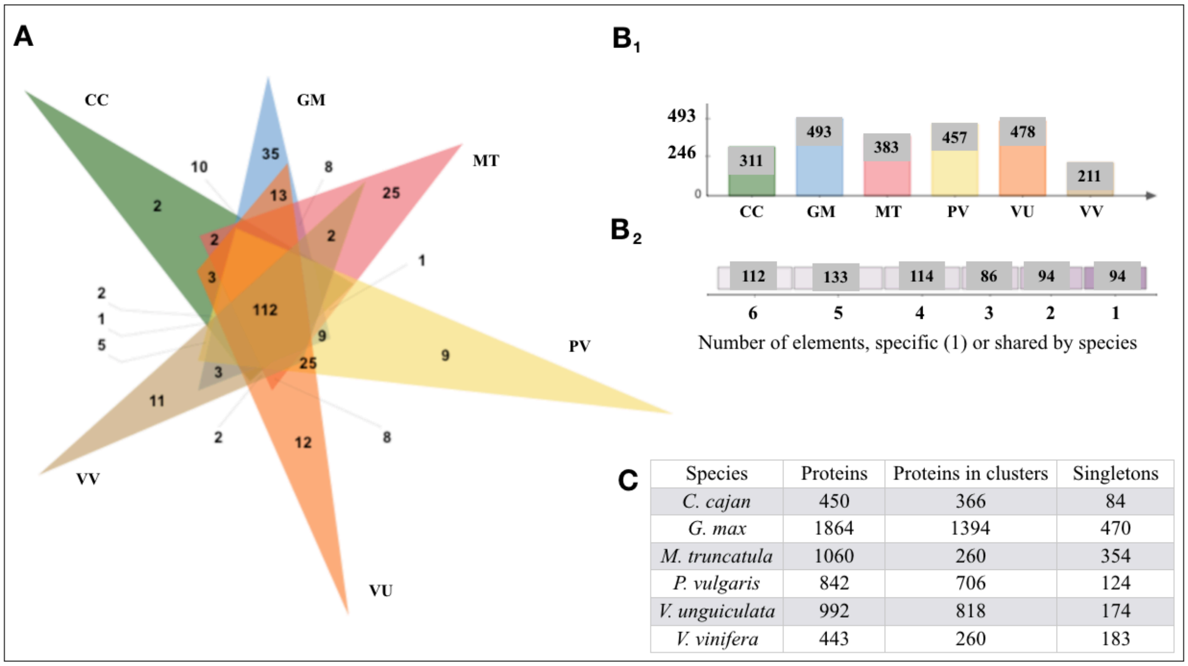
The hierarchical clustering exploration of RLK and RLP followed a different approach using different criteria for species inclusion. All of the plasma membrane proteins predicted in Chapter 2 among the legumes/non-legumes were included. Multiple alignments of amino acid sequences were performed using the ClustalO tool (default parameters) [136]. The alignments were adjusted using trimAL [137] with the parameters–noallgaps, -gt 0.9, and -cons 60. The first parameter removed columns with at least one gap and columns with only gaps (where the whole column was empty). The second and third parameters removed all positions in the alignment with gaps in 10% or more of the sequences unless this left less than 60% of positions. This left as an output the 60% best (with fewer gaps) positions. The sequence alignment and trimal exploration were executed in Google Cloud Service using a virtual machine (VM Debian8) with 22 cores, 60 GB RAM, and 200 GB SSD in free trial mode.

The phylogenetic trees of the legume species were constructed using different phylogenetic methods on the Cipres platform [138]. A maximum likelihood approach using RAxML [139] with the parameters -c 25, -p 12345, -m PROTCATJTT, -f a, -N 100, -x 12345, and --asc-corr lewis was used to obtain the results. Due to the size of the datasets, which can be considered small for RLP (2,445 proteins) and medium for RLK (8,278 proteins), and following the recommendations reported on the in the RAxML-v8.2.X manual [139], the GAMMA distribution was calculated independently, deploying 10 processes in parallel and a different –p parameter for the initial random seed. The best RAxML results were integrated using the CAT and GAMMA results to obtain the phylogenetic tree and its bootstrap values. The JTT matrix was used because of its suitability for plasma membrane proteins [140]. The time to construct the RLK and RLP trees was 360 hours and 19 hours, respectively (this took 50,000 CPU hours using the Cipres-Xsede platform (http://www.phylo.org/) [138]). The stability of internal nodes was determined by bootstrap analyses with 100 replicates.

## Results

### *Orthology and hierarchical clustering analyses of RLK*

Based on the criteria explained above, six species were selected for the analyses: CC, GM, PV, MT, VU, and VV. In the case of the *Vigna* genera, there were three species available; however, *V. unguiculata* or VU was selected as the genera representative because of data quality. The orthology and hierarchical clustering domain analyses of RLK for the species resulted in the formation of 633 orthologous and paralogous clusters, 539 orthologous clusters containing at least two species, and seven single-copy gene clusters. In total, all six species were found in 112 of the orthologous clusters, the outgroup included. Also, all six-species presented unique clusters, which can be formed by paralogous or protein isoforms that belong to the same gene (Figure 5: A and B2). The remaining 427 orthologous clusters were shared by at least two legume species, with 87 orthologous clusters shared by all five-legume species. *G. max* was the species with the most singletons (Figure 5: C) and proteins present in orthologous clusters (Figure 5: B1). The co-orthologs (recent descent and duplication), in-paralogs (recent duplication), and orthologs (recent descent) were also calculated. The orthology results for *V. radiata*, *V. angularis*, *A. thaliana, V. vinifera* and *S. lycopersicum* are included in Figure A1. The RLK-nonRD were not excluded in the analysis to evaluate the whole set of RLK predicted (Figure 5).

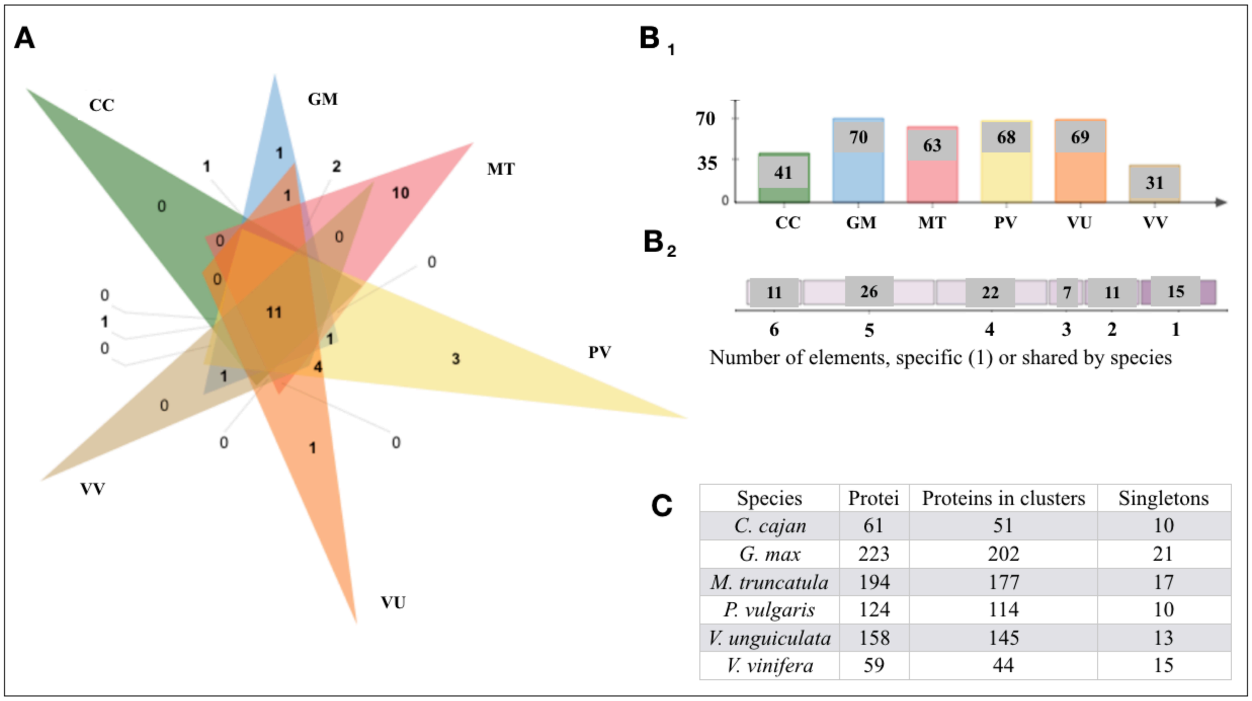


| Species | Proteins | Proteins in clusters | Singletons |
|---|---|---|---|
| *C. cajan* | 450 | 366 | 84 |
| *G. max* | 1864 | 1394 | 470 |
| *M. truncatula* | 1060 | 260 | 354 |
| *P. vulgaris* | 842 | 706 | 124 |
| *V. unguiculata* | 992 | 818 | 174 |
| *V. vinifera* | 443 | 260 | 183 |

**Figure 5. Summary of the RLK orthology analysis.**
**A**. Venn diagram showing the distribution of shared gene families (orthologous clusters) among CC, GM, PV, MT, VU, and VV. **B1**. The numbers refer to all the clusters in the species, including orthologs and in-paralogs **B2**. Distribution of the number of species present in orthologous clusters, one or share elements among species. **C**. Summary of the total number of proteins, clusters, and singletons within each species.

### ***Orthology of RLK-nonRD***

The results for the RLK-nonRD were included in the RLK orthology analysis (Figure 5), primarily to expose their distribution among CC, GM, PV, MT, VU, and VV; the RLK-nonRD are shown isolated in Figure 6. The RLK-nonRD formed 92 orthologous and paralogous clusters, 77 orthologous clusters containing at least two species, and two single-copy gene clusters. In total, 11 orthologous clusters were identified that were shared by all six species, the outgroup included. PV, GM, MT, and VU showed unique orthologous clusters; notably, the unique clusters identified in MTR were diverse, each containing two clusters with eight proteins; the next two clusters with five and three proteins, respectively; and the remaining [number] clusters with two proteins (Figure 6: A and B2). Of those remaining clusters, 66 were shared by at least two legume species, and 13 were shared by all five legumes species. *G. max* was the species with the most singletons (Figure 6: C) and proteins present in orthologous clusters (Figure 6: B1). The clusters reported as unique were formed by paralogous or protein isoforms belonging to the same gene (Figure 6: A and B2). The co-orthologs (recent descent and duplication), in-paralogs (recent duplication), and orthologs (recent descent) were calculated. The orthology results for *V. radiata*, *V. angularis*, *A. thaliana,* and *S. lycopersicum* are included in Figure A2.



| Species | Protei | Proteins in clusters | Singletons |
|---|---|---|---|
| *C. cajan* | 61 | 51 | 10 |
| *G. max* | 223 | 202 | 21 |
| *M. truncatula* | 194 | 177 | 17 |
| *P. vulgaris* | 124 | 114 | 10 |
| *V. unguiculata* | 158 | 145 | 13 |
| *V. vinifera* | 59 | 44 | 15 |

**Figure 6. Summary of the RLK-nonRD orthology analysis.**
**A**. Venn diagram showing the distribution of shared gene families (orthologous clusters) among CCA: *C. cajan*, GMA: *G. max*, PVU: *P. vulgaris*, VVI: *V. vinifera* "outgroup," VUN: *V. unguiculata*, and MTR: *M. truncatula*. **B1**. The numbers refer to all the clusters in the species, including orthologs and in-paralogs. **B2**. Distribution of the number of species present in orthologous clusters, elements 1, or shared among species lists. **C**. Summary of the total number of proteins, clusters, and singletons within each species. The RLK-nonRD isoforms are included in this analysis.

### *Hierarchical clustering analysis of RLK*

In the RLK phylogenetic tree, the legumes/non-legumes (8,278 proteins) and the experimentally-validated RLP (64 proteins) (Table A1) were included. The branches of the tree were labeled according to the type of function related to the ectodomains previously classified in Chapter 2. The tree allowed clustering patterns among the ectodomains present in the RLK to be identified (Figure 7). Among the 10 species, most RLK-nonRD proteins were clustered in 5 groups (Figure 7: A) related to LRR, LysM, G-Lectin, WAK/EGF, and RLCK-type proteins. The proteins with the C-lectin and stress-antifung domains were located in one cluster (Figure 7). The RLCK proteins were distributed in different clusters around the tree. The RLK tree was partially resolved and showed a mixture of resolve nodes and polytomies.

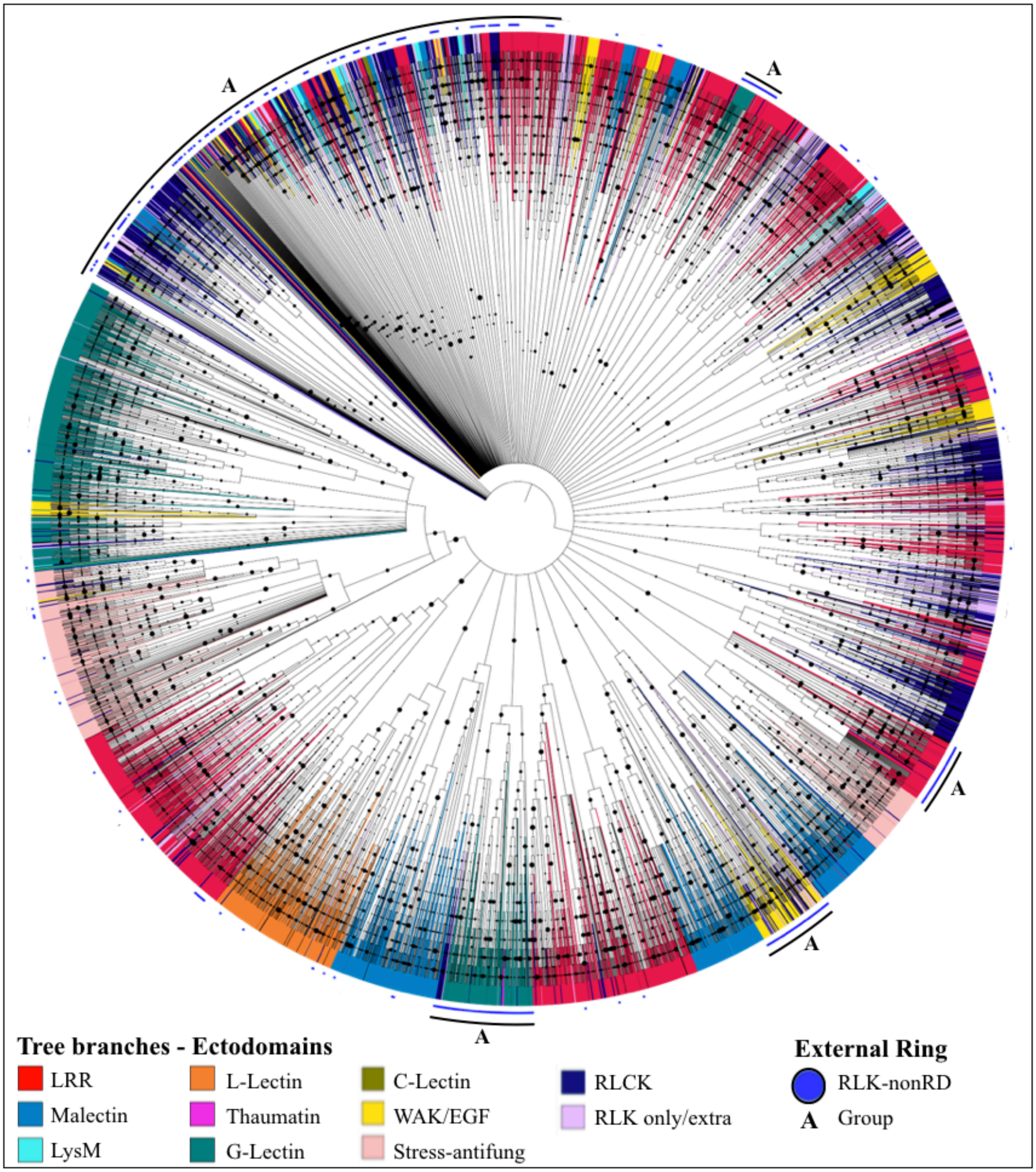
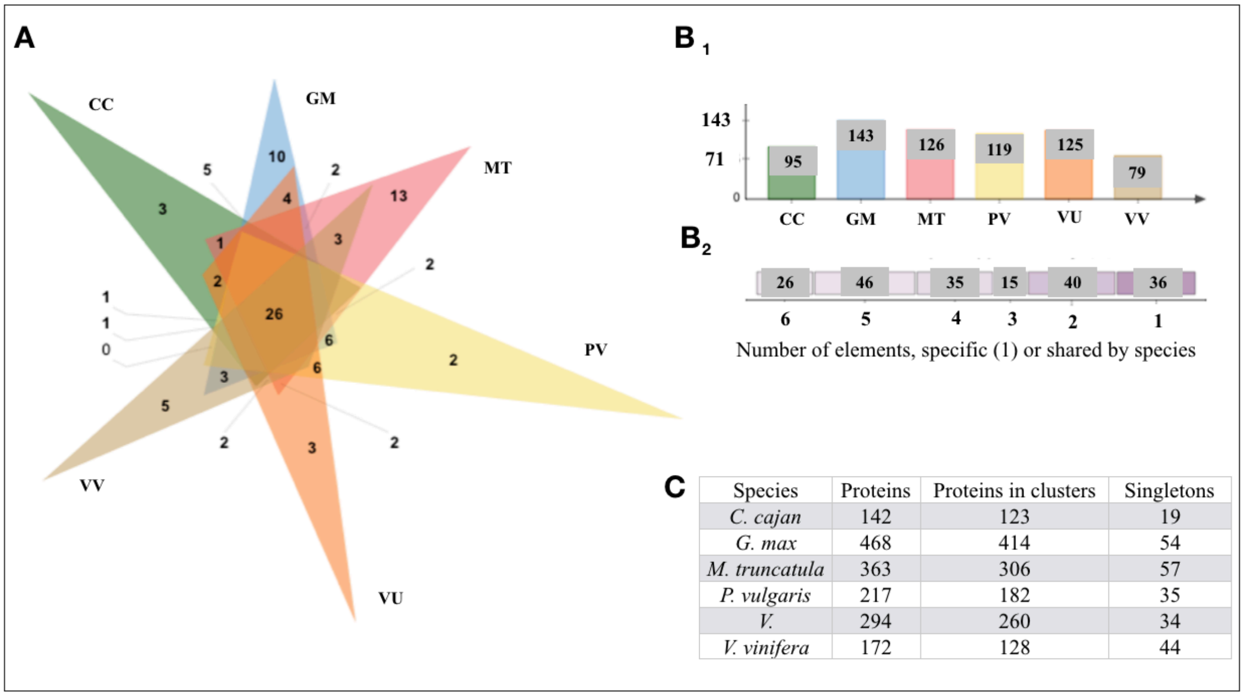


**Figure 7. RLK hierarchical tree.**
Phylogenetic tree based on protein similarity with ectodomains annotated. In the tree, the regions labelled A indicate the group location of RLK-nonRD. Black dots in the tree branches show a confidence value with a bootstrap > 50.

### ***Orthology and hierarchical clustering domain analysis of RLP***

The orthology and hierarchical clustering domain analysis of RLP for the species CC, GM, PV, MT, VU, and VV resulted in the formation of 198 orthologous and paralogous clusters, 162 orthologous clusters containing at least two species, and one single-copy gene cluster for the species CC, GM, PV, MT, VU, and VV. In total, 26 orthologous clusters were identified among the six species, the outgroup included. All species showed unique clusters (Figure 8: A and B2). The remaining 136 orthologous clusters were shared by at least two legume species, and 21 orthologous clusters were shared by all five-legume species. *M. truncatula* was the species with the most singletons overall (Figure 8: C), with *G. max* the species with the most proteins present in orthologous clusters (Figure 8: B1). The clusters reported as unique were formed by paralogous or protein isoforms belonging to the same gene (Figure 8: A and B2). The co-orthologs (recent descent and duplication), in-paralogs (recent duplication), and orthologs (recent descent) were also calculated. The orthology results for *V. radiata*, *V. angularis*, *A. thaliana,* and *S. lycopersicum* are included in Figure A3.



| Species | Proteins | Proteins in clusters | Singletons |
|---|---|---|---|
| *C. cajan* | 142 | 123 | 19 |
| *G. max* | 468 | 414 | 54 |
| *M. truncatula* | 363 | 306 | 57 |
| *P. vulgaris* | 217 | 182 | 35 |
| *V.* | 294 | 260 | 34 |
| *V. vinifera* | 172 | 128 | 44 |

**Figure 8. Summary of the RLP orthology analysis.**
**A**. Venn diagram showing the distribution of shared gene families (orthologous clusters) among CC, GM, PV, MT, VU, and VV. **B1**. The numbers refer to all the clusters in the species, including orthologs and in-paralogs. **B2**. Distribution of the number of species present in orthologous clusters, elements 1, or shared among species. **C**. Summary of the total number of proteins, clusters, and singletons within each species.

### *Hierarchical clustering analysis of RLP*

The legumes/non-legumes (2,445 proteins) and the experimentally-validated RLP (26 proteins) were included in the RLP phylogenetic tree (Table A1). The branches were labeled following the type of function related to the ectodomains previously classified in Chapter 2. Two main blocks of clusters were identified: block A with 1,280 proteins, which contained almost all the RLP-LRR proteins previously classified, and block B with 1,263 proteins, which comprised the EGF/WAK, G-lectin, L-lectin, LysM, malectin, thaumatin, and stress-antifung clusters and the remaining RLP-LRR proteins not classified in block A. The RLP tree was partially resolved and showed a mixture of resolve nodes and polytomies in both blocks (Figure 9-Blocks A and B). In block B, thaumatin and stress-antifung were independently clustered, but with the lectins

(G-lectin, malectin, L-lectin, and LysM), the clusters were assorted. Mixed protein clusters mixing proteins with the LRR and malectin domains were also identified (Figure 9: Block B).

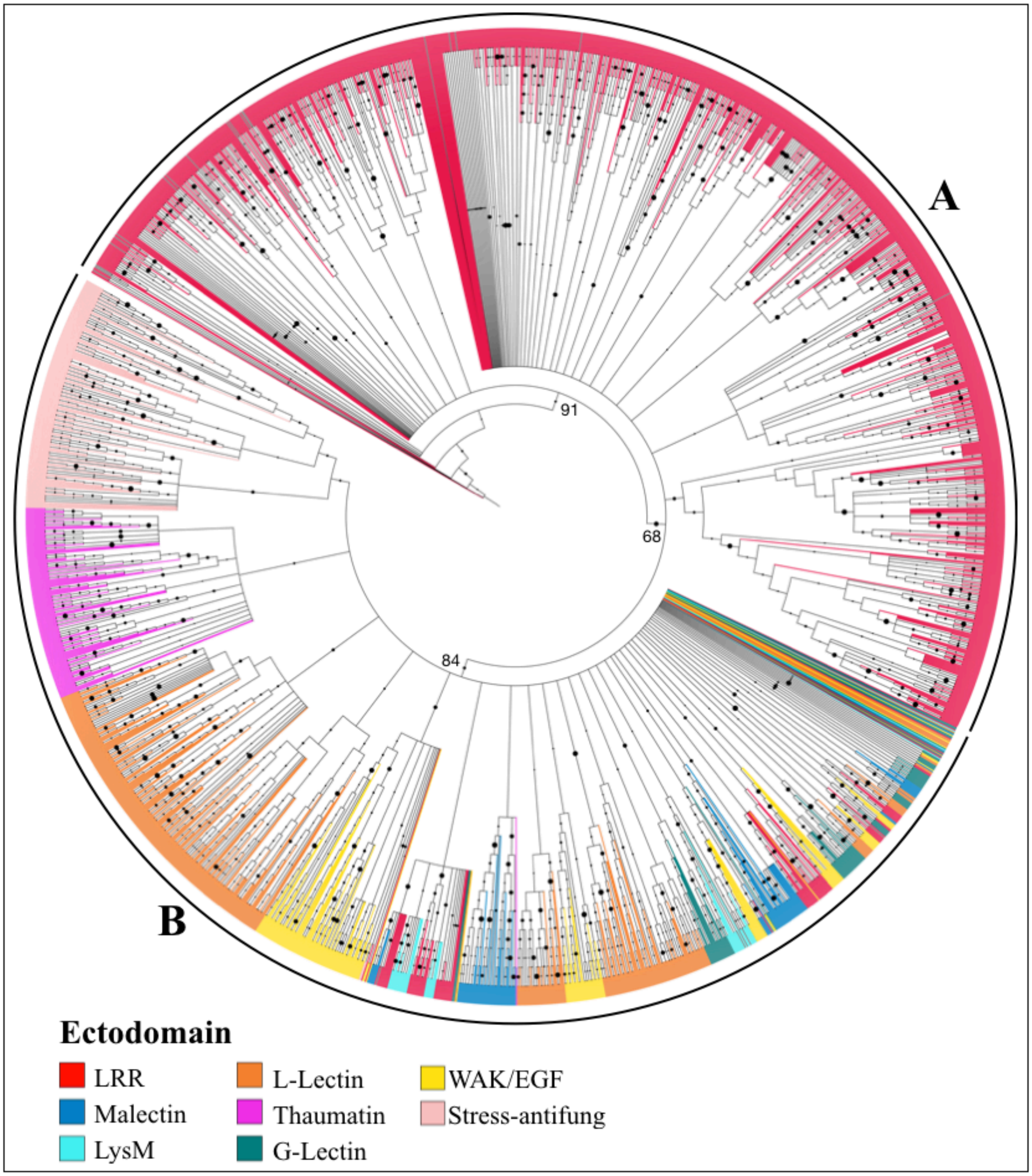


**Figure 9. RLP hierarchical tree.**
Phylogenetic tree based on protein similarity with ectodomains annotated. The tree is divided into two main blocks: block A, containing the RLP proteins with LRR domains, and block B, comprising the WAK/EGF, G-lectin, L-lectin, LysM, malectin, thaumatin and stress-antifung domains. Black dots in the tree branches show a confidence value (Bootstrap > 68) and tree branches with bootstrap < 50 were deleted (Table A5).

## Discussion

The distribution of RLK and RLP orthologous gene clusters showed that almost all evaluated proteins belong to orthologous clusters rather than a single-gene cluster. Only eight singleton clusters were discovered. This outcome suggests that WGD could contribute to the increased number of orthologous genes for RLK and RLP, corresponding to previous results reported for the identification of disease-resistant genes, or cytoplasmic R genes, in the legume family [111]. However, in order to create gene sets, single-copy gene families were identified in this study by counting the number of representatives of each species in a family. The process of identification of these families was complex due to issues with genome completeness or annotation, [141], and required high-quality genomic data to obtain reliable results. Further, the proteins allocated to the orthologous and paralogous clusters could have been redundant due to the presence of protein isoforms.

In the evolution of higher eukaryotes, WGD followed by diploidization and fractionation, including the loss of many redundant gene duplicates, has been a recurrent process [142]. Due to recent duplications among legumes (Figure 4), a high proportion of retained WGD genes in prior studies has been reported for the papilionoids. With the extra WGD of *G. max*, a higher proportion of retained genes are present in this species compared to the other legume species. The results obtained in the orthologous clustering analysis reported for RLK (Figure 5) and RLP (Figure 8) among legumes in this study showed a high proportion of duplicated proteins that belong to multiple orthologous clusters compared with the singletons proteins. Interestingly, these lineage-specific duplications increase the diversity of protein families among lineages and are often important for adaptation, especially for plants [117]; the results reported for the RLK and RLP here are part of this diversification process.

The diversity of the plasma membrane receptor results [based on the 427 and 136 orthologous clusters, seven and one single-copy gene clusters, 94 and 36 paralog clusters of RLK (including RLK-nonRD and RLCK) and RLP, respectively, and singletons identified by species among the legumes (Figures 5, 6, and 8)] suggest different forces and mechanisms of the evolutionary process [143]. These forces and mechanisms are modulated by the evolutionary rate effect of gene duplication between orthologs that have paralogs (duplicates) evolving significantly slower than singletons [144]. Further, duplicate and singleton genes have significantly different sequence properties, expression patterns, molecular functions, and biological roles [145]. The expansion of the RLK gene-family in plants was hypothesized to have accelerated evolution among domains implicated in signal reception, particularly with the extra- or intracellular LRR domain. Under this expansion, the gene-family represents a plant-specific adaptation that leads to the production of numerous and variable cell surface and cytoplasmic receptors [23], such as FLS2, FLS3, XPS1, EFR, and Xa21, all of which belong to the RLK-LRR-XII sub-family, which itself has undergone significant gene expansion [146]. However, given that the receptor configuration must arise from a fusion between an RLP and RLCK, it is plausible that these RLK with innate immunity functions were originally RLP and RLCK that fused together later on [98].

The RLK phylogenetic tree was comprised of 11 Pkinase ectodomains (Ins_P5_2-kin, RIO1, Pkinase, PI3_PI4_kinase, Pkinase_Tyr, Choline_kinase, ABC1, Pkinase_C, PIP5K, WaaY, and APH). The RLK-nonRD also included the presence of RLCK, with 5 main groups reported in Figure 7. Four out of five groups are highly dense, suggesting that further motif analysis could reveal hidden associated patterns, which are relevant because of their association with innate immune receptors. For the phylogenetic tree of all the RLP (Figure 9), two main blocks were identified; almost 99% of the LRR belonged to block A; in block B, LRR were present, but only as part of the LRR/malectin combination. Based on the distribution of ectodomains in the RLK (Figure 7) compared with the RLP tree (Figure 9), the ectodomains in the RLK tree suggest the presence of fusion patterns among the proteins. Further analysis could reveal how the mixed patterns occurred.

Over time, RLK and RLP have been exposed to a complex evolutionary process due to gene duplication and loss in plants [117]. The orthologous clustering process allowed eight single-copy gene clusters shared by MT, VU, PV, CC, GM and VV to be identified (Table A10). Notably, the single-copy genes did not comprise a random segment of the genome, but rather were highly conserved across plants and typically involved in essential housekeeping functions [147]. That a set of genes was consistently restored to single-copy status [144] among plasma membrane proteins RLK and RLP despite the large number of duplication events within the angiosperm lineage, including the legumes, was interesting. Notably, such single-copy genes have been recognized as molecular markers for inferring relationships of unresolved lineages [148]. In fact, recent research reported an optimal seed plant phylogeny resolution, but requires more than 100 single-copy genes [149].

## Conclusions

The dynamic evolution of RLK and RLP in the legume family is evidence of a complicated history of gene duplication and loss, previously described for these types of receptors [117] in relation to WGD events. Regarding gene-family expansion, the LRR-RLK/RLP comprised more than 60% of the plasma membrane receptors evaluated in the legumes in this study, which is in agreement to previous reports among other plants. This was also demonstrated by a high proportion of legume proteins belonging to orthologous groups, but at the same time having singleton clusters among their plasma membrane receptors. In the phylogenetic trees, all RLK and RLP showed multiple ectodomain conformations and combinations, which suggests that probable domain fusions could also be occurring. However, the results suggested that RLP are less diverse and more conserved compared to RLK. Further analysis must be done evaluating independently the plasma membrane proteins by functional domains to understand the domain topology and structural organization in detail.

# CHAPTER 4. SYNTENY ANALYSIS OF RLK AND RLP IN LEGUMES

## Abstract

### *Background*

Chromosomal synteny analysis is important in comparative genomics because it reveals the genomic evolution of related species. Shared synteny describes genomic fragments among chromosomes from different species that originated from a shared ancestor. Legumes are an important plant model because of their protein-rich physiology; their fundamental role in nitrogen-fixation symbiosis and mycorrhization, among other processes. Also, legumes possess unique features not found in typical plant models, such as *Arabidopsis*. They are also a good model for evaluating biotic/abiotic stimuli perceptions and plant-microbe interactions. Further, evaluating the presence of RLK/RLP among synteny blocks allows an increased understanding of the physical position and chromosomal distribution in related species.

### *Results*

The synteny block analysis in this study first calculated the whole set of blocks among the legumes/non-legumes species evaluated, targeted the blocks with the presence of the RLK and RLP previously predicted in Chapter 2, and also identified within the synteny blocks the presence of resistance plasma membrane receptors that have previously been reported by other authors. In total, one-third of the synteny blocks identified had RLK/RLP genes among the legumes *Glycine max*, *Phaseolus vulgaris*, *Medicago truncatula*, *Vigna angularis*, *Vigna radiata*, *Vigna. unguiculata*, and *Cajanus cajan* and the outgroup set *Arabidopsis thaliana*, *Solanum lycopesicum*, and *Vitis vinifera*. Among the legumes evaluated, between 75% and 98% of the RLK/RLP were present in synteny blocks. The distribution of chromosomal segments between *P. vulgaris* and *V. unguiculata* (chr 4, 7, 9, 10, and 11) was highly similar. Among the RLK/RLP synteny clusters, a set of experimentally-validated resistance RLK/RLP genes was identified in

synteny. The RLK resistant genes FLS2, BIR2, ERECTA, IOS1, and SISERK1 from *Arabidopsis* and SISEK1 from *S. lycopersicum* were present among different pairwise synteny blocks within the legume species *G. max*, *P. vulgaris*, *M. truncatula*, *V. angularis*, *V. unguiculata*, and *V. radiata*. Meanwhile, only the LYM1- RLP resistant gene from *Arabidopsis* was present in synteny blocks within *G. max*.

### *Conclusions*

The identification of RLK and RLP genes among the synteny blocks in legumes showed that multiple genome segments were highly conserved along different chromosomes. That condition is interesting because plasma membrane receptors do not represent more than 3% of the total number of genes per species. The ratio of the pairwise synteny blocks of RLK/RLP among legumes showed a 1:1 relationship; the exception was *G. max,* which had approximately a 2:1 ratio, demonstrating a similar proportion of plasma membrane proteins among the pairwise clusters. These results provide an overview of the structural organization and distribution of the RLK/RLP among legumes and add to the annotation record of the genes present or absent in synteny blocks.

**Keywords:** Dicots, target synteny blocks, plasma membrane genes, legumes/non-legumes.

### Background of synteny in legumes

After Orchidaceae and Asteraceae, the Fabaceae family of legumes is considered the largest flowering plant family. It is ecologically important because most species in the family fix atmospheric nitrogen through symbiosis [150] and agriculturally important as legumes are major food crops and used for forage and green manure. In fact, soybean (*Glycine max* (L) Merrill), common bean (*Phaseolus vulgaris* L.), peanut (*Arachis hypogaea* L.) and chickpea (*Cicer arietinum* L.) together account for more than 20% of primary crop production worldwide [151]. Due to its agricultural importance and ability to fix nitrogen through symbiosis with rhizobial bacteria, Fabaceae is one of the most studied plant families. It is also one of the largest, comprising 19,000 species and 727 genera, divided into three subfamilies: Papilionoideae, Caesalpinioideae, and Mimosoideae [152].

Synteny analysis is a useful strategy to investigate evolutionary relationships and to identify genes that could be functionally related [153]. Syntenic blocks are defined as blocks of sequences that exhibit conserved gene order across genomes [154]. The identification of homology supports synteny analysis within or across genomes. Homologous genes are classified as orthologous (speciation) or paralogous (gene duplication) [155]. Interestingly, structural homologies can be evaluated at the micro- or macrosynteny level. Microsynteny analysis evaluates continuous and short regions, while macrosynteny analysis focuses on whole genomes [156]. Recently, synteny comparisons between closely-related eukaryotic species determined that homologous genes remained on corresponding chromosomes, demonstrating synteny [17]. Today, a common strategy to infer function from homology is directly related to ortholog identification [157]. Any synteny definition must use homology or orthology, and most synteny tools today use both: homology as a matter of principle and orthology as a result of practical constraints [154].

In contrast to highly-related species, synteny conservation offers details about genomic segments that are conserved at the chromosomal level, where orthologous gene order is maintained [158]. The core of comparative genomics lies in how to compare genomes to reveal species' evolutionary relationships. Comparisons of genomic synteny between and within species have provided an opportunity to study evolutionary processes that lead to a diversity of chromosome number and structural lineages across multiple species. Interestingly, many tools use orthologous relationships between protein-coding genes as anchors to position statistically significant local alignments [159]. In the end, the identification of syntenic regions between non-legume and legume species is an efficient strategy for identifying patterns of evolutionary conservation and divergence across genomes [12].

Among legumes, it has been reported that macrosynteny in species such as *M. truncatula* and *G. max* can be as long as the chromosome arms or span most of the euchromatin region of the two genomes. Each *M. truncatula* region and its homeologue typically show similarity to three *V. vinifera* regions via the pre-rosid whole genome hexaploidy [3]. Within the millettioid clade, the species pigeonpea (*C. cajan*) diverged from the soybean species~20–30 Mya. Interestingly, after this long period of divergence, high levels of synteny are observed between the two species [6]. Each pigeonpea chromosome shows extensive synteny with two or more soybean chromosomes, likely due to an independent soybean duplication event [1]. Also, the genome comparison of *V. radiata* var. *radiata* with *A. thaliana*, *Cicer arietinum*, *C. cajan*, *G. max*, *Lotus japonicas,* and *M. truncatula* revealed well-conserved macrosynteny blocks, although these blocks were highly dispersed among plant species with different numbers of chromosomes [4].

To understand the structural relationships between the common bean and soybean genome, McClean *et al*. in 2010 [153] identified gene-rich regions for all soybean chromosomes within precise regions of the common bean genetic map. The research concluded that, relative to common bean, soybean is segmentally rearranged, exhibiting evidence of a one-to-two relationship, respectively [153]. Among the *Vigna* genus, cowpea (*V. unguiculata*) shares a high degree of collinearity with *P. vulgaris* [160]. Muñoz-Amatriaín *et al.* in 2017 explored the genetic diversity along each linkage group among *V. unguiculata* and *P. vulgaris* and found the groups to have macrosynteny [161]. In contrast, given the close relationship of *Vigna* to *Glycine*, most of the *V. radiata* var. *radiata* genes were found in synteny to *G. max*. Of the 18,378 genes on pseudo-chromosomes, 14,569 were located in 1,059 synteny blocks of orthologues or paralogues [4]. It was also reported that 11,853 mungbean genes were in synteny with the *C. cajan* genome [4].

A synteny block is defined as a genomic region spanning a number of genes that are orthologous and colinear with another genome, without consideration for orientation and with the minimum number of co-arranged orthologs defined following a microsynteny approach. These parameters, however, vary among different studies [159]. This inter-species synteny analysis estimated the whole set of synteny blocks among seven species of legumes (*P. vulgaris*, *G max*, *C. cajan, M. trucatula, V. unguiculata, V. radiata,* and *V. angularis)*, including three outgroups (*V. vinifera*, *S. lycopersicum*, and *A. thaliana*), and targeted only the blocks with the presence of RLK/RLP predicted in Chapter 2. The purpose of this analysis was to evaluate the distribution, conservation, and divergence of the pairwise synteny blocks associated with RLK/RLP. It also used the experimentally-validated RLK/RLP resistance genes (Table A1) to target synteny blocks. The analysis evaluated the chromosomal segment distribution of syntenic blocks with RLK/RLP among the species to identify patterns of evolutionary conservation and divergence and also used *P. vulgaris* chromosomes as a reference model for the comparison of RLK/RLP synteny blocks among the legume/non-legume species.

## Methods

### *Datasets*

Different datasets were used and adjusted for the synteny analysis. The whole protein dataset and the gene annotation for each genome of the evaluated species were collected. That dataset was used as an input to build a blast database using only the genes allocated to chromosomes “Chr.” The genes present in the chloroplast chromosome (ChrC), mitochondria chromosome (ChrM), and scaffolds were excluded. Two genomic databases were used to obtain the legume/non-legume genomes: the NCBI database for three of the species and the Phytozome repository database for the other seven species (Table 10).

**Table 10. Summary of genomes**

| Species | Database | File name | N. of genes | N. of proteins | N. of chr |
|---|---|---|---|---|---|
| *V. radiata* | NCBI | GCF_000741045.1_Vradiata_ver6 | 34,911 | 35,143 | 11 |
| *C. cajan* | NCBI | GCA_000340665.1_C.cajan_V1.0 | 23,374 | 48,331 | 11 |
| *V. angularis* | NCBI | annotation release 100 | 22,276 | 37,769 | 11 |
| *G. max* | Phytozome | gmax_275_wm82.a2.v1 | 55,589 | 88,647 | 20 |
| *M. truncatula* | Phytozome | Mtruncatula_285_Mt4.0v1 | 48,338 | 62,319 | 8 |
| *P. vulgaris* | Phytozome | Pvulgaris_442_v2.1 | 27,012 | 36,995 | 11 |
| *V. unguiculata* | Phytozome | Vunguiculata_469_v1.1 | 28,881 | 42,287 | 11 |
| *A. thaliana* | Phytozome | Athaliana_167_TAIR10 | 27,206 | 35,386 | 5 |
| *S. lycopersicum* | Phytozome | Slycopersicum_390_ITAG2.4 | 33,838 | 34,725 | 12 |
| *V. vinifera* | Phytozome | Vvinifera_145_Genoscope.12X | 23,647 | 26,346 | 19 |

The RLK and RLP predicted in Chapter 2 (Table 4) were used as a reference to target the synteny blocks initially calculated, with only the blocks with plasma membrane receptors evaluated further. The target blocks were used to evaluate synteny of RLK/RLP in legume/non-legume species following a comparative genomic approach, which also allowed patterns of conservation and/or divergence among and between them to be identified. At the same time, the resistance RLK/RLP proteins were also used as a reference to track the presence of synteny blocks containing experimentally-validated RLK and RLP among legumes/non-legumes. This approach does not necessarily imply that the presence of an experimentally-validated RLK/RLP protein (Table A1) must be shared among species, but if the interspecies synteny blocks were shared, the match was reported.

### *Interspecies identification of synteny blocks*

The database and the blastp for the calculation of the synteny block input were built using a ncbi-blast-2.7.1+ package (makeblastdb and blastp). The input used for the makeblastdb script was the whole set of proteins reported for the legume/non-legume species. The parameters for the blastp were blastp -outfmt 6 -evalue 1e-10 -max_target_seqs 5. The output obtained from the blast process and the GFF annotation of the 10 species (seven legumes/three non-legumes) were used as the input for the synteny blocks calculation. The interspecies syntenic blocks were calculated using an MCScanX tool [162] and the following parameters: match-score, final score = match_score + num_gaps * gap_penalty (default: 50); gap-penalty, gap penalty (default: -1); match-size, the number of genes required to call a collinear block (default: 5); E-value, alignment significance 1e-10; max-gaps, maximum gaps allowed (default: 25); and overlap-window, maximum distance 10,000 (number of nucleotides among genes) to collapse blast matches (default: 5) and the patterns of collinear blocks: 1 inter-species. Also, the script dissect_multiple_alignment was used to subset all results and obtain a reference among the species compared. For the figures, the circle_plotter and bar_plotter that come with the MCScanX package were employed. After the whole set of synteny blocks were calculated, in-house scripts were developed to subset the MCScanX collinearity output file, isolating the synteny blocks with RLK/RLP.

With the goal of identifying synteny blocks among the species with identical and/or highly identical resistance RLK/RLP blocks, an identity clustering analysis was applied. The analysis compared the predicted RLK/RLP reported in Chapter 2 (Table 3) against the experimentally-validated resistance RLK/RLP (Table A1) using the CD-HIT [92] tool. Specifically, the script “cd-hit-2d” with the parameters -c 0.9 -n 5 was applied. Approximately 75% of the experimentally-validated resistance RLK/RLP that belong to *Arabidopsis* and *S. lycopersicum*. This approach allowed the proteins that share more than 90% of their identity to be determined. The identical and highly identical resistance RLK/RLP proteins were used to identify synteny blocks in the non-legume species *Arabidopsis* and *S. lycopersicum* that were shared among the legumes. In-house scripts (https://github.com/drestmont/plant_rlk_rlp/) were used to identify the presence of resistance genes among the synteny blocks. The whole process required 10 servers (each one with 16 core and 32 GB ram) running in parallel for 2 weeks. BLASTP used 95% of the computational time. The process was run at the scientific cluster at Universidad Nacional de Colombia.

## Results

The entire process of synteny block estimation reported a total of 690,397 matches imported (853,808 discarded), 6,252 pairwise comparisons, and 9,011 alignments or pairwise clusters, and 3,592 alignments with RLK/RLP proteins. These results represent the whole set of synteny blocks shared among the legumes/non-legumes evaluated (two or more species could share a pairwise synteny block that had at least five genes shared with an E-value > 1e-10 in a maximum range of 10,000 nucleotides). The whole synteny block set was split using the RLK/RLP genes as a reference to target or identify the sets of synteny blocks with the presence of plasma membrane proteins. The relation between the genes initially processed, the RLK/RLP identified in the chromosomes, and the final set of genes belonging to different synteny blocks is reported in Table A11.

The presence and distribution of RLK/RLP genes in the interspecies synteny blocks and the identification of the plasma membrane proteins and their general distribution among the species are shown in Table 10. In most cases, the number of legume/non-legume genes belonging to one or more synteny blocks per species was higher compared with those genes that do not belong to the blocks. The exceptions belong to the non-legumes for the RLK gene AT, for the RLK-nonRD genes AT and SL, and for the RLP genes VV, AT, and SL. All legumes (CC, GM, MT, PV, VA, VU, and VR) showed a higher proportion of genes involved in synteny blocks compared with RLK/RLP genes unrelated to the blocks. All RLK-nonRD genes present in the PV and VU genomes were in synteny blocks, and among legumes, the MT species had fewer RLK/RLPs present in blocks (Table 9).

The RLK/RLP gene frequency range described the number of times a gene could be present in different synteny blocks based on the pairwise comparison (Table 11). The RLK and RLP frequency range of genes in pairwise synteny blocks in the species comparison showed values between 1 and 13, with the exception of AT, which showed a low frequency range (1 to 3) in both plasma membrane classes. Interestingly, the RLK-nonRD that belong to the RLK class showed a frequency range among legumes/non-legumes with a higher proportion of genes in synteny blocks (approximately 1 to 9) compared with non-legumes (approximately 1 to 5), and fewer non-legume RLK-nonRD genes present in synteny blocks were shared compared with the legume RLK-nonRD genes (Table 10).

**Table 11. Summary of RLK and RLP genes among species in synteny blocks.**

| Species | RLK no blocks | RLK in blocks* | Freq. range | RLK-nonRD no blocks | RLK-nonRD in blocks* | Freq. range | RLP no blocks | RLP in blocks * | Freq. range |
|---|---|---|---|---|---|---|---|---|---|
| CC | 38 | 166 | 1 to 9 | 8 | 21 | 1 to 7 | 16 | 50 | 1 to 8 |
| GM | 57 | 907 | 1 to 13 | 17 | 148 | 1 to 9 | 54 | 336 | 1 to 9 |
| MT | 116 | 525 | 1 to 9 | 88 | 95 | 1 to 9 | 154 | 164 | 1 to 7 |
| PV | 7 | 520 | 1 to 10 | 0 | 113 | 1 to 8 | 6 | 193 | 1 to 8 |
| VA | 8 | 427 | 1 to 9 | 3 | 78 | 1 to 7 | 13 | 136 | 1 to 8 |
| VR | 8 | 329 | 1 to 9 | 3 | 62 | 1 to 7 | 12 | 141 | 1 to 9 |
| VU | 8 | 563 | 1 to 10 | 0 | 142 | 1 to 8 | 6 | 257 | 1 to 9 |
| VV | 75 | 259 | 1 to 10 | 21 | 37 | 1 to 5 | 87 | 72 | 1 to 8 |
| AT | 333 | 96 | 1 to 3 | 35 | 10 | 1 to 2 | 119 | 28 | 1 to 3 |
| SL | 145 | 245 | 1 to 7 | 47 | 36 | 1 to 3 | 91 | 70 | 1 to 7 |

The “Frequency range” column describes the number of times a gene could be present in different synteny blocks among species. Because the synteny blocks were pairwise calculated and the same gene can be present multiple times, these values give a frequency reference. The RLK genes were split into two RLK and RLK-nonRD classes. *Gene numbers reported are non-redundant; however, a gene can be present in one or more synteny blocks.

### ***Species synteny analysis***

The identification of interspecies synteny blocks was calculated using a pairwise approach (MCScanX) to identify the blocks that have RLK and RLP. The genes previously predicted in Chapter 2 were used as a reference to select the blocks by species. The subset of synteny blocks of each species containing the plasma membrane proteins as a target did not automatically imply RLK or RLP transitivity, or a transitivity relation among the synteny blocks; at the same time, the presence of an RLK/RLP in one of the species did not automatically imply their presence in the other species. Different criteria were used to split the legume/non-legume synteny block comparison to give an overview of the results. Also, all sets included VV because the divergence of grape occurred early in eudicot evolution and allows the split to Papilionoids to be estimated, making it ideal for phylogenetic discrimination among legumes [109]. The four sets were: 1) PV, GM: because PV is considered a diploid model for GM [163]; 2) MT, CC: because MT is considered a cool-season legume model [164] compared with CC, which is considered an orphan legume crop [6]; 3) VR, VU, VA: because this can be used as a reference to compare the legume *Vigna* genus; 4) SL, AT, VV: because these sets correspond to the non-legumes species included and were subset as a reference to compare distribution, conservation, and divergence among the outgroups (Figures 10, 11, 12, 13, and 14).

### ***RLK among species synteny blocks***

Among the 10-species evaluated, a total of 3,049 pairwise alignments had the presence of RLK (RLK-nonRD excluded). The blocks were split by the presence of RLK, but could also have RLK-nonRD and/or RLP present. The pairwise ratios of RLK genes present in synteny blocks among species were: 843 GM to 496 PV, 258 GM to 157 VV, and 91 PV to 60 VV. Among the GM and VU legumes, a 2:1 gene ratio of RLK/RLP was found in synteny blocks, and the plasma membrane receptors were distributed in multiple regions among all chromosomes. In the GM and PV legume comparison to VV, a notable reduction (70% or less) of RLK/RLP genes in synteny blocks was reported, and the VV-Chr 5 did not share RLK/RLP synteny blocks (Figure 10:A). The pairwise gene ratios of RLK among the MT and CC legumes were: 93 MT to 64 CC, 46 MT to 31 VV, and 19 CC to 21 VV. The MT and CC legumes had approximately a 1:1 gene ratio of RLK/RLP shared in blocks and shared synteny fragments among almost all chromosomes, with the exception of CC-Chr5. The legume versus the outgroup did not show shared synteny blocks with CC-Chr3, 5, 9, and 15 (Figure 10:B).

For the pairwise gene ratio of RLK evaluated in synteny blocks among the *Vigna* genus (Figure 10:C), the identified plasma membrane receptors present in synteny blocks were: 292 VR to 257 VA, 324 VR to 438 VU, 43 VR to 39 VV, 430 VA to 490 VU, 35 VA to 51 VV, and 86 VU to 98 VV. The legumes in this comparison set followed a 1:1 pairwise gene ratio, and almost all RLK/RLP genes (Table 11) were in synteny and distributed fragments among all chromosomes. The pairwise ratio comparison of legumes against the outgroup show about a 90% reduction in RLK/RLP synteny. No synteny blocks were shared with VV-Chr 2, 3, 12, and 15. In contrast, the non-legume pairwise gene ratio reported in Figure 10:D shows 24 SL to 17 AT, 195 SL to 221 VV, and 46 AT to 84 VV. The RLK shared among SL and VV was proportionally higher compared with the other species evaluated in this study. All chromosomes for the non-legumes were reported to have RLK synteny blocks.

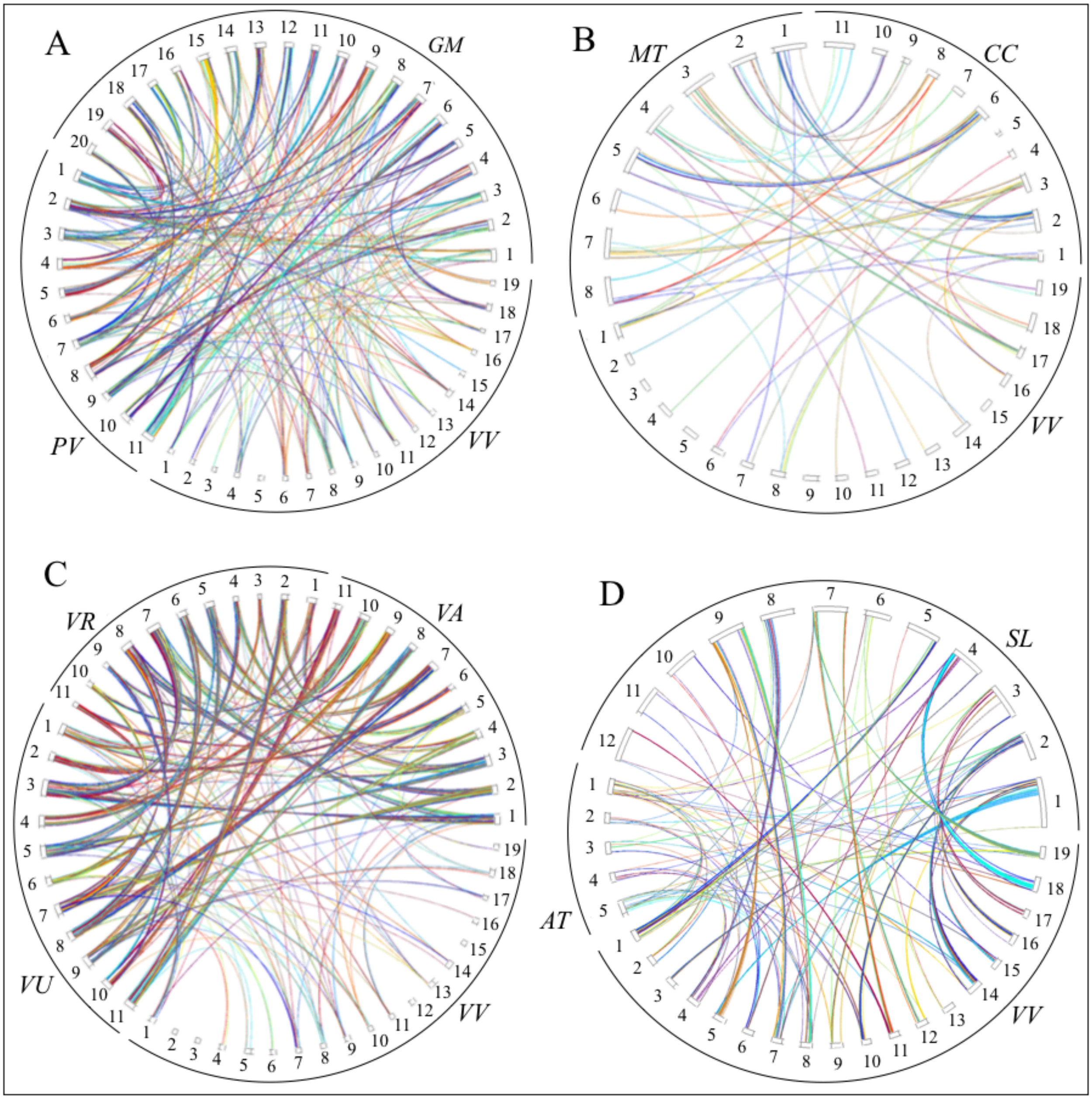


**Figure 10. Distribution of RLK present in synteny blocks.**
Chromosomes of the species evaluated. For visual purposes, the RLK identified in a synteny block were used as a reference to plot the circles. The RLK-nonRD were excluded in the figure. VV was included in all figures as an outgroup for legumes and also to compare results among AT and SL. A) *G. max* GM, *P. vulgaris* PV, and *V. vinifera* VV. B) *M. truncatula* MT, *C. cajan* CC, and VV. C) *V. radiata* VR, *V. angularis* VA, *V. unguiculata* VU, and VV. D) *A. thaliana* AT, *S. lycopersicum* SL, and VV.

***RLK-nonRD among legume/non-legume synteny blocks***

In the 10-species evaluated, a total of 715 alignments had the presence of RLK-nonRD. The predicted RLK-nonRD were used as a reference to target the synteny blocks. The alignments were not exclusive for the plasma membrane class and could also have the presence of RLK and/or RLP. The number of RLK-nonRD genes in a pairwise ratio among the synteny blocks were: 114 GM to 82 PV, 14 GM to 17 VV, and 5 PV to 17 VV. Among the GM and PV legumes, the RLK-nonRD ratio was 1:1, and all chromosomes had RLK/RLP genes present in synteny. In the legume/non-legume comparison, the proportion of RLK-nonRD genes in synteny was very low; also, 8 out of 19 chromosomes did not share synteny (Figure 11:A). The pairwise gene ratio comparisons among MT and CC and the non-legume VV were: 18 MT to 9 CC, 2 CC to 3 VV, and 5 VV to 4 VV. In relation to the other legumes, MT and CC showed the lowest number of RLK-nonRD genes in synteny and, technically, only six blocks were shared with the VV non-legume species (5 of 19 VV-Chr involved). Five chromosomes among the legume species were missing in the RLK-nonRD synteny blocks (Figure 11:B).

The RLK-nonRD pairwise gene ratios identified among the *Vigna* genus were: 44 VR to 40 VA, 56 VR to 94 VU, 1 VR to 0 VV, 68 VA to 11 VU, 4 VA to 2 VV, and 10 VU to 7 VV. The *Vigna* species showed a synteny distribution of RLK-nonRD among all the chromosomes, and only eight synteny blocks were shared with the non-legume VV; 11 out of 19 VV-Chr did not share synteny, (Figure 11:C). The non-legumes in Figure 11:D showed pairwise gene ratios of 4 SL to 2 AT, 24 SL to 24 VV, and 3 AT to 6 VV. Like with the RLK, the proportion of RLK-nonRD shared among SL and VV was higher compared with the other species evaluated in this study. Not all non-legumes reported RLK-nonRD synteny in all chromosomes.

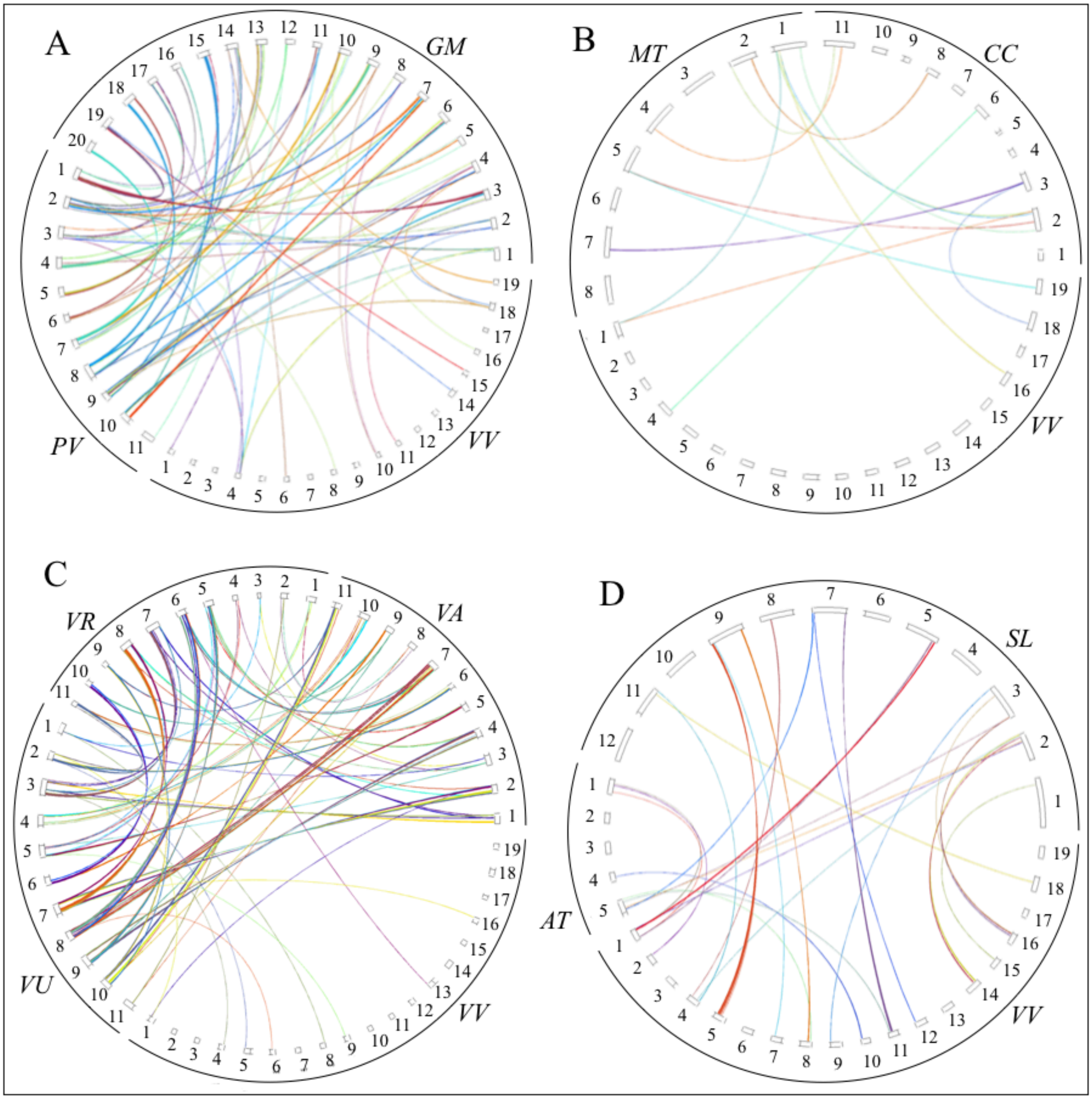


**Figure 11. Distribution of RLK-nonRD present in synteny blocks.**
Chromosomes of the species evaluated. For visual purposes, the RLK identified in a synteny block were used as a reference to plot the circles. VV was included in all figures as an outgroup for legumes and also to compare results among AT and SL. A) *G. max* GM, *P. vulgaris* PV, and *V. vinifera* VV. B) *M. truncatula* MT, *C. cajan* CC, and VV. C) *V. radiata* VR, *V. angularis* VA, *V. unguiculata* VU, and VV. D) *A. thaliana* AT, *S. lycopersicum* SL, and VV.

### ***RLP among synteny blocks***

Among the 10-species evaluated, a total of 1,361 alignments had the presence of RLP. The predicted RLP set was used as a reference to target the synteny blocks. The alignments were not exclusive for this plasma membrane class and could have also the presence of other RLK and/or RLP. The pairwise ratios of RLP genes identified among the synteny blocks in Figure 12: A were: 252 GM to 159 PV, 57 GM to 6 VV, and 11 PV to 1 VV. The RLP distribution among the GM and PV legumes involved fragments in all chromosomes. Like in the RLK ratio, the RLP had approximately a 2:1 ratio. The legume/non-legume ratio for RLP genes present in synteny blocks was low; only seven VV genes were in synteny compared with 57 GM and 11 PV genes. In total, four VV-chromosomes were not in synteny with any of the legume species (Figure 12:A). The pairwise ratio comparisons of RLP genes among the MT and the CC legumes were: 16 MT to 12 CC, 15 MT to 1 VV, and 5 CC to 1 VV. Among the MT and CC legumes, not all chromosomes shared RLP synteny blocks, and, compared with the non-legume species, 10 out of 19 VV-Chr did not share synteny (Figure 12:B).

The *Vigna* genus reported pairwise ratios of RLP genes in synteny of: 14 VR genes to 105 VA genes, 13 VR genes to 42 VU genes, 1 VR gene to 1 VV gene, 120 VA genes to 28 VU genes, 6 VA genes to 0 VV genes, and 0 VU genes to 4 VV genes. Once again, like with RLK, all *Vigna* chromosomes shared fragments of synteny with RLP, whereas with the non-legumes, nine out of 19 VV-chromosomes did not display synteny (Figure 12:C). The non-legumes presented in Figure 9: D showed RLP pairwise ratios of: 7 SL genes to 2 AT genes, 53 SL genes to 6 VV genes, and 14 AT genes to 2 VV genes. With the exception of three out of 36 chromosomes in total (SL-Chr 12 and VV-Chr 9 and 13), synteny fragments occurred among all non-legume species (Figure 12:D).

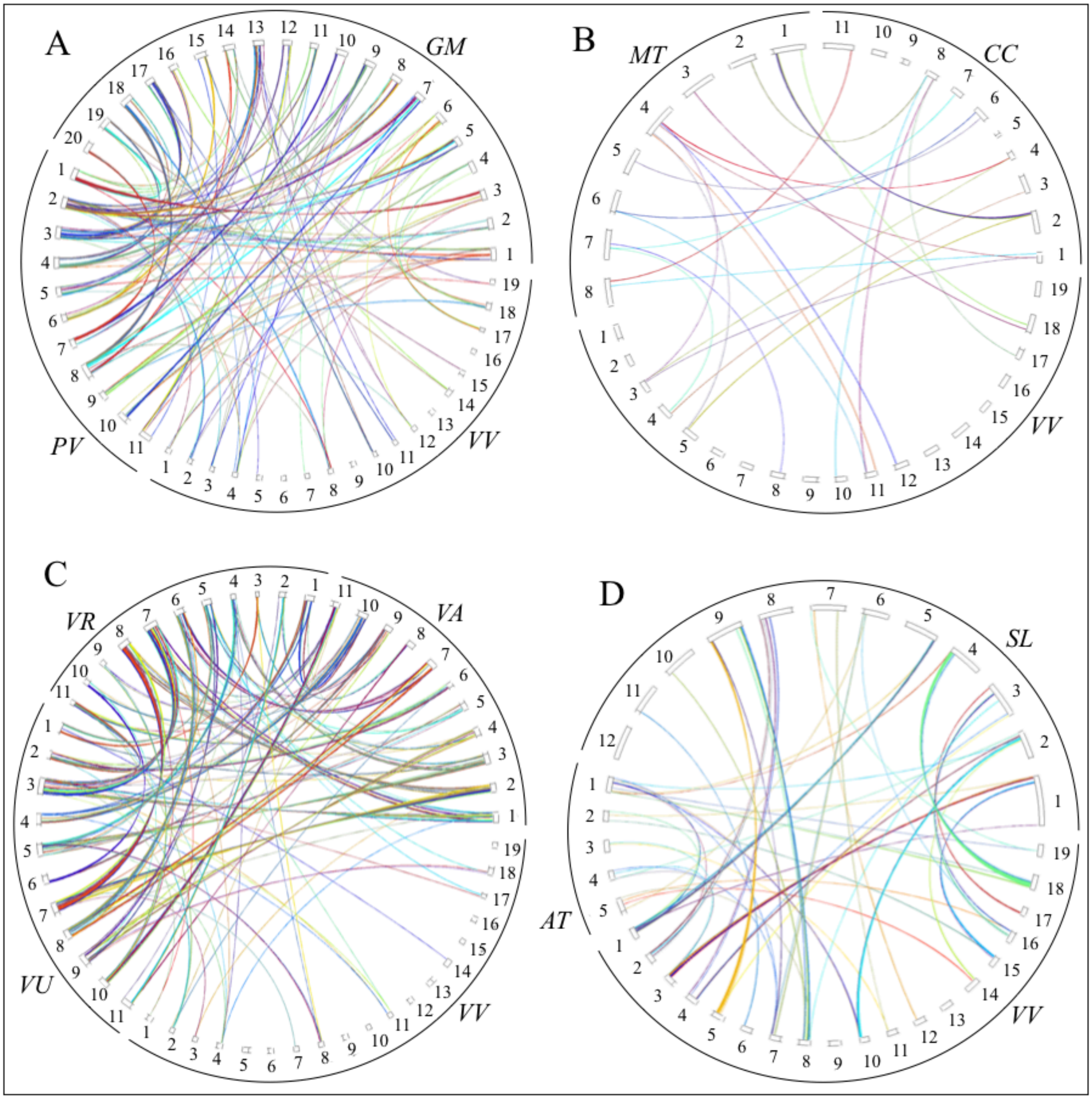


**Figure 12. Distribution of RLP present in synteny blocks.**
Chromosomes of the species evaluated. For visual purposes, the RLK identified in a synteny block were used as a reference to plot the circles. The RLK were excluded, and VV was included in all figures as an outgroup for legumes and also to compare results among AT and SL. A) *G. max* GM, *P. vulgaris* PV, and *V. vinifera* VV. B) *M. truncatula* MT, *C. cajan* CC, and VV. C) *V. radiata* VR, *V. angularis* VA, *V. unguiculata* VU, and VV. D) *A. thaliana* AT, *S. lycopersicum* SL, and VV.

### *P. vulgaris RLK and RLP synteny blocks as a model to compare the legume and non-legume species*

PV was used as a model to evaluate the RLK/RLP synteny block distribution among the legume/non-legume species. Fragments of the 11 PV chromosomes were distributed along 18 out of 20 GM-Chrs sharing more than two PV-Chrs fragments; the PV-Chr7 was only present in GM-Chr10 and Chr20. Nine of 11 CC-Chrs had more than two PV-Chrs fragments, and PV-Chr11 and PV-Chr2 only shared multiple fragments with CC-Chr4 and CC-Chr5, respectively. Seven out of eight MT-Chrs had more than two synteny blocks from different PV-Chrs, and the MT-Chr6 only had spread blocks that belonged to PV-Chr 4. The comparison of common bean to the *Vigna* species showed that nine out of 11 PV-Chrs shared more than two PV-Chrs synteny regions. The PV-Chr7 and PV-Chr9 matched only with long fragments of VA-Chr2 and VA-Chr4, respectively. All VR-Chrs shared between two to three synteny blocks from different PV-Chrs. The last comparison between PV and a *Vigna* species reported six out of 11 VU-Chrs sharing two long synteny blocks with PV-Chrs. The PV-Chr4, Chr7, Chr9, Chr10, and Chr11 showed a long fragment match with VU-Chr4, Chr7, Chr9, Chr10, and Chr11. The PV:VU chromosome distribution was notably similar. Also, the chromosome fragment evaluation between PV and the non-legumes had these results: nine out of 12 SL-Chrs shared small synteny regions with 10 PV-Chrs, and only PV-Chr5 was not present. Regarding AT, five out of five AT-Chrs shared small regions with PV-Chr1, Chr3, and Chr4. Finally, 15 out of 19 VV-Chrs shared small synteny fragments with PV-Chrs regions, and only PV-Chr11 shared any fragment with VV (Figure 13).

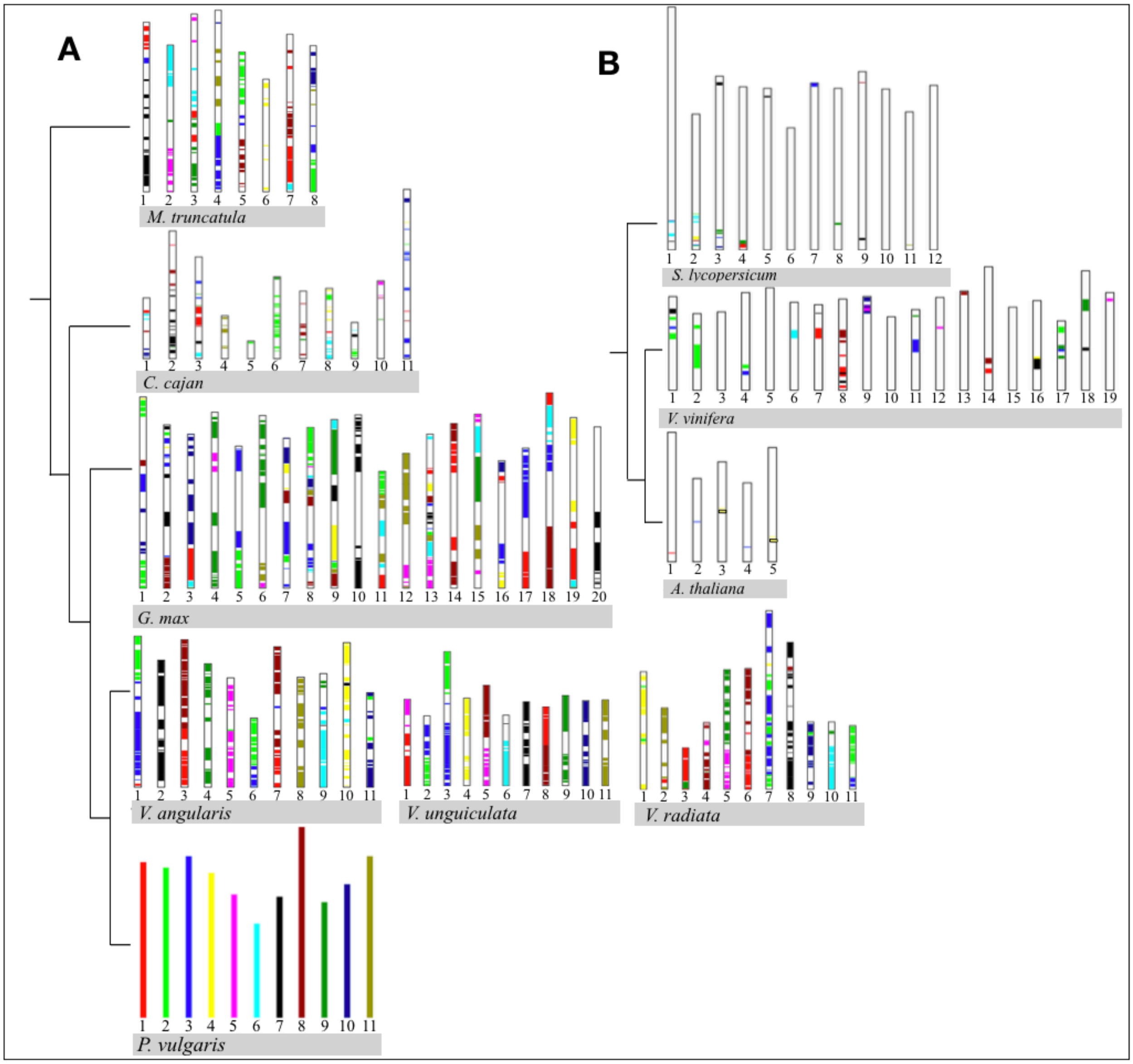


**Figure 13. Distribution of *P. vulgaris* chromosome fragments with RLK/RLP.**
A. Legumes and B. non-legumes. The PV blocks include all RLK/RLP present in legumes and also all other proteins associated with the synteny blocks as a comparison reference.

### *V. vinifera RLK and RLP synteny blocks as a model to compare the legume and non-legume species*

VV was also used as a model to evaluate the RLK/RLP synteny block distribution among the legume/non-legume species. Among the 10 species, VV shared more synteny blocks with the GM and the non-legume SL. Fragments of the 19 VV chromosomes were distributed along the

20 GM-Chrs sharing more than two VV-Chrs fragments. 10 out of 12 SL-Chrs share two or more VV-Chr fragments. Only 7 VV-Chrs fragments are share with AT-Chrs, in fact, this is the species that shared fewer synteny blocks among the 10-species compared (Figure 14).

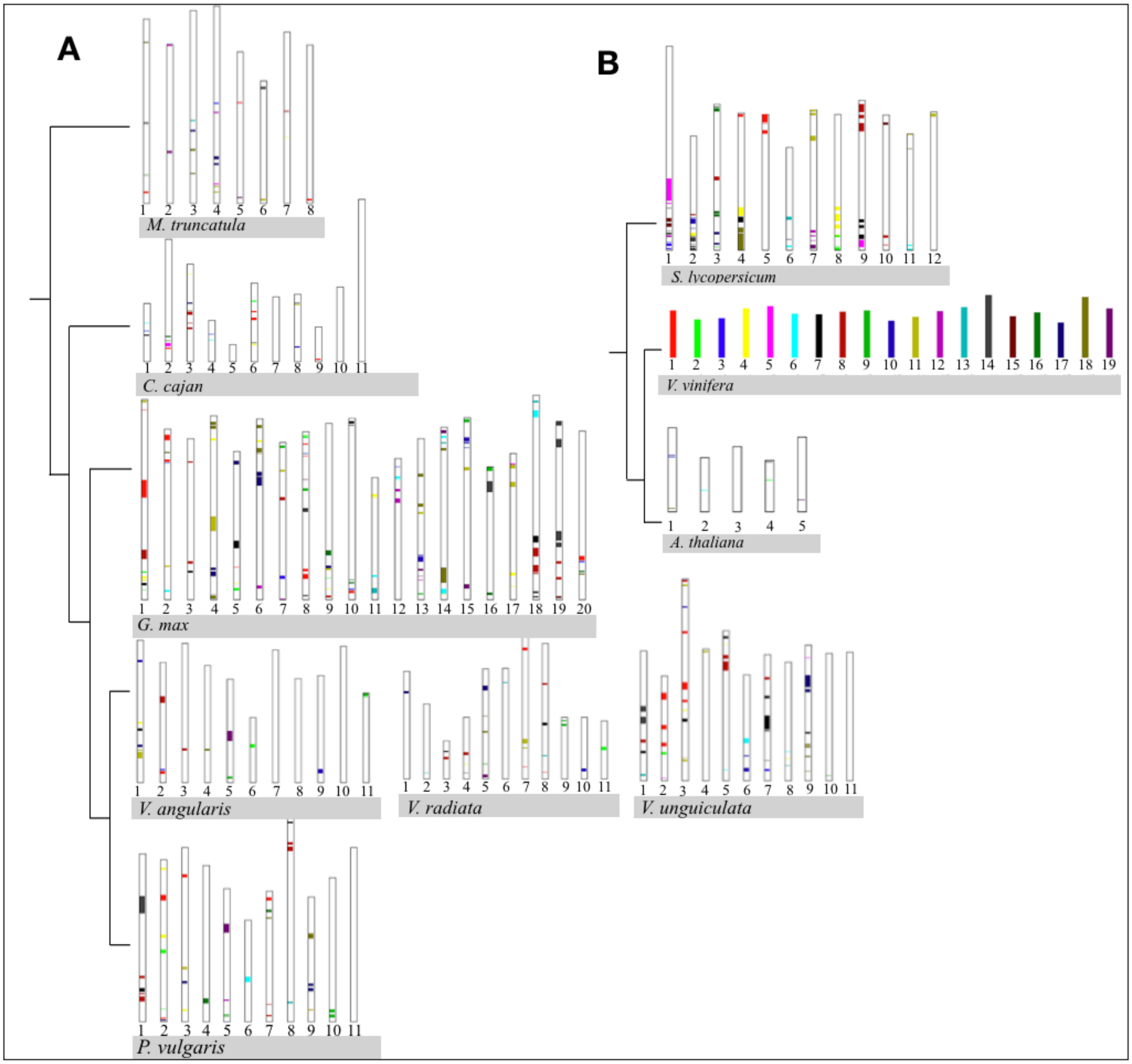


**Figure 14. Distribution of *V. vinifera* chromosome fragments with RLK/RLP.**
A. Legumes and B. non-legumes. The PV blocks include all RLK/RLP present in legumes and also all other proteins associated with the synteny blocks as a comparison reference.

### *Identification of resistance RLK and RLP genes among legumes/non-legumes*

This part of the analysis focused on determining if experimentally-validated resistance RLK/RLP (64 RLK and 26 RLP proteins in Table A1) were present in synteny among the legumes/non-legumes. In the pairwise comparison, the presence of the resistance plasma membrane in one species in a synteny block did not necessarily implicate the presence of the same experimentally-validated RLK/RLP in the other species. Still, the synteny block must have had at least one RLK/RLP, and due to the required presence of at least five genes in common, the synteny block was a valuable indicator of conserved synteny (Figure 15).

Among the RLK proteins experimentally validated in *A. thaliana*, the FLS2 gene/protein was present in a pairwise synteny block with *S. lycopersicum*; the BIR1 gene/protein in a synteny block shared with *P. vulgaris* and *S. lycopersicum*; the ERECTA gene/protein in a synteny block shared with *G. max* and *V. vinifera*; and the IOS1 gene/protein in a synteny block shared with *G. max* (Figure 11: RLK). For *S. lycopersicum*, the SISEK gene/protein or a highly identical (<90%) set of genes/proteins was present in shared synteny blocks among GM, PV, VR, VU, and VA; interestingly, the block was not shared with SL (Figure 15*). Further, for the RLP experimentally validated in *A. thaliana*, the LYM2 gene/protein was present in synteny blocks shared with *S. lycopersicum*, and the LYM1 gene/protein was in a shared synteny block with *G. max.* Finally, the RLP experimentally validated in *S. lycopersicum* and the target synteny blocks with the Ve1 and Ve2 genes/proteins were shared in a synteny block with *V. vinifera* (Figure 15: RLP).

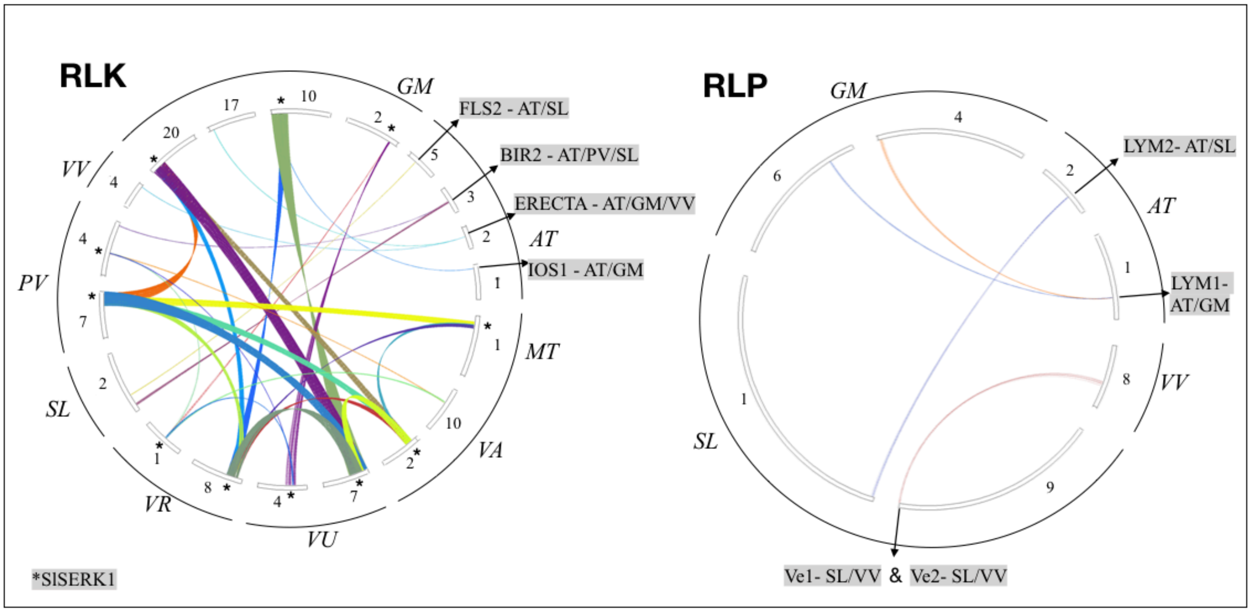


**Figure 15. Synteny blocks with resistance RLK/RLP proteins among legumes/non-legumes.**
*Blocks that showed the presence of a gene/protein with more than < 90% identity to SlSERK1 (*S. lycopersicum*). Different copies of the gene/protein were present in different species sharing the synteny blocks, but no blocks were shared with *S. lycopersicum*. *G. max* GM, *M. truncatula* MT, *V. angularis* VA, *V. unguiculata* VU, *V. radiata* VR, *S. lycopersicum* SL, *P. vulgaris* PV, *V. vinifera* VV, and *A. thaliana* AT.

## Discussion

Synteny analysis directly depends on the quality of the assembly of sequences. Missing sequences usually leads to missed gene annotations and orthologous relationships [165]. Nevertheless, the completeness of the chromosomes in the draft genomes was diverse among the legume/non-legume species evaluated. The inclusion of high-quality draft genomes that followed a comprehensive and uniform annotation and analysis in the Phytozome repository [166], such as AT, GM, MT, or PV, allowed synteny blocks to be estimated and the presence of RLK/RLP in the pairwise clusters to be targeted. Synteny analysis is regularly performed on fragmented assembled sequences, ignoring the fact that most methods were developed using high-quality draft genomes, and identifying to what extent draft assemblies lead to errors in synteny analysis is complex [159].

In evaluating the pairwise alignments calculated by MCScanX to identify then synteny blocks, about one-third of the alignments among the species showed the presence of RLK/RLP. Also, more than 75% RLK/RLP legumes genes are in synteny, the exception is MT-RLP, which was about 50% (Table 11). These results suggest that a high proportion of the plasma membrane receptors in synteny were conserved among the legumes evaluated. Interestingly, the RLK/RLP not present in synteny blocks could be orthologs or singletons; in fact, according to the results reported for RLK/RLP proteins identified in Chapter 2, 65% of RLK (Figure 5) and 91% of RLP (Figure 8) belonged to orthologous clusters, with the remaining genes classified as singletons. As expected, among legumes, not only were a higher number of orthologous proteins compared to non-legumes, but they were also present in synteny blocks. The proportion of plasma membrane proteins shared was lower among the AT, SL, and VV species.

The common bean is considered the diploid relative of the soybean. The common bean sequences are considered a reference linking two duplicate soybean regions [167]; notably, the ratio of RLK/RLP present in synteny blocks among GM:PV was approximately 2:1 (Figures 10:A, 11:A, and 12:A), suggesting that the ratio of the RLK/RLP present in the synteny blocks was also conserved in these plasma membrane receptors. Even though soybean has undergone a major duplication event [167], any sequence or sequence block unique to the soybean lineage will not have a common bean sequence signal, and any associated sequence duplication will not be uncovered [153]. This 2:1 ratio condition seems to be identified among the RLK/RLP in Chapter 2, the ortholog groups identified in Chapter 3, and the RLK/RLP in synteny blocks in Chapter 4 (Figures 10:A, 11:A, and 12:A). At different levels of the evaluation, the patterns were conserved among these two-legume species.

The established RLK/RLP ratio among MT:CC and CC:VV was 1:1, while MT:VV was 2:1 (Figures 10:B, 11:B, and 12:B). Interestingly, the total number of RLK/RLP pairwise blocks between MT and CC compared with the synteny blocks shared among legumes was the lowest; this also occurred with the outgroup VV (Figures 10:B, 11:B, and 12:B).

The 1:1 ratio of RLK/RLP was shared among VR:VA, VR:VU, VR:VV, VA:VU, VA:VV, and VU:VV. The density of the RLK/RLP spread among the pairwise synteny blocks of VR, VA, and VU was above 98% (Table 10) for all plasma membrane genes that were shared (Figures 10: C, 11: C, and 12: C), and these phenomena were also supported by the orthologous clusters identified in Chapter 3 (Figures 5 and 8). The RLK/RLP synteny blocks among the outgroup species showed a higher synteny block density among RLK compared with RLP. The 1:1 ratio was shared by SL:VV and AT:VV, and SL:AT showed a 2:1 ratio (Figures 10:D, 11:D, and 12:D). Even so, while the gene ratios among the species suggested a balanced relationship, the number of RLK/RLP genes shared among the legumes/non-legumes decreased among species with longer divergence times (Figure 4). Regarding the non-legume RLK/RLP genes in synteny, the AT gene frequency was lower compared with the results obtained for VV and SL.

The evaluation of RLK/RLP fragment distribution in synteny blocks using the PV chromosomes among the legume/non-legume species as a reference revealed diverse patterns of segmentation among species (Figure 13 and 14), which is also evident in the synteny fragment comparison in the Figure 13 using as a reference the *V. vinifera* genome. Particularly, with PV and GM, the RLK/RLP synteny block distribution was similar to the shared chromosome fragment distribution reported by McClean et al. in 2010 [153]. Also, VU had a higher conservation with the common bean than with the other legume species. Most of the chromosomes between the adzuki bean (VA) and the common bean aligned in a way similar to what Yank *et al.* reported in 2015 [5]. This corresponded with other studies where the adzuki bean (VA) species was shown to have a highly similar relationship to the common bean compared to soybean, pigeonpea, *Medicago*, chickpea, and lotus [5]. Despite the highly-related chromosome fragment distribution of PV compared with the *Vigna* species, the PV chromosome fragments showed the highest chromosome fragment relation with VU. These results were previously reported, establishing a high degree of collinearity [160, 161] (Figure 13). Notably, all RLK-nonRD genes present in the PV and VU were in synteny blocks.

Because syntenic genes are orthologs, they are often considered to share similar functions [167]. The synteny blocks reported in Figure 15 showed the experimentally-validated resistance proteins for five RLK and four RLP shared among different pairwise blocks; 24 RLK and 5 RLP synteny blocks were identified. This result suggests that further analysis must be applied to functionally evaluate the genes/proteins present in those pairwise blocks and that they do not necessarily show the presence of the same resistance RLK/RLP proteins; the presence of the other proteins that belong to the blocks could relate to the functional association to resistance, but this hypothesis needs to be confirmed. Even so, the synteny blocks reported could be used as a reference to build a targeted co-expression network and infer probable functional interactions among the genes based on the RLK/RLP proteins.

## Conclusions

In this study, the seven-legume species shared more RLK/RLP genes among synteny blocks compared with the three non-legume species (AT, SL, and VV), suggesting patterns of evolutionary conservation among them and divergence related to the non-legumes. This finding was supported by the ortholog clusters identified in Chapter 3. The identified synteny of the RLK and RLP genes in Chapter 2 was not only important as a computational annotation strategy, but also showed how those plasma membrane receptors were distributed and shared in fragments among dicots and between legumes/non-legumes. Among the legumes, the species with the fewest RLK/RLP present in synteny blocks was *M. truncatula*. For the synteny blocks shared among legumes/non-legumes, AT showed the lowest number of RLK/RLP genes present (Table

10). The GM:PV 2:1 ratio of RLK/RLP among the synteny blocks also suggests that these types of plasma receptors typically follow this ratio, which has been previously reported among these legumes [153]. The *Vigna* genus shared long fragments of chromosomes with RLK/RLP in synteny with PV. Further, *P. vulgaris* and *V. unguiculata* displayed the most similar RLK/RLP chromosome fragment distribution among all legume/non-legume comparisons. This result of their high collinearity has been reported by other authors [160]. Significantly, all RLK-nonRD genes present in the PV and VU were in synteny blocks, suggesting a highly-conserved relationship among this type of RLK between these species. Furthermore, this could suggest that these two species have the most complete genome sequences. These results suggest that as the time from species divergence increases, the species share fewer chromosome fragments with RLK/RLP genes in synteny; evidence of this process can be shown by comparing legumes/non-legumes. Interestingly the orthologous RLK/RLP could be present in the species, but not in synteny blocks. Further analysis is required to validate if the RLK/RLP synteny blocks with the presence of experimentally-validated RLK/RLP resistance genes can be used as a strategy to track blocks of resistance.

# REFERENCES

1. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J *et al*: **Genome sequence of the palaeopolyploid soybean**. *Nature* 2010, **463**(7278):178-183.
2. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu S, Song Q, Chavarro C *et al*: **A reference genome for common bean and genome-wide analysis of dual domestications**. *Nature genetics* 2014, **46**(7):707-713.
3. Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KFX, Gouzy J, Schoof H *et al*: **The Medicago genome provides insight into the evolution of rhizobial symbioses**. *Nature* 2011.
4. Kang YJ, Kim SK, Kim MY, Lestari P, Kim KH, Ha B-K, Jun TH, Hwang WJ, Lee T, Lee J *et al*: **Genome sequence of mungbean and insights into evolution within Vigna species**. *Nature Communications* 2014, **5**:5443.
5. Yang K, Tian Z, Chen C, Luo L, Zhao B, Wang Z, Yu L, Li Y, Sun Y, Li W *et al*: **Genome sequencing of adzuki bean (Vigna angularis) provides insight into high starch and low fat accumulation and domestication**. *Proceedings of the National Academy of Sciences of the United States of America* 2015, **112**(43):13213-13218.
6. Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, Donoghue MT, Azam S, Fan G, Whaley AM *et al*: **Draft genome sequence of pigeonpea (Cajanus cajan), an orphan legume crop of resource-poor farmers**. *Nature biotechnology* 2012, **30**(1):83-89.
7. NSF U, USAID, DOE-JGI, : **Vigna unguiculata v1.0**. In. phytozome.jgi.doe.gov/; 2018.
8. Pandey MK, Roorkiwal M, Singh VK, Ramalingam A, Kudapa H, Thudi M, Chitikineni A, Rathore A, Varshney RK: **Emerging Genomic Tools for Legume Breeding: Current Status and Future Prospects**. *Front Plant Sci* 2016, **7**:455.
9. Smýkal P, Coyne CJ, Ambrose MJ, Maxted N, Schaefer H, Blair MW, Berger J, Greene SL, Nelson MN, Besharat N *et al*: **Legume Crops Phylogeny and Genetic Diversity for Science and Breeding**. *Critical Reviews in Plant Sciences* 2015, **34**:43-104.
10. Moray C, Game ET, Maxted N: **Prioritising in situ conservation of crop resources: a case study of African cowpea (Vigna unguiculata)**. *Sci Rep* 2014, **4**:5247.
11. Gupta S, Nadarajan N, Gupta DS: **Legumes in Omic Era: Retrospects and Prospects**. In: *Legumes in the Omic Era.* Edited by Gupta S, Nadarajan N, Gupta DS. New York, NY: Springer New York; 2014: 1-14.
12. Li J, Dai X, Zhuang Z, Zhao PX: **LegumeIP 2.0--a platform for the study of gene function and genome evolution in legumes**. *Nucleic acids research* 2016, **44**(D1):D1189-1194.
13. Dash S, Campbell JD, Cannon EK, Cleary AM, Huang W, Kalberer SR, Karingula V, Rice AG, Singh J, Umale PE *et al*: **Legume information system (LegumeInfo.org): a key component of a set of federated data resources for the legume family**. *Nucleic acids research* 2016, **44**(D1):D1181-1188.
14. Liu J, Chen N, Grant JN, Cheng ZM, Stewart CN, Jr., Hewezi T: **Soybean kinome: functional classification and gene expression patterns**. *Journal of experimental botany* 2015, **66**(7):1919-1934.
15. Song W, Wang B, Li X, Wei J, Chen L, Zhang D, Zhang W, Li R: **Identification of Immune Related LRR-Containing Genes in Maize (** ***Zea mays*** **L.) by Genome-Wide Sequence Analysis**. *International Journal of Genomics* 2015, **2015**:1-11.

16. Sekhwal MK, Li P, Lam I, Wang X, Cloutier S, You FM: **Disease Resistance Gene Analogs (RGAs) in Plants**. *International Journal of Molecular Sciences* 2015, **16**:19248-19290.
17. Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH: **Synteny and collinearity in plant genomes**. *Science* 2008, **320**(5875):486-488.
18. Monaghan J, Zipfel C: **Plant pattern recognition receptor complexes at the plasma membrane**. *Current Opinion in Plant Biology* 2012, **15**:349-357.
19. Shiu SH, Bleecker AB: **Plant receptor-like kinase gene family: diversity, function, and signaling**. *Science's STKE : signal transduction knowledge environment* 2001, **2001**(113):re22.
20. Tor M, Lotze MT, Holton N: **Receptor-mediated signalling in plants: molecular patterns and programmes**. *Journal of experimental botany* 2009, **60**(13):3645-3654.
21. Sanabria NM, Huang JC, Dubery IA: **Self/nonself perception in plants in innate immunity and defense**. *Self Nonself* 2010, **1**(1):40-54.
22. Lin W, Ma X, Shan L, He P: **Big Roles of Small Kinases: The Complex Functions of Receptor-Like Cytoplasmic Kinases in Plant Immunity and Development**. *Journal of Integrative Plant Biology* 2013, **55**:1188-1197.
23. Afzal AJ, Wood AJ, Lightfoot DA: **Plant receptor-like serine threonine kinases: roles in signaling and plant defense**. *Molecular plant-microbe interactions : MPMI* 2008, **21**(5):507-517.
24. Vihinen M: **How to evaluate performance of prediction methods? Measures and their interpretation in variation effect analysis**. *BMC genomics* 2012, **13 Suppl 4**:S2.
25. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M *et al*: **The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools**. *Nucleic acids research* 2012, **40**(Database issue):D1202-1210.
26. Sato S, Tabata S, Hirakawa H, Asamizu E, Shirasawa K, Isobe S, Kaneko T, Nakamura Y, Shibata D, Aoki K *et al*: **The tomato genome sequence provides insights into fleshy fruit evolution**. *Nature* 2012, **485**(7400):635-641.
27. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C *et al*: **The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla**. *Nature* 2007, **449**(7161):463-467.
28. Tang H, Wang X, Bowers JE, Ming R, Alam M, Paterson AH: **Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps**. *Genome Research* 2008, **18**(12):1944-1954.
29. Zhou J-M, Tang D, Wang G: **Receptor kinases in plant pathogen interactions: more than pattern recognition**. *The Plant Cell* 2017:tpc.00891.02016.
30. Jones JD, Dangl JL: **The plant immune system**. *Nature* 2006, **444**(7117):323-329.
31. Dalio RJD, Magalhaes DM, Rodrigues CM, Arena GD, Oliveira TS, Souza-Neto RR, Picchi SC, Martins PMM, Santos PJC, Maximo HJ *et al*: **PAMPs, PRRs, effectors and R-genes associated with citrus-pathogen interactions**. *Ann Bot* 2017, **119**(5):749-774.
32. Thomma BP, Nurnberger T, Joosten MH: **Of PAMPs and effectors: the blurred PTI-ETI dichotomy**. *Plant Cell* 2011, **23**(1):4-15.
33. Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK, Park SW: **Plant disease resistance genes: Current status and future directions**. *Physiological and Molecular Plant Pathology* 2012, **78**:51-65.
34. Jinn TL, Stone JM, Walker JC: **HAESA, an Arabidopsis leucine-rich repeat receptor kinase, controls floral organ abscission**. *Genes & development* 2000, **14**(1):108-117.

35. Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y: **Arabidopsis CLV3 peptide directly binds CLV1 ectodomain**. *Science* 2008, **319**(5861):294.
36. Murase K, Shiba H, Iwano M, Che FS, Watanabe M, Isogai A, Takayama S: **A membrane-anchored protein kinase involved in Brassica self-incompatibility signaling**. *Science* 2004, **303**(5663):1516-1519.
37. Leslie ME, Lewis MW, Youn JY, Daniels MJ, Liljegren SJ: **The EVERSHED receptor-like kinase modulates floral organ shedding in Arabidopsis**. *Development* 2010, **137**(3):467-476.
38. Nadeau JA, Sack FD: **Control of stomatal distribution on the Arabidopsis leaf surface**. *Science* 2002, **296**(5573):1697-1700.
39. Bayer M, Nawy T, Giglione C, Galli M, Meinnel T, Lukowitz W: **Paternal control of embryonic patterning in Arabidopsis thaliana**. *Science* 2009, **323**(5920):1485-1488.
40. Lu D, Wu S, Gao X, Zhang Y, Shan L, He P: **A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(1):496-501.
41. Liu Z, Wu Y, Yang F, Zhang Y, Chen S, Xie Q, Tian X, Zhou JM: **BIK1 interacts with PEPRs to mediate ethylene-induced immunity**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(15):6205-6210.
42. Claus Lucas Alves N, Savatin Daniel V, Russinova E: **The crossroads of receptor-mediated signaling and endocytosis in plants**. *Journal of Integrative Plant Biology* 2018, **0**(ja).
43. Shiu S-h, Karlowski W, Pan R: **Comparative analysis of the receptor-like kinase family in Arabidopsis and rice**. *The Plant Cell …* 2004, **16**:1220-1234.
44. Dufayard J-F, Bettembourg M, Fischer I, Droc G, Guiderdoni E, Périn C, Chantret N, Diévart A: **New Insights on Leucine-Rich Repeats Receptor-Like Kinase Orthologous Relationships in Angiosperms**. *Frontiers in Plant Science* 2017, **08**:1-18.
45. Bouwmeester K, Govers F: **Arabidopsis L-type lectin receptor kinases: Phylogeny, classification, and expression profiles**. *Journal of experimental botany* 2009, **60**:4383-4396.
46. Dambuza IM, Brown GD: **C-type lectins in immunity: Recent developments**. In: *Current Opinion in Immunology.* vol. 32; 2015: 21-27.
47. Lannoo N, Van Damme EJM: **Lectin domains at the frontiers of plant defense**. *Frontiers in Plant Science* 2014, **5**:1-16.
48. Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Gronlund M, Sato S, Nakamura Y, Tabata S, Sandal N *et al*: **Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases**. *Nature* 2003, **425**(6958):585-592.
49. Knogge W, Scheel D: **LysM receptors recognize friend and foe**. *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**(29):10829-10830.
50. Mulder L, Lefebvre B, Cullimore J, Imberty A: **LysM domains of Medicago truncatula NFP protein involved in Nod factor perception. Glycosylation state, molecular modeling and docking of chitooligosaccharides and Nod factors**. *Glycobiology* 2006, **16**(9):801-809.
51. Desaki Y, Miyata K, Suzuki M, Shibuya N, Kaku H: **Plant immunity and symbiosis signaling mediated by LysM receptors**. *Innate Immunity* 2017:175342591773888.
52. Li C, Wu H-m, Cheung AY: **FERONIA and her pals: functions and mechanisms**. *Plant Physiology* 2016, **171**:pp.00667.02016.

53. Petre B, Major I, Rouhier N, Duplessis S: **Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar**. *BMC plant biology* 2011, **11**:33.
54. Maleck K, Levine a, Eulgem T, Morgan a, Schmid J, Lawton Ka, Dangl JL, Dietrich Ra: **The transcriptome of Arabidopsis thaliana during systemic acquired resistance.** *Nature genetics* 2000, **26**:403-410.
55. Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM: **Coordinated plant defense responses in Arabidopsis revealed by microarray analysis.** *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**:11655-11660.
56. Quistad SD, Traylor-Knowles N: **Precambrian origins of the TNFR superfamily**. *Cell Death Discovery* 2016, **2**:16058.
57. Nikonorova N, Vu LD, Czyzewicz N, Gevaert K, De Smet I: **A phylogenetic approach to study the origin and evolution of the CRINKLY4 family**. *Frontiers in Plant Science* 2015, **6**.
58. Bourdais G, Burdiak P, Gauthier A, Nitsch L, Salojärvi J, Rayapuram C, Idänheimo N, Hunter K, Kimura S, Merilo E *et al*: **Large-Scale Phenomics Identifies Primary and Fine-Tuning Roles for CRKs in Responses Related to Oxidative Stress**. *PLoS Genetics* 2015, **11**:1-36.
59. Frescatada-Rosa M, Robatzek S, Kuhn H: **Should I stay or should I go? Traffic control for plant pattern recognition receptors**. *Current Opinion in Plant Biology* 2015, **28**:23-29.
60. Trdá L, Boutrot F, Claverie J, Brulé D, Dorey S, Poinssot B: **Perception of pathogenic or beneficial bacteria and their evasion of host immunity: pattern recognition receptors in the frontline**. *Frontiers in Plant Science* 2015, **6**.
61. Singh P, Zimmerli L: **Lectin receptor kinases in plant innate immunity**. *Front Plant Sci* 2013, **4**:124.
62. Petre B, Major I, Rouhier N, Duplessis S: **Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar**. *BMC plant biology* 2011, **11**:33.
63. Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A, Hinata K: **The S receptor kinase determines self-incompatibility in Brassica stigma**. *Nature* 2000, **403**(6772):913-916.
64. Tax F, Kermmerling B: **Receptor-like Kinases in Plants**, vol. 13. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012.
65. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S: **The protein kinase complement of the human genome**. *Science* 2002, **298**(5600):1912-1934.
66. Lehti-Shiu MD, Shiu SH: **Diversity, classification and function of the plant protein kinase superfamily**. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 2012, **367**(1602):2619-2639.
67. Morillo SA, Tax FE: **Functional analysis of receptor-like kinases in monocots and dicots**. *Curr Opin Plant Biol* 2006, **9**(5):460-469.
68. Dardick C, Schwessinger B, Ronald P: **Non-arginine-aspartate (non-RD) kinases are associated with innate immune receptors that recognize conserved microbial signatures**. *Current Opinion in Plant Biology* 2012, **15**:358-366.
69. Schwessinger B, Roux M, Kadota Y, Ntoukakis V, Sklenar J, Jones A, Zipfel C: **Phosphorylation-dependent differential regulation of plant growth, cell death, and**

**innate immunity by the regulatory receptor-like kinase BAK1**. *PLoS Genetics* 2011, **7**.

70. Osuna-Cruz CM, Paytuvi-Gallart A, Di Donato A, Sundesha V, Andolfo G, Aiese Cigliano R, Sanseverino W, Ercolano MR: **PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes**. *Nucleic acids research* 2018, **46**(D1):D1197-D1201.
71. Bi G, Liebrand TW, Cordewener JH, America AH, Xu X, Joosten MH: ***Arabidopsis thaliana*** **receptor-like protein** ***At*** **RLP23 associates with the receptor-like kinase** ***At*** **SOBIR1**. *Plant Signaling & Behavior* 2014, **9**:e27937.
72. Shao Z-Q, Zhang Y-M, Hang Y-Y, Xue J-Y, Zhou G-C, Wu P, Wu X-Y, Wu X-Z, Wang Q, Wang B *et al*: **Long-Term Evolution of Nucleotide-Binding Site-Leucine-Rich Repeat Genes: Understanding Gained from and beyond the Legume Family**. *Plant Physiology* 2014, **166**:217-234.
73. Michelmore RW: **The impact zone: genomics and breeding for durable disease resistance**. *Curr Opin Plant Biol* 2003, **6**(4):397-404.
74. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW: **Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis**. *Plant Cell* 2003, **15**(4):809-834.
75. Shiu SH, Karlowski WM, Pan R, Tzeng YH, Mayer KF, Li WH: **Comparative analysis of the receptor-like kinase family in Arabidopsis and rice**. *Plant Cell* 2004, **16**(5):1220-1234.
76. Yang S, Zhang X, Yue JX, Tian D, Chen JQ: **Recent duplications dominate NBS-encoding gene expansion in two woody species**. *Mol Genet Genomics* 2008, **280**(3):187-198.
77. Li P, Quan X, Jia G, Xiao J, Cloutier S, You FM: **RGAugury: a pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants**. *BMC genomics* 2016, **17**(1):852.
78. Osorno JM, P.E. McClean, and T. Close.: **Advanced breeding techniques for grain legumes in the genomics era**. In: *Achieving sustainable cultivation of grain legumes Volume 1: Advances in breeding and cultivation techniques,*. Edited by Sivasankar Sea. Cambridge, UK: Burleigh Dodds Science Publishing; 2018.
79. Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: discriminating signal peptides from transmembrane regions**. *Nature methods* 2011, **8**(10):785-786.
80. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes**. *Journal of molecular biology* 2001, **305**(3):567-580.
81. Finn RD, Clements J, Eddy SR: **HMMER web server: interactive sequence similarity searching**. *Nucleic acids research* 2011, **39**(Web Server issue):W29-37.
82. The UniProt C: **UniProt: the universal protein knowledgebase**. *Nucleic acids research* 2017, **45**(D1):D158-D169.
83. Sanseverino W, Roma G, De Simone M, Faino L, Melito S, Stupka E, Frusciante L, Ercolano MR: **PRGdb: a bioinformatics platform for plant resistance gene analysis**. *Nucleic acids research* 2010, **38**(Database issue):D814-821.
84. Tang D, Wang G, Zhou JM: **Receptor Kinases in Plant-Pathogen Interactions: More Than Pattern Recognition**. *Plant Cell* 2017, **29**(4):618-637.
85. Wu Y, Zhou J-M: **Receptor-Like Kinases in Plant Innate Immunity**. *Journal of Integrative Plant Biology* 2013, **55**:1271-1286.

86. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J *et al*: **Pfam: the protein families database**. *Nucleic acids research* 2014, **42**(Database issue):D222-230.
87. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R: **The EMBL-EBI bioinformatics web and programmatic tools framework**. *Nucleic acids research* 2015, **43**(W1):W580-584.
88. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J *et al*: **The Pfam protein families database**. *Nucleic acids research* 2012, **40**(Database issue):D290-301.
89. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A *et al*: **The Pfam protein families database: towards a more sustainable future**. *Nucleic acids research* 2016, **44**(D1):D279-285.
90. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: **MEME SUITE: tools for motif discovery and searching**. *Nucleic acids research* 2009, **37**(Web Server issue):W202-208.
91. Sarris PF, Cevik V, Dagdas G, Jones JD, Krasileva KV: **Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens**. *BMC Biol* 2016, **14**:8.
92. Li W, Godzik A: **Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences**. *Bioinformatics* 2006, **22**(13):1658-1659.
93. Lavin M, Herendeen PS, Wojciechowski MF: **Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary**. *Systematic biology* 2005, **54**(4):575-594.
94. Roulin A, Auer PL, Libault M, Schlueter J, Farmer A, May G, Stacey G, Doerge RW, Jackson SA: **The fate of duplicated genes in a polyploid plant genome**. *The Plant journal : for cell and molecular biology* 2013, **73**(1):143-153.
95. Lehti-Shiu MD, Shiu S-H: **Diversity, classification and function of the plant protein kinase superfamily**. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2012, **367**:2619-2639.
96. Johnson LN, Noble MEM, Owen DJ: **Active and inactive protein kinases: Structural basis for regulation**. *Cell* 1996, **85**:149-158.
97. Vij S, Giri J, Dansana PK, Kapoor S, Tyagi AK: **The receptor-like cytoplasmic kinase (OsRLCK) gene family in rice: organization, phylogenetic relationship, and expression during development and stress**. *Mol Plant* 2008, **1**(5):732-750.
98. Lehti-Shiu MD, Zou C, Shiu S-H: **Origin, Diversity, Expansion History, and Functional Evolution of the Plant Receptor-Like Kinase/Pelle Family**. In: *Receptor-like Kinases in Plants: From Development to Defense.* Edited by Tax F, Kemmerling B. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012: 1-22.
99. Letunic I, Doerks T, Bork P: **SMART: recent updates, new developments and status in 2015**. *Nucleic acids research* 2015, **43**(Database issue):D257-260.
100. Liu PL, Du L, Huang Y, Gao SM, Yu M: **Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants**. *BMC Evol Biol* 2017, **17**(1):47.
101. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL *et al*: **The Pfam protein families database**. *Nucleic acids research* 2008, **36**(Database issue):D281-288.
102. Lewis GP: **Legumes of the world**. Richmond, UK: Royal Botanic Gardens, Kew; 2005.

103. Abbo S, Lev-Yadun S, Gopher A: **Plant Domestication and Crop Evolution in the Near East: On Events and Processes**. *Critical Reviews in Plant Sciences* 2012, **31**:241-257.

104. Cannon SB, McKain MR, Harkess A, Nelson MN, Dash S, Deyholos MK, Peng Y, Joyce B, Stewart CN, Jr., Rolf M *et al*: **Multiple polyploidy events in the early radiation of nodulating and nonnodulating legumes**. *Molecular biology and evolution* 2015, **32**(1):193-210.

105. Benedito VA, Torres-Jerez I, Murray JD, Andriankaja A, Allen S, Kakar K, Wandrey M, Verdier J, Zuber H, Ott T *et al*: **A gene expression atlas of the model legume Medicago truncatula**. *The Plant journal : for cell and molecular biology* 2008, **55**(3):504-513.

106. Ehlers JD, Hall aE: **Cowpea (Vigna unguiculata L. Walp.)**. *Field Crops Research* 1997, **53**(1-3):187-204.

107. Tang H, Wang X, Bowers JE, Ming R, Alam M, Paterson AH: **Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps**. *Genome Research* 2008, **18**:1944-1954.

108. Van de Peer Y, Fawcett JA, Proost S, Sterck L, Vandepoele K: **The flowering world: a tale of duplications**. *Trends in plant science* 2009, **14**(12):680-688.

109. Severin AJ, Cannon SB, Graham MM, Grant D, Shoemaker RC: **Changes in twelve homoeologous genomic regions in soybean following three rounds of polyploidy**. *Plant Cell* 2011, **23**(9):3129-3136.

110. Shao ZQ, Zhang YM, Hang YY, Xue JY, Zhou GC, Wu P, Wu XY, Wu XZ, Wang Q, Wang B *et al*: **Long-term evolution of nucleotide-binding site-leucine-rich repeat genes: understanding gained from and beyond the legume family**. *Plant Physiol* 2014, **166**(1):217-234.

111. Zheng F, Wu H, Zhang R, Li S, He W, Wong FL, Li G, Zhao S, Lam HM: **Molecular phylogeny and dynamic evolution of disease resistance genes in the legume family**. *BMC genomics* 2016, **17**:402.

112. Foyer CH, Lam HM, Nguyen HT, Siddique KH, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA, Hodgson JM *et al*: **Neglecting legumes has compromised human health and sustainable food production**. *Nat Plants* 2016, **2**:16112.

113. Fitch WM: **Distinguishing homologous from analogous proteins**. *Syst Zool* 1970, **19**(2):99-113.

114. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ: **Proteinortho: detection of (co-)orthologs in large-scale analysis**. *BMC bioinformatics* 2011, **12**:124.

115. Wang Y, Coleman-Derr D, Chen G, Gu YQ: **OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species**. *Nucleic acids research* 2015, **43**(W1):W78-84.

116. Fischer S, Brunk BP, Chen F, Gao X, Harb OS, Iodice JB, Shanmugam D, Roos DS, Stoeckert CJ, Jr.: **Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups**. *Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]* 2011, **Chapter 6**:Unit 6 12 11-19.

117. Shi T, Huang H, Sanderson MJ, Tax FE: **Evolutionary dynamics of leucine-rich repeat receptor-like kinases and related genes in plants: A phylogenomic approach**. *Journal of Integrative Plant Biology* 2014, **56**(7):648-662.

118. Jamieson PA, Shan L, He P: **Plant Science Plant cell surface molecular cypher : Receptor-like proteins and their roles in immunity and development**. *Plant Science* 2018, **274**(February):242-251.

119. Macho AP, Zipfel C: **Plant PRRs and the activation of innate immune signaling**. *Molecular cell* 2014, **54**(2):263-272.
120. Bohm H, Albert I, Fan L, Reinhard A, Nurnberger T: **Immune receptor complexes at the plant cell surface**. *Curr Opin Plant Biol* 2014, **20**:47-54.
121. Tavormina P, De Coninck B, Nikonorova N, De Smet I, Cammue BP: **The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions**. *Plant Cell* 2015, **27**(8):2095-2118.
122. Lin W, Ma X, Shan L, He P: **Big roles of small kinases: the complex functions of receptor-like cytoplasmic kinases in plant immunity and development**. *J Integr Plant Biol* 2013, **55**(12):1188-1197.
123. Liu J, Ding P, Sun T, Nitta Y, Dong O, Huang X, Yang W, Li X, Botella JR, Zhang Y: **Heterotrimeric G proteins serve as a converging point in plant defense signaling activated by multiple receptor-like kinases**. *Plant Physiol* 2013, **161**(4):2146-2158.
124. Zhang J, Li W, Xiang T, Liu Z, Laluk K, Ding X, Zou Y, Gao M, Zhang X, Chen S *et al*: **Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a Pseudomonas syringae effector**. *Cell Host Microbe* 2010, **7**(4):290-301.
125. Breiden M, Simon R: **Q&amp;A: How does peptide signaling direct plant development?** *BMC Biology* 2016, **14**:58.
126. Belkhadir Y, Yang L, Hetzel J, Dangl JL, Chory J: **The growth-defense pivot: Crisis management in plants mediated by LRR-RK surface receptors**. *Trends in Biochemical Sciences* 2014, **39**:447-456.
127. Jordan IK, Makarova KS, Spouge JL, Wolf YI, Koonin EV: **Lineage-specific gene expansions in bacterial and archaeal genomes**. *Genome Res* 2001, **11**(4):555-565.
128. Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu SH: **Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli**. *Plant Physiol* 2008, **148**(2):993-1003.
129. Shiu SH, Bleecker AB: **Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis**. *Plant Physiol* 2003, **132**(2):530-543.
130. Bellande K, Bono JJ, Savelli B, Jamet E, Canut H: **Plant lectins and lectin receptor-like kinases: How do they sense the outside?** *International Journal of Molecular Sciences* 2017, **18**.
131. Wel H, Loeve K: **Isolation and Characterization of Thaumatin I and II, the Sweet-Tasting Proteins from Thaumatococcus daniellii Benth**. *European Journal of Biochemistry* 1972, **31**:221-225.
132. Shatters RG, Boykin LM, Lapointe SL, Hunter WB, Weathersbee AA: **Phylogenetic and structural relationships of the PR5 gene family reveal an ancient multigene family conserved in plants and select animal taxa**. *Journal of Molecular Evolution* 2006, **63**:12-29.
133. Zhang S, Chen C, Li L, Meng L, Singh J: **Evolutionary expansion, gene structure, and expression of the rice wall-associated kinase gene family**. *Plant …* 2005, **139**:1107-1124.
134. Tocquard K, Lafon-Placette C, Auguin D, Muries B, Bronner G, Lopez D, Fumanal B, Franchel J, Bourgerie S, Maury S *et al*: **In silico study of wall-associated kinase family reveals large-scale genomic expansion potentially connected with functional diversification in Populus**. *Tree Genetics and Genomes* 2014, **10**:1135-1147.
135. Li L, Stoeckert CJ, Jr., Roos DS: **OrthoMCL: identification of ortholog groups for eukaryotic genomes**. *Genome Res* 2003, **13**(9):2178-2189.

136. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J *et al*: **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega**. *Mol Syst Biol* 2011, **7**:539.
137. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T: **trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses**. *Bioinformatics* 2009, **25**(15):1972-1973.
138. Miller MA, Pfeiffer W, Schwartz T: **Creating the CIPRES Science Gateway for inference of large phylogenetic trees**. *2010 Gateway Computing Environments Workshop, GCE 2010* 2010.
139. Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies**. *Bioinformatics* 2014, **30**(9):1312-1313.
140. Jones DT, Taylor WR, Thornton JM: **The rapid generation of mutation data matrices from protein sequences**. *Comput Appl Biosci* 1992, **8**(3):275-282.
141. Creevey CJ, Muller J, Doerks T, Thompson JD, Arendt D, Bork P: **Identifying single copy orthologs in Metazoa**. *PLoS Comput Biol* 2011, **7**(12):e1002269.
142. Xu C, Nadon BD, Kim KD, Jackson SA: **Genetic and epigenetic divergence of duplicate genes in two legume species**. *Plant, cell & environment* 2018.
143. Fang G, Bhardwaj N, Robilotto R, Gerstein MB: **Getting started in gene orthology and functional analysis**. *PLoS Comput Biol* 2010, **6**(3):e1000703.
144. Jordan IK, Wolf YI, Koonin EV: **Duplicated genes evolve slower than singletons despite the initial rate increase**. *BMC Evol Biol* 2004, **4**:22.
145. Panchy N, Lehti-Shiu M, Shiu SH: **Evolution of Gene Duplication in Plants**. *Plant Physiol* 2016, **171**(4):2294-2316.
146. Shiu SH, Li WH: **Origins, lineage-specific expansions, and multiple losses of tyrosine kinases in eukaryotes**. *Molecular biology and evolution* 2004, **21**(5):828-840.
147. De Smet R, Adams KL, Vandepoele K, Van Montagu MC, Maere S, Van de Peer Y: **Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(8):2898-2903.
148. Duarte JM, Wall PK, Edger PP, Landherr LL, Ma H, Pires JC, Leebens-Mack J, dePamphilis CW: **Identification of shared single copy nuclear genes in Arabidopsis, Populus, Vitis and Oryza and their phylogenetic utility across various taxonomic levels**. *BMC Evol Biol* 2010, **10**:61.
149. Li Z, De La Torre AR, Sterck L, Cánovas FM, Avila C, Merino I, Cabezas JA, Cervera MT, Ingvarsson PK, Van De Peer Y: **Single-copy genes as molecularmarkers for phylogenomic studies in seed plants**. *Genome Biology and Evolution* 2017, **9**(5):1130-1147.
150. Hofer JM, Noel Ellis TH: **Developmental specialisations in the legume family**. *Curr Opin Plant Biol* 2014, **17**:153-158.
151. Schaefer H, Hechenleitner P, Santos-Guerra A, Menezes de Sequeira M, Pennington RT, Kenicer G, Carine MA: **Systematics, biogeography, and character evolution of the legume tribe Fabeae with special focus on the middle-Atlantic island lineages**. *BMC Evol Biol* 2012, **12**:250.
152. Doyle JJ, Luckow MA: **The rest of the iceberg. Legume diversity and evolution in a phylogenetic context**. *Plant Physiol* 2003, **131**(3):900-910.
153. McClean PE, Mamidi S, McConnell M, Chikara S, Lee R: **Synteny mapping between common bean and soybean reveals extensive blocks of shared loci**. *BMC genomics* 2010, **11**:184.

154. Ghiurcuta CG, Moret BM: **Evaluating synteny for improved comparative studies**. *Bioinformatics* 2014, **30**(12):i9-18.
155. Coghlan A, Eichler EE, Oliver SG, Paterson AH, Stein L: **Chromosome evolution in eukaryotes: a multi-kingdom perspective**. *Trends in genetics : TIG* 2005, **21**(12):673-682.
156. Kevei Z, Seres A, Kereszt A, Kalo P, Kiss P, Toth G, Endre G, Kiss GB: **Significant microsynteny with new evolutionary highlights is detected between Arabidopsis and legume model plants despite the lack of macrosynteny**. *Mol Genet Genomics* 2005, **274**(6):644-657.
157. Pearson WR: **An introduction to sequence similarity ("homology") searching**. *Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]* 2013, **Chapter 3**:Unit3 1.
158. Tekaia F: **Inferring Orthologs: Open Questions and Perspectives**. *Genomics Insights* 2016, **9**:17-28.
159. Liu D, Hunt M, Tsai IJ: **Inferring synteny between genome assemblies: a systematic evaluation**. *BMC bioinformatics* 2018, **19**(1):26.
160. Vasconcelos EV, de Andrade Fonseca AF, Pedrosa-Harand A, de Andrade Bortoleti KC, Benko-Iseppon AM, da Costa AF, Brasileiro-Vidal AC: **Intra- and interchromosomal rearrangements between cowpea [Vigna unguiculata (L.) Walp.] and common bean (Phaseolus vulgaris L.) revealed by BAC-FISH**. *Chromosome Res* 2015, **23**(2):253-266.
161. Munoz-Amatriain M, Mirebrahim H, Xu P, Wanamaker SI, Luo M, Alhakami H, Alpert M, Atokple I, Batieno BJ, Boukar O *et al*: **Genome resources for climate-resilient cowpea, an essential crop for food security**. *The Plant journal : for cell and molecular biology* 2017, **89**(5):1042-1054.
162. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H *et al*: **MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity**. *Nucleic acids research* 2012, **40**(7):e49.
163. McClean PE, Lavin M, Gepts P, Jackson Sa: **Phaseolus vulgaris : A Diploid Model for Soybean**. *Plant Genetics and Genomics: Crops and Models* 2008, **2**:55-76.
164. Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzbittel L, Childs KL, Yandell M, Gundlach H *et al*: **An improved genome release (version Mt4.0) for the model legume Medicago truncatula**. *BMC genomics* 2014, **15**:312.
165. Bhutkar A, Russo S, Smith TF, Gelbart WM: **Techniques for Multi-Genome Synteny Analysis to Overcome Assembly Limitations**. *Genome Informatics* 2006, **17**(2):152-161.
166. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N *et al*: **Phytozome: a comparative platform for green plant genomics**. *Nucleic acids research* 2012, **40**(Database issue):D1178-1186.
167. Cheng F, Wu J, Fang L, Wang X: **Syntenic gene analysis between Brassica rapa and other Brassicaceae species**. *Frontiers in Plant Science* 2012, **3**(August):1-6.

# APPENDIX

**Table A1. List of RLK, RLP, and cytoplasmic resistance proteins experimentally validated**.

| Name | Species | Type | Class | Uniprot ID or GeneBank ID |
|---|---|---|---|---|
| FLS2 | *Arabidopsis thaliana* | LRR RLK | RLK | Q9FL28 |
| FLS3 | *Solanum lycopersicum* | LRR RLK | RLK | Q9FFQ5 |
| LORE/SD1-29 | *A. thaliana* | Lec RLK | RLK | O64782 |
| XPS1 | *A. thaliana* | LRR RLK | RLK | Q9SD62 |
| EFR | *A. thaliana* | LRR RLK | RLK | C0LGT6 |
| CERK1 | *A.thaliana* | LysM RLK | RLK | A8R7E6 |
| OsCERK1 | *O. sativa* | LysM RLK | RLK | D7UPN3 |
| LYM3 | *O. sativa* | LysM RLK | RLK | Q6NPN4 |
| LYP4 | *O. sativa* | LysM RLK | RLK | Q67UE8 |
| LYP6 | *O. sativa* | LysM RLK | RLK | Q69T51 |
| WAK1 | *A. thaliana* | WAK RLK | RLK | Q39191 |
| PEPR1 | *A.thaliana* | LRR RLK | RLK | Q9SSL9 |
| PEPR2 | *A.thaliana* | LRR RLK | RLK | Q9FZ59 |
| BAK1 | *A.thaliana* | LRR RLK | RLK | Q94F62 |
| SOBIR1 | *S. lycopersicum* | LRR RLK | RLK | K4C8Q3 |
| PSKR1 | *A. thaliana* | LRR RLK | RLK | Q9ZVR7 |
| PSY1R | *A. thaliana* | LRR RLK | RLK | Q9C7S5 |
| BIR1 | *A. thaliana* | LRR RLK | RLK | A0A178UN61 |
| BIR2 | *A. thaliana* | LRR RLK | RLK | Q9LSI9 |
| CRK28 | *A. thaliana* | LRR RLK | RLK | O65405 |
| IOS1 | *A. thaliana* | LRR RLK | RLK | Q9C8I6 |
| SOBIR1 | *A. thaliana* | LRR RLK | RLK | Q9SKB2 |
| ERECTA | *A. thaliana* | LRR RLK | RLK | Q42371 |
| SRF3 | *A. thaliana* | LRR RLK | RLK | Q6R2K3 |
| XA21 | *Oryza sativa* | LRR RLK | RLK | Q1MX30 |
| Xa3/Xa26 | *O. sativa* | LRR RLK | RLK | Q2EZ14 |
| OsSERK1 | *O. sativa* | LRR RLK | RLK | Q6Z4U4 |
| OsSERK1-v1 | *O. sativa* | LRR RLK | RLK | Q5VQL9 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or GeneBank ID |
|---|---|---|---|---|
| ds1 | *Sorghum bicolor* | LRR RLK | RLK | K0IXC4 |
| SlSERK1 | *S. lycopersicum* | LRR RLK | RLK | G0XZA3 |
| NbSERK1 | *Nicotiana benthamiana* | LRR RLK | RLK | G0XZA2 |
| LYK4 | *A. thaliana* | LysM RLK | RLK | O64825 |
| LYK5 | *A. thaliana* | LysM RLK | RLK | O22808 |
| Bti9 | *S. lycopersicum* | LysM RLK | RLK | E2FYC4 |
| SlLyk13 | *S. lycopersicum* | LysM RLK | RLK | G3KGB4 |
| THE1 | *A. thaliana* | CrRLK1L RLK | RLK | Q9LK35 |
| FER | *A. thaliana* | CrRLK1L RLK | RLK | Q9SCZ4 |
| Pi-d2 | *O. sativa* | Lec RLK | RLK | Q651J1 |
| LecRK-I.9 | *A. thaliana* | Lec RLK | RLK | Q9LSR8 |
| LecRK-V.5 | *A. thaliana* | Lec RLK | RLK | Q96285 |
| LecRK-VI.2 | *A. thaliana* | Lec RLK | RLK | Q9M021 |
| NgRLK1 | *Nicotiana glutinosa* | Lec RLK | RLK | B6EB06 |
| LecRK1 | *Nicotiana attenuate* | Lec RLK | RLK | G4XXY5 |
| NbLRK1 | *N. benthamiana* | Lec RLK | RLK | B3XWM9 |
| WAKL22 | *A.thaliana* | WAK RLK | RLK | Q8RY17 |
| WAKL10 | *A. thaliana* | WAK RLK | RLK | Q8VYA3 |
| OsWAK1 | *O. sativa* | WAK RLK | RLK | A0A0P0Y5F3 |
| TaRLK-R1 | *Triticum aestivum* | RLK | RLK | Q00NV2 |
| TaRLK-R2 | *T. aestivum* | RLK | RLK | Q00NV1 |
| TaRLK-R3 | *T. aestivum* | RLK | RLK | Q6VTH8 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or GeneBank ID |
|---|---|---|---|---|
| SNC4 | *A. thaliana* | RLK | RLK | D7SFH9 |
| LRK10 | *Triticum aestivum (Wheat)* | RLK | RLK | P93604 |
| PBS1 | *A. thaliana* | RLK | RLK | Q9FE20 |
| SIRK | *A. thaliana* | RLK | RLK | O64483 |
| AtPEP2 | *A. thaliana* | RLK | RLK | Q9LV88 |
| DIPM1 | *Malus domestica* | RLK | RLK | Q2LJM3 |
| DIMP2 | *M. domestica* | RLK | RLK | Q2LJM2 |
| DIMP3 | *M. Domestica* | RLK | RLK | Q2LJM1 |
| DIMP4 | *M. Domestica* | RLK | RLK | Q2LJM0 |
| PR5K | *A.thaliana* | RLK | RLK | Q38925 |
| Rpg1 | *Hordeum. Vulgare* | RLK | RLK | Q8L3P8 |
| XB3 “Xa21D” | *O. sativa* | RLK | RLK | Q65XV2 |
| BRI1 | *A.thaliana* | RLK | RLK | O22476 |
| BKK1 | *A.thaliana* | RLK | RLK | Q9SKG5 |
| SlLyk11 - Lyk11 | *S. lycopersicum* | RLK | RLK | G3KGB2 |
| SlLyk12 - Lyk12 | *S. lycopersicum* | RLK | RLK | G3KGB3 |
| RLP | species | RLP type | Class | Uniprot ID |
| NbCSPR | *N. benthamiana* | LRR RP | RLP | NbCSPR |
| ELR | *Solanum tuberosum* | LRR RLP | RLP | ELR |
| CEBiP | *O. sativa* | LysM RLP | RLP | Q8H8C7 |
| LYM1 | *O. sativa* | LysM RLP | RLP | Q93ZH0 |
| LYM2 | *O. sativa* | RLP | RLP | O23006 |
| EIX2 -LeEix2 | *S. lycopersicum* | LRR RLP | RLP | Q6JN46 |
| RLP42 | *A. thaliana* | LRR RLP | RLP | Q9LJS0 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or GeneBank ID |
|---|---|---|---|---|
| ReMAX | *A. thaliana* | LRR RLP | RLP | ReMAX |
| RLK1 | *A. thaliana* | LRR RLP | RLP | A0A178WLY3 |
| Ve1 | *S. lycopersicum* | LRR RLP | RLP | Q94G61 |
| Ve2 | *S. lycopersicum* | LRR RLP | RLP | Q93V91 |
| Cf-2 | *S. pimpinellifolium* | LRR RLP | RLP | Q41397 |
| Cf-4 | *S. pimpinellifolium* | LRR RLP | RLP | O50025 |
| Cf-9 | *S. pimpinellifolium* | LRR RLP | RLP | Q40235 |
| Cf-9B | *S. pimpinellifolium* | LRR RLP | RLP | Q84UH0 |
| Cf-5 | *S. pimpinellifolium* | LRR RLP | RLP | Q9ZTK1 |
| LeEix1 - Eix1 | *S. lycopersicum* | LRR RLP | RLP | Q6JN47 |
| RLP51 / SNC2 | *A.thaliana* | LRR RLP | RLP | Q9SN38 |
| RLP30 / AtRLP30 | *A. thaliana* | LRR RLP | RLP | Q9MA83 |
| RLP52 / AtRLP52 | *A. thaliana* | LRR RLP | RLP | Q7FZR1 |
| Hcr9-4E | *Solanum habrochaites* | RLP | RLP | O50023 |
| Xa27 | *O. sativa* | RLP | RLP | Q4QSA3 |
| Hs1pro-1 | *Patellifolia procumbens* | RLP | RLP | Q9LY61 |
| Hs1$^{pro-1}$ | *P.procumbens* | RLP | RLP | Q45EZ3 |
| LepR3 | *Brassica napus* | RLP | RLP | I7C3X3 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or GeneBank ID |
|---|---|---|---|---|
| LepR3 - variant | *B. napus* | RLP | RLP | J7HQI5 |
| RLM2 | *B. napus* | RLP | RLP | A0A0B5L618 |
| RLP23 | *A. thaliana* | RLP | RLP | O48849 |
| sc-1 | *S. lycopersicum* | Other | R Gene | AF198177 |
| Bs2 | *Capsicum chacoense* | CNL | R Gene | AF202179 |
| Bs4 | *S. lycopersicum* | TNL | R Gene | AY438027 |
| Gpa2 | *S. tuberosum* | CNL | R Gene | AF195939 |
| Gro1.4 | *S. tuberosum* | TNL | R Gene | AY196151 |
| Hero | *S. lycopersicum* | CNL | R Gene | AJ457052 |
| I-2 | *S.lycopersicum* | NL | R Gene | AF118127 |
| Mi1.2 | *S. lycopersicum* | CNL | R Gene | AF039682 |
| N | *Nicotiana glutinosa* | TNL | R Gene | U15605 |
| Prf | *Solanum pimpinellifolium* | CNL | R Gene | AF220602 |
| R1 | *Solanum demissum* | CNL | R Gene | AF447489 |
| R3a | *S. tuberosum* | NL | R Gene | AY849382 |
| Rpi-blb1 | *Solanum bulbocastanum* | CNL | R Gene | AY336128 |
| Rpi-blb2 | *S. bulbocastanum* | CNL | R Gene | DQ122125 |
| Rx2 | *Solanum acaule* | CNL | R Gene | AJ249448 |
| RY-1 | *Solanum tuberosum subsp andigena* | TNL | R Gene | AJ300266 |
| Sw-5 | *S, lycopersicum* | CNL | R Gene | AY007366 |
| Tm-2 | *S. lycopersicum* | CNL | R Gene | AF536200 |
| Dm3 (RGC2B) | *Lactuca sativa* | CNL | R Gene | AH007213 |
| At1 | *Cucumis melo* | Other | R Gene | AY066012 |
| At2 | *C. melo* | Other | R Gene | AF461048 |
| Hm1 | *Zea mays* | Other | R Gene | NM_001112450 |
| HRT | *A. thaliana* | CNL | R Gene | AF234174 |
| Hs1 | *Beta vulgaris* | Other | R Gene | U79733 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or Gene Bank ID |
|---|---|---|---|---|
| RPM1 | *A. thaliana* | CNL | R Gene | NM_111584 |
| RPP13 | *A. thaliana* | CNL | R Gene | NM_114520 |
| RPP5 | *A. thaliana* | TNL | R Gene | NM_117798 |
| Rps2 | *A. thaliana* | CNL | R Gene | NM_118742 |
| Rps4 | *A. thaliana* | TNL | R Gene | NM_123893 |
| RPS5 | *A. thaliana* | CNL | R Gene | NM_101094 |
| Hm2 | *Z. mays* | Other | R Gene | EU367521 |
| Rps1-k-2 | *Glycine max* | NL | R Gene | EU450800 |
| Rps1-k-1 | *G. max* | NL | R Gene | EU450800 |
| Bs3-E | *Capsicum annuum* | Other | R Gene | EU078683 |
| Mlo | *H. vulgare* | Other | R Gene | Z83834 |
| P2 | *Linum usitatissimum* | TNL | R Gene | AF310960 |
| PIB | *O. sativa* | CNL | R Gene | AB013449 |
| XA1 | *O. sativa* | CNL | R Gene | AB002266 |
| RPP1 | *A. thaliana* | TNL | R Gene | NM_114316 |
| RPP4 | *A. thaliana* | TNL | R Gene | NM_117790 |
| RPW8.1 | *A. thaliana* | Other | R Gene | AF273059 |
| RPW8.2 | *A. thaliana* | Other | R Gene | AF273059 |
| RTM1 | *A. thaliana* | Other | R Gene | NM_100456 |
| RTM2 | *A. thaliana* | Other | R Gene | NM_120571 |
| MLA1 | *H. vulgare* | CNL | R Gene | GU245961 |
| Mla12 | *Hordeum vulgare subsp. vulgare* | CNL | R Gene | AY196347 |
| Pi36 | *Oryza sativa Indica Group* | CNL | R Gene | DQ900896 |
| Rp1-D | *Z. mays* | CNL | R Gene | AF107293 |
| Pm3 | *Triticum aestivum* | CNL | R Gene | AY325736 |
| Lr10 | *T. aestivum* | CNL | R Gene | AY270157 |
| Pl8 | *Helianthus annuus* | CNL | R Gene | AY490793 |
| SSI4 | *A. thaliana* | TNL | R Gene | AY179750 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or GeneBankID |
|---|---|---|---|---|
| Xa5 | *O. sativa Indica Group* | Other | R Gene | AY643716 |
| xa27 | *O. sativa Indica Group* | Other | R Gene | AY986491 |
| Xa13 | *O. sativa Indica Group* | Other | R Gene | DQ421396 |
| IVR | *Nicotiana tabacum* | Other | R Gene | AJ009684 |
| Pikm1-TS | *Oryza sativa Japonica Group* | CNL | R Gene | AB462324 |
| Pikm2-TS | *O. sativa Japonica Group* | CNL | R Gene | AB462325 |
| Rdg2a | *Hordeum vulgare subsp. vulgare* | CNL | R Gene | HM124452 |
| Pid3 | *O. sativa Japonica Group* | CNL | R Gene | FJ773286 |
| Pi5-1 | *O. sativa Japonica Group* | CNL | R Gene | EU869185 |
| Pi5-2 | *O. sativa Japonica Group* | CNL | R Gene | EU869186 |
| Pit | *O. sativa Japonica Group* | CNL | R Gene | AB379816 |
| KR1 | *G. max* | TNL | R Gene | AF327903 |
| FOM-2 | *C. melo* | CNL | R Gene | DQ287965 |
| RLM3 | *A. thaliana* | TN | R Gene | NM_001036575 |
| RAC1 | *A. thaliana* | TNL | R Gene | AY522496 |
| Lr21 | *Triticum aestivum* | CNL | R Gene | FJ876280 |
| Lr1 | *T. aestivum* | CNL | R Gene | EF439840 |
| Lr34 | *T. aestivum* | Other | R Gene | HM775493 |
| VAT | *C. melo* | CNL | R Gene | AIU36098 |
| I2C-1 | *Solanum lycopersicum* | NR | R Gene | O24015 |
| NDR1<br>At3g20600<br>K10D20.24 | *A. thaliana* | NR | R Gene | O48915 |
|  | *A. thaliana* | NR | R Gene | Q9ZSN6 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or GeneBank ID |
|---|---|---|---|---|
| RPP5 | *A. thaliana* | NR | R Gene | O04264 |
| RPP13 | *A. thaliana* | NR | R Gene | C0KRS6 |
| Mla1 | *H. vulgare* | NR | R Gene | Q9FE39 |
| Yr10 | *Triticum aestivum* | NR | R Gene | Q9FR64 |
|  | *Linum usitatissimum* | NR | R Gene | Q9ATZ2 |
| Pi37 | *O. sativa Japonica Group* | NR | R Gene | Q06AJ9 |
| WKS2 | *Triticum dicoccoides* | NR | R Gene | B9UN36 |
| RRS1 RCH2 RRS1-R RSH4 SLH1 WRKY52 | *A. thaliana* | NR | R Gene | E1B328 |
| Os11gRGA5 | *O. sativa Japonica Group* | NR | R Gene | F7J0M8 |
| PTI1 | *S. lycopersicum* | NR | R Gene | Q41328 |
| pi54 | *Oryza sativa subsp. indica* | NR | R Gene | J7Q114 |
| Xa10 | *O. sativa subsp. indica* | NR | R Gene | W5QN15 |
| Sr35 | *Triticum monococcum subsp. monococcum* | NR | R Gene | S5ABD6 |
| Sr35-TA | *Triticum aestivum* | NR | R Gene | S5DII7 |
| xa25 | *O. sativa subsp. indica* | NR | R Gene | S5UIA5 |
| xa25-SP | *Solanum pimpinellifolium* | NR | R Gene | A0A060D304 |
| L6 | *Linum usitatissimum* | TNL | R Gene | U27081 |
| M | *L. usitatissimum* | TNL | R Gene | U73916 |
| Pi-ta | *O. sativa* | CNL | R Gene | AY196754 |
| Piz-t | *Oryza sativa Japonica Group* | CNL | R Gene | DQ352040 |

**Table A1. List of RLK, RLP and cytoplasmic resistance proteins experimentally validated. (Continued)**

| Name | Species | Type | Class | Uniprot ID or GeneBank ID |
|---|---|---|---|---|
| Cre1 | *Aegilops tauschii* | CNL | R Gene | AY124651 |
| RPR1<br>LOC_Os11g12340<br>Os11g0229500<br>OsJ_33440<br>OSNPB_110229500 | *Oryza sativa subsp. Japonica* | NR | R Gene | Q9ZQT3 |
| L | *Linum usitatissimum* | NR | R Gene | Q9XEH5 |

NR: not reported

**Table A2. Functional Pfam 31 domains to target RLK and RLP receptors.**

| Clan/Families | Domains - Pfam notation |
|---|---|
| LRR | DUF285, FNIP, LRR_1, LRR_2, LRR_3, LRR_4, LRR_5, LRR_6, LRR_8, LRR_9, Recep_L_domain, LRRNT_2 |
| Pkinase | ABC1, AceK, Act-Frag_cataly, Alpha_kinase, APH, APH_6_hur, Choline_kinase, CotH, DUF1679, DUF2252, DUF4135, EcKinase, Fam20C, Fructosamin_kin, FTA2, Haspin_kinase, HipA_C, Ins_P5_2-kin, IPK, IucA_IucC, Kdo, Kinase-like, KIND, PI3_PI4_kinase, PIP49_C, PIP5K, Pkinase, Pkinase_Tyr, Pox_ser-thr_kin, RIO1, Seadorna_VP7, UL97, WaaY, YrbL-PhoP_reg, YukC, Pkinase_C |
| L-lectin | Alginate_lyase2, ArabFuran-catal, Bac_rhamnosid, Calreticulin, Cleaved_Adhesin, DUF1080, DUF1349, DUF1583, DUF1961, DUF2401, DUF3472, DUF4975, Exotox-A_bind, Gal-bind_lectin, Glyco_hydro_11, Glyco_hydro_12, Glyco_hydro_16, Glyco_hydro_32C, Glyco_hydro_7, Laminin_G_1, Laminin_G_2, Laminin_G_3, Lectin_leg-like, Lectin_legB, MAM, Methyltransf_FA, Neuralized, Pentaxin, Peptidase_A4, Polysacc_lyase, PRY, Reoviridae_Vp9, Sial-lect-inser, Sialidase, SKN1, Spike_NTD, SPRY, TgMIC1, Toxin_R_bind_N, TSP_C, VP4_haemagglut, XET_C, YrpD |
| C-lectin | Lectin_C |
| G-lectin-B | B_lectin |
| G-lectin-S | S_locus_glycop |
| PAN | AMA-1, MANEC, PAN_1, PAN_2, PAN_3, PAN_4 |
| LysM | LysM, OapA, Phage_tail_X |
| PR5K | Thaumatin |
| TNFR | BaffR-Tall_bind, BCMA-Tall_bind, NCD3G, stn_TNFRSF12A, TACI-CRD2, TNFR_c6 |
| WAK | WAK, GUB_WAK_bind, WAK_assoc |
| Malectin | Malectin, Malectin_like |
| EGF | cEGF, CFC, DSL, EGF, EGF_2, EGF_3, EGF_alliinase, EGF_CA, EGF_MSP1_1, FOLN, FXa_inhibition, Gla, hEGF, Laminin_EGF, Plasmod_Pvs28, Sushi, Sushi_2, Tme5_EGF_like |

**Table A2. Functional Pfam 31 domains to target RLK and RLP receptors. (Continued)**

| Clan/families | Domains Pfam notation |
|---|---|
| Stress_antifung | Stress-antifung |
| NB-ARC | 6PF2K, AAA, AAA-ATPase_like, AAA_10, AAA_11, AAA_12,AAA_13, AAA_14, AAA_15, AAA_16, AAA_17, AAA_18,AAA_19, AAA_2, AAA_21, AAA_22, AAA_23, AAA_24,AAA_25, AAA_26, AAA_27, AAA_28, AAA_29, AAA_3, AAA_30,AAA_31, AAA_32, AAA_33, AAA_34, AAA_35, AAA_5, AAA_6,AAA_7, AAA_8, AAA_9, AAA_PrkA, ABC_ATPase, ABC_tran,ABC_tran_Xtn, Adeno_IVa2, Adenylsucc_synt, ADK,AFG1_ATPase, AIG1, APS_kinase, Arf, ArgK, ArsA_ATPase,ATP-synt_ab, ATP_bind_1, ATP_bind_2, ATPase, ATPase_2,Bac_DnaA, BCA_ABC_TP_C, Beta-Casp, Cas_Csn2, Cas_St_Csn2,CbiA, CBP_BcsQ, CDC73_C, CENP-M, CFTR_R, CLP1_P, CMS1,CoaE, CobA_CobO_BtuR, CobU, cobW, CPT, CSM2,CTP_synth_N, Cytidylate_kin, Cytidylate_kin2, DAP3, DEAD,DEAD_2, DLIC, DNA_pack_C, DNA_pack_N, DNA_pol3_delta,DNA_pol3_delta2, DnaB_C, dNK, DUF1611, DUF1726,DUF2075, DUF2326, DUF2478, DUF257, DUF2791, DUF2813,DUF3584, DUF463, DUF815, DUF853, DUF87, DUF927,Dynamin_N, Dynein_heavy, ERCC3_RAD25_C, Exonuc_V_gamma,FeoB_N, Fer4_NifH, Flavi_DEAD, FTHFS, FtsK_SpoIIIE,G-alpha, Gal-3-0_sulfotr, GBP, GBP_C, GTP_EFTU,Gtr1_RagA, Guanylate_kin, GvpD, HDA2-3, Helicase_C,Helicase_C_2, Helicase_C_4, Helicase_RecD, Herpes_Helicase,Herpes_ori_bp, Herpes_TK, Hydin_ADK, IIGP, IPPT, IPT,IstB_IS21, KAP_NTPase, KdpD, Kinesin, KTI12, LAP1C,Lon_2, LpxK, MCM, MEDS, Mg_chelatase, Microtub_bd,MipZ, MMR_HSR1, MMR_HSR1_C, MobB, MukB, MutS_V,Myosin_head, NACHT, NB-ARC, NOG1, NTPase_1, NTPase_P4,ORC3_N, ParA, Parvo_NS1, PAXNEB, PduV-EutP, PhoH,PIF1, Podovirus_Gp16, Polyoma_lg_T_C, Pox_A32, PPK2,PPV_E1_C, PRK, PSY3, Rad17, Rad51, Ras, RecA,ResIII, RHD3, RHSP, RNA12, RNA_helicase, Roc,RsgA_GTPase, RuvB_N, SbcCD_C, SecA_DEAD, Septin,Sigma54_activ_2, Sigma54_activat, SKI, SMC_N, SNF2_N,Spore_IV_A, SRP54, SRPRB, SulA, Sulfotransfer_1,Sulfotransfer_2, Sulfotransfer_3, Sulphotransf, T2SSE,T4SS-DNA_transf, Terminase_1, Terminase_3, Terminase_6,Terminase_GpA, Thymidylate_kin, TIP49, TK, TniB, Torsin,TraG-D_C, tRNA_lig_kinase, TrwB_AAD_bind, TsaE, UvrB,UvrD-helicase, UvrD_C, UvrD_C_2, Viral_helicase1, VirC1,VirE, Zeta_toxin, Zot |

The Pfam 31 families were selected in June 20 – 2017.

**Table A3. RLK and RLP performance evaluation of the proposed approach.**

| | | Datasets | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Process | | RLK (no RLCK) | | | | RLP | | | | cytoplasmic resistance proteins | | | |
| Initial # proteins | | 66 | | | | 28 | | | | 97 | | | |
| CD-HIT (90% identity reduction) | | 64 | | | | 26 | | | | 97 | | | |
| SignalP | | Pos | | Neg | | Pos | | Neg | | Pos | | Neg | |
| | | 54 | | 10 | | 21 | | 5 | | 4 | | 93 | |
| TMHMM | | P | A | P | A | P | A | P | A | P | A | P | A |
| | | 51 | 3 | 5 | 5 | 20 | 1 | 3 | 2 | 0 | 4 | 11 | 82 |
| RLK prediction | rlk_lrr | 24 | - | 3 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_l-lectine | 4 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_c-lectin | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_g-lectin | 3 | - | 1 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_lysm | 4 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_pr5k | 1 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_tnfr | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_wak | 3 | - | 1 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_malectin | 4 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_egf | 1 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_stress_antifug | 1 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_extra | 9 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlk_total | 54* | - | 5 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| RLP prediction | rlp_lrr | 0 | - | 0 | - | 17 | - | 2 | - | 0 | - | 0 | - |
| | rlp_l-lectine | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_c-lectin | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_g-lectin | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_lysm | 0 | - | 0 | - | 3 | - | 0 | - | 0 | - | 0 | - |
| | rlp_pr5k | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_tnfr | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |

**Table A3. RLK and RLP performance evaluation of the proposed approach. (Continued)**

| | | Datasets | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RLP prediction | rlp_wak | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_malectin | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_egf | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_stress_antifug | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - | 0 | - |
| | rlp_total | 0 | - | 0 | - | 20 | - | 2 | - | 0 | - | 0 | - |

P: presence, A: absence, Symbol "-": Not Applicable. Redundant*: rlk_lrr-malectin: sp|O64483|, sp|Q9C8I6|; rlk_wak-egf: sp|Q39191|, RLK total non-redundant: 51. True positives: TPs-Protein predicted as RLK* being an RLK*. True negatives: TNs-Protein predicted as non-RLK* being non-RLK*. False positives: FPs-Proteins predicted as RLK* being non-RLK*. False negatives: FNs-Protein predicted as non-RLK* being RLK*, defined by the contingency matrix.

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes.**

| SP | Protein IDs |
|---|---|
| CC | KYP31256.1, KYP32195.1, KYP32335.1, KYP32585.1, KYP32710.1, KYP32720.1, KYP32785.1, KYP32945.1, KYP33015.1, KYP33093.1, KYP33136.1, KYP33297.1, KYP33576.1, KYP33714.1, KYP33801.1, KYP33885.1, KYP33944.1, KYP34018.1, KYP34039.1, KYP34144.1, KYP34244.1, KYP34373.1, KYP34388.1, KYP34445.1, KYP34446.1, KYP35096.1, KYP35443.1, KYP35494.1, KYP35495.1, KYP35525.1, KYP35526.1, KYP35529.1, KYP35636.1, KYP35781.1, KYP35808.1, KYP35814.1, KYP35827.1, KYP35855.1, KYP35876.1, KYP35942.1, KYP36073.1, KYP36121.1, KYP36165.1, KYP36387.1, KYP36430.1, KYP36452.1, KYP36960.1, KYP37049.1, KYP37252.1, KYP37288.1, KYP37300.1, KYP37508.1, KYP37520.1, KYP37535.1, KYP37606.1, KYP37745.1, KYP38129.1, KYP38134.1, KYP38183.1, KYP38204.1, KYP38233.1, KYP39083.1, KYP39085.1, KYP39683.1, KYP39741.1, KYP39912.1, KYP39987.1, KYP40006.1, KYP40063.1, KYP40093.1, KYP40097.1, KYP40210.1, KYP40221.1, KYP40227.1, KYP40228.1, KYP40353.1, KYP40366.1, KYP40368.1, KYP40489.1, KYP40555.1, KYP40683.1, KYP40783.1, KYP40903.1, KYP41214.1, KYP41368.1, KYP41708.1, KYP41872.1, KYP42384.1, KYP42388.1, KYP42394.1, KYP42687.1, KYP42794.1, KYP42964.1, KYP43180.1, KYP43193.1, KYP43195.1, KYP43367.1, KYP43510.1, KYP43646.1, KYP43931.1, KYP44116.1, KYP44286.1, KYP44551.1, KYP44686.1, KYP44714.1, KYP44953.1, KYP45044.1, KYP45613.1, KYP45645.1, KYP45708.1, KYP45725.1, KYP46021.1, KYP46087.1, KYP46128.1, KYP46134.1, KYP46138.1, KYP46186.1, KYP46290.1, KYP46641.1, KYP46642.1, KYP46649.1, KYP46650.1, KYP46759.1, KYP46973.1, KYP47124.1, KYP47226.1, KYP47340.1, KYP47588.1, KYP47615.1, KYP47938.1, KYP48053.1, KYP48120.1, KYP48470.1, KYP48555.1, KYP48687.1, KYP49267.1, KYP49578.1, KYP49602.1, KYP49935.1, KYP50006.1, KYP50028.1, KYP50038.1, KYP50209.1, KYP50310.1, KYP50371.1, KYP50376.1, KYP50390.1, KYP50432.1, KYP50436.1, KYP50467.1, KYP50590.1, KYP50651.1, KYP50703.1, KYP51046.1, KYP51074.1, KYP51141.1, KYP51256.1, KYP51314.1, KYP51443.1, KYP51453.1, KYP51615.1, KYP51666.1, KYP51889.1, KYP51907.1, KYP52032.1, KYP52109.1, KYP52236.1, KYP52775.1, KYP52986.1, KYP53378.1, KYP53605.1, KYP53881.1, KYP54064.1, KYP54108.1, KYP54138.1, KYP54193.1, KYP54221.1, KYP54334.1, KYP54555.1, KYP54561.1, KYP54586.1, KYP54822.1, KYP54891.1, KYP54906.1, KYP55186.1, KYP55299.1, KYP55455.1, KYP55493.1, KYP55673.1, KYP56034.1, KYP56449.1, KYP56487.1, KYP56627.1, KYP56629.1, KYP56764.1, KYP56802.1, KYP57042.1, KYP57100.1, KYP57249.1, KYP57303.1, KYP57429.1, KYP57545.1, KYP57822.1, KYP57988.1, KYP58327.1, KYP58350.1, KYP58374.1, KYP58594.1, KYP58768.1, KYP58792.1, KYP59058.1, KYP59304.1, KYP59535.1, KYP60048.1, KYP60216.1, KYP60552.1, KYP60697.1, KYP60951.1, KYP61088.1, KYP61120.1, KYP61126.1, KYP61211.1, KYP61214.1, KYP61234.1, KYP61274.1, KYP61360.1, KYP61400.1, KYP61455.1, KYP61494.1, KYP61708.1, KYP61754.1, KYP61816.1, KYP61829.1, KYP61831.1, KYP61934.1, KYP61948.1, KYP62409.1, KYP62423.1, KYP62486.1, KYP62488.1, KYP62579.1, KYP62580.1, KYP62729.1, KYP62732.1, KYP62748.1, KYP62912.1, KYP62986.1, KYP63091.1, KYP63178.1, KYP63285.1, KYP63286.1, KYP63298.1, KYP63368.1, KYP63372.1, KYP63624.1, KYP64044.1, KYP64138.1, KYP64257.1, KYP64749.1, KYP64842.1, KYP64858.1, KYP64907.1, KYP65019.1, KYP65188.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| CC | KYP65610.1, KYP65954.1, KYP66071.1, KYP66135.1, KYP66142.1, KYP66160.1, KYP66267.1, KYP66519.1, KYP66589.1, KYP66686.1, KYP66704.1, KYP66745.1, KYP66833.1, KYP66886.1, KYP66978.1, KYP67205.1, KYP67218.1, KYP67321.1, KYP67551.1, KYP67577.1, KYP67879.1, KYP67928.1, KYP68267.1, KYP68357.1, KYP68441.1, KYP68549.1, KYP68672.1, KYP68785.1, KYP68786.1, KYP68835.1, KYP69249.1, KYP69270.1, KYP69341.1, KYP69368.1, KYP69465.1, KYP69577.1, KYP69721.1, KYP69786.1, KYP69809.1, KYP70032.1, KYP70121.1, KYP70226.1, KYP70247.1, KYP70261.1, KYP70273.1, KYP70523.1, KYP70559.1, KYP70570.1, KYP70591.1, KYP70762.1, KYP70838.1, KYP70949.1, KYP70963.1, KYP71025.1, KYP71273.1, KYP71527.1, KYP71578.1, KYP71602.1, KYP71791.1, KYP71861.1, KYP72017.1, KYP72152.1, KYP72155.1, KYP72415.1, KYP72773.1, KYP73004.1, KYP73008.1, KYP73081.1, KYP73099.1, KYP73346.1, KYP73477.1, KYP73622.1, KYP74022.1, KYP74123.1, KYP74155.1, KYP74355.1, KYP74482.1, KYP74494.1, KYP74610.1, KYP74764.1, KYP74819.1, KYP74967.1, KYP75089.1, KYP75093.1, KYP75139.1, KYP75140.1, KYP75141.1, KYP75171.1, KYP75175.1, KYP75215.1, KYP75225.1, KYP75226.1, KYP75227.1, KYP75229.1, KYP75233.1, KYP75273.1, KYP75327.1, KYP75388.1, KYP75389.1, KYP75723.1, KYP75784.1, KYP75863.1, KYP75987.1, KYP76039.1, KYP76040.1, KYP76087.1, KYP76104.1, KYP76322.1, KYP76463.1, KYP76571.1, KYP76668.1, KYP76807.1, KYP76845.1, KYP76893.1, KYP77140.1, KYP78105.1, KYP78240.1, KYP78852.1 |
| GM | Glyma.01G001800.1.p, Glyma.01G007400.1.p, Glyma.01G007500.1.p, Glyma.01G013500.1.p, Glyma.01G027100.1.p, Glyma.01G028000.1.p, Glyma.01G028100.1.p, Glyma.01G028100.2.p, Glyma.01G028100.3.p, Glyma.01G028100.4.p, Glyma.01G028700.1.p, Glyma.01G030200.1.p, Glyma.01G033600.1.p, Glyma.01G043700.1.p, Glyma.01G062900.1.p, Glyma.01G071700.1.p, Glyma.01G083200.1.p, Glyma.01G090500.1.p, Glyma.01G093100.1.p, Glyma.01G099900.1.p, Glyma.01G101000.1.p, Glyma.01G118400.1.p, Glyma.01G118400.2.p, Glyma.01G118400.3.p, Glyma.01G118400.4.p, Glyma.01G125200.1.p, Glyma.01G125800.1.p, Glyma.01G125800.2.p, Glyma.01G133300.1.p, Glyma.01G133300.2.p, Glyma.01G150700.1.p, Glyma.01G150700.2.p, Glyma.01G150700.3.p, Glyma.01G155600.1.p, Glyma.01G155600.2.p, Glyma.01G157800.1.p, Glyma.01G168100.1.p, Glyma.01G175000.1.p, Glyma.01G175000.2.p, Glyma.01G179000.1.p, Glyma.01G179100.1.p, Glyma.01G182400.1.p, Glyma.01G182400.2.p, Glyma.01G187300.1.p, Glyma.01G197600.1.p, Glyma.01G197800.1.p, Glyma.01G206500.1.p, Glyma.01G211900.1.p, Glyma.01G211900.2.p, Glyma.01G213600.1.p, Glyma.01G213600.2.p, Glyma.01G218800.2.p, Glyma.01G224000.1.p, Glyma.01G224000.2.p, Glyma.01G240800.1.p, Glyma.01G240900.1.p, Glyma.01G240900.2.p, Glyma.01G241000.1.p, Glyma.01G241000.2.p, Glyma.01G241000.3.p, Glyma.02G000400.1.p, Glyma.02G003200.1.p, Glyma.02G011400.1.p, Glyma.02G011400.2.p, Glyma.02G019800.1.p, Glyma.02G019800.2.p, Glyma.02G024400.1.p, Glyma.02G035300.1.p, Glyma.02G036500.1.p, Glyma.02G036500.2.p, Glyma.02G037100.1.p, Glyma.02G037100.2.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDS |
|---|---|
| GM | Glyma.02G037100.3.p, Glyma.02G037100.4.p, Glyma.02G037100.5.p, Glyma.02G037100.6.p, Glyma.02G037100.7.p, Glyma.02G037100.8.p, Glyma.02G037200.1.p, Glyma.02G042900.1.p, Glyma.02G043000.1.p, Glyma.02G044600.1.p, Glyma.02G050300.1.p, Glyma.02G057700.1.p, Glyma.02G059700.1.p, Glyma.02G061700.1.p, Glyma.02G061700.2.p, Glyma.02G061700.3.p, Glyma.02G061700.4.p, Glyma.02G061700.5.p, Glyma.02G061700.6.p, Glyma.02G076100.1.p, Glyma.02G088700.1.p, Glyma.02G097400.1.p, Glyma.02G097400.2.p, Glyma.02G100300.1.p, Glyma.02G100400.1.p, Glyma.02G102600.1.p, Glyma.02G120800.1.p, Glyma.02G121900.1.p, Glyma.02G122000.1.p, Glyma.02G122000.2.p, Glyma.02G127800.1.p, Glyma.02G127800.2.p, Glyma.02G127800.3.p, Glyma.02G127800.4.p, Glyma.02G129300.1.p, Glyma.02G150600.1.p, Glyma.02G156400.1.p, Glyma.02G172200.1.p, Glyma.02G175500.1.p, Glyma.02G177000.1.p, Glyma.02G177000.2.p, Glyma.02G183300.1.p, Glyma.02G197000.1.p, Glyma.02G203600.1.p, Glyma.02G206300.1.p, Glyma.02G207400.1.p, Glyma.02G207400.2.p, Glyma.02G207400.3.p, Glyma.02G219800.1.p, Glyma.02G221900.1.p, Glyma.02G228300.1.p, Glyma.02G229500.1.p, Glyma.02G237000.1.p, Glyma.02G237000.2.p, Glyma.02G237000.3.p, Glyma.02G237200.1.p, Glyma.02G237300.1.p, Glyma.02G237300.2.p, Glyma.02G244400.1.p, Glyma.02G257300.1.p, Glyma.02G261400.1.p, Glyma.02G263900.1.p, Glyma.02G269200.1.p, Glyma.02G269600.1.p, Glyma.02G270700.1.p, Glyma.02G270800.1.p, Glyma.02G274500.1.p, Glyma.02G274500.10.p, Glyma.02G274500.11.p, Glyma.02G274500.2.p, Glyma.02G274500.3.p, Glyma.02G274500.4.p, Glyma.02G274500.5.p, Glyma.02G274500.6.p, Glyma.02G274500.7.p, Glyma.02G274500.8.p, Glyma.02G274500.9.p, Glyma.02G275900.1.p, Glyma.02G281100.1.p, Glyma.02G286000.1.p, Glyma.02G286000.10.p, Glyma.02G286000.11.p, Glyma.02G286000.9.p, Glyma.02G297100.1.p, Glyma.02G302600.1.p, Glyma.02G308600.1.p, Glyma.03G020700.1.p, Glyma.03G023700.1.p, Glyma.03G023700.2.p, Glyma.03G027200.1.p, Glyma.03G034200.1.p, Glyma.03G047100.1.p, Glyma.03G047100.2.p, Glyma.03G049300.1.p, Glyma.03G051100.1.p, Glyma.03G055000.1.p, Glyma.03G068600.1.p, Glyma.03G069300.1.p, Glyma.03G090200.1.p, Glyma.03G090500.1.p,Glyma.03G095700.1.p, Glyma.03G096000.1.p, Glyma.03G096100.1.p, Glyma.03G096600.1.p, Glyma.03G137300.1.p, Glyma.03G140700.1.p, Glyma.03G149000.1.p, Glyma.03G149000.2.p, Glyma.03G165700.1.p, Glyma.03G165800.1.p, Glyma.03G166200.1.p, Glyma.03G166300.1.p, Glyma.03G167500.1.p, Glyma.03G169200.1.p, Glyma.03G188700.1.p, Glyma.03G189800.1.p, Glyma.03G202000.1.p, Glyma.03G220500.1.p, Glyma.03G220500.2.p, Glyma.03G220500.3.p, Glyma.03G228800.1.p, Glyma.03G247800.1.p, Glyma.03G253600.2.p, Glyma.03G261900.1.p, Glyma.03G261900.2.p, Glyma.03G261900.3.p, Glyma.03G262500.1.p, Glyma.03G262700.1.p, Glyma.04G012800.1.p, Glyma.04G012800.2.p, Glyma.04G012800.3.p, Glyma.04G012800.4.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.04G012900.1.p, Glyma.04G013300.1.p, Glyma.04G026700.1.p, Glyma.04G035100.1.p, Glyma.04G035100.2.p, Glyma.04G041400.1.p, Glyma.04G041400.2.p, Glyma.04G041400.3.p, Glyma.04G041400.4.p, Glyma.04G041400.5.p, Glyma.04G063100.1.p, Glyma.04G063100.2.p, Glyma.04G063100.3.p, Glyma.04G066700.1.p, Glyma.04G077100.1.p, Glyma.04G080100.1.p, Glyma.04G085200.1.p, Glyma.04G086700.1.p, Glyma.04G088700.1.p, Glyma.04G088800.1.p, Glyma.04G115700.1.p, Glyma.04G117200.1.p, Glyma.04G127100.1.p, Glyma.04G127100.2.p, Glyma.04G127100.3.p, Glyma.04G149800.1.p, Glyma.04G165400.1.p, Glyma.04G170000.1.p, Glyma.04G174700.1.p, Glyma.04G180300.1.p, Glyma.04G190400.1.p, Glyma.04G218300.1.p, Glyma.04G220200.1.p, Glyma.04G220200.2.p, Glyma.04G220200.3.p, Glyma.04G220400.1.p, Glyma.04G222800.1.p, Glyma.04G223800.1.p, Glyma.04G223800.2.p, Glyma.04G223800.3.p, Glyma.04G223800.4.p, Glyma.04G230300.1.p, Glyma.04G230500.1.p, Glyma.04G239000.1.p, Glyma.04G239000.2.p, Glyma.04G239700.1.p, Glyma.04G244100.1.p, Glyma.05G005100.1.p, Glyma.05G009900.1.p, Glyma.05G011600.1.p, Glyma.05G024000.1.p, Glyma.05G030100.1.p, Glyma.05G039900.1.p, Glyma.05G041300.1.p, Glyma.05G052700.1.p, Glyma.05G071100.1.p, Glyma.05G075800.1.p, Glyma.05G085500.1.p, Glyma.05G097400.1.p, Glyma.05G099900.1.p, Glyma.05G100000.1.p, Glyma.05G104200.1.p, Glyma.05G110400.1.p, Glyma.05G114100.1.p, Glyma.05G116100.1.p, Glyma.05G119500.1.p, Glyma.05G119500.2.p, Glyma.05G119600.1.p, Glyma.05G119600.2.p, Glyma.05G119600.3.p, Glyma.05G119600.4.p, Glyma.05G119600.5.p, Glyma.05G124200.1.p, Glyma.05G126400.1.p, Glyma.05G126400.2.p, Glyma.05G128200.1.p, Glyma.05G134800.1.p, Glyma.05G136900.1.p, Glyma.05G144400.1.p, Glyma.05G144400.2.p, Glyma.05G144400.3.p, Glyma.05G144400.4.p, Glyma.05G150800.1.p, Glyma.05G158800.1.p, Glyma.05G158800.2.p, Glyma.05G158800.3.p, Glyma.05G159200.1.p, Glyma.05G168800.1.p, Glyma.05G170600.1.p, Glyma.05G177500.1.p, Glyma.05G177500.2.p, Glyma.05G177500.3.p, Glyma.05G177500.4.p, Glyma.05G177500.5.p, Glyma.05G183100.1.p, Glyma.05G206300.1.p, Glyma.05G214300.1.p, Glyma.05G214300.2.p, Glyma.05G214300.3.p, Glyma.05G220900.1.p, Glyma.05G237100.1.p, Glyma.05G247300.1.p, Glyma.06G012700.1.p, Glyma.06G012700.2.p, Glyma.06G012800.1.p, Glyma.06G026600.1.p, Glyma.06G035000.1.p, Glyma.06G042500.1.p, Glyma.06G056400.2.p, Glyma.06G064400.1.p, Glyma.06G064400.2.p, Glyma.06G064400.3.p, Glyma.06G068100.1.p, Glyma.06G081800.1.p, Glyma.06G081800.2.p, Glyma.06G081800.3.p, Glyma.06G086800.1.p, Glyma.06G086800.2.p, Glyma.06G088400.1.p, Glyma.06G090700.1.p, Glyma.06G090800.1.p, Glyma.06G110700.1.p, Glyma.06G119200.1.p, Glyma.06G124200.1.p, Glyma.06G124700.1.p, Glyma.06G124700.2.p, Glyma.06G134700.1.p, Glyma.06G141100.1.p, Glyma.06G141100.2.p, Glyma.06G141100.3.p, Glyma.06G142500.1.p, Glyma.06G145500.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.06G145700.1.p, Glyma.06G145700.2.p, Glyma.06G147600.1.p, Glyma.06G171700.1.p, Glyma.06G171700.2.p, Glyma.06G171700.3.p, Glyma.06G175100.1.p, Glyma.06G184400.1.p, Glyma.06G190000.1.p, Glyma.06G190000.2.p, Glyma.06G197600.1.p, Glyma.06G211600.1.p, Glyma.06G215600.1.p, Glyma.06G216800.2.p, Glyma.06G221400.1.p, Glyma.06G230800.1.p, Glyma.06G238700.1.p, Glyma.06G243800.1.p, Glyma.06G256500.1.p, Glyma.06G265600.1.p, Glyma.06G265600.2.p, Glyma.06G265600.3.p, Glyma.06G271500.1.p, Glyma.06G288600.1.p, Glyma.06G292100.1.p, Glyma.06G296800.1.p, Glyma.06G312000.1.p, Glyma.06G312000.2.p, Glyma.06G320600.1.p, Glyma.07G001100.1.p, Glyma.07G004600.1.p, Glyma.07G004700.1.p, Glyma.07G009800.1.p, Glyma.07G009800.2.p, Glyma.07G015300.1.p, Glyma.07G041200.1.p, Glyma.07G046800.1.p, Glyma.07G047200.1.p, Glyma.07G082600.1.p, Glyma.07G085600.1.p, Glyma.07G094100.1.p, Glyma.07G094200.1.p, Glyma.07G094500.1.p, Glyma.07G094500.2.p, Glyma.07G094500.3.p, Glyma.07G094500.4.p, Glyma.07G094500.6.p, Glyma.07G094700.1.p, Glyma.07G095100.1.p, Glyma.07G095400.1.p, Glyma.07G095500.1.p, Glyma.07G095800.2.p, Glyma.07G095900.1.p, Glyma.07G103500.1.p, Glyma.07G130700.1.p, Glyma.07G135400.1.p, Glyma.07G135600.1.p, Glyma.07G136900.1.p, Glyma.07G136900.2.p, Glyma.07G137000.1.p, Glyma.07G147800.1.p, Glyma.07G147800.2.p, Glyma.07G154100.1.p, Glyma.07G155800.1.p, Glyma.07G156000.1.p, Glyma.07G173000.1.p, Glyma.07G173100.1.p, Glyma.07G183900.1.p, Glyma.07G184000.1.p, Glyma.07G189200.1.p, Glyma.07G189200.2.p, Glyma.07G192200.1.p, Glyma.07G201600.1.p, Glyma.07G201700.1.p, Glyma.07G213900.1.p, Glyma.07G234500.1.p, Glyma.07G270700.1.p, Glyma.07G270800.1.p, Glyma.08G013200.1.p, Glyma.08G013200.2.p, Glyma.08G020800.1.p, Glyma.08G020800.2.p, Glyma.08G020800.3.p, Glyma.08G020800.4.p, Glyma.08G020800.5.p, Glyma.08G020800.6.p, Glyma.08G020800.7.p, Glyma.08G020800.8.p, Glyma.08G027400.1.p, Glyma.08G044200.1.p, Glyma.08G044300.1.p, Glyma.08G048700.1.p, Glyma.08G055700.1.p, Glyma.08G060100.1.p, Glyma.08G060100.2.p, Glyma.08G060100.3.p, Glyma.08G061900.1.p, Glyma.08G061900.10.p, Glyma.08G061900.2.p, Glyma.08G061900.3.p, Glyma.08G061900.4.p, Glyma.08G061900.5.p, Glyma.08G061900.6.p, Glyma.08G061900.7.p, Glyma.08G061900.8.p, Glyma.08G061900.9.p, Glyma.08G062800.1.p, Glyma.08G062900.1.p, Glyma.08G065500.1.p, Glyma.08G065800.1.p, Glyma.08G065900.1.p, Glyma.08G066000.2.p, Glyma.08G066000.3.p, Glyma.08G066100.1.p, Glyma.08G066200.1.p, Glyma.08G074500.1.p, Glyma.08G074500.2.p, Glyma.08G075200.1.p, Glyma.08G079400.1.p |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.08G083000.1.p, Glyma.08G083300.1.p, Glyma.08G090000.1.p, Glyma.08G092200.1.p, Glyma.08G093100.1.p, Glyma.08G094400.1.p, Glyma.08G094400.2.p, Glyma.08G094400.3.p, Glyma.08G100800.1.p, Glyma.08G100800.2.p, Glyma.08G100800.3.p, Glyma.08G107700.1.p, Glyma.08G107700.2.p, Glyma.08G116600.1.p, Glyma.08G116600.2.p, Glyma.08G116600.3.p, Glyma.08G116600.4.p, Glyma.08G117000.1.p, Glyma.08G123800.1.p, Glyma.08G125800.1.p, Glyma.08G125800.2.p, Glyma.08G125800.3.p, Glyma.08G127400.1.p, Glyma.08G128800.1.p, Glyma.08G128900.1.p, Glyma.08G134600.1.p, Glyma.08G134600.2.p, Glyma.08G134600.3.p, Glyma.08G134600.4.p, Glyma.08G134600.5.p, Glyma.08G140900.1.p, Glyma.08G140900.2.p, Glyma.08G174500.1.p, Glyma.08G179500.1.p, Glyma.08G180800.1.p, Glyma.08G180800.2.p, Glyma.08G187400.1.p, Glyma.08G187400.2.p, Glyma.08G193000.1.p, Glyma.08G197700.1.p, Glyma.08G197800.1.p, Glyma.08G200800.1.p, Glyma.08G225900.1.p, Glyma.08G225900.2.p, Glyma.08G229100.1.p, Glyma.08G229100.2.p, Glyma.08G231100.1.p, Glyma.08G246100.1.p, Glyma.08G248900.1.p, Glyma.08G249200.1.p, Glyma.08G249400.1.p, Glyma.08G252100.1.p, Glyma.08G253100.1.p, Glyma.08G255300.1.p, Glyma.08G256200.1.p, Glyma.08G256200.2.p, Glyma.08G257700.1.p, Glyma.08G266100.1.p, Glyma.08G266100.2.p, Glyma.08G275400.1.p, Glyma.08G283300.1.p, Glyma.08G284100.1.p, Glyma.08G284100.2.p, Glyma.08G284200.1.p, Glyma.08G284200.2.p, Glyma.08G286600.1.p, Glyma.08G297400.1.p, Glyma.08G302900.1.p, Glyma.08G308900.1.p, Glyma.08G308900.2.p, Glyma.08G308900.3.p, Glyma.08G333200.1.p, Glyma.08G357200.1.p, Glyma.08G357400.1.p, Glyma.08G357400.2.p, Glyma.08G357500.1.p, Glyma.09G003100.1.p, Glyma.09G007200.1.p, Glyma.09G018800.1.p, Glyma.09G018800.2.p, Glyma.09G018900.1.p, Glyma.09G024700.1.p, Glyma.09G024700.2.p, Glyma.09G024900.1.p, Glyma.09G024900.2.p, Glyma.09G024900.3.p, Glyma.09G027500.1.p, Glyma.09G027600.1.p, Glyma.09G027700.1.p, Glyma.09G027800.1.p, Glyma.09G027900.1.p, Glyma.09G048300.1.p, Glyma.09G050300.1.p, Glyma.09G055900.1.p, Glyma.09G055900.2.p, Glyma.09G062500.1.p, Glyma.09G062500.4.p, Glyma.09G063200.1.p, Glyma.09G063200.2.p, Glyma.09G063200.3.p, Glyma.09G063200.4.p, Glyma.09G063200.5.p, Glyma.09G063200.6.p, Glyma.09G066600.1.p, Glyma.09G072500.1.p, Glyma.09G072500.2.p, Glyma.09G081800.1.p, Glyma.09G081800.2.p, Glyma.09G081800.3.p, , Glyma.09G097500.1.p, Glyma.09G097500.2.p, Glyma.09G110500.1.p, Glyma.09G110700.1.p |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.09G115300.1.p, Glyma.09G115300.2.p, Glyma.09G116300.1.p, Glyma.09G116300.2.p, Glyma.09G119600.1.p, Glyma.09G133000.1.p, Glyma.09G150100.1.p, Glyma.09G150100.2.p, Glyma.09G150100.3.p, Glyma.09G150100.4.p, Glyma.09G150500.1.p, Glyma.09G151400.1.p, Glyma.09G151400.2.p, Glyma.09G151400.3.p, Glyma.09G151400.4.p, Glyma.09G151400.5.p, Glyma.09G151400.6.p, Glyma.09G151400.7.p, Glyma.09G151400.8.p, Glyma.09G151400.9.p, Glyma.09G151700.1.p, Glyma.09G151700.2.p, Glyma.09G151700.3.p, Glyma.09G151700.4.p, Glyma.09G152400.1.p, Glyma.09G152400.2.p, Glyma.09G154600.1.p, Glyma.09G160600.1.p, Glyma.09G161100.1.p, Glyma.09G173800.1.p, Glyma.09G181600.1.p, Glyma.09G191300.1.p, Glyma.09G215700.1.p, Glyma.09G215700.2.p, Glyma.09G215700.3.p, Glyma.09G215700.4.p, Glyma.09G216400.1.p, Glyma.09G216400.2.p, Glyma.09G217200.1.p, Glyma.09G217500.1.p, Glyma.09G230300.1.p, Glyma.09G239700.1.p, Glyma.09G241500.1.p, Glyma.09G243600.1.p, Glyma.09G246600.1.p, Glyma.09G246600.2.p, Glyma.09G253200.1.p, Glyma.09G255700.1.p, Glyma.09G272300.1.p, Glyma.09G272300.2.p, Glyma.09G272300.3.p, Glyma.09G272300.4.p, Glyma.09G272300.5.p, Glyma.09G272300.6.p, Glyma.09G272300.7.p, Glyma.09G272300.8.p, Glyma.09G272900.1.p, Glyma.09G272900.2.p, Glyma.09G272900.3.p, Glyma.09G273300.1.p, Glyma.09G274600.1.p, Glyma.10G012000.1.p, Glyma.10G023400.1.p, Glyma.10G040500.1.p, Glyma.10G041300.1.p, Glyma.10G041300.2.p, Glyma.10G041300.3.p, Glyma.10G048800.1.p, Glyma.10G052600.1.p, Glyma.10G065200.1.p, Glyma.10G069500.1.p, Glyma.10G069500.2.p, Glyma.10G081300.1.p, Glyma.10G091400.1.p, Glyma.10G109200.1.p, Glyma.10G109200.3.p, Glyma.10G126700.1.p, Glyma.10G126700.2.p, Glyma.10G126700.3.p, Glyma.10G144200.1.p, Glyma.10G144200.2.p, Glyma.10G156200.1.p, Glyma.10G156200.2.p, Glyma.10G163200.1.p, Glyma.10G164700.1.p, Glyma.10G177300.1.p, Glyma.10G195700.1.p, Glyma.10G195700.2.p, Glyma.10G218800.1.p, Glyma.10G218800.2.p, Glyma.10G218800.3.p, Glyma.10G220700.1.p, Glyma.10G220700.2.p, Glyma.10G222600.1.p, Glyma.10G227000.1.p, Glyma.10G231500.1.p, Glyma.10G237900.1.p, Glyma.10G241500.1.p, Glyma.10G242300.1.p, Glyma.10G242300.2.p, Glyma.10G252700.1.p, Glyma.10G252700.3.p, Glyma.10G252800.1.p, Glyma.10G252800.2.p, Glyma.10G252800.3.p, Glyma.10G253000.1.p, Glyma.10G253100.1.p, Glyma.10G253200.1.p, Glyma.10G253300.1.p, Glyma.10G253300.2.p, Glyma.10G253800.1.p, Glyma.10G253900.1.p, Glyma.10G254300.1.p, Glyma.10G261600.1.p, Glyma.10G270600.1.p, Glyma.10G271200.2.p, Glyma.10G271200.3.p, Glyma.10G271200.4.p, Glyma.10G271200.5.p, Glyma.10G271200.6.p, Glyma.10G271200.7.p, Glyma.10G271400.1.p, Glyma.10G271600.1.p, Glyma.10G271900.1.p, Glyma.11G002800.1.p, Glyma.11G019100.1.p, Glyma.11G028300.1.p, Glyma.11G030000.1.p, Glyma.11G036700.1.p, Glyma.11G044100.1.p, Glyma.11G054800.1.p, Glyma.11G063100.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Proteins IDs |
|---|---|
| GM | Glyma.11G063200.1.p, Glyma.11G067200.1.p, Glyma.11G067200.2.p, Glyma.11G075200.1.p, Glyma.11G087200.1.p, Glyma.11G089100.1.p, Glyma.11G105000.1.p, Glyma.11G105000.2.p, Glyma.11G114100.1.p, Glyma.11G118200.1.p, Glyma.11G144800.1.p, Glyma.11G144800.10.p, Glyma.11G144800.11.p, Glyma.11G144800.12.p, Glyma.11G144800.13.p, Glyma.11G144800.14.p, Glyma.11G144800.2.p, Glyma.11G144800.3.p, Glyma.11G144800.4.p, Glyma.11G144800.5.p, Glyma.11G144800.6.p, Glyma.11G144800.7.p, Glyma.11G144800.8.p, Glyma.11G144800.9.p, Glyma.11G156900.1.p, Glyma.11G156900.2.p, Glyma.11G156900.3.p, Glyma.11G167100.1.p, Glyma.11G174800.1.p, Glyma.11G179200.1.p, Glyma.11G196800.1.p, Glyma.11G199700.1.p, Glyma.11G199700.2.p, Glyma.11G199700.3.p, Glyma.11G199700.4.p, Glyma.11G200300.1.p, Glyma.11G200300.2.p, Glyma.11G204500.1.p, Glyma.11G204700.1.p, Glyma.11G204800.1.p, Glyma.11G204800.2.p, Glyma.11G204800.3.p, Glyma.11G204800.4.p, Glyma.11G204800.5.p, Glyma.11G204800.6.p, Glyma.11G204800.7.p, Glyma.11G205200.1.p, Glyma.11G205300.1.p, Glyma.11G205400.1.p, Glyma.11G206100.1.p, Glyma.11G206100.10.p, Glyma.11G206100.11.p, Glyma.11G206100.2.p, Glyma.11G206100.3.p, Glyma.11G206100.4.p, Glyma.11G206100.5.p, Glyma.11G206100.6.p, Glyma.11G206100.7.p, Glyma.11G206100.8.p, Glyma.11G206100.9.p, Glyma.11G206300.1.p, Glyma.11G206400.1.p, Glyma.11G206400.2.p, Glyma.11G206700.1.p, Glyma.11G206700.3.p, Glyma.11G206700.5.p, Glyma.11G206700.6.p, Glyma.11G206900.1.pGlyma.11G206900.2.p, Glyma.11G206900.3.p, Glyma.11G206900.4.p, Glyma.11G207000.1.p, Glyma.11G207000.2.p, Glyma.11G207300.1.p, Glyma.11G207500.1.p, Glyma.11G211100.1.p, Glyma.11G211100.3.p, Glyma.11G213100.1.p, Glyma.11G214400.1.p, Glyma.11G220900.1.p, Glyma.11G223200.1.p, Glyma.11G230600.1.p, Glyma.11G232200.1.p, Glyma.11G233500.1.p, Glyma.11G240900.1.p, Glyma.11G240900.2.p, Glyma.11G240900.3.p, Glyma.11G240900.4.p, Glyma.11G246200.1.p, Glyma.11G246200.2.p, Glyma.11G246200.3.p, Glyma.11G246200.4.p, Glyma.11G246200.5.p, Glyma.11G246200.6.p, Glyma.11G246200.7.p, Glyma.11G246200.8.p, Glyma.11G246200.9.p, Glyma.11G254000.1.p, Glyma.12G002400.1.p, Glyma.12G002500.1.p, Glyma.12G002500.2.p, Glyma.12G006300.1.p, Glyma.12G007000.1.p, Glyma.12G007300.1.p, Glyma.12G030000.1.p, Glyma.12G030000.2.p, Glyma.12G030000.3.p, Glyma.12G040000.1.p, Glyma.12G043600.1.p, Glyma.12G052300.1.p, Glyma.12G074600.1.p, Glyma.12G074600.10.p, Glyma.12G074600.2.p, Glyma.12G074600.3.p, Glyma.12G074600.4.p, Glyma.12G074600.5.p, Glyma.12G074600.6.p, Glyma.12G074600.7.p, Glyma.12G074600.8.p, Glyma.12G074600.9.p, Glyma.12G077300.1.p, Glyma.12G077300.10.p, Glyma.12G077300.2.p, Glyma.12G077300.3.p, Glyma.12G077300.4.p, Glyma.12G077300.5.p, Glyma.12G077300.6.p, Glyma.12G077300.7.p, Glyma.12G077300.8.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.12G077300.9.p, Glyma.12G077500.1.p, Glyma.12G077500.2.p, Glyma.12G094200.1.p, Glyma.12G108400.1.p, Glyma.12G108400.2.p, Glyma.12G108400.3.p, Glyma.12G114100.1.p, Glyma.12G118600.1.p, Glyma.12G133500.1.p, Glyma.12G133500.2.p, Glyma.12G133500.3.p, Glyma.12G133500.4.p, Glyma.12G137100.1.p, Glyma.12G143300.1.p, Glyma.12G144200.1.p, Glyma.12G148200.1.p, Glyma.12G152400.1.p, Glyma.12G172700.1.p, Glyma.12G172700.2.p, Glyma.12G172700.3.p, Glyma.12G172700.4.p, Glyma.12G175300.1.p, Glyma.12G175300.10.p, Glyma.12G175300.11.p, Glyma.12G175300.12.p, Glyma.12G175300.2.p, Glyma.12G175300.3.p, Glyma.12G175300.4.p, Glyma.12G175300.5.p, Glyma.12G175300.6.p, Glyma.12G175300.7.p, Glyma.12G175300.8.p, Glyma.12G175300.9.p, Glyma.12G188100.1.p, Glyma.12G188100.2.p, Glyma.12G188100.3.p, Glyma.12G205100.1.p, Glyma.12G205100.2.p, Glyma.12G206800.1.p, Glyma.12G211800.1.p, Glyma.12G211800.2.p, Glyma.12G211800.3.p, Glyma.12G216000.1.p, Glyma.12G216000.2.p, Glyma.12G220400.1.p, Glyma.12G220400.2.p, Glyma.12G225800.1.p, Glyma.12G233100.1.p, Glyma.12G235900.1.p, Glyma.12G240300.1.p, Glyma.13G007200.1.p, Glyma.13G031800.1.p, Glyma.13G032000.1.p, Glyma.13G032100.1.p, Glyma.13G032600.1.p, Glyma.13G032600.2.p, Glyma.13G032700.1.p, Glyma.13G033000.1.p, Glyma.13G033000.3.p, Glyma.13G033000.4.p, Glyma.13G033000.5.p, Glyma.13G033200.1.p, Glyma.13G033400.1.p, Glyma.13G033400.2.p, Glyma.13G033500.1.p, Glyma.13G033500.2.p, Glyma.13G033800.1.p, Glyma.13G033900.1.p, Glyma.13G035700.1.p, Glyma.13G039700.1.p, Glyma.13G039900.1.p, Glyma.13G050200.1.p, Glyma.13G050200.2.p, Glyma.13G050200.3.p, Glyma.13G053600.1.p, Glyma.13G053600.2.p, Glyma.13G053700.1.p, Glyma.13G053800.1.p, Glyma.13G054200.1.p, Glyma.13G054300.1.p, Glyma.13G054400.1.p, Glyma.13G056200.1.p, Glyma.13G104300.1.p, Glyma.13G104300.10.p, Glyma.13G104300.11.p, Glyma.13G104300.12.p, Glyma.13G104300.13.p, Glyma.13G104300.14.p, Glyma.13G104300.15.p, Glyma.13G104300.16.p, Glyma.13G104300.17.p, Glyma.13G104300.18.p, Glyma.13G104300.2.p, Glyma.13G104300.3.p, Glyma.13G104300.4.p, Glyma.13G104300.5.p, Glyma.13G104300.6.p, Glyma.13G104300.7.p, Glyma.13G104300.8.p, Glyma.13G104300.9.p, Glyma.13G111800.1.p, Glyma.13G127100.1.p, Glyma.13G127100.2.p, Glyma.13G128200.1.p, Glyma.13G128200.2.p, Glyma.13G128200.3.p, Glyma.13G136700.1.p, Glyma.13G136700.2.p, Glyma.13G136700.3.p, Glyma.13G139800.1.p, Glyma.13G150000.1.p, Glyma.13G161700.1.p, Glyma.13G161700.2.p, Glyma.13G161700.3.p, Glyma.13G174800.1.p, Glyma.13G174900.1.p, Glyma.13G184200.1.p, Glyma.13G184200.2.p, Glyma.13G184200.3.p, Glyma.13G184200.4.p, Glyma.13G188900.1.p, Glyma.13G189000.1.p, Glyma.13G201400.1.p, Glyma.13G224300.22.p, Glyma.13G228000.1.p, Glyma.13G228000.2.p, Glyma.13G228000.3.p, Glyma.13G228000.4.p, Glyma.13G228000.5.p, Glyma.13G228300.1.p, Glyma.13G228700.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.13G228700.2.p, Glyma.13G228700.3.p, Glyma.13G228700.4.p, Glyma.13G228700.5.p, Glyma.13G228700.6.p, Glyma.13G234800.1.p, Glyma.13G235000.1.p, Glyma.13G238700.1.p, Glyma.13G244200.1.p, Glyma.13G247300.1.p, Glyma.13G247300.2.p, Glyma.13G247300.3.p, Glyma.13G247300.4.p, Glyma.13G251900.1.p, Glyma.13G253300.1.p, Glyma.13G255500.1.p, Glyma.13G265700.1.p, Glyma.13G265700.2.p, Glyma.13G265700.3.p, Glyma.13G265700.4.p, Glyma.13G265700.5.p, Glyma.13G265700.6.p, Glyma.13G266000.1.p, Glyma.13G268100.1.p, Glyma.13G268100.2.p, Glyma.13G275100.1.p, Glyma.13G285500.1.p, Glyma.13G285500.2.p, Glyma.13G285500.3.p, Glyma.13G289900.1.p, Glyma.13G289900.2.p, Glyma.13G289900.3.p, Glyma.13G294100.1.p, Glyma.13G296200.1.p, Glyma.13G296300.1.p, Glyma.13G300000.1.p, Glyma.13G300000.2.p, Glyma.13G300000.3.p, Glyma.13G300000.4.p, Glyma.13G300000.5.p, Glyma.13G300000.6.p, Glyma.13G313100.1.p, Glyma.13G326100.1.p, Glyma.13G326100.2.p, Glyma.13G326100.3.p, Glyma.13G326100.4.p, Glyma.13G344400.1.p, Glyma.13G350000.1.p, Glyma.13G352700.2.p, Glyma.13G352800.1.p, Glyma.13G352900.5.p, Glyma.13G352900.6.p, Glyma.13G353000.1.p, Glyma.13G354400.1.p, Glyma.13G354400.2.p, Glyma.13G358900.1.p, Glyma.13G358900.2.p, Glyma.13G358900.3.p, Glyma.13G358900.4.p, Glyma.13G358900.5.p, Glyma.13G365400.1.p, Glyma.13G365400.2.p, Glyma.13G369100.1.p, Glyma.13G370700.1.p, Glyma.13G371200.1.p, Glyma.13G371200.2.p, Glyma.14G004500.1.p, Glyma.14G011500.1.p, Glyma.14G013300.1.p, Glyma.14G016500.1.p, Glyma.14G026500.1.p, Glyma.14G029000.1.p, Glyma.14G029000.2.p, Glyma.14G029000.3.p, Glyma.14G029000.4.p, Glyma.14G029000.5.p, Glyma.14G033500.1.p, Glyma.14G040200.1.p, Glyma.14G041600.1.p, Glyma.14G041600.2.p, Glyma.14G041600.3.p, Glyma.14G046200.1.p, Glyma.14G047900.1.p, Glyma.14G048100.1.p, Glyma.14G048300.1.p, Glyma.14G055900.1.p, Glyma.14G057400.1.p, Glyma.14G058900.2.p, Glyma.14G059100.1.p, Glyma.14G060300.1.p, Glyma.14G060400.1.p, Glyma.14G074700.1.p, Glyma.14G077700.1.p, Glyma.14G100300.1.p, Glyma.14G100400.1.p, Glyma.14G100600.1.p, Glyma.14G100700.1.p, Glyma.14G100800.1.p, Glyma.14G111900.1.p, Glyma.14G111900.2.p, Glyma.14G111900.3.p, Glyma.14G111900.4.p, Glyma.14G111900.5.p, Glyma.14G111900.6.p, Glyma.14G111900.7.p, Glyma.14G111900.8.p, Glyma.14G113200.1.p, Glyma.14G116000.1.p, Glyma.14G116500.1.p, Glyma.14G118700.1.p, Glyma.14G118800.1.p, Glyma.14G124400.1.p, Glyma.14G124700.1.p, Glyma.14G137700.1.p, Glyma.14G144300.1.p, Glyma.14G157800.1.p, Glyma.14G157900.1.p, Glyma.14G187900.1.p, Glyma.14G187900.2.p, Glyma.14G187900.3.p, Glyma.14G189100.1.p, Glyma.14G195300.1.p, Glyma.14G197000.1.p, Glyma.14G197000.2.p, Glyma.14G206000.1.p, Glyma.14G206000.2.p, Glyma.14G206200.1.p, Glyma.14G206200.2.p, Glyma.14G206200.3.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.14G206200.4.p, Glyma.14G206300.1.p, Glyma.14G206300.2.p, Glyma.14G211200.1.p, Glyma.14G212100.1.p, Glyma.14G214700.1.p, Glyma.15G000600.1.p, Glyma.15G001500.1.p, Glyma.15G002900.1.p, Glyma.15G004700.1.p, Glyma.15G004700.2.p, Glyma.15G007800.1.p, Glyma.15G007800.2.p, Glyma.15G015100.1.p, Glyma.15G015100.2.p, Glyma.15G015100.3.p, Glyma.15G019800.1.p, Glyma.15G019800.2.p, Glyma.15G019800.3.p, Glyma.15G019800.4.p, Glyma.15G021400.1.p, Glyma.15G024300.1.p, Glyma.15G024300.2.p, Glyma.15G030000.1.p, Glyma.15G042900.1.p, Glyma.15G042900.2.p, Glyma.15G045600.1.p, Glyma.15G051600.1.p, Glyma.15G052900.1.p, Glyma.15G059200.1.p, Glyma.15G059300.1.p, Glyma.15G066700.1.p, Glyma.15G066700.2.p, Glyma.15G066700.3.p, Glyma.15G066700.4.p, Glyma.15G066700.5.p, Glyma.15G066700.6.p, Glyma.15G069500.1.p, Glyma.15G069500.2.p, Glyma.15G074600.1.p, Glyma.15G078200.1.p, Glyma.15G083800.1.p, Glyma.15G084200.1.p, Glyma.15G111300.1.p, Glyma.15G111600.1.p, Glyma.15G124900.1.p, Glyma.15G130900.1.p, Glyma.15G155600.1.p, Glyma.15G161300.1.p, Glyma.15G161500.1.p, Glyma.15G161600.1.p, Glyma.15G161700.1.p, Glyma.15G162000.1.p, Glyma.15G169000.1.p, Glyma.15G169000.2.p, Glyma.15G170000.1.p, Glyma.15G170000.10.p, Glyma.15G170000.11.p, Glyma.15G170000.12.p, Glyma.15G170000.13.p, Glyma.15G170000.14.p, Glyma.15G170000.15.p, Glyma.15G170000.2.p, Glyma.15G170000.3.p, Glyma.15G170000.4.p, Glyma.15G170000.5.p, Glyma.15G170000.6.p, Glyma.15G170000.7.p, Glyma.15G170000.8.p, Glyma.15G170000.9.p, Glyma.15G179300.1.p, Glyma.15G181400.1.p, Glyma.15G181400.2.p, Glyma.15G189200.1.p, Glyma.15G202700.1.p, Glyma.15G203600.1.p, Glyma.15G206400.1.p, Glyma.15G210200.1.p, Glyma.15G212300.1.p, Glyma.15G213600.2.p, Glyma.15G215600.1.p, Glyma.15G218200.1.p, Glyma.15G226500.1.p, Glyma.15G226500.2.p, Glyma.15G226800.1.p, Glyma.15G227000.1.p, Glyma.15G227000.2.p, Glyma.15G237200.1.p, Glyma.15G242800.1.p, Glyma.15G242800.2.p, Glyma.15G242800.3.p, Glyma.15G252600.1.p, Glyma.15G252600.2.p, Glyma.15G266400.1.p, Glyma.16G009900.1.p, Glyma.16G015000.1.p, Glyma.16G015400.1.p, Glyma.16G015400.2.p, Glyma.16G015400.3.p, Glyma.16G032700.1.p, Glyma.16G037500.1.p, Glyma.16G037500.2.p, Glyma.16G037500.3.p, Glyma.16G037500.4.p, Glyma.16G037500.5.p, Glyma.16G047100.1.p, Glyma.16G047100.2.p, Glyma.16G047100.3.p, Glyma.16G047100.4.p, Glyma.16G047100.5.p, Glyma.16G047300.1.p, Glyma.16G059200.1.p, Glyma.16G059700.1.p, Glyma.16G059700.2.p, Glyma.16G059700.3.p, Glyma.16G059700.4.p, Glyma.16G059700.5.p, Glyma.16G064100.1.p, Glyma.16G064200.1.p, Glyma.16G064400.1.p, Glyma.16G064800.1.p, Glyma.16G064900.1.p, Glyma.16G065000.1.p, Glyma.16G065200.1.p, Glyma.16G065400.1.p, Glyma.16G065500.1.p, Glyma.16G065600.1.p, Glyma.16G065700.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.16G078800.1.p, Glyma.16G078900.1.p, Glyma.16G079000.1.p, Glyma.16G079200.1.p, Glyma.16G096900.1.p, Glyma.16G096900.2.p, Glyma.16G096900.3.p, Glyma.16G101800.1.p, Glyma.16G101800.2.p, Glyma.16G111800.1.p, Glyma.16G119800.1.p, Glyma.16G122100.1.p, Glyma.16G130400.1.p, Glyma.16G140700.1.p, Glyma.16G144000.1.p, Glyma.16G144000.2.p, Glyma.16G144000.3.p, Glyma.16G144000.4.p, Glyma.16G156100.1.p, Glyma.16G156100.2.p, Glyma.16G156200.1.p, Glyma.16G179600.1.p, Glyma.16G201500.1.p, Glyma.16G202200.1.p, Glyma.16G202200.4.p, Glyma.16G202400.1.p, Glyma.16G203300.1.p, Glyma.16G203300.2.p, Glyma.16G203300.3.p, Glyma.16G205100.1.p, Glyma.16G210100.1.p, Glyma.16G210500.1.p, Glyma.17G003100.1.p, Glyma.17G040000.1.p, Glyma.17G047900.1.p, Glyma.17G047900.2.p, Glyma.17G055900.1.p, Glyma.17G055900.2.p, Glyma.17G070300.1.p, Glyma.17G071700.1.p, Glyma.17G085000.1.p, Glyma.17G087000.1.p, Glyma.17G088000.1.p, Glyma.17G096500.1.p, Glyma.17G096500.2.p, Glyma.17G096500.3.p, Glyma.17G102600.1.p, Glyma.17G103400.1.p, Glyma.17G109600.1.p, Glyma.17G109600.2.p, Glyma.17G109600.3.p, Glyma.17G109600.4.p, Glyma.17G109600.5.p, Glyma.17G117800.1.p, Glyma.17G119800.1.p, Glyma.17G135000.1.p, Glyma.17G150600.1.p, Glyma.17G156300.1.p, Glyma.17G166200.1.p, Glyma.17G167600.1.p, Glyma.17G169300.1.p, Glyma.17G187200.1.p, Glyma.17G187200.2.p, Glyma.17G187200.3.p, Glyma.17G196200.1.p, Glyma.17G211200.1.p, Glyma.17G211700.1.p, Glyma.17G214100.1.p, Glyma.17G214100.3.p, Glyma.17G214100.4.p, Glyma.17G214200.1.p, Glyma.17G214400.1.p, Glyma.17G214400.2.p, Glyma.17G214400.3.p, Glyma.17G214400.4.p, Glyma.17G214500.1.p, Glyma.17G214600.1.p, Glyma.17G214600.2.p, Glyma.17G214600.3.p, Glyma.17G214600.4.p, Glyma.17G214600.5.p, Glyma.17G214600.6.p, Glyma.17G214600.7.p, Glyma.17G214600.8.p, Glyma.17G214600.9.p, Glyma.17G214700.1.p, Glyma.17G214800.1.p, Glyma.17G214900.1.p, Glyma.17G214900.2.p, Glyma.17G214900.3.p, Glyma.17G215900.1.p, Glyma.17G215900.2.p, Glyma.17G218500.1.p, Glyma.17G224200.1.p, Glyma.17G224300.1.p, Glyma.17G224300.2.p, Glyma.17G224500.1.p, Glyma.17G224600.1.p, Glyma.17G247800.1.p, Glyma.17G250800.1.p, Glyma.17G250800.2.p, Glyma.18G003000.1.p, Glyma.18G003000.2.p, Glyma.18G011000.1.p, Glyma.18G016400.1.p, Glyma.18G023500.1.p, Glyma.18G025100.1.p, Glyma.18G026700.1.p, Glyma.18G042200.1.p, Glyma.18G043500.1.p, Glyma.18G046100.1.p, Glyma.18G046100.2.p, Glyma.18G046200.1.p, Glyma.18G046300.1.p, Glyma.18G046400.1.p, Glyma.18G046600.1.p, Glyma.18G050700.1.p, Glyma.18G050700.2.p, Glyma.18G050700.3.p, Glyma.18G051100.1.p, Glyma.18G051100.2.p, Glyma.18G051100.3.p, Glyma.18G051100.4.p, Glyma.18G051100.5.p, Glyma.18G051100.6.p, Glyma.18G059600.1.p, Glyma.18G062500.1.p, Glyma.18G073600.1.p, Glyma.18G073600.2.p, Glyma.18G076200.1.p, Glyma.18G107600.1.p, Glyma.18G107600.2.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.18G107600.3.p, Glyma.18G110400.1.p, Glyma.18G116800.1.p, Glyma.18G124500.1.p, Glyma.18G134100.1.p, Glyma.18G137600.1.p, Glyma.18G137600.2.p, Glyma.18G137600.3.p, Glyma.18G137600.4.p, Glyma.18G138700.1.p, Glyma.18G141500.1.p, Glyma.18G141700.1.p, Glyma.18G141700.2.p, Glyma.18G142100.1.p, Glyma.18G150300.1.p, Glyma.18G175700.1.p, Glyma.18G185400.1.p, Glyma.18G185500.1.p Glyma.18G198000.1.p, Glyma.18G198000.2.p, Glyma.18G198200.1.p, Glyma.18G198800.1.p, Glyma.18G199000.1.p, Glyma.18G199000.2.p, Glyma.18G199200.1.p, Glyma.18G205000.1.p, Glyma.18G206400.1.p, Glyma.18G206400.2.p, Glyma.18G213600.1.p, Glyma.18G215800.1.p, Glyma.18G216200.1.p, Glyma.18G216200.2.p, Glyma.18G216800.1.p, Glyma.18G216800.2.p, Glyma.18G217000.1.p, Glyma.18G217000.2.p, Glyma.18G217000.3.p, Glyma.18G217000.4.p, Glyma.18G219200.1.p, Glyma.18G219300.1.p, Glyma.18G219300.2.p, Glyma.18G219300.3.p, Glyma.18G219600.1.p, Glyma.18G219600.3.p, Glyma.18G219700.1.p, Glyma.18G219800.1.p, Glyma.18G237100.1.p, Glyma.18G237900.1.p, Glyma.18G237900.2.p, Glyma.18G239600.1.p, Glyma.18G239700.1.p, Glyma.18G246400.1.p, Glyma.18G249700.1.p, Glyma.18G249900.1.p, Glyma.18G250500.1.p, Glyma.18G250900.1.p, Glyma.18G254000.1.p, Glyma.18G254300.1.p, Glyma.18G254400.1.p, Glyma.18G254500.1.p, Glyma.18G254600.1.p, Glyma.18G254700.1.p, Glyma.18G257800.1.p, Glyma.18G258200.1.p, Glyma.18G267000.1.p, Glyma.18G268000.1.p, Glyma.18G269900.1.p, Glyma.18G270100.1.p, Glyma.18G270600.1.p, Glyma.18G270700.1.p, Glyma.18G270800.1.p, Glyma.18G270900.1.p, Glyma.18G271000.1.p, Glyma.18G271100.1.p, Glyma.18G271200.1.p, Glyma.18G275700.1.p, Glyma.18G278200.1.p, Glyma.18G278200.2.p, Glyma.18G278200.3.p, Glyma.18G280200.1.p, Glyma.18G280200.2.p, Glyma.18G282100.1.p, Glyma.18G284100.1.p, Glyma.18G294400.1.p, Glyma.18G294400.2.p, Glyma.18G294400.3.p, Glyma.18G294800.1.p, Glyma.19G005700.1.p, Glyma.19G010700.1.p, Glyma.19G024100.1.p, Glyma.19G030400.1.p, Glyma.19G033100.1.p, Glyma.19G036600.1.p, Glyma.19G036600.2.p, Glyma.19G039100.1.p, Glyma.19G039100.2.p, Glyma.19G040400.1.p, Glyma.19G060800.1.p, Glyma.19G062300.1.p, Glyma.19G064600.1.p, Glyma.19G069300.1.p, Glyma.19G069300.2.p, Glyma.19G069400.1.p, Glyma.19G070600.1.p, Glyma.19G073200.1.p, Glyma.19G073200.3.p, Glyma.19G074900.1.p, Glyma.19G074900.2.p, Glyma.19G079500.1.p, Glyma.19G081200.1.p, Glyma.19G086900.1.p, Glyma.19G087800.1.p, Glyma.19G088300.1.p, Glyma.19G092000.1.p, Glyma.19G104600.1.p, Glyma.19G104600.2.p, Glyma.19G104600.3.p, Glyma.19G115800.1.p, Glyma.19G115800.2.p, Glyma.19G115800.3.p, Glyma.19G115800.4.p, Glyma.19G115800.5.p, Glyma.19G115800.6.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.19G140100.1.p, Glyma.19G140100.2.p, Glyma.19G143300.1.p, Glyma.19G152200.1.p, Glyma.19G152300.1.p, Glyma.19G166100.3.p, Glyma.19G167300.1.p, Glyma.19G167400.1.p, Glyma.19G168600.1.p, Glyma.19G170500.1.p, Glyma.19G188900.1.p, Glyma.19G190200.1.p, Glyma.19G217300.1.p, Glyma.19G217300.2.p, Glyma.19G261000.1.p, Glyma.19G261700.1.p, Glyma.19G261700.2.p, Glyma.19G261700.3.p, Glyma.20G035400.1.p, Glyma.20G035400.2.p, Glyma.20G054500.1.p, Glyma.20G054500.2.p, Glyma.20G072500.1.p, Glyma.20G078100.1.p, Glyma.20G078100.2.p, Glyma.20G093200.1.p, Glyma.20G118200.1.p, Glyma.20G118300.1.p, Glyma.20G118500.1.p, Glyma.20G118600.1.p, Glyma.20G118700.1.p, Glyma.20G118900.1.p, Glyma.20G118900.2.p, Glyma.20G118900.3.p, Glyma.20G119000.1.p, Glyma.20G119300.1.p, Glyma.20G119400.1.p, Glyma.20G119600.1.p, Glyma.20G119700.1.p, Glyma.20G120100.1.p, Glyma.20G120200.1.p, Glyma.20G120200.2.p, Glyma.20G120200.3.p, Glyma.20G120200.4.p, Glyma.20G129200.1.p, Glyma.20G137300.1.p, Glyma.20G137300.2.p, Glyma.20G137300.3.p, Glyma.20G137400.1.p, Glyma.20G137500.1.p, Glyma.20G137700.1.p, Glyma.20G137900.1.p, Glyma.20G137900.2.p, Glyma.20G138100.1.p, Glyma.20G138400.1.p, Glyma.20G138400.2.p, Glyma.20G138500.1.p, Glyma.20G138600.1.p, Glyma.20G138600.2.p, Glyma.20G138700.1.p, Glyma.20G138800.1.p, Glyma.20G138900.1.p, Glyma.20G138900.2.p, Glyma.20G139000.1.p, Glyma.20G139000.4.p, Glyma.20G139000.6.p, Glyma.20G139100.1.p, Glyma.20G139200.1.p, Glyma.20G139200.2.p, Glyma.20G139300.1.p, Glyma.20G139400.1.p, Glyma.20G139500.1.p, Glyma.20G139500.2.p, Glyma.20G139500.3.p, Glyma.20G139600.1.p, Glyma.20G139600.2.p, Glyma.20G139600.3.p, Glyma.20G139600.4.p, Glyma.20G139600.5.p, Glyma.20G139600.6.p, Glyma.20G139600.7.p, Glyma.20G139600.8.p, Glyma.20G139700.1.p, Glyma.20G139900.1.p, Glyma.20G139900.2.p, Glyma.20G139900.3.p, Glyma.20G139900.4.p, Glyma.20G140100.1.p, Glyma.20G140200.1.p, Glyma.20G140400.1.p, Glyma.20G140500.1.p, Glyma.20G140600.1.p, Glyma.20G140700.1.p, Glyma.20G140700.2.p, Glyma.20G151800.1.p, Glyma.20G151800.4.p, Glyma.20G151800.5.p, Glyma.20G152800.1.p, Glyma.20G156700.1.p, Glyma.20G162300.1.p, Glyma.20G169000.1.p, Glyma.20G170900.1.p, Glyma.20G173000.1.p, Glyma.20G173000.2.p, Glyma.20G194400.1.p, Glyma.20G212800.1.p, Glyma.20G225800.1.p, Glyma.20G226900.1.p, Glyma.20G232100.1.p, Glyma.20G246600.1.p, Glyma.U022900.1.p, Glyma.U025000.1.p, Glyma.U025000.2.p, Glyma.U027000.1.p, Glyma.U027100.1.p, Glyma.U032400.1.p, Glyma.U033500.1.p, Glyma.U033500.2.p, Glyma.U033500.3.p, Glyma.U033500.4.p |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| MT | Medtr0007s0020.1, Medtr0007s0390.1, Medtr0015s0030.1, Medtr0015s0090.1, Medtr0049s0070.1, Medtr0049s0070.2, Medtr0049s0070.3, Medtr0070s0020.1, Medtr0074s0060.1, Medtr0194s0030.1, Medtr0289s0020.1, Medtr0289s0030.1, Medtr0341s0020.1, Medtr0365s0020.1, Medtr0365s0030.1, Medtr0400s0030.1, Medtr0400s0040.1, Medtr0491s0010.1, Medtr0491s0030.1, Medtr0830s0010.1, Medtr1g008980.1, Medtr1g009270.1, Medtr1g012550.1, Medtr1g012610.1, Medtr1g012610.2, Medtr1g012610.3, Medtr1g013040.1, Medtr1g013040.2, Medtr1g014240.1, Medtr1g016300.2, Medtr1g021610.1, Medtr1g021630.1, Medtr1g021632.1, Medtr1g021635.1, Medtr1g021642.1, Medtr1g021845.1, Medtr1g022265.1, Medtr1g022265.2, Medtr1g026940.1, Medtr1g027040.1, Medtr1g027150.1, Medtr1g027200.1, Medtr1g027370.1, Medtr1g027410.1, Medtr1g027420.1, Medtr1g027440.1, Medtr1g027500.1, Medtr1g027540.1, Medtr1g027680.1, Medtr1g027690.1, Medtr1g027740.1, Medtr1g027820.1, Medtr1g027890.1, Medtr1g027960.1, Medtr1g027970.1, Medtr1g027990.1, Medtr1g028020.1, Medtr1g028080.1, Medtr1g028100.1, Medtr1g028130.1, Medtr1g028170.1, Medtr1g028220.1, Medtr1g028280.1, Medtr1g028280.2, Medtr1g028290.1, Medtr1g028890.1, Medtr1g028890.2, Medtr1g029610.1, Medtr1g029930.1, Medtr1g029940.1, Medtr1g029950.1, Medtr1g031200.1, Medtr1g031280.1, Medtr1g031450.1, Medtr1g031510.1, Medtr1g031520.1, Medtr1g031540.1, Medtr1g031540.2, Medtr1g031560.1, Medtr1g031580.1, Medtr1g032930.1, Medtr1g033000.1, Medtr1g033010.1, Medtr1g033040.1, Medtr1g038890.1, Medtr1g039220.1, Medtr1g039240.1, Medtr1g039310.1, Medtr1g040073.1, Medtr1g040077.1, Medtr1g040200.1, Medtr1g040525.1, Medtr1g040545.1, Medtr1g040555.1, Medtr1g040615.1, Medtr1g047670.1, Medtr1g048360.1, Medtr1g052275.1, Medtr1g052425.1, Medtr1g052530.2, Medtr1g053525.1, Medtr1g061590.1, Medtr1g064560.1, Medtr1g069340.1, Medtr1g069470.1, Medtr1g079520.1, Medtr1g080440.1, Medtr1g082580.1, Medtr1g084790.1, Medtr1g086870.1, Medtr1g088930.1, Medtr1g088935.1, Medtr1g088940.1, Medtr1g088940.2, Medtr1g089600.1, Medtr1g090520.1, Medtr1g096260.1, Medtr1g096270.1, Medtr1g097160.1, Medtr1g097580.1, Medtr1g098360.1, Medtr1g099260.1, Medtr1g100110.1, Medtr1g100787.1, Medtr1g102500.1, Medtr1g104890.1, Medtr1g104890.2, Medtr1g104890.3, Medtr1g104890.4, Medtr1g105585.1, Medtr1g105600.1, Medtr1g105615.1, Medtr1g105615.2, Medtr1g105630.1, Medtr1g105640.1, Medtr1g105640.2, Medtr1g105640.3, Medtr1g105650.1, Medtr1g105655.1, Medtr1g105700.1, Medtr1g105710.1, Medtr1g105725.1, Medtr1g105750.1, Medtr1g105755.1, Medtr1g105800.1, Medtr1g105820.1, Medtr1g105840.1, Medtr1g107460.1, Medtr1g109580.1, Medtr1g110110.1, Medtr1g110120.1, Medtr1g110180.1, Medtr1g110230.1, Medtr1g110260.1, Medtr1g110280.1, Medtr1g115485.1, Medtr2g005810.1, Medtr2g006910.1, Medtr2g006910.2, Medtr2g008360.1, Medtr2g008370.1, Medtr2g008390.1, Medtr2g008400.1, Medtr2g008740.1, Medtr2g010470.1, Medtr2g011640.1, Medtr2g012670.1, Medtr2g013720.1, Medtr2g014560.1, Medtr2g014650.1, Medtr2g014960.1, Medtr2g016320.1, Medtr2g016360.1, Medtr2g016500.1, Medtr2g016530.1, Medtr2g016580.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| MT | Medtr2g024290.1, Medtr2g024290.2, Medtr2g024330.1, Medtr2g028580.1, Medtr2g028580.2, Medtr2g029010.1, Medtr2g030310.1, Medtr2g030310.2, Medtr2g030380.1, Medtr2g030380.2, Medtr2g030380.3, Medtr2g031520.1, Medtr2g031530.1, Medtr2g036430.1, Medtr2g036440.1, Medtr2g036460.1, Medtr2g036490.1, Medtr2g036500.1, Medtr2g038675.1, Medtr2g039290.1, Medtr2g039290.2, Medtr2g039290.3, Medtr2g040910.1, Medtr2g042710.1, Medtr2g043380.1, Medtr2g046130.1, Medtr2g054870.1, Medtr2g068650.1, Medtr2g070020.1, Medtr2g072620.1, Medtr2g072640.1, Medtr2g073250.1, Medtr2g073650.1, Medtr2g073650.2, Medtr2g074820.1, Medtr2g074870.1, Medtr2g074980.1, Medtr2g074990.1, Medtr2g075060.1,Medtr2g078810.1, Medtr2g078810.2, Medtr2g080080.1, Medtr2g080090.1, Medtr2g080100.1, Medtr2g080220.1, Medtr2g082430.1, Medtr2g082430.2, Medtr2g082430.3, Medtr2g084120.1, Medtr2g087230.1, Medtr2g090120.1, Medtr2g090120.2, Medtr2g090120.3, Medtr2g090120.4, Medtr2g090120.5, Medtr2g090410.1, Medtr2g090710.1, Medtr2g096160.1, Medtr2g098910.1, Medtr2g100470.1, Medtr2g100550.1, Medtr2g103810.1, Medtr2g104790.1, Medtr2g105260.1, Medtr2g105680.1, Medtr2g105900.1, Medtr2g437730.1, Medtr2g449790.1, Medtr3g007870.1, Medtr3g008040.1, Medtr3g009400.1, Medtr3g011910.1, Medtr3g011910.2, Medtr3g011930.1, Medtr3g028650.1, Medtr3g031600.1, Medtr3g047890.1, Medtr3g060880.1, Medtr3g062500.1, Medtr3g062500.2, Medtr3g062570.1, Medtr3g062570.2, Medtr3g062570.3, Medtr3g062590.1, Medtr3g062590.2, Medtr3g064080.1, Medtr3g064090.1, Medtr3g064110.1, Medtr3g068025.1, Medtr3g069050.1, Medtr3g070220.1, Medtr3g070570.1, Medtr3g070800.1, Medtr3g071480.1, Medtr3g072800.1, Medtr3g075440.1, Medtr3g078250.1, Medtr3g079850.1, Medtr3g079850.2, Medtr3g080050.1, Medtr3g080170.1, Medtr3g082750.1, Medtr3g082750.2, Medtr3g082750.3, Medtr3g082920.1, Medtr3g084510.1, Medtr3g086120.1, Medtr3g086120.2, Medtr3g087060.1, Medtr3g087060.2, Medtr3g087060.3, Medtr3g088930.1, Medtr3g088930.2, Medtr3g090480.1, Medtr3g090660.1, Medtr3g090660.2, Medtr3g092360.1, Medtr3g092390.1, Medtr3g092420.1, Medtr3g093710.1, Medtr3g093710.2, Medtr3g093710.3, Medtr3g093710.4, Medtr3g093930.1, Medtr3g094710.1, Medtr3g094740.1, Medtr3g094740.3, Medtr3g095100.1, Medtr3g096590.1, Medtr3g102180.1, Medtr3g106320.1, Medtr3g107070.1, Medtr3g109820.1, Medtr3g110450.1, Medtr3g110840.1, Medtr3g110860.1, Medtr3g110860.2, Medtr3g111510.1, Medtr3g113140.1, Medtr3g116450.1, Medtr3g116590.1, Medtr3g116590.2, Medtr3g116640.1, Medtr3g116640.2, Medtr3g415610.1, Medtr3g437630.1, Medtr3g449390.1, Medtr3g449540.1, Medtr3g460810.1, Medtr3g464080.1, Medtr3g465470.1, Medtr4g005130.1, Medtr4g014070.1, Medtr4g014350.1, Medtr4g014900.1, Medtr4g014900.2, Medtr4g014900.3, Medtr4g028090.1, Medtr4g029710.1, Medtr4g035180.1, Medtr4g036505.1, Medtr4g036575.1, Medtr4g037015.1, Medtr4g037720.1, Medtr4g044393.1, Medtr4g046113.1, Medtr4g046113.2, Medtr4g046113.3, Medtr4g051605.1, Medtr4g058570.1, Medtr4g058710.1, Medtr4g058710.2, Medtr4g058710.3, Medtr4g058710.4, Medtr4g058710.5, Medtr4g061330.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| MT | Medtr4g061833.1, Medtr4g061833.2, Medtr4g061930.1, Medtr4g065087.1, Medtr4g067230.1, Medtr4g069970.1, Medtr4g069970.2, Medtr4g069970.3, Medtr4g070970.1, Medtr4g073220.1, Medtr4g073230.1, Medtr4g074080.1, Medtr4g085810.1, Medtr4g088320.1, Medtr4g091510.1, Medtr4g091510.2, Medtr4g091510.3, Medtr4g091510.4, Medtr4g092070.1, Medtr4g092520.1, Medtr4g092550.1, Medtr4g093040.1, Medtr4g093050.1, Medtr4g093070.1, Medtr4g093080.1, Medtr4g093110.1, Medtr4g093140.1, Medtr4g094415.2, Medtr4g094462.1, Medtr4g094610.1, Medtr4g094615.3, Medtr4g094790.1, Medtr4g094885.1, Medtr4g094958.1, Medtr4g095012.1, Medtr4g095018.1, Medtr4g095042.1, Medtr4g095045.1, Medtr4g103850.1, Medtr4g105070.1, Medtr4g105370.1, Medtr4g105540.1, Medtr4g107620.1, Medtr4g109010.1, Medtr4g109170.1, Medtr4g111925.1, Medtr4g113100.1, Medtr4g113710.1, Medtr4g114250.1, Medtr4g114270.1, Medtr4g114280.1, Medtr4g117020.1, Medtr4g117030.1, Medtr4g117040.1, Medtr4g117490.2, Medtr4g117800.1, Medtr4g123880.1, Medtr4g123880.2, Medtr4g124990.1, Medtr4g126270.1, Medtr4g126930.1, Medtr4g130210.1, Medtr4g130210.2, Medtr4g130920.1, Medtr5g005450.1, Medtr5g005480.1, Medtr5g005520.1, Medtr5g005520.2, Medtr5g005530.1, Medtr5g009660.1, Medtr5g010440.2, Medtr5g011410.1, Medtr5g011840.1, Medtr5g011840.2, Medtr5g011840.3, Medtr5g014720.1, Medtr5g017080.1, Medtr5g017080.2, Medtr5g018570.1, Medtr5g019040.1, Medtr5g019050.1, Medtr5g019070.1, Medtr5g019940.1, Medtr5g021670.1, Medtr5g024450.1, Medtr5g024510.1, Medtr5g024510.3, Medtr5g025020.1, Medtr5g025180.1, Medtr5g025840.1, Medtr5g025850.1, Medtr5g025860.1, Medtr5g025880.1, Medtr5g025890.1, Medtr5g025930.1, Medtr5g025950.1, Medtr5g026000.1, Medtr5g026010.1, Medtr5g026090.1, Medtr5g026150.1, Medtr5g026160.1, Medtr5g026200.1, Medtr5g026510.1, Medtr5g026510.2, Medtr5g026760.1, Medtr5g033490.1, Medtr5g033690.1, Medtr5g033820.1, Medtr5g035910.1, Medtr5g037410.1, Medtr5g038600.1, Medtr5g042440.1, Medtr5g044680.1Medtr5g044680.2, Medtr5g045910.1, Medtr5g047060.1, Medtr5g047070.1, Medtr5g047110.1, Medtr5g047120.1, Medtr5g053450.1, Medtr5g055310.1, Medtr5g055470.1, Medtr5g059910.1, Medtr5g065130.1, Medtr5g068190.1, Medtr5g068210.1, Medtr5g068250.1, Medtr5g068260.1, Medtr5g068770.1, Medtr5g075630.1, Medtr5g075650.1, Medtr5g075650.2, Medtr5g075650.3, Medtr5g077100.1, Medtr5g077430.1, Medtr5g078080.1, Medtr5g082270.1, Medtr5g082290.1, Medtr5g082420.1, Medtr5g082460.1, Medtr5g082630.1, Medtr5g082920.1, Medtr5g083480.1, Medtr5g083480.2, Medtr5g085700.1, Medtr5g085790.1, Medtr5g086030.1, Medtr5g086040.1, Medtr5g086040.2, Medtr5g086040.3, Medtr5g086080.1, Medtr5g086090.1, Medtr5g086120.1, Medtr5g086130.1, Medtr5g086130.2, Medtr5g086310.1, Medtr5g086540.1, Medtr5g087340.1, Medtr5g087350.1, Medtr5g087360.1, Medtr5g087360.2, Medtr5g087780.1, Medtr5g088350.2, Medtr5g088400.2, Medtr5g089890.1, Medtr5g090100.1, Medtr5g091380.1, Medtr5g091380.2, Medtr5g091380.3, Medtr5g091380.4, Medtr5g091380.5, Medtr5g091380.6, Medtr5g091380.7, Medtr5g091380.8, Medtr5g094380.1, Medtr5g095970.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| MT | Medtr5g096530.1, Medtr6g007690.1, Medtr6g009370.1, Medtr6g015190.1, Medtr6g015265.1, Medtr6g015265.2, Medtr6g015805.1, Medtr6g016040.1, Medtr6g016040.2, Medtr6g016495.1, Medtr6g016495.2, Medtr6g023390.1, Medtr6g027720.1, Medtr6g036780.1, Medtr6g036790.1, Medtr6g036840.1, Medtr6g036870.1, Medtr6g036890.1, Medtr6g040210.1, Medtr6g040230.1, Medtr6g044740.1, Medtr6g044830.1, Medtr6g045030.1, Medtr6g048090.1, Medtr6g052970.1, Medtr6g053000.1, Medtr6g055960.1, Medtr6g057750.1, Medtr6g057770.1, Medtr6g060230.1, Medtr6g068970.1, Medtr6g069030.1, Medtr6g082870.1, Medtr6g082930.1, Medtr6g082950.1, Medtr6g083760.1, Medtr6g083780.1, Medtr6g088610.1, Medtr6g088755.1, Medtr6g088785.1, Medtr6g088785.2, Medtr6g088790.1, Medtr6g093050.1, Medtr6g445600.1, Medtr6g463630.1, Medtr6g463700.1, Medtr6g463710.1, Medtr6g463830.1, Medtr6g470960.1, Medtr7g007550.1, Medtr7g007560.1, Medtr7g007620.1, Medtr7g007630.1, Medtr7g009940.1, Medtr7g009970.1, Medtr7g010000.1, Medtr7g010010.1, Medtr7g013680.1, Medtr7g015230.1, Medtr7g015240.1, Medtr7g015250.1, Medtr7g015280.1, Medtr7g015290.1, Medtr7g015310.1, Medtr7g015320.1, Medtr7g015390.1, Medtr7g015420.1, Medtr7g015510.1, Medtr7g015550.1, Medtr7g015620.1, Medtr7g015670.1, Medtr7g015800.1, Medtr7g018200.1, Medtr7g018200.2, Medtr7g018890.1, Medtr7g018890.2, Medtr7g018890.3, Medtr7g020850.1, Medtr7g020850.2, Medtr7g021570.1, Medtr7g021570.2, Medtr7g022160.1, Medtr7g028432.1, Medtr7g029650.1, Medtr7g030070.1, Medtr7g031470.1, Medtr7g033135.1, Medtr7g038690.1, Medtr7g038690.2, Medtr7g039330.1, Medtr7g045510.1, Medtr7g045710.1, Medtr7g050990.1, Medtr7g051040.1, Medtr7g056653.1, Medtr7g059225.1, Medtr7g059285.1, Medtr7g062660.1, Medtr7g062680.1, Medtr7g062700.1, Medtr7g062730.1, Medtr7g062750.1, Medtr7g062770.1, Medtr7g062890.1, Medtr7g062920.1, Medtr7g062940.1, Medtr7g062950.1, Medtr7g062990.1, Medtr7g063010.1, Medtr7g063030.1, Medtr7g063030.2, Medtr7g067530.1, Medtr7g070200.1, Medtr7g073290.1, Medtr7g073530.1, Medtr7g073660.1, Medtr7g073710.1, Medtr7g074010.1, Medtr7g074010.2, Medtr7g074010.3, Medtr7g078730.1, Medtr7g079350.1, Medtr7g080810.1, Medtr7g080810.2, Medtr7g081410.1, Medtr7g081410.2, Medtr7g081480.1, Medtr7g081570.1, Medtr7g081720.1, Medtr7g081780.1, Medtr7g081840.1, Medtr7g082110.1, Medtr7g082280.1, Medtr7g082300.1, Medtr7g082360.1, Medtr7g082430.1, Medtr7g082460.1, Medtr7g082470.1, Medtr7g082490.1, Medtr7g082510.1, Medtr7g082530.1, Medtr7g082530.2, Medtr7g083960.1, Medtr7g084220.1, Medtr7g084250.1, Medtr7g084250.2, Medtr7g086420.1, Medtr7g086420.2, Medtr7g086420.3, Medtr7g086420.4, Medtr7g091680.1, Medtr7g092880.1, Medtr7g092910.1, Medtr7g094100.1, Medtr7g094100.2, Medtr7g096980.1, Medtr7g098610.1, Medtr7g099220.1, Medtr7g100630.1, Medtr7g103180.1, Medtr7g103440.1, Medtr7g106210.1, Medtr7g109670.1, Medtr7g111690.1, Medtr7g111690.2, Medtr7g111690.3, Medtr7g114300.4, Medtr7g115300.1, Medtr7g115740.1, Medtr7g115740.2, Medtr7g116130.1, Medtr7g116650.1, Medtr7g117520.1, Medtr7g118110.1, Medtr7g407090.1, Medtr7g407130.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | **Protein IDs** |
|---|---|
| MT | Medtr8g010180.1, Medtr8g010180.2, Medtr8g010180.3, Medtr8g011370.1, Medtr8g011440.1, Medtr8g014500.1, Medtr8g014500.2, Medtr8g014500.3, Medtr8g014500.4, Medtr8g014500.5, Medtr8g014790.1Medtr8g014860.1, Medtr8g014860.2, Medtr8g015100.2, Medtr8g015170.1, Medtr8g015200.1, Medtr8g016330.1, Medtr8g021350.1, Medtr8g023560.1, Medtr8g023560.2, Medtr8g023720.1, Medtr8g024060.1, Medtr8g028065.1, Medtr8g028695.1, Medtr8g035560.1, Medtr8g037700.1, Medtr8g041650.1, Medtr8g041660.1, Medtr8g041660.2, Medtr8g041670.1, Medtr8g041690.1, Medtr8g041710.1, Medtr8g041880.1, Medtr8g041890.1, Medtr8g041910.1, Medtr8g044230.1, Medtr8g046290.1, Medtr8g047160.1, Medtr8g047210.1, Medtr8g047220.1, Medtr8g047230.1, Medtr8g051540.1, Medtr8g051600.1, Medtr8g052050.1, Medtr8g052060.1, Medtr8g052120.1, Medtr8g052190.1, Medtr8g052190.2, Medtr8g052200.1, Medtr8g052290.1, Medtr8g052490.1, Medtr8g064690.1, Medtr8g066700.1, Medtr8g067630.1, Medtr8g067690.1, Medtr8g067720.1, Medtr8g067735.1, Medtr8g067930.1, Medtr8g068050.1, Medtr8g068050.2, Medtr8g068540.1, Medtr8g070880.1, Medtr8g070910.1, Medtr8g074920.1, Medtr8g077850.1, Medtr8g078300.1, Medtr8g078940.1, Medtr8g078940.2, Medtr8g083240.1, Medtr8g086390.1, Medtr8g086390.2, Medtr8g086490.1, Medtr8g087420.1, Medtr8g087740.1, Medtr8g088760.1, Medtr8g088780.1, Medtr8g089200.1, Medtr8g089210.1, Medtr8g090140.1, Medtr8g090140.2, Medtr8g090140.3, Medtr8g091710.1, Medtr8g098510.1, Medtr8g098510.2, Medtr8g098510.4, Medtr8g099195.1, Medtr8g099195.2, Medtr8g101260.1, Medtr8g104520.1, Medtr8g106100.1, Medtr8g107470.1, Medtr8g442270.1, Medtr8g442290.1, Medtr8g442370.1, Medtr8g461120.1, Medtr8g461120.2, Medtr8g461120.3, Medtr8g464860.1, Medtr8g465340.1, Medtr8g467150.1, Medtr8g468500.1, Medtr8g469570.1, Medtr8g469600.1, Medtr8g469650.1, Medtr8g469780.1, Medtr8g469830.1, Medtr8g469980.1, Medtr8g470050.1, Medtr8g470370.1, Medtr8g470400.1, Medtr8g470560.1, Medtr8g470950.1, |
| PV | Phvul.001G003300.2.p, Phvul.001G014400.1.p, Phvul.001G017000.1.p, Phvul.001G040000.1.p, Phvul.001G040100.1.p, Phvul.001G040300.1.p, Phvul.001G040400.1.p, Phvul.001G040500.1.p, Phvul.001G040600.1.p, Phvul.001G040700.1.p, Phvul.001G040800.1.p, Phvul.001G043001.1.p, Phvul.001G043001.2.p, Phvul.001G045400.1.p, Phvul.001G047100.1.p, Phvul.001G047100.2.p, Phvul.001G047900.1.p, Phvul.001G048050.1.p, Phvul.001G048200.1.p, Phvul.001G048500.1.p, Phvul.001G048600.1.p, Phvul.001G048800.1.p, Phvul.001G048900.1.p, Phvul.001G049000.1.p, Phvul.001G050700.1.p, Phvul.001G051100.1.p, Phvul.001G068500.1.p, Phvul.001G068800.1.p, Phvul.001G069216.1.p, Phvul.001G069232.1.p, Phvul.001G069248.1.p, Phvul.001G069264.1.p, Phvul.001G069300.1.p, Phvul.001G069600.1.p, Phvul.001G070000.1.p, Phvul.001G070200.1.p, Phvul.001G071100.1.p, Phvul.001G071200.1.p, Phvul.001G071300.1.p, Phvul.001G071401.1.p, Phvul.001G075600.1.p, Phvul.001G075800.2.p, Phvul.001G076100.1.p, Phvul.001G076200.1.p, Phvul.001G076300.1.p, Phvul.001G083100.1.p, Phvul.001G083300.1.p, Phvul.001G095500.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| PV | Phvul.001G115100.1.p, Phvul.001G115100.2.p, Phvul.001G130800.2.p, Phvul.001G134600.1.p, Phvul.001G134600.2.p, Phvul.001G134600.3.p, Phvul.001G137300.1.p, Phvul.001G137300.2.p, Phvul.001G146900.1.p, Phvul.001G154300.1.p, Phvul.001G163700.1.p, Phvul.001G163800.1.p, Phvul.001G165100.1.p, Phvul.001G167700.1.p, Phvul.001G174500.1.p, Phvul.001G174500.2.p, Phvul.001G184100.1.p, Phvul.001G185400.1.p, Phvul.001G195500.1.p, Phvul.001G214000.1.p, Phvul.001G214000.2.p, Phvul.001G214000.3.p, Phvul.001G214000.4.p, Phvul.001G222700.1.p, Phvul.001G234200.1.p, Phvul.001G236600.1.p, Phvul.001G243800.1.p, Phvul.001G245300.1.p, Phvul.001G246800.2.p, Phvul.001G252800.1.p, Phvul.001G252800.2.p, Phvul.001G252800.3.p, Phvul.001G252800.4.p, Phvul.001G265500.1.p, Phvul.002G000200.1.p, Phvul.002G007800.1.p, Phvul.002G017700.1.p, Phvul.002G017700.2.p, Phvul.002G023400.1.p, Phvul.002G025500.1.p, Phvul.002G025700.1.p, Phvul.002G029200.1.p, Phvul.002G029200.2.p, Phvul.002G049301.1.p, Phvul.002G049400.2.p, Phvul.002G049500.1.p, Phvul.002G055600.1.p, Phvul.002G059500.1.p, Phvul.002G060200.1.p, Phvul.002G063600.2.p, Phvul.002G063700.1.p, Phvul.002G063700.2.p, Phvul.002G063900.1.p, Phvul.002G064400.1.p, Phvul.002G070900.1.p, Phvul.002G080400.1.p, Phvul.002G089300.1.p, Phvul.002G111100.1.p, Phvul.002G117600.1.p, Phvul.002G117600.2.p, Phvul.002G119900.1.p, Phvul.002G123200.1.p, Phvul.002G123500.1.p, Phvul.002G125100.1.p, Phvul.002G126600.1.p, Phvul.002G127800.1.p, Phvul.002G152100.1.p, Phvul.002G162400.1.p, Phvul.002G177200.1.p, Phvul.002G182500.1.p, Phvul.002G184800.1.p, Phvul.002G187300.1.p, Phvul.002G196200.1.p, Phvul.002G201400.1.p, Phvul.002G206900.1.p, Phvul.002G206900.2.p, Phvul.002G206900.3.p, Phvul.002G207000.1.p, Phvul.002G214900.1.p, Phvul.002G215000.1.p, Phvul.002G215100.1.p, Phvul.002G215200.1.p, Phvul.002G215300.1.p, Phvul.002G215400.1.p, Phvul.002G218800.1.p, Phvul.002G224200.1.p, Phvul.002G224200.2.p, Phvul.002G224200.3.p, Phvul.002G232600.1.p, Phvul.002G232700.1.p, Phvul.002G238900.1.p, Phvul.002G238900.2.p, Phvul.002G238900.3.p, Phvul.002G239400.1.p, Phvul.002G242600.1.p, Phvul.002G242600.2.p, Phvul.002G249500.1.p, Phvul.002G251400.1.p, Phvul.002G257000.1.p, Phvul.002G257000.2.p, Phvul.002G264200.1.p, Phvul.002G295600.1.p, Phvul.002G302600.1.p, Phvul.002G318200.1.p, Phvul.002G318300.1.p, Phvul.002G324600.1.p, Phvul.002G331400.1.p, Phvul.003G015800.1.p, Phvul.003G015900.1.p, Phvul.003G024000.1.p, Phvul.003G027300.1.p, Phvul.003G027600.1.p, Phvul.003G027600.2.p, Phvul.003G029700.1.p, Phvul.003G030000.2.p, Phvul.003G038700.1.p, Phvul.003G038800.1.p, Phvul.003G040200.1.p, Phvul.003G044400.1.p, Phvul.003G044400.2.p, Phvul.003G044400.3.p, Phvul.003G052300.2.p, Phvul.003G060600.1.p, Phvul.003G062600.1.p, Phvul.003G062700.1.p, Phvul.003G062700.2.p, Phvul.003G063700.1.p, Phvul.003G068400.1.p, Phvul.003G090000.1.p, Phvul.003G090000.2.p, Phvul.003G121000.1.p, Phvul.003G121000.2.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| PV | Phvul.003G129400.1.p, Phvul.003G137700.1.p, Phvul.003G137700.2.p, Phvul.003G152900.1.p, Phvul.003G156366.1.p, Phvul.003G158600.1.p, Phvul.003G158700.1.p, Phvul.003G159700.1.p, Phvul.003G161500.1.p, Phvul.003G187200.1.p, Phvul.003G188000.1.p, Phvul.003G193100.1.p, Phvul.003G202000.1.p, Phvul.003G203100.1.p, Phvul.003G204500.1.p, Phvul.003G214800.1.p, Phvul.003G231350.1.p, Phvul.003G231350.2.p, Phvul.003G231350.3.p, Phvul.003G231350.4.p, Phvul.003G231400.1.p, Phvul.003G239300.2.p, Phvul.003G239400.2.p, Phvul.003G239500.1.p, Phvul.003G240400.1.p, Phvul.003G241400.1.p, Phvul.003G249300.1.p, Phvul.003G252900.1.p, Phvul.003G252900.2.p, Phvul.003G260300.1.p, Phvul.003G267000.1.p, Phvul.003G271000.3.p, Phvul.003G273700.1.p, Phvul.003G287101.1.p, Phvul.004G011000.1.p, Phvul.004G023700.1.p, Phvul.004G037500.2.p, Phvul.004G037600.2.p, Phvul.004G037600.3.p, Phvul.004G038800.1.p, Phvul.004G039200.1.p, Phvul.004G039400.1.p, Phvul.004G039501.1.p, Phvul.004G039600.1.p, Phvul.004G039700.1.p, Phvul.004G039800.1.p, Phvul.004G039900.1.p, Phvul.004G040000.1.p, Phvul.004G040300.2.p, Phvul.004G043800.1.p, Phvul.004G043800.2.p, Phvul.004G043800.4.p, Phvul.004G043800.5.p, Phvul.004G046800.1.p, Phvul.004G055400.1.p, Phvul.004G072600.1.p, Phvul.004G084688.1.p, Phvul.004G086300.1.p, Phvul.004G086300.2.p, Phvul.004G109500.1.p, Phvul.004G124400.1.p, Phvul.004G124900.1.p, Phvul.004G124900.2.p, Phvul.004G125200.1.p, Phvul.004G126100.1.p, Phvul.004G128100.1.p, Phvul.004G136100.1.p, Phvul.004G136500.1.p, Phvul.004G149100.1.p, Phvul.004G154800.1.p, Phvul.004G154900.1.p, Phvul.004G155000.1.p, Phvul.004G155300.1.p, Phvul.004G155400.2.p, Phvul.004G155500.1.p, Phvul.004G155600.1.p, Phvul.004G164000.1.p, Phvul.004G175700.1.p, Phvul.004G175800.1.p, Phvul.004G175900.1.p, Phvul.004G177900.1.p, Phvul.005G005900.1.p, Phvul.005G014900.1.p, Phvul.005G015100.1.p, Phvul.005G054300.1.p, Phvul.005G064700.1.p, Phvul.005G064700.2.p, Phvul.005G064700.3.p, Phvul.005G073800.1.p, Phvul.005G077400.1.p, Phvul.005G077600.1.p, Phvul.005G077900.1.p, Phvul.005G078400.1.p, Phvul.005G078400.2.p, Phvul.005G078500.1.p, Phvul.005G078900.2.p, Phvul.005G079733.1.p, Phvul.005G079733.2.p, Phvul.005G079733.3.p, Phvul.005G079733.4.p, Phvul.005G085600.1.p, Phvul.005G085600.2.p, Phvul.005G090000.1.p, Phvul.005G090000.2.p, Phvul.005G090000.3.p, Phvul.005G094300.1.p, Phvul.005G100400.1.p, Phvul.005G103200.1.p, Phvul.005G103300.1.p, Phvul.005G106800.1.p, Phvul.005G110300.3.p, Phvul.005G120300.1.p, Phvul.005G133700.1.p, Phvul.005G135500.1.p, Phvul.005G135500.2.p, Phvul.005G135500.3.p, Phvul.005G153100.1.p, Phvul.005G159200.1.p, Phvul.005G159200.2.p, Phvul.005G159200.3.p, Phvul.005G159200.4.p, Phvul.005G162300.1.p, Phvul.005G162600.1.p, Phvul.005G163000.1.p, Phvul.005G163200.1.p, Phvul.005G163400.1.p, Phvul.005G163800.10.p, Phvul.005G163800.11.p, Phvul.005G163800.12.p, Phvul.005G163800.13.p, Phvul.005G163800.2.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| PV | Phvul.005G163800.3.pPhvul.005G163800.4.p, Phvul.005G163800.5.p, Phvul.005G163800.6.p, Phvul.005G163800.7.p, Phvul.005G163800.8.p, Phvul.005G163800.9.p, Phvul.005G163901.1.p, Phvul.005G164100.1.p, Phvul.005G164200.1.p, Phvul.005G165900.1.p, Phvul.005G165900.2.p, Phvul.005G167700.1.p, Phvul.005G177900.1.p, Phvul.005G181000.1.p, Phvul.005G181000.2.p, Phvul.005G181000.3.p, Phvul.005G181000.4.p, Phvul.005G182900.1.p, Phvul.005G183300.1.p, Phvul.005G184700.1.p, Phvul.005G185800.1.p, Phvul.006G006700.1.p, Phvul.006G006700.3.p, Phvul.006G006700.4.p, Phvul.006G006700.5.p, Phvul.006G006700.6.p, Phvul.006G006800.3.p, Phvul.006G029000.1.p, Phvul.006G032300.1.p, Phvul.006G033300.1.p, Phvul.006G057900.1.p, Phvul.006G068400.1.p, Phvul.006G069600.1.p, Phvul.006G071600.1.p, Phvul.006G071600.2.p, Phvul.006G076200.2.p, Phvul.006G076200.3.p, Phvul.006G080200.1.p, Phvul.006G080500.1.p, Phvul.006G084400.1.p, Phvul.006G084500.1.p, Phvul.006G084600.1.p, Phvul.006G084700.1.p, Phvul.006G084800.1.p, Phvul.006G084900.3.p, Phvul.006G087700.1.p, Phvul.006G108300.1.p, Phvul.006G125300.1.p, Phvul.006G125300.2.p, Phvul.006G127700.1.p, Phvul.006G127700.2.p, Phvul.006G127900.1.p, Phvul.006G133200.1.p, Phvul.006G133400.1.p, Phvul.006G133400.2.p, Phvul.006G137300.1.p, Phvul.006G146100.1.p, Phvul.006G146400.1.p, Phvul.006G148000.1.p, Phvul.006G170500.1.p, Phvul.006G170500.2.p, Phvul.006G170500.3.p, Phvul.006G170500.4.p, Phvul.006G170500.5.p, Phvul.006G174400.1.p, Phvul.006G174400.2.p, Phvul.006G174700.1.p, Phvul.006G175100.1.p, Phvul.006G175100.2.p, Phvul.006G175100.3.p, Phvul.006G175100.4.p, Phvul.006G180200.1.p, Phvul.006G181600.1.p, Phvul.006G181800.2.p, Phvul.006G185000.1.p, Phvul.006G189800.1.p, Phvul.006G193000.1.p, Phvul.006G193000.2.p, Phvul.006G198200.1.p, Phvul.006G200500.1.p, Phvul.006G200800.1.p, Phvul.006G207000.1.p, Phvul.006G208100.1.p, Phvul.006G215500.1.p, Phvul.007G029900.1.p, Phvul.007G030000.1.p, Phvul.007G030300.1.p, Phvul.007G030800.1.p, Phvul.007G030900.1.p, Phvul.007G031000.1.p, Phvul.007G031100.1.p, Phvul.007G031100.2.p, Phvul.007G031300.1.p, Phvul.007G040400.1.p, Phvul.007G048500.1.p, Phvul.007G048600.1.p, Phvul.007G048700.1.p, Phvul.007G048800.1.p, Phvul.007G048900.1.p, Phvul.007G049000.1.p, Phvul.007G049000.2.p, Phvul.007G049100.1.p, Phvul.007G049400.1.p, Phvul.007G049500.1.p, Phvul.007G049600.1.p, Phvul.007G049900.5.p, Phvul.007G050000.1.p, Phvul.007G050200.1.p, Phvul.007G050300.1.p, Phvul.007G050400.1.p, Phvul.007G050500.1.p, Phvul.007G050600.1.p, Phvul.007G050700.1.p, Phvul.007G051000.1.p, Phvul.007G051100.1.p, Phvul.007G051200.1.p, Phvul.007G051300.1.p, Phvul.007G051500.1.p, Phvul.007G052500.1.p, Phvul.007G063200.1.p, Phvul.007G063200.2.p, Phvul.007G067700.1.p, Phvul.007G074000.1.p, Phvul.007G074000.2.p, Phvul.007G078200.1.p, Phvul.007G082500.1.p, Phvul.007G084400.1.p, Phvul.007G086800.1.p, Phvul.007G093900.1.p, Phvul.007G112100.1.p, Phvul.007G137800.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| PV | Phvul.007G138200.1.p, Phvul.007G154500.1.p, Phvul.007G163000.1.p, Phvul.007G168100.1.p, Phvul.007G181600.1.p, Phvul.007G204300.1.p, Phvul.007G204300.2.p, Phvul.007G213400.1.p, Phvul.007G220700.1.p, Phvul.007G221800.1.p, Phvul.007G224300.1.p, Phvul.007G234400.1.p, Phvul.007G252500.1.p, Phvul.007G258000.1.p, Phvul.007G258000.2.p, Phvul.007G260300.1.p, Phvul.007G260400.2.p, Phvul.007G260500.1.p, Phvul.007G268200.1.p, Phvul.007G280200.1.p, Phvul.008G000200.2.p, Phvul.008G007900.1.p, Phvul.008G007900.2.p, Phvul.008G007900.3.p, Phvul.008G017400.1.p, Phvul.008G019200.1.p, Phvul.008G019200.2.p, Phvul.008G022700.1.p, Phvul.008G022700.2.p, Phvul.008G025200.1.p, Phvul.008G030200.1.p, Phvul.008G030400.1.p, Phvul.008G030700.1.p, Phvul.008G030800.1.p, Phvul.008G044400.1.p, Phvul.008G044600.1.p, Phvul.008G047600.1.p, Phvul.008G047800.1.p, Phvul.008G050700.1.p, Phvul.008G050700.2.p, Phvul.008G056800.1.p, Phvul.008G056900.1.p, Phvul.008G058600.1.p, Phvul.008G058600.2.p, Phvul.008G058600.3.p, Phvul.008G058600.4.p, Phvul.008G058600.5.p, Phvul.008G058600.6.p, Phvul.008G058600.7.p, Phvul.008G058700.1.p, Phvul.008G058700.2.p, Phvul.008G058800.1.p, Phvul.008G059400.1.p, Phvul.008G077500.1.p, Phvul.008G077500.2.p, Phvul.008G077800.1.p, Phvul.008G079800.1.p, Phvul.008G080000.2.p, Phvul.008G080500.1.p, Phvul.008G081000.1.p, Phvul.008G082400.1.p, Phvul.008G082500.1.p, Phvul.008G093000.1.p, Phvul.008G093200.1.p, Phvul.008G093200.2.p, Phvul.008G093200.3.p, Phvul.008G094500.1.p, Phvul.008G105600.1.p, Phvul.008G106500.1.p, Phvul.008G106600.1.p, Phvul.008G115500.1.p, Phvul.008G117700.1.p, Phvul.008G117800.1.p, Phvul.008G118500.1.p, Phvul.008G122200.1.p, Phvul.008G122400.1.p, Phvul.008G122400.2.p, Phvul.008G127237.1.p, Phvul.008G146000.1.p, Phvul.008G146600.1.p, Phvul.008G152500.1.p, Phvul.008G154200.1.p, Phvul.008G156400.1.p, Phvul.008G156400.2.p, Phvul.008G158100.1.p, Phvul.008G159500.1.p, Phvul.008G160700.1.p, Phvul.008G160700.2.p, Phvul.008G174300.1.p, Phvul.008G174900.2.p, Phvul.008G175100.1.p, Phvul.008G175200.1.p, Phvul.008G186900.1.p, Phvul.008G186900.2.p, Phvul.008G188700.1.p, Phvul.008G192900.1.p, Phvul.008G192900.2.p, Phvul.008G207200.2.p, Phvul.008G207200.3.p, Phvul.008G211200.1.p, Phvul.008G212500.1.p, Phvul.008G212900.1.p, Phvul.008G213200.1.p, Phvul.008G220100.1.p, Phvul.008G221900.1.p, Phvul.008G223900.1.p, Phvul.008G224900.1.p, Phvul.008G225000.1.p, Phvul.008G236900.1.p, Phvul.008G238600.1.p, Phvul.008G238600.2.p, Phvul.008G239600.1.p, Phvul.008G239800.1.p, Phvul.008G239800.2.p, Phvul.008G244400.1.p, Phvul.008G245000.1.p, Phvul.008G245600.1.p, Phvul.008G245600.2.p, Phvul.008G245600.3.p, Phvul.008G245600.4.p, Phvul.008G245600.5.p, Phvul.008G245900.1.p, Phvul.008G245900.2.p, Phvul.008G258700.1.p, Phvul.008G259700.1.p, Phvul.008G275000.1.p, Phvul.008G279300.1.p, Phvul.008G281500.3.p, Phvul.008G288500.1.p, Phvul.009G014500.1.p, Phvul.009G023800.1.p, Phvul.009G023900.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| PV | Phvul.009G067200.1.p, Phvul.009G079300.3.p, Phvul.009G079300.4.p, Phvul.009G024200.1.p, Phvul.009G042300.1.p, Phvul.009G042300.2.p, Phvul.009G043600.1.p, Phvul.009G043600.2.p, Phvul.009G043600.3.p, Phvul.009G046600.1.p, Phvul.009G046900.1.p, Phvul.009G061350.1.p, Phvul.009G089500.1.p, Phvul.009G089500.2.p, Phvul.009G089500.3.p, Phvul.009G089500.4.p, Phvul.009G092900.1.p, Phvul.009G103100.1.p, Phvul.009G106200.1.p, Phvul.009G110600.1.p, Phvul.009G112200.1.p, Phvul.009G114400.1.p, Phvul.009G114500.1.p, Phvul.009G122600.1.p, Phvul.009G136200.1.p, Phvul.009G136800.1.p, Phvul.009G136800.2.p, Phvul.009G136800.3.p, Phvul.009G138532.1.p, Phvul.009G140500.1.p, Phvul.009G167600.1.p, Phvul.009G167600.2.p, Phvul.009G167600.3.p, Phvul.009G171200.1.p, Phvul.009G180700.1.p, Phvul.009G184500.2.p, Phvul.009G190400.1.p, Phvul.009G190400.2.p, Phvul.009G190400.3.p, Phvul.009G190400.4.p, Phvul.009G195900.1.p, Phvul.009G198800.1.p, Phvul.009G208500.1.p, Phvul.009G209200.1.p, Phvul.009G219300.1.p, Phvul.009G224500.1.p, Phvul.009G225700.1.p, Phvul.009G232100.1.p, Phvul.009G232100.2.p, Phvul.009G232100.3.p, Phvul.009G232100.4.p, Phvul.009G232100.5.p, Phvul.009G232100.6.p, Phvul.009G232100.7.p, Phvul.009G232100.8.p, Phvul.009G233300.1.p, Phvul.009G233300.2.p, Phvul.009G239500.1.p, Phvul.009G239700.1.p, Phvul.009G239800.1.p, Phvul.009G257000.1.p, Phvul.009G259800.1.p, Phvul.009G260500.1.p, Phvul.009G260900.1.p, Phvul.009G260900.3.p, Phvul.010G015800.1.p, Phvul.010G028300.1.p, Phvul.010G031900.1.p, Phvul.010G032000.1.p, Phvul.010G035800.1.p, Phvul.010G035800.2.p, Phvul.010G060800.1.p, Phvul.010G064300.1.p, Phvul.010G064900.1.p, Phvul.010G070400.1.p, Phvul.010G082400.1.p, Phvul.010G082800.1.p, Phvul.010G086600.1.p, Phvul.010G103100.1.p, Phvul.010G122100.1.p, Phvul.010G122100.2.p, Phvul.010G127500.1.p, Phvul.010G138900.1.p, Phvul.010G143100.2.p, Phvul.010G143100.3.p, Phvul.010G143100.4.p, Phvul.010G145100.1.p, Phvul.010G149300.1.p, Phvul.010G165200.1.p, Phvul.010G165200.2.p, Phvul.011G003100.1.p, Phvul.011G003200.1.p, Phvul.011G003200.2.p, Phvul.011G007200.1.p, Phvul.011G007800.1.p, Phvul.011G007900.1.p, Phvul.011G008100.1.p, Phvul.011G033000.2.p, Phvul.011G033000.3.p, Phvul.011G033000.4.p, Phvul.011G042000.1.p, Phvul.011G045600.1.p, Phvul.011G054300.1.p, Phvul.011G067401.1.p, Phvul.011G069600.1.p, Phvul.011G081200.1.p, Phvul.011G081800.1.p, Phvul.011G112700.1.p, Phvul.011G119200.1.p, Phvul.011G123500.1.p, Phvul.011G142200.1.p, Phvul.011G148700.1.p, Phvul.011G161000.1.p, Phvul.011G166400.1.p, Phvul.011G169300.1.p, Phvul.011G169332.1.p, Phvul.011G169332.2.p, Phvul.011G169332.3.p, Phvul.011G169332.4.p, Phvul.011G169332.5.p, Phvul.011G175700.1.p, Phvul.011G175700.2.p, Phvul.011G176400.1.p, Phvul.011G176400.2.p, Phvul.011G176400.3.p, Phvul.011G176700.1.p, Phvul.011G181900.1.p, Phvul.011G182300.1.p, Phvul.011G185500.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| PV | Phvul.011G193200.1.p, Phvul.011G193300.1.p, Phvul.011G193300.2.p, Phvul.011G194200.2.p, Phvul.011G194300.3.p, Phvul.011G194300.4.p, Phvul.011G194401.1.p, Phvul.011G194401.2.p, Phvul.011G194600.1.p, Phvul.011G194700.1.p, Phvul.011G196200.1.p, Phvul.011G197500.1.p, Phvul.011G197700.1.p, Phvul.011G197700.2.p, Phvul.011G198200.1.p, Phvul.011G198500.1.p, Phvul.011G198500.2.p, Phvul.011G198500.3.p, Phvul.011G200400.1.p, Phvul.011G210400.1.p, Phvul.011G214400.1.p, Phvul.L002151.1.p, Phvul.L003081.1.p |
| VA | gi\|1044521823\|ref\|XP_017442814.1\|, gi\|1044521839\|ref\|XP_017442823.1\|, gi\|1044521847\|ref\|XP_017442828.1\|, gi\|1044522253\|ref\|XP_017405392.1\|, gi\|1044522373\|ref\|XP_017405461.1\|, gi\|1044522375\|ref\|XP_017405462.1\|, gi\|1044522637\|ref\|XP_017405606.1\|, gi\|1044522798\|ref\|XP_017405697.1\|, gi\|1044523405\|ref\|XP_017406017.1\|, gi\|1044523597\|ref\|XP_017406123.1\|, gi\|1044523599\|ref\|XP_017406124.1\|, gi\|1044523601\|ref\|XP_017406125.1\|, gi\|1044523603\|ref\|XP_017406126.1\|, gi\|1044523605\|ref\|XP_017406127.1\|, gi\|1044523607\|ref\|XP_017406129.1\|, gi\|1044523609\|ref\|XP_017406130.1\|, gi\|1044523692\|ref\|XP_017406177.1\|, gi\|1044523702\|ref\|XP_017406182.1\|, gi\|1044523704\|ref\|XP_017406183.1\|, gi\|1044524074\|ref\|XP_017406386.1\|, gi\|1044524362\|ref\|XP_017406545.1\|, gi\|1044524364\|ref\|XP_017406546.1\|, gi\|1044524366\|ref\|XP_017406547.1\|, gi\|1044524490\|ref\|XP_017406616.1\|, gi\|1044524564\|ref\|XP_017406655.1\|, gi\|1044524566\|ref\|XP_017406656.1\|, gi\|1044524586\|ref\|XP_017406666.1\|, gi\|1044524624\|ref\|XP_017406687.1\|, gi\|1044524626\|ref\|XP_017406688.1\|, gi\|1044524737\|ref\|XP_017406750.1\|, gi\|1044524967\|ref\|XP_017406876.1\|, gi\|1044525017\|ref\|XP_017406904.1\|, gi\|1044525403\|ref\|XP_017407113.1\|, gi\|1044525462\|ref\|XP_017407147.1\|, gi\|1044525480\|ref\|XP_017407156.1\|, gi\|1044525574\|ref\|XP_017407207.1\|, gi\|1044526082\|ref\|XP_017407489.1\|, gi\|1044526290\|ref\|XP_017407597.1\|, gi\|1044526292\|ref\|XP_017407598.1\|, gi\|1044526435\|ref\|XP_017407677.1\|, gi\|1044526437\|ref\|XP_017407678.1\|, gi\|1044526539\|ref\|XP_017407735.1\|, gi\|1044527133\|ref\|XP_017408056.1\|, gi\|1044527260\|ref\|XP_017408125.1\|, gi\|1044527312\|ref\|XP_017408153.1\|, gi\|1044528057\|ref\|XP_017408558.1\|, gi\|1044528059\|ref\|XP_017408559.1\|, gi\|1044528061\|ref\|XP_017408560.1\|, gi\|1044528095\|ref\|XP_017408579.1\|, gi\|1044528097\|ref\|XP_017408580.1\|, gi\|1044528099\|ref\|XP_017408581.1\|, gi\|1044528101\|ref\|XP_017408582.1\|, gi\|1044528103\|ref\|XP_017408583.1\|, gi\|1044528105\|ref\|XP_017408584.1\|, gi\|1044528107\|ref\|XP_017408586.1\|, gi\|1044528109\|ref\|XP_017408587.1\|, gi\|1044528111\|ref\|XP_017408588.1\|, gi\|1044528176\|ref\|XP_017408618.1\|, gi\|1044528178\|ref\|XP_017408619.1\|, gi\|1044528216\|ref\|XP_017408644.1\|, gi\|1044528673\|ref\|XP_017408895.1\|, gi\|1044528725\|ref\|XP_017408925.1\|, gi\|1044528727\|ref\|XP_017408926.1\|, gi\|1044528991\|ref\|XP_017409066.1\|, gi\|1044528993\|ref\|XP_017409067.1\|, gi\|1044528995\|ref\|XP_017409068.1\|, gi\|1044529161\|ref\|XP_017409154.1\|, gi\|1044529163\|ref\|XP_017409155.1\|, gi\|1044529288\|ref\|XP_017409224.1\|, gi\|1044529338\|ref\|XP_017409251.1\|, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044529429|ref|XP_017409301.1|, gi|1044529431|ref|XP_017409302.1|, gi|1044529705|ref|XP_017409446.1|, gi|1044530337|ref|XP_017409789.1|, gi|1044530466|ref|XP_017409859.1|, gi|1044530507|ref|XP_017409883.1|, gi|1044530509|ref|XP_017409884.1|, gi|1044530511|ref|XP_017409885.1|, gi|1044530658|ref|XP_017409966.1|, gi|1044530694|ref|XP_017409985.1|, gi|1044531049|ref|XP_017410178.1|, gi|1044531051|ref|XP_017410179.1|, gi|1044531053|ref|XP_017410181.1|, gi|1044531055|ref|XP_017410182.1|, gi|1044531057|ref|XP_017410183.1|, gi|1044531258|ref|XP_017410293.1|, gi|1044531352|ref|XP_017410345.1|, gi|1044531519|ref|XP_017410440.1|, gi|1044531825|ref|XP_017410613.1|, gi|1044531900|ref|XP_017410653.1|, gi|1044531902|ref|XP_017410654.1|, gi|1044531904|ref|XP_017410655.1|, gi|1044531906|ref|XP_017410656.1|, gi|1044532202|ref|XP_017410821.1|, gi|1044532204|ref|XP_017410822.1|, gi|1044532351|ref|XP_017410905.1|, gi|1044532526|ref|XP_017410998.1|, gi|1044532544|ref|XP_017411008.1|, gi|1044532617|ref|XP_017411048.1|, gi|1044532647|ref|XP_017411065.1|, gi|1044532649|ref|XP_017411066.1|, gi|1044532673|ref|XP_017411080.1|, gi|1044532675|ref|XP_017411082.1|, gi|1044532753|ref|XP_017411125.1|, gi|1044532779|ref|XP_017411140.1|, gi|1044532781|ref|XP_017411141.1|, gi|1044532913|ref|XP_017411212.1|, gi|1044532915|ref|XP_017411213.1|, gi|1044532997|ref|XP_017411257.1|, gi|1044533099|ref|XP_017411315.1|, gi|1044533225|ref|XP_017411380.1|, gi|1044533626|ref|XP_017411595.1|, gi|1044533700|ref|XP_017411637.1|, gi|1044533827|ref|XP_017411706.1|, gi|1044533983|ref|XP_017411794.1|, gi|1044534223|ref|XP_017411927.1|, gi|1044534249|ref|XP_017411939.1|, gi|1044534378|ref|XP_017412696.1|, gi|1044534398|ref|XP_017413506.1|, gi|1044534648|ref|XP_017425296.1|, gi|1044534690|ref|XP_017427331.1|, gi|1044534832|ref|XP_017434333.1|, gi|1044535260|ref|XP_017412427.1|, gi|1044535276|ref|XP_017412515.1|, gi|1044535278|ref|XP_017412522.1|, gi|1044535280|ref|XP_017412534.1|, gi|1044535678|ref|XP_017414695.1|, gi|1044535698|ref|XP_017414913.1|, gi|1044535754|ref|XP_017415415.1|, gi|1044535770|ref|XP_017415637.1|, gi|1044535772|ref|XP_017415650.1|, gi|1044535774|ref|XP_017415660.1|, gi|1044535790|ref|XP_017415795.1|, gi|1044535976|ref|XP_017417256.1|, gi|1044536048|ref|XP_017417913.1|, gi|1044536086|ref|XP_017418161.1|, gi|1044536278|ref|XP_017419735.1|, gi|1044536352|ref|XP_017420244.1|, gi|1044536750|ref|XP_017422206.1|, gi|1044537198|ref|XP_017424440.1|, gi|1044537264|ref|XP_017424753.1|, gi|1044538231|ref|XP_017429557.1|, gi|1044538233|ref|XP_017429565.1|, gi|1044538235|ref|XP_017429574.1|, gi|1044538336|ref|XP_017430087.1|, gi|1044538745|ref|XP_017432058.1|, gi|1044539315|ref|XP_017435142.1|, gi|1044539528|ref|XP_017436294.1|, gi|1044539546|ref|XP_017436398.1|, gi|1044539550|ref|XP_017436420.1|, gi|1044539684|ref|XP_017437124.1|, gi|1044539688|ref|XP_017437144.1|, gi|1044539690|ref|XP_017437152.1|, gi|1044539698|ref|XP_017437199.1|, gi|1044539704|ref|XP_017437235.1|, gi|1044540120|ref|XP_017438876.1|, gi|1044540122|ref|XP_017438884.1|,</td></tr>
</table>

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VA | gi\|1044540124\|ref\|XP_017438894.1\|, gi\|1044540126\|ref\|XP_017438902.1\|, gi\|1044540128\|ref\|XP_017438910.1\|, gi\|1044540504\|ref\|XP_017440743.1\|, gi\|1044540570\|ref\|XP_017441040.1\|, gi\|1044540572\|ref\|XP_017441045.1\|, gi\|1044540574\|ref\|XP_017441050.1\|, gi\|1044540576\|ref\|XP_017441053.1\|, gi\|1044540578\|ref\|XP_017441055.1\|, gi\|1044540672\|ref\|XP_017441329.1\|, gi\|1044541066\|ref\|XP_017405574.1\|, gi\|1044541511\|ref\|XP_017407558.1\|, gi\|1044541535\|ref\|XP_017407683.1\|, gi\|1044541630\|ref\|XP_017408118.1\|, gi\|1044542047\|ref\|XP_017410065.1\|, gi\|1044542061\|ref\|XP_017410126.1\|, gi\|1044542063\|ref\|XP_017410132.1\|, gi\|1044542361\|ref\|XP_017411513.1\|, gi\|1044542363\|ref\|XP_017411520.1\|, gi\|1044542655\|ref\|XP_017412055.1\|, gi\|1044542752\|ref\|XP_017412103.1\|, gi\|1044542805\|ref\|XP_017412131.1\|, gi\|1044542849\|ref\|XP_017412156.1\|, gi\|1044542859\|ref\|XP_017412161.1\|, gi\|1044542915\|ref\|XP_017412193.1\|, gi\|1044543334\|ref\|XP_017437633.1\|, gi\|1044543375\|ref\|XP_017437656.1\|, gi\|1044543387\|ref\|XP_017437662.1\|, gi\|1044543774\|ref\|XP_017437863.1\|, gi\|1044543776\|ref\|XP_017437864.1\|, gi\|1044543809\|ref\|XP_017437882.1\|, gi\|1044543859\|ref\|XP_017437908.1\|, gi\|1044543861\|ref\|XP_017437909.1\|, gi\|1044543931\|ref\|XP_017437949.1\|, gi\|1044543933\|ref\|XP_017437950.1\|, gi\|1044544001\|ref\|XP_017437992.1\|, gi\|1044544159\|ref\|XP_017438086.1\|, gi\|1044544209\|ref\|XP_017438118.1\|, gi\|1044544233\|ref\|XP_017438132.1\|, gi\|1044544457\|ref\|XP_017438273.1\|, gi\|1044544717\|ref\|XP_017438435.1\|, gi\|1044544719\|ref\|XP_017438436.1\|, gi\|1044544959\|ref\|XP_017438575.1\|, gi\|1044545771\|ref\|XP_017439012.1\|, gi\|1044546034\|ref\|XP_017439146.1\|, gi\|1044546252\|ref\|XP_017439264.1\|, gi\|1044546342\|ref\|XP_017439314.1\|, gi\|1044546387\|ref\|XP_017439339.1\|, gi\|1044546749\|ref\|XP_017439533.1\|, gi\|1044546751\|ref\|XP_017439534.1\|, gi\|1044546916\|ref\|XP_017439616.1\|, gi\|1044547172\|ref\|XP_017439748.1\|, gi\|1044547206\|ref\|XP_017439766.1\|, gi\|1044547208\|ref\|XP_017439767.1\|, gi\|1044547265\|ref\|XP_017439797.1\|, gi\|1044547317\|ref\|XP_017439825.1\|, gi\|1044547319\|ref\|XP_017439826.1\|, gi\|1044547321\|ref\|XP_017439827.1\|, gi\|1044547323\|ref\|XP_017439828.1\|, gi\|1044547325\|ref\|XP_017439829.1\|, gi\|1044547327\|ref\|XP_017439830.1\|, gi\|1044547655\|ref\|XP_017440004.1\|, gi\|1044547657\|ref\|XP_017440005.1\|, gi\|1044547852\|ref\|XP_017440111.1\|, gi\|1044547941\|ref\|XP_017440160.1\|, gi\|1044547955\|ref\|XP_017440166.1\|, gi\|1044548514\|ref\|XP_017440466.1\|, gi\|1044548631\|ref\|XP_017440528.1\|, gi\|1044548701\|ref\|XP_017440567.1\|, gi\|1044548703\|ref\|XP_017440568.1\|, gi\|1044548705\|ref\|XP_017440569.1\|, gi\|1044548836\|ref\|XP_017440638.1\|, gi\|1044548933\|ref\|XP_017440693.1\|, gi\|1044549141\|ref\|XP_017440804.1\|, gi\|1044549165\|ref\|XP_017440816.1\|, gi\|1044549223\|ref\|XP_017440846.1\|, gi\|1044549493\|ref\|XP_017440986.1\|, gi\|1044549683\|ref\|XP_017441108.1\|, gi\|1044549957\|ref\|XP_017441275.1\|, gi\|1044549993\|ref\|XP_017441296.1\|, gi\|1044550846\|ref\|XP_017441764.1\|, gi\|1044551707\|ref\|XP_017442221.1\|, gi\|1044552067\|ref\|XP_017442417.1\|, gi\|1044552101\|ref\|XP_017442436.1\|, gi\|1044552347\|ref\|XP_017442568.1\|, gi\|1044552401\|ref\|XP_017442597.1\|, gi\|1044552403\|ref\|XP_017442598.1\|, gi\|1044552405\|ref\|XP_017442599.1\|, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VA | gi\|1044552407\|ref\|XP_017442600.1\|, gi\|1044552463\|ref\|XP_017442632.1\|, |
| | gi\|1044552495\|ref\|XP_017442650.1\|, gi\|1044552497\|ref\|XP_017442651.1\|, |
| | gi\|1044552499\|ref\|XP_017442652.1\|, gi\|1044552733\|ref\|XP_017412351.1\|, |
| | gi\|1044552814\|ref\|XP_017412395.1\|, gi\|1044552816\|ref\|XP_017412396.1\|, |
| | gi\|1044552818\|ref\|XP_017412397.1\|, gi\|1044552961\|ref\|XP_017412476.1\|, |
| | gi\|1044552963\|ref\|XP_017412477.1\|, gi\|1044552965\|ref\|XP_017412478.1\|, |
| | gi\|1044553187\|ref\|XP_017412597.1\|, gi\|1044553541\|ref\|XP_017412781.1\|, |
| | gi\|1044553566\|ref\|XP_017412794.1\|, gi\|1044553659\|ref\|XP_017412847.1\|, |
| | gi\|1044553719\|ref\|XP_017412884.1\|, gi\|1044553775\|ref\|XP_017412917.1\|, |
| | gi\|1044554183\|ref\|XP_017413162.1\|, gi\|1044554243\|ref\|XP_017413198.1\|, |
| | gi\|1044554245\|ref\|XP_017413199.1\|, gi\|1044554247\|ref\|XP_017413200.1\|, |
| | gi\|1044554249\|ref\|XP_017413201.1\|, gi\|1044554253\|ref\|XP_017413203.1\|, |
| | gi\|1044554255\|ref\|XP_017413205.1\|, gi\|1044554275\|ref\|XP_017413216.1\|, |
| | gi\|1044554277\|ref\|XP_017413217.1\|, gi\|1044554279\|ref\|XP_017413218.1\|, |
| | gi\|1044554285\|ref\|XP_017413222.1\|, gi\|1044554287\|ref\|XP_017413223.1\|, |
| | gi\|1044554289\|ref\|XP_017413224.1\|, gi\|1044554527\|ref\|XP_017413359.1\|, |
| | gi\|1044554529\|ref\|XP_017413360.1\|, gi\|1044554638\|ref\|XP_017413421.1\|, |
| | gi\|1044554702\|ref\|XP_017413457.1\|, gi\|1044554715\|ref\|XP_017413464.1\|, |
| | gi\|1044554745\|ref\|XP_017413481.1\|, gi\|1044555047\|ref\|XP_017413644.1\|, |
| | gi\|1044555165\|ref\|XP_017413705.1\|, gi\|1044556039\|ref\|XP_017414166.1\|, |
| | gi\|1044556130\|ref\|XP_017414212.1\|, gi\|1044556134\|ref\|XP_017414214.1\|, |
| | gi\|1044556136\|ref\|XP_017414215.1\|, gi\|1044556138\|ref\|XP_017414216.1\|, |
| | gi\|1044556140\|ref\|XP_017414217.1\|, gi\|1044556184\|ref\|XP_017414239.1\|, |
| | gi\|1044556221\|ref\|XP_017414259.1\|, gi\|1044557026\|ref\|XP_017414655.1\|, |
| | gi\|1044557099\|ref\|XP_017414693.1\|, gi\|1044557256\|ref\|XP_017414775.1\|, |
| | gi\|1044557515\|ref\|XP_017414910.1\|, gi\|1044557517\|ref\|XP_017414911.1\|, |
| | gi\|1044557767\|ref\|XP_017415041.1\|, gi\|1044557804\|ref\|XP_017415061.1\|, |
| | gi\|1044557841\|ref\|XP_017415081.1\|, gi\|1044558008\|ref\|XP_017415167.1\|, |
| | gi\|1044558398\|ref\|XP_017415367.1\|, gi\|1044558610\|ref\|XP_017415476.1\|, |
| | gi\|1044558766\|ref\|XP_017415554.1\|, gi\|1044558845\|ref\|XP_017415595.1\|, |
| | gi\|1044558928\|ref\|XP_017415630.1\|, gi\|1044558979\|ref\|XP_017415656.1\|, |
| | gi\|1044558981\|ref\|XP_017415657.1\|, gi\|1044558983\|ref\|XP_017415658.1\|, |
| | gi\|1044559133\|ref\|XP_017415737.1\|, gi\|1044559438\|ref\|XP_017415894.1\|, |
| | gi\|1044559542\|ref\|XP_017415950.1\|, gi\|1044559561\|ref\|XP_017415959.1\|, |
| | gi\|1044559595\|ref\|XP_017415978.1\|, gi\|1044560021\|ref\|XP_017416198.1\|, |
| | gi\|1044560274\|ref\|XP_017416324.1\|, gi\|1044560330\|ref\|XP_017416353.1\|, |
| | gi\|1044560384\|ref\|XP_017416376.1\|, gi\|1044560386\|ref\|XP_017416377.1\|, |
| | gi\|1044560422\|ref\|XP_017416396.1\|, gi\|1044560424\|ref\|XP_017416397.1\|, |
| | gi\|1044560430\|ref\|XP_017416400.1\|, gi\|1044560464\|ref\|XP_017416419.1\|, |
| | gi\|1044560570\|ref\|XP_017416477.1\|, gi\|1044560620\|ref\|XP_017416504.1\|, |
| | gi\|1044560660\|ref\|XP_017416526.1\|, gi\|1044560684\|ref\|XP_017416540.1\|, |
| | gi\|1044560856\|ref\|XP_017416635.1\|, gi\|1044560888\|ref\|XP_017416652.1\|, |
| | gi\|1044560894\|ref\|XP_017416655.1\|, gi\|1044560948\|ref\|XP_017416685.1\|, |
| | gi\|1044561022\|ref\|XP_017416727.1\|, gi\|1044561509\|ref\|XP_017416989.1\|, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VA | gi\|1044561703\|ref\|XP_017417091.1\|, gi\|1044561737\|ref\|XP_017417109.1\|, |
| | gi\|1044561964\|ref\|XP_017417225.1\|, gi\|1044562541\|ref\|XP_017417525.1\|, |
| | gi\|1044562651\|ref\|XP_017417580.1\|, gi\|1044562695\|ref\|XP_017417603.1\|, |
| | gi\|1044562737\|ref\|XP_017417625.1\|, gi\|1044562830\|ref\|XP_017417673.1\|, |
| | gi\|1044562976\|ref\|XP_017417746.1\|, gi\|1044563305\|ref\|XP_017417918.1\|, |
| | gi\|1044563596\|ref\|XP_017418070.1\|, gi\|1044563863\|ref\|XP_017418204.1\|, |
| | gi\|1044563865\|ref\|XP_017418205.1\|, gi\|1044563908\|ref\|XP_017418226.1\|, |
| | gi\|1044564215\|ref\|XP_017418381.1\|, gi\|1044564217\|ref\|XP_017418382.1\|, |
| | gi\|1044564254\|ref\|XP_017418400.1\|, gi\|1044564297\|ref\|XP_017418424.1\|, |
| | gi\|1044564299\|ref\|XP_017418425.1\|, gi\|1044564309\|ref\|XP_017418431.1\|, |
| | gi\|1044564606\|ref\|XP_017418583.1\|, gi\|1044564678\|ref\|XP_017418619.1\|, |
| | gi\|1044565098\|ref\|XP_017418834.1\|, gi\|1044565264\|ref\|XP_017418915.1\|, |
| | gi\|1044565266\|ref\|XP_017418916.1\|, gi\|1044565400\|ref\|XP_017418989.1\|, |
| | gi\|1044565543\|ref\|XP_017419062.1\|, gi\|1044565601\|ref\|XP_017419095.1\|, |
| | gi\|1044565603\|ref\|XP_017419096.1\|, gi\|1044565700\|ref\|XP_017419146.1\|, |
| | gi\|1044565748\|ref\|XP_017419173.1\|, gi\|1044565773\|ref\|XP_017419186.1\|, |
| | gi\|1044565815\|ref\|XP_017419209.1\|, gi\|1044565837\|ref\|XP_017419221.1\|, |
| | gi\|1044565839\|ref\|XP_017419222.1\|, gi\|1044565841\|ref\|XP_017419223.1\|, |
| | gi\|1044565937\|ref\|XP_017419273.1\|, gi\|1044566385\|ref\|XP_017419500.1\|, |
| | gi\|1044566446\|ref\|XP_017419533.1\|, gi\|1044566448\|ref\|XP_017419534.1\|, |
| | gi\|1044566481\|ref\|XP_017419551.1\|, gi\|1044566483\|ref\|XP_017419552.1\|, |
| | gi\|1044566571\|ref\|XP_017419599.1\|, gi\|1044567017\|ref\|XP_017419838.1\|, |
| | gi\|1044567241\|ref\|XP_017419954.1\|, gi\|1044567393\|ref\|XP_017420033.1\|, |
| | gi\|1044567395\|ref\|XP_017420034.1\|, gi\|1044567459\|ref\|XP_017420066.1\|, |
| | gi\|1044567463\|ref\|XP_017420069.1\|, gi\|1044567465\|ref\|XP_017420070.1\|, |
| | gi\|1044567519\|ref\|XP_017420097.1\|, gi\|1044567587\|ref\|XP_017420133.1\|, |
| | gi\|1044567621\|ref\|XP_017420152.1\|, gi\|1044567711\|ref\|XP_017420204.1\|, |
| | gi\|1044567801\|ref\|XP_017420257.1\|, gi\|1044568009\|ref\|XP_017420387.1\|, |
| | gi\|1044568065\|ref\|XP_017420420.1\|, gi\|1044568117\|ref\|XP_017420452.1\|, |
| | gi\|1044568227\|ref\|XP_017420514.1\|,gi\|1044568782\|ref\|XP_017420815.1\|, |
| | gi\|1044568839\|ref\|XP_017420845.1\|, gi\|1044568841\|ref\|XP_017420846.1\|, |
| | gi\|1044569119\|ref\|XP_017420997.1\|, gi\|1044569229\|ref\|XP_017421053.1\|, |
| | gi\|1044569326\|ref\|XP_017421107.1\|, gi\|1044569460\|ref\|XP_017421180.1\|, |
| | gi\|1044569992\|ref\|XP_017421471.1\|, gi\|1044570046\|ref\|XP_017421496.1\|, |
| | gi\|1044570139\|ref\|XP_017421542.1\|, gi\|1044570141\|ref\|XP_017421543.1\|, |
| | gi\|1044570153\|ref\|XP_017421548.1\|, gi\|1044570155\|ref\|XP_017421549.1\|, |
| | gi\|1044570157\|ref\|XP_017421550.1\|, gi\|1044570159\|ref\|XP_017421551.1\|, |
| | gi\|1044570161\|ref\|XP_017421552.1\|, gi\|1044570163\|ref\|XP_017421553.1\|, |
| | gi\|1044570165\|ref\|XP_017421554.1\|, gi\|1044570167\|ref\|XP_017421555.1\|, |
| | gi\|1044570169\|ref\|XP_017421557.1\|, gi\|1044570751\|ref\|XP_017421871.1\|, |
| | gi\|1044570799\|ref\|XP_017421897.1\|, gi\|1044570825\|ref\|XP_017421912.1\|, |
| | gi\|1044570827\|ref\|XP_017421913.1\|, gi\|1044570829\|ref\|XP_017421915.1\|, |
| | gi\|1044571053\|ref\|XP_017422030.1\|, gi\|1044571289\|ref\|XP_017422151.1\|, |
| | gi\|1044571610\|ref\|XP_017422320.1\|, gi\|1044571754\|ref\|XP_017422396.1\|, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044572005|ref|XP_017422531.1|, gi|1044572161|ref|XP_017422615.1|, gi|1044572173|ref|XP_017422622.1|, gi|1044572175|ref|XP_017422623.1|, gi|1044572542|ref|XP_017422818.1|, gi|1044572639|ref|XP_017422871.1|, gi|1044572751|ref|XP_017422934.1|, gi|1044572753|ref|XP_017422935.1|, gi|1044572755|ref|XP_017422936.1|, gi|1044572899|ref|XP_017423011.1|, gi|1044572901|ref|XP_017423012.1|, gi|1044572903|ref|XP_017423013.1|, gi|1044572905|ref|XP_017423014.1|, gi|1044572907|ref|XP_017423016.1|, gi|1044572909|ref|XP_017423017.1|, gi|1044572911|ref|XP_017423018.1|, gi|1044572929|ref|XP_017423027.1|, gi|1044572981|ref|XP_017423057.1|, gi|1044572983|ref|XP_017423058.1|, gi|1044572985|ref|XP_017423060.1|, gi|1044573050|ref|XP_017423094.1|, gi|1044573257|ref|XP_017423206.1|, gi|1044573350|ref|XP_017423256.1|, gi|1044573352|ref|XP_017423258.1|, gi|1044573503|ref|XP_017423333.1|, gi|1044573505|ref|XP_017423334.1|, gi|1044573757|ref|XP_017423477.1|, gi|1044573935|ref|XP_017423588.1|, gi|1044574213|ref|XP_017423751.1|, gi|1044574247|ref|XP_017423770.1|, gi|1044574251|ref|XP_017423772.1|, gi|1044574259|ref|XP_017423777.1|, gi|1044574273|ref|XP_017423784.1|, gi|1044574275|ref|XP_017423786.1|, gi|1044574277|ref|XP_017423787.1|, gi|1044574279|ref|XP_017423788.1|, gi|1044574281|ref|XP_017423789.1|, gi|1044574283|ref|XP_017423790.1|, gi|1044574285|ref|XP_017423791.1|, gi|1044574293|ref|XP_017423796.1|, gi|1044574295|ref|XP_017423797.1|, gi|1044574653|ref|XP_017423991.1|, gi|1044574846|ref|XP_017424095.1|, gi|1044574848|ref|XP_017424096.1|, gi|1044574850|ref|XP_017424097.1|, gi|1044574852|ref|XP_017424099.1|, gi|1044574854|ref|XP_017424100.1|, gi|1044575209|ref|XP_017424292.1|, gi|1044575211|ref|XP_017424293.1|, gi|1044575213|ref|XP_017424294.1|, gi|1044575215|ref|XP_017424295.1|, gi|1044575217|ref|XP_017424296.1|, gi|1044575219|ref|XP_017424297.1|, gi|1044575221|ref|XP_017424298.1|, gi|1044575581|ref|XP_017424491.1|, gi|1044575680|ref|XP_017424544.1|, gi|1044576110|ref|XP_017424767.1|, gi|1044576953|ref|XP_017425215.1|, gi|1044577070|ref|XP_017425277.1|, gi|1044577106|ref|XP_017425297.1|, gi|1044577304|ref|XP_017425403.1|, gi|1044577409|ref|XP_017425457.1|, gi|1044577494|ref|XP_017425504.1|, gi|1044577549|ref|XP_017425533.1|, gi|1044577559|ref|XP_017425539.1|, gi|1044577645|ref|XP_017425584.1|, gi|1044577744|ref|XP_017425635.1|, gi|1044578056|ref|XP_017425803.1|, gi|1044578058|ref|XP_017425804.1|, gi|1044578060|ref|XP_017425805.1|, gi|1044578502|ref|XP_017426048.1|, gi|1044578688|ref|XP_017426149.1|, gi|1044578690|ref|XP_017426150.1|, gi|1044578708|ref|XP_017426160.1|, gi|1044578902|ref|XP_017426275.1|, gi|1044579004|ref|XP_017426335.1|, gi|1044579006|ref|XP_017426336.1|, gi|1044579756|ref|XP_017426749.1|, gi|1044580489|ref|XP_017427135.1|, gi|1044580491|ref|XP_017427136.1|, gi|1044580493|ref|XP_017427137.1|, gi|1044580495|ref|XP_017427138.1|, gi|1044580642|ref|XP_017427218.1|, gi|1044580904|ref|XP_017427357.1|, gi|1044580938|ref|XP_017427375.1|, gi|1044581144|ref|XP_017427482.1|, gi|1044581148|ref|XP_017427484.1|, gi|1044581250|ref|XP_017427540.1|,</td></tr>
</table>

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044581252|ref|XP_017427541.1|, gi|1044581254|ref|XP_017427542.1|, gi|1044581302|ref|XP_017427568.1|, gi|1044581502|ref|XP_017427678.1|, gi|1044581504|ref|XP_017427679.1|, gi|1044581528|ref|XP_017427692.1|, gi|1044581578|ref|XP_017427720.1|, gi|1044581580|ref|XP_017427722.1|, gi|1044581582|ref|XP_017427723.1|, gi|1044581678|ref|XP_017427776.1|, gi|1044581878|ref|XP_017427886.1|, gi|1044581927|ref|XP_017427910.1|, gi|1044581992|ref|XP_017427940.1|, gi|1044582033|ref|XP_017427963.1|, gi|1044582035|ref|XP_017427964.1|, gi|1044582216|ref|XP_017428063.1|, gi|1044582549|ref|XP_017428233.1|, gi|1044582562|ref|XP_017428239.1|, gi|1044582564|ref|XP_017428240.1|, gi|1044582566|ref|XP_017428241.1|, gi|1044582568|ref|XP_017428242.1|, gi|1044582576|ref|XP_017428247.1|, gi|1044582659|ref|XP_017428292.1|, gi|1044582695|ref|XP_017428312.1|, gi|1044582761|ref|XP_017428346.1|, gi|1044582884|ref|XP_017428408.1|, gi|1044582894|ref|XP_017428413.1|, gi|1044582902|ref|XP_017428418.1|, gi|1044582904|ref|XP_017428419.1|, gi|1044582906|ref|XP_017428420.1|, gi|1044582908|ref|XP_017428421.1|, gi|1044583111|ref|XP_017428538.1|, gi|1044583371|ref|XP_017428693.1|, gi|1044583531|ref|XP_017428794.1|, gi|1044583537|ref|XP_017428797.1|, gi|1044583607|ref|XP_017428842.1|, gi|1044583659|ref|XP_017428872.1|, gi|1044583735|ref|XP_017428917.1|, gi|1044583737|ref|XP_017428918.1|, gi|1044583739|ref|XP_017428919.1|, gi|1044583741|ref|XP_017428920.1|, gi|1044584066|ref|XP_017429099.1|, gi|1044584133|ref|XP_017429136.1|, gi|1044584358|ref|XP_017429253.1|, gi|1044584374|ref|XP_017429261.1|, gi|1044584382|ref|XP_017429267.1|, gi|1044584384|ref|XP_017429268.1|, gi|1044584400|ref|XP_017429276.1|, gi|1044584418|ref|XP_017429285.1|, gi|1044584420|ref|XP_017429286.1|, gi|1044584677|ref|XP_017429425.1|, gi|1044584717|ref|XP_017429447.1|, gi|1044584719|ref|XP_017429448.1|, gi|1044584837|ref|XP_017429508.1|, gi|1044584996|ref|XP_017429595.1|, gi|1044585210|ref|XP_017429711.1|, gi|1044585323|ref|XP_017429772.1|, gi|1044585362|ref|XP_017429794.1|, gi|1044585556|ref|XP_017429898.1|, gi|1044585613|ref|XP_017429929.1|, gi|1044585635|ref|XP_017429940.1|, gi|1044585637|ref|XP_017429941.1|, gi|1044585808|ref|XP_017430023.1|, gi|1044585922|ref|XP_017430081.1|, gi|1044585924|ref|XP_017430082.1|, gi|1044585926|ref|XP_017430083.1|, gi|1044585928|ref|XP_017430084.1|, gi|1044585930|ref|XP_017430085.1|, gi|1044586018|ref|XP_017430131.1|, gi|1044586124|ref|XP_017430187.1|, gi|1044586437|ref|XP_017430354.1|, gi|1044586439|ref|XP_017430355.1|, gi|1044586465|ref|XP_017430370.1|, gi|1044586495|ref|XP_017430386.1|, gi|1044586597|ref|XP_017430441.1|, gi|1044586754|ref|XP_017430525.1|, gi|1044586845|ref|XP_017430571.1|, gi|1044586847|ref|XP_017430572.1|, gi|1044586849|ref|XP_017430573.1|, gi|1044586857|ref|XP_017430577.1|, gi|1044586906|ref|XP_017430600.1|, gi|1044587409|ref|XP_017430871.1|, gi|1044587642|ref|XP_017430997.1|,gi|1044587738|ref|XP_017431047.1|, gi|1044587874|ref|XP_017431121.1|, gi|1044587966|ref|XP_017431169.1|, gi|1044588012|ref|XP_017431195.1|, gi|1044588148|ref|XP_017431268.1|,</td></tr>
</table>

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044588337|ref|XP_017431373.1|, gi|1044588341|ref|XP_017431376.1|,</td></tr>
<tr><td></td><td>gi|1044588430|ref|XP_017431422.1|, gi|1044588432|ref|XP_017431423.1|,</td></tr>
<tr><td></td><td>gi|1044588553|ref|XP_017431484.1|, gi|1044588555|ref|XP_017431485.1|,</td></tr>
<tr><td></td><td>gi|1044588557|ref|XP_017431486.1|, gi|1044588569|ref|XP_017431494.1|,</td></tr>
<tr><td></td><td>gi|1044588614|ref|XP_017431518.1|, gi|1044588662|ref|XP_017431545.1|,</td></tr>
<tr><td></td><td>gi|1044588887|ref|XP_017431666.1|, gi|1044588939|ref|XP_017431695.1|,</td></tr>
<tr><td></td><td>gi|1044589065|ref|XP_017431769.1|, gi|1044589069|ref|XP_017431771.1|,</td></tr>
<tr><td></td><td>gi|1044589169|ref|XP_017431831.1|, gi|1044589345|ref|XP_017431941.1|,</td></tr>
<tr><td></td><td>gi|1044589395|ref|XP_017431973.1|, gi|1044589641|ref|XP_017432119.1|,</td></tr>
<tr><td></td><td>gi|1044589983|ref|XP_017432305.1|, gi|1044590457|ref|XP_017432560.1|,</td></tr>
<tr><td></td><td>gi|1044590515|ref|XP_017432592.1|, gi|1044590615|ref|XP_017432645.1|,</td></tr>
<tr><td></td><td>gi|1044590688|ref|XP_017432684.1|, gi|1044590698|ref|XP_017432690.1|,</td></tr>
<tr><td></td><td>gi|1044590728|ref|XP_017432703.1|, gi|1044590730|ref|XP_017432704.1|,</td></tr>
<tr><td></td><td>gi|1044591155|ref|XP_017432918.1|, gi|1044591486|ref|XP_017433091.1|,</td></tr>
<tr><td></td><td>gi|1044591488|ref|XP_017433092.1|, gi|1044591490|ref|XP_017433093.1|,</td></tr>
<tr><td></td><td>gi|1044591492|ref|XP_017433094.1|, gi|1044591522|ref|XP_017433110.1|,</td></tr>
<tr><td></td><td>gi|1044591730|ref|XP_017433221.1|, gi|1044591732|ref|XP_017433222.1|,</td></tr>
<tr><td></td><td>gi|1044591734|ref|XP_017433223.1|, gi|1044591736|ref|XP_017433224.1|,</td></tr>
<tr><td></td><td>gi|1044591738|ref|XP_017433225.1|, gi|1044591740|ref|XP_017433226.1|,</td></tr>
<tr><td></td><td>gi|1044591742|ref|XP_017433227.1|, gi|1044591744|ref|XP_017433228.1|,</td></tr>
<tr><td></td><td>gi|1044591746|ref|XP_017433229.1|, gi|1044591812|ref|XP_017433264.1|,</td></tr>
<tr><td></td><td>gi|1044592180|ref|XP_017433464.1|, gi|1044592212|ref|XP_017433481.1|,</td></tr>
<tr><td></td><td>gi|1044592214|ref|XP_017433482.1|, gi|1044592364|ref|XP_017433557.1|,</td></tr>
<tr><td></td><td>gi|1044592366|ref|XP_017433558.1|, gi|1044592488|ref|XP_017433622.1|,</td></tr>
<tr><td></td><td>gi|1044592732|ref|XP_017433752.1|, gi|1044592977|ref|XP_017433882.1|,</td></tr>
<tr><td></td><td>gi|1044592979|ref|XP_017433883.1|, gi|1044593112|ref|XP_017433957.1|,</td></tr>
<tr><td></td><td>gi|1044593264|ref|XP_017434039.1|, gi|1044593331|ref|XP_017434074.1|,</td></tr>
<tr><td></td><td>gi|1044593422|ref|XP_017434124.1|, gi|1044593722|ref|XP_017434281.1|,</td></tr>
<tr><td></td><td>gi|1044593790|ref|XP_017434316.1|, gi|1044593862|ref|XP_017434356.1|,</td></tr>
<tr><td></td><td>gi|1044593864|ref|XP_017434357.1|, gi|1044593926|ref|XP_017434389.1|,</td></tr>
<tr><td></td><td>gi|1044593928|ref|XP_017434390.1|, gi|1044593930|ref|XP_017434391.1|,</td></tr>
<tr><td></td><td>gi|1044593967|ref|XP_017434410.1|, gi|1044593971|ref|XP_017434413.1|,</td></tr>
<tr><td></td><td>gi|1044593995|ref|XP_017434425.1|, gi|1044594002|ref|XP_017434427.1|,</td></tr>
<tr><td></td><td>gi|1044594046|ref|XP_017434454.1|, gi|1044594231|ref|XP_017434561.1|,</td></tr>
<tr><td></td><td>gi|1044594299|ref|XP_017434601.1|, gi|1044594325|ref|XP_017434618.1|,</td></tr>
<tr><td></td><td>gi|1044594457|ref|XP_017434696.1|, gi|1044594545|ref|XP_017434744.1|,</td></tr>
<tr><td></td><td>gi|1044594553|ref|XP_017434748.1|, gi|1044594573|ref|XP_017434759.1|,</td></tr>
<tr><td></td><td>gi|1044594575|ref|XP_017434760.1|, gi|1044594625|ref|XP_017434787.1|,</td></tr>
<tr><td></td><td>gi|1044594717|ref|XP_017434838.1|, gi|1044594729|ref|XP_017434845.1|,</td></tr>
<tr><td></td><td>gi|1044594731|ref|XP_017434846.1|, gi|1044594751|ref|XP_017434856.1|,</td></tr>
<tr><td></td><td>gi|1044595218|ref|XP_017435106.1|, gi|1044595430|ref|XP_017435216.1|,</td></tr>
<tr><td></td><td>gi|1044595514|ref|XP_017435258.1|, gi|1044595919|ref|XP_017435474.1|,</td></tr>
<tr><td></td><td>gi|1044596115|ref|XP_017435576.1|, gi|1044596248|ref|XP_017435649.1|,</td></tr>
<tr><td></td><td>gi|1044596250|ref|XP_017435650.1|, gi|1044596611|ref|XP_017435841.1|,</td></tr>
</table>

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044596640|ref|XP_017435856.1|, gi|1044596882|ref|XP_017435984.1|, gi|1044596909|ref|XP_017435999.1|, gi|1044596958|ref|XP_017436022.1|, gi|1044597019|ref|XP_017436053.1|, gi|1044597545|ref|XP_017436336.1|, gi|1044597547|ref|XP_017436337.1|, gi|1044597807|ref|XP_017436477.1|, gi|1044598252|ref|XP_017436718.1|, gi|1044598254|ref|XP_017436719.1|, gi|1044598256|ref|XP_017436720.1|, gi|1044598258|ref|XP_017436721.1|, gi|1044598314|ref|XP_017436751.1|, gi|1044598316|ref|XP_017436752.1|, gi|1044598318|ref|XP_017436753.1|, gi|1044598384|ref|XP_017436790.1|, gi|1044598420|ref|XP_017436810.1|, gi|1044598422|ref|XP_017436811.1|, gi|1044598707|ref|XP_017436967.1|, gi|1044598709|ref|XP_017436968.1|, gi|1044598711|ref|XP_017436969.1|, gi|1044598875|ref|XP_017437055.1|, gi|1044598877|ref|XP_017437057.1|, gi|1044598879|ref|XP_017437058.1|, gi|1044599112|ref|XP_017437186.1|, gi|1044599156|ref|XP_017437210.1|, gi|1044599160|ref|XP_017437213.1|, gi|1044599162|ref|XP_017437214.1|, gi|1044599164|ref|XP_017437215.1|, gi|1044599166|ref|XP_017437216.1|, gi|1044599168|ref|XP_017437217.1|, gi|1044599170|ref|XP_017437218.1|, gi|1044599172|ref|XP_017437219.1|, gi|1044599245|ref|XP_017437259.1|, gi|1044599247|ref|XP_017437260.1|, gi|1044599249|ref|XP_017437262.1|, gi|1044599251|ref|XP_017437263.1|, gi|1044599318|ref|XP_017437297.1|, gi|1044599320|ref|XP_017437298.1|, gi|1044599322|ref|XP_017437299.1|, gi|1044599324|ref|XP_017437300.1|</td></tr>
<tr><td>VR</td><td>XP_014489574.1, XP_014489674.1, XP_014489980.1, XP_014489981.1, XP_014489982.1, XP_014489988.1, XP_014489991.1, XP_014490021.1, XP_014490022.1, XP_014490023.1, XP_014490024.1, XP_014490068.1, XP_014490069.1, XP_014490154.1, XP_014490321.1, XP_014490324.1, XP_014490413.1, XP_014490414.1, XP_014490499.1, XP_014490511.1, XP_014490561.1, XP_014490603.1, XP_014490612.1, XP_014490630.1, XP_014490660.1, XP_014490681.1, XP_014490717.1, XP_014490733.1, XP_014490854.1, XP_014490997.1, XP_014491057.1, XP_014491088.1, XP_014491307.1, XP_014491327.1, XP_014491328.1, XP_014491329.1, XP_014491330.1, XP_014491331.1, XP_014491332.1, XP_014491383.1, XP_014491451.1, XP_014491668.1, XP_014491686.1, XP_014491760.1, XP_014491768.1, XP_014491809.1, XP_014491813.1, XP_014491824.1, XP_014491825.1, XP_014491829.1, XP_014491834.1, XP_014491835.1, XP_014491836.1, XP_014491837.1, XP_014491889.1, XP_014492060.1, XP_014492070.1, XP_014492079.1, XP_014492134.1, XP_014492136.1, XP_014492145.1, XP_014492314.1, XP_014492350.1, XP_014492351.1, XP_014492352.1, XP_014492369.1, XP_014492434.1, XP_014492544.1, XP_014492624.1, XP_014492683.1, XP_014492769.1, XP_014492785.1, XP_014492790.1, XP_014492796.1, XP_014492813.1, XP_014493008.1, XP_014493063.1, XP_014493190.1, XP_014493201.1, XP_014493209.1, XP_014493485.1, XP_014493557.1, XP_014493652.1, XP_014493859.1, XP_014493900.1, XP_014493925.1, XP_014493982.1, XP_014493983.1, XP_014493985.1, XP_014493986.1, XP_014494005.1, XP_014494023.1,</td></tr>
</table>

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VR | XP_014494024.1, XP_014494029.1, XP_014494030.1, XP_014494031.1, XP_014494044.1, XP_014494070.1, XP_014494072.1, XP_014494164.1, XP_014494245.1, XP_014494297.1, XP_014494299.1, XP_014494300.1, XP_014494346.1, XP_014494347.1, XP_014494348.1, XP_014494349.1, XP_014494350.1, XP_014494384.1, XP_014494385.1, XP_014494431.1, XP_014494509.1, XP_014494521.1, XP_014494552.1, XP_014494553.1, XP_014494587.1, XP_014494610.1, XP_014494611.1, XP_014494655.1, XP_014494766.1, XP_014494798.1, XP_014494929.1, XP_014494969.1, XP_014495167.1, XP_014495209.1, XP_014495296.1, XP_014495309.1, XP_014495403.1, XP_014495523.1, XP_014495534.1, XP_014495638.1, XP_014495676.1, XP_014495699.1, XP_014495842.1, XP_014495873.1, XP_014495982.1, XP_014495984.1, XP_014496063.1, XP_014496064.1, XP_014496065.1, XP_014496112.1, XP_014496113.1, XP_014496126.1, XP_014496148.1, XP_014496184.1, XP_014496186.1, XP_014496244.1, XP_014496390.1, XP_014496391.1, XP_014496468.1, XP_014496650.1, XP_014496748.1, XP_014496772.1, XP_014496775.1, XP_014496838.1, XP_014496839.1, XP_014496858.1, XP_014496882.1, XP_014496883.1, XP_014496885.1, XP_014496926.1, XP_014496928.1, XP_014496929.1, XP_014496942.1, XP_014496943.1, XP_014497026.1, XP_014497308.1, XP_014497316.1, XP_014497459.1, XP_014497488.1, XP_014497489.1, XP_014497795.1, XP_014497847.1, XP_014497920.1, XP_014497928.1, XP_014497994.1, XP_014498038.1, XP_014498047.1, XP_014498163.1, XP_014498218.1, XP_014498219.1, XP_014498222.1, XP_014498275.1, XP_014498276.1, XP_014498298.1, XP_014498341.1, XP_014498348.1, XP_014498349.1, XP_014498350.1, XP_014498351.1, XP_014498354.1, XP_014498355.1, XP_014498357.1, XP_014498358.1, XP_014498359.1, XP_014498360.1, XP_014498361.1, XP_014498362.1, XP_014498363.1, XP_014498364.1, XP_014498365.1, XP_014498366.1, XP_014498444.1, XP_014498481.1, XP_014498515.1, XP_014498568.1, XP_014498850.1, XP_014498889.1, XP_014498890.1, XP_014498945.1, XP_014498973.1, XP_014498993.1, XP_014499038.1, XP_014499061.1, XP_014499097.1, XP_014499098.1, XP_014499387.1, XP_014499427.1, XP_014499439.1, XP_014499508.1, XP_014499509.1, XP_014499512.1, XP_014499704.1, XP_014500063.1, XP_014500065.1, XP_014500066.1, XP_014500297.1, XP_014500298.1, XP_014500351.1, XP_014500450.1, XP_014500547.1, XP_014500850.1, XP_014501148.1, XP_014501149.1, XP_014501150.1, XP_014501151.1, XP_014501152.1, XP_014501153.1, XP_014501278.1, XP_014501279.1, XP_014501284.1, XP_014501356.1, XP_014501357.1, XP_014501390.1, XP_014501398.1, XP_014501399.1, XP_014501400.1, XP_014501401.1, XP_014501518.1, XP_014501549.1, XP_014501550.1, XP_014501551.1, XP_014501664.1, XP_014501665.1, XP_014501678.1, XP_014501844.1, XP_014501845.1, XP_014501846.1, XP_014501847.1, XP_014501848.1, XP_014501925.1, XP_014501973.1, , |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDS |
|---|---|
| VR | XP_014502030.1, XP_014502031.1, XP_014502115.1, XP_014502190.1, XP_014502194.1, XP_014502216.1, XP_014502230.1, XP_014502268.1, XP_014502460.1, XP_014502560.1, XP_014502561.1, XP_014502569.1, XP_014502570.1, XP_014502576.1, XP_014502586.1, XP_014502589.1, XP_014502602.1, XP_014502648.1, XP_014502685.1, XP_014502686.1, XP_014502689.1, XP_014502690.1, XP_014502724.1, XP_014502730.1, XP_014502834.1, XP_014502897.1, XP_014502964.1, XP_014503161.1, XP_014503164.1, XP_014503232.1, XP_014503273.1, XP_014503297.1, XP_014503309.1, XP_014503311.1, XP_014503324.1, XP_014503367.1, XP_014503375.1, XP_014503376.1, XP_014503381.1, XP_014503382.1, XP_014503383.1, XP_014503420.1, XP_014503430.1, XP_014503607.1, XP_014503608.1, XP_014503609.1, XP_014503610.1, XP_014503708.1, XP_014503819.1, XP_014503878.1, XP_014503935.1, XP_014503939.1, XP_014504026.1, XP_014504042.1, XP_014504058.1, XP_014504127.1, XP_014504295.1, XP_014504308.1, XP_014504370.1, XP_014504383.1, XP_014504395.1, XP_014504460.1, XP_014504462.1, XP_014504562.1, XP_014504564.1, XP_014504565.1, XP_014504566.1, XP_014504616.1, XP_014504617.1, XP_014504717.1, XP_014504731.1, XP_014504784.1, XP_014504814.1, XP_014504817.1, XP_014504818.1, XP_014504998.1, XP_014504999.1, XP_014505012.1, XP_014505013.1, XP_014505027.1, XP_014505075.1, XP_014505078.1, XP_014505242.1, XP_014505288.1, XP_014505340.1, XP_014505341.1, XP_014505406.1, XP_014505432.1, XP_014505760.1, XP_014505926.1, XP_014505927.1, XP_014505928.1, XP_014506045.1, XP_014506119.1, XP_014506127.1, XP_014506148.1, XP_014506187.1, XP_014506388.1, XP_014506405.1, XP_014506427.1, XP_014506445.1, XP_014506477.1, XP_014506558.1, XP_014506559.1, XP_014506560.1, XP_014506561.1, XP_014506562.1, XP_014506630.1, XP_014506690.1, XP_014506697.1, XP_014506718.1, XP_014506794.1, XP_014506923.1, XP_014506943.1, XP_014506944.1, XP_014507238.1, XP_014507434.1, XP_014507461.1, XP_014507462.1, XP_014507711.1, XP_014507729.1, XP_014507736.1, XP_014508086.1, XP_014508087.1, XP_014508105.1, XP_014508336.1, XP_014508424.1, XP_014508479.1, XP_014508692.1, XP_014508994.1, XP_014508995.1, XP_014508996.1, XP_014508997.1, XP_014509240.1, XP_014509241.1, XP_014509242.1, XP_014509243.1, XP_014509244.1, XP_014509263.1, XP_014509264.1, XP_014509399.1, XP_014509513.1, XP_014509672.1, XP_014509730.1, XP_014509761.1, XP_014509822.1, XP_014510007.1XP_014510056.1, XP_014510057.1, XP_014510285.1, XP_014510312.1, XP_014510313.1, XP_014510412.1, XP_014510488.1, XP_014510522.1, XP_014510523.1, XP_014510544.1, XP_014510583.1, XP_014510590.1, XP_014510976.1, XP_014511080.1, XP_014511239.1, XP_014511280.1, XP_014511302.1, XP_014511368.1, XP_014511372.1, XP_014511397.1, XP_014511419.1, XP_014511421.1, XP_014511423.1, XP_014511467.1, XP_014511529.1, XP_014511585.1, XP_014511595.1, XP_014511779.1, XP_014511843.1, , |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VR | XP_014511923.1, XP_014511951.1, XP_014511952.1, XP_014511953.1, XP_014511954.1, XP_014511955.1, XP_014511956.1, XP_014511959.1, XP_014511963.1, XP_014511990.1, XP_014512249.1, XP_014512292.1, XP_014512301.1, XP_014512474.1, XP_014512499.1, XP_014512699.1, XP_014512738.1, XP_014512745.1, XP_014512753.1, XP_014512792.1, XP_014512822.1, XP_014512966.1, XP_014512982.1, XP_014513045.1, XP_014513122.1, XP_014513304.1, XP_014513305.1, XP_014513306.1, XP_014513309.1, XP_014513323.1, XP_014513324.1, XP_014513327.1, XP_014513340.1, XP_014513341.1, XP_014513342.1, XP_014513343.1, XP_014513345.1, XP_014513376.1, XP_014513440.1, XP_014513507.1, XP_014513586.1, XP_014513640.1, XP_014513641.1, XP_014513729.1, XP_014513732.1, XP_014513743.1, XP_014513752.1, XP_014513792.1, XP_014513793.1, XP_014513794.1, XP_014513795.1, XP_014513912.1, XP_014514115.1, XP_014514166.1, XP_014514218.1, XP_014514515.1, XP_014514537.1, XP_014514619.1, XP_014514744.1, XP_014514798.1, XP_014514800.1, XP_014514813.1, XP_014514832.1, XP_014514845.1, XP_014514846.1, XP_014514948.1, XP_014514963.1, XP_014515274.1, XP_014515299.1, XP_014515487.1, XP_014515497.1, XP_014515582.1, XP_014515640.1, XP_014515854.1, XP_014515869.1, XP_014516100.1, XP_014516329.1, XP_014516330.1, XP_014516383.1, XP_014516384.1, XP_014516413.1, XP_014516516.1, XP_014516635.1, XP_014516636.1, XP_014516637.1, XP_014516638.1, XP_014516639.1, XP_014516640.1, XP_014516641.1, XP_014516667.1, XP_014516669.1, XP_014516728.1, XP_014516729.1, XP_014516751.1, XP_014516848.1, XP_014516974.1, XP_014516995.1, XP_014517209.1, XP_014517257.1, XP_014517302.1, XP_014517330.1, XP_014517331.1, XP_014517367.1, XP_014517386.1, XP_014517403.1, XP_014517408.1, XP_014517486.1, XP_014517487.1, XP_014517492.1, XP_014517512.1, XP_014517513.1, XP_014517515.1, XP_014517516.1, XP_014517532.1, XP_014517655.1, XP_014517671.1, XP_014517733.1, XP_014517744.1, XP_014517803.1, XP_014517967.1, XP_014518002.1, XP_014518199.1, XP_014518253.1, XP_014518355.1, XP_014518356.1, XP_014518384.1, XP_014518385.1, XP_014518386.1, XP_014518388.1, XP_014518488.1, XP_014518542.1, XP_014518543.1, XP_014518544.1, XP_014518632.1, XP_014518633.1, XP_014518634.1, XP_014518635.1, XP_014518637.1, XP_014518743.1, XP_014518770.1, XP_014518809.1, XP_014518919.1, XP_014518920.1, XP_014518988.1, XP_014518989.1, XP_014519007.1, XP_014519057.1, XP_014519324.1, XP_014519361.1, XP_014519398.1, XP_014519429.1, XP_014519431.1, XP_014519527.1, XP_014519539.1, XP_014519619.1, XP_014519669.1, XP_014519854.1, XP_014520139.1, XP_014520260.1, XP_014520474.1, XP_014520481.1, XP_014520724.1, XP_014520785.1, XP_014520889.1, XP_014520896.1, XP_014520899.1, XP_014520976.1, XP_014521050.1, XP_014521051.1, XP_014521056.1, XP_014521113.1, XP_014521139.1, XP_014521155.1, XP_014521156.1, XP_014521193.1, XP_014521227.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VR | XP_014521281.1, XP_014521392.1, XP_014521393.1, XP_014521395.1, XP_014521396.1, XP_014521397.1, XP_014521496.1, XP_014521505.1, XP_014521522.1, XP_014521617.1, XP_014521619.1, XP_014521730.1, XP_014521799.1, XP_014521802.1, XP_014521803.1, XP_014521887.1, XP_014521950.1, XP_014522053.1, XP_014522219.1, XP_014522234.1, XP_014522258.1, XP_014522320.1, XP_014522387.1, XP_014522388.1, XP_014522389.1, XP_014522401.1, XP_014522629.1, XP_014522660.1, XP_014522709.1, XP_014522710.1, XP_014522839.1, XP_014522912.1, XP_014523025.1, XP_014523047.1, XP_014523048.1, XP_014523050.1, XP_014523051.1, XP_014523067.1, XP_014523068.1, XP_014523141.1, XP_014523217.1, XP_014523251.1, XP_014523328.1, XP_014523329.1, XP_014523458.1XP_014523459.1, XP_014523460.1, XP_014523461.1, XP_014523504.1, XP_014523580.1, XP_014523588.1, XP_014523680.1, XP_014523815.1, XP_014524226.1, XP_014524227.1, XP_014524228.1, XP_014524280.1, XP_014524358.1 |
| VU | Vigun01g006300.1.p, Vigun01g006300.2.p, Vigun01g006300.3.p, Vigun01g006300.4.p, Vigun01g014500.1.p, Vigun01g014700.1.p, Vigun01g061900.1.p, Vigun01g067100.1.p, Vigun01g080500.1.p, Vigun01g080500.2.p, Vigun01g080500.3.p, Vigun01g080500.4.p, Vigun01g080700.1.p, Vigun01g091500.1.p, Vigun01g091500.2.p, Vigun01g091500.3.p, Vigun01g091500.4.p, Vigun01g114200.1.p, Vigun01g116700.1.p, Vigun01g119600.1.p, Vigun01g128800.1.p, Vigun01g129000.1.p, Vigun01g145500.1.p, Vigun01g145700.1.p, Vigun01g145800.1.p, Vigun01g147000.1.p, Vigun01g149300.1.p, Vigun01g149300.2.p, Vigun01g156800.1.p, Vigun01g167100.1.p, Vigun01g168300.1.p, Vigun01g179800.1.p, Vigun01g198100.1.p, Vigun01g198100.2.p, Vigun01g198100.3.p, Vigun01g207000.1.p, Vigun01g221200.1.p, Vigun01g228300.1.p, Vigun01g229900.1.p, Vigun01g231500.1.p, Vigun01g237300.1.p, Vigun01g237300.2.p, Vigun01g243500.1.p, Vigun01g250900.1.p, Vigun01g251000.1.p, Vigun02g017100.1.p, Vigun02g017200.1.p, Vigun02g025400.1.p, Vigun02g025400.2.p, Vigun02g025400.3.p, Vigun02g025400.4.p, Vigun02g028600.1.p, Vigun02g028600.2.p, Vigun02g034700.1.p, Vigun02g034900.1.p, Vigun02g037800.1.p, Vigun02g044500.1.p, Vigun02g044600.1.p, Vigun02g044700.1.p, Vigun02g044800.1.p, Vigun02g046500.1.p, Vigun02g067400.1.p, Vigun02g077500.1.p, Vigun02g077500.2.p, Vigun02g077500.3.p, Vigun02g079200.1.p, Vigun02g079700.1.p, Vigun02g079700.2.p, Vigun02g079700.3.p, Vigun02g079700.4.p, Vigun02g080500.1.p, Vigun02g084800.1.p, Vigun02g085000.1.p, Vigun02g085100.1.p, Vigun02g085200.1.p, Vigun02g085300.1.p, Vigun02g088500.1.p, Vigun02g103200.1.p, Vigun02g104600.1.p, Vigun02g105200.1.p, Vigun02g105200.2.p, Vigun02g105900.1.p, Vigun02g106700.1.p, Vigun02g107100.1.p, Vigun02g110500.1.p, Vigun02g112400.1.p, Vigun02g124500.1.p, Vigun02g132500.1.p, Vigun02g136300.1.p, Vigun02g136400.1.p, Vigun02g138500.1.p, Vigun02g138500.2.p, Vigun02g144800.1.p, Vigun02g154800.1.p, Vigun02g162600.1.p, Vigun02g162700.1.p, Vigun02g169300.1.p, Vigun02g170800.1.p, Vigun02g180700.1.p, Vigun02g199100.1.p, Vigun02g199200.1.p, Vigun02g199300.1.p, Vigun03g002100.1.p, Vigun03g008700.1.p, Vigun03g015300.1.p, Vigun03g015400.1.p, Vigun03g015400.2.p, Vigun03g031700.1.p, Vigun03g038500.1.p, Vigun03g038500.2.p, Vigun03g071100.1.p, Vigun03g077300.1.p, Vigun03g077300.2.p, Vigun03g077300.3.p, Vigun03g077300.4.p, Vigun03g084000.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VU | Vigun03g084100.1.p, Vigun03g086100.1.p, Vigun03g094400.1.p, Vigun03g097900.1.p, Vigun03g098300.1.p, Vigun03g098300.2.p, Vigun03g098300.3.p, Vigun03g098300.4.p, Vigun03g103600.1.p, Vigun03g103700.1.p, Vigun03g112900.1.p, Vigun03g118600.1.p, Vigun03g118600.2.p, Vigun03g118600.3.p, Vigun03g120300.1.p, Vigun03g123100.1.p, Vigun03g123200.1.p, Vigun03g123300.1.p, Vigun03g123400.1.p, Vigun03g123500.1.p, Vigun03g123600.1.p, Vigun03g133400.1.p, Vigun03g133600.1.p, Vigun03g134200.1.p, Vigun03g145600.1.p, Vigun03g154300.1.p, Vigun03g158000.1.p, Vigun03g160100.1.p, Vigun03g171700.1.p, Vigun03g188100.1.p, Vigun03g211800.1.p, Vigun03g221300.1.p, Vigun03g221300.2.p, Vigun03g227700.1.p, Vigun03g232900.1.p, Vigun03g234000.1.p, Vigun03g234000.2.p, Vigun03g234000.3.p, Vigun03g234000.4.p, Vigun03g234000.5.p, Vigun03g234000.6.p, Vigun03g234100.1.p, Vigun03g234100.2.p, Vigun03g234100.3.p, Vigun03g234100.4.p, Vigun03g236500.1.p, Vigun03g236500.2.p, Vigun03g237300.1.p, Vigun03g245500.1.p, Vigun03g245500.2.p, Vigun03g245500.3.p, Vigun03g248800.1.p, Vigun03g248800.2.p, Vigun03g254400.1.p, Vigun03g260600.1.p, Vigun03g260600.2.p, Vigun03g267700.1.p, Vigun03g271100.1.p, Vigun03g271100.2.p, Vigun03g271100.3.p, Vigun03g271100.4.p, Vigun03g280400.1.p, Vigun03g281500.1.p, Vigun03g282400.1.p, Vigun03g282500.1.p, Vigun03g282600.1.p, Vigun03g291100.1.p, Vigun03g291200.1.p, Vigun03g308600.1.p, Vigun03g308600.2.p, Vigun03g320500.1.p, Vigun03g321800.1.p, Vigun03g322800.1.p, Vigun03g332600.1.p, Vigun03g338200.1.p, Vigun03g350500.1.p, Vigun03g352400.1.p, Vigun03g368900.1.p, Vigun03g368900.2.p, Vigun03g368900.3.p, Vigun03g368900.4.p, Vigun03g369000.1.p, Vigun03g371200.1.p, Vigun03g373900.1.p, Vigun03g375300.1.p, Vigun03g375300.2.p, Vigun03g390800.1.p, Vigun03g390800.2.p, Vigun03g399300.1.p, Vigun03g407800.1.p, Vigun03g435500.1.p, Vigun04g001000.1.p, Vigun04g007200.1.p, Vigun04g007300.1.p, Vigun04g007400.1.p, Vigun04g014700.1.p, Vigun04g024400.1.p, Vigun04g024500.1.p, Vigun04g024700.1.p, Vigun04g024700.2.p, Vigun04g024900.1.p, Vigun04g024900.2.p, Vigun04g025000.1.p, Vigun04g025100.1.p, Vigun04g033300.1.p, Vigun04g047000.1.p, Vigun04g060200.1.p, Vigun04g060300.1.p, Vigun04g061600.1.p, Vigun04g062100.1.p, Vigun04g062300.1.p, Vigun04g062400.1.p, Vigun04g062600.1.p, Vigun04g062700.1.p, Vigun04g065600.1.p, Vigun04g068600.1.p, Vigun04g082200.1.p, Vigun04g094400.1.p, Vigun04g094400.2.p, Vigun04g094400.3.p, Vigun04g094400.4.p, Vigun04g094400.5.p, Vigun04g115900.1.p, Vigun04g119400.1.p, Vigun04g119400.2.p, Vigun04g149100.1.p, Vigun04g150000.1.p, Vigun04g164700.1.p, Vigun04g165100.1.p, Vigun04g165100.2.p, Vigun04g165100.3.p, Vigun04g165100.4.p, Vigun04g165200.1.p, Vigun04g165900.1.p, Vigun04g167500.1.p, Vigun04g174400.1.p, Vigun04g174800.1.p, Vigun04g195500.1.p, Vigun05g008200.1.p, Vigun05g008500.1.p, Vigun05g008600.1.p, Vigun05g018700.1.p, Vigun05g020600.1.p, Vigun05g020600.2.p, Vigun05g024300.1.p, Vigun05g026800.1.p, Vigun05g030800.1.p, Vigun05g030900.1.p, Vigun05g031300.1.p, Vigun05g045900.1.p, Vigun05g046000.1.p, Vigun05g048600.1.p, Vigun05g048800.1.p, Vigun05g048800.2.p, Vigun05g048800.3.p, Vigun05g049100.1.p, Vigun05g052100.1.p, Vigun05g052100.2.p, Vigun05g058600.1.p, Vigun05g058700.1.p, Vigun05g060900.1.p, Vigun05g060900.2.p, Vigun05g061600.1.p, Vigun05g081100.1.p, Vigun05g081100.2.p, Vigun05g081200.1.p, Vigun05g081200.2.p, Vigun05g081200.3.p, Vigun05g081500.1.p, Vigun05g083600.1.p, Vigun05g083600.2.p, Vigun05g083600.3.p, Vigun05g083600.4.p, Vigun05g083600.5.p, Vigun05g083600.6.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VU | Vigun05g083800.1.p, Vigun05g083800.2.p, Vigun05g084300.1.p, Vigun05g084300.2.p, Vigun05g084700.1.p, Vigun05g085900.1.p, Vigun05g086000.1.p, Vigun05g095200.1.p, Vigun05g095200.2.p, Vigun05g095400.1.p, Vigun05g096900.1.p, Vigun05g102800.1.p, Vigun05g103400.1.p, Vigun05g104300.1.p, Vigun05g104400.1.p, Vigun05g104500.1.p, Vigun05g114100.1.p, Vigun05g114100.2.p, Vigun05g114100.3.p, Vigun05g116100.1.p, Vigun05g116200.1.p, Vigun05g119600.1.p, Vigun05g125400.1.p, Vigun05g153100.1.p, Vigun05g155700.1.p, Vigun05g174500.1.p, Vigun05g174500.2.p, Vigun05g179600.1.p, Vigun05g179800.1.p, Vigun05g179800.2.p, Vigun05g179800.3.p, Vigun05g179900.1.p, Vigun05g179900.2.p, Vigun05g179900.3.p, Vigun05g180500.1.p, Vigun05g180600.1.p, Vigun05g180600.2.p, Vigun05g183100.1.p, Vigun05g191100.1.p, Vigun05g198200.1.p, Vigun05g202600.1.p, Vigun05g202600.2.p, Vigun05g207500.1.p, Vigun05g212400.1.p, Vigun05g214700.1.p, Vigun05g214800.1.p, Vigun05g218000.1.p, Vigun05g232400.1.p, Vigun05g232400.2.p, Vigun05g247100.1.p, Vigun05g247100.2.p, Vigun05g247100.3.p, Vigun05g247100.4.p, Vigun05g248800.1.p, Vigun05g253900.1.p, Vigun05g258500.1.p, Vigun05g258500.2.p, Vigun05g268300.1.p, Vigun05g275300.1.p, Vigun05g277900.1.p, Vigun05g277900.2.p, Vigun05g278000.1.p, Vigun05g278400.1.p, Vigun05g278500.1.p, Vigun05g278600.1.p, Vigun05g278700.1.p, Vigun05g278800.1.p, Vigun05g279000.1.p, Vigun05g280500.1.p, Vigun05g285500.1.p, Vigun05g294900.1.p, Vigun05g294900.2.p, Vigun05g294900.3.p, Vigun05g294900.4.p, Vigun05g298100.1.p, Vigun05g300400.1.p, Vigun05g300900.1.p, Vigun05g302600.1.p, Vigun05g303900.1.p, Vigun06g009200.1.p, Vigun06g010700.1.p, Vigun06g011000.1.p, Vigun06g030400.1.p, Vigun06g039600.1.p, Vigun06g039600.2.p, Vigun06g064100.1.p, Vigun06g076600.1.p, Vigun06g079700.1.p, Vigun06g079700.2.p, Vigun06g079700.3.p, Vigun06g089100.1.p, Vigun06g089100.2.p, Vigun06g089400.1.p, Vigun06g094100.1.p, Vigun06g094200.1.p, Vigun06g094300.1.p, Vigun06g094400.1.pVigun06g094500.1.p, Vigun06g094600.1.p, Vigun06g094700.1.p, Vigun06g097200.1.p, Vigun06g119200.1.p, Vigun06g119200.2.p, Vigun06g136500.1.p, Vigun06g136500.2.p, Vigun06g136600.1.p, Vigun06g136700.1.p, Vigun06g139600.1.p, Vigun06g139800.1.p, Vigun06g139800.2.p, Vigun06g144600.1.p, Vigun06g144700.1.p, Vigun06g149200.1.p, Vigun06g158000.1.p, Vigun06g158300.1.p, Vigun06g158300.2.p, Vigun06g159800.1.p, Vigun06g185100.1.p, Vigun06g185100.2.p, Vigun06g189200.1.p, Vigun06g189200.2.p, Vigun06g189200.3.p, Vigun06g189500.1.p, Vigun06g190000.1.p, Vigun06g190000.2.p, Vigun06g190000.3.p, Vigun06g190000.4.p, Vigun06g190000.5.p, Vigun06g190000.6.p, Vigun06g195800.1.p, Vigun06g195800.2.p, Vigun06g197400.1.p, Vigun06g197600.1.p, Vigun06g200900.1.p, Vigun06g205900.1.p, Vigun06g209000.1.p, Vigun06g209000.2.p, Vigun06g209000.3.p, Vigun06g209000.4.p, Vigun06g209000.5.p, Vigun06g209000.6.p, Vigun06g213000.1.p, Vigun06g214300.1.p, Vigun06g214300.2.p, Vigun06g216800.1.p, Vigun06g217100.1.p, Vigun06g224000.1.p, Vigun06g225500.1.p, Vigun06g233300.1.p, Vigun07g001200.1.p, Vigun07g014200.1.p, Vigun07g024600.1.p, Vigun07g024700.1.p, Vigun07g024800.1.p, Vigun07g024900.1.p, Vigun07g032800.1.p, Vigun07g033200.1.p, Vigun07g039800.1.p, Vigun07g039900.1.p, Vigun07g040000.1.p, Vigun07g054000.1.p, Vigun07g065700.1.p, Vigun07g068900.1.p, Vigun07g068900.2.p, Vigun07g070000.1.p, Vigun07g079100.1.p, Vigun07g089000.1.p, Vigun07g089000.2.p, Vigun07g090100.1.p, Vigun07g090200.1.p, Vigun07g114300.1.p, Vigun07g126400.1.p, Vigun07g136400.1.p, Vigun07g136400.2.p, Vigun07g148600.1.p, Vigun07g156100.1.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VU | Vigun07g157400.1.p, Vigun07g168500.1.p, Vigun07g185700.1.p, Vigun07g187600.1.p, Vigun07g203100.1.p, Vigun07g209900.1.p, Vigun07g212300.1.p, Vigun07g214100.1.p, Vigun07g218900.1.p, Vigun07g222700.1.p, Vigun07g222700.2.p, Vigun07g222700.3.p, Vigun07g222700.4.p, Vigun07g233500.1.p, Vigun07g233500.2.p, Vigun07g233500.3.p, Vigun07g245400.1.p, Vigun07g245500.1.p, Vigun07g245600.1.p, Vigun07g245700.1.p, Vigun07g245800.1.p, Vigun07g245900.1.p, Vigun07g246000.1.p, Vigun07g246400.1.p, Vigun07g246500.1.p, Vigun07g246600.1.p, Vigun07g246700.1.p, Vigun07g246800.1.p, Vigun07g246900.1.p, Vigun07g247000.1.p, Vigun07g247000.2.p, Vigun07g247000.3.p, Vigun07g247100.1.p, Vigun07g247200.1.p, Vigun07g247500.1.p, Vigun07g247600.1.p, Vigun07g247600.2.p, Vigun07g247700.1.p, Vigun07g248000.1.p, Vigun07g248100.1.p, Vigun07g248200.1.p, Vigun07g248200.2.p, Vigun07g248500.1.p, Vigun07g248600.1.p, Vigun07g256900.1.p, Vigun07g265700.1.p, Vigun07g266500.1.p, Vigun07g266600.1.p, Vigun07g266600.2.p, Vigun07g266700.1.p, Vigun07g266700.2.p, Vigun07g266800.1.p, Vigun07g266900.1.p, Vigun07g267000.1.p, Vigun07g267300.2.p, Vigun07g267400.1.p, Vigun07g267500.1.p, Vigun07g267800.1.p, Vigun07g268000.1.p, Vigun08g015600.1.p, Vigun08g018600.1.p, Vigun08g042700.1.p, Vigun08g042800.1.p, Vigun08g042900.1.p, Vigun08g043000.1.p, Vigun08g043100.1.p, Vigun08g043200.1.p, Vigun08g043300.1.p, Vigun08g043400.1.p, Vigun08g043500.1.p, Vigun08g043600.1.p, Vigun08g043700.1.p, Vigun08g049000.1.p, Vigun08g050600.1.p, Vigun08g051600.1.p, Vigun08g052300.1.p, Vigun08g052400.1.p, Vigun08g053000.1.p, Vigun08g053100.1.p, Vigun08g053300.1.p, Vigun08g054700.1.p, Vigun08g054700.2.p, Vigun08g055100.1.p, Vigun08g064800.1.p, Vigun08g067700.1.p, Vigun08g070700.2.p, Vigun08g072200.1.p, Vigun08g072400.1.p, Vigun08g072800.1.p, Vigun08g072900.1.p, Vigun08g073000.1.p, Vigun08g073100.1.p, Vigun08g073400.1.p, Vigun08g074000.1.p, Vigun08g074100.1.p, Vigun08g074500.1.p, Vigun08g074600.1.p, Vigun08g081300.1.p, Vigun08g081300.2.p, Vigun08g081300.3.p, Vigun08g081400.1.p, Vigun08g081600.1.p, Vigun08g081900.1.p, Vigun08g082000.1.p, Vigun08g082200.1.p, Vigun08g085200.1.p, Vigun08g086900.1.p, Vigun08g101200.1.p, Vigun08g101300.1.p, Vigun08g101400.1.p, Vigun08g101500.1.p, Vigun08g101700.1.p, Vigun08g101800.1.p, Vigun08g101900.1.p, Vigun08g102000.1.p, Vigun08g102100.1.p, Vigun08g113500.1.p, Vigun08g113500.2.p, Vigun08g115200.1.p, Vigun08g121500.1.p, Vigun08g122600.2.p,Vigun08g135800.1.p, Vigun08g135800.2.p, Vigun08g136100.1.p, Vigun08g136100.2.p, Vigun08g136100.3.p, Vigun08g136100.4.p, Vigun08g136100.5.p, Vigun08g136800.1.p, Vigun08g137000.1.p, Vigun08g137000.2.p, Vigun08g142300.1.p, Vigun08g144000.1.p, Vigun08g158400.1.p, Vigun08g158500.1.p, Vigun08g158500.2.p, Vigun08g158600.1.p, Vigun08g158700.1.p, Vigun08g158800.1.p, Vigun08g158900.1.p, Vigun08g159000.1.p, Vigun08g159100.1.p, Vigun08g159200.1.p, Vigun08g159300.1.p, Vigun08g159400.1.p, Vigun08g159500.1.p, Vigun08g159600.1.p, Vigun08g159700.1.p, Vigun08g160900.1.p, Vigun08g163200.1.p, Vigun08g165700.1.p, Vigun08g173200.1.p, Vigun08g173500.1.p, Vigun08g173900.1.p, Vigun08g175200.1.p, Vigun08g175200.2.p, Vigun08g175300.1.p, Vigun08g178300.1.p, Vigun08g178300.2.p, Vigun08g186400.1.p, Vigun08g192800.1.p, Vigun08g192800.2.p, Vigun08g192800.3.p, Vigun08g192800.4.p, Vigun08g192800.5.p, Vigun08g192800.6.p, Vigun08g192800.7.p, Vigun08g192800.8.p, Vigun08g206500.1.p, Vigun08g210400.1.p, Vigun08g212500.1.p, Vigun08g220100.1.p, Vigun08g220100.2.p, Vigun08g220100.3.p, Vigun09g002600.1.p, Vigun09g002600.2.p, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VU | Vigun09g003000.1.p, Vigun09g004000.1.p, Vigun09g018000.1.p, Vigun09g025000.1.p, Vigun09g025100.1.p, Vigun09g025200.1.p, Vigun09g025400.1.p, Vigun09g032500.1.p, Vigun09g032500.3.p, Vigun09g032500.4.p, Vigun09g033200.1.p, Vigun09g033200.2.p, Vigun09g033200.3.p, Vigun09g033200.4.p, Vigun09g039900.1.p, Vigun09g041100.1.p, Vigun09g045800.1.p, Vigun09g054500.1.p, Vigun09g055100.1.p, Vigun09g064000.1.p, Vigun09g067100.1.p, Vigun09g077700.1.p, Vigun09g078800.1.p, Vigun09g078800.2.p, Vigun09g078800.3.p, Vigun09g083400.1.p, Vigun09g096400.1.p, Vigun09g101400.1.p, Vigun09g101400.2.p, Vigun09g128000.1.p, Vigun09g129900.1.p, Vigun09g130100.1.p, Vigun09g133300.1.p, Vigun09g133300.2.p, Vigun09g134700.1.p, Vigun09g134700.2.p, Vigun09g134700.3.p, Vigun09g134700.4.p, Vigun09g142300.1.p, Vigun09g142500.1.p, Vigun09g142700.1.p, Vigun09g154000.1.p, Vigun09g154600.1.p, Vigun09g157600.1.p, Vigun09g168700.1.p, Vigun09g187800.1.p, Vigun09g187900.1.p, Vigun09g189800.1.p, Vigun09g191400.1.p, Vigun09g196200.1.p, Vigun09g199300.1.p, Vigun09g209900.1.p, Vigun09g213400.1.p, Vigun09g213400.2.p, Vigun09g235900.1.p, Vigun09g242900.1.p, Vigun09g242900.2.p, Vigun09g242900.3.p, Vigun09g251200.1.p, Vigun09g264900.1.p, Vigun09g265200.1.p, Vigun09g265300.1.p, Vigun10g025900.1.p, Vigun10g025900.2.p, Vigun10g025900.3.p, Vigun10g025900.4.p, Vigun10g025900.5.p, Vigun10g025900.6.p, Vigun10g029600.1.p, Vigun10g042400.1.p, Vigun10g042900.1.p, Vigun10g049200.1.p, Vigun10g049200.2.p, Vigun10g051800.1.p, Vigun10g085600.1.p, Vigun10g087500.1.p, Vigun10g088300.1.p, Vigun10g099000.1.p, Vigun10g099000.2.p, Vigun10g110100.1.p, Vigun10g110100.2.p, Vigun10g110200.1.p, Vigun10g110300.1.p, Vigun10g110300.2.p, Vigun10g110400.1.p, Vigun10g110500.1.p, Vigun10g110800.1.p, Vigun10g110900.1.p, Vigun10g115700.1.p, Vigun10g133700.1.p, Vigun10g155600.1.p, Vigun10g156000.1.p, Vigun10g156000.2.p, Vigun10g162000.1.p, Vigun10g173300.1.p, Vigun10g173300.2.p, Vigun10g179100.1.p, Vigun10g179100.2.p, Vigun10g183600.1.p, Vigun10g199700.1.p, Vigun11g002600.1.p, Vigun11g006300.1.p, Vigun11g015000.1.p, Vigun11g015000.2.p, Vigun11g019300.1.p, Vigun11g019300.2.p, Vigun11g019300.3.p, Vigun11g019300.4.p, Vigun11g020500.1.p, Vigun11g020500.2.p, Vigun11g020600.1.p, Vigun11g020600.2.p, Vigun11g020800.1.p, Vigun11g020900.1.p, Vigun11g021100.1.p, Vigun11g021200.1.p, Vigun11g021300.1.p, Vigun11g021400.1.p, Vigun11g021500.1.p, Vigun11g021700.1.p, Vigun11g021800.1.p, Vigun11g022000.1.p, Vigun11g022000.2.p, Vigun11g022100.1.p, Vigun11g022200.1.p, Vigun11g022400.1.p, Vigun11g029100.1.p, Vigun11g033500.1.p, Vigun11g038600.1.p, Vigun11g038700.1.p, Vigun11g039700.1.p, Vigun11g039800.1.p, Vigun11g039800.2.p, Vigun11g039800.3.p, Vigun11g039900.1.p, Vigun11g047400.1.p, Vigun11g063400.1.p, Vigun11g068300.2.p, Vigun11g068300.3.p, Vigun11g074500.1.p, Vigun11g074500.2.p, Vigun11g074500.3.p, Vigun11g074500.4.p, Vigun11g074500.5.p, Vigun11g074500.6.p,Vigun11g074500.7.p, Vigun11g074500.8.p, Vigun11g074500.9.p, Vigun11g081700.1.p, Vigun11g081700.2.p, Vigun11g099500.1.p, Vigun11g103200.1.p, Vigun11g108800.1.p, Vigun11g108800.2.p, Vigun11g140700.1.p, Vigun11g150300.1.p, Vigun11g150300.2.p, Vigun11g165900.1.p, Vigun11g175200.1.p, Vigun11g178500.1.p, Vigun11g190000.1.p, Vigun11g190000.2.p, Vigun11g190000.3.p, Vigun11g190000.4.p, Vigun11g218200.1.p, Vigun11g218400.1.p, Vigun11g219200.1.p, Vigun11g223600.1.p, Vigun11g223700.1.p |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VV | GSVIVT01000017001, GSVIVT01000058001, GSVIVT01000200001, GSVIVT01000210001, GSVIVT01000211001, GSVIVT01000212001, GSVIVT01000222001, GSVIVT01000352001, GSVIVT01001044001, GSVIVT01001061001, GSVIVT01001078001, GSVIVT01001600001, GSVIVT01001679001, GSVIVT01001699001, GSVIVT01001702001, GSVIVT01001811001, GSVIVT01001815001, GSVIVT01002472001, GSVIVT01002515001, GSVIVT01003162001, GSVIVT01003778001, GSVIVT01003848001, GSVIVT01004004001, GSVIVT01004465001, GSVIVT01004673001, GSVIVT01004806001, GSVIVT01004912001, GSVIVT01005158001, GSVIVT01005163001, GSVIVT01005211001, GSVIVT01005222001, GSVIVT01005266001, GSVIVT01005456001, GSVIVT01005704001, GSVIVT01005722001, GSVIVT01006063001, GSVIVT01006118001, GSVIVT01006159001, GSVIVT01006299001, GSVIVT01006317001, GSVIVT01006319001, GSVIVT01006320001, GSVIVT01006430001, GSVIVT01006471001, GSVIVT01006472001, GSVIVT01006526001, GSVIVT01006683001, GSVIVT01006740001, GSVIVT01006872001, GSVIVT01006942001, GSVIVT01007136001, GSVIVT01007137001, GSVIVT01007370001, GSVIVT01007371001, GSVIVT01007478001, GSVIVT01007661001, GSVIVT01007688001, GSVIVT01007706001, GSVIVT01007713001, GSVIVT01007755001, GSVIVT01007793001, GSVIVT01008070001, GSVIVT01008212001, GSVIVT01008214001, GSVIVT01008308001, GSVIVT01008327001, GSVIVT01008328001, GSVIVT01008432001, GSVIVT01008436001, GSVIVT01008509001, GSVIVT01008747001, GSVIVT01009256001, GSVIVT01009414001, GSVIVT01009415001, GSVIVT01009420001, GSVIVT01009422001, GSVIVT01009444001, GSVIVT01009829001, GSVIVT01009941001, GSVIVT01010040001, GSVIVT01010052001, GSVIVT01010078001, GSVIVT01010159001, GSVIVT01010298001, GSVIVT01010376001, GSVIVT01010436001, GSVIVT01010596001, GSVIVT01011370001, GSVIVT01011373001, GSVIVT01011374001, GSVIVT01011375001, GSVIVT01011615001, GSVIVT01011677001, GSVIVT01011842001, GSVIVT01011849001, GSVIVT01011850001, GSVIVT01011854001, GSVIVT01011857001, GSVIVT01011952001, GSVIVT01012008001, GSVIVT01012107001, GSVIVT01012617001, GSVIVT01012618001, GSVIVT01012662001, GSVIVT01012686001, GSVIVT01013402001, GSVIVT01013461001, GSVIVT01013705001, GSVIVT01013741001, GSVIVT01013787001, GSVIVT01013796001, GSVIVT01013798001, GSVIVT01013799001, GSVIVT01014024001, GSVIVT01014051001, GSVIVT01014070001, GSVIVT01014112001, GSVIVT01014115001, GSVIVT01014128001, GSVIVT01014140001, GSVIVT01014149001, GSVIVT01014189001, GSVIVT01014478001, GSVIVT01014494001, GSVIVT01014519001, GSVIVT01014533001, GSVIVT01014571001, GSVIVT01014693001, GSVIVT01014750001, GSVIVT01015046001, GSVIVT01015198001, GSVIVT01015298001, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VV | GSVIVT01015383001, GSVIVT01015460001, GSVIVT01015489001, GSVIVT01015513001, GSVIVT01015799001, GSVIVT01015977001, GSVIVT01016054001, GSVIVT01016080001, GSVIVT01016126001, GSVIVT01016473001, GSVIVT01016502001, GSVIVT01016648001, GSVIVT01016654001, GSVIVT01016722001, GSVIVT01016857001, GSVIVT01016910001, GSVIVT01016926001, GSVIVT01016933001, GSVIVT01016975001, GSVIVT01017060001, GSVIVT01017096001, GSVIVT01017326001, GSVIVT01017518001, GSVIVT01017568001, GSVIVT01017681001, GSVIVT01017779001, GSVIVT01017953001, GSVIVT01018462001, GSVIVT01018490001, GSVIVT01018536001, GSVIVT01018814001, GSVIVT01019038001, GSVIVT01019039001, GSVIVT01019205001, GSVIVT01019259001, GSVIVT01019630001, GSVIVT01019727001, GSVIVT01020055001, GSVIVT01020131001, GSVIVT01020514001, GSVIVT01020556001, GSVIVT01020728001, GSVIVT01020730001, GSVIVT01020735001,GSVIVT01020778001, GSVIVT01020857001, GSVIVT01020860001, GSVIVT01021118001, GSVIVT01021119001, GSVIVT01021228001, GSVIVT01021232001, GSVIVT01021233001, GSVIVT01021278001, GSVIVT01021280001, GSVIVT01021283001, GSVIVT01021285001, GSVIVT01021286001, GSVIVT01021289001, GSVIVT01021293001, GSVIVT01021295001, GSVIVT01021297001, GSVIVT01021407001, GSVIVT01021409001, GSVIVT01021492001, GSVIVT01021635001, GSVIVT01021645001, GSVIVT01021647001, GSVIVT01021662001, GSVIVT01021667001, GSVIVT01021685001, GSVIVT01021687001, GSVIVT01021694001, GSVIVT01021777001, GSVIVT01022022001, GSVIVT01022070001, GSVIVT01022209001, GSVIVT01022210001, GSVIVT01022212001, GSVIVT01022261001, GSVIVT01022393001, GSVIVT01022729001, GSVIVT01022868001, GSVIVT01023037001, GSVIVT01023048001, GSVIVT01023113001, GSVIVT01023349001, GSVIVT01023362001, GSVIVT01023365001, GSVIVT01023369001, GSVIVT01023658001, GSVIVT01023665001, GSVIVT01023673001, GSVIVT01023677001, GSVIVT01023681001, GSVIVT01023710001, GSVIVT01023846001, GSVIVT01024236001, GSVIVT01024382001, GSVIVT01024538001, GSVIVT01024545001, GSVIVT01024731001, GSVIVT01025061001, GSVIVT01025185001, GSVIVT01025212001, GSVIVT01025346001, GSVIVT01025552001, GSVIVT01025870001, GSVIVT01026000001, GSVIVT01026007001, GSVIVT01026016001, GSVIVT01026021001, GSVIVT01026124001, GSVIVT01026196001, GSVIVT01026205001, GSVIVT01026418001, GSVIVT01026551001, GSVIVT01026898001, GSVIVT01026982001, GSVIVT01027054001, GSVIVT01027459001, GSVIVT01027564001, GSVIVT01027831001, GSVIVT01028135001, GSVIVT01028400001, GSVIVT01028401001, GSVIVT01028463001, GSVIVT01028542001, GSVIVT01028547001, GSVIVT01028727001, GSVIVT01028783001, GSVIVT01028785001, GSVIVT01028801001, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VV | GSVIVT01028858001, GSVIVT01029038001, GSVIVT01029148001, GSVIVT01029233001, GSVIVT01029462001, GSVIVT01029529001, GSVIVT01029718001, GSVIVT01029797001, GSVIVT01029798001, GSVIVT01029816001, GSVIVT01029829001, GSVIVT01029870001, GSVIVT01029950001, GSVIVT01030099001, GSVIVT01030106001, GSVIVT01030179001, GSVIVT01030183001, GSVIVT01030206001, GSVIVT01030309001, GSVIVT01030310001, GSVIVT01030482001, GSVIVT01030565001, GSVIVT01030992001, GSVIVT01030995001, GSVIVT01031315001, GSVIVT01031618001, GSVIVT01031658001, GSVIVT01031661001, GSVIVT01031699001, GSVIVT01031722001, GSVIVT01031962001, GSVIVT01031972001, GSVIVT01032263001, GSVIVT01032435001, GSVIVT01032499001, GSVIVT01032570001, GSVIVT01032610001, GSVIVT01032611001, GSVIVT01032670001, GSVIVT01032740001, GSVIVT01032768001, GSVIVT01032769001, GSVIVT01032772001, GSVIVT01032832001, GSVIVT01032879001, GSVIVT01032883001, GSVIVT01032888001, GSVIVT01032908001, GSVIVT01032970001, GSVIVT01033338001, GSVIVT01033343001, GSVIVT01033355001, GSVIVT01033490001, GSVIVT01033522001, GSVIVT01033727001, GSVIVT01034081001, GSVIVT01034082001, GSVIVT01034108001, GSVIVT01034326001, GSVIVT01034573001, GSVIVT01034604001, GSVIVT01034766001, GSVIVT01034992001, GSVIVT01035007001, GSVIVT01035186001, GSVIVT01035188001, GSVIVT01035211001, GSVIVT01035242001, GSVIVT01035303001, GSVIVT01035304001, GSVIVT01035307001, GSVIVT01035311001, GSVIVT01035313001, GSVIVT01035367001, GSVIVT01035397001, GSVIVT01035476001, GSVIVT01035532001, GSVIVT01035611001, GSVIVT01035842001, GSVIVT01035870001, GSVIVT01035915001, GSVIVT01036465001, GSVIVT01036495001, GSVIVT01036527001, GSVIVT01036637001, GSVIVT01036639001, GSVIVT01036853001, GSVIVT01036936001, GSVIVT01036966001, GSVIVT01037563001, GSVIVT01037982001, GSVIVT01038010001, GSVIVT01038348001, GSVIVT01038466001, GSVIVT01038628001, GSVIVT01038671001, GSVIVT01038699001, GSVIVT01038703001, GSVIVT01038718001, GSVIVT01038720001, GSVIVT01038721001, GSVIVT01038723001, GSVIVT01038764001, |
| AT | AT1G01450.1, AT1G01540.1, AT1G01540.2, AT1G06840.1, AT1G07150.1, AT1G07870.2, AT1G08590.1, AT1G09440.1, AT1G09970.1, AT1G09970.2, AT1G10620.1, AT1G10850.1, AT1G11050.1, AT1G11130.1, AT1G11390.1, AT1G12460.1, AT1G14390.1, AT1G15530.1, AT1G16110.1, AT1G16120.1, AT1G16130.1, AT1G16150.1, AT1G16160.1, AT1G16260.1, AT1G16260.2, AT1G17230.1, AT1G17750.1, AT1G17910.1, AT1G18390.1, AT1G18390.2, AT1G19090.1, AT1G19390.1, AT1G23540.1, AT1G24650.1, AT1G25320.1, AT1G25390.1, AT1G26150.1, AT1G27190.1, AT1G27570.2, AT1G28440.1, AT1G29720.1, AT1G29730.1, AT1G29740.1, AT1G29750.1, AT1G29750.2, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| AT | AT1G30570.1, AT1G31420.1, AT1G31420.2, AT1G34110.1, AT1G34210.1, AT1G34420.1, AT1G35710.1, AT1G48480.1, AT1G49270.1, AT1G49730.1, AT1G49730.2, AT1G49730.3, AT1G49730.4, AT1G50610.1, AT1G51790.1, AT1G51870.1, AT1G51880.1, AT1G51890.2, AT1G51910.1, AT1G51940.1, AT1G52290.1, AT1G52310.1, AT1G53420.1, AT1G53440.1, AT1G53730.1, AT1G53730.2, AT1G54820.1, AT1G55610.1, AT1G55610.2, AT1G56145.1, AT1G56145.2, AT1G56720.1, AT1G56720.2, AT1G56720.3, AT1G58643.1, AT1G58936.1, AT1G59312.1, AT1G60630.1, AT1G60800.1, AT1G61475.1, AT1G61640.2, AT1G62950.1, AT1G63430.1, AT1G63430.2, AT1G64210.1, AT1G65950.1, AT1G66150.1, AT1G66830.1, AT1G66880.1, AT1G66910.1, AT1G66930.1, AT1G66980.1, AT1G67000.1, AT1G68400.1, AT1G68690.1, AT1G69270.1, AT1G69730.1, AT1G69910.1, AT1G69990.1, AT1G70110.1, AT1G70130.1, AT1G70460.1, AT1G70520.1, AT1G70530.1, AT1G71830.1, AT1G72180.1, AT1G72300.1, AT1G72460.1, AT1G73080.1, AT1G74360.1, AT1G75640.1, AT1G75820.1, AT1G78530.1, AT1G78980.1, AT1G79620.1, AT1G79670.1, AT1G79670.2, AT1G79680.1, AT1G80640.1, AT1G80640.2, AT1G80870.1, AT2G01820.1, AT2G01950.1, AT2G02220.1, AT2G02780.1, AT2G07040.1, AT2G13790.1, AT2G13800.1, AT2G14510.1, AT2G15300.1, AT2G16250.1, AT2G18470.1, AT2G18530.1, AT2G20300.1, AT2G20850.1, AT2G21480.1, AT2G23200.1, AT2G23300.1, AT2G23450.1, AT2G23450.2, AT2G23770.1, AT2G23950.1, AT2G24230.1, AT2G25220.1, AT2G25220.2, AT2G25790.1, AT2G26330.1, AT2G26730.1, AT2G27060.1, AT2G28250.1, AT2G28250.2, AT2G28990.1, AT2G29000.1, AT2G29220.1, AT2G29250.1, AT2G30040.1, AT2G30940.1, AT2G30940.2, AT2G31880.1, AT2G33170.1, AT2G33580.1, AT2G35620.1, AT2G35620.2, AT2G36570.1, AT2G37710.1, AT2G38490.1, AT2G39180.1, AT2G39360.1, AT2G40090.1, AT2G40270.1, AT2G40270.2, AT2G41820.1, AT2G41910.1, AT2G42960.1, AT2G43690.1, AT2G43700.1, AT2G45340.1, AT2G45590.1, AT2G46850.1, AT2G48010.1, AT3G01840.1, AT3G02130.1, AT3G02880.1, AT3G03770.1, AT3G03770.2, AT3G04690.1, AT3G07700.1, AT3G07700.2, AT3G07700.3, AT3G08680.1, AT3G08680.2, AT3G08870.1, AT3G11870.1, AT3G13065.1, AT3G13380.1, AT3G14350.1, AT3G14350.2, AT3G14350.3, AT3G17420.1, AT3G17840.1, AT3G18810.1, AT3G19300.1, AT3G19700.1, AT3G20190.1, AT3G21630.1, AT3G23750.1, AT3G24240.1, AT3G24540.1, AT3G24550.1, AT3G24660.1, AT3G25490.1, AT3G25560.1, AT3G25560.2, AT3G25560.3, AT3G26700.1, AT3G28040.1, AT3G28450.1, AT3G42880.1, AT3G45330.1, AT3G45390.1, AT3G45410.1, AT3G45420.1, AT3G45430.1, AT3G45440.1, AT3G45860.1, AT3G46290.1, AT3G47090.1, AT3G47110.1, AT3G47570.1, AT3G47580.1, AT3G49670.1, AT3G50230.1, AT3G51550.1, AT3G51740.1, AT3G53380.1, AT3G53810.1, AT3G53840.1, AT3G55450.2, AT3G55550.1, AT3G55950.1, AT3G56050.1, AT3G56050.2, AT3G56100.1, AT3G56370.1, AT3G57730.1, AT3G57830.1, AT3G58690.1, AT3G59110.1, AT3G59420.1, AT3G59700.1, AT3G59730.1, AT3G59740.1, AT3G59750.1, AT3G63340.1, AT3G63340.2, AT4G00970.1, AT4G01330.1, AT4G01330.2, AT4G02010.1, AT4G02410.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| AT | AT4G02420.1, AT4G02630.1, AT4G03390.1, AT4G04490.1, AT4G04500.1, AT4G04510.1, AT4G04540.1, AT4G04570.1, AT4G04570.2, AT4G04960.1, AT4G05200.1, AT4G08850.1, AT4G08850.2, AT4G09760.3, AT4G11460.1, AT4G11470.1, AT4G11480.1, AT4G11490.1, AT4G11530.1, AT4G18250.1, AT4G18640.1, AT4G20140.1, AT4G20270.1, AT4G20790.1, AT4G20940.1, AT4G21230.1, AT4G21400.1, AT4G21410.1, AT4G22130.1, AT4G22130.2, AT4G22730.1, AT4G23130.1, AT4G23130.2, AT4G23140.1, AT4G23140.2,AT4G23150.1, AT4G23160.1, AT4G23180.1, AT4G23190.1, AT4G23210.1, AT4G23210.2, AT4G23210.3, AT4G23220.1, AT4G23230.1, AT4G23250.1, AT4G23260.1, AT4G23260.2, AT4G23270.1, AT4G23280.1, AT4G23290.1, AT4G23290.2, AT4G23300.1, AT4G23310.1, AT4G23320.1, AT4G23740.1, AT4G25390.1, AT4G25390.2, AT4G26540.1, AT4G28490.1, AT4G28670.1, AT4G29050.1, AT4G30520.1, AT4G31110.1, AT4G31250.1, AT4G32000.1, AT4G32000.2, AT4G32300.1, AT4G33430.1, AT4G33430.2, AT4G34220.1, AT4G34440.1, AT4G34500.1, AT4G36180.1, AT4G37250.1, AT4G38830.1, AT4G39110.1, AT4G39270.1, AT4G39270.2, AT4G39400.1, AT5G01540.1, AT5G01550.1, AT5G01560.1, AT5G01890.1, AT5G01950.1, AT5G02070.1, AT5G03140.1, AT5G05160.1, AT5G06740.1, AT5G06820.1, AT5G06940.1, AT5G07150.1, AT5G07180.1, AT5G07280.1, AT5G07620.1, AT5G10020.1, AT5G10020.2, AT5G10290.1, AT5G10530.1, AT5G11020.1, AT5G13290.1, AT5G13290.2, AT5G13290.3, AT5G14210.1, AT5G15730.1, AT5G15730.2, AT5G16000.1, AT5G16590.1, AT5G16900.1, AT5G18500.1, AT5G18500.2, AT5G20050.1, AT5G20480.1, AT5G20690.1, AT5G24010.1, AT5G24080.1, AT5G24100.1, AT5G24360.1, AT5G24360.2, AT5G24360.3, AT5G24970.2, AT5G25930.1, AT5G28680.1, AT5G35370.1, AT5G35390.1, AT5G37350.2, AT5G37450.1, AT5G38210.1, AT5G38260.1, AT5G38280.1, AT5G38560.1, AT5G38990.1, AT5G39000.1, AT5G39020.1, AT5G39030.1, AT5G39390.1, AT5G40380.1, AT5G42120.1, AT5G42440.1, AT5G43020.1, AT5G44700.1, AT5G45780.1, AT5G45800.1, AT5G45840.1, AT5G45840.2, AT5G46080.1, AT5G46330.1, AT5G47850.1, AT5G48380.1, AT5G48940.1, AT5G49660.1, AT5G49760.1, AT5G49770.1, AT5G49780.1, AT5G51350.1, AT5G51560.1, AT5G53320.1, AT5G53890.1, AT5G54380.1, AT5G54590.1, AT5G54590.2, AT5G55830.1, AT5G56040.1, AT5G56040.2, AT5G56890.1, AT5G58150.1, AT5G58300.1, AT5G58300.2, AT5G58540.1, AT5G58540.2, AT5G58540.3, AT5G59260.1, AT5G59270.1, AT5G59660.1, AT5G59700.1, AT5G60270.1, AT5G60280.1, AT5G60300.1, AT5G60300.2, AT5G60300.3, AT5G60310.1, AT5G60320.1, AT5G61350.1, AT5G61480.1, AT5G61570.1, AT5G61570.2, AT5G62230.1, AT5G62230.2, AT5G62710.1, AT5G63410.1, AT5G63930.1, AT5G64940.1, AT5G64940.2, AT5G65240.1, AT5G65240.2, AT5G65600.1, AT5G65700.1, AT5G65700.2, AT5G65710.1, AT5G66790.1, AT5G67200.1, AT5G67280.1 |
| SL | Solyc00g009090.2.1, Solyc00g110870.2.1, Solyc00g153680.1.1, Solyc01g005030.2.1, Solyc01g006530.1.1, Solyc01g006590.1.1, Solyc01g007130.2.1, Solyc01g007960.2.1, Solyc01g007970.1.1, Solyc01g007990.2.1, Solyc01g008140.2.1, Solyc01g008500.2.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| SL | Solyc01g010030.2.1, Solyc01g014520.1.1, Solyc01g059910.2.1, Solyc01g067020.2.1, Solyc01g067510.2.1, Solyc01g068360.2.1, Solyc01g079340.2.1, Solyc01g080770.2.1, Solyc01g091230.2.1, Solyc01g094020.2.1, Solyc01g094940.2.1, Solyc01g096940.2.1, Solyc01g098410.2.1, Solyc01g098740.2.1, Solyc01g101100.2.1, Solyc01g102680.2.1, Solyc01g102700.2.1, Solyc01g102710.2.1, Solyc01g103530.2.1, Solyc01g104050.2.1, Solyc01g104970.2.1, Solyc01g105080.2.1, Solyc01g106160.1.1, Solyc01g107650.2.1, Solyc01g108000.2.1, Solyc01g108840.2.1, Solyc01g109530.2.1, Solyc01g109590.2.1, Solyc01g109650.2.1, Solyc01g109950.2.1, Solyc02g014030.1.1, Solyc02g021590.2.1, Solyc02g023950.2.1, Solyc02g031790.1.1, Solyc02g062790.2.1, Solyc02g063210.1.1, Solyc02g065520.1.1, Solyc02g067560.1.1, Solyc02g067780.2.1, Solyc02g068060.2.1, Solyc02g068300.2.1, Solyc02g068370.2.1, Solyc02g068660.1.1, Solyc02g068820.1.1, Solyc02g068830.1.1, Solyc02g069750.1.1, Solyc02g069970.2.1, Solyc02g070000.2.1, Solyc02g070890.2.1, Solyc02g070910.1.1, Solyc02g071810.2.1, Solyc02g071860.2.1, Solyc02g071880.2.1, Solyc02g072070.2.1, Solyc02g072310.2.1, Solyc02g072400.1.1, Solyc02g072440.2.1, Solyc02g072470.2.1, Solyc02g072480.2.1, Solyc02g072520.2.1, Solyc02g076660.2.1, Solyc02g077630.2.1, Solyc02g078170.1.1, Solyc02g078750.2.1, Solyc02g078780.2.1, Solyc02g079990.2.1, Solyc02g080010.2.1, Solyc02g080030.1.1, Solyc02g080040.2.1, Solyc02g080070.2.1, Solyc02g080080.2.1, Solyc02g081040.2.1, Solyc02g081050.2.1, Solyc02g081070.2.1, Solyc02g081500.2.1, Solyc02g082470.2.1,Solyc02g083410.2.1, Solyc02g084370.1.1, Solyc02g085430.2.1, Solyc02g086210.2.1, Solyc02g086270.2.1, Solyc02g086590.2.1, Solyc02g087460.1.1, Solyc02g089010.1.1, Solyc02g089090.2.1, Solyc02g089290.2.1, Solyc02g089550.2.1, Solyc02g089900.1.1, Solyc02g089920.1.1, Solyc02g090110.2.1, Solyc02g091170.2.1, Solyc02g091590.2.1, Solyc02g091840.2.1, Solyc02g091860.2.1, Solyc02g092940.2.1, Solyc02g093100.2.1, Solyc02g093970.2.1, Solyc02g094010.1.1, Solyc03g006030.2.1, Solyc03g006080.2.1, Solyc03g006100.2.1, Solyc03g006130.2.1, Solyc03g006300.1.1, Solyc03g006660.2.1, Solyc03g006920.2.1, Solyc03g007050.2.1, Solyc03g007210.2.1, Solyc03g007910.1.1, Solyc03g019830.2.1, Solyc03g019980.1.1, Solyc03g026040.2.1, Solyc03g031980.2.1, Solyc03g033610.1.1, Solyc03g034060.2.1, Solyc03g043710.1.1, Solyc03g044160.1.1, Solyc03g051800.1.1, Solyc03g059020.1.1, Solyc03g059070.2.1, Solyc03g059490.1.1, Solyc03g062660.2.1, Solyc03g064010.2.1, Solyc03g078520.2.1, Solyc03g080060.1.1, Solyc03g082450.2.1, Solyc03g082470.2.1, Solyc03g083470.2.1, Solyc03g093330.2.1, Solyc03g093460.2.1, Solyc03g095490.2.1, Solyc03g096190.1.1, Solyc03g098400.1.1, Solyc03g111540.1.1, Solyc03g111670.2.1, Solyc03g111800.2.1, Solyc03g112310.1.1, Solyc03g112580.2.1, Solyc03g112730.2.1, Solyc03g113450.2.1, Solyc03g114080.1.1, Solyc03g115610.2.1, Solyc03g115710.1.1, Solyc03g117180.2.1, Solyc03g117550.1.1, Solyc03g118510.2.1, Solyc03g119240.2.1, Solyc03g120760.2.1, Solyc03g121050.2.1, Solyc03g121610.2.1, Solyc03g122230.1.1, Solyc03g123740.2.1, Solyc03g123780.2.1, Solyc03g123860.2.1, Solyc03g124050.2.1, Solyc04g005390.1.1, Solyc04g005910.2.1, Solyc04g006930.2.1, Solyc04g008300.1.1, Solyc04g008430.1.1, Solyc04g008650.2.1, Solyc04g009040.2.1, Solyc04g009640.2.1, Solyc04g012090.1.1, Solyc04g012100.1.1, Solyc04g012190.1.1, Solyc04g014650.2.1, Solyc04g015460.2.1, Solyc04g015600.2.1, Solyc04g015980.2.1, Solyc04g024970.1.1, Solyc04g039730.2.1, Solyc04g045600.2.1, Solyc04g050170.2.1, Solyc04g050940.1.1, Solyc04g051510.1.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDS |
|---|---|
| SL | Solyc04g054190.2.1, Solyc04g064940.2.1, Solyc04g071000.1.1, Solyc04g071680.1.1, Solyc04g071870.1.1, Solyc04g072570.2.1, Solyc04g074000.2.1, Solyc04g074020.2.1, Solyc04g074030.2.1, Solyc04g074050.2.1, Solyc04g076990.2.1, Solyc04g077010.2.1, Solyc04g077330.2.1, Solyc04g077380.2.1, Solyc04g078590.2.1, Solyc04g078700.2.1, Solyc04g079710.2.1, Solyc04g081080.1.1, Solyc04g081590.2.1, Solyc04g082620.2.1, Solyc05g005070.2.1, Solyc05g005140.2.1, Solyc05g005660.1.1, Solyc05g006570.1.1, Solyc05g007120.2.1, Solyc05g007230.2.1, Solyc05g008860.2.1, Solyc05g008930.2.1, Solyc05g008940.2.1, Solyc05g008950.2.1, Solyc05g008960.2.1, Solyc05g008980.2.1, Solyc05g008990.2.1, Solyc05g009040.2.1, Solyc05g009050.2.1, Solyc05g009090.1.1, Solyc05g009100.2.1, Solyc05g010140.2.1, Solyc05g010400.2.1, Solyc05g014240.2.1, Solyc05g015150.2.1, Solyc05g023760.2.1, Solyc05g047570.2.1, Solyc05g051640.2.1, Solyc05g052350.2.1, Solyc05g052590.1.1, Solyc05g053010.1.1, Solyc05g054680.1.1, Solyc05g054860.1.1, Solyc05g056370.2.1, Solyc06g005880.2.1, Solyc06g006010.1.1, Solyc06g008890.2.1, Solyc06g009540.1.1, Solyc06g048560.1.1, Solyc06g048620.2.1, Solyc06g048740.1.1, Solyc06g048950.2.1, Solyc06g050270.1.1, Solyc06g050280.1.1, Solyc06g050290.1.1, Solyc06g050300.1.1, Solyc06g050560.2.1, Solyc06g051030.2.1, Solyc06g051540.2.1, Solyc06g063150.2.1, Solyc06g063360.2.1, Solyc06g065260.2.1, Solyc06g068910.2.1, Solyc06g069610.1.1, Solyc06g069650.2.1, Solyc06g071810.1.1, Solyc06g074070.2.1, Solyc06g076910.1.1, Solyc06g082610.2.1, Solyc06g082920.2.1, Solyc07g005010.2.1, Solyc07g005540.1.1, Solyc07g005920.1.1, Solyc07g006110.2.1, Solyc07g006480.2.1, Solyc07g006770.2.1, Solyc07g006940.1.1, Solyc07g007780.2.1, Solyc07g008400.1.1, Solyc07g017230.2.1, Solyc07g018180.1.1, Solyc07g018190.2.1, Solyc07g039340.2.1, Solyc07g045420.2.1, Solyc07g049180.2.1, Solyc07g053600.2.1, Solyc07g054500.2.1, Solyc07g055180.2.1, Solyc07g055560.2.1, Solyc07g055650.1.1, Solyc07g056270.2.1, Solyc07g056410.2.1, Solyc07g062170.1.1, Solyc07g063000.2.1, Solyc07g063950.2.1,Solyc07g064820.1.1, Solyc07g065240.1.1, Solyc07g065610.1.1, Solyc07g065860.2.1, Solyc07g066230.2.1, Solyc08g016730.1.1, Solyc08g059730.1.1, Solyc08g061560.2.1, Solyc08g066210.2.1, Solyc08g066270.1.1, Solyc08g066300.1.1, Solyc08g066320.2.1, Solyc08g066490.2.1, Solyc08g068920.2.1, Solyc08g069170.1.1, Solyc08g074760.2.1, Solyc08g075330.1.1, Solyc08g075590.1.1, Solyc08g075600.1.1, Solyc08g075610.1.1, Solyc08g075620.1.1, Solyc08g079460.2.1, Solyc08g079740.2.1, Solyc08g080830.2.1, Solyc08g081940.2.1, Solyc09g005000.1.1, Solyc09g007110.2.1, Solyc09g007510.1.1, Solyc09g007750.2.1, Solyc09g008640.1.1, Solyc09g008860.2.1, Solyc09g011060.2.1, Solyc09g011070.1.1, Solyc09g011200.1.1, Solyc09g011990.1.1, Solyc09g012000.1.1, Solyc09g015170.2.1, Solyc09g015230.1.1, Solyc09g015520.2.1, Solyc09g015830.1.1, Solyc09g030450.2.1, Solyc09g055180.2.1, Solyc09g055570.2.1, Solyc09g057680.2.1, Solyc09g060110.2.1, Solyc09g061930.2.1, Solyc09g061940.1.1, Solyc09g064270.2.1, Solyc09g072810.2.1, Solyc09g074240.1.1, Solyc09g075720.2.1, Solyc09g083210.2.1, Solyc09g090210.2.1, Solyc09g091260.2.1, Solyc09g091400.2.1, Solyc09g098290.2.1, Solyc09g098420.1.1, Solyc10g005140.1.1, Solyc10g006790.2.1, Solyc10g006870.1.1, Solyc10g008960.1.1, Solyc10g047140.1.1, Solyc10g047680.1.1, Solyc10g047700.1.1, Solyc10g047810.1.1, Solyc10g051330.1.1, Solyc10g054050.1.1, Solyc10g074890.1.1, Solyc10g074900.1.1, Solyc10g076530.1.1, Solyc10g079170.1.1, Solyc10g080510.1.1, Solyc10g081420.1.1, Solyc10g081910.1.1, Solyc10g084250.1.1, |

**Table A4. Protein IDs of RLK predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| SL | Solyc10g084860.1.1, Solyc10g085110.1.1, Solyc10g085120.1.1, Solyc10g085670.1.1, Solyc11g006040.1.1, Solyc11g008960.1.1, Solyc11g011020.1.1, Solyc11g011870.1.1, Solyc11g011880.1.1, Solyc11g017270.1.1, Solyc11g020230.1.1, Solyc11g020260.1.1, Solyc11g020280.1.1, Solyc11g039590.1.1, Solyc11g044460.1.1, Solyc11g044940.1.1, Solyc11g069630.1.1, Solyc11g069960.1.1, Solyc11g071880.1.1, Solyc11g072910.1.1, Solyc12g007110.1.1, Solyc12g008400.1.1, Solyc12g008500.1.1, Solyc12g009190.1.1, Solyc12g009360.1.1, Solyc12g010740.1.1, Solyc12g014350.1.1, Solyc12g016050.1.1, Solyc12g016090.1.1, Solyc12g016130.1.1, Solyc12g036330.1.1, Solyc12g055720.1.1, Solyc12g056300.1.1, Solyc12g056730.1.1, Solyc12g088040.1.1, Solyc12g088750.1.1, Solyc12g089020.1.1, Solyc12g089160.1.1, Solyc12g096710.1.1, Solyc12g098100.1.1, Solyc12g098570.1.1 |

SP: Species

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes**

| SP | Protein IDs |
|---|---|
| CC | KYP31440.1, KYP32286.1, KYP32288.1, KYP32336.1, KYP32412.1, KYP32511.1, KYP33486.1, KYP34202.1, KYP35018.1, KYP36210.1, KYP36276.1, KYP36367.1, KYP36424.1, KYP36440.1, KYP36443.1, KYP36447.1, KYP37524.1, KYP37653.1, KYP37807.1, KYP37968.1, KYP37970.1, KYP38084.1, KYP38310.1, KYP38510.1, KYP38553.1, KYP38912.1, KYP39082.1, KYP39400.1, KYP40853.1, KYP41797.1, KYP41800.1, KYP41943.1, KYP42120.1, KYP42682.1, KYP42884.1, KYP42992.1, KYP43245.1, KYP43941.1, KYP44037.1, KYP44574.1, KYP44712.1, KYP44788.1, KYP45165.1, KYP45426.1, KYP46485.1, KYP46665.1, KYP47636.1, KYP47671.1, KYP47780.1, KYP48207.1, KYP48310.1, KYP48467.1, KYP48541.1, KYP48816.1, KYP48985.1, KYP49051.1, KYP49129.1, KYP49307.1, KYP49513.1, KYP50276.1, KYP50368.1, KYP50921.1, KYP51038.1, KYP51049.1, KYP51301.1, KYP51579.1, KYP51721.1, KYP51847.1, KYP52587.1, KYP52989.1, KYP53016.1, KYP53149.1, KYP53800.1, KYP54115.1, KYP54606.1, KYP55399.1, KYP55637.1, KYP56346.1, KYP56573.1, KYP56874.1, KYP57449.1, KYP58262.1, KYP58701.1, KYP59506.1, KYP59509.1, KYP59693.1, KYP59711.1, KYP60284.1, KYP60285.1, KYP60304.1, KYP60531.1, KYP60767.1, KYP61537.1, KYP61898.1, KYP61904.1, KYP62046.1, KYP62087.1, KYP62415.1, KYP62743.1, KYP63546.1, KYP63716.1, KYP64187.1, KYP64240.1, KYP64398.1, KYP65949.1, KYP66127.1, KYP66146.1, KYP66939.1, KYP67084.1, KYP67194.1, KYP67399.1, KYP67404.1, KYP67455.1, KYP67700.1, KYP68340.1, KYP68341.1, KYP69882.1, KYP70259.1, KYP71043.1, KYP73566.1, KYP73567.1, KYP73569.1, KYP74127.1, KYP74128.1, KYP74500.1, KYP74802.1, KYP74826.1, KYP74995.1, KYP75073.1, KYP75234.1, KYP75286.1, KYP75544.1, KYP75599.1, KYP75952.1, KYP76356.1, KYP76925.1, KYP78156.1, KYP78587.1 |
| GM | Glyma.01G013900.1.p, Glyma.01G039800.1.p, Glyma.01G039900.1.p, Glyma.01G048500.1.p, Glyma.01G048500.2.p, Glyma.01G048500.3.p, Glyma.01G048500.4.p, Glyma.01G078000.1.p, Glyma.01G078000.2.p, Glyma.01G087700.1.p, Glyma.01G090200.1.p, Glyma.01G109200.1.p, Glyma.01G117000.1.p, Glyma.01G117200.1.p, Glyma.01G118300.1.p, Glyma.01G118300.2.p, Glyma.01G120000.1.p, Glyma.01G120300.1.p, Glyma.01G137600.1.p, Glyma.01G146000.1.p, Glyma.01G152300.1.p, Glyma.01G165400.1.p, Glyma.01G165600.1.p, Glyma.01G196000.1.p, Glyma.01G211000.1.p, Glyma.01G242800.1.p, Glyma.01G244600.1.p, Glyma.02G025000.1.p, Glyma.02G047400.1.p, Glyma.02G068900.1.p, Glyma.02G077600.1.p, Glyma.02G099800.1.p, Glyma.02G101900.1.p, Glyma.02G101900.2.p, Glyma.02G101900.3.p, Glyma.02G108100.1.p, Glyma.02G132800.1.p, Glyma.02G150900.1.p, Glyma.02G150900.2.p, Glyma.02G156800.1.p, Glyma.02G183600.1.p, Glyma.02G220900.1.p, Glyma.02G220900.2.p, Glyma.02G242700.1.p, Glyma.02G244200.1.p, Glyma.02G269100.1.p, Glyma.03G029500.1.p, Glyma.03G052400.1.p, Glyma.03G053400.1.p, Glyma.03G066500.1.p, Glyma.03G074300.1.p, Glyma.03G087400.1.p, Glyma.03G139300.1.p, Glyma.03G166900.1.p, Glyma.03G193700.1.p, Glyma.03G197400.1.p, Glyma.04G034300.1.p, Glyma.04G037600.1.p, Glyma.04G042600.1.p, Glyma.04G176600.1.p, Glyma.04G185400.1.p, Glyma.04G202900.1.p, Glyma.04G202900.2.p, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.04G216900.1.p, Glyma.04G225800.1.p, Glyma.05G014600.1.p, Glyma.05G031600.1.p, Glyma.05G039100.1.p, Glyma.05G109600.1.p, Glyma.05G133700.1.p, Glyma.05G144000.1.p, Glyma.05G150400.1.p, Glyma.05G159900.1.p, Glyma.05G163900.1.p, Glyma.05G199200.1.p, Glyma.05G217000.1.p, Glyma.05G235200.1.p, Glyma.06G038500.1.p, Glyma.06G042300.1.p, Glyma.06G139000.1.p, Glyma.06G162600.1.p, Glyma.06G162600.2.p, Glyma.06G162600.3.p, Glyma.06G188500.1.p, Glyma.06G189100.1.p, Glyma.06G215000.1.p, Glyma.06G237200.1.p, Glyma.06G290700.1.p, Glyma.06G296800.2.p, Glyma.06G302300.1.p, Glyma.06G318500.1.p, Glyma.06G319700.1.p, Glyma.07G056200.1.p, Glyma.07G079600.1.p, Glyma.07G080600.1.p, Glyma.07G135100.1.p, Glyma.07G141400.1.p, Glyma.07G142900.1.p, Glyma.07G143000.1.p, Glyma.07G143500.1.p, Glyma.07G143900.1.p, Glyma.07G144100.1.p, Glyma.07G144100.2.p, Glyma.07G144100.3.p, Glyma.07G153200.1.p, Glyma.07G153200.2.p, Glyma.07G153500.1.p, Glyma.07G155000.1.p, Glyma.07G190500.1.p, Glyma.07G212200.1.p, Glyma.07G212200.2.p, Glyma.07G215900.1.p, Glyma.07G219800.1.p, Glyma.07G262800.1.p, Glyma.07G262800.2.p, Glyma.08G006500.1.pGlyma.08G022900.1.p, Glyma.08G039400.1.p, Glyma.08G042600.1.p, Glyma.08G058700.1.p, Glyma.08G065700.1.p, Glyma.08G093900.1.p, Glyma.08G093900.2.p, Glyma.08G100300.1.p, Glyma.08G100300.2.p, Glyma.08G100300.3.p, Glyma.08G107300.1.p, Glyma.08G121200.1.p, Glyma.08G198000.1.p, Glyma.08G205100.1.p, Glyma.08G274600.1.p, Glyma.08G294200.1.p, Glyma.08G294300.1.p, Glyma.08G309600.1.p, Glyma.08G350000.1.p, Glyma.09G062600.1.p, Glyma.09G125900.1.p, Glyma.09G145600.1.p, Glyma.09G193500.1.p, Glyma.09G231500.1.p, Glyma.09G231500.2.p, Glyma.09G248600.1.p, Glyma.09G248600.2.p, Glyma.09G251800.1.p, Glyma.09G263500.1.p, Glyma.09G272100.1.p, Glyma.09G272100.2.p, Glyma.09G274300.1.p, Glyma.09G274300.2.p, Glyma.09G274300.3.p, Glyma.09G278200.1.p, Glyma.10G023100.1.p, Glyma.10G038800.1.p, Glyma.10G099600.1.p, Glyma.10G110100.1.p, Glyma.10G129000.1.p, Glyma.10G129400.1.p, Glyma.10G133900.1.p, Glyma.10G133900.2.p, Glyma.10G165400.1.p, Glyma.10G190500.1.p, Glyma.10G228000.1.p, Glyma.10G228100.1.p, Glyma.10G228200.1.p, Glyma.10G228600.1.p, Glyma.10G228800.1.p, Glyma.10G251700.1.p, Glyma.10G253400.1.p, Glyma.10G254100.1.p, Glyma.10G254500.1.p, Glyma.10G257300.1.p, Glyma.11G001500.1.p, Glyma.11G025600.1.p, Glyma.11G044800.1.p, Glyma.11G045700.1.p, Glyma.11G045700.2.p, Glyma.11G077700.1.p, Glyma.11G106100.1.p, Glyma.11G106100.4.p, Glyma.11G114200.1.p, Glyma.11G126100.1.p, Glyma.11G140800.1.p, Glyma.11G225900.1.p, Glyma.11G253900.1.p, Glyma.12G031000.1.p, Glyma.12G031200.1.p, Glyma.12G064300.1.p, Glyma.12G101800.1.p, Glyma.12G116200.1.p, Glyma.12G123100.1.p, Glyma.12G123300.1.p, Glyma.12G123400.1.p, Glyma.12G233600.1.p, Glyma.12G233700.1.p, Glyma.13G003500.1.p, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDS |
|---|---|
| GM | Glyma.13G005800.1.p, Glyma.13G050500.1.p, Glyma.13G067100.1.p, Glyma.13G067200.1.p, Glyma.13G077700.1.p, Glyma.13G082700.1.p, Glyma.13G095200.1.p, Glyma.13G095300.1.p, Glyma.13G126700.1.p, Glyma.13G221000.1.p, Glyma.13G227700.1.p, Glyma.13G251500.1.p, Glyma.13G270100.1.p, Glyma.13G305900.1.p, Glyma.13G331300.1.p, Glyma.13G331300.2.p, Glyma.13G365900.1.p, Glyma.14G015500.1.p, Glyma.14G042100.1.p, Glyma.14G042300.1.p, Glyma.14G042500.1.p, Glyma.14G042700.1.p, Glyma.14G042900.1.p, Glyma.14G043000.1.p, Glyma.14G043300.1.p, Glyma.14G044400.1.p, Glyma.14G046000.1.p, Glyma.14G077200.1.p, Glyma.14G077300.1.p, Glyma.14G077300.2.p, Glyma.14G096200.1.p, Glyma.14G106400.1.p, Glyma.14G113100.1.p, Glyma.14G151900.1.p, Glyma.14G163700.1.p, Glyma.14G173300.1.p, Glyma.14G173400.1.p, Glyma.14G173600.1.p, Glyma.14G173600.2.p, Glyma.14G173600.3.p, Glyma.14G173600.4.p, Glyma.14G173900.1.p, Glyma.14G188400.1.p, Glyma.14G213600.1.p, Glyma.15G021600.1.p, Glyma.15G076900.2.p, Glyma.15G084800.1.p, Glyma.15G088400.1.p, Glyma.15G091500.1.p, Glyma.15G093000.1.p, Glyma.15G110400.1.p, Glyma.15G152800.1.p, Glyma.15G168900.1.p, Glyma.15G169100.1.p, Glyma.15G254300.1.p, Glyma.16G025100.1.p, Glyma.16G065500.2.p, Glyma.16G105800.1.p, Glyma.16G105900.1.p, Glyma.16G106200.1.p, Glyma.16G125300.1.p, Glyma.16G125600.1.p, Glyma.16G125800.1.p, Glyma.16G126000.1.p, Glyma.16G126100.1.p, Glyma.16G126900.1.p, Glyma.16G129300.1.p, Glyma.16G157500.1.p, Glyma.16G162300.1.p, Glyma.16G163400.1.p, Glyma.16G168700.1.p, Glyma.16G169000.1.p, Glyma.16G169100.1.p, Glyma.16G169200.1.p, Glyma.16G169300.1.p, Glyma.16G169400.1.p, Glyma.16G169500.1.p, Glyma.16G169700.1.p, Glyma.16G169900.1.p, Glyma.16G170700.1.p, Glyma.16G170800.1.p, Glyma.16G170900.1.p, Glyma.16G171000.1.p, Glyma.16G171100.1.p, Glyma.16G171200.1.p, Glyma.16G171400.1.p, Glyma.16G171500.1.p, Glyma.16G171600.1.p, Glyma.16G171800.1.p, Glyma.16G172100.1.p, Glyma.16G172200.1.p, Glyma.16G172300.1.p, Glyma.16G173900.1.p, Glyma.16G173900.2.p, Glyma.16G174000.1.p, Glyma.16G174100.1.p, Glyma.16G174500.1.p, Glyma.16G174600.1.p, Glyma.16G174700.1.p, Glyma.16G174800.1.p, Glyma.16G175000.1.p, Glyma.16G175100.1.p, Glyma.16G175100.2.p, Glyma.16G176400.1.p, Glyma.16G176600.1.p, Glyma.16G176800.1.p, Glyma.16G176900.1.p, Glyma.16G182700.1.p, Glyma.16G183000.1.p, Glyma.16G183300.1.p, Glyma.16G183400.1.p, Glyma.16G183500.1.p, Glyma.16G183600.1.p, Glyma.16G184000.1.p, Glyma.16G184200.1.p, Glyma.16G184300.1.p, Glyma.16G184700.1.p, Glyma.16G184800.1.p, Glyma.16G185100.1.p, Glyma.16G185300.1.p, Glyma.16G185400.1.p,Glyma.16G185800.1.p, Glyma.16G186100.1.p, Glyma.16G186200.1.p, Glyma.16G186400.1.p, Glyma.16G186500.1.p, Glyma.16G186600.1.p, Glyma.16G186700.1.p, Glyma.16G186900.1.p, Glyma.16G187200.1.p, Glyma.16G187400.1.p, Glyma.16G187600.1.p, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Domain IDs |
|---|---|
| GM | Glyma.16G187700.1.p, Glyma.16G187800.1.p, Glyma.16G187900.1.p, Glyma.16G188000.1.p, Glyma.16G188100.1.p, Glyma.16G188300.1.p, Glyma.16G188700.1.p, Glyma.16G190000.1.p, Glyma.16G190200.1.p, Glyma.16G190400.1.p, Glyma.16G190500.1.p, Glyma.16G190600.1.p, Glyma.16G191000.1.p, Glyma.16G191500.1.p, Glyma.16G191700.1.p, Glyma.16G191800.1.p, Glyma.16G192100.1.p, Glyma.16G192300.1.p, Glyma.16G192600.1.p, Glyma.16G192800.1.p, Glyma.16G192900.1.p, Glyma.16G193300.1.p, Glyma.16G193500.1.p, Glyma.16G193600.1.p, Glyma.16G193700.1.p, Glyma.16G193900.1.p, Glyma.16G201900.1.p, Glyma.16G202200.2.p, Glyma.16G202200.3.p, Glyma.16G202400.2.p, Glyma.17G011300.1.p, Glyma.17G038100.1.p, Glyma.17G038100.2.p, Glyma.17G038100.3.p, Glyma.17G038100.4.p, Glyma.17G038100.5.p, Glyma.17G038100.6.p, Glyma.17G038600.1.p, Glyma.17G064900.1.p, Glyma.17G087700.1.p, Glyma.17G095300.1.p, Glyma.17G095300.2.p, Glyma.17G123100.1.p, Glyma.17G157400.1.p, Glyma.17G243300.1.p, Glyma.17G248300.1.p, Glyma.17G248400.1.p, Glyma.17G248400.2.p, Glyma.17G248400.3.p, Glyma.17G248400.4.p, Glyma.17G258500.1.p, Glyma.17G258600.1.p, Glyma.18G031400.1.p, Glyma.18G036700.1.p, Glyma.18G040800.1.p, Glyma.18G040800.2.p, Glyma.18G128600.1.p, Glyma.18G128600.2.p, Glyma.18G150100.1.p, Glyma.18G150200.1.p, Glyma.18G159300.1.p, Glyma.18G191700.1.p, Glyma.18G193800.1.p, Glyma.18G193900.1.p, Glyma.18G195200.1.p, Glyma.18G195200.10.p, Glyma.18G195200.2.p, Glyma.18G195200.3.p, Glyma.18G195200.4.p, Glyma.18G195200.5.p, Glyma.18G195200.6.p, Glyma.18G195200.7.p, Glyma.18G195200.8.p, Glyma.18G195200.9.p, Glyma.18G204300.1.p, Glyma.18G204300.2.p, Glyma.18G204400.1.p, Glyma.18G204500.1.p, Glyma.18G204600.1.p, Glyma.18G205500.1.p, Glyma.18G205600.1.p, Glyma.18G210600.1.p, Glyma.18G213400.1.p, Glyma.18G228600.1.p, Glyma.18G236000.1.p, Glyma.18G240800.1.p, Glyma.18G244100.1.p, Glyma.18G273000.1.p, Glyma.19G017300.1.p, Glyma.19G045600.1.p, Glyma.19G045600.2.p, Glyma.19G066300.1.p, Glyma.19G100700.1.p, Glyma.19G100800.1.p, Glyma.19G115000.1.p, Glyma.19G145200.1.p, Glyma.19G168000.1.p, Glyma.19G195400.1.p, Glyma.19G255700.1.p, Glyma.20G012300.1.p, Glyma.20G012900.1.p, Glyma.20G015700.1.p, Glyma.20G080700.1.p, Glyma.20G083000.1.p, Glyma.20G083000.2.p, Glyma.20G133800.1.p, Glyma.20G137000.1.p, Glyma.20G137800.1.p, Glyma.20G138000.1.p, Glyma.20G140000.1.p, Glyma.20G140300.1.p, Glyma.20G141800.1.p, Glyma.20G173500.1.p, Glyma.20G186100.1.p, Glyma.20G200200.1.p, Glyma.20G227400.1.p, Glyma.U002100.1.p, Glyma.U002200.1.p, Glyma.U002300.1.p, Glyma.U002400.1.p, Glyma.U006500.1.p, Glyma.U006600.1.p, Glyma.U007900.1.p, Glyma.U023800.1.p, Glyma.U035200.1.p, Glyma.U037600.1.p, Glyma.U043600.1.p, Glyma.U044500.1.p |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| MT | Medtr0017s0230.1, Medtr0061s0050.1, Medtr0087s0020.1, Medtr0087s0030.1, Medtr0087s0040.1, Medtr0087s0050.1, Medtr0087s0070.1, Medtr0087s0090.1, Medtr0163s0020.1, Medtr0271s0040.1, Medtr0274s0010.1, Medtr0274s0050.1, Medtr0289s0040.1, Medtr0640s0020.1, Medtr0684s0020.1, Medtr0721s0020.1, Medtr0722s0020.1, Medtr1395s0010.1, Medtr1g009240.1, Medtr1g019780.1, Medtr1g021825.1, Medtr1g021950.1, Medtr1g027610.1, Medtr1g033420.1, Medtr1g044580.1, Medtr1g047190.1, Medtr1g052730.1, Medtr1g062340.1, Medtr1g062630.1, Medtr1g073320.1, Medtr1g080350.1, Medtr1g080720.1, Medtr1g090930.1, Medtr1g090937.1, Medtr1g090957.1, Medtr1g090960.1, Medtr1g093630.1, Medtr1g098980.1, Medtr1g099240.1, Medtr1g101250.1, Medtr1g104890.5, Medtr1g105495.1, Medtr1g105855.1, Medtr1g105860.3, Medtr1g105900.1, Medtr1g106940.1, Medtr1g110090.1, Medtr1g114010.3, Medtr2g010710.1, Medtr2g010720.1, Medtr2g010730.1, Medtr2g016200.1, Medtr2g016590.1, Medtr2g016620.1, Medtr2g017420.1, Medtr2g017470.1, Medtr2g017480.1, Medtr2g017495.1, Medtr2g017805.1, Medtr2g018770.1, Medtr2g018770.2, Medtr2g019170.1, Medtr2g023970.1, Medtr2g031660.1, Medtr2g032560.1, Medtr2g038705.1, Medtr2g038860.1, Medtr2g046780.1, Medtr2g055360.1, Medtr2g055690.1, Medtr2g068655.1, Medtr2g076590.1, Medtr2g078260.1, Medtr2g078420.1, Medtr2g088930.1,Medtr2g095800.1, Medtr2g450870.1, Medtr3g008530.1, Medtr3g027280.1, Medtr3g027330.1, Medtr3g031520.1, Medtr3g041560.1, Medtr3g041560.2, Medtr3g045020.1, Medtr3g047140.1, Medtr3g048440.1, Medtr3g048470.1, Medtr3g048590.1, Medtr3g048740.1, Medtr3g048760.1, Medtr3g048785.1, Medtr3g048860.1, Medtr3g048910.1, Medtr3g065370.1, Medtr3g069910.1, Medtr3g082130.1, Medtr3g089540.1, Medtr3g091200.1, Medtr3g095600.1, Medtr3g098430.1, Medtr3g098430.2, Medtr3g101480.1, Medtr3g102120.1, Medtr3g108560.2, Medtr3g111620.1, Medtr3g437820.1, Medtr3g449240.1, Medtr3g449520.1, Medtr3g451090.1, Medtr3g451890.1, Medtr3g452210.1, Medtr3g452730.1, Medtr3g452750.1, Medtr3g452760.1, Medtr3g452770.1, Medtr3g452790.1, Medtr3g452800.1, Medtr3g452850.1, Medtr3g452880.1, Medtr3g452900.1, Medtr3g452970.1, Medtr4g006280.1, Medtr4g009930.1, Medtr4g011310.1, Medtr4g011860.1, Medtr4g013315.1, Medtr4g015930.1, Medtr4g015930.10, Medtr4g015930.11, Medtr4g015930.12, Medtr4g015930.13, Medtr4g015930.14, Medtr4g015930.2, Medtr4g015930.3, Medtr4g015930.4, Medtr4g015930.5, Medtr4g015930.6, Medtr4g015930.7, Medtr4g015930.8, Medtr4g015930.9, Medtr4g015960.1, Medtr4g016780.1, Medtr4g016800.1, Medtr4g016820.1, Medtr4g016850.1, Medtr4g016870.1, Medtr4g016910.1, Medtr4g017260.1, Medtr4g017280.1, Medtr4g017350.1, Medtr4g017370.1, Medtr4g017490.1, Medtr4g017600.1, Medtr4g017640.1, Medtr4g017690.1, Medtr4g017700.1, Medtr4g017710.1, Medtr4g017720.1, Medtr4g017730.1, Medtr4g017780.1, Medtr4g018910.1, Medtr4g018920.1, Medtr4g018930.1, Medtr4g018940.1, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Domain IDs |
|---|---|
| MT | Medtr4g018970.1, Medtr4g019010.1, Medtr4g019030.1, Medtr4g019080.1, Medtr4g023810.1, Medtr4g032320.1, Medtr4g035200.1, Medtr4g035200.2, Medtr4g035200.3, Medtr4g040330.1, Medtr4g040360.1, Medtr4g040420.1, Medtr4g046920.1, Medtr4g056320.1, Medtr4g056360.1, Medtr4g056380.1, Medtr4g057450.1, Medtr4g063630.1, Medtr4g064760.1, Medtr4g070950.1, Medtr4g073730.2, Medtr4g073730.3, Medtr4g078535.1, Medtr4g085480.1, Medtr4g091000.1, Medtr4g091010.1, Medtr4g091020.1, Medtr4g094730.1, Medtr4g094730.2, Medtr4g095022.1, Medtr4g101630.1, Medtr4g105520.1, Medtr4g107360.1, Medtr4g123090.1, Medtr4g128400.1, Medtr4g128410.1, Medtr4g128570.1, Medtr4g128580.1, Medtr4g128590.1, Medtr4g132450.1, Medtr4g417260.1, Medtr4g417270.1, Medtr5g005100.1, Medtr5g009200.1, Medtr5g015050.1, Medtr5g022310.1, Medtr5g022310.2, Medtr5g022350.1, Medtr5g022350.2, Medtr5g023850.1, Medtr5g025900.1, Medtr5g025910.1, Medtr5g029100.1, Medtr5g037300.1, Medtr5g041210.1, Medtr5g041210.2, Medtr5g043800.1, Medtr5g046350.1, Medtr5g047390.1, Medtr5g053610.1, Medtr5g063740.1, Medtr5g063760.1, Medtr5g069580.1, Medtr5g078030.1, Medtr5g079980.1, Medtr5g080000.1, Medtr5g081920.1, Medtr5g082370.1, Medtr5g085910.1, Medtr5g085920.1, Medtr5g085930.1, Medtr5g085970.1, Medtr5g086530.1, Medtr5g086550.1, Medtr5g086570.1, Medtr5g086620.1, Medtr5g086630.1, Medtr5g086810.1, Medtr5g086945.1, Medtr5g087070.1, Medtr5g087080.1, Medtr5g087090.1, Medtr5g087320.1, Medtr5g089160.1, Medtr5g091730.1, Medtr5g092670.1, Medtr5g095120.1, Medtr5g095200.1, Medtr5g096320.1, Medtr5g096340.1, Medtr5g096360.1, Medtr6g009480.1, Medtr6g015795.1, Medtr6g016130.1, Medtr6g016140.1, Medtr6g016195.1, Medtr6g016200.1, Medtr6g016210.1, Medtr6g022500.1, Medtr6g037750.1, Medtr6g038190.1, Medtr6g038670.1, Medtr6g038700.1, Medtr6g038730.1, Medtr6g038760.1, Medtr6g038790.1, Medtr6g038790.2, Medtr6g038910.1, Medtr6g038930.1, Medtr6g038940.1, Medtr6g038980.1, Medtr6g039110.1, Medtr6g039180.1, Medtr6g051800.1, Medtr6g082820.1, Medtr6g082920.1, Medtr6g082940.1, Medtr6g082980.1, Medtr6g083000.1, Medtr6g083010.1, Medtr6g088320.1, Medtr6g090080.1, Medtr6g090605.1, Medtr6g471240.1, Medtr7g006870.1, Medtr7g009450.1, Medtr7g009470.1, Medtr7g009510.1, Medtr7g009540.1, Medtr7g009560.1, Medtr7g009570.1, Medtr7g009580.1, Medtr7g009790.1, Medtr7g010730.1, Medtr7g014430.1, Medtr7g023730.1, Medtr7g023740.1, Medtr7g056090.1, Medtr7g062610.1, Medtr7g062850.1, Medtr7g066210.1, Medtr7g066590.1, Medtr7g072460.1, Medtr7g078220.1, Medtr7g079550.1, Medtr7g082537.1, Medtr7g086800.1, Medtr7g092740.1, Medtr7g098240.1, Medtr7g098410.1, Medtr7g115260.1, Medtr7g117350.1, Medtr7g117350.2, Medtr8g007350.1, Medtr8g014840.1, Medtr8g023445.1, Medtr8g023510.1, Medtr8g027745.1, Medtr8g032890.1, Medtr8g032940.1, Medtr8g037890.1, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| MT | Medtr8g040910.1, Medtr8g040920.1, Medtr8g040925.1, Medtr8g041100.1, Medtr8g041150.1, Medtr8g041190.1, Medtr8g041870.1, Medtr8g043660.1, Medtr8g046150.1, Medtr8g046170.1, Medtr8g061960.1, Medtr8g067940.1, Medtr8g068030.1, Medtr8g068040.1, Medtr8g075510.1, Medtr8g075510.2, Medtr8g077105.1, Medtr8g081040.1, Medtr8g086590.1, Medtr8g088970.1, Medtr8g089000.1, Medtr8g095680.1, Medtr8g096900.1, Medtr8g099515.1, Medtr8g102220.1, Medtr8g104990.1, Medtr8g107140.1, Medtr8g469670.1, Medtr8g469820.1 |
| PV | Phvul.001G005000.1.p, Phvul.001G005000.2.p, Phvul.001G005100.1.p, Phvul.001G016500.1.p, Phvul.001G016600.1.p, Phvul.001G042800.1.p, Phvul.001G043000.1.p, Phvul.001G043450.1.p, Phvul.001G043901.1.p, Phvul.001G049300.1.p, Phvul.001G164400.1.p, Phvul.001G229800.1.p, Phvul.001G229800.2.p, Phvul.001G239200.1.p, Phvul.001G265300.1.p, Phvul.002G008400.1.p, Phvul.002G009000.1.p, Phvul.002G052000.1.p, Phvul.002G056100.1.p, Phvul.002G061200.1.p, Phvul.002G079900.1.p, Phvul.002G084400.1.p, Phvul.002G107800.1.p, Phvul.002G107900.1.p, Phvul.002G136200.1.p, Phvul.002G155400.1.p, Phvul.002G177500.1.p, Phvul.002G188800.1.p, Phvul.002G222400.1.p, Phvul.002G240100.1.p, Phvul.002G244000.1.p, Phvul.002G276504.1.p, Phvul.002G280200.1.p, Phvul.002G280300.1.p, Phvul.002G298300.1.p, Phvul.002G314200.1.p, Phvul.002G319900.1.p, Phvul.003G014400.1.p, Phvul.003G027500.1.p, Phvul.003G071700.1.p, Phvul.003G080900.1.p, Phvul.003G119000.1.p, Phvul.003G119000.2.p, Phvul.003G147300.1.p, Phvul.003G147400.1.p, Phvul.003G147500.1.p, Phvul.003G194900.1.p, Phvul.003G194900.2.p, Phvul.003G194900.3.p, Phvul.003G202100.1.p, Phvul.003G231900.1.p, Phvul.003G263400.1.p, Phvul.004G023600.1.p, Phvul.004G044800.1.p, Phvul.004G044900.1.p, Phvul.004G063000.3.p, Phvul.004G063000.4.p, Phvul.004G084600.1.p, Phvul.004G085076.1.p, Phvul.004G094094.1.p, Phvul.004G094094.2.p, Phvul.004G094094.3.p, Phvul.004G095900.1.p, Phvul.004G096200.1.p, Phvul.004G096400.1.p, Phvul.004G099100.1.p, Phvul.004G099400.1.p, Phvul.004G099500.1.p, Phvul.004G099600.1.p, Phvul.004G099900.1.p, Phvul.004G100000.1.p, Phvul.004G100200.1.p, Phvul.004G100300.1.p, Phvul.004G103000.1.p, Phvul.004G103200.1.p, Phvul.004G103300.1.p, Phvul.004G103600.1.p, Phvul.004G103700.1.p, Phvul.004G104700.1.p, Phvul.004G104900.1.p, Phvul.004G105100.1.p, Phvul.004G105400.1.p, Phvul.004G114700.1.p, Phvul.004G114800.1.p, Phvul.004G115100.1.p, Phvul.004G115200.1.p, Phvul.004G115400.1.p, Phvul.004G115600.1.p, Phvul.004G115700.1.p, Phvul.004G115700.2.p, Phvul.004G125100.2.p, Phvul.004G125100.3.p, Phvul.004G142500.1.p, Phvul.004G142600.1.p, Phvul.004G142700.1.p, Phvul.004G168700.1.p, Phvul.004G174100.1.p, Phvul.005G007100.1.p, Phvul.005G007200.2.p, Phvul.005G008200.1.p, Phvul.005G036600.1.p, Phvul.005G061160.1.p, Phvul.005G065700.1.p, Phvul.005G065700.2.p, Phvul.005G077100.1.p, Phvul.005G130900.1.p, Phvul.006G001500.1.p, Phvul.006G003400.2.p, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| PV | Phvul.006G033800.1.p, Phvul.006G112600.1.p, Phvul.006G147700.1.p, Phvul.006G165901.1.p, Phvul.006G167800.1.p, Phvul.006G174100.1.p, Phvul.006G196800.1.p, Phvul.007G038700.1.p, Phvul.007G045100.1.p, Phvul.007G049300.1.p, Phvul.007G050800.2.p, Phvul.007G051401.1.p, Phvul.007G052800.1.p, Phvul.007G070400.1.p, Phvul.007G077500.1.p, Phvul.007G087200.1.p, Phvul.007G087500.1.p, Phvul.007G087550.1.p, Phvul.007G087600.1.p, Phvul.007G087700.1.p, Phvul.007G087800.1.p, Phvul.007G102400.1.p, Phvul.007G117000.1.p, Phvul.007G167800.1.p, Phvul.007G189600.1.p, Phvul.007G231300.1.p, Phvul.007G246600.1.p, Phvul.007G255200.1.p, Phvul.008G043400.1.p, Phvul.008G055400.2.p, Phvul.008G060000.1.p, Phvul.008G069500.1.p, Phvul.008G077300.3.p, Phvul.008G077300.4.p, Phvul.008G077300.5.p, Phvul.008G077700.1.p, Phvul.008G079700.2.p, Phvul.008G080101.1.p, Phvul.008G082200.1.p, Phvul.008G082200.2.p, Phvul.008G087300.1.p, Phvul.008G093900.1.p, Phvul.008G094000.1.p, Phvul.008G095200.1.p, Phvul.008G096900.1.p, Phvul.008G108400.1.p, Phvul.008G108500.2.p, Phvul.008G109100.1.p, Phvul.008G109300.1.p, Phvul.008G109600.1.p, Phvul.008G111500.1.p, Phvul.008G111800.1.p, Phvul.008G118200.1.p, Phvul.008G154300.1.p, Phvul.008G163700.1.p, Phvul.008G173600.1.p, Phvul.008G188000.1.p, Phvul.008G210400.1.p, Phvul.008G210800.1.p, Phvul.008G210814.1.p, Phvul.008G237100.2.p, Phvul.008G273600.1.p, Phvul.008G273700.1.p, Phvul.008G273801.1.p, Phvul.008G273900.1.p, Phvul.008G274000.1.p, Phvul.008G274100.1.p, Phvul.008G276100.1.p,Phvul.008G277310.1.p, Phvul.008G277352.1.p, Phvul.008G281000.1.p, Phvul.008G281100.1.p, Phvul.009G054000.1.p, Phvul.009G063800.1.p, Phvul.009G067000.1.p, Phvul.009G176400.1.p, Phvul.009G183900.1.p, Phvul.009G233200.1.p, Phvul.010G006300.1.p, Phvul.010G013600.1.p, Phvul.010G013700.1.p, Phvul.010G014101.1.p, Phvul.010G073300.1.p, Phvul.010G079800.1.p, Phvul.010G088600.1.p, Phvul.010G089500.1.p, Phvul.010G113700.1.p, Phvul.011G005900.1.p, Phvul.011G034100.1.p, Phvul.011G034200.1.p, Phvul.011G042100.1.p, Phvul.011G065300.1.p, Phvul.011G107000.1.p, Phvul.011G115300.1.p, Phvul.011G175601.1.p, Phvul.011G194501.1.p, Phvul.L001744.1.p |
| VA | gi\|1044521676\|ref\|XP_017442734.1\|, gi\|1044521821\|ref\|XP_017442813.1\|, gi\|1044522243\|ref\|XP_017405386.1\|, gi\|1044522595\|ref\|XP_017405582.1\|, gi\|1044522700\|ref\|XP_017405640.1\|, gi\|1044523839\|ref\|XP_017406259.1\|, gi\|1044524752\|ref\|XP_017406757.1\|, gi\|1044525042\|ref\|XP_017406917.1\|, gi\|1044525701\|ref\|XP_017407276.1\|, gi\|1044525703\|ref\|XP_017407277.1\|, gi\|1044526502\|ref\|XP_017407715.1\|, gi\|1044526508\|ref\|XP_017407719.1\|, gi\|1044527844\|ref\|XP_017408440.1\|, gi\|1044527846\|ref\|XP_017408441.1\|, gi\|1044527927\|ref\|XP_017408485.1\|, gi\|1044528132\|ref\|XP_017408596.1\|, gi\|1044528134\|ref\|XP_017408597.1\|, gi\|1044528266\|ref\|XP_017408671.1\|, gi\|1044528893\|ref\|XP_017409012.1\|, gi\|1044529984\|ref\|XP_017409595.1\|, gi\|1044530000\|ref\|XP_017409604.1\|, gi\|1044532005\|ref\|XP_017410712.1\|, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044532409|ref|XP_017410936.1|, gi|1044532755|ref|XP_017411126.1|, gi|1044532757|ref|XP_017411128.1|, gi|1044532893|ref|XP_017411203.1|, gi|1044532897|ref|XP_017411205.1|, gi|1044533471|ref|XP_017411511.1|, gi|1044533791|ref|XP_017411687.1|, gi|1044533846|ref|XP_017411716.1|, gi|1044533939|ref|XP_017411768.1|, gi|1044534167|ref|XP_017411895.1|, gi|1044535974|ref|XP_017417244.1|, gi|1044536379|ref|XP_017420323.1|, gi|1044537017|ref|XP_017423554.1|, gi|1044537107|ref|XP_017423942.1|, gi|1044537109|ref|XP_017423955.1|, gi|1044537757|ref|XP_017427322.1|, gi|1044538182|ref|XP_017429299.1|, gi|1044538477|ref|XP_017430907.1|, gi|1044539568|ref|XP_017436516.1|, gi|1044539570|ref|XP_017436530.1|, gi|1044541778|ref|XP_017408774.1|, gi|1044542005|ref|XP_017409863.1|, gi|1044542007|ref|XP_017409872.1|, gi|1044542009|ref|XP_017409880.1|, gi|1044542011|ref|XP_017409889.1|, gi|1044542013|ref|XP_017409895.1|, gi|1044542682|ref|XP_017412068.1|, gi|1044542684|ref|XP_017412069.1|, gi|1044542686|ref|XP_017412070.1|, gi|1044543857|ref|XP_017437907.1|, gi|1044544145|ref|XP_017438078.1|, gi|1044544147|ref|XP_017438079.1|, gi|1044544247|ref|XP_017438142.1|, gi|1044544285|ref|XP_017438166.1|, gi|1044544287|ref|XP_017438167.1|, gi|1044544293|ref|XP_017438170.1|, gi|1044544299|ref|XP_017438175.1|, gi|1044544303|ref|XP_017438177.1|, gi|1044544311|ref|XP_017438182.1|, gi|1044544313|ref|XP_017438183.1|, gi|1044544319|ref|XP_017438187.1|, gi|1044544337|ref|XP_017438197.1|, gi|1044544339|ref|XP_017438198.1|, gi|1044544351|ref|XP_017438204.1|, gi|1044544363|ref|XP_017438212.1|, gi|1044544367|ref|XP_017438214.1|, gi|1044544369|ref|XP_017438215.1|, gi|1044544477|ref|XP_017438286.1|, gi|1044544581|ref|XP_017438351.1|, gi|1044545534|ref|XP_017438883.1|, gi|1044545779|ref|XP_017439017.1|, gi|1044545860|ref|XP_017439056.1|, gi|1044545876|ref|XP_017439065.1|, gi|1044545922|ref|XP_017439085.1|, gi|1044545924|ref|XP_017439086.1|, gi|1044546076|ref|XP_017439170.1|, gi|1044546173|ref|XP_017439223.1|, gi|1044546206|ref|XP_017439240.1|, gi|1044546557|ref|XP_017439432.1|, gi|1044546706|ref|XP_017439512.1|, gi|1044548498|ref|XP_017440457.1|, gi|1044549392|ref|XP_017440936.1|, gi|1044549410|ref|XP_017440945.1|, gi|1044549519|ref|XP_017441003.1|, gi|1044549577|ref|XP_017441039.1|, gi|1044549579|ref|XP_017441041.1|, gi|1044550267|ref|XP_017441454.1|, gi|1044550802|ref|XP_017441740.1|, gi|1044550888|ref|XP_017441788.1|, gi|1044550890|ref|XP_017441789.1|, gi|1044551149|ref|XP_017441930.1|, gi|1044551827|ref|XP_017442287.1|, gi|1044552368|ref|XP_017442579.1|, gi|1044553033|ref|XP_017412512.1|, gi|1044553035|ref|XP_017412513.1|, gi|1044553759|ref|XP_017412908.1|, gi|1044553787|ref|XP_017412923.1|, gi|1044553793|ref|XP_017412927.1|, gi|1044554161|ref|XP_017413150.1|, gi|1044554187|ref|XP_017413164.1|, gi|1044554189|ref|XP_017413165.1|, gi|1044554241|ref|XP_017413196.1|, gi|1044554595|ref|XP_017413397.1|, gi|1044554749|ref|XP_017413483.1|,</td></tr>
</table>

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Domain IDs |
|---|---|
| VA | gi\|1044555003\|ref\|XP_017413620.1\|, gi\|1044555139\|ref\|XP_017413692.1\|, gi\|1044555141\|ref\|XP_017413693.1\|, gi\|1044555374\|ref\|XP_017413819.1\|, gi\|1044555460\|ref\|XP_017413865.1\|, gi\|1044555462\|ref\|XP_017413866.1\|, gi\|1044555925\|ref\|XP_017414106.1\|, gi\|1044555978\|ref\|XP_017414135.1\|, gi\|1044556079\|ref\|XP_017414186.1\|, gi\|1044556515\|ref\|XP_017414403.1\|, gi\|1044556517\|ref\|XP_017414405.1\|, gi\|1044557101\|ref\|XP_017414694.1\|, gi\|1044558084\|ref\|XP_017415209.1\|, gi\|1044558086\|ref\|XP_017415211.1\|, gi\|1044558088\|ref\|XP_017415212.1\|, gi\|1044558435\|ref\|XP_017415387.1\|, gi\|1044558926\|ref\|XP_017415629.1\|, gi\|1044559300\|ref\|XP_017415829.1\|, gi\|1044559302\|ref\|XP_017415830.1\|, gi\|1044559304\|ref\|XP_017415831.1\|, gi\|1044560326\|ref\|XP_017416351.1\|, gi\|1044560328\|ref\|XP_017416352.1\|, gi\|1044560416\|ref\|XP_017416393.1\|, gi\|1044560494\|ref\|XP_017416435.1\|, gi\|1044560628\|ref\|XP_017416508.1\|, gi\|1044560976\|ref\|XP_017416701.1\|, gi\|1044561412\|ref\|XP_017416941.1\|, gi\|1044562050\|ref\|XP_017417271.1\|, gi\|1044562617\|ref\|XP_017417560.1\|, gi\|1044562918\|ref\|XP_017417717.1\|, gi\|1044562978\|ref\|XP_017417747.1\|, gi\|1044562980\|ref\|XP_017417748.1\|, gi\|1044562982\|ref\|XP_017417749.1\|, gi\|1044563202\|ref\|XP_017417863.1\|, gi\|1044563204\|ref\|XP_017417864.1\|, gi\|1044564193\|ref\|XP_017418371.1\|, gi\|1044564195\|ref\|XP_017418372.1\|, gi\|1044564197\|ref\|XP_017418373.1\|, gi\|1044564199\|ref\|XP_017418374.1\|, gi\|1044564201\|ref\|XP_017418375.1\|, gi\|1044565276\|ref\|XP_017418921.1\|, gi\|1044565897\|ref\|XP_017419252.1\|, gi\|1044565998\|ref\|XP_017419306.1\|, gi\|1044566757\|ref\|XP_017419701.1\|, gi\|1044566864\|ref\|XP_017419758.1\|, gi\|1044566870\|ref\|XP_017419762.1\|, gi\|1044566872\|ref\|XP_017419763.1\|, gi\|1044567977\|ref\|XP_017420366.1\|, gi\|1044568288\|ref\|XP_017420547.1\|, gi\|1044568915\|ref\|XP_017420886.1\|, gi\|1044569593\|ref\|XP_017421254.1\|, gi\|1044572074\|ref\|XP_017422568.1\|, gi\|1044573557\|ref\|XP_017423358.1\|, gi\|1044573559\|ref\|XP_017423359.1\|, gi\|1044574215\|ref\|XP_017423752.1\|, gi\|1044574549\|ref\|XP_017423934.1\|, gi\|1044575254\|ref\|XP_017424317.1\|, gi\|1044575298\|ref\|XP_017424338.1\|, gi\|1044578310\|ref\|XP_017425945.1\|, gi\|1044578492\|ref\|XP_017426043.1\|, gi\|1044579388\|ref\|XP_017426548.1\|, gi\|1044580242\|ref\|XP_017427004.1\|, gi\|1044580962\|ref\|XP_017427387.1\|, gi\|1044580964\|ref\|XP_017427388.1\|, gi\|1044581123\|ref\|XP_017427472.1\|, gi\|1044581588\|ref\|XP_017427726.1\|, gi\|1044581590\|ref\|XP_017427727.1\|, gi\|1044581592\|ref\|XP_017427728.1\|, gi\|1044581774\|ref\|XP_017427828.1\|, gi\|1044582174\|ref\|XP_017428038.1\|, gi\|1044582176\|ref\|XP_017428039.1\|, gi\|1044583021\|ref\|XP_017428486.1\|, gi\|1044583067\|ref\|XP_017428512.1\|, gi\|1044583069\|ref\|XP_017428513.1\|, gi\|1044583313\|ref\|XP_017428657.1\|, gi\|1044583493\|ref\|XP_017428770.1\|, gi\|1044583731\|ref\|XP_017428914.1\|, gi\|1044584877\|ref\|XP_017429529.1\|, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044584879|ref|XP_017429530.1|, gi|1044584881|ref|XP_017429531.1|, gi|1044584883|ref|XP_017429532.1|, gi|1044584885|ref|XP_017429533.1|, gi|1044584887|ref|XP_017429535.1|, gi|1044584895|ref|XP_017429539.1|, gi|1044584897|ref|XP_017429540.1|, gi|1044585245|ref|XP_017429730.1|, gi|1044586435|ref|XP_017430353.1|, gi|1044588628|ref|XP_017431526.1|, gi|1044589245|ref|XP_017431879.1|, gi|1044589902|ref|XP_017432262.1|, gi|1044590283|ref|XP_017432467.1|, gi|1044590373|ref|XP_017432515.1|, gi|1044590469|ref|XP_017432566.1|, gi|1044590625|ref|XP_017432651.1|, gi|1044594333|ref|XP_017434623.1|, gi|1044594505|ref|XP_017434722.1|, gi|1044594727|ref|XP_017434844.1|, gi|1044594963|ref|XP_017434972.1|, gi|1044595270|ref|XP_017435134.1|, gi|1044595528|ref|XP_017435265.1|, gi|1044595585|ref|XP_017435295.1|, gi|1044595623|ref|XP_017435315.1|, gi|1044598093|ref|XP_017436631.1|, gi|1044599592|ref|XP_017437442.1|</td></tr>
<tr><td>VR</td><td>XP_014489578.1, XP_014490586.1, XP_014490649.1, XP_014490664.1, XP_014491012.1, XP_014491013.1, XP_014491014.1, XP_014491070.1, XP_014491447.1, XP_014491828.1, XP_014491854.1, XP_014491855.1, XP_014491857.1, XP_014491994.1, XP_014492495.1, XP_014492538.1, XP_014492539.1, XP_014492711.1, XP_014492974.1, XP_014492975.1, XP_014492987.1, XP_014492991.1, XP_014493019.1, XP_014493268.1, XP_014493862.1, XP_014494622.1, XP_014495268.1, XP_014495379.1, XP_014495525.1, XP_014495526.1, XP_014496570.1, XP_014496571.1, XP_014496646.1, XP_014496654.1, XP_014496656.1, XP_014496820.1, XP_014496925.1, XP_014497447.1, XP_014497484.1, XP_014497700.1, XP_014497704.1, XP_014497909.1, XP_014498048.1, XP_014498313.1, XP_014498314.1, XP_014498524.1, XP_014498576.1, XP_014498856.1, XP_014498868.1, XP_014498882.1, XP_014498936.1, XP_014498989.1, XP_014499048.1, XP_014499060.1, XP_014499084.1, XP_014499107.1, XP_014499120.1, XP_014499491.1, XP_014499531.1, XP_014499545.1, XP_014499714.1, XP_014499737.1, XP_014499764.1, XP_014499804.1, XP_014499951.1, XP_014499963.1, XP_014499975.1, XP_014500238.1, XP_014500249.1, XP_014500295.1, XP_014500304.1, XP_014500544.1, XP_014500692.1, XP_014500735.1, XP_014500768.1, XP_014501010.1, XP_014501046.1, XP_014501332.1, XP_014501346.1, XP_014501359.1, XP_014501520.1, XP_014501593.1, XP_014501652.1, XP_014501699.1, XP_014501732.1, XP_014501918.1, XP_014501965.1, XP_014501976.1, XP_014501996.1, XP_014502103.1, XP_014502138.1, XP_014502151.1, XP_014502474.1, XP_014502899.1, XP_014502902.1, XP_014502903.1, XP_014502957.1, XP_014502958.1, XP_014502959.1, XP_014502960.1, XP_014503168.1, XP_014503208.1, XP_014503269.1, XP_014503355.1, XP_014503356.1, XP_014503378.1, XP_014503744.1, XP_014503779.1, XP_014503823.1, XP_014503824.1, XP_014503988.1, XP_014503989.1, XP_014503990.1, XP_014504002.1, XP_014504004.1, XP_014504005.1, XP_014504102.1, XP_014504221.1, XP_014504702.1, XP_014504714.1, XP_014504828.1, XP_014505105.1, XP_014505629.1, XP_014505844.1,</td></tr>
</table>

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VR | XP_014505845.1, XP_014506246.1, XP_014506249.1, XP_014506255.1, XP_014506401.1, XP_014506419.1, XP_014506453.1, XP_014506458.1, XP_014506465.1, XP_014506862.1, XP_014506975.1, XP_014507659.1, XP_014507660.1, XP_014507833.1, XP_014507927.1, XP_014508015.1, XP_014508043.1, XP_014508104.1, XP_014508324.1, XP_014508325.1, XP_014508661.1, XP_014508663.1, XP_014508664.1, XP_014508665.1, XP_014508666.1, XP_014508813.1, XP_014509368.1, XP_014509417.1, XP_014509519.1, XP_014510070.1, XP_014510807.1, XP_014511093.1, XP_014511460.1, XP_014511463.1, XP_014511501.1, XP_014511504.1, XP_014511509.1, XP_014511643.1, XP_014511648.1, XP_014511797.1, XP_014511873.1, XP_014511909.1, XP_014511910.1, XP_014511924.1, XP_014511925.1, XP_014511933.1, XP_014511934.1, XP_014511962.1, XP_014512130.1, XP_014512131.1, XP_014512271.1, XP_014512385.1, XP_014512468.1, XP_014512565.1, XP_014512695.1, XP_014512728.1, XP_014513297.1, XP_014513325.1, XP_014514200.1, XP_014514271.1, XP_014514882.1XP_014515704.1, XP_014516175.1, XP_014516602.1, XP_014516670.1, XP_014516671.1, XP_014517496.1, XP_014517662.1, XP_014517969.1, XP_014518077.1, XP_014518245.1, XP_014519279.1, XP_014519334.1, XP_014519335.1, XP_014519412.1, XP_014519414.1, XP_014519641.1, XP_014519642.1, XP_014519643.1, XP_014519677.1, XP_014520193.1, XP_014520276.1, XP_014520358.1, XP_014520500.1, XP_014520780.1, XP_014521052.1, XP_014521074.1, XP_014521099.1, XP_014521362.1, XP_014521367.1, XP_014521407.1, XP_014521426.1, XP_014521870.1, XP_014522458.1, XP_014522460.1, XP_014522470.1, XP_014522923.1, XP_014523046.1, XP_014523122.1, XP_014523123.1, XP_014523124.1, XP_014523155.1, XP_014523166.1, XP_014523167.1, XP_014523299.1, XP_014523323.1, XP_014523634.1, XP_014523728.1, XP_014523886.1, XP_014524389.1, XP_014524532.1 |
| VU | Vigun01g007400.1.p, Vigun01g008800.1.p, Vigun01g058400.1.p, Vigun01g083400.1.p, Vigun01g089100.1.p, Vigun01g089300.1.p, Vigun01g090000.1.p, Vigun01g146400.1.p, Vigun01g215000.1.p, Vigun01g223300.1.p, Vigun02g015600.1.p, Vigun02g037700.1.p, Vigun02g059900.1.p, Vigun02g068100.1.p, Vigun02g068200.1.p, Vigun02g081900.1.p, Vigun02g084900.1.p, Vigun02g095300.1.p, Vigun02g121800.1.p, Vigun02g121900.1.p, Vigun02g122100.1.p, Vigun02g122300.1.p, Vigun02g122400.1.p, Vigun02g122500.1.p, Vigun02g153600.1.p, Vigun02g154200.1.p, Vigun02g163200.1.p, Vigun02g168400.1.p, Vigun02g174300.1.p, Vigun02g202100.1.p, Vigun03g004100.1.p, Vigun03g013200.1.p, Vigun03g013200.2.p, Vigun03g019900.1.p, Vigun03g036000.1.p, Vigun03g054100.1.p, Vigun03g054100.2.p, Vigun03g054200.1.p, Vigun03g085100.1.p, Vigun03g097400.1.p, Vigun03g115000.1.p, Vigun03g152600.1.p, Vigun03g172300.1.p, Vigun03g190700.1.p, Vigun03g215200.1.p, Vigun03g227600.1.p, Vigun03g257700.1.p, Vigun03g290600.1.p, Vigun03g322700.1.p, Vigun03g322900.1.p, Vigun03g331200.1.p, Vigun03g331200.2.p, Vigun03g381000.1.p, Vigun03g381300.1.p, Vigun03g410000.1.p, Vigun03g444400.1.p, Vigun04g010200.1.p, Vigun04g017600.1.p, Vigun04g065400.1.p, Vigun04g066200.1.p, Vigun04g066400.1.p, Vigun04g093500.1.p, Vigun04g115800.1.p, Vigun04g116900.1.p, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VU | Vigun04g117000.1.p, Vigun04g117200.1.p, Vigun04g117300.1.p, Vigun04g117400.1.p, Vigun04g126400.1.p, Vigun04g126600.1.p, Vigun04g132900.1.p, Vigun04g133000.1.p, Vigun04g133100.1.p, Vigun04g133200.1.p, Vigun04g133300.1.p, Vigun04g133500.1.p, Vigun04g133600.1.p, Vigun04g133700.1.p, Vigun04g133800.1.p, Vigun04g133900.1.p, Vigun04g134000.1.p, Vigun04g134100.1.p, Vigun04g134200.1.p, Vigun04g134300.1.p, Vigun04g136300.1.p, Vigun04g136400.1.p, Vigun04g136500.1.p, Vigun04g136700.1.p, Vigun04g136800.1.p, Vigun04g136900.1.p, Vigun04g137100.1.p, Vigun04g137200.1.p, Vigun04g137300.1.p, Vigun04g137400.1.p, Vigun04g137500.1.p, Vigun04g137600.1.p, Vigun04g137800.1.p, Vigun04g137900.1.p, Vigun04g138000.1.p, Vigun04g138200.1.p, Vigun04g138400.1.p, Vigun04g138600.1.p, Vigun04g138800.1.p, Vigun04g138900.1.p, Vigun04g139000.1.p, Vigun04g139100.1.p, Vigun04g139400.1.p, Vigun04g139400.2.p, Vigun04g139500.1.p, Vigun04g139600.1.p, Vigun04g139700.1.p, Vigun04g139800.1.p, Vigun04g140000.1.p, Vigun04g142500.1.p, Vigun04g142500.2.p, Vigun04g142700.1.p, Vigun04g142900.1.p, Vigun04g143100.1.p, Vigun04g143200.1.p, Vigun04g143300.1.p, Vigun04g143500.1.p, Vigun04g143600.1.p, Vigun04g144600.1.p, Vigun04g144700.1.p, Vigun04g144900.1.p, Vigun04g145000.1.p, Vigun04g145100.1.p, Vigun04g145400.1.p, Vigun04g145500.1.p, Vigun04g146800.1.p, Vigun04g147000.1.p, Vigun04g154100.1.p, Vigun04g154200.1.p, Vigun04g154800.1.p, Vigun04g155100.1.p, Vigun04g155200.1.p, Vigun04g155500.1.p, Vigun04g155600.1.p, Vigun04g155700.1.p, Vigun04g156500.1.p, Vigun05g000400.1.p, Vigun05g000400.2.p, Vigun05g000400.3.p, Vigun05g044700.1.p, Vigun05g057100.1.p, Vigun05g057100.2.p, Vigun05g062200.1.p, Vigun05g073200.1.p, Vigun05g081100.3.p, Vigun05g081100.6.p, Vigun05g081200.4.p, Vigun05g081200.5.p, Vigun05g081200.6.p, Vigun05g083300.1.p, Vigun05g083400.1.p, Vigun05g083500.1.p, Vigun05g085700.1.p, Vigun05g090000.1.p, Vigun05g096400.1.p, Vigun05g097600.1.p, Vigun05g097800.1.p, Vigun05g097800.2.p, Vigun05g099000.1.p, Vigun05g099100.1.p, Vigun05g106100.1.p, Vigun05g106100.2.p, Vigun05g106200.1.p, Vigun05g106800.1.p, Vigun05g107000.1.p, Vigun05g107200.1.p, Vigun05g107200.2.p, Vigun05g107300.1.p, Vigun05g107400.1.p, Vigun05g107600.1.p, Vigun05g107600.2.p, Vigun05g107600.3.p, Vigun05g109300.1.p, Vigun05g109400.1.p, Vigun05g132500.1.p, Vigun05g139400.1.p, Vigun05g155800.1.p, Vigun05g165300.1.p, Vigun05g165400.1.p, Vigun05g178900.1.p, Vigun05g179200.1.p, Vigun05g185100.1.p, Vigun05g277900.3.p, Vigun05g278700.2.p, Vigun05g295400.1.p, Vigun06g041200.1.p, Vigun06g123400.1.p, Vigun06g159500.1.p, Vigun06g179000.1.p, Vigun06g180600.1.p, Vigun06g212800.1.p, Vigun07g030700.1.p, Vigun07g039700.1.p, Vigun07g048400.1.p, Vigun07g049200.1.p, Vigun07g055700.1.p, Vigun07g055800.1.p, Vigun07g057900.1.p, Vigun07g082700.1.p, Vigun07g082800.1.p, Vigun07g092000.1.p, Vigun07g114600.1.p, Vigun07g157900.1.p, Vigun07g181500.1.p, Vigun07g185300.1.p, Vigun07g195200.1.p, Vigun07g209100.1.p, Vigun07g209300.1.p, Vigun07g219600.1.p, Vigun07g219700.1.p, Vigun07g226300.1.p, Vigun07g226400.1.p, Vigun07g226600.1.p, Vigun07g244200.1.p, Vigun07g246200.1.p, Vigun07g246300.1.p, Vigun07g248300.1.p, Vigun07g248400.1.p, Vigun07g252200.1.p, Vigun07g253700.1.p, Vigun08g005100.1.p, Vigun08g018100.1.p, Vigun08g018200.1.p, Vigun08g046100.1.p, Vigun08g046200.1.p, Vigun08g046300.1.p, Vigun08g046400.1.p, Vigun08g046500.1.p, Vigun08g069800.1.p, Vigun08g072300.1.p, Vigun08g114500.1.p, Vigun08g114500.2.p, Vigun08g143800.1.p, Vigun08g143800.2.p, Vigun08g205600.1.p, Vigun08g205700.1.p, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VU | Vigun08g207900.1.p, Vigun08g208100.1.p, Vigun09g032600.1.p, Vigun09g079800.1.p, Vigun09g088600.1.p, Vigun09g137400.1.p, Vigun09g150900.1.p, Vigun09g150900.2.p, Vigun09g236100.1.p, Vigun09g239600.1.p, Vigun09g239600.2.p, Vigun09g239600.3.p, Vigun09g243800.1.p, Vigun09g251100.1.p, Vigun10g054700.1.p, Vigun10g054900.1.p, Vigun10g055000.1.p, Vigun10g055000.2.p, Vigun10g055000.3.p, Vigun10g055200.1.p, Vigun10g055300.1.p, Vigun10g055400.1.p, Vigun10g071300.1.p, Vigun10g071300.2.p, Vigun10g081300.1.p, Vigun10g081400.1.p, Vigun10g082500.1.p, Vigun10g082900.1.p, Vigun10g083000.1.p, Vigun10g083200.1.p, Vigun10g084200.1.p, Vigun10g084200.2.p, Vigun10g084200.3.p, Vigun10g106800.1.p, Vigun10g106800.2.p, Vigun10g106800.3.p, Vigun10g117800.1.p, Vigun10g118700.1.p, Vigun10g145600.1.p, Vigun11g039000.1.p, Vigun11g040000.1.p, Vigun11g096700.1.p, Vigun11g154200.1.p, Vigun11g178400.1.p, Vigun11g189000.1.p |
| VV | GSVIVT01000416001, GSVIVT01001120001, GSVIVT01001407001, GSVIVT01001412001, GSVIVT01001418001, GSVIVT01001697001, GSVIVT01001802001, GSVIVT01001803001, GSVIVT01002369001, GSVIVT01002803001, GSVIVT01003682001, GSVIVT01003865001, GSVIVT01005164001, GSVIVT01005167001, GSVIVT01005169001, GSVIVT01005680001, GSVIVT01005786001, GSVIVT01005906001, GSVIVT01006040001, GSVIVT01006297001, GSVIVT01007390001, GSVIVT01008279001, GSVIVT01009981001, GSVIVT01010391001, GSVIVT01011042001, GSVIVT01011424001, GSVIVT01011675001, GSVIVT01011948001, GSVIVT01012224001, GSVIVT01012235001, GSVIVT01012274001, GSVIVT01012680001, GSVIVT01012806001, GSVIVT01012819001, GSVIVT01013055001, GSVIVT01013340001, GSVIVT01013353001, GSVIVT01013974001, GSVIVT01013995001, GSVIVT01014120001, GSVIVT01014414001, GSVIVT01014856001, GSVIVT01015345001, GSVIVT01015659001, GSVIVT01015897001, GSVIVT01015972001, GSVIVT01016000001, GSVIVT01016012001, GSVIVT01016019001, GSVIVT01017385001, GSVIVT01017390001, GSVIVT01017396001, GSVIVT01017399001, GSVIVT01017416001, GSVIVT01017419001, GSVIVT01017641001, GSVIVT01018463001, GSVIVT01018682001, GSVIVT01018692001, GSVIVT01018767001, GSVIVT01018769001, GSVIVT01018835001, GSVIVT01019130001, GSVIVT01019530001, GSVIVT01019836001, GSVIVT01019838001, GSVIVT01019841001, GSVIVT01019973001, GSVIVT01020111001, GSVIVT01020334001, GSVIVT01020771001, GSVIVT01020793001 |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VV | GSVIVT01020945001, GSVIVT01021521001, GSVIVT01021556001, GSVIVT01021675001, GSVIVT01023937001, GSVIVT01024050001, GSVIVT01024424001, GSVIVT01024911001, GSVIVT01025060001, GSVIVT01025348001, GSVIVT01025349001, GSVIVT01025477001, GSVIVT01025479001, GSVIVT01025484001, GSVIVT01025825001, GSVIVT01025885001, GSVIVT01025912001, GSVIVT01026074001, GSVIVT01026086001, GSVIVT01026633001, GSVIVT01026990001, GSVIVT01027003001, GSVIVT01027094001, GSVIVT01027526001, GSVIVT01027651001, GSVIVT01027913001, GSVIVT01028413001, GSVIVT01028764001, GSVIVT01028766001, GSVIVT01028773001, GSVIVT01028799001, GSVIVT01028922001, GSVIVT01028924001, GSVIVT01028928001, GSVIVT01028966001, GSVIVT01028967001, GSVIVT01028970001, GSVIVT01029278001, GSVIVT01029646001, GSVIVT01029712001, GSVIVT01030062001, GSVIVT01030071001, GSVIVT01030223001, GSVIVT01030308001, GSVIVT01030335001, GSVIVT01030338001, GSVIVT01030341001, GSVIVT01030345001, GSVIVT01030353001, GSVIVT01030637001, GSVIVT01030681001, GSVIVT01030915001, GSVIVT01031187001, GSVIVT01031843001, GSVIVT01032973001, GSVIVT01033116001, GSVIVT01033182001, GSVIVT01033488001, GSVIVT01034031001, GSVIVT01034086001, GSVIVT01034124001, GSVIVT01034126001, GSVIVT01034283001, GSVIVT01034545001, GSVIVT01034995001, GSVIVT01035314001, GSVIVT01035859001, GSVIVT01035937001, GSVIVT01036442001, GSVIVT01036793001, GSVIVT01037018001, GSVIVT01037075001, GSVIVT01037079001, GSVIVT01037080001, GSVIVT01037155001, GSVIVT01037158001, GSVIVT01037845001, GSVIVT01038307001, GSVIVT01038316001, GSVIVT01038331001, GSVIVT01038458001, GSVIVT01038464001, GSVIVT01038679001, GSVIVT01038702001, GSVIVT01038717001 |
| AT | AT1G04520.1, AT1G07390.1, AT1G07390.2, AT1G07390.3, AT1G08450.1, AT1G08450.2, AT1G08450.3, AT1G09210.1, AT1G11545.1, AT1G11915.1, AT1G12240.1, AT1G13910.1, AT1G16090.1, AT1G16905.1, AT1G17240.1, AT1G17250.1, AT1G18250.1, AT1G18250.2, AT1G19320.1, AT1G21880.2, AT1G24485.3, AT1G25570.1, AT1G27120.1, AT1G34060.1, AT1G34290.1, AT1G45616.1, AT1G47890.1, AT1G54470.1, AT1G54470.2, AT1G55120.2, AT1G55660.1, AT1G55660.2, AT1G56610.1, AT1G58190.1, AT1G58190.2, AT1G61750.1, AT1G62660.1, AT1G63550.1, AT1G65310.1, AT1G65380.1, AT1G66940.1, AT1G66940.2, AT1G66940.3, AT1G70690.1, AT1G71390.1, AT1G71400.1, AT1G74170.1, AT1G74180.1, AT1G74190.1, AT1G74800.1, AT1G75030.1, AT1G75040.1, AT1G75050.1, AT1G77700.1, AT1G78230.1, AT1G78830.1, AT1G80080.1, AT2G01660.1, AT2G01660.2, AT2G01850.1, AT2G06850.1, AT2G14620.1, AT2G14720.1, AT2G14720.2, AT2G14740.1, AT2G14740.2, AT2G15042.1, AT2G15080.1, AT2G15080.2, AT2G17120.1, AT2G17860.1, AT2G18800.1, AT2G19780.1, AT2G24810.1, AT2G25440.1, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| AT | AT2G25470.1, AT2G26860.2, AT2G32680.1, AT2G33020.1, AT2G33050.1, AT2G33060.1, AT2G33330.1, AT2G46494.1, AT2G46495.1, AT3G04370.1, AT3G04370.2, AT3G05360.1, AT3G05370.1, AT3G05650.1, AT3G05660.1, AT3G09035.1, AT3G09190.1, AT3G11010.1, AT3G11080.1, AT3G13790.1, AT3G13790.2, AT3G21900.1, AT3G21910.1, AT3G21940.1, AT3G21960.2, AT3G21980.1, AT3G21990.1, AT3G22010.1, AT3G22040.1, AT3G22060.1, AT3G22800.1, AT3G23010.1, AT3G23110.1, AT3G23120.1, AT3G24480.1, AT3G24900.1, AT3G24982.1, AT3G28890.1, AT3G28890.2, AT3G46240.1, AT3G46270.1, AT3G46280.1, AT3G49750.1, AT3G51710.1, AT3G53240.1, AT3G54080.1, AT3G59510.1, AT3G60720.1, AT4G00300.1, AT4G00300.2, AT4G03010.1, AT4G11650.1, AT4G13340.1, AT4G13810.1, AT4G13810.2, AT4G13880.1, AT4G13920.1, AT4G14130.1, AT4G18670.1, AT4G18760.1, AT4G20670.1, AT4G21060.1, AT4G21060.2, AT4G24180.1, AT4G25820.1, AT4G29240.1, AT4G30280.1, AT4G36010.1, AT4G36010.2, AT4G38660.1, AT5G01090.1, AT5G07340.1, AT5G07340.2, AT5G12940.1, AT5G18470.1, AT5G24620.1, AT5G25910.1, AT5G27060.1, AT5G37660.1, AT5G37660.2, AT5G38660.1, AT5G38660.2, AT5G40170.1, AT5G43980.1, AT5G45770.1, AT5G49290.1, AT5G53110.1, AT5G57530.1, AT5G57550.1, AT5G61790.1, AT5G62620.1, AT5G62620.2, AT5G65730.1, AT5G65830.1 |
| SL | Solyc01g005120.2.1, Solyc01g005720.2.1, Solyc01g005730.2.1, Solyc01g005760.2.1, Solyc01g005780.1.1, Solyc01g006550.2.1, Solyc01g008390.1.1, Solyc01g008410.1.1, Solyc01g009690.1.1, Solyc01g009700.1.1, Solyc01g009930.1.1, Solyc01g087510.1.1, Solyc01g088500.2.1, Solyc01g091040.2.1, Solyc01g094920.2.1, Solyc01g095810.2.1, Solyc01g095830.2.1, Solyc01g095840.1.1, Solyc01g095850.1.1, Solyc01g095860.2.1, Solyc01g098370.1.1, Solyc01g098680.2.1, Solyc01g098690.2.1, Solyc01g099250.1.1, Solyc01g100380.2.1, Solyc01g102870.1.1, Solyc01g106500.2.1, Solyc01g106650.2.1, Solyc01g107130.2.1, Solyc01g107670.2.1, Solyc01g111330.2.1, Solyc02g062190.1.1, Solyc02g067770.2.1, Solyc02g070400.2.1, Solyc02g072250.1.1, Solyc02g080160.2.1, Solyc02g080900.2.1, Solyc02g083760.2.1, Solyc02g083790.2.1, Solyc02g091920.2.1, Solyc03g005690.2.1, Solyc03g026360.1.1, Solyc03g026370.1.1, Solyc03g031550.1.1, Solyc03g031800.2.1, Solyc03g033490.1.1, Solyc03g034170.2.1, Solyc03g063750.2.1, Solyc03g071520.1.1, Solyc03g082780.1.1, Solyc03g083480.2.1, Solyc03g083510.2.1, Solyc03g083910.2.1, Solyc03g093080.2.1, Solyc03g093120.2.1, Solyc03g111530.2.1, Solyc03g112680.1.1, Solyc03g113520.2.1, Solyc03g118040.2.1, Solyc03g118780.2.1, Solyc04g007310.1.1, Solyc04g007950.2.1, Solyc04g026140.1.1, Solyc04g056640.1.1, Solyc04g076450.1.1, Solyc04g079890.2.1, Solyc04g081550.2.1, Solyc04g081560.2.1, Solyc05g005050.2.1, Solyc05g005680.2.1, Solyc05g006680.2.1, Solyc05g006770.2.1, Solyc05g007930.2.1, Solyc05g009000.2.1, Solyc05g009800.2.1, Solyc05g026240.1.1, Solyc05g046290.2.1, Solyc05g050700.1.1, Solyc05g051270.2.1, Solyc05g053280.2.1, Solyc05g055190.1.1, Solyc05g055930.1.1, Solyc05g056230.2.1, Solyc06g008270.2.1, Solyc06g008300.2.1, Solyc06g033920.1.1, Solyc06g068700.2.1, Solyc06g071160.2.1, Solyc06g073000.2.1, Solyc06g083400.1.1, Solyc06g083750.2.1, Solyc07g005150.1.1, Solyc07g006850.1.1, Solyc07g008590.1.1, Solyc07g008600.1.1, Solyc07g008620.1.1, Solyc07g008630.1.1, Solyc07g008640.1.1, Solyc07g009380.2.1, Solyc07g015990.1.1, |

**Table A5. Protein IDs of RLP predicted among legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| SL | Solyc07g054710.1.1, Solyc07g055820.1.1, Solyc07g055990.2.1, Solyc08g016210.2.1, Solyc08g016270.1.1, Solyc08g076080.2.1, Solyc08g077740.1.1, Solyc08g079080.2.1, Solyc08g080620.1.1, Solyc08g080640.1.1, Solyc08g080670.1.1, Solyc08g082310.2.1, Solyc09g005080.1.1, Solyc09g005090.1.1, Solyc09g010090.2.1, Solyc09g011840.1.1, Solyc09g015200.1.1, Solyc09g015840.2.1, Solyc09g018490.2.1, Solyc09g057960.1.1, Solyc09g072630.2.1, Solyc09g082530.1.1, Solyc09g083200.2.1, Solyc09g090680.2.1, Solyc09g092520.2.1, Solyc10g050470.1.1, Solyc10g076490.1.1, Solyc10g083290.1.1, Solyc10g083300.1.1, Solyc10g084840.1.1, Solyc10g085360.1.1, Solyc11g011000.1.1, Solyc11g011180.1.1, Solyc11g013300.1.1, Solyc12g005620.1.1, Solyc12g006020.1.1, Solyc12g007250.1.1, Solyc12g007260.1.1, Solyc12g007270.1.1, Solyc12g009510.1.1, Solyc12g009520.1.1, Solyc12g009690.1.1, Solyc12g009720.1.1, Solyc12g009730.1.1, Solyc12g009740.1.1, Solyc12g009750.1.1, Solyc12g009770.1.1, Solyc12g009780.1.1, Solyc12g013680.1.1, Solyc12g013740.1.1, Solyc12g042760.1.1, Solyc12g049190.1.1, Solyc12g056360.1.1, Solyc12g099870.1.1, Solyc12g099980.1.1, Solyc12g100030.1.1 |

SP: Species

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes**.

| SP | Protein IDs |
|---|---|
| CC | KYP51453.1, KYP50432.1, KYP37606.1, KYP63624.1, KYP59535.1, KYP48687.1, KYP32335.1, KYP37049.1, KYP75389.1, KYP75388.1, KYP55693.1, KYP50158.1, KYP76928.1, KYP76668.1, KYP55692.1, KYP51907.1, KYP40903.1, KYP66135.1, KYP46138.1, KYP34144.1, KYP34018.1, KYP74831.1, KYP55186.1, KYP50005.1, KYP47214.1, KYP47204.1, KYP45644.1, KYP45645.1, KYP33714.1, KYP32785.1, KYP61784.1, KYP56802.1, KYP38183.1, KYP58767.1, KYP50310.1, KYP57042.1, KYP37300.1, KYP75139.1, KYP72155.1, KYP75140.1, KYP72152.1, KYP69721.1, KYP63285.1, KYP34388.1, KYP46128.1, KYP46649.1, KYP66886.1, KYP58768.1, KYP55455.1, KYP40783.1, KYP51141.1, KYP33015.1, KYP60951.1, KYP62748.1, KYP38129.1, KYP63286.1, KYP45708.1, KYP33136.1, KYP57429.1, KYP70523.1, KYP50376.1 |
| GM | Glyma.14G060300.1.p, Glyma.05G128200.1.p, Glyma.14G060400.1.p, Glyma.01G125200.1.p, Glyma.03G049300.1.p, Glyma.20G152800.1.p, Glyma.10G241500.1.p, Glyma.08G128800.1.p, Glyma.08G083300.1.p, Glyma.08G044300.1.p, Glyma.08G128900.1.p, Glyma.05G170600.1.p, Glyma.13G371200.2.p, Glyma.13G371200.1.p, Glyma.17G071700.1.p, Glyma.02G206300.1.p, Glyma.18G199200.1.p, Glyma.07G147800.2.p, Glyma.07G147800.1.p, Glyma.07G155800.1.p, Glyma.18G206400.1.p, Glyma.10G270600.1.p, Glyma.05G126400.2.p, Glyma.05G126400.1.p, Glyma.03G020700.1.p, Glyma.13G370700.1.p, Glyma.15G002900.1.p, Glyma.07G082600.1.p, Glyma.09G150100.4.p, Glyma.09G150100.3.p, Glyma.09G150100.2.p, Glyma.09G150100.1.p, Glyma.12G240300.1.p, Glyma.08G355200.1.p, Glyma.08G355000.1.p, Glyma.08G355100.1.p, Glyma.08G354900.1.p, Glyma.11G036700.1.p, Glyma.03G090200.1.p, Glyma.03G003000.1.p, Glyma.03G090500.1.p, Glyma.03G002900.1.p, Glyma.03G002700.1.p, Glyma.02G000400.1.p, Glyma.13G268100.2.p, Glyma.13G268100.1.p, Glyma.07G080700.1.p, Glyma.07G108400.1.p, Glyma.09G003100.1.p, Glyma.01G090500.1.p, Glyma.01G083200.1.p, Glyma.01G206500.1.p, Glyma.05G071100.1.p, Glyma.20G246600.1.p, Glyma.03G096600.1.p, Glyma.15G258400.1.p, Glyma.15G202700.1.p, Glyma.09G050300.1.p, Glyma.09G217500.1.p, Glyma.09G216400.2.p, Glyma.09G216400.1.p, Glyma.09G217200.1.p, Glyma.16G034900.1.p, Glyma.16G157400.1.p, Glyma.20G164200.1.p, Glyma.02G075600.1.p, Glyma.07G173000.1.p, Glyma.07G069000.1.p, Glyma.10G228900.1.p, Glyma.06G110700.1.p, Glyma.19G064600.1.p, Glyma.10G228900.3.p, Glyma.10G228900.2.p, Glyma.20G118900.2.p, Glyma.20G118900.3.p, Glyma.20G118900.1.p, Glyma.20G118500.1.p, Glyma.20G118300.1.p, Glyma.20G118600.1.p, Glyma.19G069300.2.p, Glyma.19G069300.1.p, Glyma.02G100300.1.p, Glyma.13G354400.2.p, Glyma.13G354400.1.p, Glyma.13G033000.4.p, Glyma.13G033000.3.p, Glyma.13G033000.5.p, Glyma.13G033000.1.p, Glyma.13G033500.2.p, Glyma.13G033500.1.p, Glyma.13G032600.2.p, Glyma.13G032600.1.p, Glyma.13G033400.1.p, |

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.07G094200.1.p, Glyma.07G094100.1.p, Glyma.07G095500.1.p, Glyma.07G094500.2.p, Glyma.07G094500.1.p, Glyma.17G214600.8.p, Glyma.17G214600.9.p, Glyma.17G214600.1.p, Glyma.17G214100.1.p, Glyma.17G214900.1.p, Glyma.17G214400.4.p, Glyma.17G214400.1.p, Glyma.08G044200.1.p, Glyma.14G113200.1.p, Glyma.14G118800.1.p, Glyma.05G237100.1.p, Glyma.20G119000.1.p, Glyma.20G118700.1.p, Glyma.13G033800.1.p, Glyma.13G033400.2.p, Glyma.07G095800.2.p, Glyma.07G095400.1.p, Glyma.07G094500.6.p, Glyma.07G094500.4.p, Glyma.07G094500.3.p, Glyma.17G214600.2.p, Glyma.17G214600.3.p, Glyma.17G214600.7.p, Glyma.17G214600.6.p, Glyma.17G214600.5.p, Glyma.17G214600.4.p, Glyma.17G214100.4.p, Glyma.17G214100.3.p, Glyma.17G214900.2.p, Glyma.17G214900.3.p, Glyma.17G214400.2.p, Glyma.17G214400.3.p, Glyma.10G271600.1.p, Glyma.18G142100.1.p, Glyma.06G043200.1.p, Glyma.04G042400.1.p, Glyma.04G042500.1.p, Glyma.13G139800.1.p, Glyma.U027000.1.p, Glyma.13G054200.1.p, Glyma.10G052600.1.p, Glyma.06G300600.1.p, Glyma.12G104100.1.p, Glyma.12G198600.1.p, Glyma.08G307400.1.p, Glyma.08G307300.1.p, Glyma.20G173600.1.p, Glyma.13G365400.2.p, Glyma.13G365400.1.p, Glyma.15G007800.2.p, Glyma.15G007800.1.p, Glyma.07G108200.1.p, Glyma.18G110400.1.p, Glyma.12G198600.2.p, Glyma.12G198600.3.p, Glyma.08G307400.2.p, Glyma.06G134700.1.p, Glyma.12G198800.1.p, Glyma.13G303400.1.p, Glyma.06G068100.1.p, Glyma.14G116500.1.p, Glyma.04G066700.1.pGlyma.17G211200.1.p, Glyma.13G127100.1.p, Glyma.13G127100.2.p, Glyma.16G130400.1.p, Glyma.02G050300.1.p, Glyma.15G019800.3.p, Glyma.15G019800.2.p, Glyma.15G019800.4.p, Glyma.15G019800.1.p, Glyma.06G026600.1.p, Glyma.01G168100.1.p, Glyma.11G075200.1.p, Glyma.04G026700.1.p, Glyma.15G161600.1.p, Glyma.15G162000.1.p, Glyma.15G161500.1.p, Glyma.09G055900.2.p, Glyma.09G055900.1.p, Glyma.19G033100.1.p, Glyma.07G156000.1.p, Glyma.07G095100.1.p, Glyma.13G224300.22.p, Glyma.04G230500.1.p, Glyma.09G097500.1.p, Glyma.09G097500.2.p, Glyma.11G063200.1.p, Glyma.02G059700.1.p, Glyma.01G179000.1.p, Glyma.15G161700.1.p, Glyma.14G100700.1.p, Glyma.06G215600.1.p, Glyma.05G009900.1.p, Glyma.04G149800.1.p, Glyma.17G117800.1.p, Glyma.19G069400.1.p, Glyma.18G073600.2.p, Glyma.18G073600.1.p, Glyma.08G333200.1.p, Glyma.14G011500.1.p, Glyma.02G302600.1.p, Glyma.14G118700.1.p, Glyma.02G100400.1.p, Glyma.13G033900.1.p, Glyma.13G032700.1.p, Glyma.13G033200.1.p, Glyma.17G214200.1.p, Glyma.17G214500.1.p, Glyma.17G214700.1.p, Glyma.17G214800.1.p, Glyma.04G230300.1.p, Glyma.15G203600.1.p, Glyma.03G137300.1.p, Glyma.19G140100.1.p, Glyma.19G140100.2.p, Glyma.14G195300.1.p, Glyma.02G228300.1.p, Glyma.06G216800.2.p, Glyma.13G228700.2.p, Glyma.13G228700.3.p, |

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| GM | Glyma.13G228700.4.p, Glyma.13G228700.5.p, Glyma.13G228700.6.p, Glyma.13G228700.1.p |
| MT | Medtr5g024450.1, Medtr4g094610.1, Medtr8g068540.1, Medtr8g089210.1, Medtr3g079850.2, Medtr3g079850.1, Medtr5g025860.1, Medtr8g089200.1, Medtr5g082460.1, Medtr6g036780.1, Medtr1g029950.1, Medtr5g082270.1, Medtr5g082920.1, Medtr5g025890.1, Medtr5g025950.1, Medtr7g067530.1, Medtr4g114270.1, Medtr4g094858.1, Medtr6g036870.1, Medtr1g088930.1, Medtr1g088940.1, Medtr5g019070.1, Medtr5g026760.1, Medtr5g026000.1, Medtr5g013130.1, Medtr5g026010.1, Medtr5g025850.1, Medtr5g013070.1, Medtr5g025930.1, Medtr7g058550.1, Medtr7g058860.1, Medtr7g058830.1, Medtr7g053050.1, Medtr7g058530.1, Medtr8g011370.1, Medtr8g465340.1, Medtr2g067990.1, Medtr2g067980.1, Medtr3g031480.1, Medtr3g070220.1, Medtr3g031490.1, Medtr3g031580.1, Medtr3g031640.1, Medtr3g031600.1, Medtr3g031610.1, Medtr3g031470.1, Medtr4g029710.1, Medtr6g036790.1, Medtr6g040210.1, Medtr6g036890.1, Medtr6g036840.1, Medtr1g088935.1, Medtr1g088940.2, Medtr5g026160.1, Medtr5g026200.1, Medtr5g044680.2, Medtr5g044680.1, Medtr5g026150.1, Medtr5g025180.1, Medtr5g026090.1, Medtr0070s0020.1, Medtr8g470400.1, Medtr8g470950.1, Medtr8g469650.1, Medtr8g470560.1, Medtr8g469780.1, Medtr8g470370.1, Medtr8g469600.1, Medtr8g011440.1, Medtr2g016530.1, Medtr2g067970.1, Medtr2g040910.1, Medtr2g068650.1, Medtr3g449540.1, Medtr3g031500.1, Medtr5g082290.1, Medtr5g082420.1, Medtr4g114250.1, Medtr5g006160.1, Medtr2g103810.1, Medtr3g102400.1, Medtr1g029940.1, Medtr1g104555.1, Medtr6g043510.1, Medtr6g043790.1, Medtr1g032930.1, Medtr1g109580.1, Medtr7g070200.1, Medtr8g030500.1, Medtr2g089440.1, Medtr3g072800.1, Medtr1g099400.1, Medtr2g089440.2, Medtr8g066700.1, Medtr4g014350.1, Medtr8g035560.1, Medtr2g105260.1, Medtr6g082870.1, Medtr6g083760.1, Medtr1g031520.1, Medtr1g031540.1, Medtr1g031510.1, Medtr1g027990.1, Medtr1g028280.1, Medtr1g028280.2, Medtr1g031280.1, Medtr1g027440.1, Medtr1g028290.1, Medtr1g031560.1, Medtr8g104520.1, Medtr4g073220.1, Medtr6g045030.1, Medtr6g083780.1, Medtr6g082950.1, Medtr1g031540.2, Medtr1g027680.1, Medtr1g031200.1, Medtr1g027410.1, Medtr1g027540.1, Medtr1g028100.1, Medtr1g110230.1, Medtr1g110260.1, Medtr1g028170.1, Medtr1g028020.1, Medtr5g025840.1, Medtr3g092420.1, Medtr8g469830.1, Medtr8g469980.1, Medtr2g016580.1, Medtr1g012550.1, Medtr0280s0040.1, Medtr2g080080.1, Medtr2g089360.1, Medtr2g073250.1, Medtr3g107070.1, Medtr0007s0390.1, Medtr1g031580.1, Medtr2g072640.1, Medtr2g072620.1, Medtr4g073230.1, Medtr5g042440.1, Medtr1g064560.1, Medtr4g044393.1, Medtr5g021670.1, Medtr7g082430.1, Medtr7g082510.1, Medtr7g082110.1, Medtr7g082460.1, Medtr3g113140.1, Medtr7g082470.1, Medtr7g082300.1, Medtr5g096530.1, Medtr3g092390.1, Medtr5g087360.1, Medtr5g087360.2, Medtr2g080090.1, Medtr2g080100.1, Medtr6g445600.1, Medtr1g027150.1, Medtr2g054870.1, |

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| MT | Medtr2g036460.1, Medtr2g036490.1, Medtr2g036500.1, Medtr2g036440.1, Medtr2g036430.1, Medtr5g019050.1, Medtr1g027960.1, Medtr1g028220.1, Medtr1g027500.1, Medtr1g027040.1, Medtr1g027820.1, Medtr1g026940.1, Medtr1g027970.1, Medtr1g028130.1, Medtr1g027740.1, Medtr1g027690.1, Medtr1g027420.1, Medtr3g028650.1, Medtr4g113710.1, Medtr1g027890.1, Medtr1g027200.1, Medtr6g012810.2, Medtr6g012810.1, Medtr1g115485.1, Medtr1g021845.1, Medtr3g060880.1, Medtr1g028080.1, Medtr1g027370.1, Medtr7g091680.1, Medtr2g029010.1, Medtr2g449790.1, Medtr4g094615.3, Medtr3g071480.1, Medtr6g040230.1 |
| PV | Phvul.002G318200.1.p, Phvul.002G215100.1.p, Phvul.002G251400.1.p, Phvul.008G093000.1.p, Phvul.008G224900.1.p, Phvul.008G225000.1.p, Phvul.008G093200.3.p, Phvul.008G093200.2.p, Phvul.008G105600.1.p, Phvul.008G118500.1.p, Phvul.008G093200.1.p, Phvul.002G055600.1.p, Phvul.002G000200.1.p, Phvul.008G124300.1.p, Phvul.011G214400.1.p, Phvul.010G057000.1.p, Phvul.010G057300.1.p, Phvul.010G057600.1.p, Phvul.010G057100.1.p, Phvul.010G057500.1.p, Phvul.010G031900.1.p, Phvul.006G033300.1.p, Phvul.008G124800.1.p, Phvul.010G032000.1.p, Phvul.002G196200.1.p, Phvul.002G318300.1.p, Phvul.002G127800.1.p, Phvul.001G048050.1.p, Phvul.001G048600.1.p, Phvul.001G075800.2.p, Phvul.001G048200.1.p, Phvul.001G075600.1.p, Phvul.001G048800.1.p, Phvul.001G049000.1.p, Phvul.001G047900.1.p, Phvul.001G048900.1.p, Phvul.007G030300.1.p, Phvul.007G030000.1.p, Phvul.007G252500.1.p, Phvul.004G155400.2.p, Phvul.004G155000.1.p, Phvul.004G155300.1.p, Phvul.004G154900.1.p, Phvul.004G155600.1.p, Phvul.008G175100.1.p, Phvul.008G174900.2.p, Phvul.011G197700.2.p, Phvul.011G197700.1.p, Phvul.011G198500.1.p, Phvul.011G200400.1.p, Phvul.011G197500.1.p, Phvul.002G123200.1.p, Phvul.001G069248.1.p, Phvul.001G068800.1.p, Phvul.001G070200.1.p, Phvul.001G076100.1.p, Phvul.001G048500.1.p, Phvul.001G069216.1.p, Phvul.004G124400.1.p, Phvul.011G198200.1.p, Phvul.011G198500.3.p, Phvul.011G198500.2.p, Phvul.005G182900.1.p, Phvul.009G209200.1.p, Phvul.010G086600.1.p, Phvul.010G028300.1.p, Phvul.007G280200.1.p, Phvul.007G077300.1.p, Phvul.004G085800.1.p, Phvul.009G068300.1.p, Phvul.009G068100.1.p, Phvul.009G122600.1.p, Phvul.003G164800.1.p, Phvul.003G015900.1.p, Phvul.010G100400.1.p, Phvul.005G177900.1.p, Phvul.002G046100.2.p, Phvul.002G046100.1.p, Phvul.005G165900.1.p, Phvul.005G165900.2.p, Phvul.002G215200.1.p, Phvul.006G137300.1.p, Phvul.010G064900.1.p, Phvul.009G046900.1.p, Phvul.009G046600.1.p, Phvul.007G220700.1.p, Phvul.005G110400.1.p, Phvul.005G110300.1.p, Phvul.011G109100.2.p, Phvul.005G110300.3.p, Phvul.005G110300.2.p, Phvul.002G046500.2.p, Phvul.001G051100.1.p, Phvul.009G092900.1.p, Phvul.002G123500.1.p, Phvul.003G260300.1.p, Phvul.009G014500.1.p, Phvul.002G025700.1.p, Phvul.009G239700.1.p, Phvul.009G239500.1.p, Phvul.004G039400.1.p, Phvul.004G039501.1.p, |

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>PV</td><td>Phvul.005G183300.1.p, Phvul.009G239800.1.p, Phvul.001G069232.1.p, Phvul.001G076200.1.p, Phvul.001G069264.1.p, Phvul.001G070000.1.p, Phvul.001G068500.1.p, Phvul.001G069600.1.p, Phvul.001G069300.1.p, Phvul.009G198800.1.p, Phvul.003G161500.1.p, Phvul.008G281500.3.p, Phvul.009G208500.1.p, Phvul.006G148000.1.p, Phvul.006G069600.1.p, Phvul.008G174300.1.p, Phvul.008G175200.1.p, Phvul.008G259700.1.p, Phvul.001G134600.3.p, Phvul.001G134600.1.p, Phvul.001G134600.2.p, Phvul.007G093900.1.p</td></tr>
<tr><td>VA</td><td>gi|1044542655|ref|XP_017412055.1|, gi|1044539315|ref|XP_017435142.1|, gi|1044538336|ref|XP_017430087.1|, gi|1044558845|ref|XP_017415595.1|, gi|1044553187|ref|XP_017412597.1|, gi|1044552814|ref|XP_017412395.1|, gi|1044552816|ref|XP_017412396.1|, gi|1044552818|ref|XP_017412397.1|, gi|1044557767|ref|XP_017415041.1|, gi|1044557468|ref|XP_017414886.1|, gi|1044557471|ref|XP_017414887.1|, gi|1044555047|ref|XP_017413644.1|, gi|1044554285|ref|XP_017413222.1|, gi|1044554287|ref|XP_017413223.1|, gi|1044554289|ref|XP_017413224.1|, gi|1044553541|ref|XP_017412781.1|, gi|1044565748|ref|XP_017419173.1|, gi|1044562695|ref|XP_017417603.1|, gi|1044565601|ref|XP_017419095.1|, gi|1044565603|ref|XP_017419096.1|, gi|1044559542|ref|XP_017415950.1|, gi|1044567017|ref|XP_017419838.1|, gi|1044569326|ref|XP_017421107.1|, gi|1044569119|ref|XP_017420997.1|, gi|1044567519|ref|XP_017420097.1|, gi|1044567587|ref|XP_017420133.1|, gi|1044569992|ref|XP_017421471.1|, gi|1044572161|ref|XP_017422615.1|, gi|1044570046|ref|XP_017421496.1|, gi|1044567801|ref|XP_017420257.1|, gi|1044567393|ref|XP_017420033.1|, gi|1044567395|ref|XP_017420034.1|, gi|1044569229|ref|XP_017421053.1|, gi|1044572639|ref|XP_017422871.1|, gi|1044576953|ref|XP_017425215.1|, gi|1044575680|ref|XP_017424544.1|, gi|1044576168|ref|XP_017424796.1|, gi|1044576166|ref|XP_017424795.1|, gi|1044572751|ref|XP_017422934.1|, gi|1044572753|ref|XP_017422935.1|, gi|1044572755|ref|XP_017422936.1|, gi|1044573257|ref|XP_017423206.1|, gi|1044574247|ref|XP_017423770.1|, gi|1044579756|ref|XP_017426749.1|, gi|1044578902|ref|XP_017426275.1|, gi|1044581878|ref|XP_017427886.1|, gi|1044582659|ref|XP_017428292.1|, gi|1044587874|ref|XP_017431121.1|, gi|1044583111|ref|XP_017428538.1|, gi|1044586124|ref|XP_017430187.1|, gi|1044582906|ref|XP_017428420.1|, gi|1044582908|ref|XP_017428421.1|, gi|1044582902|ref|XP_017428418.1|, gi|1044582904|ref|XP_017428419.1|, gi|1044583531|ref|XP_017428794.1|, gi|1044585323|ref|XP_017429772.1|, gi|1044583537|ref|XP_017428797.1|, gi|1044584400|ref|XP_017429276.1|, gi|1044585930|ref|XP_017430085.1|, gi|1044585926|ref|XP_017430083.1|, gi|1044585922|ref|XP_017430081.1|, gi|1044585924|ref|XP_017430082.1|, gi|1044584719|ref|XP_017429448.1|, gi|1044584717|ref|XP_017429447.1|, gi|1044584133|ref|XP_017429136.1|, gi|1044584382|ref|XP_017429267.1|, gi|1044584384|ref|XP_017429268.1|, gi|1044586018|ref|XP_017430131.1|, gi|1044584837|ref|XP_017429508.1|, gi|1044588148|ref|XP_017431268.1|,</td></tr>
</table>

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

<table>
<tr><th>SP</th><th>Protein IDs</th></tr>
<tr><td>VA</td><td>gi|1044584374|ref|XP_017429261.1|, gi|1044587966|ref|XP_017431169.1|, gi|1044592977|ref|XP_017433882.1|, gi|1044592979|ref|XP_017433883.1|, gi|1044588377|ref|XP_017431394.1|, gi|1044594046|ref|XP_017434454.1|, gi|1044594048|ref|XP_017434455.1|, gi|1044595731|ref|XP_017435372.1|, gi|1044593264|ref|XP_017434039.1|, gi|1044594573|ref|XP_017434759.1|, gi|1044593971|ref|XP_017434413.1|, gi|1044594575|ref|XP_017434760.1|, gi|1044593926|ref|XP_017434389.1|, gi|1044593928|ref|XP_017434390.1|, gi|1044593930|ref|XP_017434391.1|, gi|1044593513|ref|XP_017434173.1|, gi|1044594625|ref|XP_017434787.1|, gi|1044543931|ref|XP_017437949.1|, gi|1044543859|ref|XP_017437908.1|, gi|1044543933|ref|XP_017437950.1|, gi|1044544959|ref|XP_017438575.1|, gi|1044548728|ref|XP_017440580.1|, gi|1044547852|ref|XP_017440111.1|, gi|1044549307|ref|XP_017440889.1|, gi|1044549165|ref|XP_017440816.1|, gi|1044549957|ref|XP_017441275.1|, gi|1044552347|ref|XP_017442568.1|, gi|1044549993|ref|XP_017441296.1|, gi|1044551017|ref|XP_017441857.1|, gi|1044549347|ref|XP_017440912.1|, gi|1044522253|ref|XP_017405392.1|, gi|1044523405|ref|XP_017406017.1|, gi|1044523692|ref|XP_017406177.1|, gi|1044523704|ref|XP_017406183.1|, gi|1044523702|ref|XP_017406182.1|, gi|1044524490|ref|XP_017406616.1|, gi|1044524977|ref|XP_017406882.1|, gi|1044526539|ref|XP_017407735.1|, gi|1044526719|ref|XP_017407835.1|, gi|1044526819|ref|XP_017407890.1|, gi|1044526823|ref|XP_017407892.1|, gi|1044527750|ref|XP_017408387.1|, gi|1044529338|ref|XP_017409251.1|, gi|1044529429|ref|XP_017409301.1|, gi|1044529431|ref|XP_017409302.1|, gi|1044531629|ref|XP_017410502.1|, gi|1044531641|ref|XP_017410509.1|, gi|1044531637|ref|XP_017410506.1|, gi|1044531643|ref|XP_017410510.1|, gi|1044532526|ref|XP_017410998.1|, gi|1044532544|ref|XP_017411008.1|, gi|1044532649|ref|XP_017411066.1|</td></tr>
<tr><td>VR</td><td>XP_014514948.1, XP_014511080.1, XP_014489574.1, XP_014508105.1, XP_014506697.1, XP_014506148.1, XP_014498341.1, XP_014496650.1, XP_014490154.1, XP_014506127.1, XP_014490717.1, XP_014522921.1, XP_014520481.1, XP_014516009.1, XP_014515487.1, XP_014497362.1, XP_014497291.1, XP_014496727.1, XP_014496724.1, XP_014491057.1, XP_014512982.1, XP_014492624.1, XP_014515497.1, XP_014505406.1, XP_014522912.1, XP_014519431.1, XP_014519429.1, XP_014515582.1, XP_014510590.1, XP_014498515.1, XP_014521139.1, XP_014520316.1, XP_014519535.1, XP_014516826.1, XP_014516721.1, XP_014516601.1, XP_014514619.1, XP_014501401.1, XP_014501400.1, XP_014501399.1, XP_014501398.1, XP_014500351.1, XP_014499374.1, XP_014498848.1, XP_014513881.1, XP_014494611.1, XP_014494610.1, XP_014522660.1, XP_014521802.1, XP_014513045.1, XP_014512822.1, XP_014511990.1, XP_014506794.1, XP_014504460.1, XP_014504026.1, XP_014503878.1, XP_014503375.1, XP_014503367.1, XP_014502724.1, XP_014494509.1, XP_014513343.1, XP_014505242.1, XP_014503819.1, XP_014502686.1, XP_014502685.1, XP_014514832.1, XP_014520378.1, XP_014519536.1,</td></tr>
</table>

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VR | XP_014519533.1, XP_014504395.1, XP_014502230.1, XP_014500914.1, XP_014494441.1, XP_014499569.1, XP_014510285.1, XP_014517209.1, XP_014520139.1, XP_014517386.1, XP_014519669.1, XP_014523141.1, XP_014494299.1, XP_014494297.1, XP_014494300.1, XP_014521522.1, XP_014498850.1, XP_014492796.1, XP_014492785.1, XP_014492769.1, XP_014491686.1, XP_014492813.1, XP_014511529.1, XP_014511397.1, XP_014507736.1, XP_014495699.1, XP_014504562.1, XP_014503164.1, XP_014503161.1, XP_014502648.1, XP_014502561.1, XP_014502560.1, XP_014504462.1, XP_014502589.1, XP_014502586.1, XP_014503376.1, XP_014517408.1, XP_014502194.1, XP_014518743.1, XP_014521803.1, XP_014503708.1, XP_014495984.1, XP_014495982.1, XP_014505288.1, XP_014513586.1 |
| VU | Vigun06g011000.1.p, Vigun08g158900.1.p, Vigun08g159500.1.p, Vigun08g159400.1.p, Vigun08g159100.1.p, Vigun08g159600.1.p, Vigun08g159200.1.p, Vigun08g158400.1.p, Vigun08g159700.1.p, Vigun08g158500.1.p, Vigun03g084000.1.p, Vigun03g145600.1.p, Vigun03g373900.1.p, Vigun08g159000.1.p, Vigun08g158800.1.p, Vigun08g158500.2.p, Vigun08g158600.1.p, Vigun08g159300.1.p, Vigun05g103400.1.p, Vigun05g095400.1.p, Vigun02g085300.1.p, Vigun02g085100.1.p, Vigun02g085000.1.p, Vigun02g084800.1.p, Vigun02g085200.1.p, Vigun03g084100.1.p, Vigun03g015400.2.p, Vigun03g015400.1.p, Vigun05g095200.1.p, Vigun06g040300.1.p, VigunL088600.1.p, Vigun10g029600.1.p, Vigun10g003400.1.p, Vigun10g003100.1.p, Vigun05g127700.1.p, Vigun05g127800.1.p, Vigun05g128100.1.p, Vigun05g127600.1.p, Vigun02g162700.1.p, Vigun11g002600.1.p, Vigun10g117700.1.p, Vigun05g128000.1.p, Vigun02g162600.1.p, Vigun02g105900.1.p, Vigun09g054500.1.p, Vigun02g103200.1.p, Vigun02g107100.1.p, Vigun07g265700.1.p, Vigun10g003300.1.p, Vigun04g165100.2.p, Vigun04g165100.3.p, Vigun04g165100.4.p, Vigun04g165100.1.p, Vigun10g115700.1.p, Vigun04g164700.1.p, Vigun05g300400.1.p, Vigun09g234800.1.p, Vigun09g234900.1.p, Vigun09g234700.1.p, Vigun09g235000.1.p, Vigun05g294900.4.p, Vigun05g294900.3.p, Vigun05g294900.2.p, Vigun05g294900.1.p, Vigun02g196000.1.p, Vigun10g003000.1.p, Vigun02g195800.1.p, Vigun09g168700.1.p, Vigun10g131300.1.p, Vigun04g118800.1.p, Vigun02g017200.1.p, Vigun07g219900.1.p, Vigun11g015000.1.p, Vigun07g267300.2.p, Vigun07g033200.1.p, Vigun04g024900.2.p, Vigun04g024900.1.p, Vigun04g024500.1.p, Vigun04g024400.1.p, Vigun04g025000.1.p, Vigun08g052400.1.p, Vigun08g102100.1.p, Vigun08g101400.1.p, Vigun08g072900.1.p, Vigun08g081400.1.p, Vigun08g051600.1.p, Vigun08g053100.1.p, Vigun08g064800.1.p, Vigun08g101200.1.p, Vigun08g101900.1.p, Vigun08g052300.1.p, Vigun08g073000.1.p, |

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| VU | Vigun08g053000.1.p, Vigun08g101700.1.p, Vigun08g053300.1.p, Vigun03g015300.1.p, Vigun11g015000.2.p, Vigun07g267400.1.p, Vigun07g266500.1.p, Vigun07g267500.1.p, Vigun07g267800.1.p, Vigun04g024700.2.p, Vigun04g024700.1.p, Vigun08g081900.1.p, Vigun08g073400.1.p, Vigun08g081600.1.p, Vigun08g082200.1.p, Vigun08g072200.1.p, Vigun07g001200.1.p, Vigun02g136300.1.p, Vigun09g142500.1.p, Vigun09g142300.1.p, Vigun09g142700.1.p, Vigun09g209900.1.p, Vigun08g055100.1.p, Vigun05g221900.1.p, Vigun02g195700.1.p, Vigun02g104600.1.p, Vigun02g195500.1.p, Vigun02g196100.1.p, Vigun11g112200.1.p, Vigun05g183100.1.p, Vigun02g106700.1.p, Vigun10g087500.1.p, Vigun10g042900.1.p, Vigun03g123200.1.p, Vigun07g070000.1.p, Vigun10g042400.1.p, Vigun07g247100.1.p, Vigun09g251200.1.p, Vigun05g214800.1.p, Vigun05g300900.1.p, Vigun04g062100.1.p, Vigun09g025100.1.p, Vigun09g025000.1.p, Vigun09g025400.1.p, Vigun09g055100.1.p, Vigun09g025200.1.p, Vigun08g082000.1.p, Vigun08g101800.1.p, Vigun08g081300.1.p, Vigun08g072800.1.p, Vigun08g072400.1.p, Vigun08g073100.1.p, Vigun08g081300.2.p, Vigun08g081300.3.p, Vigun08g101500.1.p, Vigun08g212500.1.p, Vigun09g064000.1.p, Vigun03g352400.1.p, Vigun06g159800.1.p, Vigun07g032800.1.p, Vigun03g260600.2.p, Vigun03g260600.1.p, Vigun08g102000.1.p, Vigun08g121500.1.p, Vigun01g116700.1.p, Vigun07g203100.1.p |
| VV | GSVIVT01016080001, GSVIVT01019205001, GSVIVT01010376001, GSVIVT01019727001, GSVIVT01026551001, GSVIVT01026547001, GSVIVT01025346001, GSVIVT01020730001, GSVIVT01028785001, GSVIVT01014328001, GSVIVT01036817001, GSVIVT01028787001, GSVIVT01028783001, GSVIVT01011620001, GSVIVT01001679001, GSVIVT01012686001, GSVIVT01011842001, GSVIVT01021645001, GSVIVT01021635001, GSVIVT01020735001, GSVIVT01029148001, GSVIVT01021407001, GSVIVT01017518001, GSVIVT01010040001, GSVIVT01009415001, GSVIVT01009414001, GSVIVT01009412001, GSVIVT01017953001, GSVIVT01012107001, GSVIVT01021647001, GSVIVT01017326001, GSVIVT01015977001, GSVIVT01000058001, GSVIVT01038723001, GSVIVT01038721001, GSVIVT01038720001, GSVIVT01038718001, GSVIVT01038703001, GSVIVT01030995001, GSVIVT01011854001, GSVIVT01038699001, GSVIVT01023349001, GSVIVT01018814001, GSVIVT01009829001, GSVIVT01020778001, GSVIVT01008328001, GSVIVT01024382001, GSVIVT01005163001, GSVIVT01014750001, GSVIVT01027831001, GSVIVT01015489001, GSVIVT01008436001, GSVIVT01019038001, GSVIVT01001044001, GSVIVT01035532001, GSVIVT01032970001, GSVIVT01028400001, GSVIVT01027459001, GSVIVT01015513001 |
| AT | AT5G20480.1, AT3G47110.1, AT5G05160.1, AT2G45590.1, AT5G51350.1, AT3G47090.1, AT5G39390.1, AT1G80870.1, AT4G25390.2, AT4G25390.1, AT1G68400.1, AT5G46330.1, AT3G47570.1, AT3G47580.1, AT1G34300.1, AT4G00340.1, AT4G32300.1, AT2G19130.1, AT5G24080.1, AT5G38260.1, AT5G39030.1, AT5G38280.1, AT5G39020.1, AT5G35370.1, AT4G18250.1, AT4G36180.1, AT5G06820.1, AT1G07150.1, AT1G75640.1, AT1G66910.1, |

**Table A6. Protein IDs of RLK-nonRD identified among the legumes/non-legumes. (Continued)**

| SP | Protein IDs |
|---|---|
| AT | AT1G66980.1, AT1G67000.1, AT2G13800.1, AT1G66930.1, AT5G56040.1, AT5G56040.2, AT4G26540.1, AT2G30040.1, AT2G41910.1, AT5G20050.1, AT3G26700.1, AT2G41820.1, AT2G18530.1, AT2G30940.1, AT2G30940.2, AT5G37350.2, AT4G09760.3, AT4G18640.1 |
| SL | Solyc07g005920.1.1, Solyc10g085110.1.1, Solyc08g075610.1.1, Solyc08g075590.1.1, Solyc04g009040.2.1, Solyc07g018190.2.1, Solyc10g085120.1.1, Solyc03g006100.2.1, Solyc06g076910.1.1, Solyc11g005630.1.1, Solyc01g006530.1.1, Solyc01g006520.2.1, Solyc03g078370.1.1, Solyc09g075920.1.1, Solyc01g094020.2.1, Solyc03g006920.2.1, Solyc03g098400.1.1, Solyc03g019980.1.1, Solyc12g008400.1.1, Solyc02g070890.2.1, Solyc07g062170.1.1, Solyc02g072470.2.1, Solyc02g072480.2.1, Solyc02g068820.1.1, Solyc02g072400.1.1, Solyc08g075600.1.1, Solyc03g006030.2.1, Solyc02g031790.1.1, Solyc02g068830.1.1, Solyc03g006080.2.1, Solyc06g006010.1.1, Solyc11g020260.1.1, Solyc05g008860.2.1, Solyc07g053220.1.1, Solyc05g015150.2.1, Solyc01g094830.2.1, Solyc03g005130.1.1, Solyc03g007790.2.1, Solyc04g078410.2.1, Solyc04g014650.2.1, Solyc06g036470.1.1, Solyc03g122230.1.1, Solyc01g014520.1.1, Solyc05g008980.2.1, Solyc05g008960.2.1, Solyc04g012100.1.1, Solyc05g008950.2.1, Solyc06g048740.1.1, Solyc04g015460.2.1, Solyc09g011330.1.1, Solyc07g006770.2.1, Solyc08g059730.1.1, Solyc04g015980.2.1, Solyc04g009640.2.1, Solyc04g012090.1.1, Solyc03g117550.1.1, Solyc09g083210.2.1, Solyc02g081500.2.1, Solyc04g012190.1.1, Solyc07g055650.1.1, Solyc02g084370.1.1, Solyc04g081080.1.1, Solyc03g033610.1.1, Solyc02g086210.2.1, Solyc05g009050.2.1, Solyc05g009090.1.1, Solyc03g083470.2.1, Solyc05g009040.2.1, Solyc01g008500.2.1, Solyc03g007210.2.1, Solyc02g072070.2.1, Solyc02g072440.2.1, Solyc07g005010.2.1, Solyc07g064820.1.1, Solyc06g008890.2.1, Solyc07g065860.2.1, Solyc02g065520.1.1, Solyc04g015600.2.1, Solyc02g089900.1.1, Solyc10g079170.1.1, Solyc02g076660.2.1, Solyc02g069750.1.1, Solyc08g075620.1.1 |
| RLK | sp\|Q9SD62\|, sp\|C0LGT6\|, tr\|K0IXC4\|, sp\|Q9FL28\|, tr\|Q38925\|, sp\|Q1MX30\|, sp\|P93604\|, tr\|Q6VTH8\|, tr\|Q00NV1\|, tr\|Q2EZ14\|, tr\|Q651J1\|, sp\|D7SFH9\|, tr\|A0A0P0Y5F3\|, tr\|G4XXY5\|, tr\|B6EB06\| |

SP: Species, the RLK correspond to the resistant RLK reported in Table A1 that were evaluated to identify the presence on the nonRD modification in the Pkinase.

**Table A7. Non-target domains identified in the RLK predicted.**

| Domains | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Domain name | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| RCC1_2 | X | X | X | X | X | X | X | X | X | X |
| RVT_2 | X | | | | | | | | X | |
| DUF3403 | X | X | X | X | X | X | X | X | X | X |
| Ribonuc_2-5A | X | X | X | X | X | X | X | X | X | X |
| DUF3453 | | | | | | | | | X | |
| NAF | | X | X | X | X | X | | X | X | X |
| PP2C | X | | | | | | | | X | |
| PRIMA1 | | | | | | | | | X | |
| DUF3660 | X | X | X | X | X | X | X | | X | X |
| GDPD | | | | | | | | | X | |
| Rve | X | | | | | X | | | | |
| DHHC | X | | | | | | | | | |
| DUF4219 | X | | | | | | | | | |
| PB1 | X | | | | | | | | | |
| Transposase_21 | X | | | | | | | | | |
| Transposase_24 | X | | | | | | | | | |
| Sec16 | X | | | | | | | | | |
| gag_pre-integrs | X | | | | | X | | | | |
| PI3Ka | X | | | | | | | | | |
| Transpos_assoc | X | | | | | | | | | |
| Glyco_hydro_18 | X | X | X | X | X | X | X | X | | |
| SCAMP | X | | | | | | | | | |
| S1FA | X | | | | | | | | | |
| 2OG-FeII_Oxy_2 | | X | | | | | | | | |
| RETICULATA-like | | X | | | | | | | | |
| MEKHLA | | X | | | | | | | | |
| HPP | | X | | | | | | | | |
| Methyltransf_29 | | | X | | | | | | | |
| MlaE | | | X | | | | | | | |
| Clathrin | | | X | | | | | | | |
| EDR1 | | | X | | | | | | | |
| zf-HIT | | | X | | | | | | | |
| DUF569 | | | X | | | | | | | |

**Table A7. Non-target domains identified in the RLK predicted. (Continued)**

| Domains | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Domain name | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| Herpes_gE | | | **X** | | | | | | | |
| Pep3_Vps18 | | | X | | | | | | | |
| Beta-lactamase | | | X | | | | | | | |
| Rio2_N | | | X | | | | | | | |
| F-box | | | X | | | | | | | |
| SHQ1 | | | X | | | | | | | |
| FBD | | | X | | | | | | | |
| MRP-L27 | | | X | | | | | | | |
| HSP20 | | | X | | | | | | | |
| Terpene_synth_C | | | X | | | | | | | |
| Nodulin_late | | | X | | | | | | | |
| Peptidase_M20 | | | X | | | | | | | |
| Ost5 | | | X | | | | | | | |
| IQ | | | | X | | | | | | |
| Adeno_E3_CR2 | | | | X | | | | | | |
| Amino_oxidase | | | | | | | | | | X |
| p450 | | | | | | | | X | | X |
| EF-hand_7 | | | | | | | | | | X |
| EMP70 | | | | | | | | | | X |
| TMEM154 | | | | | | | | | | X |
| OB_NTP_bind | | | | | | | | | | X |
| MatE | | | | | | | | | | X |
| EB | | | | | X | | | | | X |
| PQQ | | | | | | | | | | X |
| FAD_binding_4 | | | | | | | | | | X |
| WD40 | | | | | X | | | X | | |
| RVP_2 | | | | | X | | | | | |
| E1-E2_ATPase | | | | | X | X | | | | |
| Hydrolase | | | | | X | X | | | | |
| EamA | | | | | X | | | | | |
| LEA_2 | | | | | | X | | | | |
| DUF4441 | | | | | | X | | | | |
| PMR5N | | | | | | | | X | | |

**Table A7. Non-target domains identified in the RLK predicted. (Continued)**

| Domains | Species | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Domain name | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| H_PPase | | | | | | | | X | | |
| Cu_bind_like | | | | | | | | X | | |
| Sugar_tr | | | | | | | | X | | |
| Cupin_1 | | | | | | | | X | | |
| Retrotran_gag_3 | | | | | | | | X | | |
| Glyco_hydro_28 | | | | | | | | X | | |
| Aldose_epim | | | | | | | | X | | |
| DHQS | | | | | | | | X | | |
| PRA1 | | | | | | | | X | | |
| zf-RING_2 | | | | | | | | X | | |
| Pyrophosphatase | | | | | | | | X | | |
| JmjN | | | | | | | | X | | |
| JmjC | | | | | | | | X | | |
| Peptidase_M50B | | | | | | | | X | | |
| PGG | | | | | | | | X | | |
| Ank_2 | | | | | | | | X | | |
| SRF-TF | | | | | | | | X | | |
| FATC | | | | | | | | X | | |
| zf-C5HC2 | | | | | | | | X | | |
| K-box | | | | | | | | X | | |
| ATP-synt_A | | | | | | | | X | | |
| Retrotran_gag_2 | | | | | | | | X | | |
| Tmemb_14 | | | | | | | | X | | |
| SLAC1 | | | | | | | | X | | |

Present: x

**Table A8. Non-target domains identified in the RLP predicted.**

| Domains | Species | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Domain name | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| DUF2854 | x | x | | x | x | x | x | x | x | x |
| UMP1 | | | | | | | | | x | |
| Glyco_hydro_32N | x | x | x | x | x | x | x | | x | x |
| DUF3357 | x | x | x | x | x | x | x | | x | x |
| DUF604 | | | | | | x | | | x | |
| Alliinase_C | x | x | | | | | x | x | x | x |
| F-box | | | x | | | | | x | x | |
| Galactosyl_T | x | x | x | x | x | x | x | x | x | x |
| FBD | | | x | | | | | | x | |
| zf-RING_2 | x | x | x | x | x | x | x | x | x | x |
| PA | x | x | x | x | x | x | x | x | x | x |
| PRIMA1 | | | | | | | | | x | |
| Mito_carr | x | | | | | | | | | |
| MtN3_slv | | x | | | | | | | | |
| SHMT | | x | | | | | | | | |
| Peptidase_M8 | | x | x | x | x | x | x | | | |
| Exostosin | | x | x | x | x | x | x | x | | x |
| Glyco_transf_90 | | x | | | | | | | | |
| Ribosomal_S13 | | | x | | | | | | | |
| RNA_pol_Rpb2_5 | | | x | | | | | | | |
| RNA_pol_Rpb2_4 | | | x | | | | | | | |
| RNA_pol_Rpb2_7 | | | x | | | | | | | |
| RNA_pol_Rpb2_6 | | | x | | | | | | | |
| RNA_pol_Rpb2_1 | | | x | | | | | | | |
| RNA_pol_Rpb2_3 | | | x | | | | | | | |
| RNA_pol_Rpb2_2 | | | x | | | | | | | |
| LIAS_N | | | | x | | | | | | |
| DUF4216 | | | | | x | | | | | |
| DUF4218 | | | | | x | | | | | |
| F-box-like | | | | | x | | x | x | | |
| Retrotran_gag_2 | | | | | | x | | | | |
| Snf7 | | | | | | x | | | | |
| MatE | | | | | | | | x | | |

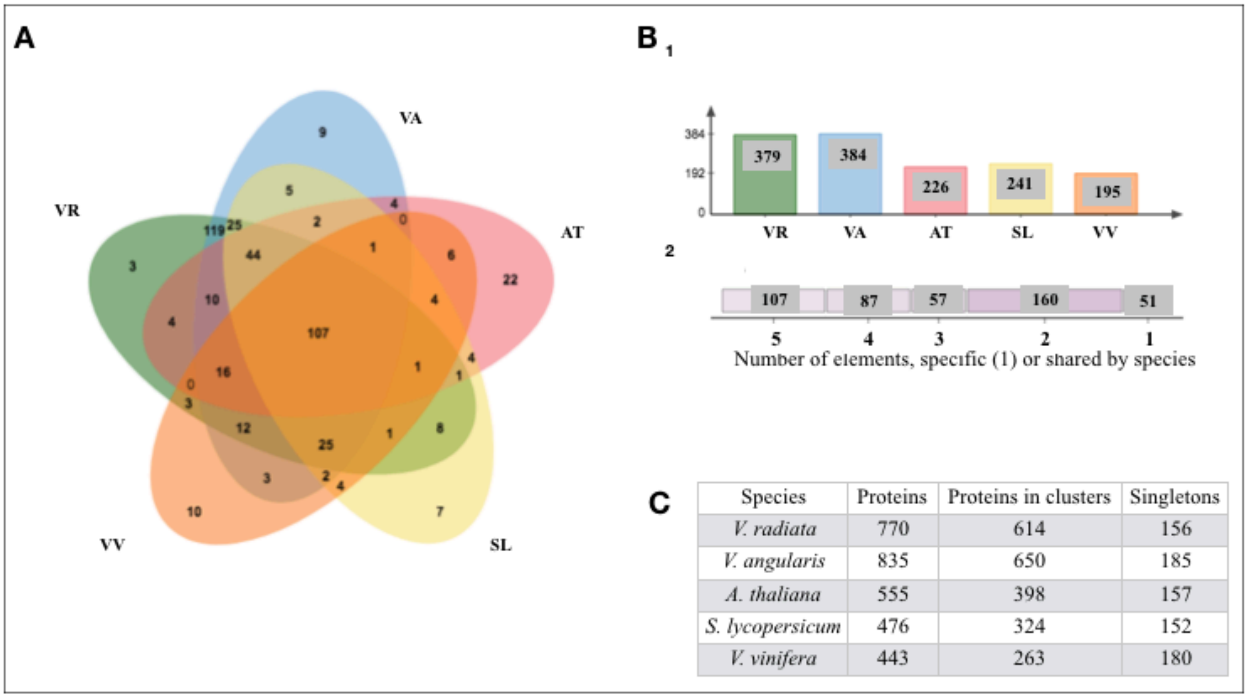
**Table A8. Non-target domains identified in the RLP predicted. (Continued)**

| Domains | Species | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Domain name | CC | GM | MT | PV | VA | VR | VU | VV | AT | SL |
| Peptidase_S10 | | | | | | | | x | | |
| zf-CCHC | | | | | | | | x | | |
| Glyoxal_oxid_N | | | | | | | | x | | |
| PEARLI-4 | | | | | | | | x | | |
| DUF1929 | | | | | | | | x | | |

Present: X

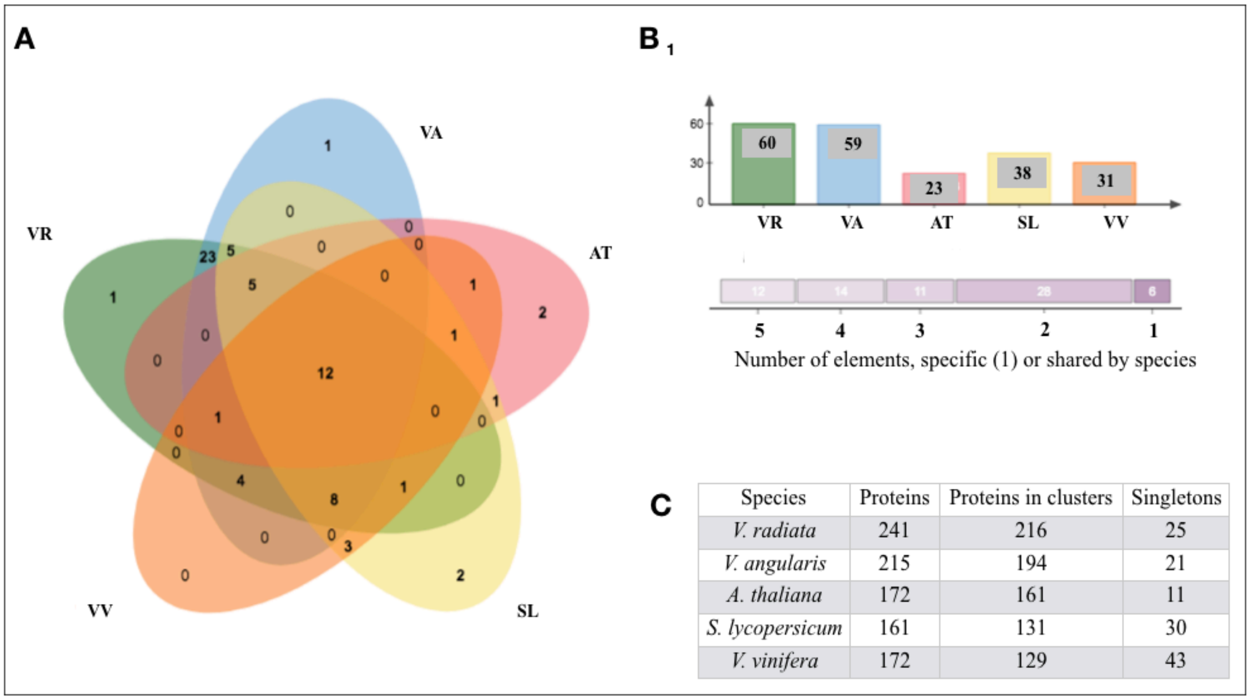
**Table A9. RLK-TNFR receptors reported in legumes/non-legumes.**

| Species | protein ID | uniprot URL | phytozome ID | class and signal peptide | Details |
|---|---|---|---|---|---|
| AT | ACCR2_ARATH | http://www.uniprot.org/uniprot/O80963 | AT2G39180.1 | psg_rlk | rlk_extra |
| | ACR4L_ARATH | http://www.uniprot.org/uniprot/Q9LX29 | AT3G59420.1 | psg_rlk | rlk_extra |
| GM | I1L8V3_SOYBN | http://www.uniprot.org/uniprot/I1L8V3 | Glyma.10G052600.1 | psg_rlk | rlk_only_kinase |
| | | | Glyma.13G139800.1 | psg_rlk | rlk_only_kinase |
| | I1LZ42_SOYBN | http://www.uniprot.org/uniprot/I1LZ42 | Glyma.13G139800.1 | psg_rlk | rlk_only_kinase |
| | | | Glyma.10G052600.1 | psg_rlk | rlk_only_kinase |
| | K7LPG1_SOYBN | http://www.uniprot.org/uniprot/K7LPG1 | Glyma.12G052300.1 | nsg_rlk | rlk_only_kinase |
| | | | Glyma.15G030000.1 | nsg_rlk | rlk_extra |
| | | | Glyma.13G344400.1 | psg_rlk | rlk_extra |
| | K7LT62_SOYBN | http://www.uniprot.org/uniprot/K7LT62 | Glyma.12G052300.1 | nsg_rlk | rlk_only_kinase |
| | | | Glyma.13G344400.1 | psg_rlk | rlk_extra |
| SL | UPI0002BC8D7F | http://www.uniprot.org/uniparc/UPI0002BC8D7F | Solyc09g007750.2.1 | psg_rlk | rlk_only_kinase |
| | | | Solyc11g044940.1.1 | psg_rlk | rlk_extra |
| | UPI0002BCA038 | http://www.uniprot.org/uniparc/UPI0002BCA038 | Solyc09g007750.2.1 | psg_rlk | rlk_only_kinase |
| VV | F6HNH5_VITVI | http://www.uniprot.org/uniprot/F6HNH5 | GSVIVT01016294001 | NO | NO |
| | F6HYS0_VITVI | http://www.uniprot.org/uniprot/F6HYS0 | GSVIVT01010759001 | NO | NO |
| | UPI00023B33D4 | http://www.uniprot.org/uniparc/UPI00023B33D4 | GSVIVT01016294001 | NO | NO |



| Species | Proteins | Proteins in clusters | Singletons |
|---|---|---|---|
| *V. radiata* | 770 | 614 | 156 |
| *V. angularis* | 835 | 650 | 185 |
| *A. thaliana* | 555 | 398 | 157 |
| *S. lycopersicum* | 476 | 324 | 152 |
| *V. vinifera* | 443 | 263 | 180 |

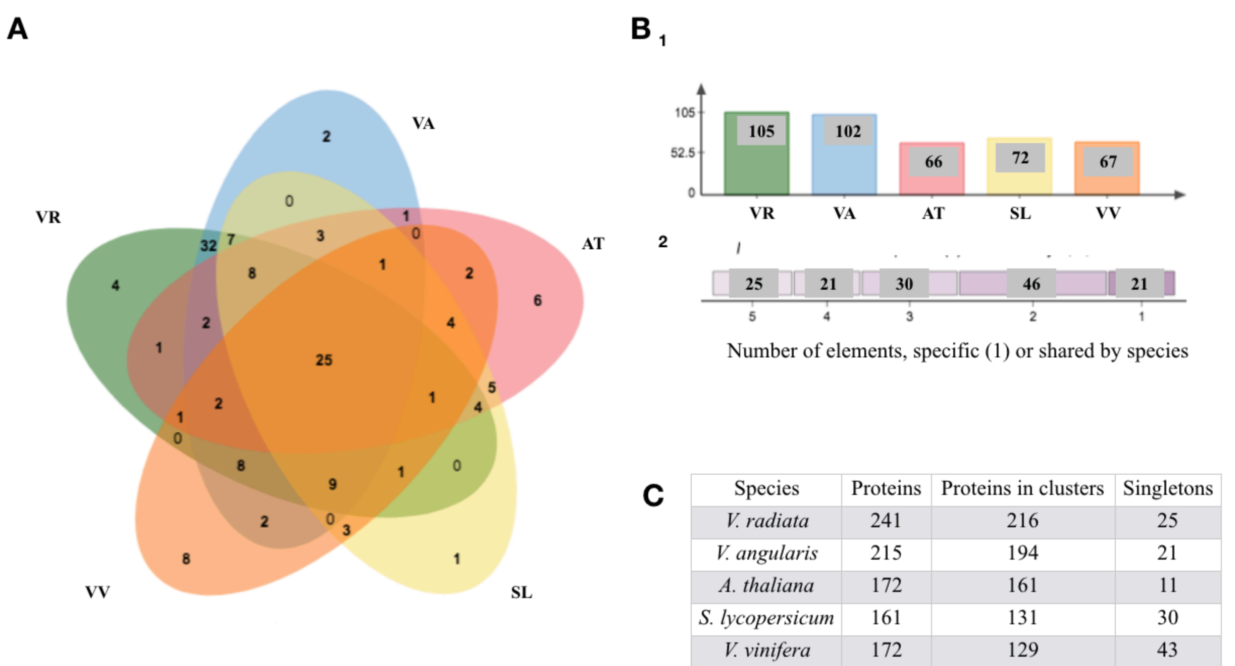
**Figure A1. Summary of the RLK orthology analysis among VR, VA, AT, SL, and VV.**
**A**. Venn diagram showing the distribution of shared gene families (orthologous clusters) among VR, VA, AT, SL, and VV. **B1**. The numbers refer to all the clusters in the species, including orthologs and in-paralogs. **B2**. Distribution of the number of species present in ortholog clusters, one or share elements among species. **C**. Summary of the total number of proteins, clusters, and singletons within each species. The RLK and its isoforms and nonRD proteins were included in this figure. 28 single–copy gene clusters were reported among the species evaluated.



| Species | Proteins | Proteins in clusters | Singletons |
|---|---|---|---|
| *V. radiata* | 241 | 216 | 25 |
| *V. angularis* | 215 | 194 | 21 |
| *A. thaliana* | 172 | 161 | 11 |
| *S. lycopersicum* | 161 | 131 | 30 |
| *V. vinifera* | 172 | 129 | 43 |

**Figure A2. Summary of the RLK-nonRD orthology analysis among VR, VA, AT, SL, and VV.**

**A**. Venn diagram showing the distribution of shared gene families (orthologous clusters) among VR, VA, AT, SL, and VV. **B1**. The numbers refer to all the clusters in the species, including orthologs and in-paralogs. **B2**. Distribution of the number of species present in orthologs clusters, one or share elements among species. **C**. Summary of the total number of proteins, clusters, and singletons within each species. The RLK and its isoforms and nonRD proteins were included in this figure. 3 single-copy gene clusters were reported among the species evaluated.



| Species | Proteins | Proteins in clusters | Singletons |
|---|---|---|---|
| *V. radiata* | 241 | 216 | 25 |
| *V. angularis* | 215 | 194 | 21 |
| *A. thaliana* | 172 | 161 | 11 |
| *S. lycopersicum* | 161 | 131 | 30 |
| *V. vinifera* | 172 | 129 | 43 |

**Figure A3. Summary of the RLP orthology analysis among VR, VA, AT, SL, and VV.** **A**. Venn diagram showing the distribution of shared gene families (orthologous clusters) among VR, VA, AT, SL, and VV. **B1**. The numbers refer to all the clusters in the species, including orthologs and in-paralogs. **B2**. Distribution of the number of species present in orthologs clusters, one or share elements among species. **C**. Summary of the total number of proteins, clusters, and singletons within each species. The RLK and its isoforms and nonRD proteins were included in this evaluation report. 4 single-copy gene clusters were reported among the species evaluated.

**Table A10. RLK/RLP single-copy gene clusters.**

| Class | Protein IDs in single-copy gene clusters |
|---|---|
| RLK | vvi\|GSVIVT01011620001;cca\|KYP50005.1;pvu\|Phvul.003G015900.1.p;vun\|Vigun02g017200.1.p;mtr\|Medtr1g032930.1;gma\|Glyma.07G173000.1.p |
| RLK | vvi\|GSVIVT01015489001;cca\|KYP51141.1;pvu\|Phvul.003G161500.1.p;vun\|Vigun03g352400.1.p;mtr\|Medtr4g113710.1;gma\|Glyma.17G117800.1.p |
| RLK | vvi\|GSVIVT01026007001;cca\|KYP68786.1;pvu\|Phvul.011G003100.1.p;gma\|Glyma.12G002400.1.p;vun\|Vigun11g223700.1.p;mtr\|Medtr4g035180.1 |
| RLK | vvi\|GSVIVT01016473001;cca\|KYP32585.1;vun\|Vigun01g221200.1.p;pvu\|Phvul.001G236600.1.p;mtr\|Medtr3g069050.1;gma\|Glyma.11G223200.1.p |
| RLK | cca\|KYP73376.1;vvi\|GSVIVT01020786001;vun\|Vigun11g152300.1.p;mtr\|Medtr5g083910.1;gma\|Glyma.02G288600.1.p;pvu\|Phvul.008G264400.1.p |
| RLK | vvi\|GSVIVT01016722001;pvu\|Phvul.010G127500.1.p;vun\|Vigun10g162000.1.p;mtr\|Medtr8g021350.1;gma\|Glyma.16G009900.1.p;cca\|KYP50467.1 |
| RLK | vvi\|GSVIVT01019039001;cca\|KYP66704.1;pvu\|Phvul.002G025500.1.p;vun\|Vigun02g136400.1.p;mtr\|Medtr5g019040.1;gma\|Glyma.11G063100.1.p |
| RLP | cca\|KYP64240.1;vvi\|GSVIVT01026990001;pvu\|Phvul.008G060000.1.p;vun\|Vigun05g062200.1.p;mtr\|Medtr7g078220.1;gma\|Glyma.18G236000.1.p |

**Table A11. Summary of the genes evaluated among legumes/non-legumes in the synteny analysis.**

| Plasma membrane class | Initial genes | Genes located in the chromosomes evaluated | Genes located in 1 or more synteny blocks* |
|---|---|---|---|
| RLK | 7103 | 4832 | 3739 |
| RLK-nonRD | 1184 | 964 | 740 |
| RLP | 2445 | 2005 | 1447 |

* Data reported is non-redundant.