Cross-seasons gene expression study on programmed cell death regulatory proteases in the secondary vascular tissues of angiosperm trees

A case study of heartwood formation in *Robinia pseudoacacia*

Dissertation for attaining the degree of Doctor of Natural Sciences at the Faculty of Mathematics, Computer Science and Natural Sciences, Department of Biology, University of Hamburg

Submitted by

Sylas K. Lau

Hamburg, 2022

Disputation Date

17th May 2023

Examiners

Prof. Dr. Elisabeth Magel
Universität Hamburg
Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Biologie
Institut für Holzwissenschaften

PD Dr. Matthias Fladung
Johann Heinrich von Thünen-Institut
Institut für Forstgenetik

Acknowledgment

For the last few years, this doctoral project has been a secure anchor for me in a stormy sea of uncertainty. As a pandemic loomed in the air and threats to world peace arose around the globe, I had the privilege to bury my head in studying how trees grow and be resilient. First I would like to thank my supervisor, Professor Elisabeth Magel, for giving me the opportunity, support, and liberty to build upon her work on this beautiful topic of heartwood formation. I would also like to thank my second supervisor, Dr. Matthias Fladung for the constructive feedback on the thesis. I am grateful to my "Ph.D. sister", Anna Bogun, for her companion and watching my back around the lab. The same is extended to Katharina Erdt, who was my compass in navigating the German work culture and university system. I also thank our technical assistant, Olaf Polak, for his support in establishing personal workstations and ensuring laboratory safety.

In addition to our direct working group, I must thank the gardeners at the Arboretum of Thünen Institute for providing the tree samples, which they have cared for many years and harvested for this project. I would also like to thank the amicable colleagues at the Thünen Institute of Wood Research. Especially, I would like to express my appreciation to Sergej Kaschuro for sharing his extensive knowledge of wood anatomy and microscopy on the daily basis, as well as the production of amazing images with the samples. Furthermore, I am grateful for the friendship and support of my peer in the neighboring departments. In particular, I am immensely grateful to the senior fellows Dr. Eliza Zhunusova and Dr. Martin Lopens for consistently checking on my development, and offering invaluable advice with encouragement throughout the years.

In addition to those who supported me during my doctoral studies, I want to take this opportunity to express my utmost gratitude to the teachers and mentors that were instrumental in my earlier journey. I am grateful to my high school Physiology teacher, Dr. Mao, for her passionate teaching that indulged my curiosity and nurtured my interest in the subject. The same is extended to Professor Stephen Robertson, who

fostered my development in Genetics. I must thank Dr. Agnes Chan and Dr. Henri Vanhanen, who went above and beyond to offer personal guidance and support. Without them, I would not have had the opportunity or the idea to move to another continent again for postgraduate studies in a field that gives me motivation and fulfillment.

Finally, I must state my sincere gratitude to my family and friends for their unconditional love, often despite the absence from my end. Their understanding and belief in me have been a source of strength and courage to take on and persevere through challenges in life.

Declarations

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

I hereby declare upon oath that I have written the present dissertation independently and have not used further resources and aids than those stated in the dissertation.

Ich versichere, dass dieses gebundene Exemplar der Dissertation und das in elektronischer Form eingereichte Dissertationsexemplar (über den Docata-Upload) und das bei der Fakultät (zuständiges Studienbüro bzw. Promotionsbüro Physik) zur Archivierung eingereichte gedruckte gebundene Exemplar der Dissertationsschrift identisch sind.

I, the undersigned, declare that this bound copy of the dissertation and the dissertation submitted in electronic form (via the Docata upload) and the printed bound copy of the dissertation submitted to the faculty (responsible Academic Office or the Doctoral Office Physics) for archiving are identical.

Hamburg, November 2022

Ort, den | City, date

Unterschrift | Signature

Summary

Heartwood is the inner stem core in trees that is deposited with natural antimicrobial preservatives. Its formation is defined by the conversion of reserve materials into heartwood substances in conjunction with programmed cell death (PCD) of the xylem parenchymas. The process resembles wounding responses and shares molecular regulatory mechanisms with wood formation. Despite the ecological significance as a crucial development and defense strategy in trees, research on heartwood formation (HWF) has remained scarce, especially concerning its aspects in PCD. The current study focuses on two families of regulatory proteases in plant PCD, type-II metacaspases (MC-IIs) and vacuolar processing enzymes (VPEs), and examines their gene expression patterns in the different wood regions across an annual cycle based on the pre-established HWF marker genes, *PAL1* and *CHS3*. In addition, the physiology of xylem parenchymas in trees and their terminal differentiation is reviewed to elaborate on the ecological significance of HWF. Taking *Robinia pseudoacacia* as a representative of angiosperm trees for the case study, heartwood-forming samples were collected monthly between April 2019 to March 2020 for gene expression analysis by real-time PCR. Complementary, weather data were included for a comprehensive overview, as well as sequences analysis with homologs of the candidate genes.

The results show the orderly co-expression patterns of maker genes across the wood regions and annual cycle. In the differentiating xylem, upsurges of gene expression were observed with three MC4-like *RpMC-IIs* and *RpMC9* starting in April and subsided by June, and the upregulation of γ-*RpVPE* occurred primarily in May and lessened in June, then *β-RpVPE* abruptly rose in June, and finally *PAL1* increased in June and gradually declined until September. When viewing the combination with the corresponding weather data, upregulation of the type-II MC coincided with the first sharp increase of sunshine duration of the year, whereas elevation of *β-RpVPE and PAL1* followed the rise in both sunshine and average temperature. In the transition zone (TZ), gene expression of *PAL1*, *CHS3*, and γ*VPE* elevated in synchrony in autumn. The observed pattern in the outer TZ intriguingly preceded that of the inner TZ by a

month in autumn, which may be an indication of the flow of HWF initiation signals. The present work demonstrated that HWF involves all living regions in the stems and occurs over the seasonal cycle. Furthermore, HWF is a complex differentiation process comprised of multiple major biochemical processes and structural changes, and γ-*VPE* appeared to be a suitable reference as the marker for PCD.

Zusammenfassung

Kernholz ist der innere Kern des Baumstamms, der natürliche antimikrobielle Konservierungsstoffe beinhaltet. Das Kernholz wird durch die Umwandlung von Reservestoffen in Kerninhaltsstoffe in Verbindung mit dem programmierten Zelltod des Xylemparenchyms gebildet. Der Prozess ähnelt der Wundreaktion im Holz und weist ähnliche molekulare Regulationsmechanismen wie die Holzbildung auf. Trotz ihrer ökologischen Bedeutung als entscheidende Entwicklungs- und Verteidigungsstrategie bei Bäumen ist die Kernholzbildung, insbesondere im Hinblick auf den programmierten Zelltod, noch wenig erforscht. Die vorliegende Studie konzentriert sich auf zwei Familien von bereits etablierten Marker-Proteasen, die den programmierten Zelltod in der Pflanzenentwicklung regulieren: Typ-II-Metacaspasen (MC-II) und vakuoläre Prozessierungsenzyme (VPE). Weiterhin wurden die Gene Phenylalanin-Ammoniak-Lyase (PAL) und Chalkonsynthase (CHS) betrachtet, als Schlüsselenzyme des Phenylpropanoid- bzw. Flavonoid-Biosynthesewegs. Ihre Genexpressionen weisen auf die Produktion phenolischer Verbindungen hin und wurden als Marker für die Produktion von Kerninhaltsstoffen identifiziert. Untersucht wurden die Genexpressionsmuster der genannten Gene in den verschiedenen Holzregionen über einen Jahreszyklus hinweg. Darüber hinaus wurde die Physiologie des Xylemparenchyms und seine terminale Differenzierung untersucht, um die ökologische Bedeutung der Kernholzbildung herauszustellen. Am Beispiel der Robinie (Robinia pseudoacacia), einem Vertreter der Angiospermen, wurden für die Genexpressionsanalyse mittels Real-Time-PCR, über ein Jahr hinweg monatlich kernholzbildende Gewebeproben gesammelt. Ergänzend wurde eine vergleichende Sequenzanalyse mit Homologen durchgeführt, um zusammen mit den Wetterdaten einen umfassenden Überblick zu erhalten.

Die Ergebnisse zeigen die zusammenhängenden Expressionsmuster der Maker-Gene über die Holzregionen und den Jahreszyklus hinweg. Im differenzierenden Xylem wurde ein Anstieg der Genexpression bei vier MC-II Genen beobachtet, der im April begann und bis Juni abfiel. Die Hochregulierung von γ-RpVPE trat hauptsächlich im Mai und weniger im Juni auf. Die Expression von β-RpVPE hingegen stieg im Juni

abrupt an. PAL1 wurde schließlich vermehrt ab Juni exprimiert und ging bis September allmählich zurück. In der kombinierten Betrachtung mit den entsprechenden Wetterdaten fiel die Hochregulierung von MC-II mit dem ersten starken Anstieg der Sonnenscheindauer des Jahres zusammen. Während der Anstieg von β-RpVPE und RpPAL1 sowohl dem Anstieg der Sonnenscheindauer als auch der Durchschnittstemperatur folgte. In der Übergangszone stieg die Genexpression von RpPAL1, RpCHS3 und γ-RpVPE synchron an. Das beobachtete Muster in der äußeren Übergangszone trat im Herbst einen Monat früher auf, als in der Inneren. Dies könnte einen Hinweis auf die Flussrichtung der Initiierungssignale der Kernholzbildung darstellen. Die vorliegende Arbeit zeigt, dass die Kernholzbildung alle lebenden Regionen des Holzes betrifft und im Laufe des Jahreszyklus stattfindet. Darüber hinaus ist die Kernholzbildung ein komplexer Differenzierungsprozess, der aus mehreren wichtigen biochemischen Prozessen und strukturellen Veränderungen besteht. Für den programmierten Zelltod scheint die Protease γ-VPE ein geeigneter Referenzmarker zu sein.

Table of Contents

List of Tables

Table of Abbreviations

Abbreviation	Full version of the word or phrase
ACE	ACGT-containing element
ACC	1-Aminocyclopropane-1-carboxylic acid
Acs	1-Aminocyclopropane-1-carboxylic acid (ACC) synthase
ACD	Accidental cell death
AEP	Asparagine endopeptidases
AP	Axial parenchymas
BLAST	Basic Local Alignment Search Tool
CARD	Caspase recruitment domain
caspase	Cysteine-dependent aspartyl specific protease
cDNA	Complementary DNA
CEP	Cysteine endopeptidases
CHS	Chalcone synthase
CODIT	Compartmentalization of decay in trees
CREs	Cis-regulatory elements
Ct	Cycle threshold
CTP	C-terminal pro-domain
DAMPs	Danger-associated molecular patterns
DED	Death effector domains
dPCD	Developmental programmed cell death
DWD	Deutscher Wetterdienst
ER	Endoplasmatic reticulum
EST	Expressed sequence tag
F3H	Flavanone 3-hydroxylase
FLS	Flavonol synthase
HR	Hypersensitive response
HRGPs	Hydroxyproline-rich glycoproteins
HWF	Heartwood formation
HWS	Heartwood substances
IAWA	International Association of Wood Anatomists
ITS	Internal transcribed spacer
ISW	Innermost sapwood
KDEL	Amino acid motif Lys-Asp-Glu-Leu
LAD	Ligase-activity determinant

Abbreviation	Full version of the word or phrase
LSAM	Legumain stabilization and activity modulation
LVs	Lytic vacuoles
MAMPs	Microbe-associated molecular patterns
MCs	Metacaspases
MLA	Marker of ligase activity
MRCA	Most recent common ancestor
MRE	MYB-binding responsive element
mRNA	Messenger RNA
MSV	Metabolite storage vacuoles
NCBI	The National Center for Biotechnology Information
NCCD	Nomenclature Committee on Cell Death
NSCs	Non-structural carbohydrates
NST	NAC secondary wall thickening promoting factor
NTP	N-terminal pro-domain
OD	Optical density
PAL	Phenylalanine ammonia lyase
PCD	Programmed cell death
PCR	Polymerase chain reaction
PPL	Poly-proline loop
PRRs	Pattern recognition receptors
PSVs	Protein storage vacuoles
qPCR	Quantitative polymerase chain reaction, Syn. Real-time PCR
RCD	Regulated cell death, Syn. programmed cell death
RE	Relative gene expression
ROS	Reactive oxygen species
RP	Ray parenchymas
RRE	R response element
SCW	Secondary cell walls
SM	Secondary (or specialized) metabolites
SNBEs	Secondary wall NAC binding elements
SND	Secondary wall associated NAC domain
SWNs	Secondary wall NAC domain transcription factors
TEM	Transmission electron microscopy
TEs	Tracheary elements
TFs	Transcription factors

Abbreviation	Full version of the word or phrase
TZ	Transition zone
UTR	Untranslated region
VND	Vascular-related NAC domain
VPEs	Vacuolar processing enzymes
VSR	Vacuolar sorting receptor
XCPs	Xylem cysteine proteases
XP	Xylem parenchyma
XPN	Xylem parenchyma network

1. INTRODUCTION

In many trees, heartwood formation (HWF) is the final stage of stem development that impart natural preservation to its core. The process occurs in the deep interior of the secondary vascular tissues, where xylem parenchyma (XP) cells undergo a secondary differentiation process similar to wounding response. Once the process is completed, all wood cells in the region are dead, but their corpses are reinforced with defense substances against biotic attacks. Heartwood poses clear economic incentives due to the natural decay resistance in timber and the products derived from their extractives, such as fragrances (Kampe & Magel, 2013; Singh & Singh, 2012). From the ecological perspective, the mechanism behind such formation reveals one of the trees' most crucial defense strategies that contributes to their resilience and remarkable longevity.

The lack of living cells in heartwood led to the debate on their physiological value in living trees (Bamber, 1976; Hillis, 1987; C. M. Stewart, 1966). Studies on the allocation of carbohydrates, deposition of extractives, and activity of key biosynthetic enzymes indicated deliberate translocation of carbon recourses to the inner region and conversion of reserve materials into heartwood substances (HWS; Bergström *et al.*, 1999; Beritognolo *et al.*, 2002; Hauch & Magel, 1998; Magel *et al.*, 1994). In addition to the genes associated with the biosynthesis of HWS, a list of genes responsible for signal transduction was up-regulated (Lim *et al.*, 2016; Yang *et al.*, 2003; Yeh *et al.*, 2020). These findings confirmed that the process of HWF is under intrinsic hormonal and genetic regulation; recent works described the formation as a natural and active developmental process and proposed the physiological roles of heartwood in trees (Kampe & Magel, 2013; Spicer, 2005).

Woody stems in trees are mainly composed of dead xylem cells. During xylogenesis, tracheary elements (TEs) and fibers undergo vigorous production of secondary cell walls (SCW) and lignification (Funada *et al.*, 2016). Upon completion of building SCW is the removal of cytoplasmic content, as TEs develop into hollow pipe-shaped conduits, and fibers mature into bricks-like structural support. These two cell types constitute the main volume of wood, and their lignified cell corpses establish the

infrastructure for the essential secondary vascular functions of long-distance water conduction. Interwoven among them is an intricate network of living XP cells, which is responsible for the metabolic activities in the mature sapwood (Morris, 2016; Spicer, 2014).

This living XP network (XPN) is essential for tree functioning and survival in response to environmental variability (von Arx *et al.*, 2015). The flow and content of the sap are regulated by the XP cells that surround the conduits (Brodersen & McElrone, 2013; Clearwater & Goldstein, 2005). Carbon resources gathered from photosynthesis are under the management of the XPN cells; the surplus is stored as reserve substances in it, and the XP cells carry out their conversion and translocation to developing organs in the time and body parts of need (Magel *et al.*, 2000). XP cell also provides dynamic responses in the events of infection and wounding for damage control and allow new growth to resume (Shigo, 1984). HWF carried out by the aged XP cell at the inner region offers preservation to the trees' internal structure, thereby sustaining continuous lateral expansion.

Like wood formation near the cambium, HWF is a complex differentiation process characterized by SCW modification and programmed cell death (PCD). PCD is a pivotal mechanism amongst multicellular organisms and, thus, a highly conserved process on the molecular level. The different xylem cell types undergo developmental PCD (dPCD) that varies in the timeframe, spatial position, and initiation triggers. However, they all undergo a large-scale synthesis of phenolic compounds and their deposition to the cell walls, which are tightly regulated in conjunction with cell death. Hence, the molecular machinery involved in these processes is comparable.

Several regulatory proteases that act as initiators in plant dPCD are examined in this work. We herein discuss how they link together the different aspects of wood and heartwood formation in the differentiation process. Specimens were collected over an annual cycle monthly for an overview of their involvement during the secondary xylem's full development. In addition, the physiological significance of the xylem parenchymas in trees is reviewed to connect our understanding of the ecological values of heartwood formation.

2. STATE OF KNOWLEDGE

2.1 Heartwood

2.1.1 Definitions

Heartwood, as defined by the International Association of Wood Anatomists (IAWA), is "the inner layers of wood, which, in the growing tree, have ceased to contain living cells, and in which the reserve materials (e.g. starch) have been removed or converted into heartwood substances." (IAWA, 1964). While converting reserve materials to antimicrobial substances also occurs during the localized wounding response induced by injury or infection, heartwood is distinct with its seasonal initiation and the anatomic position at the innermost living layers of the stem.

The region where sapwood transforms into heartwood is termed the transition zone (TZ; Hillis, 1987; Bergström, 2003), or intermediate wood in the past (IAWA, 1964). The sap- and heartwood boundary tends to loosely follow the growth rings, and the width varies significantly between species (Hillis, 1987). In some trees, the boundaries are conspicuous and abrupt, while others appear in a gradient. In the research experiments that employed methods in molecular biology, the TZ is often defined by the active production of HWS (Beritognolo *et al.*, 2002; Lim *et al.*, 2016; Yang *et al.*, 2003). For example, in the gene expression profiling study for HWF with scot pine, TZ was defined as the region where stilbene synthase encoding transcripts are most highly expressed (Lim *et al.,* 2016). In the recent two decades, a great variety of advanced microscopes have been employed to allow visual assessment of the subcellular location of HWS depositions, which is reviewed in Kampe and Magel (2013). The specimens for microscopic studies are dead in most cases; they are often air-dried, and TZ is defined by the deposition of HWS.

HWF takes heavy investment in trees in terms of carbon resources and is a complex developmental process involving multiple major cellular processes across the different regions within the stems. In its biological relevance, the transformation focuses on the two essential cellular processes that define its formation: (1) the conversion and

deposition of reserve materials into bioactive substances and (2) the cessation of containing living cells. Compared to the amount of research studying HWS production and deposition, investigation focusing on the aspects of PCD during the process is rare. The various cellular processes of HWF are generally not considered separately. However, recent evidence proposed how various component of HWF occurs in different seasons. Taking the softwood example of *Larix kaempferi*, Nakada and Fukatsu (2012) defined three components of HWF and concluded that the progress of HWF at a given place follows the order in the disappearance of free water from tracheid lumen, then the ray parenchyma cells die between April and July when the cambium is most active, and finally HWS deposition in late autumn to early winter.

2.1.2 Heartwood color

The color of heartwood is determined by the species-specific HWS, which varies within the same species by the growing conditions of the trees (Moya & Marín, 2011; Taylor *et al.*, 2002). Some heartwoods are conspicuous, such as that in pine (*Pinus silvestris* L.), larch (*Larix decidua* Mill.), yew (*Taxus baccata* L.), redwood (*Sequoia sempervirens* Endl.), and black locust (*Robinia pseudacacia* L.). However, some trees produce uncolored heartwood, such as spruce (*Picea abies*), hornbeam (*Carpinus betulus* L.), and linden (*Tilia cordata* Mill.). The major contributing factor to the darker shade in heartwood is the higher contents of tannins and phenolic compounds, while the molecules typically associated with wood durability fall into one of several polyphenolic classes, such as flavonoids and stilbenes. The color or shade of heartwood does not define its durability, but it has an aesthetic appeal and association with other desired heartwood properties, including higher density and decay resistance. Therefore, color makes a major determining factor for economic values in timber trading.

There is little difference in the chemical composition of HWS between juvenile heartwood formed by the younger trees and mature heartwood in older trees (Dünisch *et al.,* 2010; Magel *et al.,* 2001). Diameter growth and tree age have been excluded as the major contributing factors (Nelson *et al.*, 1969). Recent evidence

suggested that parameters of wood color are influenced by soil properties such as moisture content, acidity, available phosphate, total nitrogen, and exchangeable metal ions. Specifically, soil magnesium content has been proposed as an accelerator of phenolic accumulation (C. P. Li *et al.*, 2017), and high soil magnesium content has been observed to result in more reddish and yellowish heartwood in Juglans species (Makoto *et al.*, 2021).

2.1.3 Models and types of heartwood formation

Different types of HWF have been described based on three angiosperm models, with *Robinia pseudoacacia* for Type-I (Magel *et al.*, 1994), *Juglans nigra* for type II (Beritognolo *et al.*, 2002), and *Santalum album* for a debatable type III HWF (Celedon & Bohlmann, 2018). Black locust *(R. pseudoacacia*) represents the classic obligate heartwood forming species and is taken as an example to describe Type I HWF. This type of formation is characterized by a *de novo* and *in situ* synthesis of phenolic heartwood substances within the TZ (Magel *et al.*, 1994), meaning that all biochemical steps from the breakdown of reserve materials to finalizing with the end-product as HWS occurs within the narrow transitioning region. Black walnut (*Juglans nigra*) makes the exemplar for describing the Type-II, or *Juglans*-type, HWF (Beritognolo *et al.*, 2002). Instead of converting reserve materials into HWS on a singular occasion, this model features the gradual decrease in reserve substances in the aging sapwood coincides with an accumulation of glycosylated precursors, which then transform into the characteristic HWS within the TZ through oxidation and hydrolysis.

HWS are chemically diverse; many trees produce a cocktail of compounds via distinct metabolic pathways and in different cell types for the purpose. Tropical sandalwoods (*Santalum album*) and Alaska yellow cedar (*Callitropsis nootkatensis*) are examples that produce terpenes as the signature HWS in addition to phenolics. Their heartwood extractives are rich in sesquiterpenes and pose economic significance in the flavor and fragrance industry. Based on the example of sandalwoods, Celedon *et al.* (2016) describe a third model of HW formation, in which teraponids are synthesized in the metabolically active "heartwood". It is noteworthy that the authors defined heartwood

as the region with an accumulation of heartwood extractives (Celedon *et al.,* 2016). When adhered to the definition that most extant literature is based on, the region with any living parenchyma cells still undergoing the transformation would be considered the TZ; only the region that has completed the process and no longer contains living cells would be regarded as heartwood.

Nonetheless, Celedon & Bohlmann (2018) drew attention to how the various constituents of HWS may synthesize differently. Flavonoids and teraponids are synthesized in separate specialized XP cells, in which their transitions are scheduled on a different time frame and position. Phenolic substances are first produced during an earlier transition stage in the outer realm, and teraponids at a later stage in the inner region. Such observation has been reported in other trees. With a gymnosperm example of Scots pine, the signature HWS are stilbenes and resin acids (Saranpää & Nyberg, 1987), which are also produced in different temporospatial patterns (Lim *et al.,* 2016). Transcriptomic analysis indicated the *de novo* and *in situ* synthesis of stilbenes at the TZ, while resin acids are synthesized early in the spring, mainly in the sapwood (Lim *et al.,* 2016).

2.2 Wood formation and physiology

2.2.1 Physiology of the secondary vascular system

Primary growth in plants is the development and expansion at the tips, whereas secondary growth results in lateral thickening in the stems (Figure 1). In the evolution of secondary vascular development, there are great variations in the arrangement of vascular cells within the stems (Spicer & Groover, 2010). The secondary growth in typical woody stems is made by the vascular cambia, a thin meristem layer that lies at the periphery beneath the bark that regenerates perpetually. For secondary growth, it produces a complex matrix of specialized cells in a bidirectional manner that gives rise to secondary phloem to the outside centrifugally and secondary xylem to the inside centripetally (Larson, 1994). The phloem cells eventually become parts of the

bark and shed off the trees. The matrix of xylem cells accumulates throughout every growing season, resulting in radial thickening as wood. The extant forest trees around the world are predominantly gymnosperms and dicotyledonous angiosperms, both form wood in this manner. Other forms of secondary vascular development that support considerable height and mass but do not form wood are seen in tree ferns and tree-like monocots, such as palms and bananas. In comparison, woody plants show the most evolutionary success in terms of species diversification, achieving mass and height, and organismal longevity (Spicer, 2016).

This architectural arrangement of secondary vascular cells that result in wood offers a long list of advantageous adaptive traits in the territorial landscape (Spicer & Groover, 2010). There are three main types of cells in the secondary xylem: TEs (tracheids in gymnosperms and vessel elements in angiosperms), fibers, and parenchyma cells (Larson, 1994). TEs have straw-like anatomy at maturity, and their connection makes up the pipeline of vascular conduction (Escamez & Tuominen, 2014). Fibers have robust secondary cell walls, and their brick-like anatomy constitutes the main mass that offers rigidity and strength to the stems. Xylem cell types are less specialized in gymnosperm, with tracheids comprising over 90% of the total wood volume and serving the combined functions of mechanical support and water transport (Higuchi, 1997). In angiosperms, the cells are more specialized as vessels are exclusively for water conductance, while fibers establish mechanical stability. While TEs and fibers lay down the infrastructure for long-distance transport, the XP cells serve a diverse suite of living functions, which are discussed in the next chapter (2.3 The living functions of the xylem parenchyma network system). The different cell types are produced in a highly ordered arrangement that compartmentalizes the stem, which holds the capacity for long-distance water transport, water and nutrients storage, resistance to drought and freeze, and recovery from embolism (Choat & Pittermann, 2009; Domec *et al.*, 2008; Shigo, 1984; Spicer, 2016).

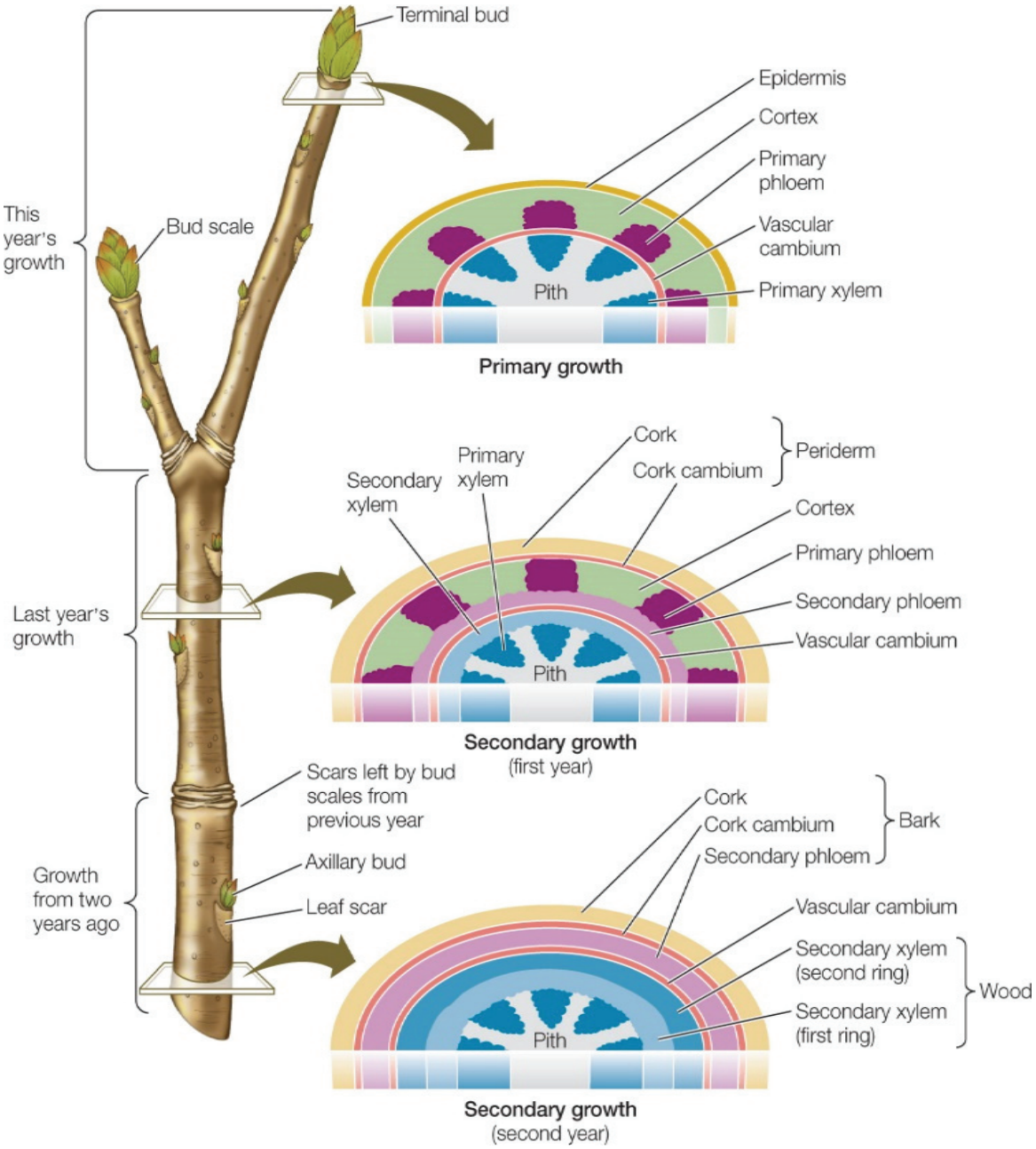


Figure 1: Schematic illustration of the woody eudicots' primary and secondary vascular growth (*Secondary Growth: The Vascular Cambium*, n.d.).

2.2.2 Molecular aspects of xylogenesis

The secondary xylem cells are formed through a complex developmental program to achieve physiological maturity. After the division from cambial mother cells and elongation, the differentiation is carried out through a meticulous orchestration of biosynthesis of secondary cell wall (SCW) components, their deposition, lignification,

and PCD (Fukuda, 1996; Higuchi, 1997). Vessels in angiosperms undergo dPCD with rapid clearance of cytoplasm shortly after their division at the vascular cambia (Bollhöner *et al.*, 2012). In the differentiating TE cell cultures of Zinnia and Arabidopsis, the cells deposit cellulose and hemicellulose into the SCW for 10–16 hours upon division. Then the cell remains alive for another 2–6 hour before committing suicide, and proceed to lignification (Iakimova & Woltering, 2017; Pesquet *et al.*, 2010, 2011). Tracheids in gymnosperms and fibers in angiosperms share close ontogeny and morphological changes during their differentiation; these cells take longer to develop multiple layers of SCW (Courtois-Moreau *et al.*, 2009). Genomic analysis revealed that the process is backed by the coordination of thousands of genes, involving functions in signal transduction, transcriptional regulation, biosynthesis of cell wall components, cytoskeletal organization, and PCD (Andersson-Gunnerås *et al.*, 2006; Dharmawardhana *et al.*, 2010; Prassinos *et al.*, 2005; Wang *et al.*, 2009). A significant part of orchestrating the expression of such a vast amount of genes in xylogenesis occurs at the transcriptional level (Zhong & Ye, 2013).

Transcription is the molecular process of copying a specific segment of DNA from the genome, usually a gene, into short RNA sequences for the downstream synthesis of functional gene products. These short RNA sequences are termed messenger RNA (mRNA), and they are synthesized by RNA polymerases (Kornberg, 2007). Multiple specific sequences outside of a gene's coding region participate in the modulation of its transcription; they are named cis-regulatory elements (CREs). Enhancers are CREs that reside at some distance from the gene, while promoters lie directly upstream of the coding region. Table 1 lists several common wood associated cis-regulatory elements (CREs; Zhong & Ye, 2013). Transcription factors (TFs) are sequence-specific DNA-binding proteins that bind to CREs to either promote or block the recruitment of RNA polymerases (Conaway & Conaway, 1997). As the prerequisite to initiating the transcription of a gene, its CREs must be bound sequentially with the appropriate set of TFs, whose assembly is facilitated by mediator proteins, for RNA polymerase to be in position and form the transcriptional preinitiation complex (Figure 2).

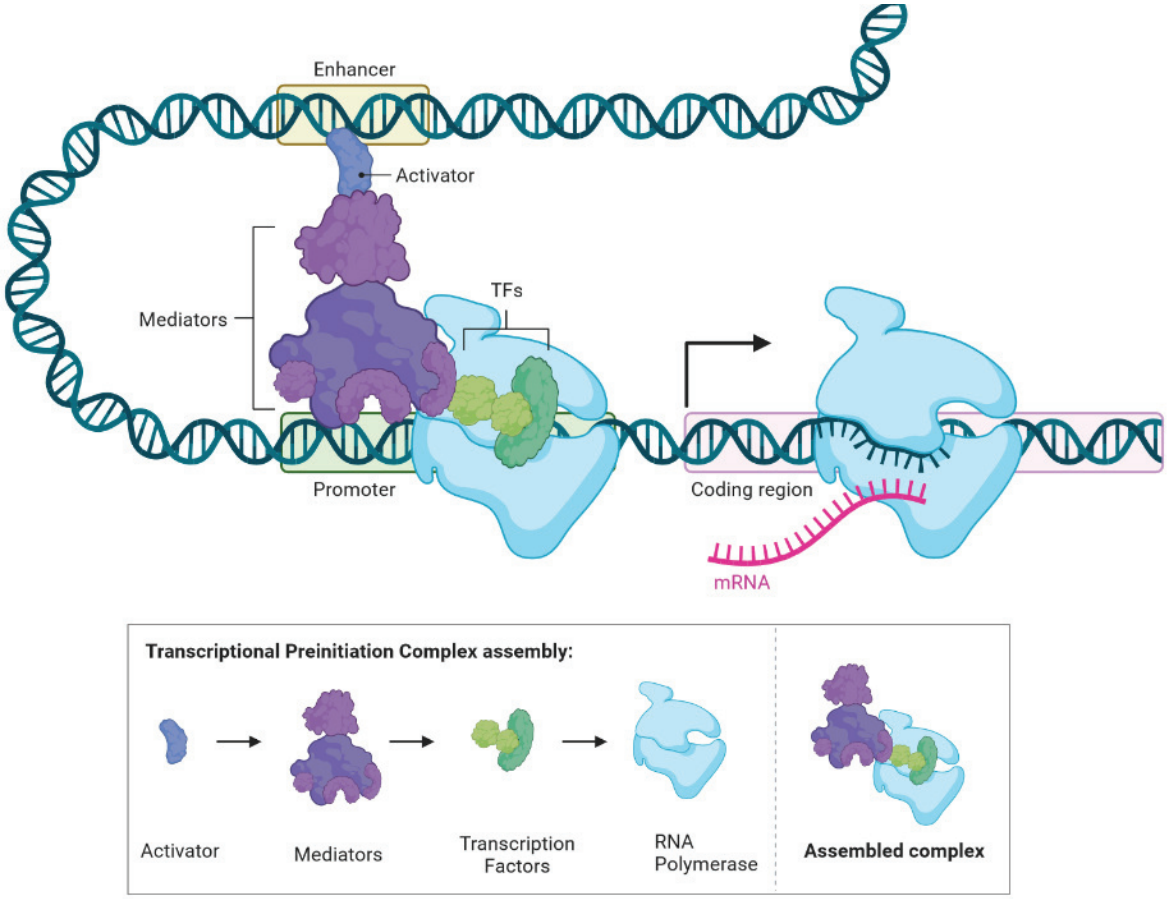


Figure 2: Schematic illustration of transcriptional initiation. The process begins with binding the appropriate set of transcription factors (TFs) to the enhancer and promoter sequences. Then, with the aid of mediator proteins, RNA polymerase is recruited to the preinitiation complex. Finally, RNA polymerase exits the assembly and starts transcribing the gene's coding region into messenger RNA (mRNA).

Functional studies revealed that gene expression for the biosynthesis of SCW components is concerted via the interaction between wood-specific TFs and CREs (Hertzberg *et al.*, 2001; Zhong & Ye, 2013). The SCW are composed of three major types of polymers: cellulose, hemicelluloses, and lignin; the biosynthesis of SCW requires the coordinated activation of all genes involved in the production of these major components. The MYB family is one of the largest TF families in plants; their members are key players in the regulatory networks in various physiological and biochemical processes. Bioinformatic analysis identified a list of MYB-TFs that regulate the genes responsible for the biosynthesis of all three main cell wall components (J. Liu *et al.*, 2015), while an additional set of TFs separately affects only lignin production (O. Wilkins *et al.*, 2009; Zhong & Ye, 2007). Lignin, the chief constituent of wood, is a complex phenylpropanoid polymer of three monolignols in

varying compositions depending on tree species (Boerjan *et al.*, 2003). The production of monolignols starts with the general phenylpropanoid pathway, and most phenylpropanoid biosynthetic enzymes contain AC-rich elements and MYB-binding responsive elements (MRE) in the promoters (Figure 3; Raes *et al.,* 2003). MYB-TFs recognize target genes through these elements, and their interaction determines the activation or suppression of their transcription. The same mechanism also controls transcriptional initiation for cellulose and hemicellulose biosynthetic genes but with a different set of wood-specific MYB-TFs. A number of putative CREs in the promoters of cellulose synthase identified via bioinformatic analysis revealed that their presence in the orthologous is evolutionarily conserved among angiosperms (Creux *et al.*, 2008).

Table 1: Examples of wood associated cis-regulatory elements (CREs).

Name of CRE	Abbr.	Size	Sequence	Reference
AC element	AC	7-bp	ACC (A/T)A(A/C)C	Hatton *et al.,* 1995
Secondary wall MYB-responsive element	SMRE	7-bp	ACC(A/T)A(A/C)(T/C)	Zhong and Ye, 2013
Secondary wall NAC binding element	SNBE	19-bp	(T/A)NN(C/T)(T/C/G)TNNNN NNNA(A/C)GN(A/C/T)(A/T)	Zhong *et al.,* 2010b

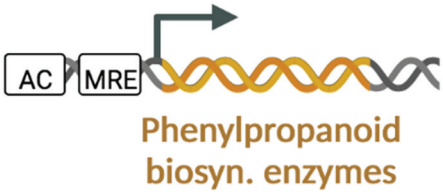


Figure 3: Schematic illustration of the genetic structure of phenylpropanoid biosynthetic enzymes. Arrow indicates the site for transcription initiation. White elements present the promotor regions: AC-rich elements (AC) and MYB-binding responsive elements (MRE).

Upstream of the MYB-TFs are the transcriptional master switches that link SCW production to PCD. While TFs bind directly to the CREs of genes to modulate their expression, transcriptional master switches control groups of TFs that are functionally and cell-type specific, they are named secondary wall NAC domain transcription factors or collectively known as the SWNs (Yamaguchi *et al.*, 2008; Zhong *et al.*, 2010). Examples of cell type-specific SWNs are vessel-specific vascular-related NAC domain (VND), fiber-specific secondary wall-associated NAC domain protein (SND), and NAC secondary wall thickening promoting factor (NST). In addition to the modulation of SCW component biosynthesis, SWNs activate a suite of genes involved in SCW

modification and PCD (Figure 4). For the deposition of secondary walls, the SWNs bind to the secondary wall NAC binding elements (SNBEs) in the promoters of target genes (Zhong *et al.*, 2021). For detailed elaboration on the SWN-mediated transcriptional network, readers are referred to Zhong & Ye (2015).

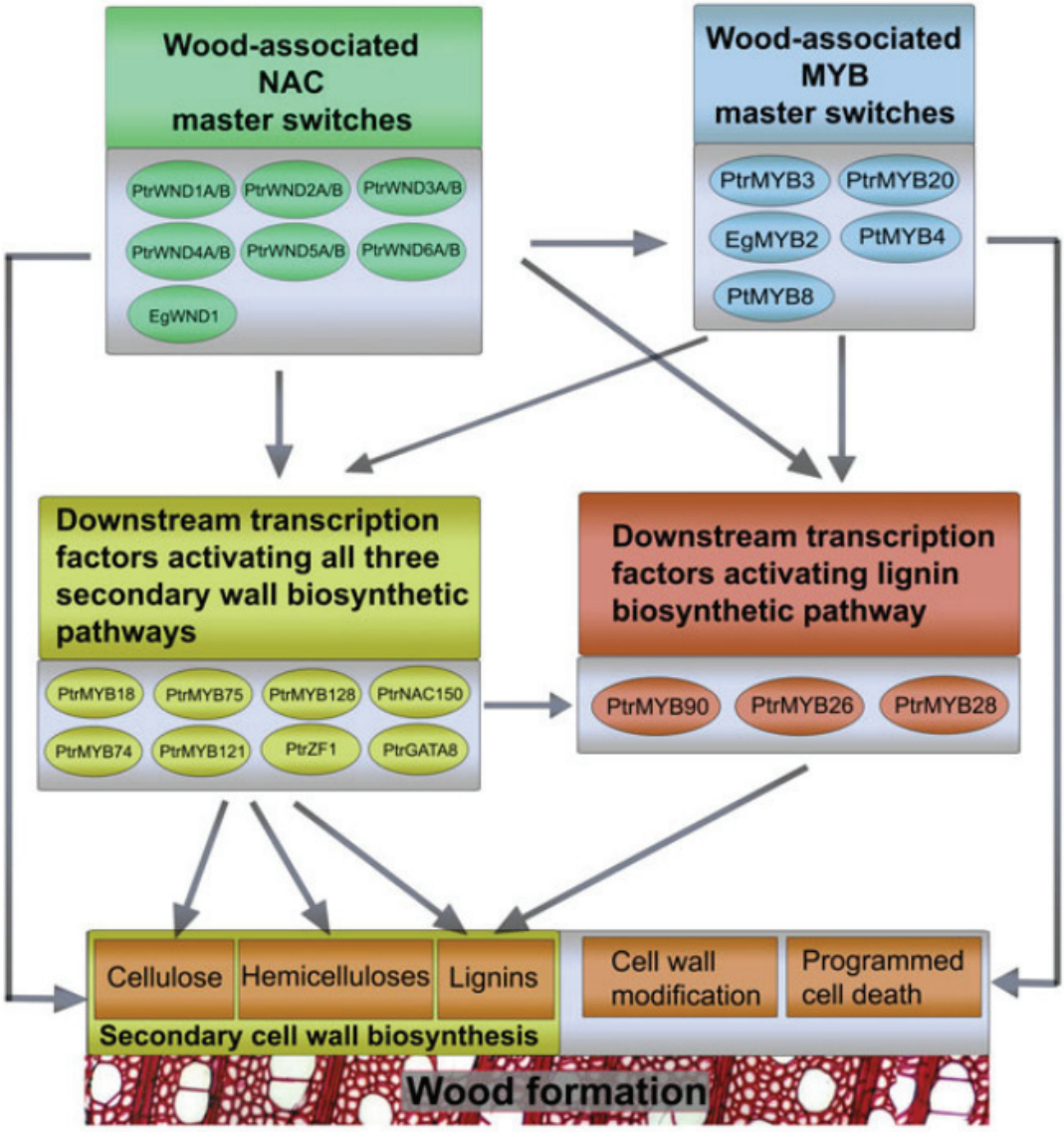


Figure 4: Schematic illustration of the transcriptional regulation network in xylogenesis (Zhong & Ye, 2015).

2.2.3 The xylem parenchyma network system

Parenchyma cells are one of the three ground tissues in plants. Except in the secondary xylem, they mostly have thin primary cell walls that offer flexibility to structural changes in the cell shape (W. N. Stewart & Rothwell, 2009). This cell type is alive and totipotent at maturity; they may contain specialist organelles, such as chloroplasts for photosynthesis and amyloplasts that accumulates starch(Simpson, 2019). Due to their association with specialist organelles and the ability to dedifferentiate, parenchyma cells have a very diverse suite of functions, including metabolic processes such as photosynthesis, storage of non-structural carbohydrates (NSCs), water reservoirs, wound healing, and regeneration (Evert, 2006).

Xylem parenchyma cells are produced by the vascular cambium, along with TEs and fibers. Ray parenchyma (RP) strands are connected ray cells developed from the ray initials of the cambium; they have symplastic connections via the cytoplasm that allow active transport of materials and inter-cellular communication. The rays run centripetally across the pith to the bark showing a wheel-like appearance in the cross-section of a stem; they connect the phloem and xylem, providing symplastic pathways for materials exchange between the two (S. Salleo *et al.*, 2004; Sokolowska & Zagórska-Marek, 2012; Spicer, 2014; van der Schoot & van Bel, 1990). The fusiform initials of the vascular cambium develop into axial parenchyma (AP) strands that orient somewhat parallel along the stem, which share connections with the vascular conduits through half-bordered pits (Chaffey & Barlow, 2001). The two types of strands run perpendicular to each other and form a highly interconnected three-dimensional XP network (XPN) system throughout the stem (Figure 5; Brodersen, 2013; Morris *et al.*, 2016; Spicer, 2014).

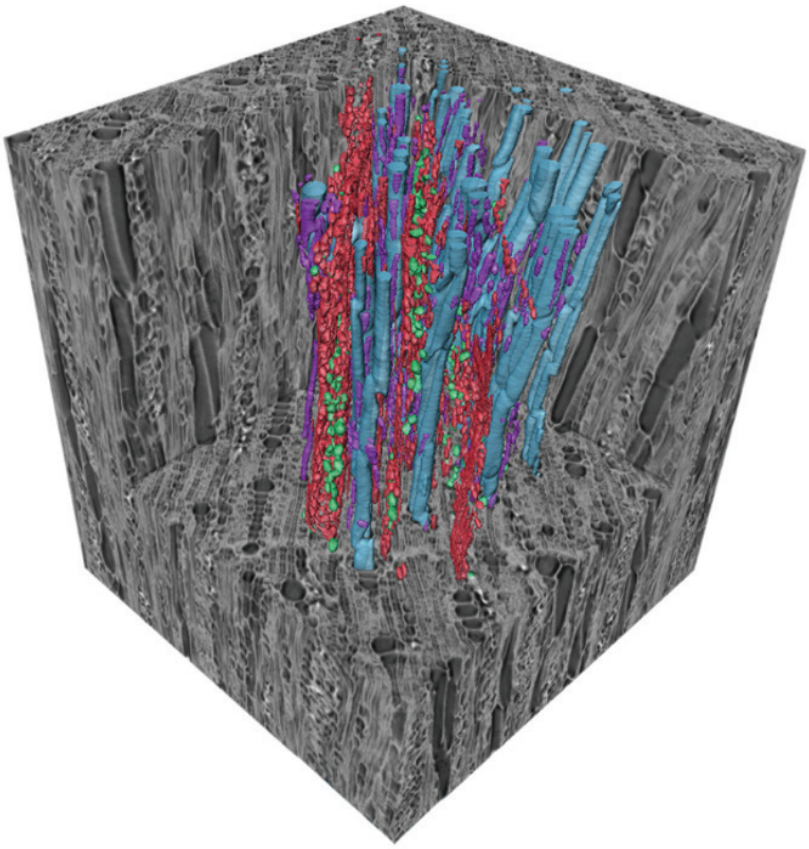


Figure 5: Three-dimensional model of wood image showing the XPN (Morris *et al.*, 2016a)**.** The grey cube shows the wood anatomy of *Ziziphus obtusifolia stem*, vessels are shown in blue, the rays are marked in red, axial parenchymas in purple, and green dots indicates the location of defense chemical distributed throughout the ray parenchymas.

2.3 The living functions of the xylem parenchyma network system

Vessels and fiber cells are dead in the mature conducting sapwood; interwoven among them is a living XPN system, which directs all live functions that require metabolic activities (Magel, 2001; Spicer & Groover, 2010). The entirety of the XPN interconnects throughout a tree, radially from pith to bark and axially from roots to the tips of branches. XP cells are categorized into contact and isolate cells; the former has direct contact with the conduits, whereas the latter is the contrary (Spicer, 2016). Although the various highly specialized XP coalesces to a lesser portion of wood by volume, the XPN system bears the critical duties in resource management throughout the seasonal fluctuation, maintaining stable vascular conduction, and recovery response against environmental damages (Carlquist, 2018; Holbrook *et al.*, 2002; Magel, 2001; Morris, 2016; Spicer & Groover, 2010).

2.3.1 Carbon resource management

The XPN system manages the carbon resources in response to seasonal fluctuations to facilitate the growth and development of various tree organs (Essiamah & Eschrich, 1985; Magel *et al.*, 2000). As perennial plants undergo different developmental stages according to seasonal clues in the annual cycle, energy demand arises in various sites with distance and time apart. These include bud bursts in spring and dormancy in winter; in angiosperms, there are also the production of flowers, fruits, and seeds. In between the photosynthetic tissues as carbon sources and developing organs as the sink, reserve substances in trees are mainly stored in the XP cells. In gymnosperm, such storages are located in the rays with efficient routes for radial translocation. Storage substances are primarily parked in the isolated cell, and their release is mediated through contact cells into the conduits. In addition to being the primary site of reserve storage and providing means of carbon translocation, the XPN also manages the interconversion between transportable soluble carbohydrates and their storage forms, such as starch, triacylglycerides, and storage proteins. The convention is carried out by a list of sucrose and starch metabolizing enzymes; readers are referred to Magel *et al.* (2000) for a comprehensive review.

2.3.2 Maintenance of hydraulic pressure and embolism repairs

Contrary to the cohesion-tension theory (Dixon & Joly, 1895), a large body of evidence indicates that long-distance water transport in the secondary xylem is not entirely a passive process (Brodersen & McElrone, 2013; Clearwater & Goldstein, 2005). While TEs and fibers make up the infrastructure in the sapwood for water transport, the regulation of sap flow requires input from living cells (Holbrook *et al.*, 2002). Although the exact mechanisms await further elaboration, evidence has demonstrated that XPN involves hydraulic maintenance by regulating the sugar content, pH level, and ion movement of sap (Nardini *et al.*, 2011; Salleo *et al.*, 2009; Secchi & Zwieniecki, 2011). In the case of embolism, the contact XP cells can release sugars into the affected vessels through the proton pump mechanism, increasing the sap osmolarity and thus xylem

pressure (Améglio *et al.*, 2004; Essiamah & Eschrich, 1985). XP cells also control the acidity of tree sap (Fromard *et al.*, 1995). The pH level of tree sap varies along the annual cycle in temperate tree species; for example, it is more acidic in early spring during bud bursts (Ferguson *et al.*, 1983; Sauter, 1988). While proof of theory awaits, it has been speculated that the lower pH level facilitates the movement of ions and organic nutrients between contact cells and vessels along a pH gradient (Larsson & Møller, 1990).

Tylosis formation has been suggested as another mechanism for maintaining hydraulic pressure carried out by the XP cells. The secondary cell walls of XP cells lack the complex laminar structure and resemble a thickened primary cell wall (Panshin & Zeeuw, 1980). Tylosis are cytoplasmic extensions from the contact XP cells that grow through the cell wall pits into the lumen of neighboring vessels; such formation renders the affected wood region impermeable to both liquids and air (Fujita *et al.*, 1978; Gerry, 1914). They are often considered in association with HWF (Chattaway, 1949), but a clear correlation has not yet been identified. No morphological or biochemical characteristics of PCD have been found in association with tylosis growth, nor any sign of specialized metabolite production. Instead, recent reports showed tylosis occurs as a reaction to embolism and water stress, and such formation results in the recovery in the conduction (Morris *et al.*, 2016; Secchi *et al.*, 2017). Such a mechanism for embolism repairs also applies to the trees' reaction to pathogenic invasions, indicating its nature of function in damage control (Clearwater & Goldstein, 2005).

2.3.3 Defense and wounding response

XPN has the pivotal role of providing dynamic responses to infections and damage in trees, which is a significant factor in explaining the remarkable longevity of perennial woody plants and their capacity to survive countless attacks (Morris *et al.*, 2016; Shigo, 1984). Living sapwood has active and passive defense mechanisms; active defenses are induced by pathogens or mechanical wounding, while passive defenses are laid down as preventive measures. A tree's survival after injury or infection

depends significantly on its ability to compartmentalize the damages and continue with growth and development afterward (Shigo, 1984). The secondary vascular tissues in woody plants comprise an intricate arrangement of cells that compartmentalize the tree in multiple degrees. The lignified SCW of the xylem makes the basic units in compartmentalization, which is an example of the built-in passive defense.

Compartmentalization of decay in trees (CODIT) describes four walls of defensive response with increasing strength (Shigo, 1984). Wall 1 and 2 are the first defensive responses to be put in place by the XPN. Wall 1 restricts the axial movement of infection or embolism by forming occlusion in the conduits with tylosis or gels (Gerry, 1914; Shigo, 1984). Wall 2 halt the radial spread of pathogens with the chemical static lignified boundary. Wall 3 is the chemical alteration carried out by the RP at the time of wounding (Hillis, 1977; Shigo, 1984). In response to microbial invasions and physical injuries, the RP cells in the affected area use their storage reserves to fuel the synthesis microbe resistant polyphenolic compounds (Hillis, 1977). Furthermore, wounding signals are transmitted to the neighboring cells to activate their responses.

Walls 1 to 3 together form the reaction zone, which results in a physical and chemical boundary that impedes the spread of infections or damage (Shigo & Tippett, 1981). Figure 6 shows an example of a reaction zone in *R. pseudoacacia*. Wall 4 of CODIT describes the post-wounding response of establishing a barrier zone that isolates the necrotic tissues (Shigo & Tippett, 1981). XP cells in the barrier zone undergo suberization, depositing a layer of antimicrobial fatty sealant to the cell wall, making the site completely impervious (Biggs, 1987; Hillis, 1977; Schmitt & Liese, 1993). In the immediate vicinity of the injured area, the new RAP forms air-tight callus tissues, further sealing the damaged regions, and separating them from subsequent new growth outside (Sharples & Gunnery, 1933; Shigo, 1984).

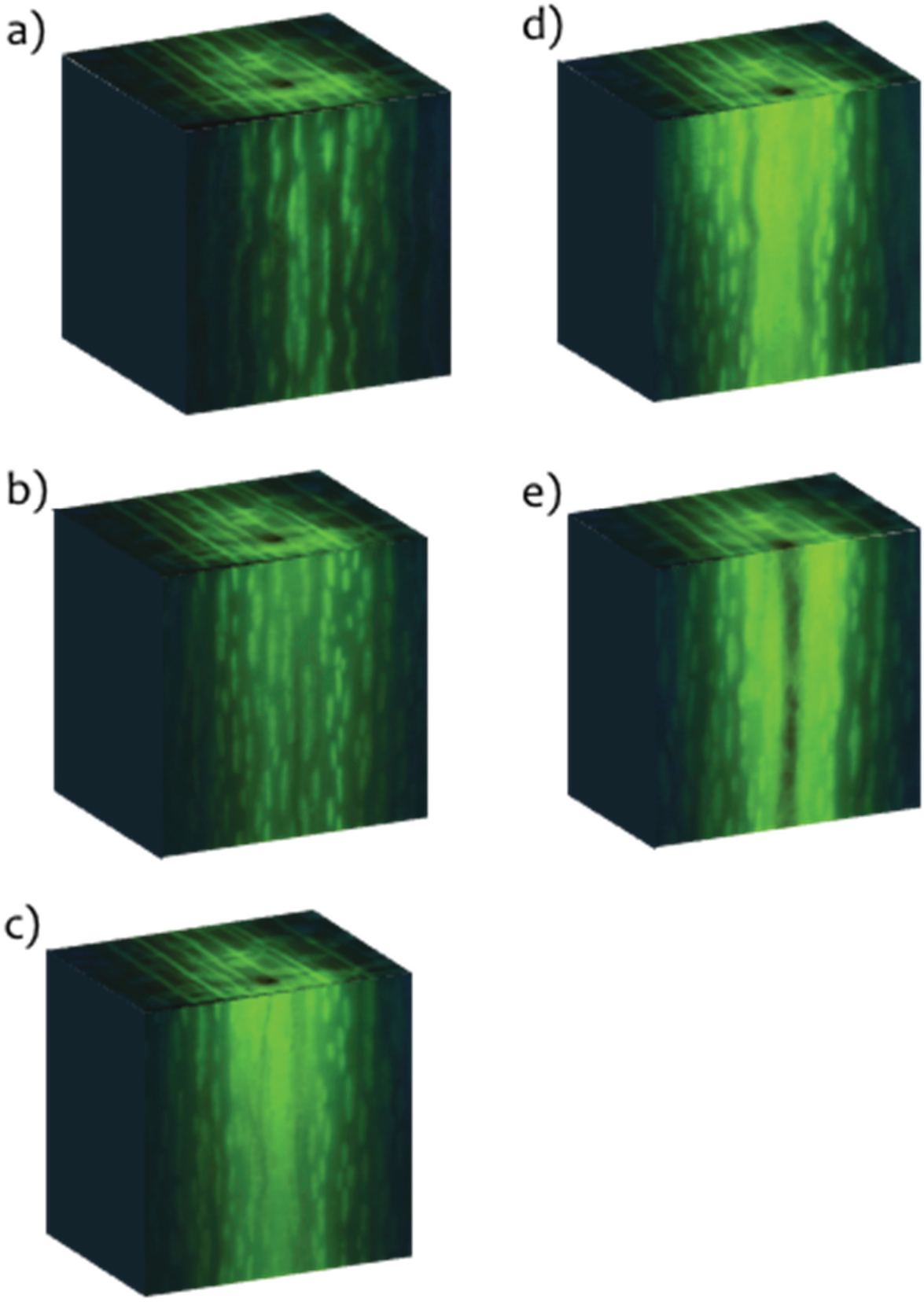


Figure 6: Wounding response in *Robinia pseudoacacia* shown with UV microscopy. The green color shows the natural fluoresces from the heartwood substances robinetin and dihydrorobinetin, indicating the chemical alternation of the parenchyma cells. Image a to e displays the tangential view at a successively closer distance to the wound site (dark hole/canal in the center of the fluorescent region) of about 50 µm increment.

2.4 Transformation into heartwood

In HW-forming trees, the innermost layer of XP cells goes through a secondary and terminal differentiation that couples with bulk SM production. The XPN has a significant role in active defense and repair, and the wood region outside its coverage is vulnerable to biotic attacks (Shain, 1995). As the innermost layer of the XPN ages, the cells face increasing stress, such as oxygen depletion and carbon dioxide enrichment (Hillis, 1968, 1987). In heartwood-forming trees, retiring cells at the inner layer of the XPN undergo a secondary differentiation similar to the wounding response, which offers permanent passive defense to the region (Hillis, 1987; Shigo, 1984). The cellular events include the breakdown of carbon reserve, biosynthesis of HWS, deposition of HWS into the cell walls, and PCD. On the organismal level, the allocation of carbon resources within the stem also occurs to facilitate the high energy demand of HWS production. We review the aspects of HFW that mainly revolve around the production of HWS in this chapter, and PCD is discussed in detail in the next (2.5 Programmed cell death).

2.4.1 Energy supply for the transformation

Starch or flavonoid precursors at the TZ make the primary source for *in situ* syntheses, while the energy required for production is predominantly supported by an import of sucrose from the outer wood region. (Magel *et al.*, 1994) The biosynthesis of HWS demands a high amount of energy in the transiting region (Hasegawa & Shiroya, 1968; Magel, 2000). During HWF, elevated activities of amylases and lipases occur in the outer wood regions. Storage substances there in the form of starch and fats are broken down into sucrose by hydrolytic cleavage, while carbohydrate transporter proteins are activated in the sapwood for their allocation through the rays to the transitioning XP cells (Figure 7; Magel 2000). The simple sugar is then further cleaved by sucrose synthase and neutral invertase in the TZ, mostly to provide for glycolytic and respiratory energy demand in the synthesis of SM (Magel *et al.*, 1994).

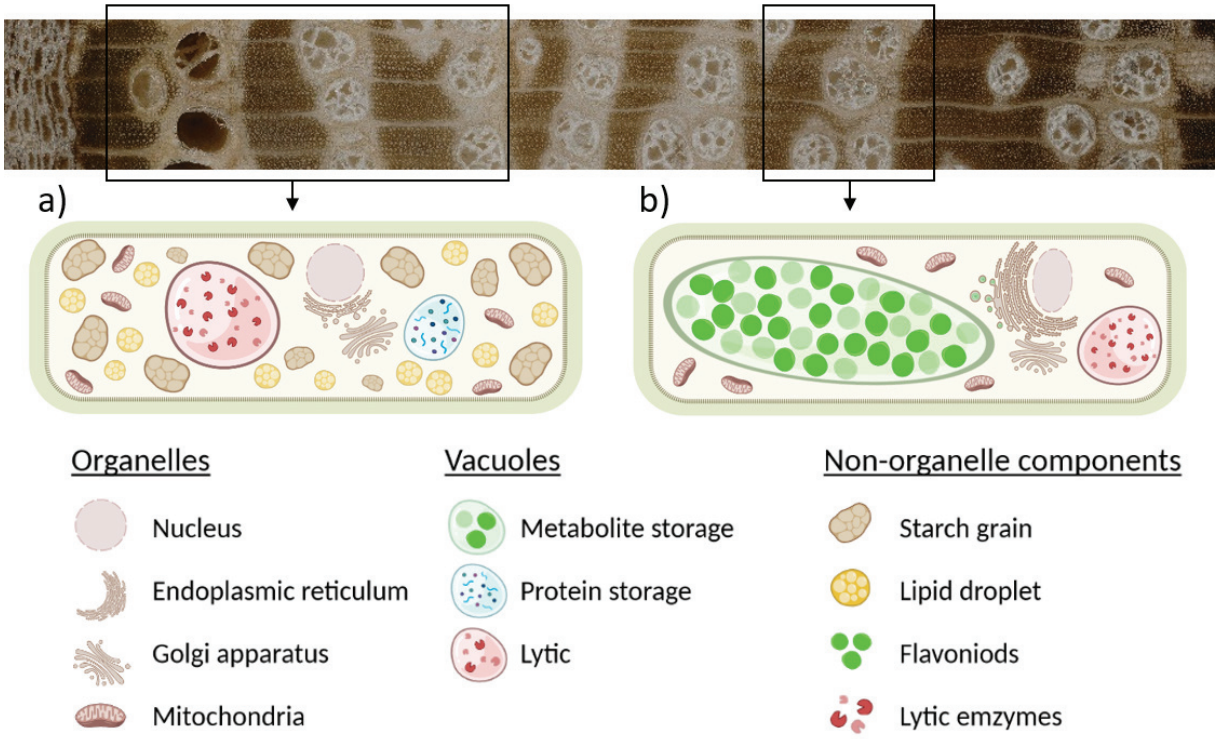


Figure 7: Illustration on the cellular states of ray parenchyma cells in relation to heartwood substance biosynthesis. a) Parenchyma cells in the outer region contain reserve materials, which are broken down into simple sugar during heartwood formation and translocated to the transition zone (TZ) through the rays. b) Parenchyma cells at the TZ under vigorous production of secondary metabolites.

2.4.2 Phytochemicals and heartwood substances

Plants are sessile on both the organismal and cellular levels; the essential life functions of communication and defense are predominately carried out with phytochemicals. With the simple sugars generated via photosynthesis as starting materials, plants synthesize a vast array of structurally diverse aromatic compounds through subsidiary pathways. These biomolecules are termed secondary or specialized metabolites (SM); they are widely distributed in various tissues and serve multiple purposes. Scents and pigments in flowers and fruits are SM accumulated in reproductive tissues to attract and interact with other organisms in the environment for propagation. Anthocyanins, for example, determine flower and fruit color to attract pollinators and seed dispersers (Landi *et al.*, 2015). However, a large portion of SM is defensive in nature; they are stored in plant tissues as deterrents and toxins to cope with the constant threats of

herbivores and infectious microorganisms (Bennett and Wallsgrove, 1994; Patten *et al.*, 2010).

All botanical lineages synthesize phenolic compounds and flavonoids, which function as developmental cues and defense substances against environmental stresses (Bennett & Wallsgrove, 1994). SM in wood are defined as everything that is not a structural polysaccharide or lignin (Obst, 1998); they are often referred to as extractives in forestry, including resin, gums, and HWS (Hillis, 1968). Both lignin and flavonoids are phenolic compounds that accumulate in the xylem; their biosynthesis occurs during the differentiation of distinct xylem cell types in response to environmental changes (Barros *et al.,* 2015).

Phenolic compounds are synthesized through the phenylpropanoid or acetate-malonate metabolic pathway; hence they share precursors and regulation mechanisms with those for lignin biosynthesis. The biosynthetic enzymes in the phenylpropanoid pathway contain MRE in the promoters of their genes, which allow the coordinated expression regulated by MYB-TFs (Raes *et al.,* 2003). The same regulatory mechanism on transcriptional initiation applies to the production of flavonols, only that they differ in combinatorial interactions. Flavonols and lignin production share the initial steps in the phenylpropanoid pathway (Figure 8). Downstream of the production of 4-coumarate-CoA, chalcone synthase (*CHS*) is the first enzyme that carries on towards the synthesis of flavonols, with the following successive steps catalyzed by chalcone flavanone isomerase (CFI), flavanone 3-hydroxylase (F3H) and flavonol synthase (FLS) (Weston & Mathesius, 2013). These four enzymes contain MRE in the promoter of their genes, the same as that of the phenylpropanoid enzymes. However, instead of the AC element, they have an ACGT-containing element (ACE) and an R response element (RRE). These CRES require the binding of a different set of TFs, such as BHLH and R2R3 MYB-TFs (Hartmann *et al.*, 2005; O. Wilkins *et al.*, 2009). The coordinated expression of *CHS*, CFI, *F3H* and *FLS* is based on functionally identical CREs recognized by the same TFs (Hartmann *et al.,* 2005).

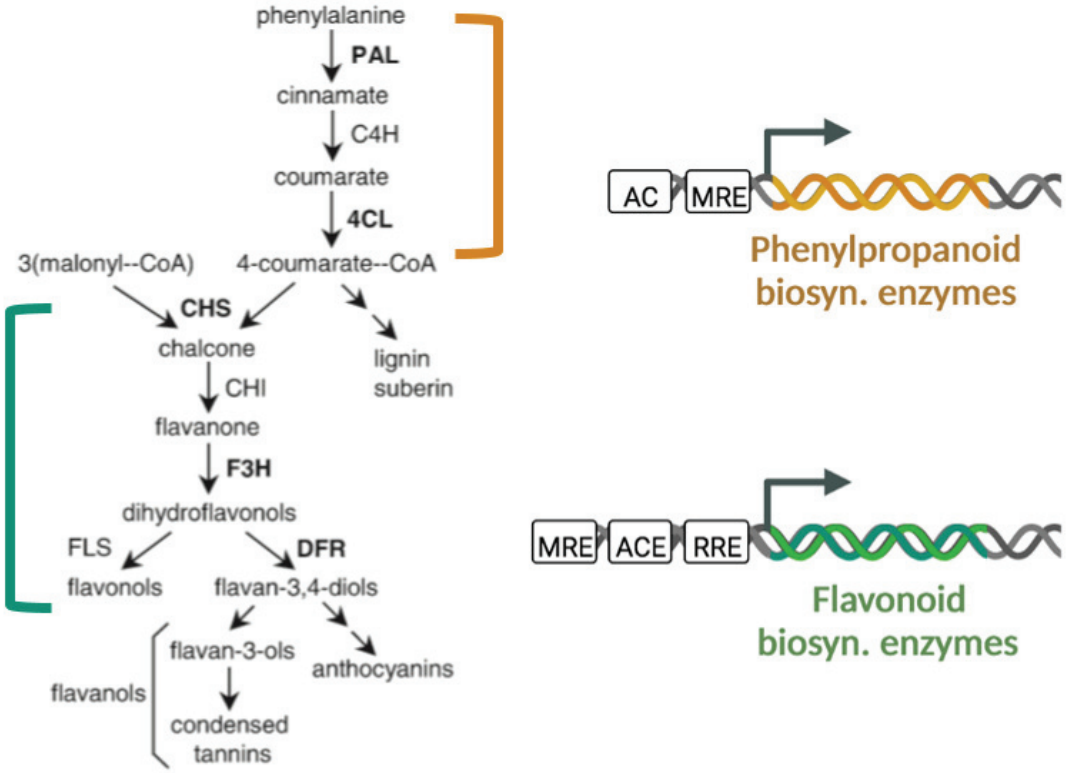


Figure 8: (Left) general phenylpropanoid and flavanols biosynthetic pathways. The biosynthesis of lignin and flavonols shared the initial steps in the phenylpropanoid pathway until 4-coumarate-CoA is produced. For flavonols, the downstream successive steps are catalyzed by chalcone synthase (CHS), chalcone flavanone isomerase (CFI), flavanone 3-hydroxylase (F3H), and flavonol synthase (FLS). **(Right) Schematic illustration of the phenylpropanoid and flavonoid biosynthetic genes.** Arrow indicates the start of the coding region. The white boxes indicate the cis-regulatory elements (CREs): AC-rich elements (AC) and MYB-binding responsive elements (MRE), ACGT-containing element (ACE), and R-response element (RRE).

2.4.3 Heartwood substances deposition

It has been well-established that HWS are most effective against wood degradation when deposited into the cell wall matrix (Hillis, 1968), but the mechanism of its deposition has remained unclear. In hardwood species, the accumulation of phenolic compounds seems to involve an enzymatic initiation, then a chemically driven copolymerization between the phenolic derivatives and the pre-existing cell wall polymers (Monties, 1991; Magel *et al.* 1995b). The genes encoding enzymes for cell wall modification and lignification in the outer wood region during wood formation were up-regulated in the TZ concomitantly with HWS biosynthesis (Bergström *et al.*, 1999; Yang *et al.*, 2003). In particular, the authors specifically pointed out the exclusive upregulation of extensins and extensins peroxidase at the TZ during HWS

biosynthesis in *R. pseudoacacia* (Yang *et al.*, 2003). In addition to a complex matrix of polysaccharides (cellulose and hemicelluloses) and polyphenols (pectin and lignin), plant cell walls also contain structural glycoproteins (Hoson, 1991). Extensins are a family of hydroxyproline-rich glycoproteins (HRGPs) that constitute the major protein components in the cell walls of dicot plants. Expression of the extensin genes is developmentally regulated and inducible by mechanical wounding, pathogen attack, or ethylene treatment (Corbin *et al.*, 1987; Parmentier *et al.*, 1995; Wycoff *et al.*, 1995).

The association of extensins in HWF is rarely documented and its relation to HWS deposition is unclear. However, its linkage with another cysteine protease associated with SCW building has been reported. The papain-like KDEL-tailed cysteine endopeptidases (CEPs) have been observed to involve SCW thickening and stem maturation during the differentiation of TEs and fibers (Han *et al.,* 2019). CPEs have broad substrate specificity, including proline and glycosylated hydroxyproline; this gives them the ability to digest extensins, which subsequently form the basic scaffold for cell wall formation (Helm *et al.,* 2008). A revisit of the topic is in chapter 2.6 (Regulatory proteases in plant programmed cell death), where its relevance to a candidate gene is further detailed.

2.5 Programmed cell death

2.5.1 Definitions and types of programmed cell death

Programmed cell death, synonymous with regulated cell death, is the process when cells commit suicide by activating an intracellular death program (Galluzzi *et al.*, 2018). It is a ubiquitous phenomenon and an indispensable mechanism in multicellular organisms but is also observed in colonies of single-cell organisms. In metazoans, an organism is an individual constituted by a highly organized cell community; the population and functionality of cells in communal settings are under strict regulation to serve the collective whole. Cell division and differentiation are tightly controlled for growth and regeneration; in complementary, cell death is strictly

regulated to remove the damaged and unwanted cells, thereby maintaining organismal homeostasis. These two cellular processes work in synergy throughout an organism's growth, development, and senescence.

PCD is also a critical mechanism of innate immune responses in plants and animals. To activate defense responses against harmful agents, the immune systems of an organism recognize the self and discriminate the foreign. Pattern recognition receptors (PRRs) at the cell membrane make the first line of defense by recognizing microbe- or danger-associated molecular patterns (MAMPs, DAMPs) to trigger immune signaling (Tang *et al.*, 2019). Virulent pathogens may evade the PRR-based defense and succeed in entering the cell, and host cells recognize the presence of their effectors and intracellular MAMPs. As post-invasion defense responses, higher eukaryotes have evolved convergent strategies in specific cell death modalities to minimize the inflicted damage, which plays out as inflammatory responses in animals and hypersensitive responses (HR) in plants (Coll *et al.*, 2011; Hofius *et al.*, 2011).

Initial classification of PCD relied on the morphology of the dead or dying cells in mammalian models; three types of cell death were established: apoptosis, autophagy, and necrosis (Galluzzi *et al.*, 2007; Schweichel & Merker, 1973). In drastically simplified terms, Type I PCD, or apoptosis, features the fragmentation of cell content via the formation of apoptotic bodies, which are then engulfed and digested by phagocytes. The second type of PCD is autophagy, which does not always result in cell death (Kroemer *et al.*, 2010). Instead, it involves a "self-digesting" mechanism that enables the recycling and repairs of cellular components and captures invading threats. Morphologically, it features the formation of double-membraned autophagic organelles. Necrosis is the third; it received declining recognition as a form of regulated death due to its chaotic nature and external activation (Berghe *et al.*, 2009). It is characterized by the breakage of the cell membrane and the subsequent inflammatory response.

Over the last three decades, experimental evidence rapidly expanded the topic on the genetic, biochemical, pharmacological, and functional aspects (Tang *et al.,* 2019). Apoptosis has been split into multiple subtypes, and many novel forms of non-

apoptotic PCD have been established. Morphology is still an important aspect as cell death manifests with morphological alterations. However, cell death modalities on the molecular level have become the primary focus. Signal transduction involved in the initiation, execution, and propagation of cell death explains the mechanistic difference between the newly established PCD and their pathophysiological relevance. The Nomenclature Committee on Cell Death (NCCD) formulates guidelines for defining and interpreting cell death as the field continues to expand. For details on the current classification system of cell death, readers are referred to Galluzzi *et al.* (2018) or the latest update from the NCCD.

2.5.2 Programmed cell death in plants

Animals and plants share evolutionary convergences in PCD, and their mechanisms differ according to their cellular structure, most distinctively the cell walls and the vacuoles. The defining criteria for apoptosis are cell fragmentation into discrete bodies and heterophonic removal; this is irrelevant to the botanical counterpart as plant cells are sessile and bound by rigid cell walls. In addition, incorporating cell corpses for physiological purposes is a common norm throughout the development of plants. Although the classification of plant PCD is still an ongoing debate, there is a consensus that plant vacuoles set the main stage.

The vacuoles occupy most of the volume in plant cells and are multifunctional organelles (Shimada *et al.*, 2018). They are parts of the secretory pathway and physiologically comparable to vesicles in animal cells, which compartmentalize different cellular components (Tohge *et al.*, 2011). An individual plant cell may contain several types of vacuoles that differ in function and structure. Protein storage vacuoles (PSVs) are the depositary for accumulating proteins, sugars, and organic acids. Metabolite storage vacuoles (MSVs) are the compartments for secondary metabolites, some of which are cytotoxic. Lytic vacuoles (LVs) are acidified environments impregnated with hydrolytic enzymes and defense proteins, and they set the main contrast between PCD modalities in animals and plants. In animals, cell debris is engulfed and degraded by phagocytic cells that patrol the body; in plants,

cytoplasmic removal is carried out by lytic enzymes in a cell-autonomous and stationary manner (Shimada *et al.*, 2018; Yamada *et al.*, 2019).

The NCCD's classification is mainly based on animal models and is not applicable to plants; the classification system for plant PCD is an ongoing discussion. One school of thought draws parallels between modalities of plant PCD with autophagy in animal cells (Floyd *et al.*, 2015; Hofius *et al.*, 2011). The concept is supported by the presence of autophagic organelles and the activation of autophagy-related genes (Liu *et al.,* 2005). The studies of autophagy in plants predominately focus on the mechanism of stress responses, enabling cell survival via nutrient recycling, pathogen termination, or clearance of cellular damage (D'Mello, 2007). The apparent resemblance between plant and animal autophagy is the formation of double-membraned vesicles, termed autophagosomes. In selective and macro-autophagy, cytoplasmic components are packaged by autophagosomes and delivered to LVs for degradation (Van Doorn & Papini, 2013). In micro-autophagy, invagination of the vacuolar membrane (or tonoplast) directly engulfs the subject to be removed, forming intra-vacuolar vesicles for digestion. In both cases, materials are degraded to small molecules and exported back to the cytosol for recycling (Figure 9). Mega-autophagy (syn. autolysis) describes the scenarios where the cells are damaged beyond salvage or destined to die for organismal development. Instead of forming membrane-bound vesicles, a central LV arises with the accumulation of hydrolases, which then collapses and releases the lytic enzymes into the cytoplasm (Van Doorn & Papini, 2013). The release of lytic enzymes is always followed by the proteolytic cascade that results in cellular demise. Therefore, tonoplast rupture is referred to as "the point of no return" and marks the moment of cell death.

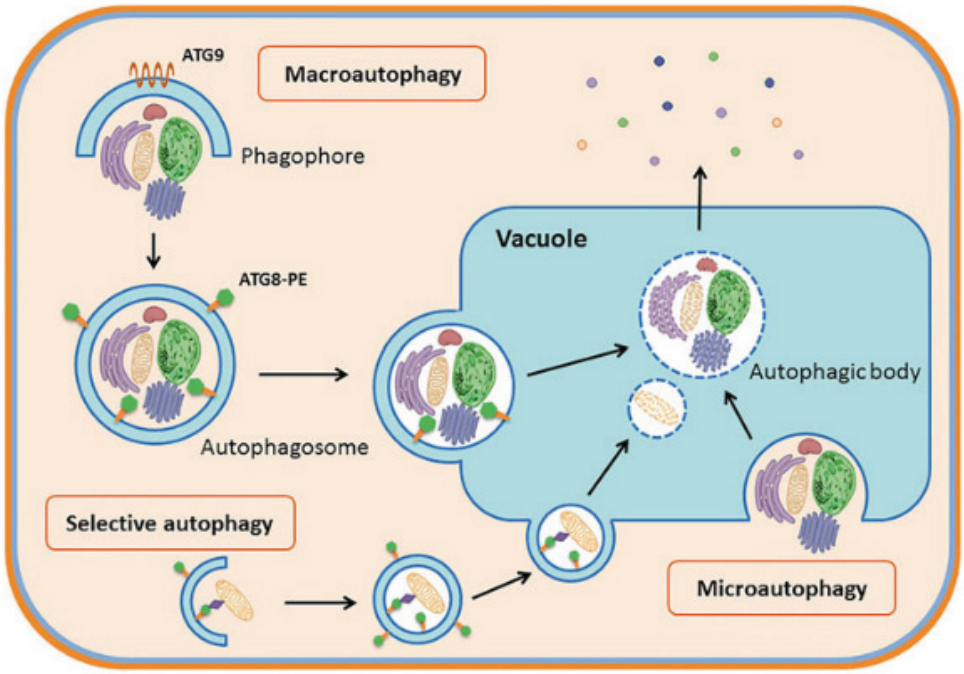


Figure 9: Schematic illustration of autophagy in plants (Floyd *et al.*, 2015).

Van Doorn *et al.* (2011) categorized botanical PCD into autolytic and non-autolytic. The former is characterized by vacuolar collapse and subsequent rapid destruction of the cytoplasm, which made an umbrella coverage for all developmental and HR-PCD. The latter does not include cytoplasm clearance and is considered equal to necrotic PCD in animals (Van Doorn *et al.,* 2011), which is increasingly recognized as accidental cell death (ACD). Another concept regarding the types of plant PCD reviews the biochemical comparables between inflammatory types of PCD in animals, such as pyroptosis and necroptosis, and HR-PCD in plants (Coll *et al.*, 2011). They lay out the comparison of the signaling cascades in the two kingdoms, which both lead to the expression of defense-related genes and ultimately reach a finale in PCD.

In addition to the well-established autolysis by tonoplast rupture, Hara-Nishimura & Hatsugai (2011) described another mode of vacuole-mediated cell death focusing on tonoplast dynamics (Figure 10). The other vacuole-mediated cell death features the fusion of vacuolar and cellular membranes, resulting in the discharge of vacuolar content directly to the extracellular space. The two modes of vacuole-mediated cell death are described along with pathogenic inducers and their subcellular locations. While tonoplast disintegration targets the internal cellular components and is effective

against viral invasion, proto-tono-plast fusion is induced by bacterial MAMPs, and defense proteins contained in the LVs are released to the extracellular space, where bacteria proliferate. Furthermore, biochemical markers were also identified. PCD via tonoplast disintegration requires caspase-1-like activity (YVAD), while proto-tono-plast fusion is marked by caspase-3-like activity (DEVD).

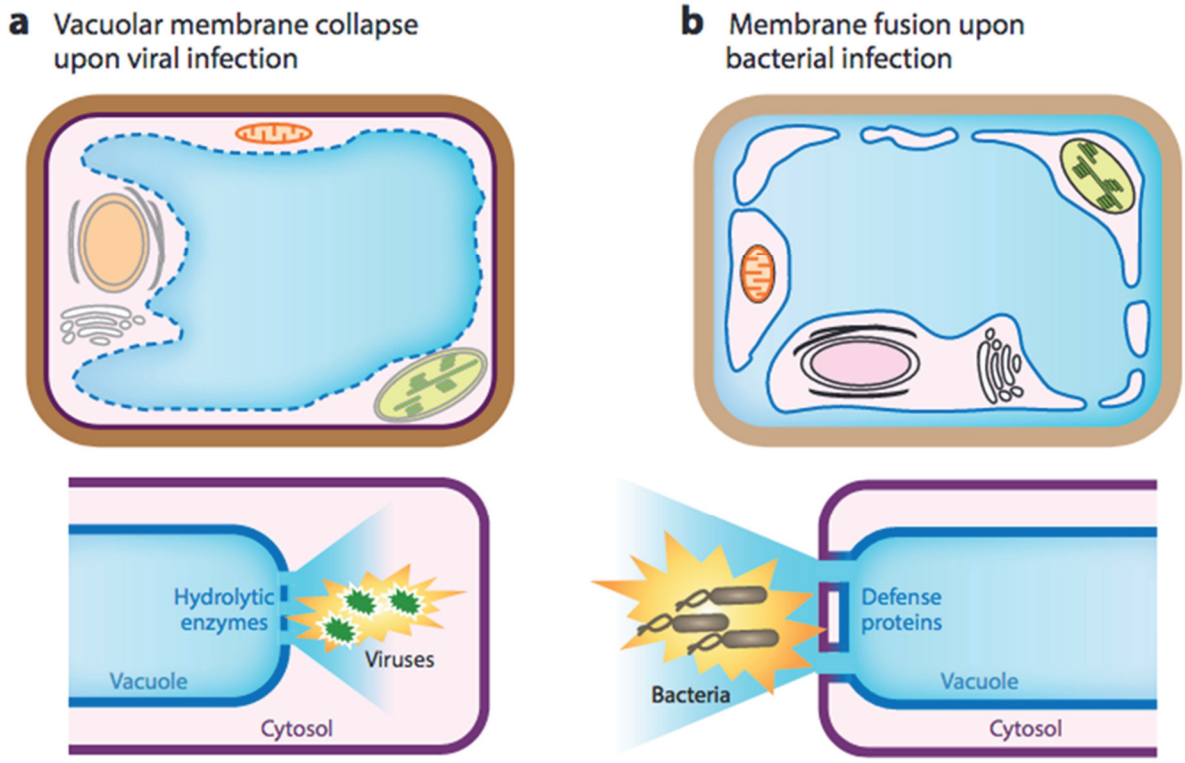


Figure 10: Schematic illustration of the two modes of vacuole-mediated hypersensitive response programmed cell death (HR-PCD) in plants induced by pathogens. (a) HR-PCD via tonoplast disintegration upon viral invasion, lytic enzymes are released to the cytoplasmic environment. (b) HR-PCD by proto-tonoplast fusion upon receiving the signal of bacterial treats, ejecting the content of the lytic vacuole to the extracellular environment (Hara-Nishimura & Hatsugai, 2011).

Concluding the different schools of concepts, there are three types of morphologically common cell deaths observed in plants, one of which is mostly considered ACD. The two modes of vacuole-mediated cell death are described as parts of plant defense mechanisms, it is a general consensus that dPCD generally follows the same pathways and morphology as that of tonoplast disintegration and exhibits caspase-1-like activities in the process.

2.5.3 Terminal differentiation of xylem cells

Noteworthy from the latest recommendations from NCCD, cellular senescence and terminal differentiation are conceptually distinct from PCD (Galluzzi *et al.,* 2018). Terminal differentiation describes the developmental process of specific tissues or cells that involve PCD toward the last phase. In animals, cornification presents an example of dPCD in which the cell corpses serve as the outermost skin barrier (Eckhart *et al.,* 2013). Cellular senescence refers to the cells permanently losing their proliferative capacity, while remaining viable and metabolically active (Zeng *et al.*, 2018). However, the same word "senescence" has a different meaning in botany; it refers to the death of body parts that are no longer necessary for the plant, and remobilizing nutrients for growth or storage elsewhere (Guiboileau *et al.,* 2010). Examples of plant senescence are leaves falling before winter or dry season, flowers shedding petals after fertilization, and forming tubers for perennation while the entire above-ground parts die off (Thomas, 2013). Adopting the recommended definitions, cell deaths observed during wood and HWF are terminal differentiation. In the different types of xylem cells investigated, autolytic PCD occurs in synergy with cell wall modification to fulfill specific physiology (Figure 11; Escamez and Tuominen 2014).

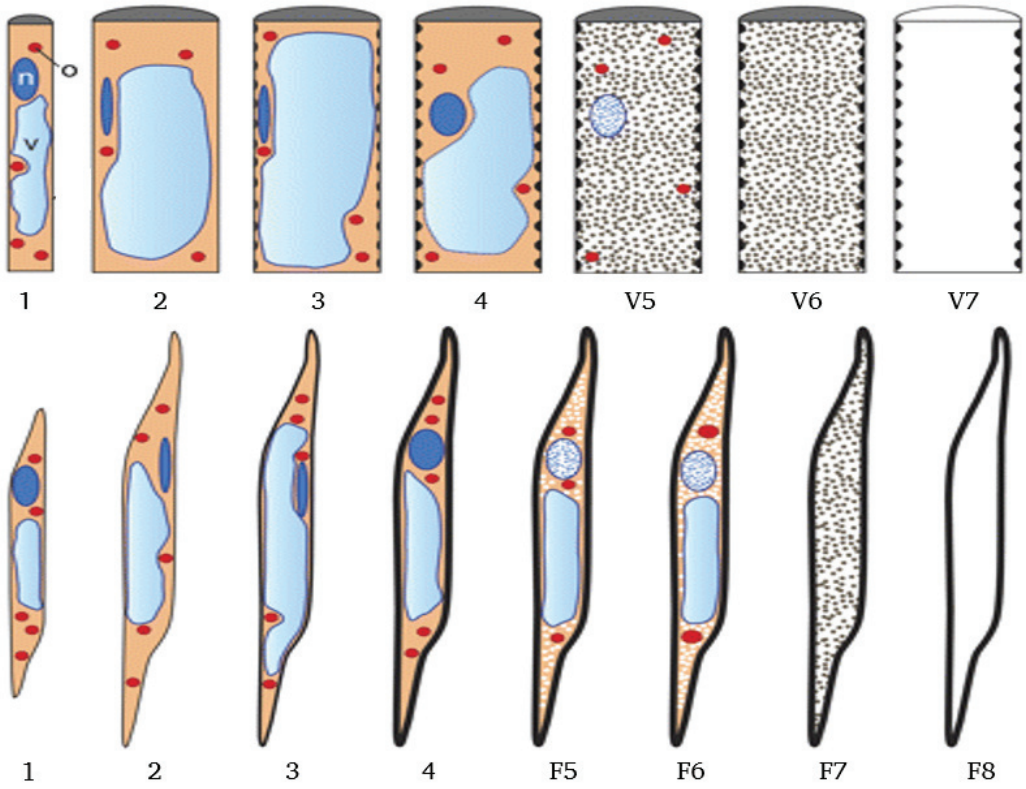


Figure 11: Schematic illustration of terminal differentiation in xylem vessel on the top and fiber at the bottom (Bollhöner *et al.*, 2012). Light blue (v) indicates the central vacuole, dark blue (n) the nucleus, and red dots (o) represent other organelles. The morphological changes at early stages in the cambial zone (1), cell expansion (2), secondary wall formation (3), and thickening (4) are similar in the two cell types. In vessels, vacuolar rupture occurs upon completion of secondary cell wall thickening, followed by rapid degradation of cytoplasmic remains (V5-V6). In fibers, degradation of the organelles starts before tonoplast rupture (F5), swelling of organelles secondary wall building continues for a longer time (F6), then the vacuole collapse and final clearing occurs (F7). In both cell types, the lignified cell walls remain after cell death and serve their physiological function.

In trees that do not form heartwood, XP cells may enter a prolonged period of cellular senescence after a certain age, and their eventual cell death does not result in physiological modification. As a natural development of organ senescence, the carbon sources stored in the dying XP cells are recycled into the tree. Contrarily, the TZ in heartwood-forming trees is the tissue for carbon sink. When XP cells undergo terminal differentiation for HWF, the large-scale SM production consumes carbon resources and their deposition modifies cell walls, resulting in a different cell type. Although the end outcomes in both cases are void of reserve materials in the inner stem core, the investment of carbon resources in heartwood-forming trees is returned as deposits of preservatives in the tree's densified internal structure.

2.6 Regulatory proteases in plant programmed cell death

PCD is executed via proteolytic cascades; in metazoa, the caspase (cysteine-dependent aspartyl specific protease) family and their hundreds of substrates constitute a significant part of the PCD degradome (Cohen, 1997). Therefore, they are often regarded as molecular markers of PCD. Caspases have a strict substrate specificity in cleaving target proteins after an aspartic acid (symbol Asp or D) residue (Cohen, 1997). They are categorized into apoptotic and inflammatory based on the type of PCD they are involved in, and the former is subdivided into initiators and executioners. The initiator caspases are activated through protein-protein interaction, whereas executioners are activated by initiators to perform the cleavage of cellular substrates necessary to carry out proteolytic cascades (Figure 12, Shi, 2004). Although no caspase has been found in the plant genome to date, studies using caspase-specific peptide inhibitors suggested that the presence of caspase-like activities also plays an integral role in plant PCD of various causes (Baskett, 2012). Hence, proteases with caspase-like structural folds are of great interest in the search for responsible enzymes that carry out such activities in plants.

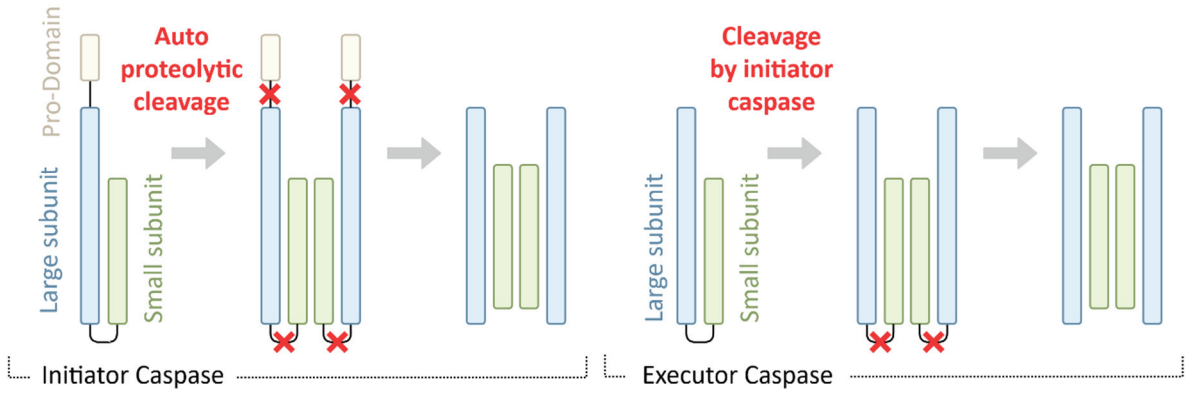


Figure 12: Activation of initiator and executor caspases, both result in the formation of heterotetramers. Cleavage sites for maturation are indicated by red crosses. The maturation of initiator caspases is carried out by auto-proteolytic cleavages, whereas the cleavages on the executor caspases at the linker region between the two subunits in conducted by initiator caspases.

Metacaspases (MCs) are the structural homologs of caspases and are only found in eukaryotes that lack caspases (Carmona-Gutierrez *et al.*, 2010). Their involvement has been observed in both developmental (Suarez *et al.*, 2004; Sundström *et al.*, 2009; Bozhkov *et al.*, 2005) and stress-induced PCD (He *et al.*, 2008; Watanabe and Lam

2011). Hence, MCs have been identified as the molecular marker for autolysis in plants (Uren *et al.*, 2000). While both caspases and MCs are mainly cytosolic, MCs differ from caspases in substrate specificity and do not exhibit typical caspase-like activities (Vercammen *et al.*, 2007; Sundström *et al.*, 2009; Yamada *et al.*, 2019).

Legumains in plants exhibit similar enzymatic properties to caspases in animals (Hara-Nishimura *et al.*, 2005; Kuroyanagi *et al.*, 2005), and its homologs are found in a diverse range of eukaryotes, including animals that have caspases. While caspases take the leading role in the cell death proteolytic cascade in the metazoan system, AEPs serve critical functions in endosomal and lysosomal degradation, remodeling of extracellular matrix, and their involvement in developmental cell death has also been shown (Dall & Brandstetter, 2016; Shirahama-Noda *et al.*, 2003). In plants, they are most commonly termed vacuolar processing enzymes (VPEs) for their localization and functions in vacuoles (Hara-Nishimura *et al.*, 1993). Homologs of VPEs are usually referred to as asparagine endopeptidases (AEPs) in animals and legumains in mammals; other terminologies applied to the enzymes are listed in Table 2. VPEs have the substrate specificity for Asp as caspases do, additionally, they also cleave after asparagine (symbol Asn or N) (Becker *et al.*, 1995; Hara-Nishimura *et al.*, 1991).

Table 2: Common terminologies of vacuolar processing enzymes (EC 3.4.22.34).

Enzyme name	Abbreviation	Species/ taxon	Explanation / localization	Reference
Legumain	LGMN	Mammals	the recommendations of the Nomenclature Committee of the IUBMB	
Asparagine/ asparaginyl endopeptidases	AEP	All taxon, more common in animals	Strict specificity for cleaving after asparagine residues	
Vacuolar processing enzymes	VPE	Plants	Localize and function in vacuoles	
Haemoglobinase		Schistosoma mansoni	Digestion of hemoglobin the midgut	(El Meanawy *et al.*, 1990)
Endoprotease B	EP-B	*Hordeum vulgare*	Localize in germinating seeds	(Marttila *et al.*, 1993)
Nucellain		*Hordeum vulgare*	Localize in nucellar cell walls	(Linnestad *et al.*, 1998)
Cysteine protease PRSC1	PRSC1	Humans		(Tanaka *et al.*, 1996)
Asparaginyl carboxypeptidase	ACP		Mono-carboxy-peptidase	(Dall & Brandstetter, 2013)
Butelase		Clitoria ternatea		(Nguyen *et al.*, 2014)
Vicilin peptidohydrolase		Legume	Cleave vicilin in cotyledons	

2.6.1 Metacaspases

Metacaspases (MCs) are a family of ancestral relatives of caspases, and they share high homology in structures (Figure 13; Tsiatsiani *et al.*, 2011). Most MCs have the same conserved His-Cys catalytic dyad as in caspases, with a few exceptions that have a His-Ser catalytic dyad (Szallies *et al.*, 2002). Their main difference from caspases is that they cleave specifically after arginine (symbol Arg or R) or lysine (symbol Lys or K) residues at the P1 position of their substrates (Tsiatsiani *et al.*, 2011; Vercammen *et al.*, 2004). Due to the difference in cleaving targets with caspases, MCs are not directly responsible for the caspase-like activities in plants (Minina *et al.*, 2013). Instead, they act upstream of the Asp-specific proteases, which indicates their upstream position to the activation of autophagy (Minina *et al.*, 2013; Tsiatsiani *et al.*, 2011).

Three types of MCs have been identified as the structural homologs of metazoan caspases, and they are found exclusively in eukaryotes with no caspases in the genome (Carmona-Gutierrez *et al.*, 2010). The domain architecture of MCs includes a catalytic

subunit named p20, a non-catalytic subunit p10, and a linker region in between (Uren *et al.*, 2000; L. Tsiatsiani *et al.*, 2011). Type-I MCs are found in diverse organisms, from single-cell protists to higher plants and fungi. Their precursors have a pro-domain at the N-terminal, resembling that of initiator pro-caspases (Uren *et al.*, 2000). Type-II MCs are similar to executioner pro-caspases in lacking this pro-domain in the precursors, and they have a more extended linker region in comparison. Type-II MCs are present in green plants exclusively; Type-III MC has a reverse order of the two subunits and has only been found in planktonic (Uren *et al.*, 2000; Vercammen *et al.*, 2007). The number of MC genes in different plant species varies considerably; angiosperms commonly have multiple homologs of each MC type (Tsiatsiani *et al.*, 2011).

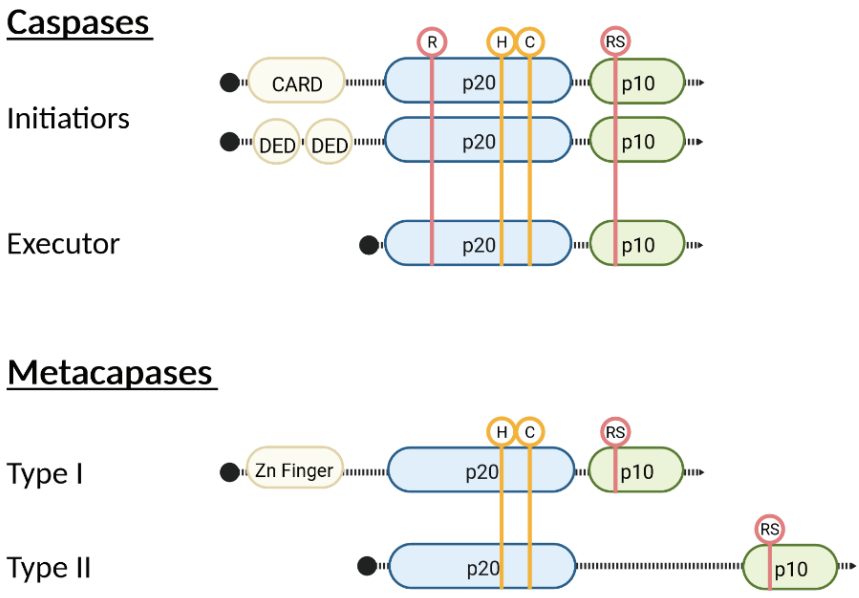


Figure 13: Illustration of the domain structures of caspases and metacaspases (MCs). Beige elements present the N-terminal pro-domains (NTP), blue elements indicate the large subunits p20, and green elements show the small subunits p10. There are two types of pro-domains in initiator caspases: the single death fold known as the caspase recruitment domain (CARD) and two death folds known as death effector domains (DED). Type I MCs contain a zinc-finger motif in the pro-domain similar to that of initiator caspases. Type-II MCs do not have a prodomain; instead, they have a longer linker region between the two domain subunits. The conserved His-Cys catalytic dyad is marked in yellow. Three of the amino acids that form the caspase-1 substrate-binding pocket (Arg179, Arg341, and Ser347 in H. sapiens) are marked in red, two of which are also present in MCs.

With little exception, MCs rely on calcium ions (Ca^{2+}) for activation (Watanabe and Lam, 2011), which suggests their involvement in the initial signaling cascades. MCs are synthesized in the ER as inactive pro-enzymes, and their activity depends on

specific cellular conditions as parts of post-translational regulatory mechanisms (Tsiatsiani *et al.*, 2013). Unlike the caspase family, MCs do not undergo dimerization to form an active protease. Instead, evidence suggests that a cleavage within the linker region carries out their maturation, and this cleavage requires the presence of Ca^{2+}(Klemenčič & Funk, 2018; Watanabe & Lam, 2011). In addition, each MC subunit contains a region rich in negatively charged amino acids, which fold into loops and form a second calcium-binding site when they come into proximity (Fortin and Lam, 2018; Klemenčič and Funk, 2018), as shown in Figure 14. Signal transduction pathways that rely on Ca^{2+} as an intracellular second messenger, so-called calcium signaling, are broadly involved in plant developmental processes and stress responses (Batistič and Kudla, 2012). In the event of coming across physiological stimuli such as pathogenic elicitors, plant hormones, and abiotic stresses, plants cells respond with the generation of Ca^{2+} signatures that differ in peak duration, intensity, and repetitions (Tuteja & Mahajan, 2007; De Vriese *et al.*, 2018). These Ca^{2+} signatures serve as messengers for inter-cellular communication, and the Ca^{2+}-dependent activation of MCs suggested that they are responsive to physiological stimuli via calcium signaling.

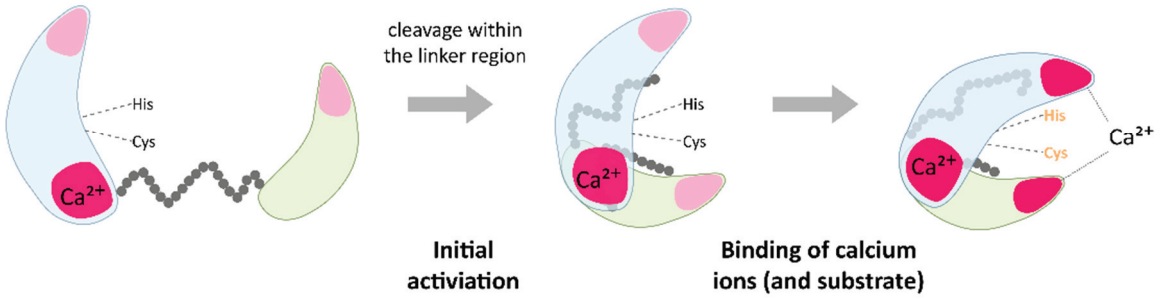


Figure 14: Schematic illustration of the calcium ions (Ca^{2+}) dependent activation in Type II Metacaspases. Blue elements indicate the large subunits p20, and the green elements show the small subunits p10.

The pH optima of MCs is another post-translation regulatory mechanism, and the combined activation prerequisite with Ca^{2+}-dependency suggests their mode of action in vacuole-mediated PCD. The cytoplasm of living cells typically has a neutral pH,

whereas LVs and the cell walls are acidic. Repeatedly observed in the timeline of autolytic cell death, an influx of Ca^{2+} arrives at the pH-neutral cytoplasmic environment immediately preceding the vacuolar collapse. After tonoplast rupture, the acidic vacuolar content is released, rendering a pH drop in the cytoplasm. *In vitro* activity assays of MCs in diverse species indicate that their pH optima are neutral to slightly basic (7.0–8.5) (Vercammen *et al.*, 2004; Bozhkov *et al.*, 2005). AtMC9 and homologs are the only exceptions in the MC family to activate in acidic conditions without Ca^{2+}-dependency (Bozhkov *et al.*, 2005). As a plant immune response, the Ca^{2+}-dependent type-II MCs mediate the processing of PROPEP, a class of immune peptides can be induced by stress-associated stimuli, such as pathogens, MAMPs, wounding, and defense hormones (Bartels & Boller, 2015; Huffaker *et al.*, 2006; Klauser *et al.*, 2015). On the other hand, *MC9* is expressed in cells undergoing developmental cell death and serves its role in the acidic cytoplasm after the vacuolar collapse (Bollhöner, 2013; Olvera-Carrillo *et al.*, 2015; Tsiatsiani *et al.*, 2013). During xylogenesis, activities of AtMC9 are required for the rapid and efficient autolysis of xylem vessel elements (Bollhöner *et al.*, 2013).

2.6.2 Vacuolar processing enzymes

The emergence and function of VPE isoforms

Analyses of plant genomes indicate that *VPEs* are widely distributed in land plants and hold significance in the emergence of seed plants (Hatsugai *et al.*, 2015). VPE was first identified in castor beans as the single responsible enzyme for seed protein maturation (Hara-Nishimura *et al.*, 1991). The enzymes have been isolated in a wide range of animals and plant species and serve a diverse suite of functions. Phylogenetic studies revealed three gene duplication events of the *VPE* along the diversification of green plants (Figure 15; Poncet *et al.*, 2015). Gymnosperms have only one *VPE* type, and two distinct isoforms are present in the most recent common ancestor (MRCA) of seed plants, *Amborella trichopoda*. In *Amborella*, monocots, and basal eudicots, there are β-type and a pro-ortholog that later give rise to α-γ-δ-type VPE. βVPE generally functions in the PSV and converts storage protein precursors into their respective

mature forms (Hara-Nishimura *et al.*, 1998). To the present knowledge, most seed proteins of various plants are produced by proteolytic cleavages by βVPE (Hara-Nishimura *et al.*, (1995). The separation of *βVPE* from other *VPE*s during angiosperm evolution coincided with the development of the double fertilization system, which involved the development of endosperm (Friedman, 1994). During the diversification of core eudicots, a vegetative γ-type *VPE*, and a seed coat specific δ-type *VPE* emerged. Expression of *δVPE* is confined in the integument cell layers, and the executor enzyme in the specific PCD leads to the formation of a hard seed coat (Rocha *et al.*, 2013; Nakaune *et al.*, 2005).

αVPE arose from the split with γ*VPE* and is only present in Brassicaceae (Poncet *et al.*, 2015; Yamada *et al.*, 2019). α- and γ*VPEs* are considered the vegetative types because they are commonly found in vegetative tissues, and they predominantly function in the LVs(Kinoshita *et al.*, 1999). Despite the name, they do participate in embryo development and reproductive growth. Conversely, while the seed-type *βVPEs* is primarily present in seeds, they have also been found in the PSV throughout plant bodies, including both reproductive and vegetative tissues (Shimada *et al.*, 2003). The model organism *A. thaliana* has one of each of four VPE types; gene knock-out studies showed that while *βVPE* plays the major part in processing seed storage proteins (Gruis *et al.*, 2002; Shimada *et al.*, 2003), γ*VPE* is the essential isoform in both developmental and stress-induced PCD (Kuroyanagi *et al.*, 2005).

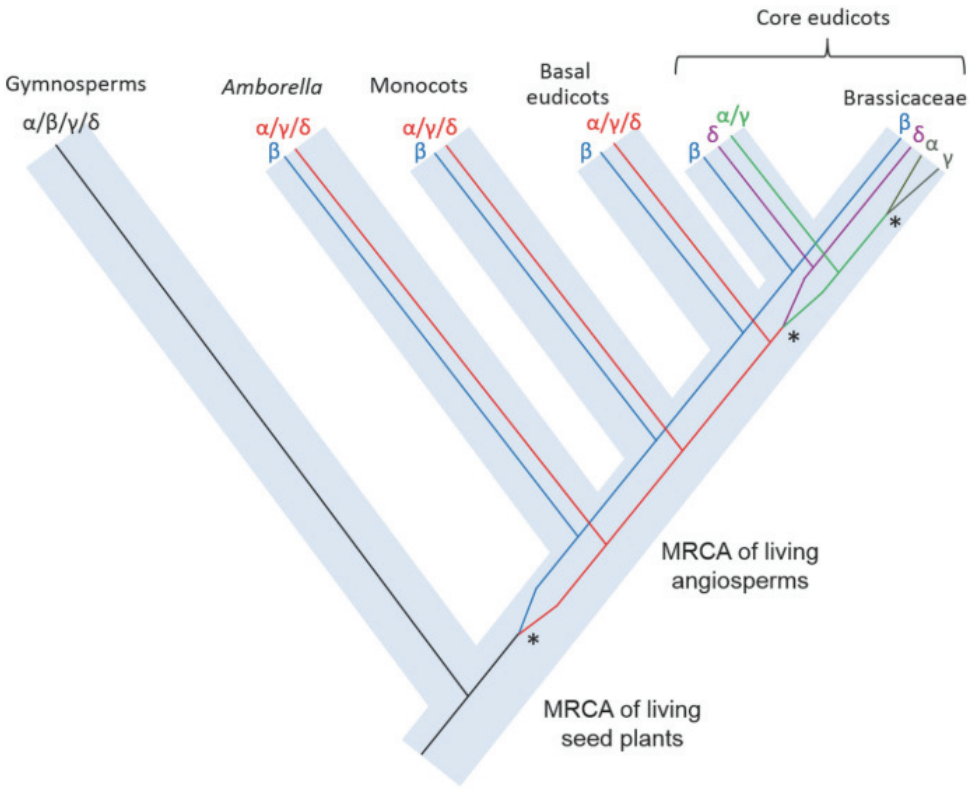


Figure 15: A schematic reconstruction of the evolution of the VPE family in plants, showing the origins of the duplication events that occurred at various evolutionary stages (Poncet *et al.*, 2015).

Mode of function in plant PCD

The vegetative type VPEs mediate PCD through their primary role in the LVs (Hara-Nishimura *et al.*, 1991; Kinoshita *et al.*, 1999). The LVs contain proteases that degrade cellular components; proteases that commonly participate in PCD include papain, aleurain, carboxypeptidase Y, and serine proteases. In addition, the LV is also a storage for proteins produced under the induction of stress, such as chitinases (Ponstein *et al.*, 1994), lectins (Gatehouse *et al.*, 1987; T. A. Wilkins & Raikhel, 1989), and cyclotides (Craik, 2012). These PCD-associated proteases and stress-induced proteins are synthesized as inactive zymogens, often with a C-terminal inhibitory extension to be cleaved at the Asn or Asp residues during maturation (Kinoshita, Nishimura, and Hara-Nishimura 1995; Yamada *et al.*, 2005). The self-activating VPE is a priming initiator in the protein processing system (Hiraiwa *et al.*, 1997; Kuroyanagi *et al.*, 2002). Once activated in the vacuolar compartments, the mature VPE recognizes the Asn or Asp residues on other precursor proteins and cleaves their peptide bonds at the carbonyl sides (Hara-Nishimura *et al.*, 1998; Neuhaus & Rogers, 1998). The extensive variety of

proteins that are post-translationally processed by VPEs is reviewed in Yamada *et al.*, (2019). This action results in the maturation of the PCD-associated proteases and stress-induced proteins in the LV, thereby initiating the proteolytic cascade leading to tonoplast rupture and the degradation of cytoplasmic structures (Fukuda 1996; Hatsugai *et al.*, 2015).

The essential role of *VPE* in vacuolar collapse was demonstrated with virus-infected leaves in tobacco (*N. benthamiana and N. tabacum*) (Hatsugai *et al.*, 2004). The *VPE*-silenced plants showed a reduction of caspase-1 activity; it did not carry out tonoplast disintegration and failed to conduced pathogen-induced PCD observed in the wild type (Hatsugai *et al.*, 2004). Furthermore, elevated gene expression of vegetative *VPE* during dPCD has also been reported in vascular cells. During the active deposition of SCW in *A. thaliana* TEs, elevated expression of *α*- and γ*VPE* occurs along with cellulose synthase genes (Kinoshita *et al.*, 1999; Turner *et al.*, 2007). The same pattern has been observed for fiber with both *β*- and γ*VPE* in poplar (Courtois-Moreau *et al.*, 2009). As previously mentioned in Chapter 2.4.3 Heartwood substances deposition, CEPs have been observed to involve SCW thickening and stem maturation during the differentiation of TEs and fibers (Han *et al.,* 2019). They have the ability to digest extensins, which form the basic scaffold for cell wall formation (Helm *et al.,* 2008). In xylem fiber cells of *A. thaliana,* CEP1 zymogens are activated by γVPE (Cheng *et al.,* 2019). In their experiment with a γ*vpe*-suppressed mutant of *A. thaliana*, the differentially expressed genes (DEGs) categorically belong to both PCD and the biosynthetic pathways for cell wall components (Cheng *et al.*, 2019). The γ*vpe* mutant had vigorous SCW biosynthesis at the transcription level, an accumulation of unprocessed proenzyme CEP1, prolonged SCW thickening, delayed PCD, and incomplete degradation of the cellular contents.

Domain structure and activating mechanism of VPEs

VPEs are synthesized as inactive zymogens at the ER, and the propeptides encode for a self-activating mechanism. The VPE pro-enzymes have a signal peptide, a short

propeptide at the N-terminus (N-terminal propeptide, NTP), and an inhibitory prodomain at the C-terminal (C-terminal propeptide, CTP) that masks the catalytic sites (Dall & Brandstetter, 2013; Kuroyanagi *et al.*, 2002). After its synthesis at the ER, the self-catalytic removal of the signal peptide releases the inactive pro-enzymes into the secretory pathway (Fischer *et al.*, 2000); they are then trafficked to the vacuoles (Kuroyanagi *et al.*, 2002) or the cell wall (Linnestad *et al.*, 1998). Arriving in these acidic environments triggers the self-catalytic removal of the inhibitory CTP, resulting in mature enzymes (Kuroyanagi *et al.*, 2002).

VPEs and their metazoan homologs are predominantly localized in an acidic subcellular environment that has a reducing redox potential, such as the plant vacuoles and nucellar cell walls (Linnestad *et al.*, 1998; Shimada *et al.*, 2018) or within the endo-lysosomal system in animal cells (Dall & Brandstetter, 2013). However, they have also been found translocating to the cytosol, into the nucleus, and the extracellular space that have incompatible pH and redox potential (Dall & Brandstetter, 2013). In the mammalian model, legumains interact with a protein named integrins at pH-neutral environments that stabilizes their tertiary structures in different subcellular localizations (Dall and Brandstetter, 2016; 2013). In the botanical model, VPE relies on the Legumain Stabilization and Activity Modulation (LSAM) domain within the CTP (Figure 16). After auto-catalytic cleavage of the inhibitory CTP, the LSAM domain remain non-covalently bound to the active domain, which confers stability at neutral pH and allows the enzymes to survive in different cell compartments (Zauner *et al.*, 2018).

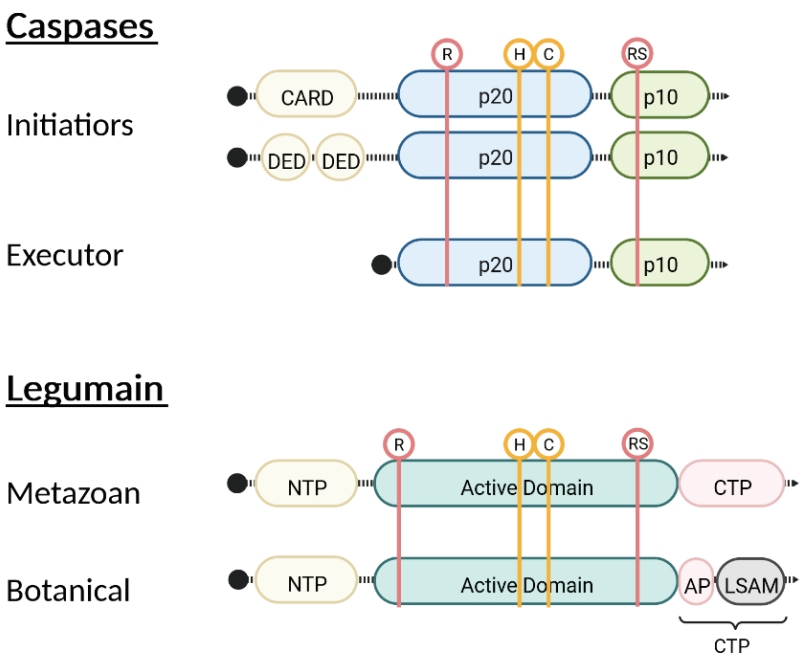


Figure 16: Illustration of the domain structures of caspases and legumains. Domain structures of caspases: beige elements present the N-terminal pro-domains (NTP), blue elements as the large subunits p20, and the green elements show the small subunits p10. There are two types of pro-domains in initiator caspases: the single death fold known as the caspase recruitment domain (CARD) and two death folds known as death effector domains (DED). Legumains contain a single active domain and sequences are comparable to that of the two subunits in caspases combined. The conserved His-Cys catalytic dyad is marked in yellow. Three of the amino acids that form the caspase-1 substrate-binding pocket (Arg179, Arg341, and Ser347 in *H. sapiens*) are also present in all identified VPEs and marked in red. Legumain in plants contains an activation peptide (AP) and the Legumain Stabilization and Activity Modulation (LSAM) domain within the C-terminal pro-domain (CTP).

Beyond the proteolytic activities

In addition to their proteolytic activities, a list of *VPE* isoforms has been identified with peptide bond ligase ability in animals and plants (Craik, 2012; Du *et al.*, 2020; Mylne *et al.*, 2011; Nguyen *et al.*, 2014). As the LSAM domain is bound to the mature VPE, it not only confers stability to the enzyme at neutral pH but also modulates the activity of the catalytic domain in favor of ligase activity (Zauner *et al.*, 2018). Among the list of identified ligases, butelase 1 from *Clitoria ternatea* and HaAEP1 from *Helianthus annuus* are two *VPE* isoforms reported with particularly high efficiency as peptide cyclases (Nguyen *et al.*, 2014; Mylne *et al.*, 2011; James, Haywood, and Mylne 2018).

Cyclic peptides (cyclotides) are miniature cyclic proteins of around 30 amino acids (Conlan & Anderson, 2012). A variety of toxic cyclotides serve defense functions in the Eudicot families plants (Dutton *et al.*, 2004; Mylne *et al.*, 2012; Poth *et al.*, 2011), such as insecticidal, nematocidal, molluscicidal, bactericidal, and antiviral (Craik, 2012; Shafee *et al.*, 2015). Their production requires cleavage between domains and then forming a new peptide bond between the carboxyl and amino terminals within the protein precursor (Conlan & Anderson, 2012). In several plant species, VPEs are responsible for both actions required for cyclotide production (Haywood *et al.*, 2018; M. A. Jackson *et al.*, 2018).

In addition to their enzymatic functions, VPE can also affect cell fate and plant development by manipulating ethylene production at the transcription level (Matarasso *et al.*, 2005). In higher plants, the first committed step in ethylene biosynthesis is carried out by the regulatory enzyme 1-Aminocyclopropane- 1-carboxylic acid (ACC) synthases. The ACC synthase (*Acs*) genes are encoded by a multigene family, and each member is differentially regulated by various environmental and developmental factors. In tomatoes (*Solanum lycopersicum*), a γVPE LeCp has been identified to control the gene expression of an *Acs* gene (Matarasso *et al.*, 2005). The ChIP analysis demonstrated the *in vivo* interaction between the LeCp protein and the Acs gene sequence inside the nuclei. In addition, they identified the AT-rich element in the *Acs2* promoter as the binding site.

The plant hormone ethylene is a prominent influencer of plant growth and development (Abeles *et al.*, 1992; M. B. Jackson, 1991; Mattoo & Handa, 2008). During the different phases of TE formation observed in Zinnia cell cultures, ethylene is produced with characteristic accumulation peaks during the initiation, SCW formation, and lignification (Pesquet & Tuominen, 2011). Ethylene is also an inducer for defense responses (Boller, 1991; Lund *et al.*, 1998); its production is closely associated with the formation of phenolic compounds in herbaceous and woody plants (Chappell *et al.*, 1984; Imaseki, 1991). Ethylene as an essential factor in HWF has been demonstrated in decades of reports, such as that in black walnut (*Juglans nigra* L.) and black cherry (*Prunus serotina* Ehrh.; Nelson *et al.*, 1981), pine (Nilsson *et al.*, 2002),

and Chinese fir (Song *et al.*, 2014). In addition, the up-regulation of an ethylene receptor homologous gene was among the DEGs during HWF in Robinia (Yang *et al.*, 2004).

2.7 Experimental design

2.7.1 Model organisms

Robinia pseudoacacia offers multiple advantages to the studies of HWF, thus selected for the current experiment as a representative model for ring-porous trees that form the classic obligate heartwood. Heartwood forms in mature xylem that are at the minimum several years of age. In many species, HWF only start in stems older than a decade. *R. pseudoacacia* is a fast-growing legume species that starts forming heartwood early, at around 4 to 6 years old (Nakaba *et al.*, 2012). It is ecologically diverse yet monophyletic, and has been widely cultivated in temperate regions for its highly durable timber and ornamental strands (Cierjacks *et al.*, 2013). The sharp boundary between the sap- and heartwood is conspicuous, with HWS that fluoresces in yellowish green under UV light (Figure 17), allowing easy identification of the region. The characteristic HWS in *Robinia* are dihydrorobinetin and robinetin, which are flavonoids generated through well-studied biochemical pathways (Magel *et al.*, 1994). These molecules share parts of biochemical pathways with lignin production, hence offering a common ground for comparing regulations on the main component in xylogenesis and HWF.

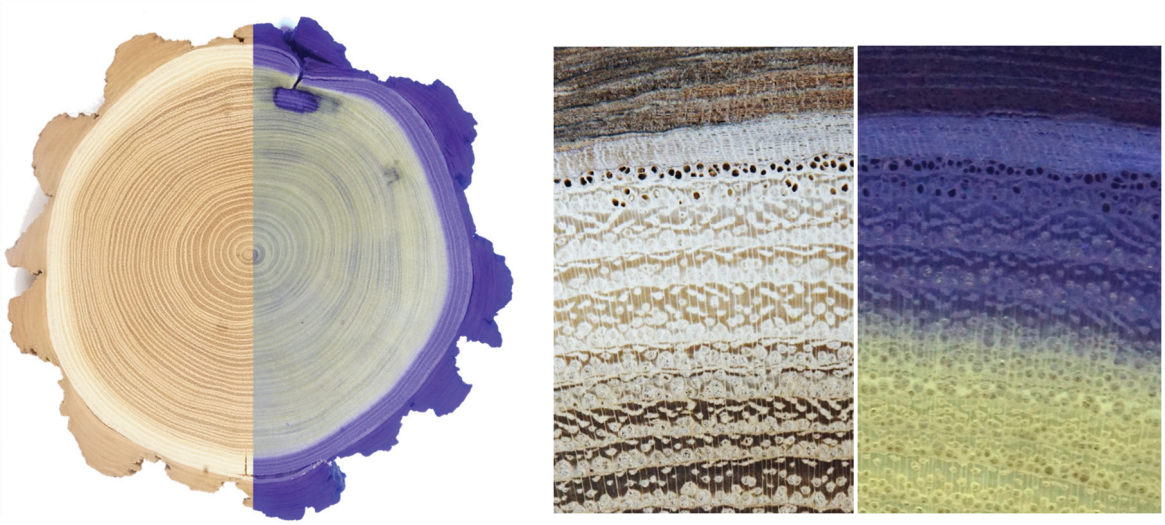


Figure 17: Photos of a cross-sectioned stem of *Robinia pseudoacacia* under normal light (left side) and UV light (right side). (Left) Photo of the full disk shows the large volume occupancy of heartwood region in the stem. (Right) Photo of the same disk focusing on the region between inner bark to outer heartwood.

2.7.2 Sampling scheme

Studies on HWF require specimens of age and from species with far longer life cycles than common model organisms for biological experiments. Furthermore, samples in the depth of the stem core are exceptionally rigid, which presents obstacles in their processing. For example, phenolic compounds present at the TZ challenges the standard extraction methods in molecular biology (De Filippis & Magel, 1998), traditional chemical fixation methods for transmission electron microscopy (TEM) that perform well for cambium cells are not applicable because the inner wood region is impermeable due to tylosis, and common cryofixation methods result in ice crystals that destroy tonoplast integrity (Takabe & Kim, 2016).

Due to sample availability and the practical cumbersome, the vast majority of research on HWF was conducted based on a few specimens. Older literature generally includes many samples, either with a diversity of species (Frey-Wyssling & Bosshard, 1959; Hillis, 1987) or across the seasons (Nobuchi *et al.*, 1982). Although technological advancement in recent years offers new opportunities for further investigation on the topic at the cellular and molecular level, the norm for HWF studies in recent years is to harvest samples from one or two trees in spring and another one to two in autumn. HWF is a complex developmental process that takes place over several months at the

minimum, involving multiple major cellular processes across all living regions in the stems of a tree. The current study aims to provide an overview on the gene expression patterns of the PCD marker genes over an annual cycle; samples from mature and healthy trees were collected monthly for a year, and all wood regions with living cells were examined.

2.7.3 Gene expression studies

Regulation of gene expression is the basis for cellular function, differentiation, and development of any organism. Naturally, gene expression is tightly regulated via multiple mechanisms and can be modulated in many steps, some of which are the epigenetic binding of DNA, transcriptional initiation, RNA processing, and post-translational modification of a protein. For the current study, the selected method for analysis is real-time polymerase chain reaction (real-time PCR, or qPCR). Real-time PCR detects gene expression at the post-transcription level and before the functional gene products as proteins are made.

The expression level of any particular gene is implied by the turnover of its transient transcripts in the form of single-stranded mRNA, and qPCR is a method that aims to measure the copy number of gene transcripts in a sample. The preparation procedure for qPCR involves immediate fixation of the specimen, total RNA extraction, and transcribing the total mRNA present in the cell into a more stable version of nucleic acid sequences, termed complementary DNA (cDNA). During a qPCR run with gene-specific primers, the amount of cDNA is recorded at every replication cycle, which serves as the basis for calculating the quantity of gene transcripts. Taking the relative gene expression approach, we measure the candidate gene relative to the stably expressed reference gene across different time points and wood regions. The relative gene expression of PCD regulatory enzymes indicates the timeframe in which the cells in the corresponding region have received the signal to prepare for PCD, which may offer clues on identifying the effectors of HWF.

3. MATERIALS AND METHODS

3.1 Plant materials

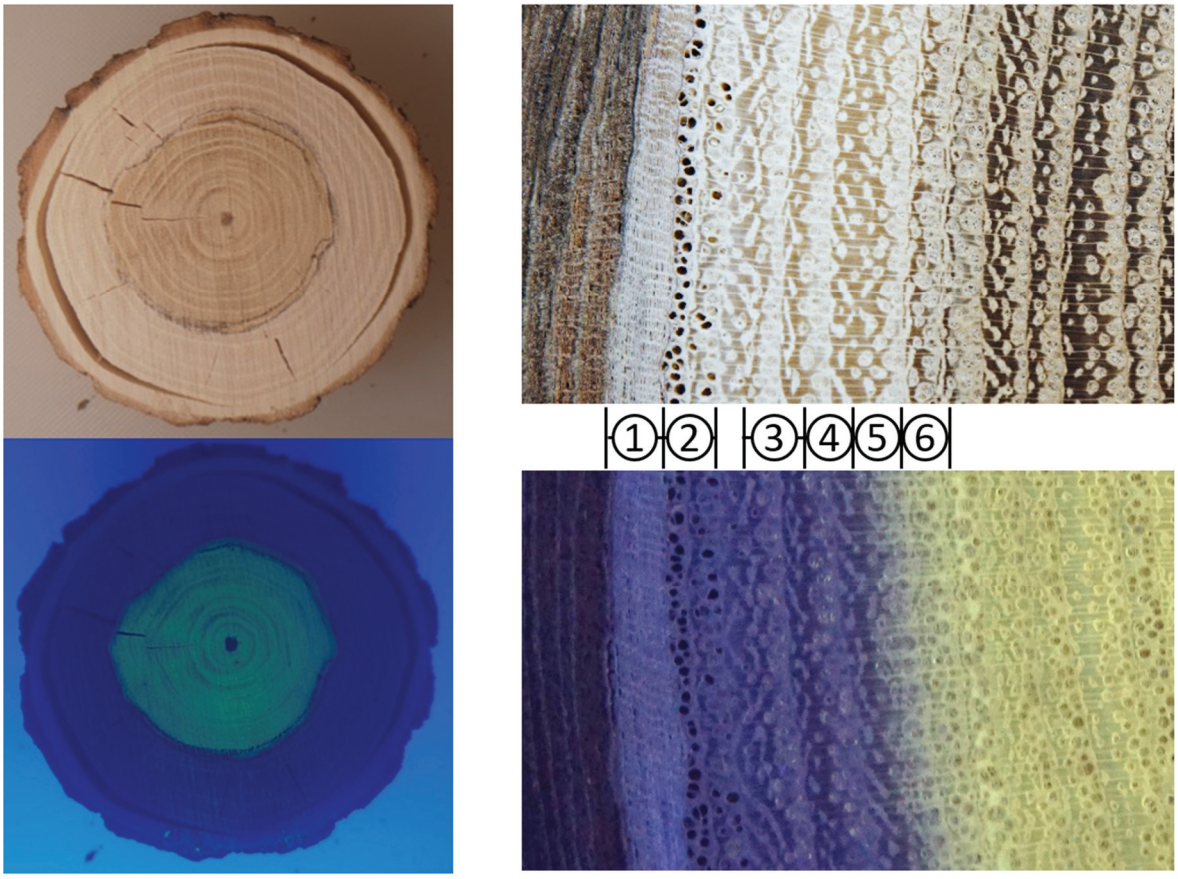
3.1.1 Species identification and sample harvesting

Six black locust trees grown in the arboretum of the Thünen Institute in Hamburg-Bergedorf, Germany, were identified for a full-year sampling scheme. They were between 15 and 20 years old, 15 to 20 meters in height, and their stem diameters at breast height were around 30 cm. Monthly between April 2019 to March 2020, a healthy branch between 10 to 20 cm in diameter was cut and flash-frozen by dry ice immediately, then transferred to an ultra-low freezer for storage until further processing.

Prior to the sampling regime, species identity was confirmed to be *Robinia pseudoacacia* L. by comparing DNA sequences at the internal transcribed spacer (ITS) region of the 18S ribosomal RNA gene to the reference sequence from GeneBank (Accession number EF494737.1).

3.1.2 Specimens processing and experimental design

After the specimen were solidly frozen, they were sawn into disks of 2 cm thickness, lyophilized, and sectioned into six regions as shown in Figure 18. There are 63 sectioned samples in total, and they were labeled with the codes by year (2 digits) – month (2 digits) – wood region (1 digit), as listed in Table 3.



Region Code	Wood Region	Sample Description
1	Phloem	Between bark and vascular cambium
2	Outermost sapwood	Outermost growth ring(s) with open vessel elements, differentiating xylem
3	Middle sapwood	Sapwood with tylosis
4	Innermost sapwood	Innermost sapwood region right next to the fluorescent area under UV light
5	Transition zone	Between area of observable fluorescence to solid fluorescence under UV light (between the pencil marks)
6	Outer heartwood	Outermost 2mm region with solid fluorescence

Figure 18 Illustration on wood regions for sample sectioning. (Left) Photos of a freeze-dried specimen disk under normal light on the top and UV light at the bottom. (Right) Indication of how the different regions were sectioned for the present study.

Table 3: Coding of samples and their availability for the present study, a total number of 63 samples.

	Sample code by wood region						
		1	2	3	4	5	6
Sample code by harvest date	2019 Apr	1904-1	1904-2	1904-3	1904-4	--	1904-6
	2019 May	1905-1	1905-2	1905-3	1905-4	--	1905-6
	2019 Jun	1906-1	1906-2	1906-3	1906-4	1906-5	--
	2019 Jul	1907-1	1907-2	1907-3	1907-4	1907-5	1907-6
	2019 Aug	1908-1	1908-2	1908-3	1908-4	1908-5	--
	2019 Sep	1909-1	1909-2	1909-3	1909-4	1909-5	--
	2019 Oct	1910-1	1910-2	1910-3	1910-4	1910-5	1910-6
	2019 Nov	1911-1	1911-2	1911-3	1911-4	1911-5	--
	2019 Dec	1912-1	1912-2	1912-3	1912-4	--	1912-6
	2020 Jan	--	--	--	--	--	--
	2020 Feb	2002-1	2002-2	2002-3	2002-4	--	2002-6
	2020 Mar	2003-1	2003-2	2003-3	2003-4	--	2003-6

* -- indicate the absence of sample

3.2 Nuclei acid isolation and reverse transcription

With each sample, 10 g of sectioned wood chips were collected and milled into flour with Retsch Mill while kept cold with liquid nitrogen. Genomic DNA was isolated with Qiagen DNeasy® Plant Kit and total RNA with RNeasy Plus Universal Kits (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Samples closer to heartwood have little RNA due to the lesser amount of living cells and higher amount of heartwood specific phenolics. Hence the following steps were inserted additionally into the extraction protocol: RNA was precipitated with isopropanol and centrifuged at 12000 x g for 10 min at 4°C to obtain pellets, washed with 70% ice-cold ethanol, resuspended in RNase-free water, and re-precipitated with 2 vol. of 100% ethanol at and 0.1 vol. of 3M Sodium Acetate (pH 8) at -20°C overnight before pooling onto RNeasy mini column for washing steps and final elution. The optical density (OD) of total RNA was determined on a Nanodrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). Removal of genomic DNA and cDNA synthesis were carried out using QuantiNova Reverse Transcription Kit (Qiagen, Hilden, Germany) with 200 ng of total RNA as templates.

3.3 Obtaining sequences of candidate genes

3.3.1 Primer design

To identify the sequences of candidate genes in *R. pseudoacacia*, DNA sequences of a homologous gene from members of Fabales were retrieved from the NCBI and UniProt database (Table 4 to 6). Sequence alignments were made in the software Geneious Prime to design degenerate primers for polymerase chain reaction (PCR; ROZEN *et al.*, 2000), and parameters were set to respect the guideline according to Pennington *et al.* (2009). Properties and compatibility of the oligonucleotides were validated with Oligo Analysis Tool available at Eurofins' website[1]. Primers validated by digital means were then purchased from biomers.net GmbH for the *laboratory* procedures.

Table 4: Homolog sequences of type-II metacaspases used for primer design.

Accession	Description
Homologs of AtMC4 in Fabales	
XM_004509944	PREDICTED: Cicer arietinum metacaspase-4-like (LOC101509880), mRNA
XM_027499192	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X1, mRNA
XM_028348014	PREDICTED: Glycine soja metacaspase-4-like (LOC114387810), mRNA
XM_028389058	PREDICTED: Glycine soja metacaspase-4-like (LOC114422611), mRNA
XM_003546612	PREDICTED: Glycine max metacaspase-4 (LOC100802658), mRNA
Homologs of AtMC9 in Fabales	
XM_027495318.1	PREDICTED: Abrus precatorius metacaspase-9 (LOC113862209), mRNA
XM_016115254.2	PREDICTED: Arachis duranensis metacaspase-9 (LOC107494198), mRNA
XM_025749698.2	PREDICTED: Arachis hypogaea metacaspase-9 (LOC112696814), mRNA
XM_016306384.2	PREDICTED: Arachis ipaensis metacaspase-9 (LOC107604728), mRNA
XM_020364252.2	PREDICTED: Cajanus cajan metacaspase-9 (LOC109802849), mRNA
XM_004493225.3	PREDICTED: Cicer arietinum metacaspase-9 (LOC101488845), mRNA
XM_003554030.5	PREDICTED: Glycine max metacaspase-9 (LOC100781510), mRNA
XM_019601616.1	PREDICTED: Lupinus angustifolius metacaspase-9 (LOC109357640), mRNA
XM_013594075.3	PREDICTED: Medicago truncatula metacaspase-9 (LOC25498997), mRNA
XM_028910894.1	PREDICTED: Prosopis alba metacaspase-9-like (LOC114724535), mRNA
XM_028916909.1	PREDICTED: Prosopis alba metacaspase-9-like (LOC114729879), mRNA
XM_017562410.1	PREDICTED: Vigna angularis metacaspase-9 (LOC108328531), mRNA
XM_014635660.2	PREDICTED: Vigna radiata var. radiata metacaspase-9 (LOC106753803), mRNA
XM_028061471.1	PREDICTED: Vigna unguiculata metacaspase-9 (LOC114176425), mRNA

[1] www.eurofinsgenomics.eu/de/ecom/tools/oligo-analysis/

Table 5: Homolog sequences of γ-type VPE in Fabales used for primer design.

UniProt ID	Species	Gene aliases	Common name
I6N3H4	*Arachis diogoi*		VPE
A0A151T0P6	*Cajanus cajan*	KK1_023041	
A0A151SVD2	*Cajanus cajan*	KK1_014184	
A0A1S2Y9H1	*Cicer arietinum*	LOC101515346	
A0A1S2Y3V1	*Cicer arietinum*	LOC101514688	
A0A060D9Z7	*Clitoria ternatea*		CtAEP, Butelase-1
A0A222UCD2	*Clitoria ternatea*		CtAEP6
A0A0P0QMG1	*Clitoria ternatea*		AEP
A0A0P0QMJ1	*Clitoria ternatea*		AEP
I1MXD5	*Glycine max*	GLYMA_17G230700	
I1K8B5	*Glycine max*	GLYMA_06G050700	
I1JTV2	*Glycine max*	GLYMA_04G049900	
Q5QL07	*Glycine max*	GLYMA_14G092800	VPE2
A0A0B2RJV7	*Glycine soja*	D0Y65_013779	
A0A1J7FPR2	*Lupinus angustifolius*	TanjilG_00365	
A0A1J7GSL9	*Lupinus angustifolius*	TanjilG_12001	
B7FI64	*Medicago truncatula*	MTR_1g016780	
G7J7K1	*Medicago truncatula*	MTR_3g104120	
A0A0L9UHD9	*Phaseolus angularis*	LR48_Vigan04g237600	
V7AW15	*Phaseolus vulgaris*	PHAVU_009G076100g	
O24325	*Phaseolus vulgaris*	PHAVU_001G033500g	VPE1, LLP
A0A2K3NJ45	*Trifolium pratense*	L195_g026370	
P49044	*Vicia sativa*		Proteinase B
A0A0S3RPM5	*Vigna angularis*	Vigan.03G256300	
A0A0S3RB86	*Vigna angularis*	Vigan.02G058100	
Q9XFZ5	*Vigna mungo*		VmPE1a
Q9XFZ4	*Vigna mungo*		VmPE1
Q9AUD9	*Vigna radiata*	LOC106764744	VrPE1
A0A1S3U5Z5	*Vigna radiata*	LOC106762174	

Table 6: Homolog sequences of β-type VPE used for primer design.

UniProt ID	Species	Gene aliases	Common name
A0A151SKW2	*Cajanus cajan*	KK1_001604	
P49046	*Canavalia ensiformis*		AEP, Legumain
A0A1S2Z637	*Cicer arietinum*	LOC101491855	
A0A0P0QM28	*Clitoria ternatea*		Butelase-2
I1K0K8	*Glycine max*	GLYMA_05G055700	
Q9LLQ5	*Glycine max*	GLYMA_05G055700	PM40
P49045	*Glycine max*	GLYMA_17G137800	VPE1
A0A0B2P9P8	*Glycine soja*	D0Y65_011377	
A0A1J7HB47	*Lupinus angustifolius*	TanjilG_20543	
G7JK51	*Medicago truncatula*	MTR_4g101730	
A0A0L9TQC1	*Phaseolus angularis*	LR48_Vigan01g216800	
O24326	*Phaseolus vulgaris*	PHAVU_003G217500g	VPE2, PvVPE
O24539	*Vicia narbonensis*		VNPB2
O82102	*Vicia sativa*		CP
A0A1S3UW21	*Vigna radiata*	LOC106769123	

3.3.2 Polymerase chain reaction and gel electrophoresis

Polymerase chain reaction *(*PCR*)* was conducted with QIAGEN Taq PCR Core Kit (Qiagen, Hilden, Germany) according to the manufacturer's instruction manual, with primers produced by biomers.net GmbH. PCR was carried out on Biometra thermocyclers T-Gradient and T-Professional (Biomedizinische Analytik GmbH, Göttingen, Germany) under thermal conditions listed in Table 7. PCR products were separated by size using gel electrophoresis and then visualized by ethidium bromide staining. PCR amplicons with the correct band sizes were processed with a QIAGEN PCR Purification kit (Qiagen, Hilden, Germany) and sent to Eurofins MWG Operon for Sanger sequencing. The resulting data were imported to Geneious Prime for sequence identification and other downstream processes.

Table 7: PCR thermal cycle condition.

Step	Temperature (°C)	Time	
Initial denaturation	94°C	4 min	
Denaturation	94°C	30 sec	x 35 cycles
Annealing	50–65°C	1 min	
Elongation	72°C	1 min	
Final extension	72°C	7 min	

3.4 Sequences analysis

Sequence identification was achieved via Basic Local Alignment Search Tool (BLAST); DNA sequences were searched against the general nucleotide collection and the non-human expressed sequence tag database (dbEST). cDNA sequences were translated into protein for BLASTp; alignments were generated in the software Geneious, sequence logos were generated with HMMER (Finn *et al.*, 2011), and phylogenetic trees were built via the neighbor-joining method of the Jukes-Cantor model within the same software. In addition to homologous reference protein sequences identified from BLASTp results, supplementary data from DePamphilis *et al.* (2013), Fortin & Lam (2018), Yamada *et al.*, (2019), and Du *et al.* (2020) were used for selecting relevant data for annotation.

3.5 Gene expression studies

The gene expression of *candidate* genes was analyzed by quantitative reverse transcription PCR (qPCR) or real-time quantitative PCR. qPCR reactions were performed using the thermal cycler (Stratagene MX3005P, Agilent Technologies) and qPCR reagent kit (Quantinova qPCR Mix, SYBR Green, low ROX, QIAGEN) according to the manufacturer's manual. The reference gene 18S has been previously reported to show stable expression for HWF studies. Hence it was used for the normalization of threshold cycles (Cq) in the current project. The number of samples exceeded the capacity of the 96-well plates for the qPCR runs. Calibrator samples in three 10-fold serial dilutions are included on each plate for inter-run calibration.

3.5.1 Primer validation

qPCR Primers were designed based on the obtained sequences of candidate genes from *R. pseudoacaia*. They were then ordered from the manufacturer biomers.net GmbH. These primers were tested with gradient PCR in a range of temperatures, and their amplicon was validated by gel electrophoresis and Sanger sequencing. A final step of disassociation curve analysis of the amplicons on the qPCR instrument was followed to validate the primers' specificity.

3.5.2 Running conditions for the qPCR instrument

The thermal profile was set up according to the manufacturer's manual of the qPCR reagent kit (Quantinova qPCR Mix, SYBR Green, low ROX, QIAGEN). The temperature for combined annealing and extension was adjusted for each primer pair according to the results of primer validation (Figure 19).

Step	Time	Temperature	Ramp rate	Additional comments
PCR initial activation step	2 min	95°C	Maximal/ fast mode	QuantiNova DNA Polymerase is activated by this heating step
2-step cycling				
Denaturation	5 s	95°C	Maximal/ fast mode	
Combined annealing/ extension	10 s*	60°C	Maximal/ fast mode	Perform fluorescence data collection
Number of cycles	35–40‡			The number of cycles depends on the amount of template DNA
Melting curve analysis§				

* If your cycler does not accept this short time for data acquisition, use the shortest acceptable time.

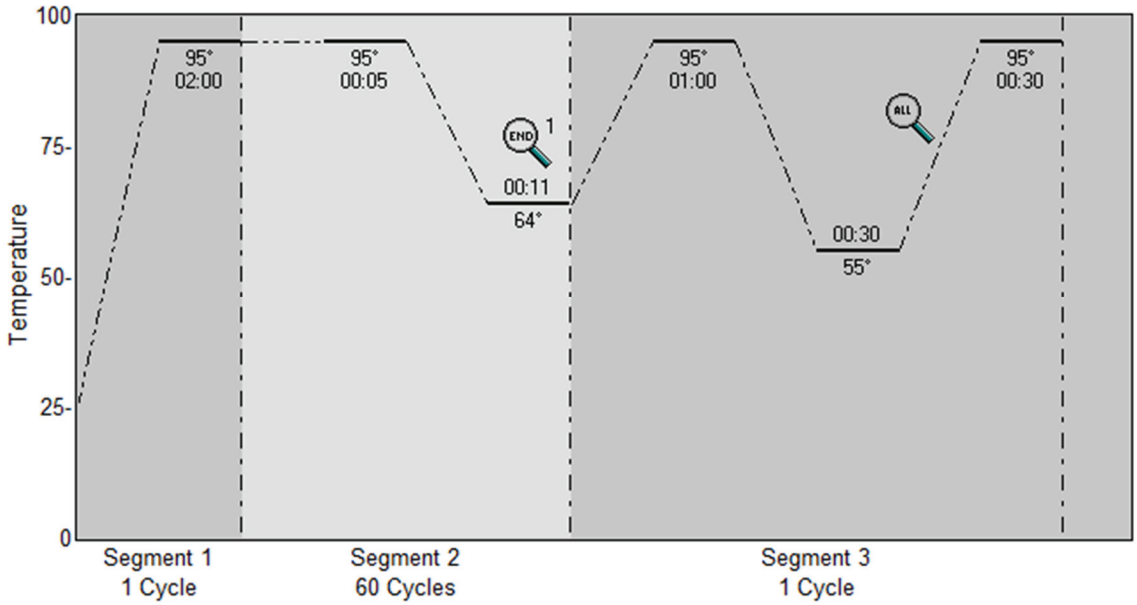


Figure 19: Thermal profile and cycling conditions for a qPCR run. (Top) Instructions on the Real-time cycler conditions from the manufacturers' manual of the qPCR reagent kit (Quantinova qPCR Mix, SYBR Green, low ROX, QIAGEN). (Bottom) Thermal profile on the qPCR instrument as shown in the Mx Pro software. Segment 2 in light grey highlights the replication cycles, with the fixed temperature at 95°C for the denaturing step. The combined annealing and extension temperature (shown as 64°C in this figure) is adjusted for each primer pair according to the results of primer validation (60°C – 65°C).

4. RESULTS

4.1 Species identification at the ITS region

All six trees for sampling were verified to be the target species, *Robinia pseudoacacia*. Figure 20 presents the sequence alignment of the PCR products from each *Robinia* tree in alignment with the *Rp18S* reference sequences from GeneBank (Accession numbers EF494737.1 and AF174637.1).

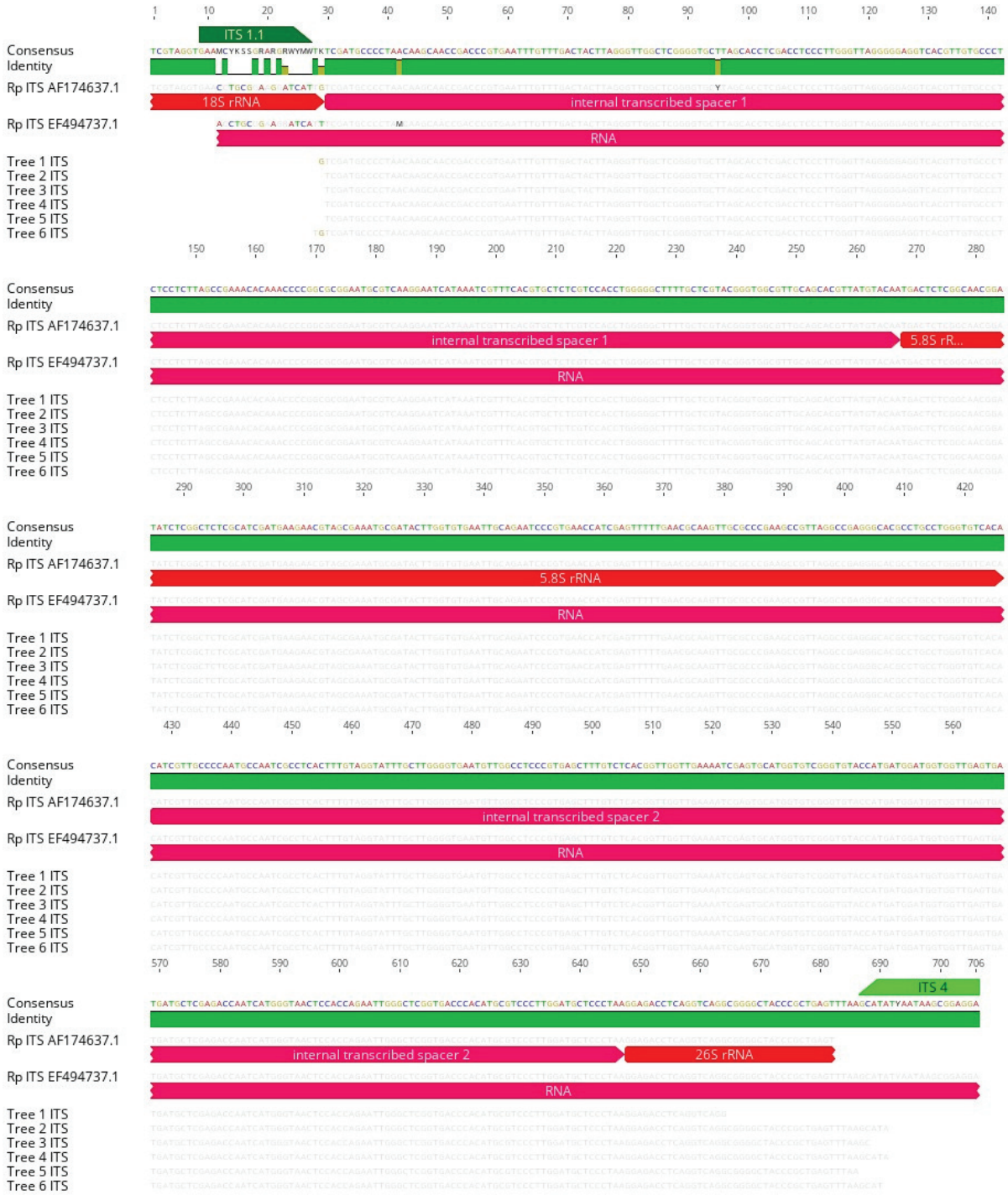


Figure 20: DNA sequence alignment of the ITS region from *Robinia pseudoacacia*. Sequences served as references are named with GeneBank accession numbers and annotated with the genes and ITS regions they encode in red bars. Tree 1 to 6 represents the individual specimen identified for the project. PCR primers in green bars are positioned at their binding regions.

4.2 DNA sequences of candidate genes in *R. pseudoacacia*

4.2.1 *RpMC-II* DNA sequences

With the degenerate primers that were designed based on the homologs of *AtMC9* in Fabales species, the initial result from PCR was an *RpMC9* sequence that contains four heterogeneous sites (Figure 21). Gene-specific primers were then designed based on these sites, resulting in two *RpMC9* sequences that cover the same region of the gene and differ by two of these heterogenic positions; information on the primers is listed in Table 8. These heterogeneous bases do not indicate any change in the amino acid sequence; an example is shown in

Figure 22. Therefore, they are unlikely to present two distinct *MC9* genes; for the current presentation, they are noted as *RpMC9a* and *RpMC9b*.

Table 8: List of primers used for obtaining the *RpMC9* sequences.

Primer name	Direction	Sequence (5' - 3')	AA Motif	Degeneracy	Melting Tm (°C)
Degenerate					
MC9 dF1	F	TGCRTGGYTGCATAAAYGATGT	RGCINDV	8	57.3 - 63.7
MC9 dR1	R	GTGATRAGATTGAARTCACARGG	PCDFNLIT	8	53.8 - 59.1
MC9 dR6	R	ACTRCAATARAGRCAAGGATG	HPCLYCS	8	50.9 - 58.1
Gene specific					
RpMC9 gsF1-c	F	TGCGTGGTTGCATAAACGATGT		0	61.7
RpMC9 gsF2	F	TGAAGCGTTTTGGGTTTGAT		0	56.4
RpMC9 gsF3	F	AACATTGAGCTCCTCACTGA		0	56.5
RpMC9 256R-A	R	ATCTCCTGCTTCAGCTTGAC		0	57.6
RpMC9 256R-G	R	ATCCCCTGCTTCAGCTTGAC		0	60.0
RpMC9 gsF6	F	TTCCGCCATGAGGAAG		0	52.8
RpMC9 gsR9-T	R	TGGGTTCATGTCTGCAGAGG		0	59.7

Figure 21: Overview of the sequencing Alignment of *RpMC9* and homologs listed in Table 4 . Primer information is listed in Table 8, and the full sequence alignment is in the appendix.

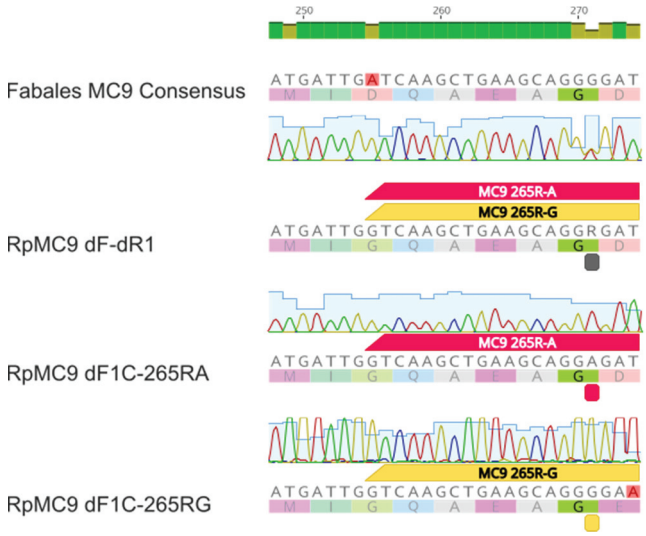
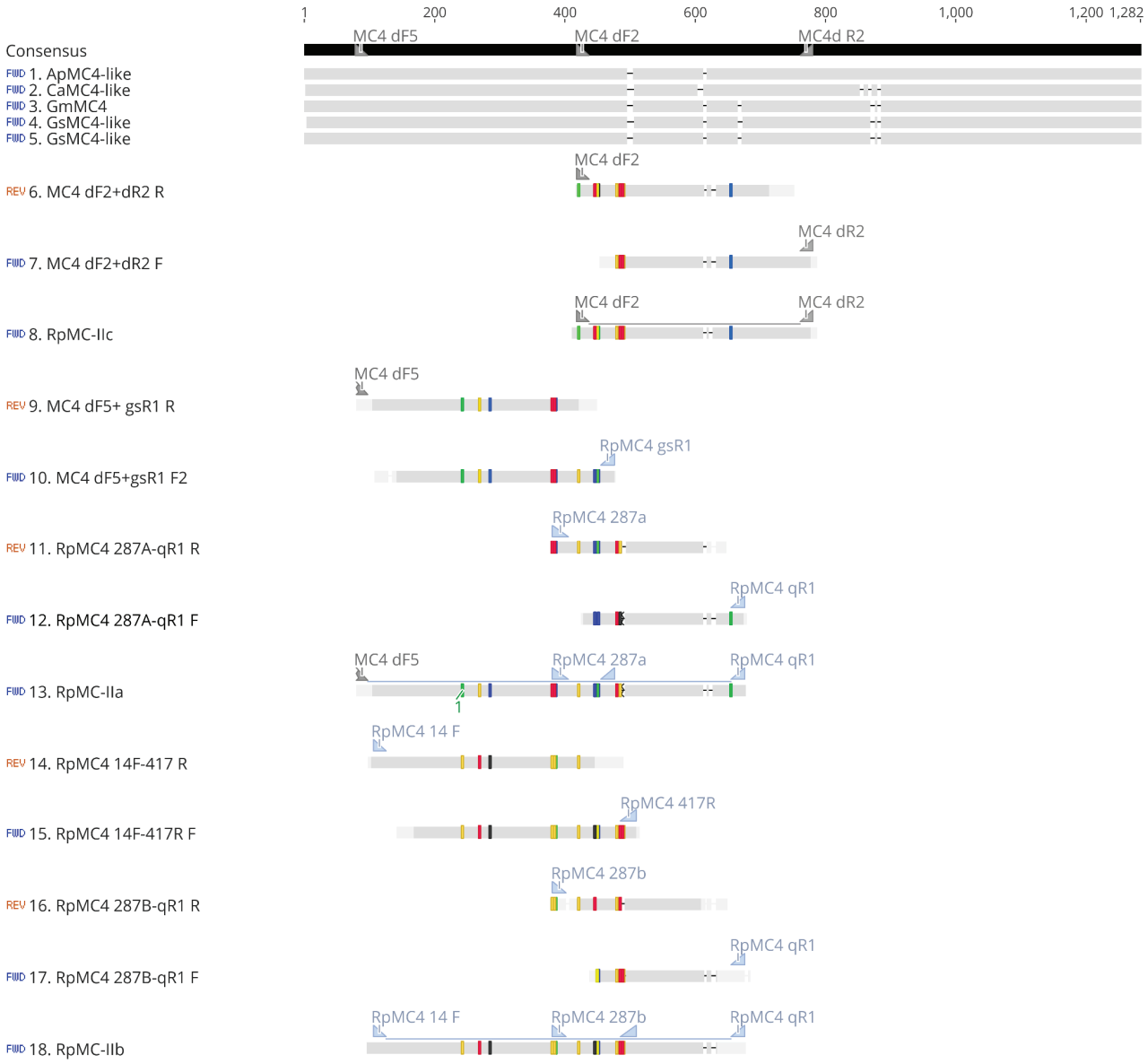


Figure 22: Heterogeneous sites in the two *RpMC9* DNA sequences encoding for the same amino acid residue.

With the degenerate primers that were designed on *AtMC4* homologs in legumes, three distinct sequences containing 17 heterogeneous sites were isolated (Figure 23). BLASTn and sequence alignment results showed that their identities are undistinguishably close to both *MC4* and *MC5*. Figure 24 shows an overview of the final sequencing results, where the obtained sequences are aligned with the homologs of both genes; all primers used to obtain these sequences are listed in Table 9. Despite the ambiguous identity, there is no dispute that they are members of type-II MCs. For the current presentation, the sequences are noted as *RpMC-IIa*, *RpMC-II*b, and *RpMC-II*c. Out of the 17 heterogeneous sites, 11 of them are associated with changes in amino acid, as shown in Figure 25. To determine whether or not these *RpMC-II* sequences are individual members of the MC gene family, further analysis based on their translated amino acid sequence is detailed in the next section, along with their gene phylogeny.

Table 9: List of primers used for obtaining the *RpMC-II* sequences.

Primer name	Direction	Sequence (5' - 3')	AA Motif	Degeneracy	Melting Tm (°C)
RpMC-IIc					
MC4 dF2	F	GATTCTTGCCAYAGTGGTG	DSCHSG	2	53.4 - 56.5
MC4 dR2	R	CRAARAGKGTAGGTCTCAG	LRPTLF	8	50.7 - 56.0
RpMC-IIa					
MC4 dF5	F	AACGACGTGWGGAGRATGCA	NDV(R/W) RM(H/Q)	4	59.7 - 62.1
RpMC4 gsR1	R	TGCTCTCTCCTATCTGCTCCT	--	0	59.5
RpMC4 287Fa	F	CAATACCCCTAGAGGTGTTAGGATC	--	0	63.0
RpMC4 qR1	R	ATAGCCATAGTTCCCATGTGG	--	0	57.9
RpMC-IIb					
RpMC4 14 F	F	TGGTCGAACGATACGGGTTC	--	0	59.3
RpMC4 417R	R	ACCAGAGCTAGAATGTTTCTCTTCTT	--	0	60.1
RpMC4 287Fb	F	CGATGTCCCTAGAGGTGTTAGG	--	0	59.4
RpMC4 qR1	R	ATAGCCATAGTTCCCATGTGG	--	0	57.9

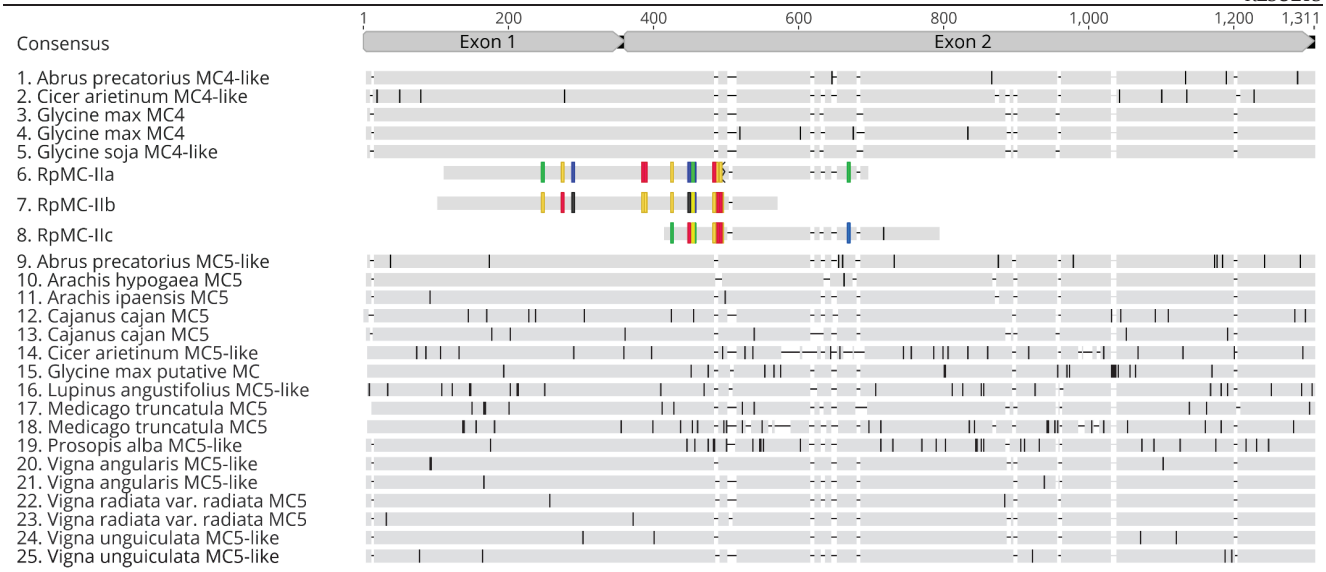


Lane	Sequence label	Sequence description	*Accession*
1	ApMC4-like	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X1, mRNA	XM_027499192.1
2	CaMC4-like	PREDICTED: Cicer arietinum metacaspase-4-like (LOC101509880), mRNA	XM_004509944.3
3	GmMC4	PREDICTED: Glycine max metacaspase-4 (LOC100802658), mRNA	XM_003546612.5
4	GsMC4-like	PREDICTED: Glycine soja metacaspase-4-like (LOC114387810), mRNA	XM_028348014
5	GsMC4-like	PREDICTED: Glycine soja metacaspase-4-like (LOC114422611), mRNA	XM_028389058
6 -7	MC4 dF2+dR2 F	Amplicons from PCR with primers MC4 dF2 and MC4 dR2	--
8	RpMC-IIc	Consensus sequence of lanes 6 and 7	--
9 - 10	MC4 dF5 – gsR1 F & R	Amplicons from PCR with primers MC4 dF5 and MC4 gsR1	--
11 - 12	RpMC4 287A-qR1 F & R	Amplicons from PCR with primers RpMC4 287Fa and RpMC4 qR1	--
13	RpMC-IIa	Consensus sequence of lanes 9 to 12	--
14 - 15	RpMC4 14F-417 F & R	Amplicons from PCR with primers RpMC4 14F and RpMC4 417R	--
16 - 17	RpMC4 287B-qR1 F & R	Amplicons from PCR with primers RpMC4 287Fb and RpMC4 qR1	--
18	RpMC-IIb	Consensus sequence of lanes 14 to 17	--

-- Indicates not applicable

Figure 23: DNA sequence alignment overview illustrates the sequencing results of *RpMC-IIs*. Grey triangles indicate the degenerate primers, and the blue ones correspond to gene specific primers listed in Table 9.

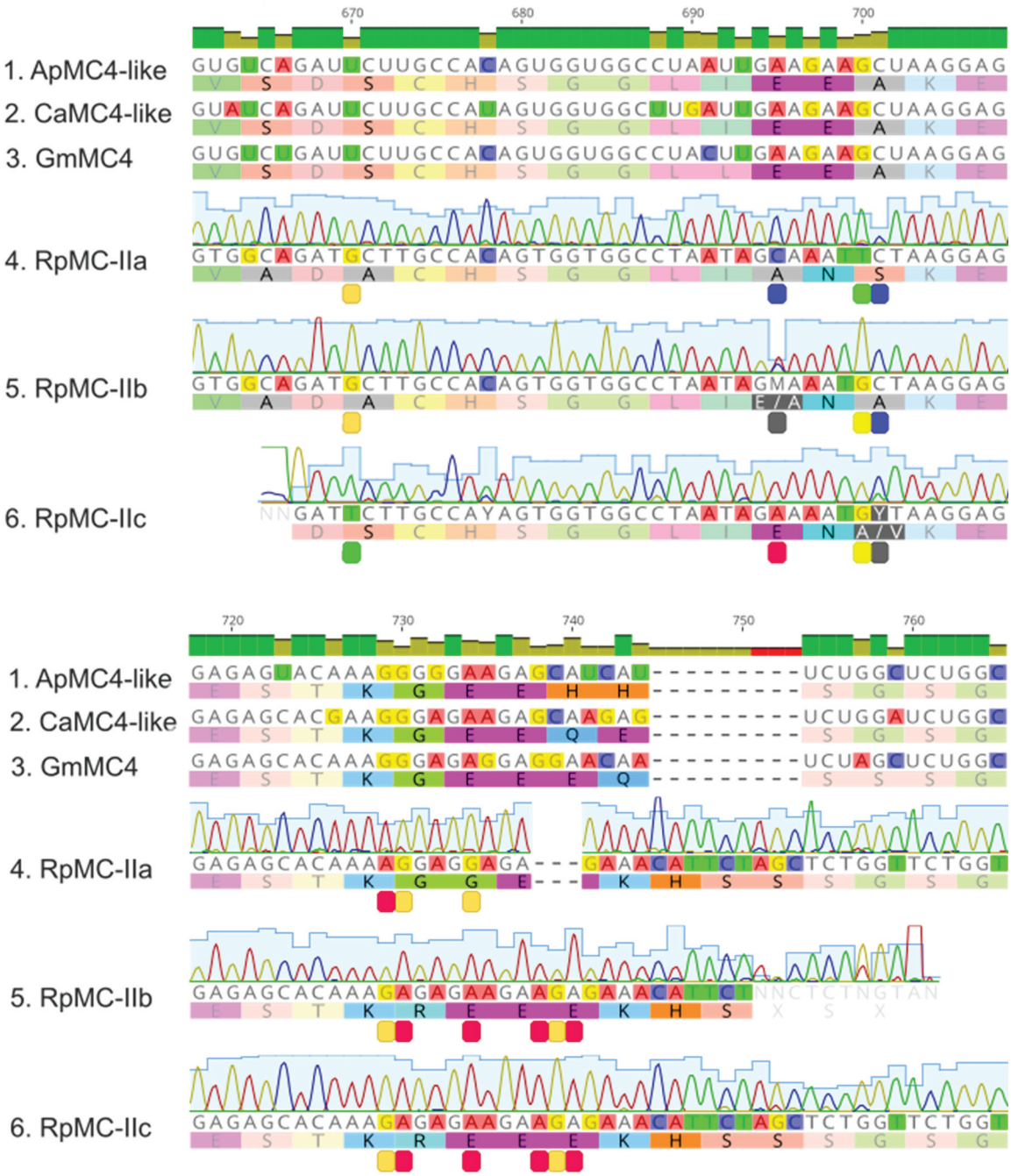
The intended size for this page is A3. The full-sized page is available in the supplement material.



Lane	Sequence label	Sequence description	*Accession*
1	Abrus precatorius MC4-like	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X1, mRNA	XM_027499192.1
2	Cicer arietinum MC4-like	PREDICTED: Cicer arietinum metacaspase-4-like (LOC101509880), mRNA	XM_004509944.3
3	Glycine max MC4	PREDICTED: Glycine max metacaspase-4 (LOC100795043), mRNA	XM_003531762.5
4	Glycine max MC4	PREDICTED: Glycine max metacaspase-4 (LOC100802658), mRNA	XM_003546612.5
5	Glycine soja MC4-like	PREDICTED: Glycine soja metacaspase-4-like (LOC114422611), mRNA	XM_028389058.1
6	RpMC-IIa	**Robinia pseudoacacia Type-II matecaspase** RpMC-IIa	--
7	RpMC-IIb	**Robinia pseudoacacia Type-II matecaspase RpMC-IIb**	--
8	RpMC-IIc	**Robinia pseudoacacia Type-II matecaspaseRpMC-IIc**	--
9	Arachis hypogaea MC5	PREDICTED: Arachis hypogaea metacaspase-5 (LOC112707436), mRNA	XM_025759180.2
10	Arachis ipaensis MC5	PREDICTED: Arachis ipaensis metacaspase-5 (LOC107635285), mRNA	XM_016338719.2
11	Abrus precatorius MC5-like	PREDICTED: Abrus precatorius metacaspase-5-like (LOC113865563), transcript variant X1, mRNA	XM_027500219.1
12	Cajanus cajan MC5	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789350), mRNA	XM_020348274.2
13	Cajanus cajan MC5	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789349), transcript variant X1, mRNA	XM_020348272.2
14	Cicer arietinum MC5-like	PREDICTED: Cicer arietinum metacaspase-5-like (LOC101509553), mRNA	XM_004488501.3
15	Glycine max putative MC	Glycine max putative metacaspase (LOC100527809), mRNA	NM_001250818.2
16	Lupinus angustifolius MC5-like	PREDICTED: Lupinus angustifolius metacaspase-5-like (LOC109352001), mRNA	XM_019593760.1
17	Medicago truncatula MC5	PREDICTED: Medicago truncatula metacaspase-5 (LOC25502442), mRNA	XM_013589850.3
18	Medicago truncatula MC5	PREDICTED: Medicago truncatula metacaspase-5 (LOC11419585), mRNA	XM_003595881.4
19	Prosopis alba MC5-like	PREDICTED: Prosopis alba metacaspase-5-like (LOC114754970), mRNA	XM_028943832.1
20	Vigna angularis MC5-like	PREDICTED: Vigna angularis metacaspase-5-like (LOC108338807), mRNA	XM_017575836.1
21	Vigna angularis MC5-like	PREDICTED: Vigna angularis metacaspase-5-like (LOC108340250), mRNA	XM_017577495.1
22	Vigna radiata var. radiata MC5	PREDICTED: Vigna radiata var. radiata metacaspase-5 (LOC106752949), mRNA	XM_014634726.2
23	Vigna radiata var. radiata MC5	PREDICTED: Vigna radiata var. radiata metacaspase-5 (LOC106752959), mRNA	XM_014634744.2
24	Vigna unguiculata MC5-like	PREDICTED: Vigna unguiculata metacaspase-5-like (LOC114169553), mRNA	XM_028054754.1
25	Vigna unguiculata MC5-like	PREDICTED: Vigna unguiculata metacaspase-5-like (LOC114170591), mRNA	XM_028056101.1

-- Indicates not applicable

Figure 24: DNA sequence alignment overview for the homologs of *AtMC4* and *AtMC5* in Fabales. Black lines in the sequences correspond to disagreement with the majority. Information on the sequences used in the alignment is listed in the table under the alignment; BLASTn result of each *RpMC-II* and the full sequence alignment is in the appendix.



Lane	Sequence label	Sequence description	*Accession*
1	*ApMC4-like*	Homolog, MC4-like in *Abrus precatorius*	XM_027499192
2	*CaMC4-like*	Homolog, MC4-like in *Cicer arietinum*	XM_004509944
3	*GmMC4*	Homolog, MC4 in *Glycine max*	XM_003546612
4	*RpMC-IIa*	Amplicon of *RpMC-IIa* obtained via PCR	--
5	*RpMC-IIb*	Amplicon of *RpMC-IIb* obtained via PCR	--
6	*RpMC-IIc*	Amplicon of *RpMC-IIc* obtained via PCR	--
			-- Indicates not applicable

Figure 25: DNA and amino acid alignment of MC4 homologs and *RpMC-II*. With the difference in DNA at base 695, 700, and 701, the corresponding base of amino acids also differ. Within bases, 730 to 755, and 3 extra amino acid bases are inserted into all the *RpMC-II* sequences compared to the homologs in other species.

4.2.2 *RpVPEs* DNA sequence

The first set of obtained sequences using degenerate primers was identified as two different types of VPE. BLASTn and sequence alignment with their homologs confirmed that one of them is β-type and the other is γ-type, and they were named *β-RpVPE* and γ-*RpVPE* accordingly. γ-*RpVPE* is identical to an expressed sequence tag (EST) that was registered in NCBI with the gene expression profiling studies Yang *et al.* (2003; 2004). The EST was then included in the design of gene-specific primers, as shown in the sequencing result overview of γ-*RpVPE* in Figure 27. Figure 26 shows the alignment overview of *β-RpVPE* DNA sequences and the coverage over its homologs, the information on the sequences used in the alignment is detailed in Table 12. The primers used for obtaining the *β-RpVPE* and γ-*RpVPE* sequences are listed in Table 10Table 11 respectively.

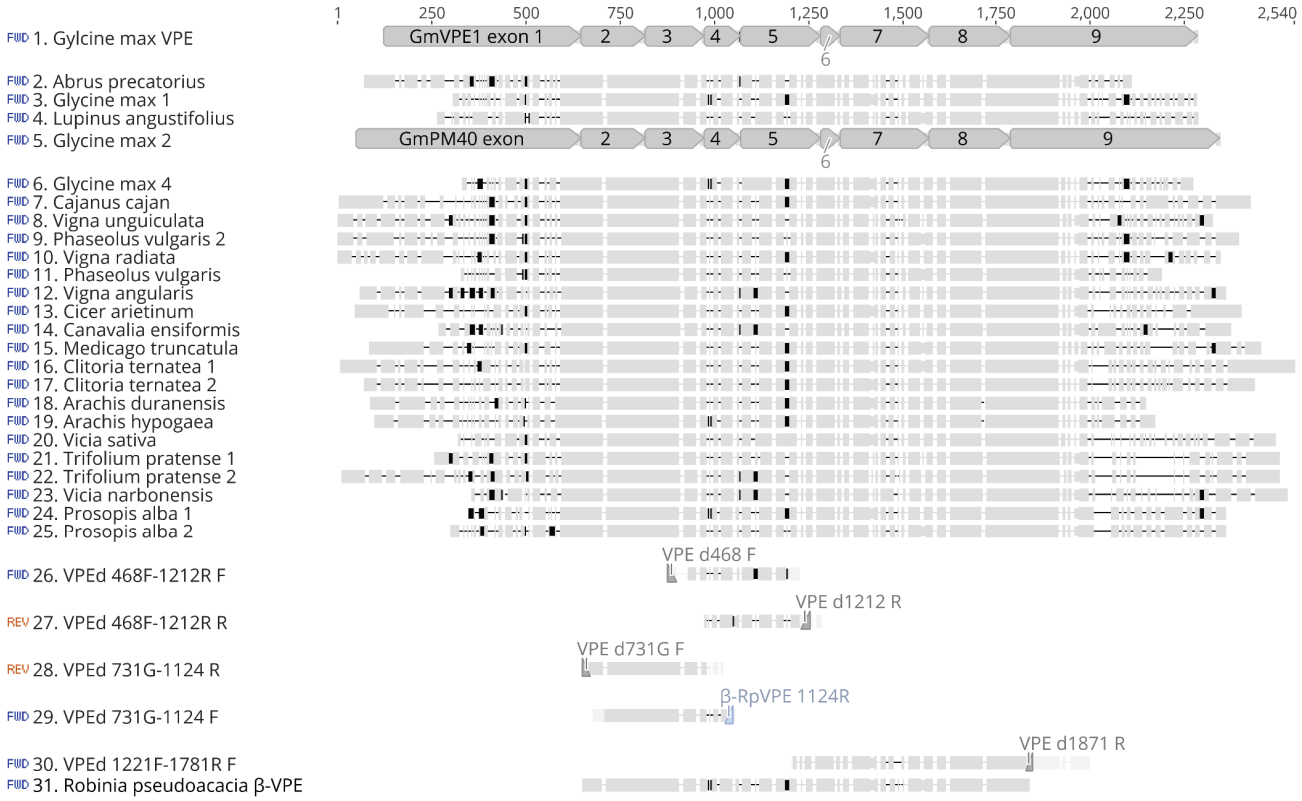
Table 10: List of primers used for obtaining the *β-RpVPE* sequences

Primer name	Direction	Sequence (5' - 3')	AA Motif	Degeneracy	Melting Tm (°C)
Degenerate					
VPE d731FG	F	GATGTGTGCCATGCDTAYCA	DVCHYQ	6	56.2 - 61.0
VPE d468F	F	TAARGTDAARGGWGGAAGTGGCAA	K(V/L)KGGSGK	24	58.3 - 64.1
VPE d1212R	R	CAAATCCCCHAGGCAAGTGATGT	YITCLGDL	3	60.6 - 63.3
VPE d1781R	R	CGCATRTGYTTCATNCCATA	YGMKHMR	16	50.8 - 59.3
Gene specific					
β-RpVPE 1221F	F	TGCATCCAATGCACAAGAGAG	--	0	58.9
β-RpVPE 1124R	R	CCAGATGCATGTTTCTTCTT	--	0	53.9

Table 11: List of primers used for obtaining the γ-*RpVPE* sequences.

Primer name	Direction	Sequence (5' - 3')	AA Motif	Degeneracy	Melting Tm (°C)
Degenerate					
VPE d731FW	F	GATGTWTGYCATGCATACCA	DVCHYQ	4	52.8 - 56.7
VPE d1339R	R	CTTCCATCCAAGMAACRCT	SVAWME	4	52.4 - 58.4
Gene specific					
γ-RpVPE 948F	F	TGGCAGTGGGAAGGTTGTGG	--	0	61.4
γ-RpVPE 1235R	R	ACTAGGGTACTCCCCAGGGC	--	0	63.4
γ-RpVPE 1024F	F	GTGCTTGGGATGCCTGTTG	--	0	59.4
γ-RpVPE YR1	R	TGGTCCTTTATCGCTTCTTATGT	--	0	57.6

The intended size for this page is A3. The full-sized page is available in the supplement material.



Lane	Sequence label	Accession	Description	Query coverage	Pairwise Identity	Bit-Score
1	Glycine max VPE	NM_001249749.2	Glycine max vacuolar-processing enzyme (VPE), mRNA	100.00%	89.6%	1421.19
2	Abrus precatorius	XM_027479555.1	PREDICTED: Abrus precatorius legumain (LOC113849561), mRNA	100.00%	89.5%	1410.11
3	Glycine max 1	D28876.1	Glycine max mRNA for cysteine proteinase, complete cds	100.00%	89.5%	1410.11
4	Lupinus angustifolius	XM_019591087.1	PREDICTED: Lupinus angustifolius legumain (LOC109349992), mRNA	100.00%	89.3%	1399.03
5	Glycine max 2	NM_001349197.1	Glycine max seed maturation protein PM40 (PM40), mRNA	100.00%	89.1%	1387.95
6	Cajanus cajan	XM_020370788.2	PREDICTED: Cajanus cajan vacuolar-processing enzyme (LOC109808005), mRNA	100.00%	88.0%	1321.47
7	Vigna unguiculata	XM_028063059.1	PREDICTED: Vigna unguiculata vacuolar-processing enzyme (LOC114177634), mRNA	100.00%	88.0%	1319.63
8	Vigna radiata	XM_014654608.2	PREDICTED: Vigna radiata var. radiata vacuolar-processing enzyme (LOC106769123), mRNA	100.00%	87.8%	1308.55
9	Phaseolus vulgaris 2	XM_007155562.1	Phaseolus vulgaris hypothetical protein (PHAVU_003G217500g) mRNA, complete cds	100.00%	87.8%	1304.85
10	Phaseolus vulgaris	Z99957.1	Phaseolus vulgaris Moldavian encoding legumain-like proteinase precursor (clone p21b)	100.00%	87.8%	1304.85
11	Vigna angularis	XM_017554660.1	PREDICTED: Vigna angularis vacuolar-processing enzyme (LOC108322532), mRNA	100.00%	87.3%	1277.16
12	Cicer arietinum	XM_004515685.3	PREDICTED: Cicer arietinum legumain (LOC101491855), mRNA	100.00%	87.0%	1260.54
13	Canavalia ensiformis	D31787.1	Canavalia ensiformis mRNA for asparaginyl endopeptidase, complete cds	100.00%	87.0%	1255.0
14	Medicago truncatula	XM_003608718.4	PREDICTED: Medicago truncatula legumain (LOC11405547), mRNA	100.00%	86.8%	1243.92
15	Clitoria ternatea 1	MT468738.1	Clitoria ternatea clone ctr29014_c2_g1_i2 asparaginyl endopeptidase 2 mRNA, complete cds	100.00%	86.8%	1243.92
16	Clitoria ternatea 2	KR912009.1	Clitoria ternatea clone Ct_flwr_1565_g1_i1 asparaginyl endopeptidase mRNA, complete cds	100.00%	86.7%	1238.38
17	Arachis hypogaea	XM_025846712.2	PREDICTED: Arachis hypogaea legumain (LOC112803196), mRNA	100.00%	85.6%	1170.05
18	Arachis duranensis	XM_016107728.2	PREDICTED: Arachis duranensis legumain (LOC107487130), mRNA	100.00%	85.6%	1170.05
19	Vicia sativa	AJ007743.1	Vicia sativa mRNA for cysteine proteinase precursor	100.00%	85.2%	1142.35
20	Trifolium pratense 2	XM_045932441.1	PREDICTED: Trifolium pratense legumain (LOC123883594), transcript variant X1, mRNA	100.00%	84.8%	1122.04
21	Trifolium pratense 1	XM_045932442.1	PREDICTED: Trifolium pratense legumain (LOC123883594), transcript variant X2, mRNA	100.00%	84.8%	1122.04
22	Vicia narbonensis	Z99174.1	Vicia narbonensis mRNA for cysteine proteinase precursor	100.00%	84.7%	1116.5
23	Prosopis alba 2	XM_028906873.1	PREDICTED: Prosopis alba legumain-like (LOC114721067), mRNA	100.00%	84.1%	1077.72
24	Prosopis alba 1	XM_028931650.1	PREDICTED: Prosopis alba legumain-like (LOC114743439), mRNA	100.00%	84.1%	1077.72
25 - 26	VPEd 468F-1212R F & R	--	Amplicons from PCR with primers VPE d468F and VPE d1212R	--	--	--
27 - 28	VPEd 731G-1124 F & R	--	Amplicons from PCR with primers VPE d731FG and β-RpVPE 1124R	--	--	--
29	VPEd 1221-1781 F	--	Amplicons from PCR with primers β-RpVPE 1221F and VPE d1781R	--	--	--
30	Robinia pseudoacacia β-VPE	--	Consensus sequence from the PCR amplicons in lanes 25 to 29	--	--	--

Figure 26: Alignment overview of *β-RpVPE* sequencing results. Information on the sequences is listed in the table herein, details of the primers are listed in Table **10**, and the full sequence alignment is in the appendix.

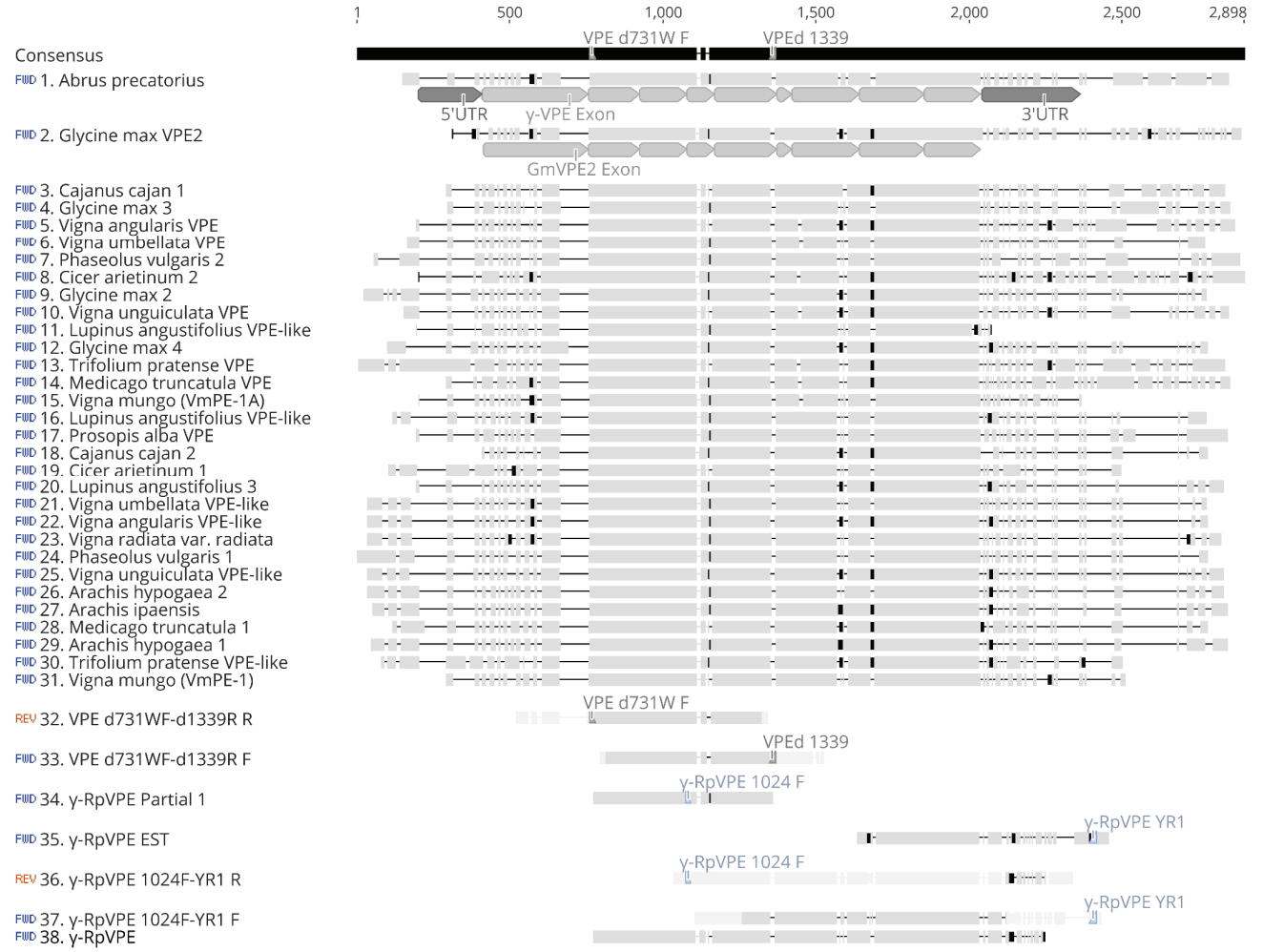


Figure 27: BLASTn results of γ-*RpVPE* in sequence alignment overview. Homologous sequences are sorted in descending order of pairwise identity percentage to *γ-RpVPE*. Grey elements present the annotations as exons provided in the NCBI database. Details on the primers are listed in Table 11, and information on the sequences is listed in Table 12 on the next page

Table 12: Sequence information for the alignment in Figure 27.

Lane	Sequence label	Accession	Description	Query coverage	Pairwise Identity	Bit-Score
1	Abrus precatorius	XM_027513192.1	PREDICTED: Abrus precatorius vacuolar-processing enzyme-like (LOC113874911), transcript variant X1, mRNA	92.05%	90.6%	1716.66
2	Glycine max (VPE2)	NM_001249635.2	Glycine max vacuolar processing enzyme 2 (VPE2), mRNA	96.05%	90.5%	1703.67
3	Cajanus cajan 1	XM_020366661.2	PREDICTED: Cajanus cajan vacuolar-processing enzyme (LOC109804800), mRNA	92.05%	89.0%	1598.47
4	Glycine max 3	XM_003550235.5	PREDICTED: Glycine max vacuolar-processing enzyme (LOC100784318), mRNA	97.16%	88.8%	1659.41
5	Vigna umbellata VPE	XM_047306939.1	PREDICTED: Vigna umbellata vacuolar-processing enzyme (LOC124832678), mRNA	91.77%	88.1%	1526.45
6	Vigna angularis VPE	XM_017574898.1	PREDICTED: Vigna angularis vacuolar-processing enzyme (LOC108338162), transcript variant X1, mRNA	91.77%	88.1%	1526.45
7	Phaseolus vulgaris 2	XM_007160923.1	Phaseolus vulgaris hypothetical protein (PHAVU_001G033500g) mRNA, complete cds	91.70%	87.8%	1506.14
8	Cicer arietinum 2	XM_027333646.1	PREDICTED: Cicer arietinum vacuolar-processing enzyme-like (LOC101514688), transcript variant X1, mRNA	92.05%	87.8%	1511.68
9	Glycine max 2	XM_003525931.5	PREDICTED: Glycine max vacuolar-processing enzyme (LOC100791675), mRNA	88.79%	87.8%	1459.97
10	Vigna unguiculata VPE	XM_028084964.1	PREDICTED: Vigna unguiculata vacuolar-processing enzyme (LOC114194621), mRNA	91.77%	87.6%	1491.37
11	Lupinus angustifolius VPE-like	XM_019573985.1	PREDICTED: Lupinus angustifolius vacuolar-processing enzyme-like (LOC109337103), mRNA	84.74%	87.1%	1345.48
12	Glycine max 4	XM_014774474.3	PREDICTED: Glycine max vacuolar-processing enzyme (LOC100807010), transcript variant X1, mRNA	88.79%	87.0%	1406.42
13	Trifolium pratense VPE	XM_045968705.1	PREDICTED: Trifolium pratense vacuolar-processing enzyme (LOC123917080), mRNA	88.86%	86.8%	1387.95
14	Medicago truncatula VPE	XM_013610451.3	PREDICTED: Medicago truncatula vacuolar-processing enzyme (LOC25482000), mRNA	91.34%	86.8%	1417.5
15	Vigna mungo (VmPE-1A)	D89972.1	Vigna mungo mRNA for asparaginyl endopeptidase (VmPE-1A), complete cds	91.77%	86.7%	1426.73
16	Lupinus angustifolius VPE-like	XM_019561080.1	PREDICTED: Lupinus angustifolius vacuolar-processing enzyme-like (LOC109327911), mRNA	89.99%	86.5%	1389.8
17	Prosopis alba VPE	XM_028940230.1	PREDICTED: Prosopis alba vacuolar-processing enzyme (LOC114751576), mRNA	87.01%	86.5%	1341.79
18	Cajanus cajan 2	XM_020368305.2	PREDICTED: Cajanus cajan vacuolar-processing enzyme (LOC109805986), mRNA	88.22%	86.1%	1336.25
19	Cicer arietinum 1	XM_004501420.3	PREDICTED: Cicer arietinum vacuolar-processing enzyme-like (LOC101515346), mRNA	88.72%	85.7%	1308.55
20	Lupinus angustifolius 3	XM_019582316.1	PREDICTED: Lupinus angustifolius vacuolar-processing enzyme-like (LOC109343833), mRNA	89.07%	85.7%	1315.93
21	Vigna umbellata VPE-like	XM_047316214.1	PREDICTED: Vigna umbellata vacuolar-processing enzyme-like (LOC124840200), mRNA	88.72%	85.6%	1310.39
22	Vigna angularis VPE-like	XM_017566611.1	PREDICTED: Vigna angularis vacuolar-processing enzyme-like (LOC108331721), mRNA	88.72%	85.6%	1304.85
23	Vigna radiata var. radiata	XM_014645928.2	PREDICTED: Vigna radiata var. radiata vacuolar-processing enzyme (LOC106762174), mRNA	88.72%	85.5%	1299.31
24	Phaseolus vulgaris 1	XM_007136747.1	Phaseolus vulgaris hypothetical protein (PHAVU_009G076100g) mRNA, complete cds	88.72%	85.4%	1293.78
25	Vigna unguiculata VPE-like	XM_028047286.1	PREDICTED: Vigna unguiculata vacuolar-processing enzyme-like (LOC114163155), mRNA	88.72%	85.2%	1279.0
26	Arachis ipaensis	XM_016308381.2	PREDICTED: Arachis ipaensis vacuolar-processing enzyme (LOC107606332), mRNA	90.70%	85.2%	1304.85
27	Arachis hypogaea 2	XM_025810236.2	PREDICTED: Arachis hypogaea vacuolar-processing enzyme (LOC112764565), mRNA	90.70%	85.2%	1304.85
28	Medicago truncatula 1	XM_003603073.4	PREDICTED: Medicago truncatula vacuolar-processing enzyme gamma-isozyme (LOC11432331), mRNA	87.51%	84.8%	1234.68
29	Arachis hypogaea 1	XM_025757656.2	PREDICTED: Arachis hypogaea vacuolar-processing enzyme (LOC112706381), mRNA	90.70%	84.7%	1271.62
30	Trifolium pratense VPE-like	XM_045944193.1	PREDICTED: Trifolium pratense vacuolar-processing enzyme-like (LOC123894249), mRNA	90.70%	84.7%	1271.62
31	Vigna mungo (VmPE-1)	D89971.1	Vigna mungo mRNA for asparaginyl endopeptidase (VmPE-1), complete cds	88.29%	84.6%	1232.84
32 - 33	VPE d731FW-d1339R F & R	--	Amplicons from PCR with primers VPE d731FW and VPE d1339R	--	--	--
34	γ-RpVPE Partial 1	--	Consensus sequence from the amplicons in lanes 32 and 33 with gene-specific primer	--	--	--
35	γ-RpVPE EST	BI677606	EST of γ-RpVPE in reference to (Yang *et al.*, 2004)	--	--	--
36 - 37	γ-RpVPE 1024F – YR1 F & R	--	Amplicons from PCR with primers γ-RpVPE 1024F and γ-RpVPE YR1	--	--	--
38	γ-RpVPE	--	Consensus sequence from the PCR amplicons in lanes 32 to 33 and 36 to 37			

4.3 Protein sequence analysis

4.3.1 Type–II Metacaspases protein sequences

Since the two *RpMC9* encode for no difference in amino acid sequence, they were combined in length and labeled as RpMC9 for protein sequence analysis, alignments, and construction of the gene trees. Taking *RpMC9* as the query, BLASTp results within the Fabales are listed in Table 13, and the result without limitation to taxa is listed in Table 14. Figure 28 shows the phylogeny of MC9 homologs from legumes; sequences with curated annotation in the database or association with empirical research are included for comparison (Table 15). There are four conserved aspartates (D) residues in the sequence that are predicted to contribute to calcium binding; in legumes, one of them is substituted by a Glutamic acid (E) in legume, including *R. pseudoacacia.*

The gene name in Norway spruce and hybrid aspen differ from the norm, hence clarified here. *PaMC5* in Norway spruce, as introduced in the earlier section, is an *AtMC9* homolog with strongly up-regulated gene expression associated with earlywood formation (Jokipii-Lukkari *et al.,* 2018). *PttMC13* and *PttMC14* from *Populous tremula x tremuloides* are *AtMC9* homologs with proposed xylem-specific function in Bollhöner *et al.* (2018).

Table 13: BLASTp results of RpMC9 limited to leguminous species.

Query	Bit-Score	Query coverage	Pairwise Identity	Organism	Description	Accession
RpMC9a	461.84	98.86%	85.4%	*Abrus precatorius*	Metacaspase-9 [Abrus precatorius]	XP_027351119
	452.595	98.86%	83.9%	*Lupinus albus*	Putative Caspase-like domain-containing protein [Lupinus albus]	KAE9599704
	450.669	98.86%	83.9%	*Lupinus angustifolius*	PREDICTED: metacaspase-9 [Lupinus angustifolius]	XP_019457161
	445.662	98.86%	84.5%	*Cicer arietinum*	Metacaspase-9 [Cicer arietinum]	XP_004493282
	431.409	99.24%	80.7%	*Medicago truncatula*	Metacaspase-9 [Medicago truncatula]	XP_013449529
	428.713	99.24%	80.3%	*Medicago truncatula*	Unknown [Medicago truncatula]	AFK49292
	428.328	98.86%	80.6%	*Trifolium pratense*	Metacaspase-9-like protein [Trifolium pratense]	PNY07712
	426.787	98.86%	80.6%	*Mucuna pruriens*	Metacaspase-9 [Mucuna pruriens]	RDY12566
	426.017	98.86%	80.2%	*Trifolium pratense*	Metacaspase-9 [Trifolium pratense]	XP_045796026
	419.083	98.86%	79.5%	*Glycine max*	Hypothetical protein JHK86_053251 [Glycine max]	KAG4912818
	418.313	98.86%	79.1%	*Glycine max*	Metacaspase-9 [Glycine max]	XP_003554078
	417.927	98.86%	79.1%	*Glycine soja*	Metacaspase-9 [Glycine soja]	XP_028217188
	417.542	98.86%	79.1%	*Glycine max*	Hypothetical protein JHK85_054186 [Glycine max]	KAG4927700
	416.387	99.24%	75.7%	*Prosopis alba*	Metacaspase-9-like [Prosopis alba]	XP_028766727
	415.616	99.24%	75.3%	*Prosopis alba*	Metacaspase-9-like [Prosopis alba]	XP_028772742
	412.149	98.86%	77.9%	*Arachis hypogaea*	Hypothetical protein Ahy_B06g083752 [Arachis hypogaea]	RYR04163
	411.764	98.86%	77.9%	*Arachis ipaensis*	Metacaspase-9 [Arachis ipaensis]	XP_016161870
	409.838	98.86%	77.5%	*Arachis duranensis*	Metacaspase-9 [Arachis duranensis]	XP_015970740
	409.453	98.86%	77.5%	*Arachis hypogaea*	Metacaspase-9 [Arachis hypogaea]	XP_025605483
	405.216	98.86%	76.4%	*Cajanus cajan*	Metacaspase-9 [Cajanus cajan]	XP_020219841
RpMC9b	392.119	99.55%	85.9%	*Abrus precatorius*	Metacaspase-9 [Abrus precatorius]	XP_027351119
	381.719	99.55%	84.1%	*Lupinus albus*	Putative Caspase-like domain-containing protein [Lupinus albus]	KAE9599704
	379.793	99.55%	85.5%	*Cicer arietinum*	Metacaspase-9 [Cicer arietinum]	XP_004493282
	376.326	99.55%	83.6%	*Lupinus angustifolius*	PREDICTED: metacaspase-9 [Lupinus angustifolius]	XP_019457161
	362.844	100.00%	81.2%	*Medicago truncatula*	Metacaspase-9 [Medicago truncatula]	XP_013449529
	362.073	99.55%	81.4%	*Glycine max*	Metacaspase-9 [Glycine max]	XP_003554078
	362.073	99.55%	81.4%	*Glycine max*	Hypothetical protein JHK86_053251 [Glycine max]	KAG4912818
	362.073	99.55%	81.4%	*Glycine max*	Hypothetical protein JHK85_054186 [Glycine max]	KAG4927700
	361.688	99.55%	81.4%	*Glycine soja*	Metacaspase-9 [Glycine soja]	XP_028217188
	360.918	99.55%	82.0%	*Mucuna pruriens*	Metacaspase-9 [Mucuna pruriens]	RDY12566
	360.147	100.00%	80.7%	*Medicago truncatula*	Unknown [Medicago truncatula]	AFK49292
	359.377	99.55%	80.6%	*Trifolium pratense*	Metacaspase-9-like protein [Trifolium pratense]	PNY07712
	357.066	99.55%	80.2%	*Trifolium pratense*	Metacaspase-9 [Trifolium pratense]	XP_045796026
	351.673	99.55%	78.4%	*Cajanus cajan*	Metacaspase-9 [Cajanus cajan]	XP_020219841
	350.132	100.00%	75.7%	*Prosopis alba*	Metacaspase-9-like [Prosopis alba]	XP_028766727
	348.977	100.00%	75.2%	*Prosopis alba*	Metacaspase-9-like [Prosopis alba]	XP_028772742
	348.977	99.55%	78.7%	*Arachis hypogaea*	Hypothetical protein Ahy_B06g083752 [Arachis hypogaea]	RYR04163
	348.206	99.55%	78.7%	*Arachis ipaensis*	Metacaspase-9 [Arachis ipaensis]	XP_016161870
	348.206	99.55%	78.7%	*Arachis hypogaea*	Metacaspase-9 [Arachis hypogaea]	XP_025605483
	348.206	99.55%	78.7%	*Arachis duranensis*	Metacaspase-9 [Arachis duranensis]	XP_015970740

Table 14: BLASTp results of RpMC9, results not limited by taxa.

Query	Bit-Score	Query coverage	Pairwise Identity	Organism	Description	Accession
RpMC9a	405601	98.86%	75.5%	*Populus alba*	metacaspase-9-like [Populus alba]	XP_034919959
	404.831	98.86%	75.5%	*Populus tomentosa*	hypothetical protein POTOM_022449 [Populus tomentosa]	KAG6771103
	403.675	98.86%	75.1%	*Populus trichocarpa*	LOW QUALITY PROTEIN: metacaspase-9 [Populus trichocarpa]	XP_002307934
	403.675	98.86%	74.7%	*Populus deltoides*	hypothetical protein H0E87_027381 [Populus deltoides]	KAH8485917
	403.29	98.86%	75.1%	*Populus tomentosa*	hypothetical protein POTOM_024633 [Populus tomentosa]	KAG6769018
	402.905	98.86%	74.7%	*Populus deltoides*	hypothetical protein H0E87_013570 [Populus deltoides]	KAH8506819
	402.519	98.86%	75.5%	*Populus euphratica*	PREDICTED: metacaspase-9-like [Populus euphratica]	XP_011040678
	401.749	98.86%	75.5%	*Corchorus capsularis*	Peptidase C14, caspase catalytic [Corchorus capsularis]	OMO53857
	401.749	98.86%	75.1%	*Xanthoceras sorbifolium*	hypothetical protein JRO89_XS02G0231200 [Xanthoceras sorbifolium]	KAH7575845
	400.208	98.86%	74.3%	*Populus trichocarpa*	metacaspase-9 [Populus trichocarpa]	XP_002322580
	399.823	98.86%	75.1%	*Corchorus olitorius*	Peptidase C14, caspase catalytic [Corchorus olitorius]	OMP08879
	399.823	98.86%	74.7%	*Salix brachista*	hypothetical protein DKX38_008939 [Salix brachista]	KAB5551628
	399.438	98.86%	74.3%	*Populus euphratica*	PREDICTED: metacaspase-9-like [Populus euphratica]	XP_011041315
	397.897	98.86%	73.2%	*Populus tomentosa*	hypothetical protein POTOM_052314 [Populus tomentosa]	KAG6743615
	397.512	98.86%	73.6%	*Populus alba*	metacaspase-9-like [Populus alba]	XP_034893917
	397.512	98.86%	73.6%	*Populus tomentosa*	hypothetical protein POTOM_051245 [Populus tomentosa]	KAG6744609
	397.512	98.86%	73.9%	*Salix dunnii*	hypothetical protein SADUNF_Sadunf06G0018100 [Salix dunnii]	KAF9679468
	396.356	98.86%	73.6%	*Populus alba*	latex-abundant family protein [Populus alba]	TKR97962
	395.971	98.86%	72.6%	*Manihot esculenta*	metacaspase-9 [Manihot esculenta]	XP_021595595
	394.045	98.86%	74.8%	*Gossypium davidsonii*	hypothetical protein [Gossypium davidsonii]	MBA0638256
RpMC9b	343.584	99.55%	75.5%	*Populus deltoides*	hypothetical protein H0E87_013570 [Populus deltoides]	KAH8506819
	341.658	99.55%	75.5%	*Populus alba*	metacaspase-9-like [Populus alba]	XP_034919959
	341.273	99.55%	74.5%	*Populus deltoides*	hypothetical protein H0E87_027381 [Populus deltoides]	KAH8485917
	340.502	99.55%	75.0%	*Corchorus capsularis*	Peptidase C14, caspase catalytic [Corchorus capsularis]	OMO53857
	340.502	99.55%	75.5%	*Populus tomentosa*	hypothetical protein POTOM_022449 [Populus tomentosa]	KAG6771103
	340.117	99.55%	74.5%	*Populus trichocarpa*	metacaspase-9 [Populus trichocarpa]	XP_002322580
	340.117	99.55%	75.0%	*Populus trichocarpa*	LOW QUALITY PROTEIN: metacaspase-9 [Populus trichocarpa]	XP_002307934
	340.117	99.55%	75.0%	*Xanthoceras sorbifolium*	hypothetical protein JRO89_XS02G0231200 [Xanthoceras sorbifolium]	KAH7575845
	339.732	99.55%	75.0%	*Populus tomentosa*	hypothetical protein POTOM_024633 [Populus tomentosa]	KAG6769018
	339.347	99.55%	75.5%	*Populus euphratica*	PREDICTED: metacaspase-9-like [Populus euphratica]	XP_011040678
	338.576	99.55%	72.8%	*Manihot esculenta*	metacaspase-9 [Manihot esculenta]	XP_021595595
	338.576	99.55%	74.5%	*Populus euphratica*	PREDICTED: metacaspase-9-like [Populus euphratica]	XP_011041315
	338.576	99.55%	74.5%	*Corchorus olitorius*	Peptidase C14, caspase catalytic [Corchorus olitorius]	OMP08879
	338.191	99.55%	73.6%	*Populus alba*	metacaspase-9-like [Populus alba]	XP_034893917
	338.191	99.55%	73.6%	*Populus tomentosa*	hypothetical protein POTOM_051245 [Populus tomentosa]	KAG6744609
	338.191	99.55%	75.5%	*Salix brachista*	hypothetical protein DKX38_008939 [Salix brachista]	KAB5551628
	337.806	99.55%	76.2%	*Nyssa sinensis*	hypothetical protein F0562_016927 [Nyssa sinensis]	KAA8516567
	337.035	99.55%	73.6%	*Populus alba*	latex-abundant family protein [Populus alba]	TKR97962
	336.265	99.55%	73.2%	*Populus tomentosa*	hypothetical protein POTOM_052314 [Populus tomentosa]	KAG6743615
	333.569	99.55%	73.6%	*Salix dunnii*	hypothetical protein SADUNF_Sadunf06G0018100 [Salix dunnii]	KAF9679468

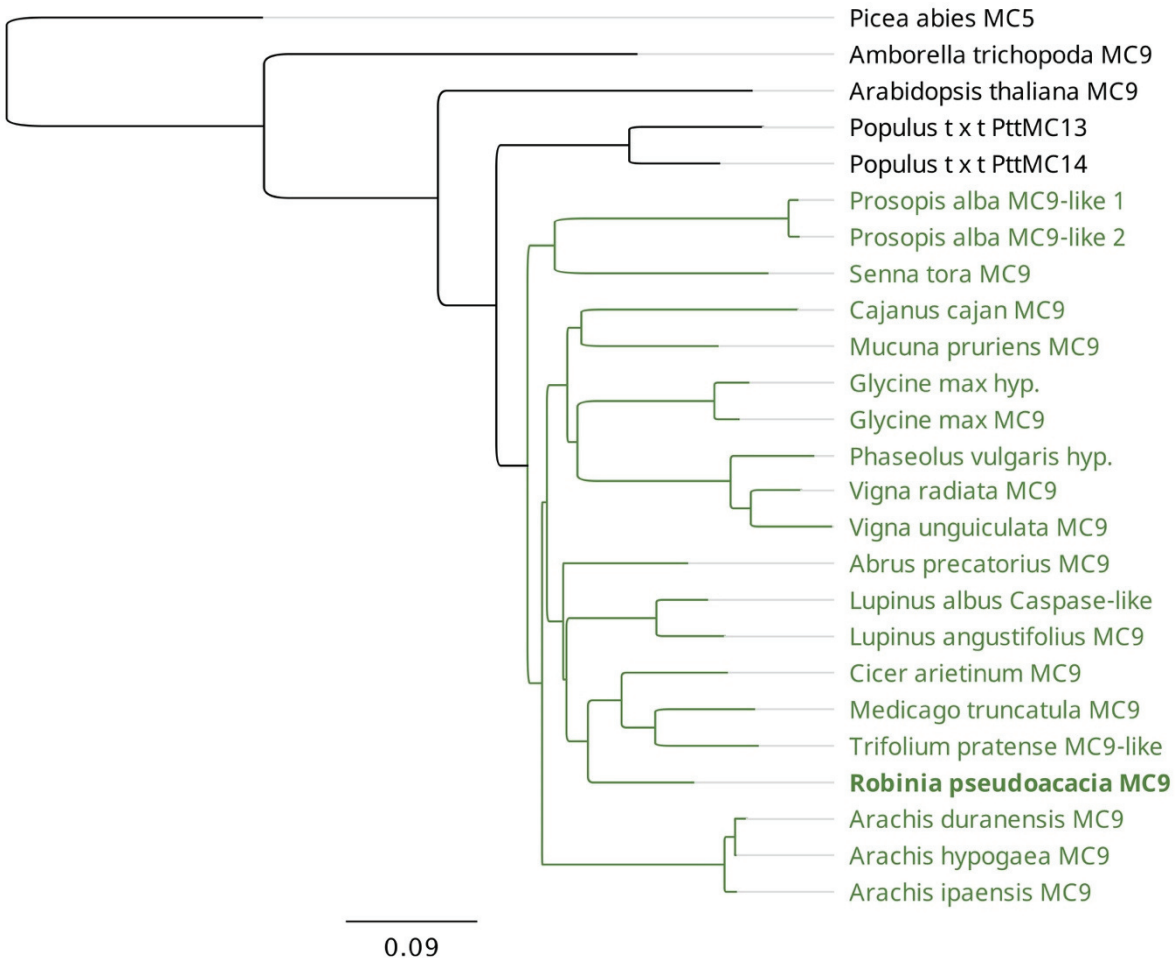


Figure 28: Phylogenetic relationship of MC9 homologs. Green nodes present the species in Fabaceae, with Robinia in bold. Reference sequences included for comparison: MC5 of *P. abies* as the representative for gymnosperm, *A. trichopoda* for basal angiosperm, A. thaliana as the model in dicots, and PttMC13 and PttMC14 in hybrid aspen (*P. tremula x tremuloides*, abbreviated as Populous t x t above) as an example of an angiosperm tree. Information in the sequences used in the gene tree construction is listed in Table 15.

As the protein alignment of the *MC9* homolog shown in Figure 29, the catalytic triad of cysteine protease and the residue at the P1 position for the autocatalytic cleavage in *RpMC9* are in consensus with the others. Among the 13 signature residues conserved among MC9 homologs reported in Fortin and Lam (2018), RpMC9 covers and matches 11 of them. Regarding the conserved aspartates predicted to contribute to calcium binding in the substrates, one of these residues is occupied by a glutamic acid instead in all surveyed legumes, as with *RpMC9*.

Table 15: Information on the sequences used in MC9 protein alignment in Figure 29 and for the construction of the gene tree in Figure 28.

	Sequence label	Description from database [species]	Tissue	Accession
1	Picea abies MC5	MC5 [Picea abies]		AJP06423.1
2	Amborella trichopoda MC9	metacaspase-9	leaf	XP_006832890.1
3	Arabidopsis thaliana MC9	metacaspase-9 (AT5G04200, AtMC9, AtMCP2f, F21E1.120, F21E1_120, MCP2f, metacaspase 2f)		NP_196040.1
4	Populus t x t PttMC13	AtMC9-like [Populous tremula x tremuloides]		Potri.006G026500.1
5	Populus t x t PttMC14	AtMC9-like [Populous tremula x tremuloides]		Potri.016G024500.1
6	Senna tora MC9	metacaspase-9 [Senna tora]	leaf	KAF7819196.1
7	Prosopis alba MC9-like 1	metacaspase-9-like [Prosopis alba]	leaf	XP_028766727.1
8	Prosopis alba MC9-like 2	metacaspase-9-like [Prosopis alba]	leaf	XP_028772742.1
9	Arachis ipaensis MC9	metacaspase-9 [Arachis ipaensis]		XP_016161870.1
10	Arachis duranensis MC9	metacaspase-9 [Arachis duranensis]	whole plant	XP_015970740.2
11	Arachis hypogaea MC9	metacaspase-9 [Arachis hypogaea]	etiolated seedling	XP_025605483.1
12	Abrus precatorius MC9	metacaspase-9 [Abrus precatorius]	young leaves	XP_027351119.1
13	Lupinus albus Caspase-like	putative Caspase-like domain-containing protein [Lupinus albus]	leaves	KAE9599704.1
14	Lupinus angustifolius MC9	PREDICTED: metacaspase-9 [Lupinus angustifolius]	whole plant	XP_019457161.1
15	Robinia pseudoacacia MC9		xylem	
16	Cicer arietinum MC9	metacaspase-9 [Cicer arietinum]	etiolated seedlings	XP_004493282.1
17	Medicago truncatula MC9	metacaspase-9 [Medicago truncatula]	leaves	XP_013449529.1
18	Trifolium pratense MC9-like	metacaspase-9-like protein [Trifolium pratense]	young leaves	PNY07712.1
19	Cajanus cajan MC9	Metacaspase-9 [Cajanus cajan]		KYP64030.1
20	Mucuna pruriens MC9	Metacaspase-9 [Mucuna pruriens]	seed	RDY12566.1
21	Glycine max hyp.	hypothetical protein JHK86_053251 [Glycine max]	leaf	KAG4912818.1
22	Glycine max MC9	metacaspase-9 [Glycine max]	callus	XP_003554078.1
23	Phaseolus vulgaris hyp.	hypothetical protein PHAVU_001G101400g [Phaseolus vulgaris]		XP_007161829.1
24	Vigna radiata MC9	metacaspase-9 [Vigna radiata var. radiata]	leaf	XP_014491146.1
25	Vigna unguiculata MC9	metacaspase-9 [Vigna unguiculata]	first leaves	XP_027917272.1

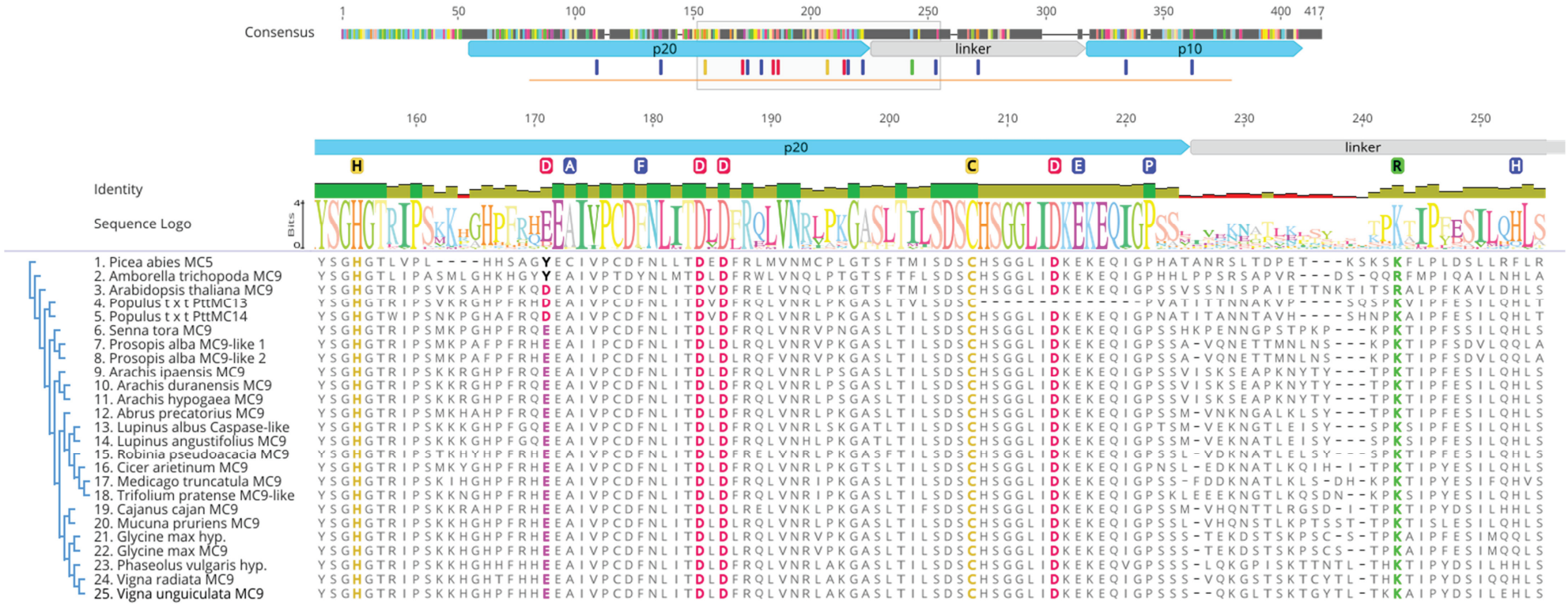


Figure 29: Amino acid sequence alignment of MC9 homologs. Overview of the consensus sequence in the complete gene length is displayed on top, showing the positions of the key sites and coverage of *Rp*MC9 as an orange line underneath. The shaded area corresponds to the alignment below. Light blue elements indicate the catalytic domain, and grey elements the linker region. The order of sequences in the alignment is sorted according to the phylogenetic tree on the left, HMM logo is displayed on the top of the alignment. Annotations: (blue) signature residues of MC9 homologs according to Fortin and Lam (2018); (orange) conserved catalytic triad of cysteine protease; (green) a lysine or an arginine or at the P1 position for autocatalytic cleavage (Vercammen *et al.,* 2004); (red) the conserved aspartates predicted to contribute to calcium binding. The exception (magenta) is the first aspartate at base 171, which is substituted by a Glutamic acid (E) in legumes.

The three RpMC-II sequences cover parts of the p20 domain and the linker but not the p10 subunit (Figure 30). BLASTp results limited to legume species are listed in Table 16, and Table 17 shows the results unlimited to taxa. Each RpMC-II shares a similar pairwise identity with both MC4 and MC5 homologs; Figure 31 shows the phylogeny of these two genes within Fabaceae.

In the protein alignment shown in Figure 32 at base 149, the conserved D is substituted by a glutamic acid (E) in the majority of legumes, and *RpMC-IIc* allows the trend. In *RpMC-IIa,* it is substituted by alanine. One of the heterogeneous sites of *RpMC-IIb* resulted in a change of amino acid residue (Figure 25 in page 77) at this base, which is neither an E nor an A.

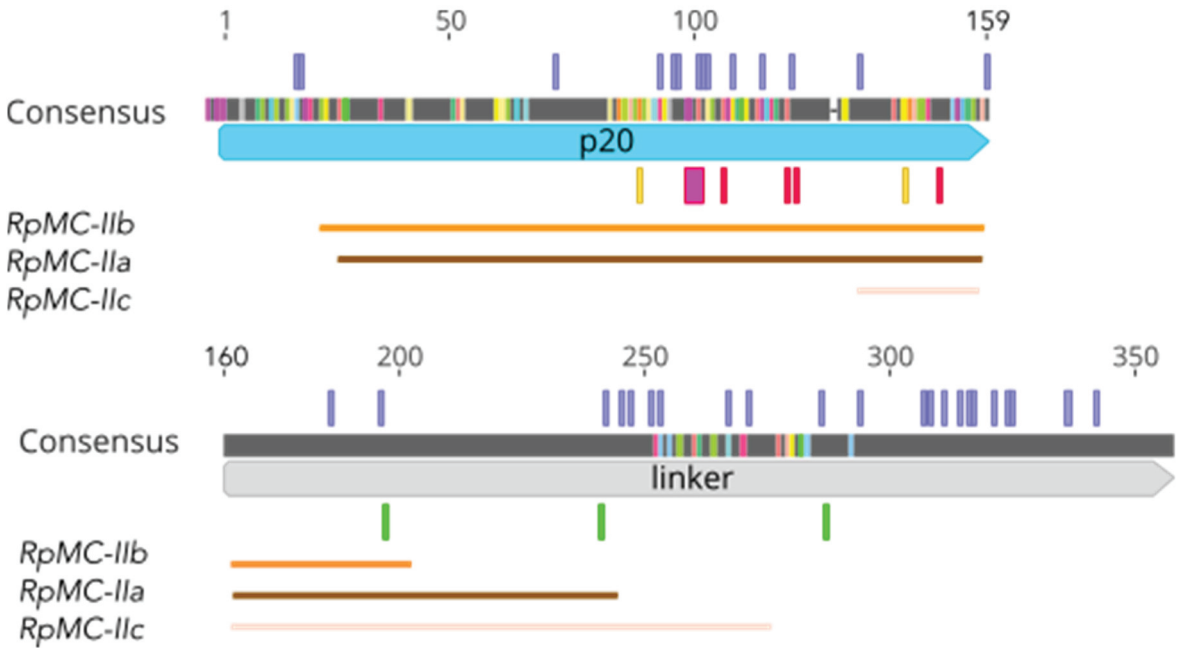


Figure 30: Overview of the amino acid sequence alignment of MC4 and MC5 homologs, showing the coverage of the positions of key sites and coverage of *RpMC-IIa* (brown line), *RpMC-IIb* (orange line), and *RpMC-IIc* (beige line). Light blue elements correspond to the p20 catalytic domain, and grey elements to the linker region. Annotations: (purple) signature residues among MC4 homologs (Fortin & Lam, 2018); (yellow) conserved catalytic residues of cysteine protease; (green) sites for autocatalytic cleavage (Vercammen *et al.,* 2004); (red) the conserved aspartates predicted to contribute to calcium binding, (magenta) negatively charged residues for calcium ions dependent activation.

Table 16: Top hits from BLASTp taking the three RpMC-II as queries. The search was limited to the pea family.

Query	Bit-Score	Query coverage	Pairwise Identity	Organism	Description	Accession
RpMC-IIa	316235	100.00%	82.4%	Vigna radiata	Metacaspase-5 [Vigna radiata var. radiata]	XP_014490230
	314.694	99.46%	83.3%	Spatholobus suberectus	Metacaspase-4 protein [Spatholobus suberectus]	TKY68542
	310.457	100.00%	79.0%	Cicer arietinum	Metacaspase-4-like [Cicer arietinum]	XP_004510001
	302.753	100.00%	79.0%	Spatholobus suberectus	Metacaspase-5 protein [Spatholobus suberectus]	TKY68543
	298.901	100.00%	78.5%	Vigna umbellata	Metacaspase-5-like [Vigna umbellata]	XP_047168269
	297.745	100.00%	77.4%	Phaseolus vulgaris	Hypothetical protein PHAVU_011G180200g [Phaseolus vulgaris]	XP_007133456
	297.745	100.00%	80.6%	Mucuna pruriens	Metacaspase-4 [Mucuna pruriens]	RDX65940
	295.819	97.85%	78.6%	Medicago truncatula	Metacaspase-5 [Medicago truncatula]	XP_013445304
	295.434	100.00%	77.4%	Vigna angularis	PREDICTED: metacaspase-5-like [Vigna angularis]	XP_017432984
	294.664	97.85%	79.1%	Trifolium pratense	Metacaspase-5-like [Trifolium pratense]	XP_045800411
	294.664	97.85%	79.1%	Trifolium pratense	Metacaspase type ii [Trifolium pratense]	PNY03882
	294.664	100.00%	76.9%	Vigna angularis	Metacaspase-5 protein [Vigna angularis]	KAG2396873
	293.123	100.00%	76.9%	Vigna unguiculata	Metacaspase-5-like [Vigna unguiculata]	XP_027911902
	292.738	100.00%	76.9%	Vigna radiata	Metacaspase-5 [Vigna radiata var. radiata]	XP_014490212
	291.967	100.00%	76.9%	Cajanus cajan	Metacaspase-5 [Cajanus cajan]	XP_020203863
	289.271	100.00%	80.7%	Abrus precatorius	Metacaspase-5-like isoform X1 [Abrus precatorius]	XP_027354993
	288.886	100.00%	82.0%	Vigna angularis	PREDICTED: metacaspase-5-like [Vigna angularis]	XP_017431325
	287.345	100.00%	81.8%	Phaseolus vulgaris	Hypothetical protein PHAVU_011G180300g [Phaseolus vulgaris]	XP_007133458
	284.263	98.92%	82.3%	Glycine max	Metacaspase-4 [Glycine max]	XP_003546660
	283.878	100.00%	80.9%	Cajanus cajan	Metacaspase-4 [Cajanus cajan]	KYP38529
RpMC-IIb	257299	100.00%	79.9%	Senna tora	Metacaspase-4 subunit p10 like [Senna tora]	KAF7822468
	252.677	100.00%	79.3%	Vigna angularis	Metacaspase-5 protein [Vigna angularis]	KAG2396873
	243.817	86.36%	88.0%	Glycine soja	Hypothetical protein JHK87_043149 [Glycine soja]	KAG4947142
	249.595	100.00%	79.3%	Trifolium pratense	Metacaspase type ii [Trifolium pratense]	PNY03882
	261.922	100.00%	83.8%	Spatholobus suberectus	Metacaspase-4 protein [Spatholobus suberectus]	TKY68542
	262.692	100.00%	83.1%	Spatholobus suberectus	Metacaspase-5 protein [Spatholobus suberectus]	TKY68543
	246.514	86.36%	88.8%	Glycine max	Metacaspase-4 [Glycine max]	XP_003546660
	254.218	100.00%	79.9%	Cicer arietinum	Metacaspase-4-like [Cicer arietinum]	XP_004510001
	254.988	100.00%	80.5%	Phaseolus vulgaris	Hypothetical protein PHAVU_011G180200g [Phaseolus vulgaris]	XP_007133456
	259.61	100.00%	82.5%	Phaseolus vulgaris	Hypothetical protein PHAVU_011G180300g [Phaseolus vulgaris]	XP_007133458
	254.603	100.00%	80.5%	Medicago truncatula	Metacaspase-5 [Medicago truncatula]	XP_013445304
	252.292	100.00%	79.3%	Vigna radiata	Metacaspase-5 [Vigna radiata var. radiata]	XP_014490212
	263.077	100.00%	83.1%	Vigna radiata	Metacaspase-5 [Vigna radiata var. radiata]	XP_014490230
	253.832	100.00%	79.9%	Vigna angularis	PREDICTED: metacaspase-5-like [Vigna angularis]	XP_017432984
	243.432	100.00%	78.6%	Cajanus cajan	Metacaspase-5 [Cajanus cajan]	XP_020203863
	261.151	100.00%	83.8%	Abrus precatorius	Metacaspase-5-like isoform X1 [Abrus precatorius]	XP_027354993
	252.677	100.00%	79.9%	Vigna unguiculata	Metacaspase-5-like [Vigna unguiculata]	XP_027911902
	241.121	100.00%	77.3%	Prosopis alba	Metacaspase-5-like [Prosopis alba]	XP_028799665
	249.595	100.00%	79.3%	Trifolium pratense	Metacaspase-5-like [Trifolium pratense]	XP_045800411
	254.218	100.00%	80.5%	Vigna umbellata	Metacaspase-5-like [Vigna umbellata]	XP_047168269
RpMC-IIc	170629	98.31%	73.3%	Vigna angularis	Metacaspase-5 protein [Vigna angularis]	KAG2396873
	171.4	100.00%	76.3%	Trifolium medium	Metacaspase-4-like [Trifolium medium]	MCH94412
	171.4	100.00%	76.3%	Trifolium pratense	Metacaspase type ii [Trifolium pratense]	PNY03882
	170.629	98.31%	74.1%	Mucuna pruriens	Metacaspase-4 [Mucuna pruriens]	RDX65940
	167.933	100.00%	77.3%	Mucuna pruriens	Metacaspase-4 [Mucuna pruriens]	RDX65940
	185.267	100.00%	79.0%	Spatholobus suberectus	Metacaspase-4 protein [Spatholobus suberectus]	TKY68542
	170.244	100.00%	72.9%	Spatholobus suberectus	Metacaspase-5 protein [Spatholobus suberectus]	TKY68543
	190.66	100.00%	78.0%	Cicer arietinum	Metacaspase-4-like [Cicer arietinum]	XP_004510001
	165.622	98.31%	71.6%	Phaseolus vulgaris	Hypothetical protein PHAVU_011G180200g [Phaseolus vulgaris]	XP_007133456
	168.318	100.00%	80.7%	Phaseolus vulgaris	Hypothetical protein PHAVU_011G180300g [Phaseolus vulgaris]	XP_007133458
	167.162	100.00%	74.6%	Medicago truncatula	Metacaspase-5 [Medicago truncatula]	XP_013445304
	166.392	98.31%	71.6%	Vigna radiata	Metacaspase-5 [Vigna radiata var. radiata]	XP_014490212
	192.971	100.00%	82.4%	Vigna radiata	Metacaspase-5 [Vigna radiata var. radiata]	XP_014490230
	163.696	100.00%	75.0%	Vigna angularis	PREDICTED: metacaspase-5-like [Vigna angularis]	XP_017431325
	170.629	98.31%	73.3%	Vigna angularis	PREDICTED: metacaspase-5-like [Vigna angularis]	XP_017432984
	171.785	100.00%	72.9%	Cajanus cajan	Metacaspase-5 [Cajanus cajan]	XP_020203863
	168.703	98.31%	72.4%	Vigna unguiculata	Metacaspase-5-like [Vigna unguiculata]	XP_027911902
	171.4	100.00%	76.3%	Trifolium pratense	Metacaspase-5-like [Trifolium pratense]	XP_045800411
	164.081	100.00%	75.0%	Vigna umbellata	Metacaspase-5-like [Vigna umbellata]	XP_047168265
	170.629	98.31%	73.3%	Vigna umbellata	Metacaspase-5-like [Vigna umbellata]	XP_047168269

Table 17: Top hits from BLASTp taking the three RpMC-II as queries. The search was not limited by taxa.

Query	Bit-Score	Query coverage	Pairwise Identity	Organism	Description	Accession
RpMC-IIa	280796	100.00%	74.2%	Rhodamnia argentea	Metacaspase-4-like [Rhodamnia argentea]	XP_030524088
	276.174	100.00%	71.1%	Hevea brasiliensis	Type II metacaspase [Hevea brasiliensis]	ADM52185
	275.404	99.46%	73.1%	Theobroma cacao	PREDICTED: metacaspase-4 [Theobroma cacao]	XP_007022751
	274.633	100.00%	71.0%	Xanthoceras sorbifolium	Hypothetical protein JRO89_XS08G0024700 [Xanthoceras sorbifolium]	KAH7565842
	274.248	100.00%	70.6%	Hevea brasiliensis	Metacaspase-4-like [Hevea brasiliensis]	XP_021692568
	273.092	99.46%	72.6%	Herrania umbratica	Metacaspase-5 [Herrania umbratica]	XP_021299552
	271.166	100.00%	71.5%	Citrus unshiu	Hypothetical protein CUMW_067430 [Citrus unshiu]	GAY42508
	270.396	100.00%	71.5%	Citrus sinensis	Metacaspase-4 [Citrus sinensis]	XP_006493737
	269.24	100.00%	70.1%	Hevea brasiliensis	Hypothetical protein GH714_007130 [Hevea brasiliensis]	KAF2302770
	268.855	100.00%	70.4%	Tripterygium wilfordii	MC2 [Tripterygium wilfordii]	KAF5733555
	268.47	100.00%	71.0%	Citrus clementina	LOW QUALITY PROTEIN: metacaspase-4 [Citrus clementina]	XP_024036599
	268.47	89.78%	75.4%	Nyssa sinensis	Hypothetical protein F0562_009810 [Nyssa sinensis]	KAA8523386
	266.929	100.00%	70.4%	Juglans microcarpa	Metacaspase-4 [Juglans microcarpa x Juglans regia]	XP_040997649
	266.544	100.00%	74.2%	Carpinus fangiana	Hypothetical protein FH972_017766 [Carpinus fangiana]	KAE8099813
	266.159	97.85%	69.2%	Cucumis sativus	Metacaspase-5 [Cucumis sativus]	XP_004139955
	265.003	87.10%	78.4%	Eucalyptus grandis	Metacaspase-4 [Eucalyptus grandis]	XP_010064092
	265.003	99.46%	70.2%	Vaccinium darrowii	Hypothetical protein Vadar_003635 [Vaccinium darrowii]	KAH7842288
	262.692	100.00%	71.0%	Psidium guajava	Uncharacterized protein J3R85_015726 [Psidium guajava]	KAI3414923
	261.922	86.56%	76.4%	Corymbia citriodora	Hypothetical protein BT93_C2363 [Corymbia citriodora subsp. variegata]	KAF8036613
	261.922	100.00%	70.4%	Camellia sinensis	Hypothetical protein HYC85_014979 [Camellia sinensis]	KAF5949022
RpMC-IIb	262307	99.35%	80.4%	Rhodamnia argentea	Metacaspase-4-like [Rhodamnia argentea]	XP_030524088
	261.536	99.35%	81.1%	Nelumbo nucifera	PREDICTED: metacaspase-4 [Nelumbo nucifera]	XP_010260805
	258.07	100.00%	81.8%	Carpinus fangiana	Hypothetical protein FH972_017766 [Carpinus fangiana]	KAE8099813
	257.299	99.35%	79.8%	Eucalyptus grandis	Metacaspase-4 [Eucalyptus grandis]	XP_010064092
	256.529	99.35%	78.5%	Corymbia citriodora	Hypothetical protein BT93_C2363 [Corymbia citriodora subsp. variegata]	KAF8036613
	255.758	100.00%	78.6%	Hevea brasiliensis	Type II metacaspase [Hevea brasiliensis]	ADM52185
	254.603	99.35%	77.8%	Psidium guajava	Uncharacterized protein J3R85_015726 [Psidium guajava]	KAI3414923
	254.603	98.70%	78.3%	Xanthoceras sorbifolium	Hypothetical protein JRO89_XS08G0024700 [Xanthoceras sorbifolium]	KAH7565842
	254.218	100.00%	79.3%	Theobroma cacao	PREDICTED: metacaspase-4 [Theobroma cacao]	XP_007022751
	253.832	100.00%	79.9%	Telopea speciosissima	Metacaspase-4-like [Telopea speciosissima]	XP_043688938
	253.447	100.00%	78.0%	Hevea brasiliensis	Metacaspase-4-like [Hevea brasiliensis]	XP_021692568
	251.521	100.00%	78.6%	Herrania umbratica	Metacaspase-5 [Herrania umbratica]	XP_021299552
	251.521	100.00%	75.4%	Asparagus officinalis	Metacaspase-4-like [Asparagus officinalis]	XP_020256681
	251.136	100.00%	76.7%	Macadamia integrifolia	Metacaspase-4 [Macadamia integrifolia]	XP_042520914
	251.136	100.00%	78.0%	Hevea brasiliensis	Hypothetical protein GH714_007130 [Hevea brasiliensis]	KAF2302770
	250.366	98.70%	78.3%	Juglans microcarpa	Metacaspase-4 [Juglans microcarpa x Juglans regia]	XP_040997649
	249.595	100.00%	74.7%	Macleaya cordata	Peptidase C14 [Macleaya cordata]	OVA03142
	249.595	100.00%	78.6%	Aristolochia fimbriata	Hypothetical protein H6P81_019466 [Aristolochia fimbriata]	KAG9439301
	248.825	100.00%	76.8%	Vaccinium darrowii	Hypothetical protein Vadar_003635 [Vaccinium darrowii]	KAH7842288
	247.669	100.00%	76.7%	Nyssa sinensis	Hypothetical protein F0562_009810 [Nyssa sinensis]	KAA8523386
RpMC-IIc	164851	100.00%	70.3%	Tripterygium wilfordii	Metacaspase-4-like [Tripterygium wilfordii]	XP_038690069
	164.851	100.00%	70.3%	Tripterygium wilfordii	Metacaspase-4-like [Tripterygium wilfordii]	XP_038689229
	159.844	100.00%	67.2%	Hevea brasiliensis	Metacaspase-4-like [Hevea brasiliensis]	XP_021692568
	159.844	100.00%	67.2%	Hevea brasiliensis	Type II metacaspase [Hevea brasiliensis]	ADM52185
	159.458	100.00%	69.5%	Rhodamnia argentea	Metacaspase-4-like [Rhodamnia argentea]	XP_030524088
	157.147	100.00%	67.2%	Theobroma cacao	PREDICTED: metacaspase-4 [Theobroma cacao]	XP_007022751
	155.992	100.00%	66.4%	Hevea brasiliensis	Hypothetical protein GH714_007130 [Hevea brasiliensis]	KAF2302770
	155.606	100.00%	66.4%	Herrania umbratica	Metacaspase-5 [Herrania umbratica]	XP_021299552
	155.606	100.00%	66.9%	Tripterygium wilfordii	MC2 [Tripterygium wilfordii]	KAF5733555
	154.451	100.00%	66.4%	Durio zibethinus	Metacaspase-4-like [Durio zibethinus]	XP_022734348
	154.451	100.00%	64.4%	Xanthoceras sorbifolium	Hypothetical protein JRO89_XS08G0024700 [Xanthoceras sorbifolium]	KAH7565842
	152.91	100.00%	64.4%	Citrus sinensis	Metacaspase-4 [Citrus sinensis]	XP_006493737
	152.14	100.00%	66.1%	Cephalotus follicularis	Peptidase_C14 domain-containing protein [Cephalotus follicularis]	GAV64892
	150.599	100.00%	65.4%	Impatiens glandulifera	Metacaspase-4-like [Impatiens glandulifera]	XP_047317572
	150.599	100.00%	63.6%	Citrus clementina	LOW QUALITY PROTEIN: metacaspase-4 [Citrus clementina]	XP_024036599
	150.214	100.00%	63.6%	Citrus unshiu	Hypothetical protein CUMW_067430 [Citrus unshiu]	GAY42508
	148.673	100.00%	64.2%	Mangifera indica	Metacaspase-5-like [Mangifera indica]	XP_044468509
	148.288	100.00%	67.8%	Psidium guajava	Uncharacterized protein J3R85_015726 [Psidium guajava]	KAI3414923
	147.902	100.00%	62.2%	Momordica charantia	Metacaspase-4-like [Momordica charantia]	XP_022147831
	147.902	100.00%	68.6%	Prunus persica	Metacaspase-4 [Prunus persica]	XP_007211951

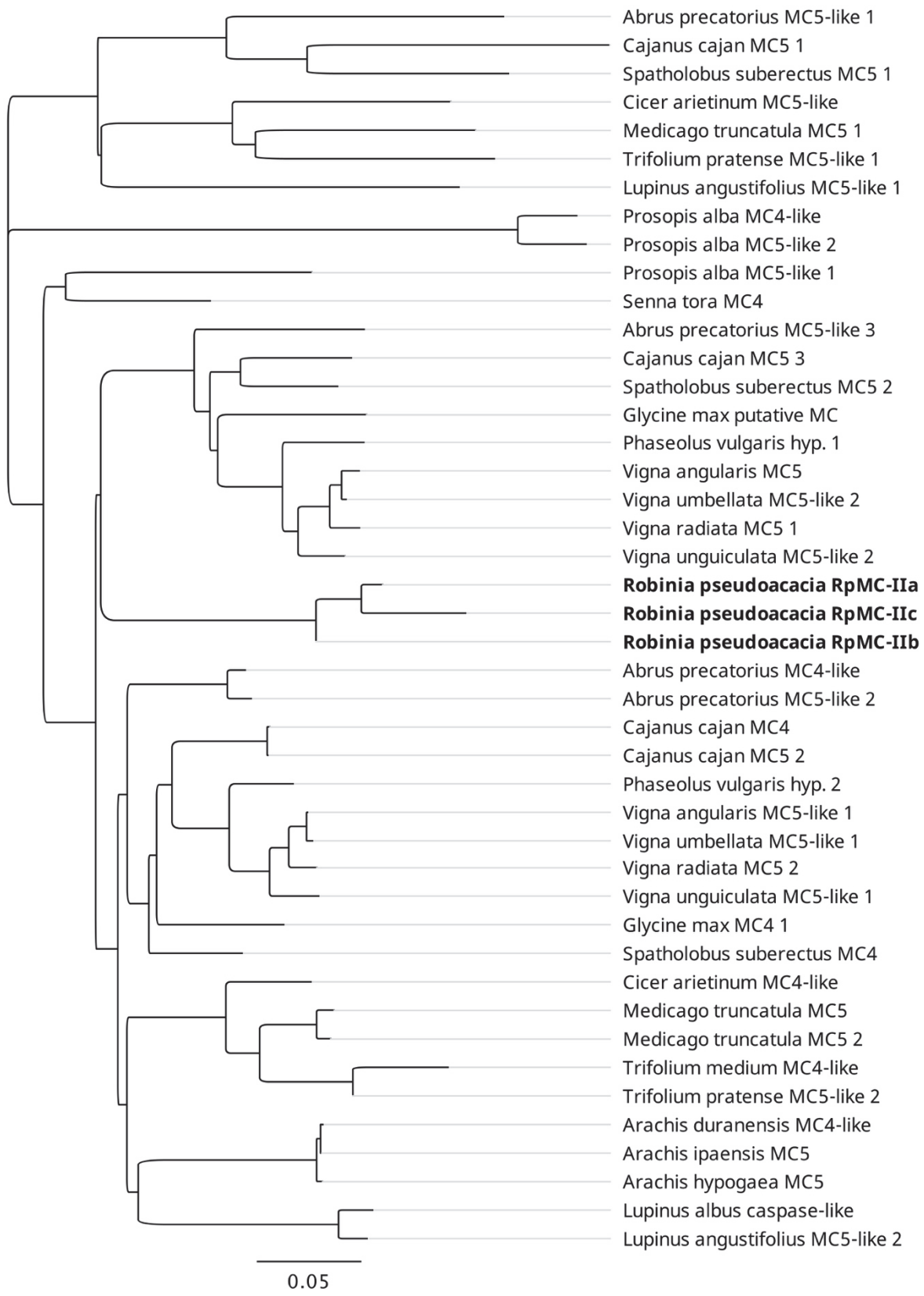


Figure 31: Phylogenic tree of metacaspase 4 and 5 (MC4 and MC5) homologs in Fabacae species. Sequences from *R. pseudoacacia* are bolded.

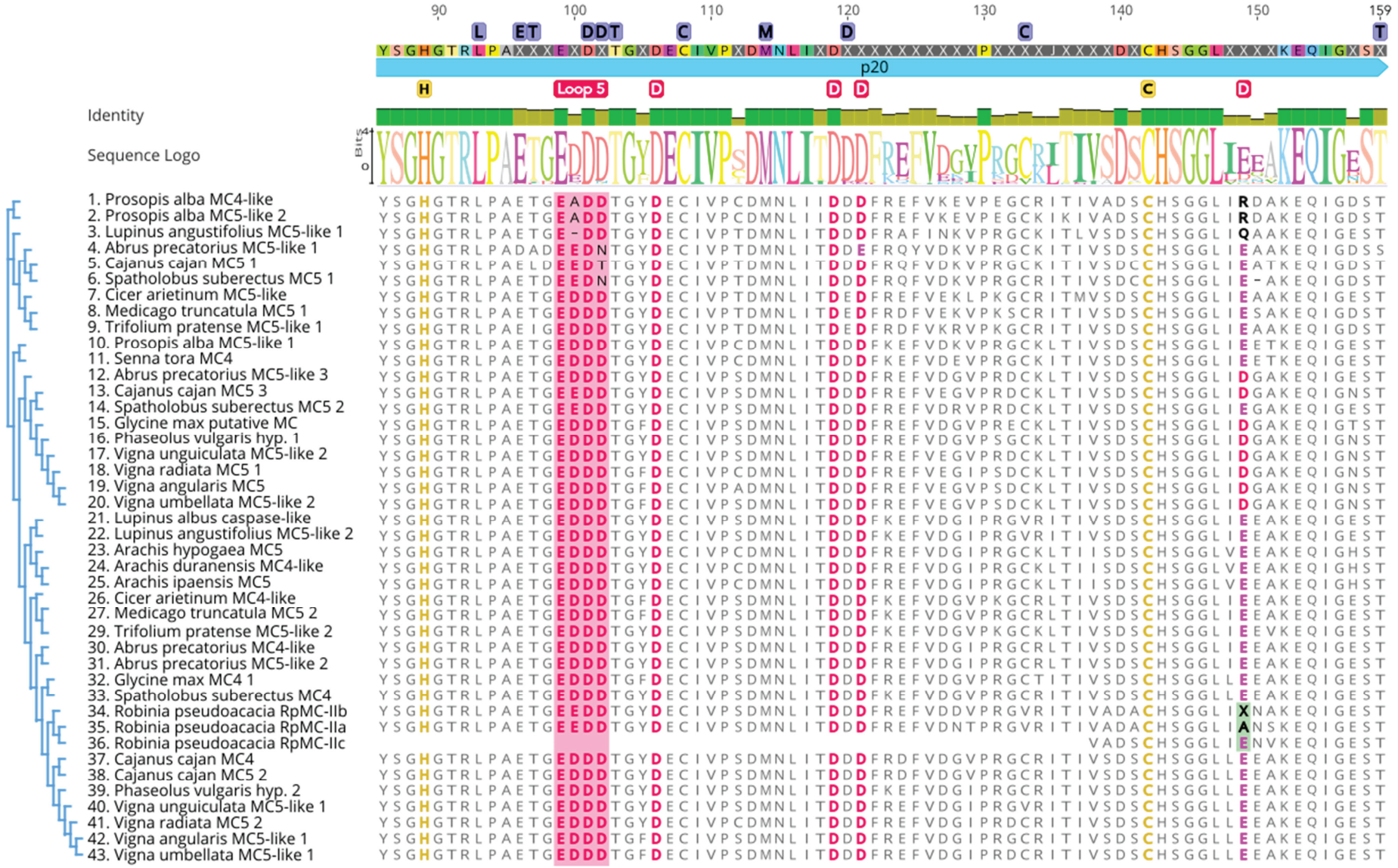


Figure 32: Amino acid sequence alignment of metacaspase 4 and 5 (MC4 and MC5) homologs in Fabales. The order of sequences in the alignment is sorted according to the phylogenetic tree on the left, and HMM logo is displayed above the alignment. Light blue elements mark the catalytic domain p20. The consensus sequence is displayed on the top with annotation on the key sites: (purple) signature residues among MC4 homologs; (yellow) conserved catalytic residues of cysteine protease; (red) the conserved aspartates (D) predicted to contribute to calcium binding. Bold text within the alignment: (red text) conserved residues for calcium binding; Shaded boxes: (magenta) Loop 5 regions with negatively charged residues for Ca2+-dependent activation and substrate processing (Zhu *et al.,* 2020). At base 149, the conserved D is substituted by a glutamic acid (E) in the majority of sequences, (green) shaded box highlights the bases in RpMC-II.

4.3.2 VPEs protein sequences

With the two translated sequences as queries, the BLASTp results are listed in Table 18 for those within the legumes family and species outside of the family in Table 19. The selection of protein sequences for the alignment was based on BLASTp results with RpVPEs and the supplementary data from Yamada *et al.* (2019), Du *et al.* (2020), and DePamphilis *et al.* (2013). Figure 35 demonstrates their phylogenetic relationship. Contrary to the phylogeny of type II MCs, the two distinct types of VPEs are closely related to their homologs in other species.

Specific regions in the protein sequences of VPEs have been identified as the determinant for their enzymatic property as ligase or protease (Du *et al.*, 2020), such as the ligase-activity determinant (LAD) gatekeepers, poly-proline loop (PPL), and marker of ligase activity (MLA). As shown in the amino acid sequence alignments (Figure 33 for βVPE and Figure 34 for γVPE), the residues in these determinant regions in β-RpVPE and γ-RpVPE are identical to the signatures of proteases.

Table 18: Top results from BLASTp in legumes, taking the two RpVPE as query.

Query	Bit-Score	Query coverage	Pairwise Identity	Organism	Description	Accession
γ-RpVPE	743.806	99.76%	83.5%	*Arachis diogoi*	Vacuolar processing enzyme [Arachis diogoi]	AEO79971
	743.806	99.76%	83.0%	*Lupinus albus*	Putative legumain protein [Lupinus albus]	KAE9609334
	775.393	99.76%	89.3%	*Glycine soja*	Hypothetical protein JHK87_048261 [Glycine soja]	KAG4934259
	776.548	99.76%	89.6%	*Glycine soja*	Hypothetical protein JHK87_039258 [Glycine soja]	KAG4953664
	776.548	99.76%	89.6%	*Glycine max*	Hypothetical protein JHK86_039459 [Glycine max]	KAG4962591
	776.933	99.76%	89.6%	*Glycine soja*	Vacuolar-processing enzyme [Glycine soja]	KHN37015
	775.778	99.76%	89.6%	*Glycine max*	Vacuolar processing enzyme 2 precursor [Glycine max]	NP_001236564
	761.911	99.76%	86.7%	*Mucuna pruriens*	Hypothetical protein CR513_60678, partial [Mucuna pruriens]	RDX61115
	782.326	99.76%	90.3%	*Spatholobus suberectus*	Vacuolar-processing enzyme [Spatholobus suberectus]	TKY57951
	742.265	99.76%	84.2%	*Glycine max*	Vacuolar-processing enzyme [Glycine max]	XP_003525979
	773.467	99.76%	89.3%	*Glycine max*	Vacuolar-processing enzyme [Glycine max]	XP_003550283
	744.962	99.76%	83.3%	*Cicer arietinum*	Vacuolar-processing enzyme-like [Cicer arietinum]	XP_004501477
	744.577	99.76%	83.5%	*Arachis duranensis*	Vacuolar-processing enzyme [Arachis duranensis]	XP_015972214
	742.265	99.76%	83.3%	*Arachis ipaensis*	Vacuolar-processing enzyme isoform X1 [Arachis ipaensis]	XP_016162949
	745.347	99.76%	83.0%	*Lupinus angustifolius*	PREDICTED: vacuolar-processing enzyme-like [Lupinus angustifolius]	XP_019416625
	766.533	99.76%	87.0%	*Cajanus cajan*	Vacuolar-processing enzyme [Cajanus cajan]	XP_020222250
	774.622	99.76%	89.1%	*Abrus precatorius*	Vacuolar-processing enzyme-like isoform X1 [Abrus precatorius]	XP_027368993
	742.651	95.64%	88.9%	*Abrus precatorius*	vacuolar-processing enzyme-like isoform X2 [Abrus precatorius]	XP_027368995
	744.191	99.76%	84.5%	*Vigna unguiculata*	Vacuolar-processing enzyme [Vigna unguiculata]	XP_027940765
	750.74	99.76%	84.7%	*Prosopis alba*	LOW QUALITY PROTEIN: vacuolar-processing enzyme [Prosopis alba]	XP_028796063
β-RpVPE	683.715	100.00%	86.9%	*Canavalia ensiformis*	Chain A, Legumain [Canavalia ensiformis]	6XT5_A
	700.664	100.00%	88.2%	*Glycine max*	Seed maturation protein PM40 [Glycine max]	AAF89646
	704.901	100.00%	88.5%	*Lupinus albus*	Legumain [Lupinus albus]	KAE9614175
	711.449	100.00%	89.3%	*Glycine max*	Hypothetical protein JHK85_012219 [Glycine max]	KAG5039743
	690.263	100.00%	83.9%	*Glycine max*	Hypothetical protein JHK82_047448 [Glycine max]	KAG5097594
	701.049	100.00%	88.2%	*Glycine max*	Vacuolar-processing enzyme precursor [Glycine max]	NP_001236678
	709.909	100.00%	89.3%	*Glycine max*	Seed maturation protein PM40 precursor [Glycine max]	NP_001336126
	699.893	100.00%	87.9%	*Glycine max*	RecName: Full=Vacuolar-processing enzyme; Short=VPE; Flags: Precursor [Glycine max]	P49045
	686.411	100.00%	87.1%	*Canavalia ensiformis*	RecName: Full=Legumain; AltName: Full=Asparaginyl endopeptidase; Flags: Precursor [Canavalia ensiformis]	P49046
	721.465	100.00%	91.4%	*Spatholobus suberectus*	Vacuolar-processing enzyme [Spatholobus suberectus]	TKY55414
	691.419	100.00%	86.1%	*Medicago truncatula*	Legumain [Medicago truncatula]	XP_003608766
	695.656	100.00%	87.4%	*Cicer arietinum*	Legumain [Cicer arietinum]	XP_004515742
	693.345	100.00%	86.9%	*Vigna radiata*	Vacuolar-processing enzyme [Vigna radiata var. radiata]	XP_014510094
	677.937	100.00%	86.1%	*Arachis duranensis*	Legumain [Arachis duranensis]	XP_015963214
	680.633	100.00%	86.6%	*Arachis ipaensis*	Legumain [Arachis ipaensis]	XP_016188950
	711.835	100.00%	90.6%	*Lupinus angustifolius*	PREDICTED: legumain [Lupinus angustifolius]	XP_019446632
	702.975	100.00%	88.5%	*Cajanus cajan*	Vacuolar-processing enzyme [Cajanus cajan]	XP_020226377
	715.301	100.00%	89.8%	*Abrus precatorius*	Legumain [Abrus precatorius]	XP_027335356
	706.442	100.00%	88.2%	*Vigna unguiculata*	Vacuolar-processing enzyme [Vigna unguiculata]	XP_027918860
	701.049	100.00%	88.2%	*Glycine soja*	Vacuolar-processing enzyme [Glycine soja]	XP_028211613

Table 19: Top hit from BLASTp outside of the Fabales family, taking the two RpVPE as queries.

Query	Bit-Score	Query coverage	Pairwise Identity	Organism	Description	Accession
γ-RpVPE	645.58	100.00%	81.8%	*Prunus persica*	Hypothetical protein PRUPE_5G076300 [Prunus persica]	ONI06720
	644.81	100.00%	81.2%	*Manihot esculenta*	Vacuolar-processing enzyme isoform X2 [Manihot esculenta]	XP_043812118
	659062	100.00%	83.1%	*Prunus armeniaca*	Hypothetical protein GBA52_015747 [Prunus armeniaca]	KAH0973848
	658292	100.00%	81.5%	*Quercus lobata*	Legumain isoform X1 [Quercus lobata]	XP_030945422
	657522	100.00%	81.5%	*Quercus lobata*	Legumain isoform X2 [Quercus lobata]	XP_030945423
	654825	99.73%	82.8%	*Prunus armeniaca*	Unnamed protein product [Prunus armeniaca]	CAB4310086
	652514	100.00%	82.6%	*Prunus dulcis*	Legumain [Prunus dulcis]	XP_034216686
	651358	100.00%	82.6%	*Prunus dulcis*	PREDICTED: vacuolar-processing enzyme [Prunus dulcis]	VVA31981
	650973	100.00%	81.2%	*Quercus suber*	Legumain [Quercus suber]	XP_023911206
	650588	100.00%	82.3%	*Prunus armeniaca*	Unnamed protein product [Prunus armeniaca]	CAB4279623
	648662	100.00%	83.1%	*Prunus avium*	Legumain isoform X1 [Prunus avium]	XP_021810342
	647892	100.00%	80.7%	*Morus notabilis*	Legumain [Morus notabilis]	XP_010093667
	647506	100.00%	80.6%	*Carpinus fangiana*	Hypothetical protein FH972_012780 [Carpinus fangiana]	KAE8055977
	646736	100.00%	81.8%	*Prunus persica*	Legumain [Prunus persica]	XP_007209951
	646351	100.00%	79.6%	*Hevea brasiliensis*	Vacuolar-processing enzyme [Hevea brasiliensis]	XP_021641659
	645966	100.00%	79.6%	*Hevea brasiliensis*	Hypothetical protein GH714_029724 [Hevea brasiliensis]	KAF2299009
	645195	100.00%	79.9%	*Ricinus communis*	Vacuolar-processing enzyme precursor [Ricinus communis]	NP_001310660
	643654	100.00%	81.2%	*Manihot esculenta*	Vacuolar-processing enzyme isoform X1 [Manihot esculenta]	XP_021609203
	643269	100.00%	79.9%	*Jatropha curcas*	Hypothetical protein JCGZ_20778 [Jatropha curcas]	KDP25622
	642884	100.00%	79.1%	*Populus euphratica*	PREDICTED: vacuolar-processing enzyme [Populus euphratica]	XP_011018251
β-RpVPE	645.58	100.00%	81.8%	*Prunus persica*	Hypothetical protein PRUPE_5G076300 [Prunus persica]	ONI06720
	644.81	100.00%	81.2%	*Manihot esculenta*	Vacuolar-processing enzyme isoform X2 [Manihot esculenta]	XP_043812118
	659.062	100.00%	83.1%	*Prunus armeniaca*	Hypothetical protein GBA52_015747 [Prunus armeniaca]	KAH0973848
	658.292	100.00%	81.5%	*Quercus lobata*	Legumain isoform X1 [Quercus lobata]	XP_030945422
	657.522	100.00%	81.5%	*Quercus lobata*	Legumain isoform X2 [Quercus lobata]	XP_030945423
	654.825	99.73%	82.8%	*Prunus armeniaca*	Unnamed protein product [Prunus armeniaca]	CAB4310086
	652.514	100.00%	82.6%	*Prunus dulcis*	Legumain [Prunus dulcis]	XP_034216686
	651.358	100.00%	82.6%	*Prunus dulcis*	PREDICTED: vacuolar-processing enzyme [Prunus dulcis]	VVA31981
	650.973	100.00%	81.2%	*Quercus suber*	Legumain [Quercus suber]	XP_023911206
	650.588	100.00%	82.3%	*Prunus armeniaca*	Unnamed protein product [Prunus armeniaca]	CAB4279623
	648.662	100.00%	83.1%	*Prunus avium*	Legumain isoform X1 [Prunus avium]	XP_021810342
	647.892	100.00%	80.7%	*Morus notabilis*	Legumain [Morus notabilis]	XP_010093667
	647.506	100.00%	80.6%	*Carpinus fangiana*	Hypothetical protein FH972_012780 [Carpinus fangiana]	KAE8055977
	646.736	100.00%	81.8%	*Prunus persica*	Legumain [Prunus persica]	XP_007209951
	646.351	100.00%	79.6%	*Hevea brasiliensis*	Vacuolar-processing enzyme [Hevea brasiliensis]	XP_021641659
	645.966	100.00%	79.6%	*Hevea brasiliensis*	Hypothetical protein GH714_029724 [Hevea brasiliensis]	KAF2299009
	645.195	100.00%	79.9%	*Ricinus communis*	Vacuolar-processing enzyme precursor [Ricinus communis]	NP_001310660
	643.654	100.00%	81.2%	*Manihot esculenta*	Vacuolar-processing enzyme isoform X1 [Manihot esculenta]	XP_021609203
	643.269	100.00%	79.9%	*Jatropha curcas*	Hypothetical protein JCGZ_20778 [Jatropha curcas]	KDP25622
	642.884	100.00%	79.1%	*Populus euphratica*	PREDICTED: vacuolar-processing enzyme [Populus euphratica]	XP_011018251

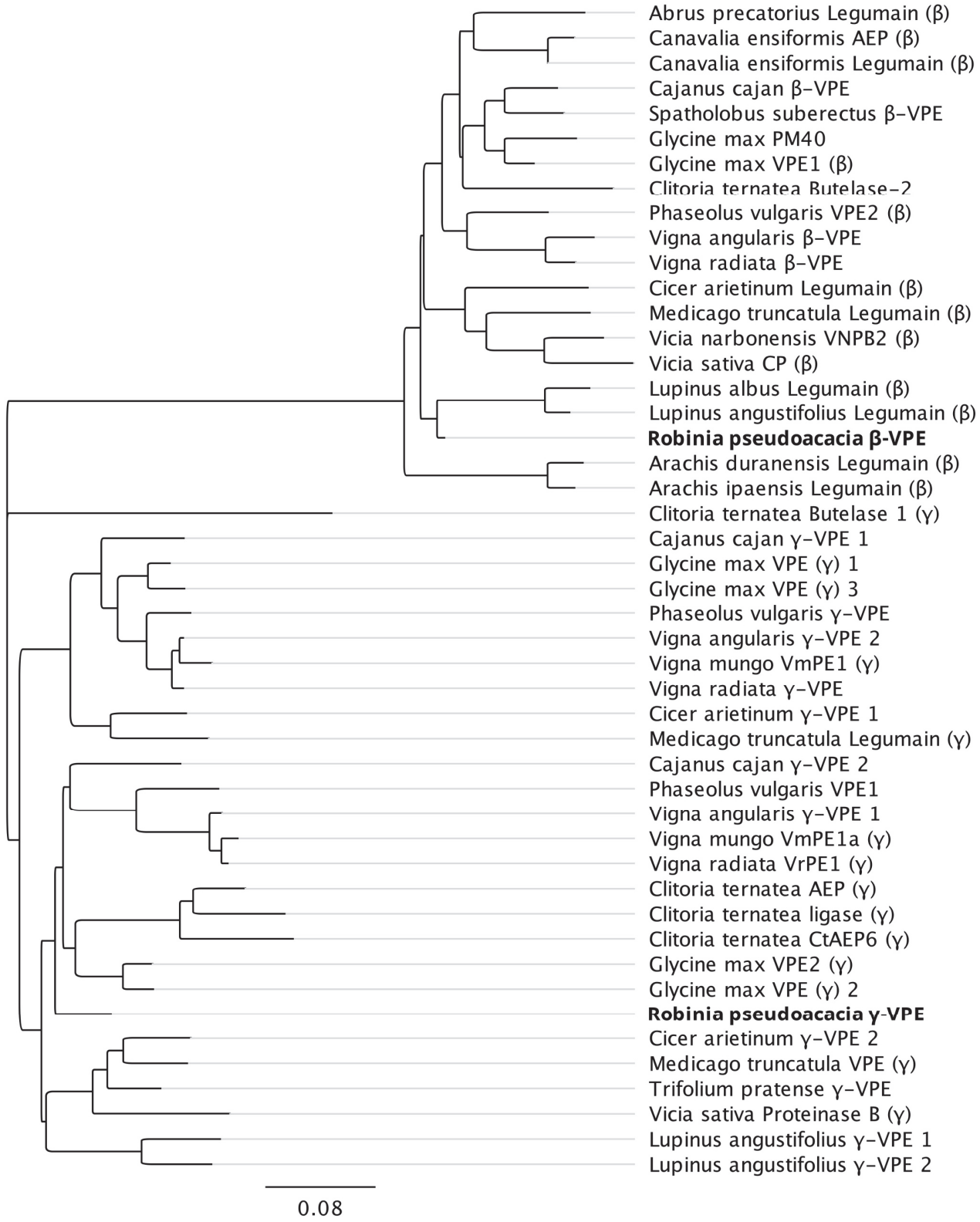


Figure 35: Phylogenetic tree constructed for vacuolar processing enzymes (VPEs) in legumes. The obtained sequences from *Robinia* are in bold.

The intended size for this page is A3. The full-sized page is available in the supplement material.

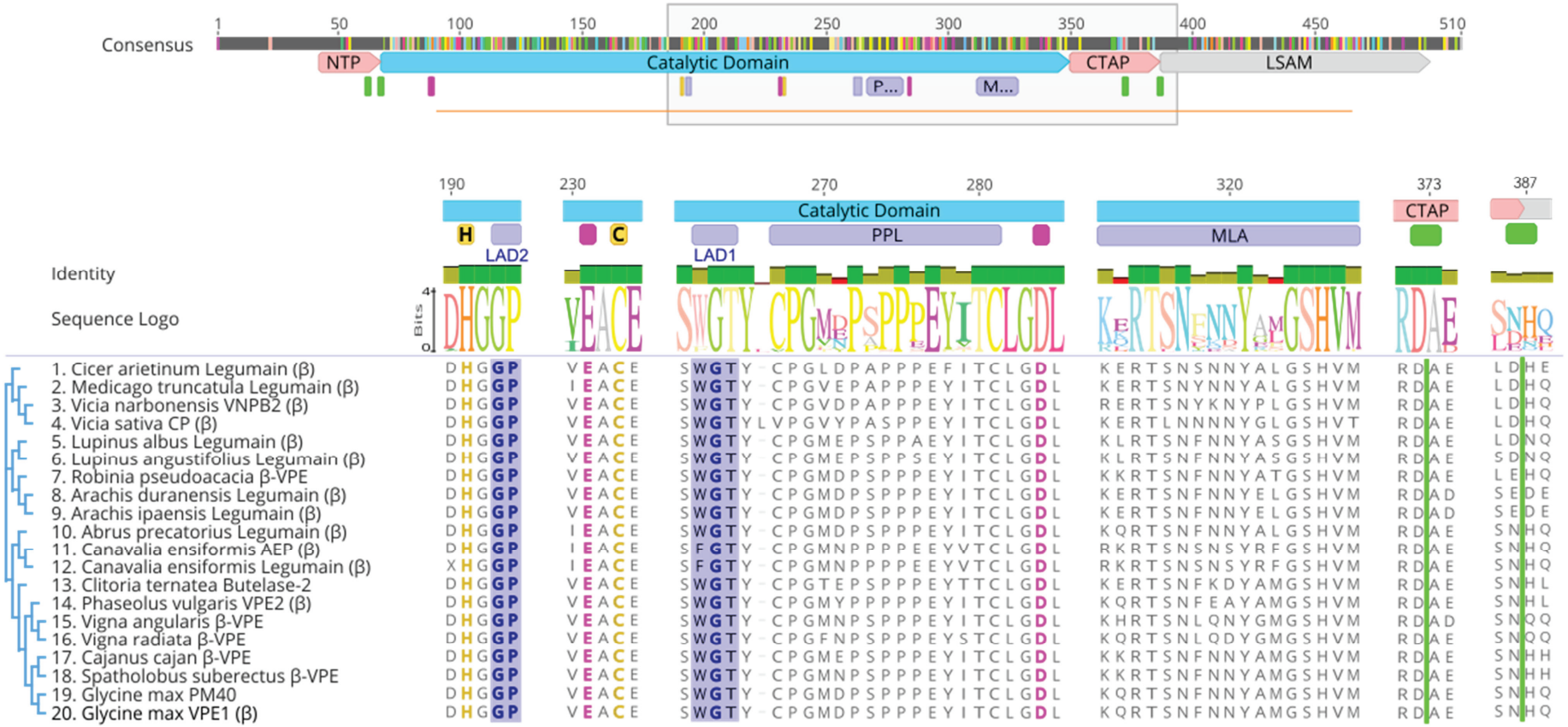


Figure 36: Amino acid sequence alignment of β-VPE in Fabaceae. Overview of the consensus sequence is displayed on the top, showing the full length of the gene and coverage of B-RpVPE represented by an orange line; key regions within the shaded box are shown in the alignment below. Blue elements correspond to the core catalytic domain, pink indicates the N-terminal propeptide (NTP) on the left and C-terminal activation propeptide (AP) on the right, and grey elements mark the Legumain Stabilization and Activity Modulation (LSAM) domain. Annotations: (yellow) conserved catalytic residues of cysteine protease; (green) autocatalytic processing sites with the vertical line indicating the cleavage; (magenta) S1-specificity residues; (lilac) determinant regions for ligase verse protease, including ligase-activity determinant (LAD) gatekeepers, poly-proline loop (PPL), marker of ligase activity (MLA). Bases shaded in lilac mark LAD2 and LAD1, residues in bolded blue text match with proposed signatures for protease according to Du *et al.* (2020).

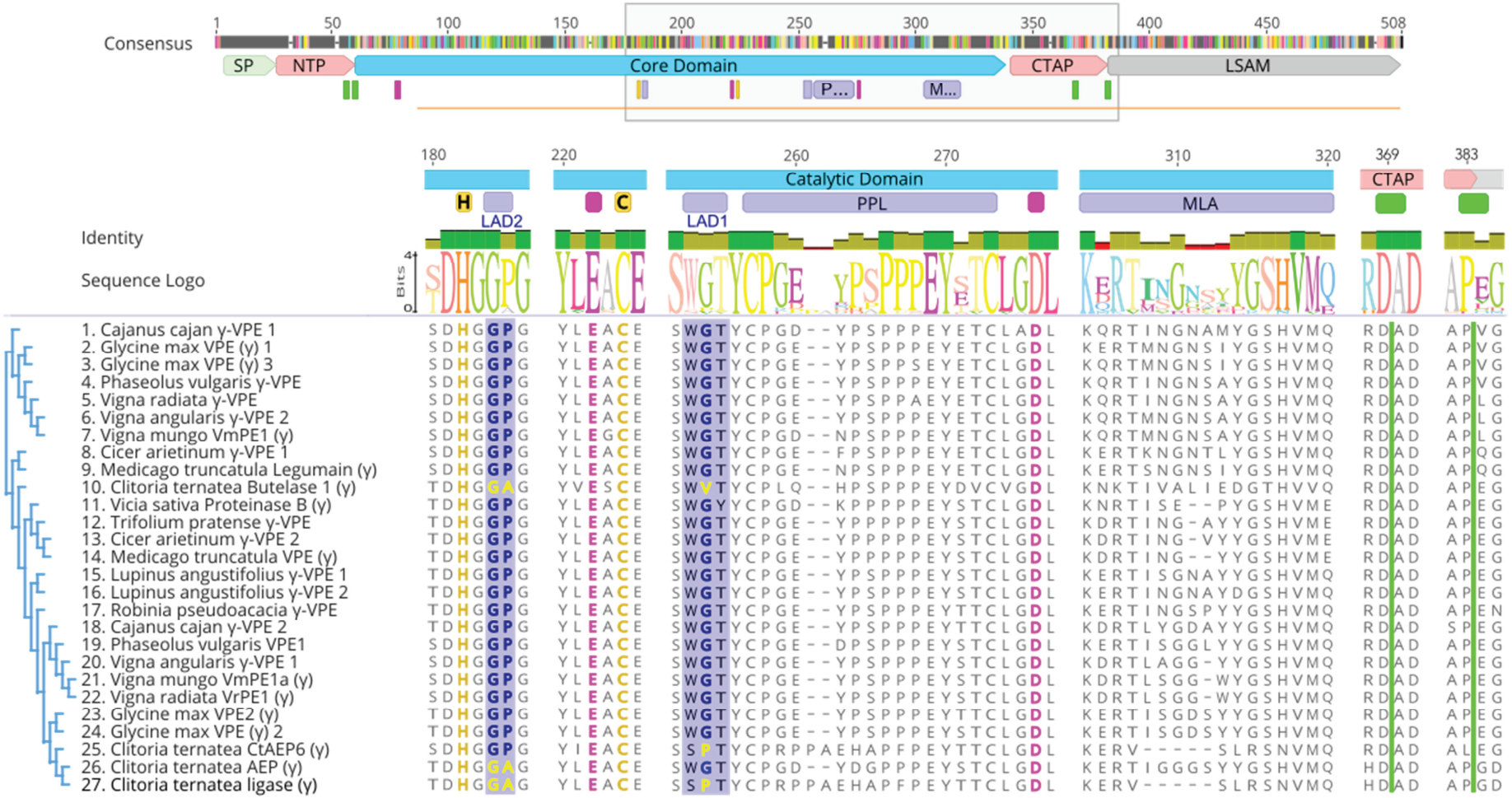


Figure 37: Amino acid sequence alignment of γVPE in Fabaceae. Overview of the consensus sequence is displayed on the top, showing the full length of the gene and coverage of γ-RpVPE as an orange line; key regions in the shaded area are shown as alignment below. Light blue elements correspond to the core catalytic domain, pink element indicates the N-terminal propeptide on the left and C-terminal activation peptide on the right, and grey elements mark the Legumain Stabilization and Activity Modulation (LSAM) domain. Annotations: (yellow) conserved catalytic residues of cysteine protease; (green) autocatalytic processing sites with the vertical line indicating the cleave; (magenta) S1-specificity residues; (lilac) determinant regions for ligase verse protease, including ligase-activity determinant (LAD) gatekeepers, poly-proline loop (PPL), marker of ligase activity (MLA). Bases shaded in lilac mark LAD2 and LAD1, residues in bolded blue text match with proposed signatures for protease and yellow ligase in Du *et al.* (2020).

4.4 Validation for gene expression analysis

4.4.1 qPCR on reference gene

The gene *Rp*18S and its qPCR primer pair have been previously validated as a stable reference for gene expression study in HWF (Lange, 2009). All samples in the set were collected in equal parts to set a standardized pooled sample. A 5-fold serial dilution was made and used for generating the standard curve by qPCR. The results from both the standard curve and disassociation analysis indicate that the primer pair perform adequately (Figure 38).

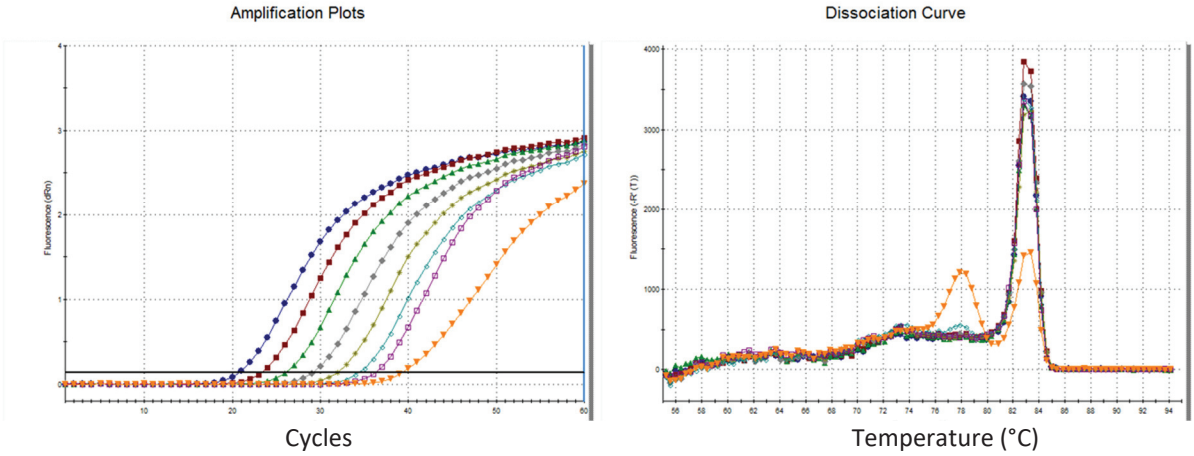


Figure 38: Amplification plots and disassociation analysis on the primer pair for reference gene *Rp18S.*

qPCR results with the reference genes offered an indication of the sample quality. Samples from Region 1 to 5 from all months passed the criteria for quality control in the dissociation curve and cycle threshold (Ct) value. With the samples from Region 6, however, sub-optimal Ct values and double peaks in the dissociation curve were observed. Samples that did not pass the quality control were omitted in further presentation.

4.4.2 Primer validation

A list of qPCR primers was designed for each candidate gene. They were tested with gradient PCR and gel electrophoresis to determine primer specificity and optimal annealing temperature. Examples of gel photos is shown and described in Figure 39.

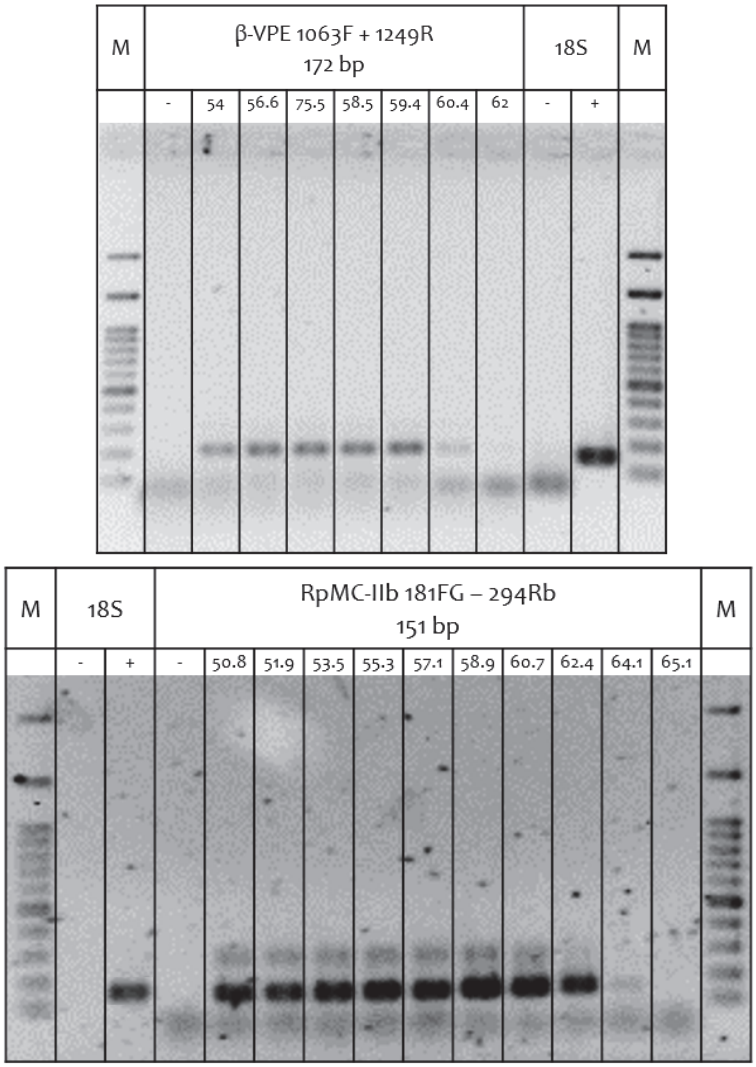


Figure 39: Photos after gradient PCR and gel electrophoresis. In the first row above the gel photo: Marker (M) shows the DNA ladder, name of the primer pair, and the amplicon size. The primer pair 18S was used as the positive (+) and negative (-) control indicated in the second row under 18S. The no template control testing primer pair is indicated by "-" within the corresponding columns. Numbers above the gel photo are the annealing temperature (°C). Two example primer pairs are shown here: *β-RpVPE* with the standard annealing temperature at 60°C and *RpMC-IIb* having the highest at 65°C

For each primer pair that resulted in a single band in the gel, its PCR amplicon was then validated by sanger sequencing (Figure 40 - Figure 47).

The intended size for this page is A3. The full-sized page is available in the supplement material.

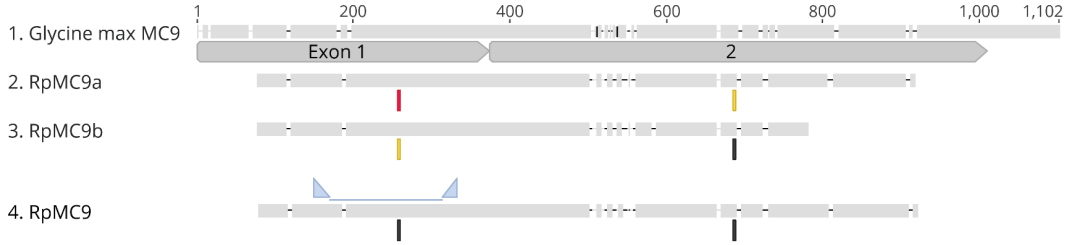


Figure 40: DNA sequence alignment overview showing the locations of *RpMC9* qPCR primers and amplicon. Grey element corresponds to the exons of *MC9* in *Glycine max*. Colored markings under the sequences of *RpMC9* are annotations on the heterogeneous sites. Blue triangles indicate the primers used in the qPCR experiment, and the lines between them represent the amplicons, which lie on exon 1.

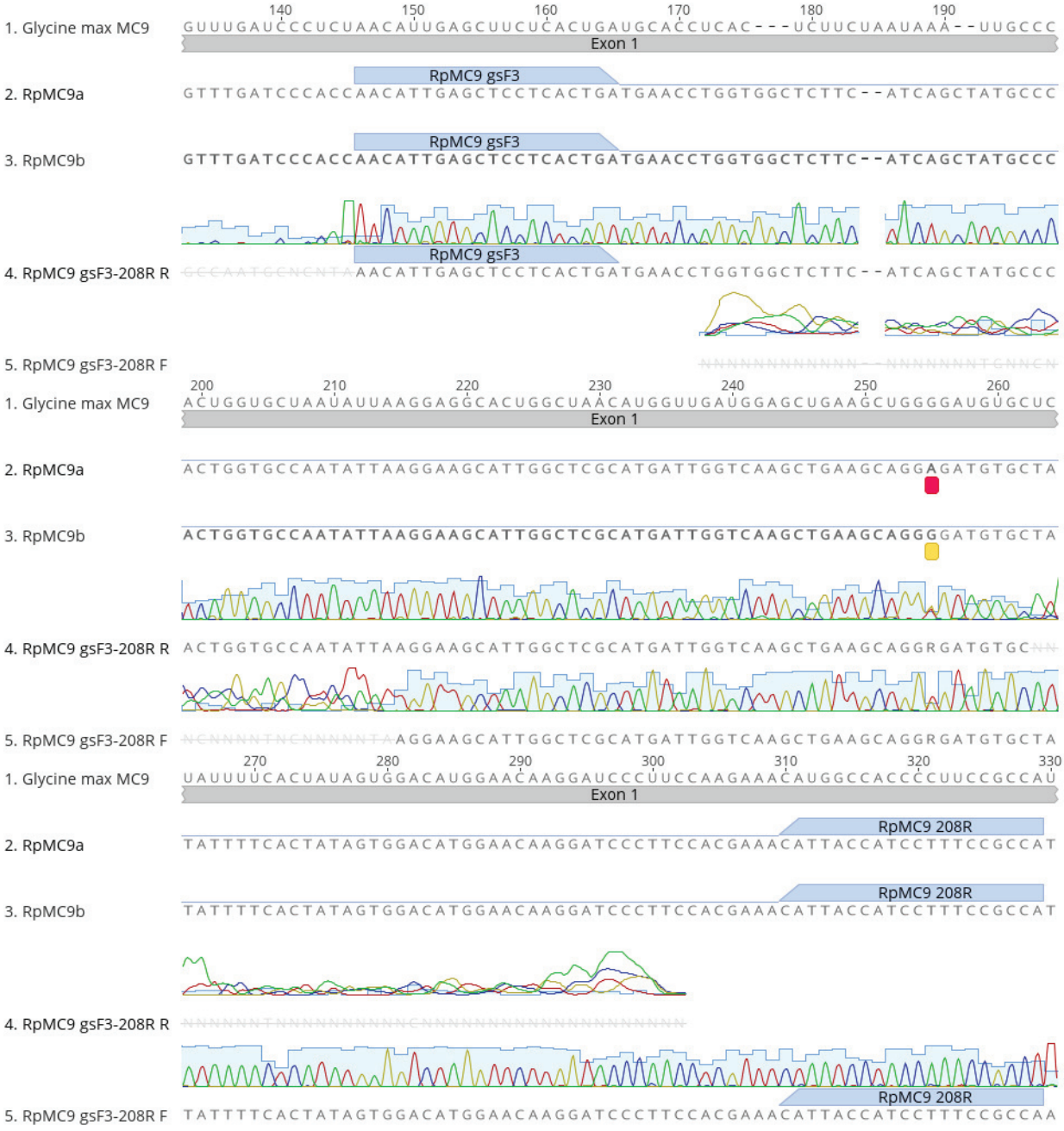


Figure 41: *RpMC9* qPCR primer validation with aligned chromatogram from sanger sequencing. Grey element corresponds to the exons of MC9 in genes in *Glycine max*, the red and yellow annotations under the sequences indicate the heterogeneous sites. The primers used in the qPCR experiment are indicated by blue elements, and the lines between them correspond to the amplicons.

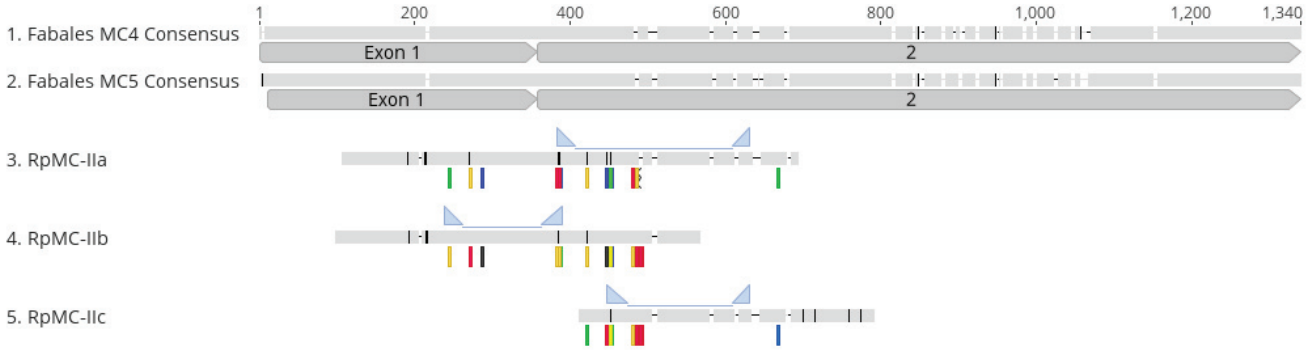


Figure 42: Locations of qPCR primers and amplicons for the three *RpMC-II* in DNA sequence alignment overview. Grey element corresponds to the exon in the consensus sequences of MC4, and MC5 homologs, the marks under sequences of *RpMC-IIs* are annotations on the heterogeneous sites. The primers used in the qPCR experiment are indicated by blue triangles, and the lines between them represent the amplicons. In both *RpMC-IIa and RpMC-IIc*, the qPCR amplicons locate on exon 2, whereas in *RpMC-IIb*, it spans across Exon 1 and 2.

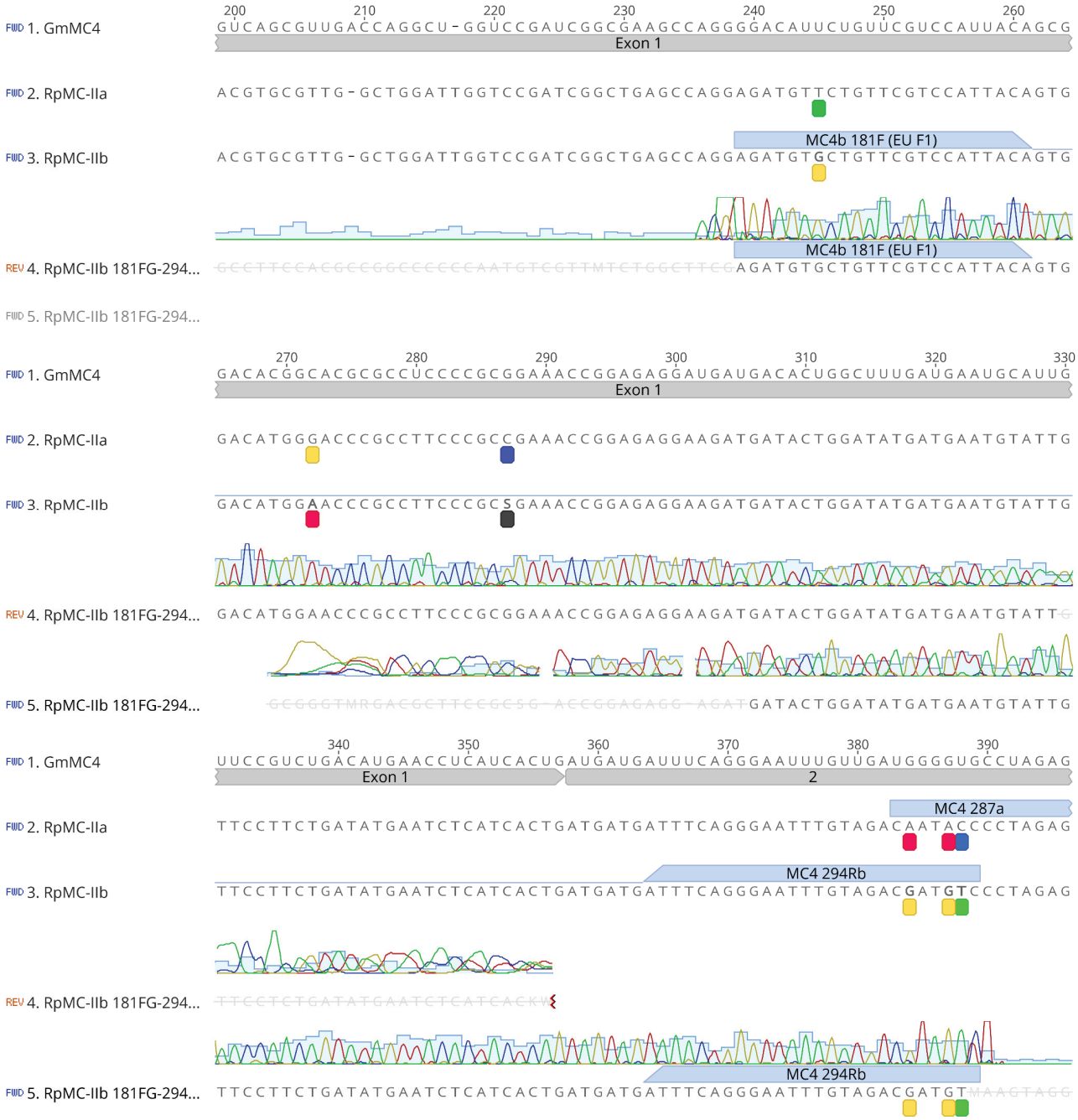


Figure 43: *RpMC9* qPCR primer validation with aligned chromatogram from sanger sequencing. Grey element corresponds to the exons of MC9 in genes in *Glycine max*, the red and yellow annotations under the sequences indicate the differing sites between the three *RpMC-II* sequences. The primers used in the qPCR experiment are indicated by blue elements, and the lines between them correspond to the amplicons.

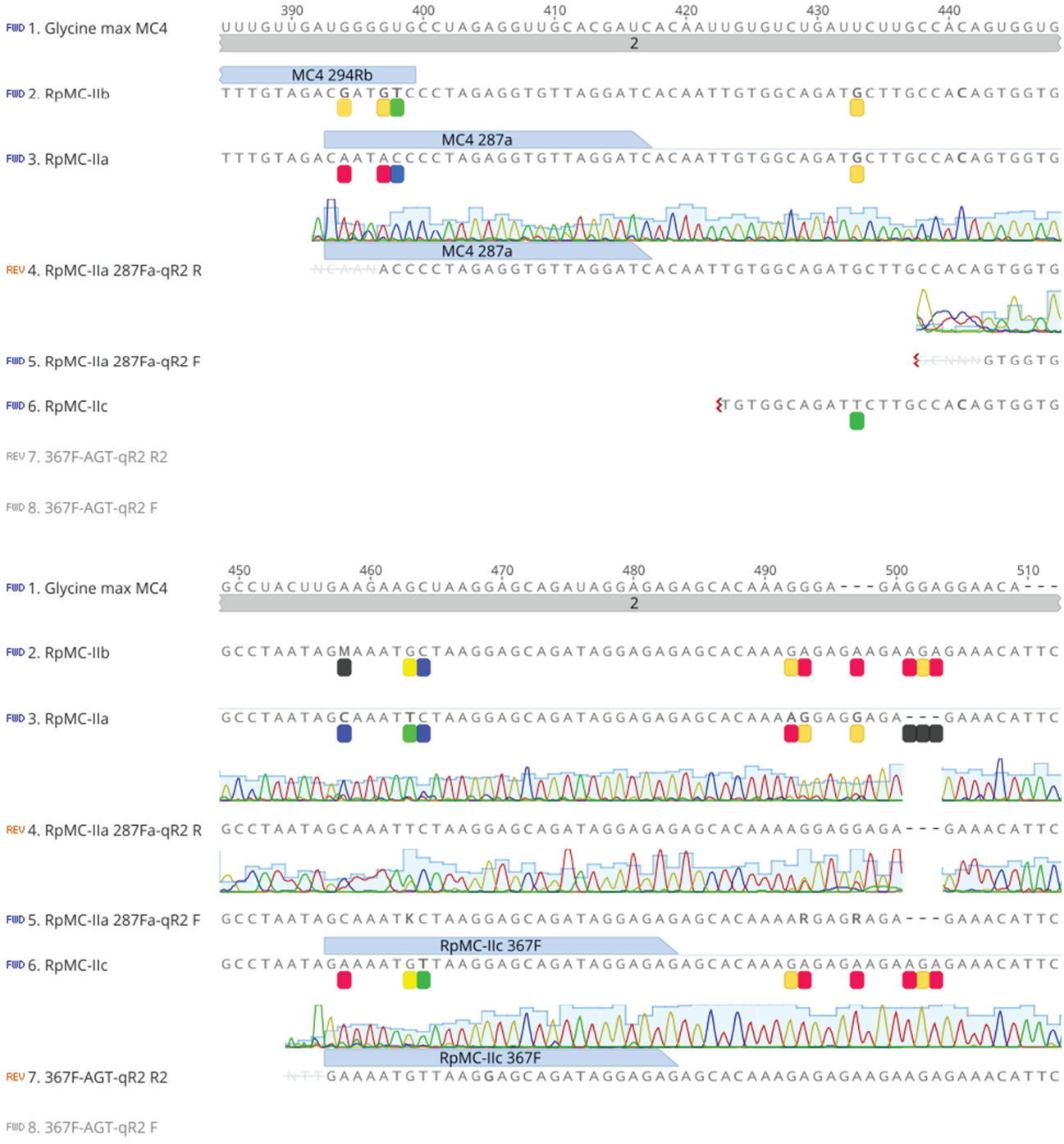


Figure 44 (first half, continue on the next page): *RpC-II a and RpMC-IIc* qPCR primer validation with aligned chromatogram from sanger sequencing. Grey element corresponds to the exons of MC4 in *Glycine max*, the annotations under the sequences indicate the differing sites between the three *RpMC-II* sequences. The primers used in the qPCR experiment are indicated by blue elements, and the lines between them correspond to the amplicons. The two pairs of primers lie in overlapping positions on the alignment, both on the second exon of their corresponding genes.

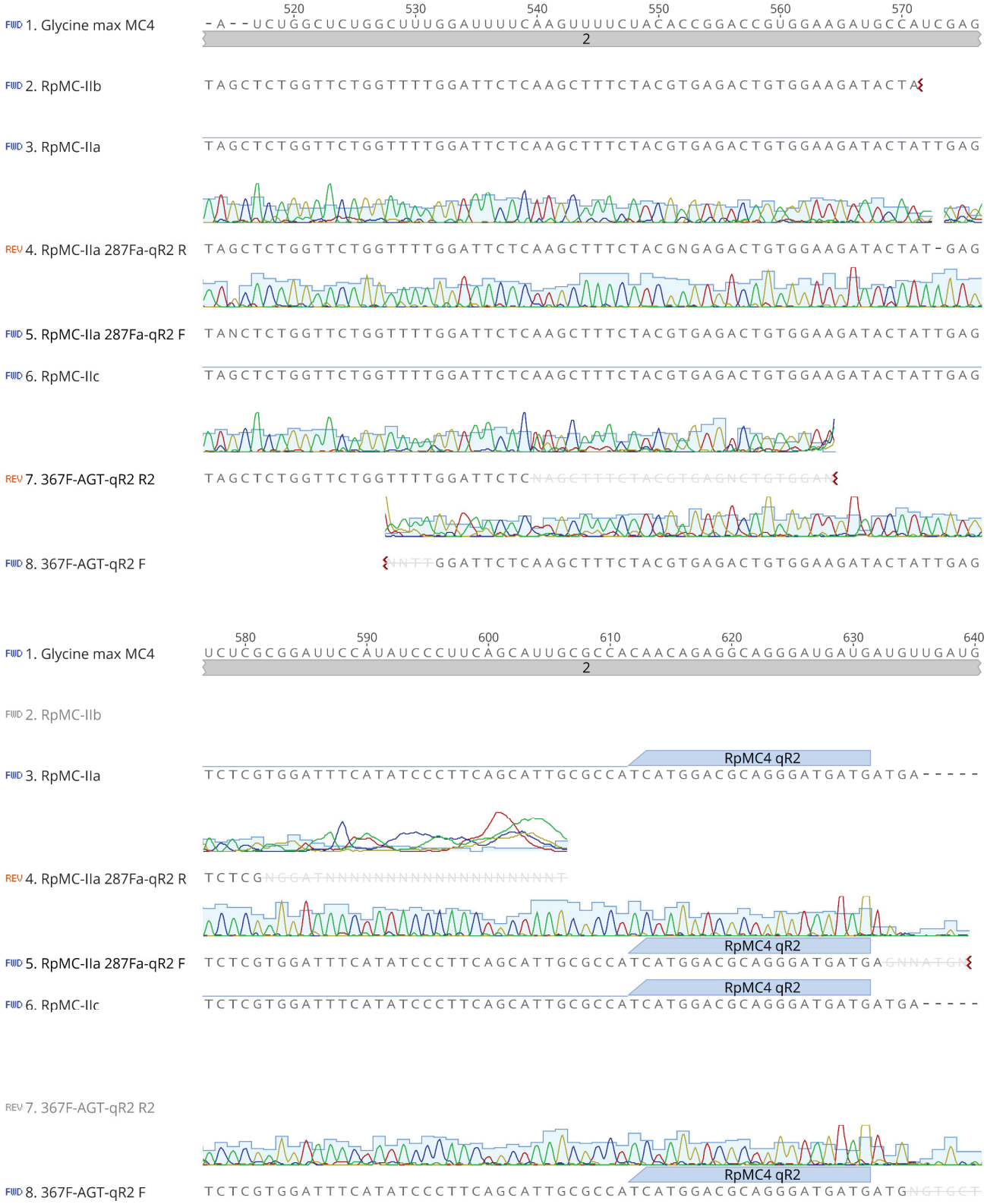
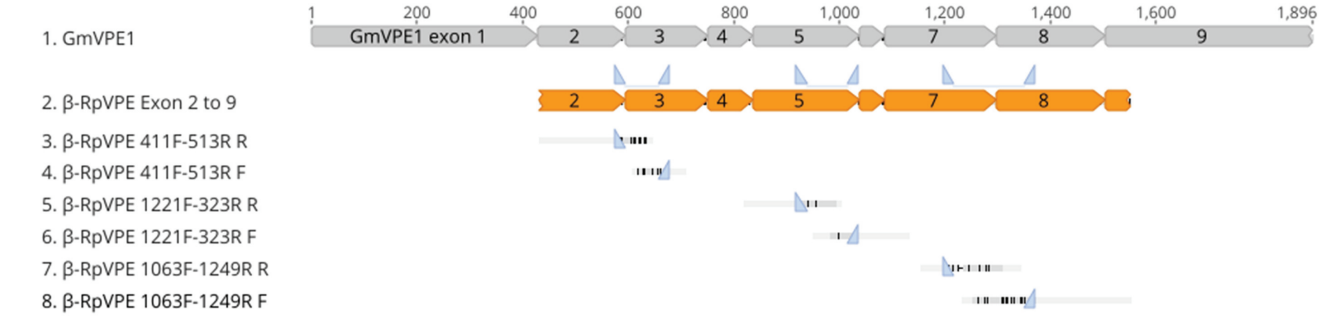


Figure 44: (second half, continue from the last page): *RpC-II a and RpMC-IIc* qPCR primer validation with aligned chromatogram from sanger sequencing. Blue elements indicate the primers used in the qPCR experiment, and the lines between them correspond to the amplicons; the two genes have the same reverse primers. The differing sites are shown in the first half of this figure on the last page.

The intended size for this page is A3. The full-sized page is available in the supplement material.



Lane	Sequence label	Sequence description	*Accession*
1	*GmVPE1*	β-type VPE in Glycine max	NM_001249749.2
2	*β-RpVPE*	β-*RpVPE* sequence obtained by PCR	--
3 - 4	*β-RpVPE* 411F and 513R F and R	Amplicon from PCR with the primer pair *β-RpVPE* 411F and 513R	--
5 - 6	*β-RpVPE* 1221F and 323R F and R	Amplicon from PCR with the primer pair *β-RpVPE* 1221F and 323R	--
7 - 8	*β-RpVPE* 1063F and 1249R F and R	Amplicon from PCR with the primer pair *β-RpVPE* 1063F and 1249R	--
			-- Indicates not applicable

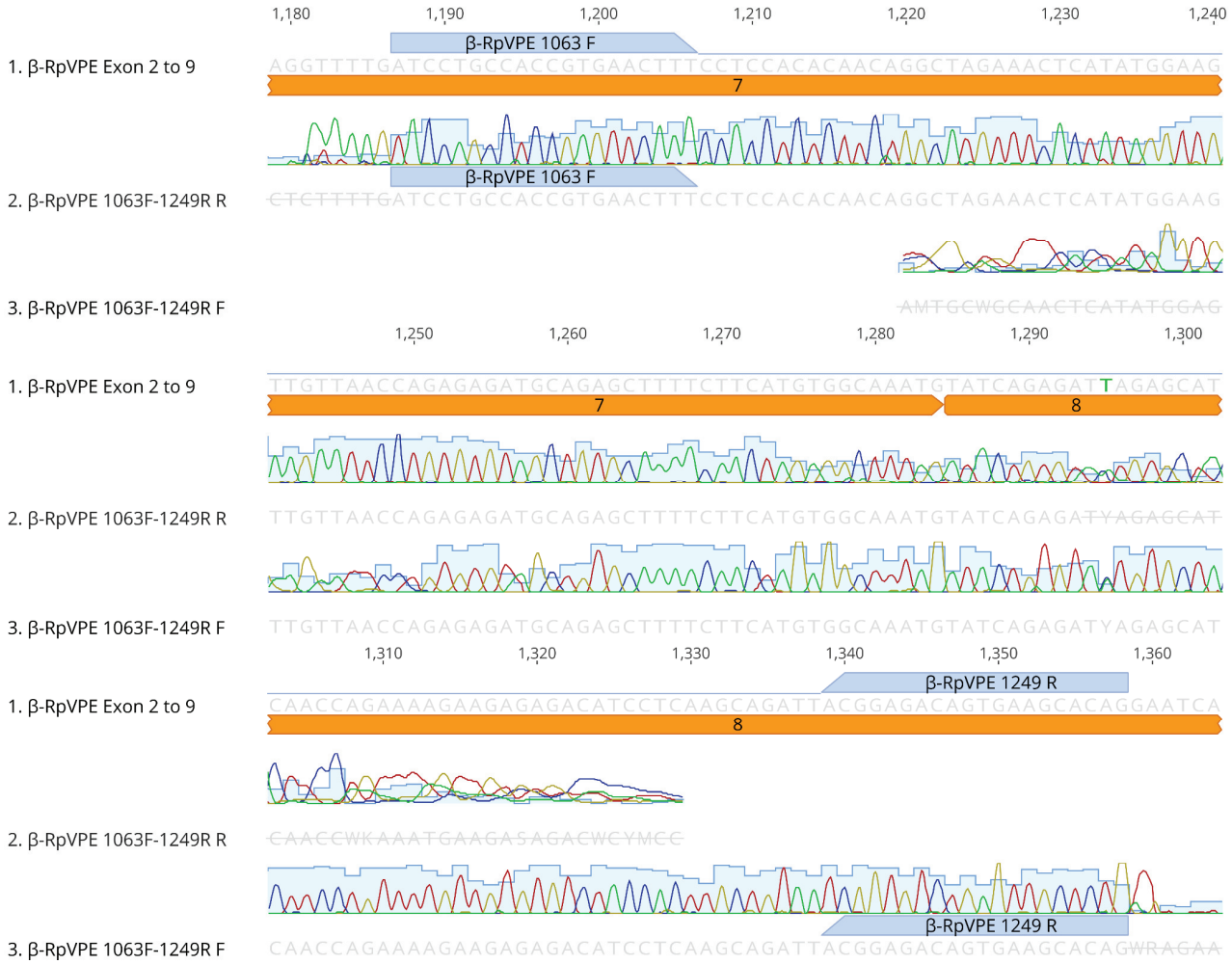
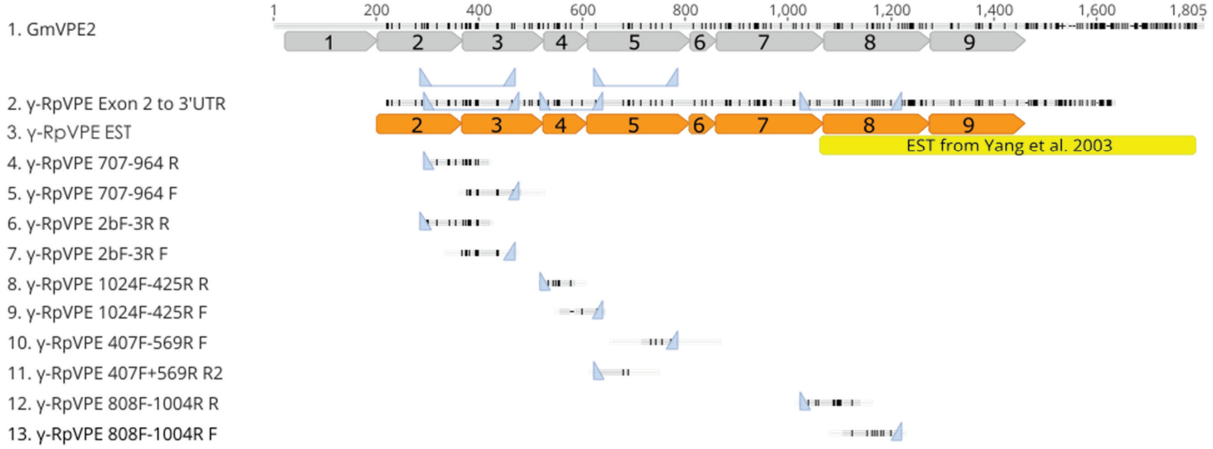


Figure 45: *β-RpVPE* qPCR primer validation with sanger sequencing. (Top) DNA sequence alignment overview with sequence details is listed in the table underneath. Blue triangles indicate the qPCR primers. The grey element corresponds to the exon regions of *GmVPE1* and the orange element for *β-RpVPE*. (Bottom) Chromatogram of the amplicon from PCR with the primer pair *β-RpVPE* 1036F and 1249R, the pair used in the qPCR experiment.



Lane	Sequence label	Sequence description	*Accession*
1	*GmVPE2*	γ-type VPE in Glycine max	NM_001249635.2
2	*γ-RpVPE*	*γ-RpVPE* sequence obtained by PCR	--
3	*γ-RpVPE* EST	EST of *γ-RpVPE* in reference to (Yang *et al.*, 2004)	BI677606
4 -5	*γ-RpVPE* 707F - 964R F and R	Amplicon from PCR with the primer pair *γ-RpVPE* 707F and 964R	--
6-7	*γ-RpVPE* 2bF -3R F and R	Amplicon from PCR with the primer pair *γ-RpVPE* 2bF and 3R	--
8-9	*γ-RpVPE* 1024F and 425R F and R	Amplicon from PCR with the primer pair *γ-RpVPE* 1024F and 425R	--
10-11	*γ-RpVPE* 407F - 569R F and R	Amplicon from PCR with the primer pair *γ-RpVPE* 407F and 569R	--
12-13	*γ-RpVPE* 808F - 1004R F and R	Amplicon from PCR with the primer pair *γ-RpVPE* 808F and 1004R	--

-- Indicates not applicable

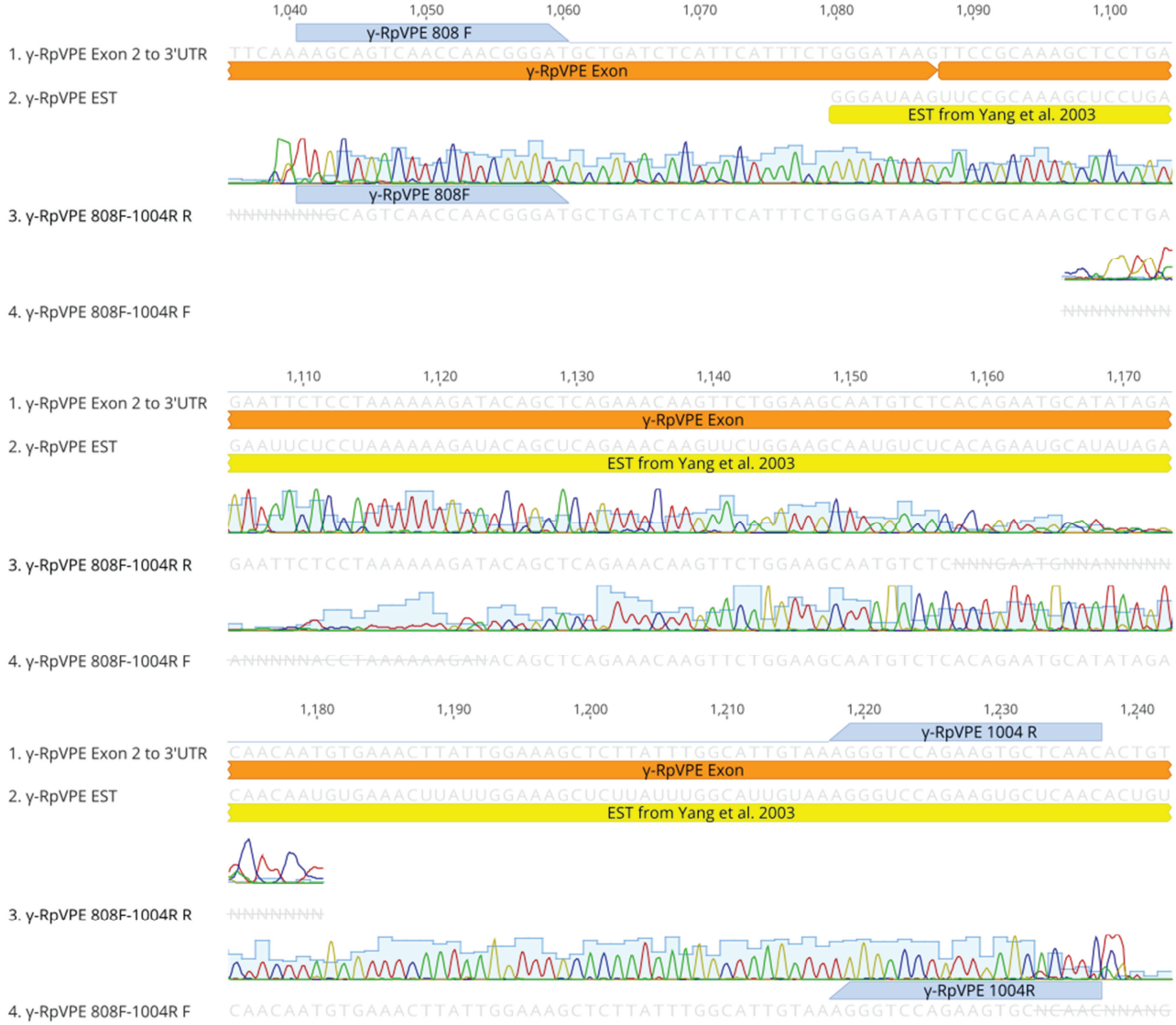
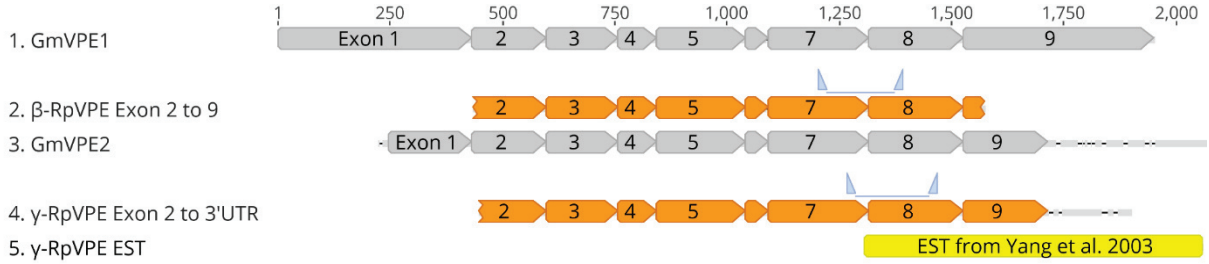


Figure 46: *γ-RpVPE* qPCR primer validation with sanger sequencing. (Top) DNA sequence alignment overview with sequence details listed in the table below. The qPCR primers are indicated by blue triangles. The Grey element corresponds to the exon regions of *GmVPE2* and the orange element for *γ-RpVPE*, yellow element shows the EST of *γ-RpVPE*. (Bottom) Chromatogram of the amplicon from PCR with the primer pair *γ-RpVPE* 808F and 1004R, which is the pair used in the qPCR experiment.



Lane	Sequence label	Sequence description	Accession
1	*GmVPE1*	β-type VPE in Glycine max	NM_001249749.2
2	*β-RpVPE*	β-*RpVPE* sequence obtained by PCR	--
3	*GmVPE2*	γ-type VPE in Glycine max	NM_001249635.2
4	*γ-RpVPE*	*γ-RpVPE* sequence obtained by PCR	--
5	*γ-RpVPE* EST	EST of *γ-RpVPE* in reference to (Yang *et al.*, 2004)	BI677606
			-- Indicates not applicable

Figure 47: Locations of qPCR primers and amplicons for the two *RpVPE* in DNA sequence alignment overview. The Grey element corresponds to the exon regions of VPE genes in *Glycine max* and the orange element for *RpVPE*. The yellow element shows the EST of *γ-RpVPE* in reference to Yang *et al.* (2003). The primers used in the qPCR experiment are indicated by blue triangles, and the lines between them represent the amplicons. In both *RpVPE*s, the amplicon span across exon 7 to 8.

A final step to validate the qPCR primers was the testing for their efficiency and specificity. Each primer pair validated by sanger sequencing were subjected to stand curve and disassociation curve analysis on the qPCR instrument, as the results described in Figure 48, and the qPCR primers that passed the quality control are listed in Table 20.

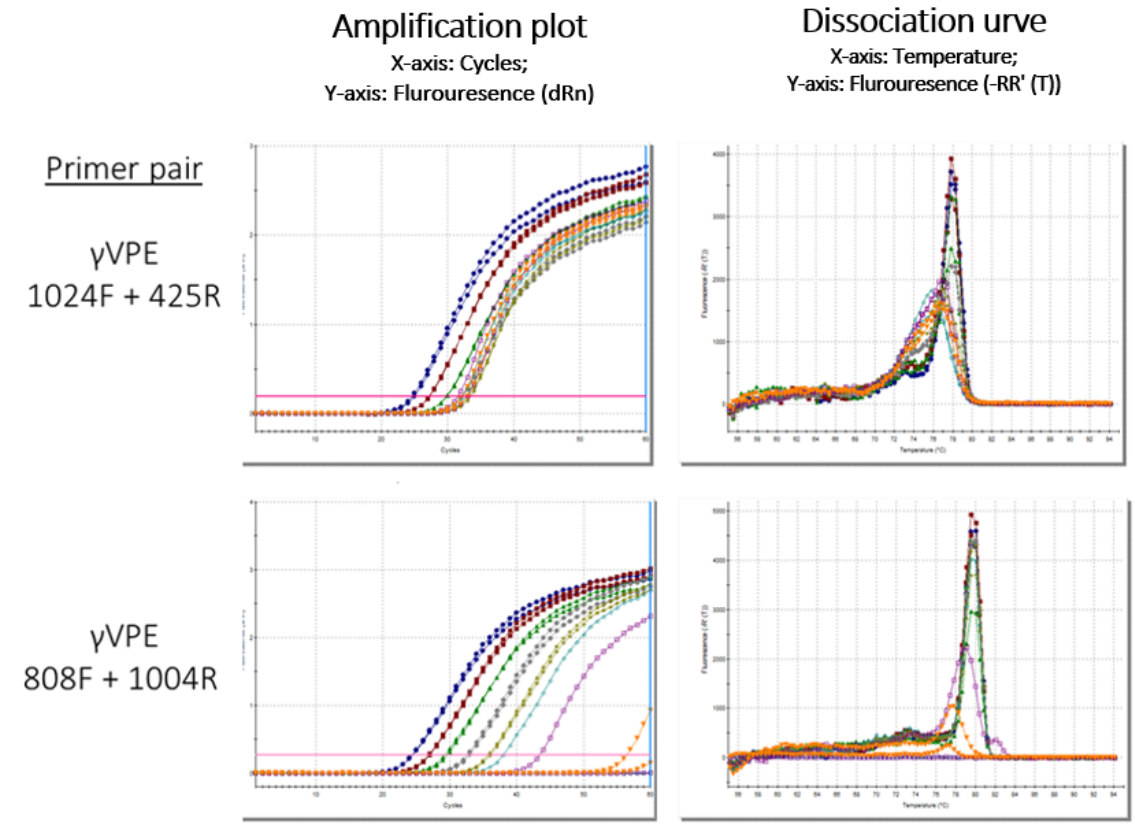


Figure 48 qPCR primer validation with amplification plots and dissociation curve, with 2 of the *γ-RpVPE* candidate primer pairs as examples. The primer pair *γ-RpVPE* 1024F-425R form primer dimers that resulted in peaks at a lower temperature when the concentration of cDNA is lower than the third sample in serial dilution, as shown in the DC on top. With the primer pair *γ-RpVPE* 808F-1004R, primer dimer only occurred in the lowest concentration and no template control.

Table 20: List of qPCR primers used in the experiment.

Gene name	qPCR Primers					Amplicon	
	Direction	Sequence (5' to 3')	Tm (°C)	Ta (°C)	Efficiency	Size (bp)	Tm (°C)
Rp18S	F	GCTTTTAGGACTCCGCTGG	58.8	60	2.02	155	83.3
	R	GGTAAGTTTCCCCGTGTTGA	57.7				
RpPAL1	F	GGAATGCCTTGGGGAGTG	57.7	60	2.04	147	78.6
	R	CGACGCTGTATTGATTTGCC	57.9				
RpCHS3	F	ACCACAGGTGAAGGACTCG	59.0	60	2.03	127	80.6
	R	GATAAAATCCAGGTCCAAAGC	54.9				
RpMC9	F	AACATTGAGCTCCTCACTGA	56.5	62	1.96	182	82.2
	R	TGGCGGAAAGGATGGTAATG	57.9				
RpMC-IIa	F	CAATACCCCTAGAGGTGTTAGGATC	59.8	64	1.94	236	82.1
	R	CATCATCCCTGCGTCCATGA	59.9				
RpMC-IIb	F	AGATGTGCTGTTCGTCCATTACA	60.3	65	1.91	151	80.8
	R	GACATCGTCTACAAATTCCCTGAAAT	59.7				
RpMC-IIc	F	GAAAATGTTAAGGAGCAGATAGGAGAG	59.6	64	2.07	175	80.8
	R	CATCATCCCTGCGTCCATGA	59.9				
β-RpVPE	F	ATCCTGCCACCGTGAACTTT	59.9	60	1.96	172	80.7
	R	CTGTGCTTCACTGTCTCCGT	60.0				
γ-RpVPE	F	AAGCAGTCAACCAACGGGAT	59.9	60	1.94	197	80.1
	R	GTTGAGCACTTCTGGACCCT	59.6				

F = forward; R = reverse; Ta = annealing temperature; Tm = melting temperature

For the quality control on qPCR data, dissociation curve analyses were conducted at the end of every qPCR run. With the sample assay that generated a Ct value without a peak at the designated temperature in the dissociation curve, they are either considered as no value or omitted, as detailed in Figure 49.

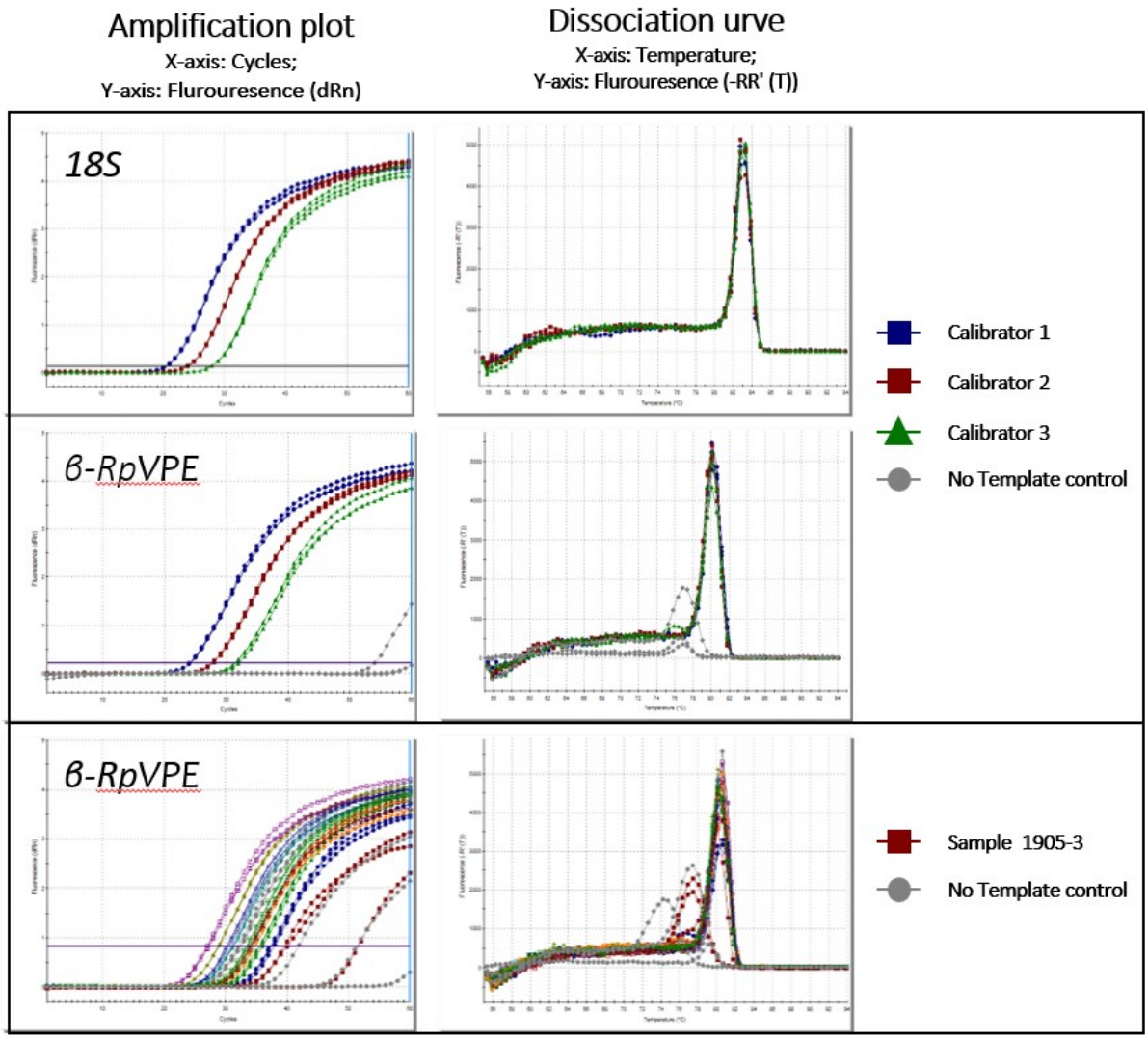


Figure 49: Example of qPCR result in amplification plot (AP) on the left and dissociation curve (DC) on the right for the demonstration of quality control criteria. (Top) AP and DC of calibrator samples in three 10-fold serial dilutions conducted with reference gene 18S. Middle: AP and DC of the same calibrator samples conducted with the candidate gene *β-RpVPE*. (Bottom) Also, AP and DC from all samples of Region-3 ran with *β-RpVPE* primer pair; all samples except 1905-3 resulted in the correct peak at the temperature between 80.1-80.6°C. The data from sample 1905-3 that peaked at 77.4°C matched with one for the no-template-control and therefore counted as no value.

4.5 Gene expression of the marker genes

Based on previous findings in (Lange, 2009), two key enzymes in the flavonoid biosynthetic pathway, *PAL* and *CHS*, were selected as the marker genes for HWF. From relative gene expression (RE) patterns of *RpPAL1* in the current results, a sharp increase is observed in Region 2 in June 2019 and gradually decreases over the summer months. A strong and abrupt rise in RE occurred in September in the innermost sapwood (ISW), then a mild and gradual increase in RE occurred in the TZ from July to November (Figure 50).

The RE of *RpCHS3* showed little observable increase in the outer wood regions. However, distinct elevation exceeding the expression level of 18S by 3-fold was detected in the ISW and TZ in the autumn months. The strongest expression of *RpCHS3* in the ISW coincides with that of *RpPAL1* in September, whereas the peak level of expression in the TZ appeared in the later months.

The intended size for this page is A3. The full-sized page is available in the supplement material.

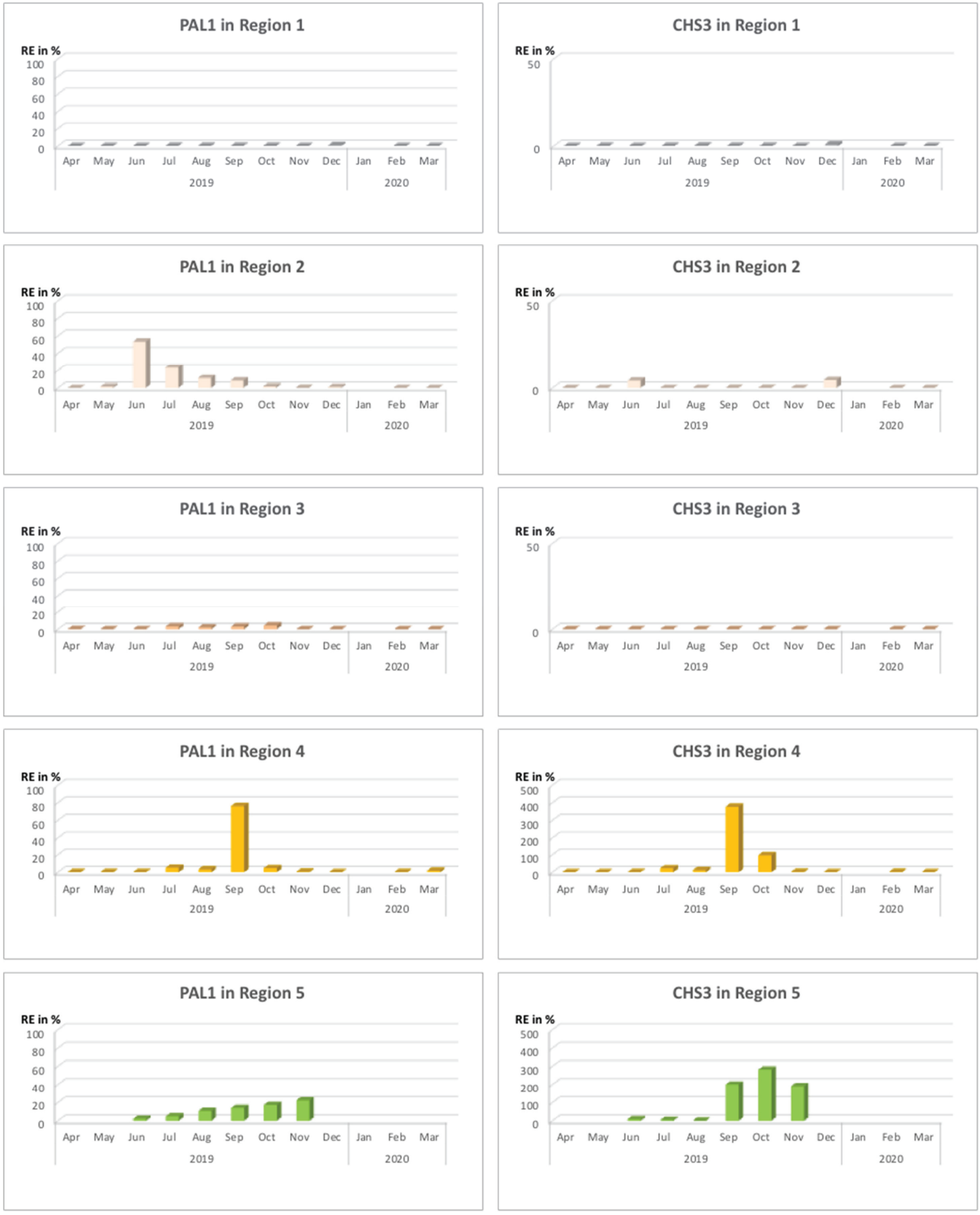


Figure 50: Relative gene expression (RE) of two heartwood formation marker genes, Phenylalanine ammonia-lyase 1 (*PAL1*) and Chalcone synthase 3 (*CHS3*), to the reference gene 18S in *Robinia pseudoacacia*. (Left) A sharp increase in RE of *PAL1* occurred in Region 2 (differentiating xylem) in June and gradually declined during the summer months, with an abrupt rise to nearly 80% in September in the innermost sapwood (ISW). (Right) the RE of *CHS* showed little observable increase in the outer wood regions, but substantial increases are detected in both the ISW and transition zone in the autumn months.

4.6 Gene expression of the candidate genes

4.6.1 Gene expression of each candidate

The RE of γ-*RpVPE* showed a mild fluctuation in all regions and time points, with the maximum elevated expression at around 20% relative to 18S. Nevertheless, elevation patterns are visible in the outer wood regions during summer and the inner regions during autumn in 2019 (Figure 51). In the same figure, an upsurge of *β-RpVPE* RE occurred in Region 2 in June; in November, an observable upregulation was detected in Regions 4 and 5. In February and March, discernible upregulation is seen in the phloem, middle and innermost sapwood.

Compared to *RpVPE*s, the RE of *RpMC*s have a more conspicuous level of fluctuation and are seen earlier in the year (Figure 52 and Figure 53). The three MC4-like *RpMC-IIs* have similar patterns of RE in the outer wood regions, and they all showed a rise in the phloem in June at varying levels. In the differentiating xylem, RE of the three *RpMC-II* and *RpMC9* are intensely up-regulated in spring starting in April, followed by gradual declines except for *RpMC-II*c, which ceased abruptly. Their RE patterns differ in the inner wood regions. Little noticeable RE was detected for *RpMC-IIa and RpMC-IIb* in Region 3, whereas a gradual rise in *RpMC-IIc* from July to October was observed. In the ISW, *RpMC-IIb* showed a distinct elevated RE in September and October. Finally, in the TZ, up-regulated RE of *RpMC-IIa* and -b are clear in November.

The intended size for this page is A3. The full-sized page is available in the supplement material.

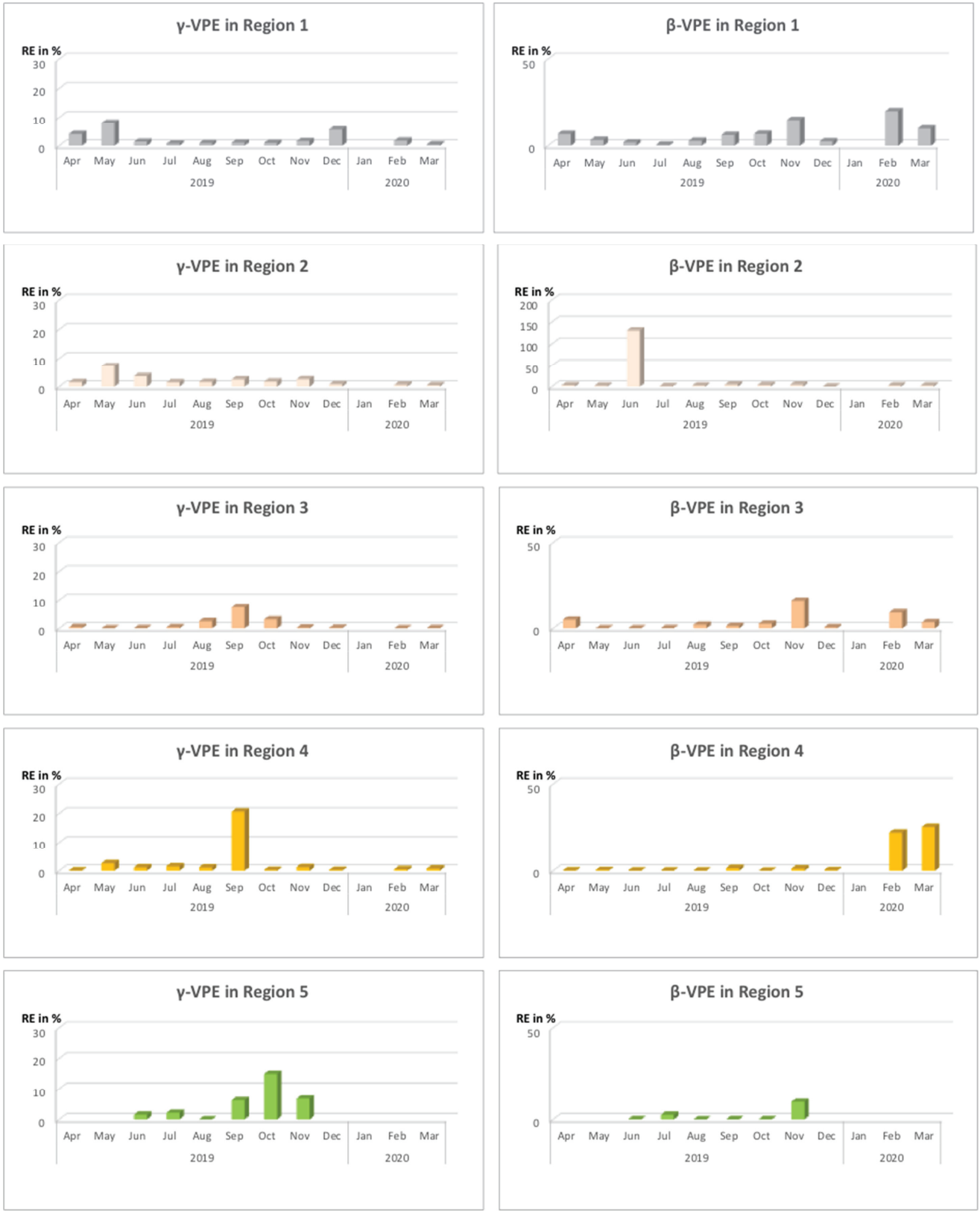


Figure 51: Relative gene expression (RE) of vacuolar processing enzymes (VPEs) to the reference gene 18S in *R. pseudoacacia*. (Left) The RE of *γ-RpVPE* showed a mild fluctuation in all regions and time points, as indicated in the Y-axes, with the maximum being only 30%. Nevertheless, elevation patterns are observable in the outer wood regions during spring and the inner regions during autumn. (Right) The RE of *βVPE* has a stronger level of fluctuation in comparison. In Region 2 (differentiating xylem), an abrupt increase in expression occurred in June. It also showed an observable rise in Region 3 (middle sapwood) in November and Region 4 in February and March.

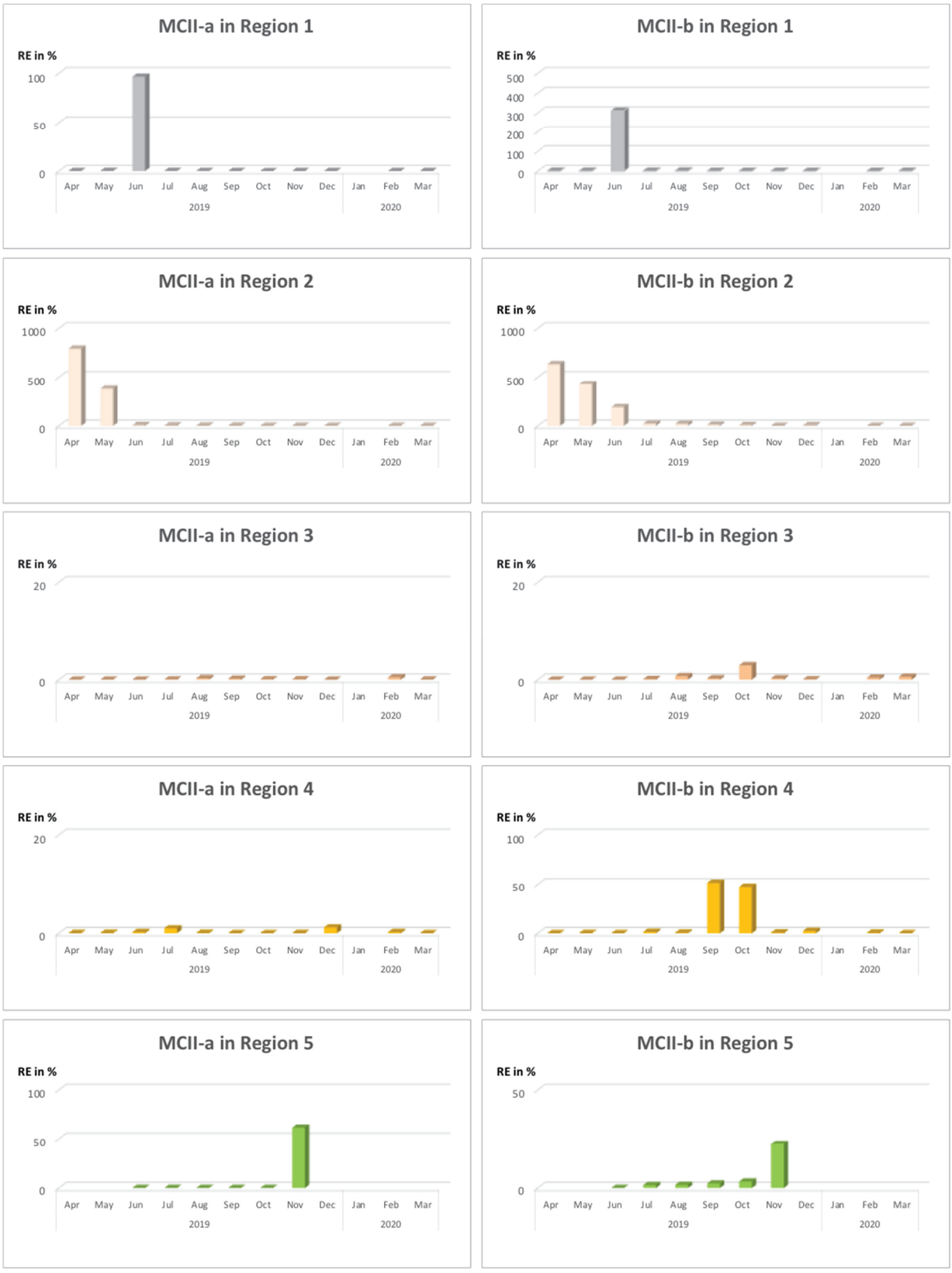


Figure 52: Gene expression (RE) of two type-II metacaspases, *RpMC-IIa* on the left and *RpMC-IIb* on the right, relative to the reference gene 18S in *Robinia pseudoacacia*. The Y- axes in each graph are individually adjusted to show the expression pattern within each sample group of the region. The patterns of RE are highly similar between the two genes, except for Region 4 (innermost sapwood). Both genes show an abruptly elevated expression in the phloem, with *RpMC-IIb* reaching a 3-fold increase. In Region 2 (differentiating xylem), both genes are strongly up-regulated in early spring, followed by gradual declines. The RE patterns of the two genes differ in Region 4; *RpMC-IIa* showed no observable changes, while *RpMC-IIb* had an apparent elevation in September and October. Both genes show an elevation in November, with *RpMC-IIa* at a higher level.

The intended size for this page is A3. The full-sized page is available in the supplement material.

Figure 53: Gene expression (RE) of two type-II Metacaspases relative to the reference gene 18S in *Robinia pseudoacacia*. The Y-axes in each graph were adjusted to show the expression pattern within each sample group (Left) RE of *RpMCII-c.* A sharp peak RE in the phloem is seen in June; in Region 2 (differentiating xylem), the RE is up-regulated over 5-folds in spring. A gradually increasing RE pattern started in July and peaked in October in Region 3 (middle sapwood). No significant change is observed in Regions 4 and 5 (innermost sapwood and transition zone). (Left) RE of *RpMC9.* No RE was detected in Regions 1 and 5. However, a substantial peak is observed in Region 2 in spring; another peak appears in July in Regions 3 to 4.

4.6.2 Gene expression patterns overviews

Weather data from the period relevant to the present study is included for an overview along with the gene expression pattern (Figure 54). In the first half of 2019, the sunshine duration fluctuated greatly; sharp increases were recorded between March and April and from May to June, and the latter coincided with a substantial rise in average temperature. Throughout the year's second half, sunshine duration and the average temperature gradually decline at an even pace. The rainfall fluctuation was somewhat even throughout the harvesting period, except for the drop in April 2019.

Figure 55 shows the gene expression patterns of the two HWF marker genes and the two *RpVPEs*, across the season and wood regions. In June 2019, there was a sharp elevation of RE with *PAL1* and *β-type* in Region 2, with the former followed by a gradual decline until October. γ-*RpVPE* also showed elevated RE in the outer regions, and they were observed in May in both the phloem and the differentiating xylem. On the other hand, expression of *CHS3* was not detectable in the outer and middle wood regions throughout the year. Intriguingly, a noticeable elevated RE of *β-RpVPE* was detected in the phloem, middle and inner sapwood in February and March; in 2020, these months were in the late winter. In Region 3 (middle sapwood), up-regulated RE of γ-*RpVPE* was seen from August to October and peaked in September. Up-regulated RE of γ-*RpVPE* occurred at a higher level and abruptly in September, and this abrupt elevation coincided with those of *PAL1* and *CHS3*. In the same region, *CHS3* and γ-*RpVPE* showed a synchronic elevation between September to November that peaked in October. In 2019, these months observed an even decline in average temperature and sunshine duration. Interestingly, this elevation pattern in Region 5 occurred a month later than that observed in Region 4.

The RE patterns of *RpMCs* are primarily observed in the outer wood regions, and are particularly conspicuous in the differentiating xylem in spring (Figure 56). In Region 2, the RE upsurges are at their highest in April, which preceded the elevation of *PAL1* and *RpVPEs*. While the elevated RE of *RpMC-IIb* and *-c* lasted until June at a low level in the same region, *RpMC-IIa* and *RpMC9* have ceased their expression by then. In

other words, the upregulated expression of *RpMC-IIs* occurred before the rise of γ-*RpVPE*, and the *PAL1* and *β-RpVPE* arrived as the upsurge of *RpMC-IIs* subsided. With the noticeable expression of *RpMC-IIs* in the phloem in early summer, *RpMC-IIb* showed the most substantial rise; it is also the only examined *RpMC-II* with distinct RE in Region 4 in Fall, matching that of the *CHS3*. RE of *RpMC9* appeared in a sharp level of contrast. There was no detectable gene expression in the phloem and the TZ, and the elevation level in the differentiating xylem reached nearly 25-fold of the reference gene.

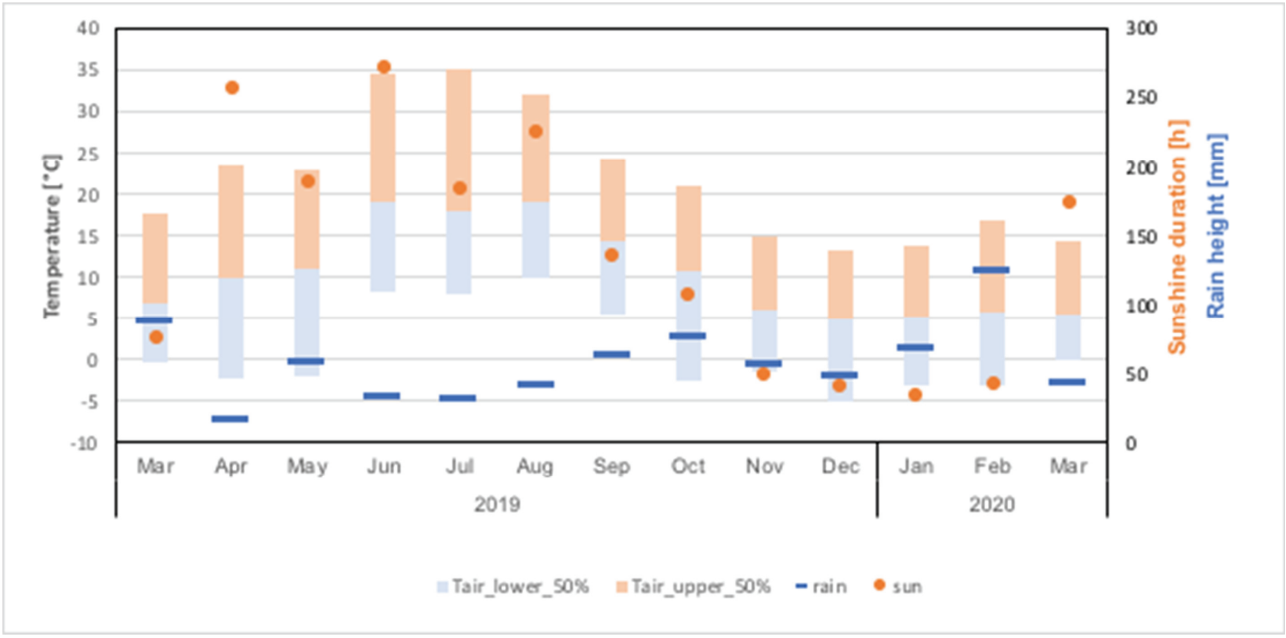


Figure 54: Weather data relevant to the harvesting period of the present study, from March 2019 to Mar 2020. Data sourced from German Meteorological Service, Deutscher Wetterdienst (DWD), station 1975 Hamburg-Fuhlsbüttel, with a temporal resolution of 10 minutes sampling interval. Scale bars indicate air temperature at 2 meters in height; data above the mean value are shown in yellow and those below are in light blue. Precipitation height and sunshine duration are shown as the monthly sum in dark blue lines and orange dots, respectively.

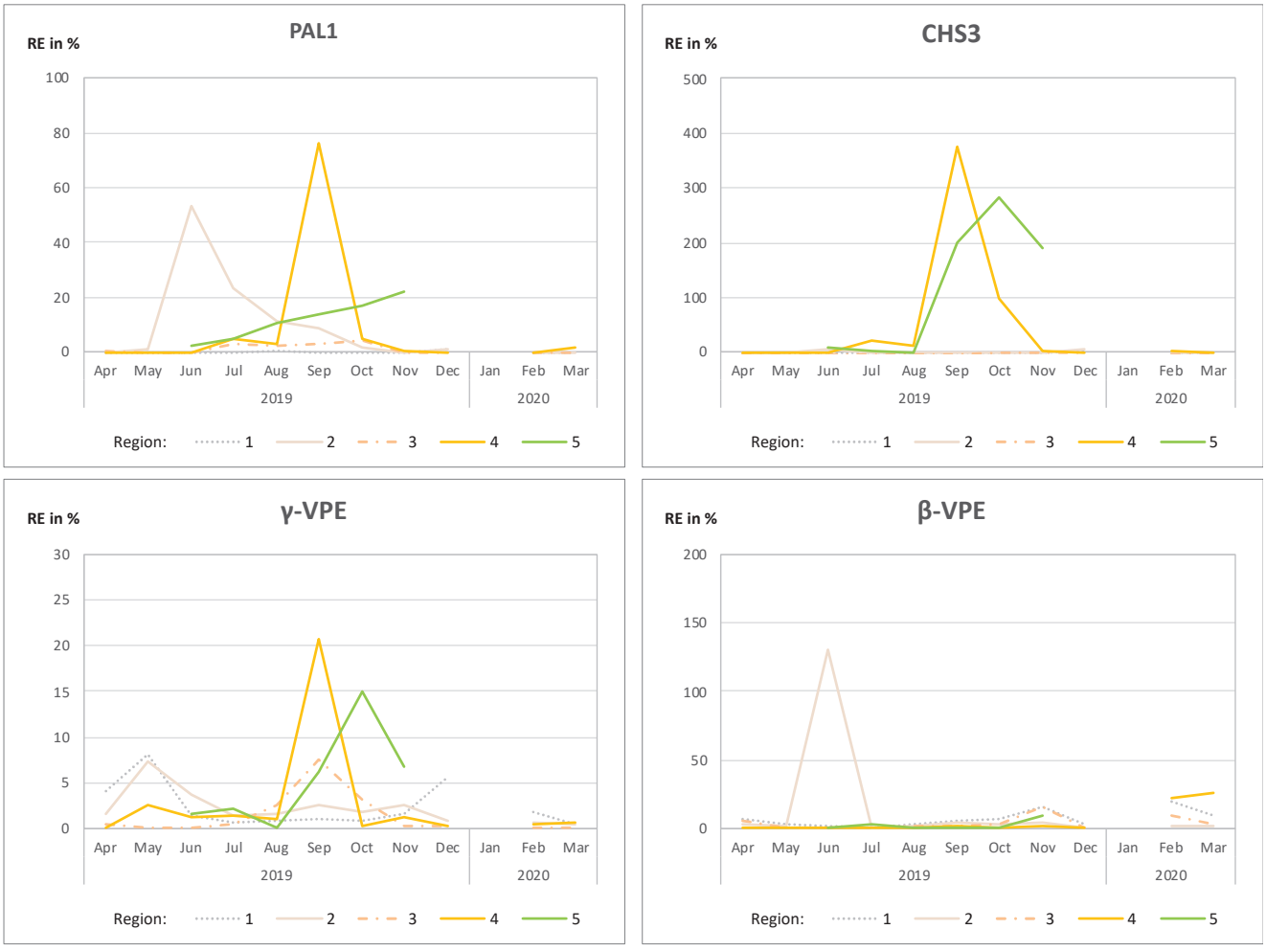


Figure 55: Relative gene expression (RE) of two heartwood formation marker-genes and vacuolar processing enzymes (VPEs) to the reference gene 18S in *Robinia pseudoacacia*. (Top Left) RE of Phenylalanine ammonia-lyase 1 (*PAL1*). A sharp increase occurred in Region 2 (differentiating xylem) in June, followed by a gradual decline through the summer months, an abrupt rise in September in Region 4 (innermost sapwood) and a gradual increase in Region 5 (transition zone) from June to November. (Top right) RE of Chalcone synthase 3 (*CHS3*). *CHS3* showed little observable expression in the outer wood regions, but sharp upregulations are observed in Regions 4 and 5 in the autumn months. ..(Bottom left) RE of *.γ-RpVPE*. Upregulation of the gene is observed in May in Regions 1 to 3. In Region 2, the upregulation is a month ahead of the elevated expression of *PAL1*. Its upregulation in Regions 4 and 5 occurred from September to November, and followed a similar pattern to that of *CHS3*. (Bottom right) RE of *β-RpVPE.* Its elevated expression in Region 2 occurred in June, coinciding with that of *PAL1*.

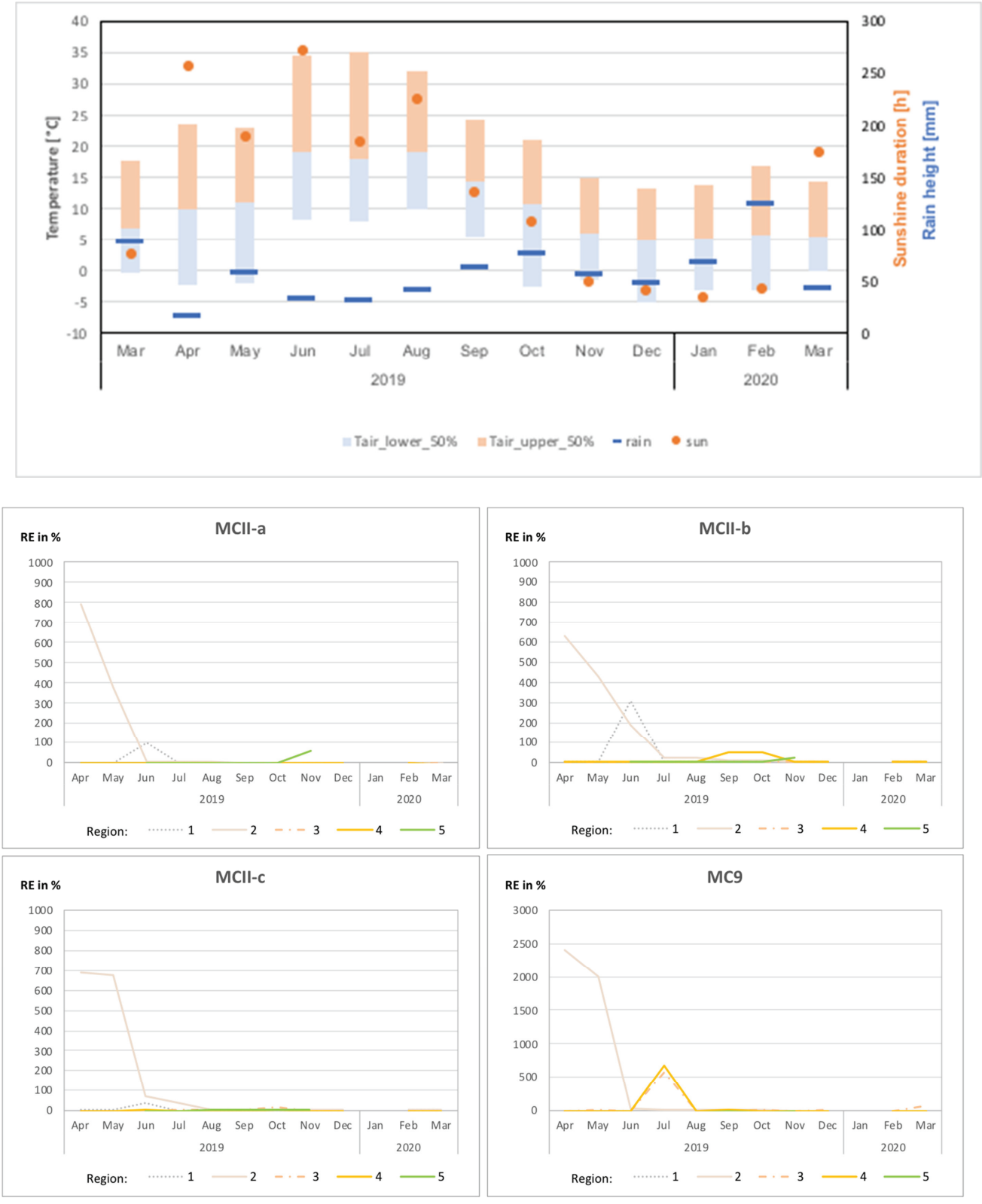


Figure 56: (Top) Weather data relevant to the harvesting period of the present study. Data sourced from German Meteorological Service, Deutscher Wetterdienst (DWD), station 1975 Hamburg-Fuhlsbüttel. Scale bars indicate air temperature at 2 meters in height; data above the mean value are shown in yellow and those below are in light blue. Precipitation height and sunshine duration are shown as the monthly sum in dark blue lines and orange dots, respectively. **(Bottom) Relative gene expression (RE) of four type-II metacaspases to the reference gene 18S in different wood regions.**

5. DISCUSSION

5.1 Limitations to the current experiment

Limited by the availability of samples and processing capacity within the timeframe of this project, only one biological replicate was harvested each month for the present study. Nonetheless, the results unprecedentedly offer an impression of the PCD marker genes expression profile across the living regions in heartwood forming stems with a monthly survey over an annual cycle. Repeating the experiment on a larger scale would be necessary for a solid conclusion. Besides biological replicates, increasing the number of technical replicates is also important for backing the reliability of the data set. In the current qPCR experiment, the sample amount in the set exceeds the well-plate capacity and must be separated onto multiple plates for individual runs. This incurs possible variations and errors between the reaction assays, such as variations between technical replicas and individually prepared qPCR reaction mixes. Although the variation has been accounted for via inter-plate calibration to a certain extent, the technical variations incurred by multiple plate runs should be considered for results interpretation. If the laboratory settings for follow-up experiments permit, using the 384-well plates instead of the 96-well ones could accommodate the complete samples set onto the same plate or minimize the number of separate runs, thus avoiding or lessening the issue.

Regarding the qPCR assay design, the probe-based method would be a better option in a follow-up experiment as it generates data with both the reference and multiple candidate genes in identical runs, which offer a direct comparison between the gene expression level, omitting the variation between the reaction master mixes in different runs and differences in the technical replicas. From the list of standard reference genes for gene-expression studies, 18S has been previously validated as the most suitable candidate, specifically regarding HWF (Lange, 2009). However, in the calculation for relative gene expression, a minuscule difference in the number of average Ct in the reference gene may result in magnified differences in the number for the percentage of relative expression in the candidate gene. Therefore, determining and including

more reference genes would greatly enhance the result reliability, as is scaling up the number of repeat runs.

The obtained genetic sequences do not cover the entire length of the genes in question, and their complete sequences are of great interest for follow-up studies. In addition to the more accurate phylogenic analysis, having their complete sequences with their promoter regions would allow the identification of their CREs and possible transcription factors, thereby gaining insight into their transcriptional regulation. Their 3'UTR of mRNAs is also of interest, as this region often contains regulatory sequences that influence gene expression post-transcriptionally (Ogorodnikov *et al.*, 2016). In addition to sequencing the remaining sequences for the candidate genes, the rest of the gene family members are excellent candidates for comparative investigation. MCs and VPEs are topics of priority in the expanding field of plant PCD, especially on how the gene family members have differentially expressed in specific cell types for development.

5.2 Results interpretation

5.2.1 Sequence analysis

With the protein sequences of MC4 and MC5 homologs retrieved from the database, phylogenetic analysis showed that these genes tend to share higher homology with each other in closely related species than they are with their own respective homologs. The signature amino acid residues that have been identified for MC4 in the literature also apply to MC5, rendering the exact identification of the three *RpMC-II* sequences a challenge. Despite that, their translated sequences cover the majority of key sites of the enzymes and allowed a decent comparison with their reference homologs, which have been investigated via experimental research. On the contrary, homologs of MC9 are distinctive in sequence identity among type-II MCs. The protein sequence alignment of RpMC9 with its homologs showed no remarkable difference; all the conserved residues matched the others, suggesting similar activation requirements

and enzymatic properties. However, analysis of the three RpMC-II sequences revealed an unexpected change of key residue in their protein peptide. One of the conserved aspartate residues predicted to contribute to calcium binding is substituted by glutamic acid in most legume species, and RpMC-IIc follows the trend. It is an alanine at this base with RpMC-IIa; a heterogeneous site in the DNA sequence of *RpMC-IIb* makes it either an alanine or a glutamic acid. The heterogeneous site may indicate two genes that were not separatable by the primers used, which warrants further investigation.

The two VPE types share less sequence similarity in comparison to MCs, and their identification presented no ambiguity. VPEs have been identified with different cellular functions, including the maturation of storage protein in the PSV, activation of lytic hydrolases that execute proteolytic cascade in PCD, formation of cyclotides through ligation, and regulation in gene expression as transcription factors (Dall & Brandstetter, 2016; Hara-Nishimura, 2013; Matarasso *et al.*, 2005). Protein sequence analysis indicates that both *RpVPEs* isolated in the present studies have the proteases-specific motif identified in Du *et al.* (2020) and other conserved residues matched with their respective homologs. Hence, the exclusion of their role as ligase can be assumed until proven otherwise, such as the presence of ligase activity. We do not have the available data to consider their possible role as a TF.

5.2.2 Gene expression analysis

PAL1 and *CHS3* are the pre-existing marker genes for HWF, and their expression patterns in the present results generally align with the extant literature. PAL is involved in the biosynthesis of lignin and a plethora of HWS that continue after the general phenylpropanoid pathway. As previous studies have shown, its increased activities occur during both vigorous lignifications in differentiating xylem and HWS biosynthesis at the TZ (Magel *et al.*, 1991). In the present results, we observed its elevated gene expression in the outermost xylem ring corresponding to the season of active xylogenesis and the inner regions during the heartwood forming season. On the other hand, CHS, the first structural enzyme in the flavonoid biosynthetic pathway,

shows no detectable gene expression except in the inner regions between August to November. Based on the weather data for the harvesting period, their up-regulation in the differentiating xylem coincides with the increase of both average temperature and sunshine duration, while the RE surge in the inner regions corresponds to the co-decline of these conditions.

At the sample preparation stage of the current study, we assumed Region 4 to be the inner sapwood and Region 5 the TZ based on the presence of fluorescence. However, the results of the reference marker genes strongly suggest that the ISW immediately adjacent to the florescent region was already in transition as the production of HWS on the molecular level has started. As the entire region transforming from sapwood to heartwood is the TZ by definition, it would be more appropriate to consider Regions 4 and 5 in the present experiment as outer TZ and inner TZ, respectively.

The RE patterns of *PAL1, CHS3, and γVPE* are particularly interesting when viewed in combination. The elevated expression of *PAL1* in early summer provides an indicating reference for vigorous lignin production during xylogenesis, and the result showed the co-expression with the VPEs. Although the rise of *γVPE* expression during TEs and fiber dPCD has been described in multiple studies (Cheng *et al.*, 2019; Courtois-Moreau *et al.*, 2009; Kinoshita *et al.*, 1999; Turner *et al.*, 2007), recording its upregulation preceding that of *PAL1* was out of our expectations. Furthermore, the co-elevation of *β-RpVPE* and *PAL1* was also not foreseeable, as the main function of *βVPE* is believed to lie in the PSV for protein maturation; its functions in the differentiating xylem await further investigation. The synchronized increased RE of *PAL1* and *CHS3* in the TZ in autumn points to HWS biosynthesis; the associated rise of *γVPE* RE observed in our results has also been reported in *Taiwania* (Yeh *et al.*, 2020) and *Robinia* (Yang *et al.*, 2003). While *γ-RpVPE* showed elevated RE before *PAL1* in the differentiating xylem, the same pattern is not observed in the TZ. As this could be an indication of how *γ-RpVPE* differentially behaves in specific xylem cell types, repeating the experiment for confirmation is warranted.

The elevated RE of *β-RpVPE* in the phloem, middle and inner sapwood in the late winter is intriguing; one possible explanation is its participation in the remobilizing

of cellular resources for the bud burst in spring. This notion also awaits further evidence for the proof of concept.

The co-expression of *PAL1*, *CHS3*, and γ-*RpVPE* in the outer TZ appeared one month ahead of time compared to that in the inner TZ. One way to interpret this result is that the transition spans more than one annual cycle, where a majority of XP cells convert into heartwood in the outer TZ starting after mid-summer, and the remaining XP cells transform in the year after and begin later in autumn. Previously demonstrated in gene expression profiling with Robinia (Yang *et al.,* 2003), over half of the DEGs in the TZ are signal transduction-related, suggesting that the cells located in the innermost part of the trunk are affected by external environmental conditions. Therefore, another interpretation is that the signals for XP cells to initiate transformation are triggered in the outer stem and slowly diffused to inner regions, hence such expression patterns within the transitioning region in the same year.

All surveyed type *RpMC-II*s showed conspicuous RE in the differentiating xylem in early spring, preceding those of *PAL1* and *VPE*s; their elevated expressions generally agree with the extant literature that type-II *MCs* are involved in xylogenesis. MCs are known to act upstream of Asp-specific proteases in the activation cascade of autolysis (Minina *et al.*, 2013; Tsiatsiani *et al.*, 2011), while VPEs are known to execute the caspase-like activities (Hatsugai & Hara-Nishimura, 2018). Our results showed that the transcription of *RpMC-II*s, thus the production of their zymogens, occurs ahead of time of the VPEs; such a pattern is not expected. For the proteolytic cascade to properly occur, all necessary components must be present in the form of zymogens in their respective subcellular location and ready for activation upon the arrival of the cell death triggering signals. Therefore, the timing difference in their elevated RE is curious.

Upregulated RE of *RpMC9* in Region 3 and 4 are nearly identical, which occurred in July 2019. This left doubt on the originating cell type of the data; it can indicate its role in secondary cell wall development in fibers, which would align with previous findings. Based on the assumption that this upregulated expression belongs to fibers, the results suggest that *RpMC9* has no role in HWF; the deposition of HWS and their

copolymerization into the cell wall is facilitated by other means. Furthermore, its lack of involvement in the terminal differentiation of XP cells implies that the activities of MC9 do not apply to all lineage of xylem cell types. Instead, it seems highly specific to water-conducting xylem cells (tracheids in gymnosperm and TEs in angiosperms) and the sibling cell type that is ontologically close (fibers in angiosperms). However, the specimens were dissected by wood regions in the present experiment, thus we cannot be certain about the cell type the gathered data represent. For confirmation in follow-up studies, separating cell types with techniques such as laser micro-dissection would clarify the matter.

RpMC-IIa and *-b* also showed elevated RE in the inner regions, and distinctively in the TZ. Similar to that of *PAL1*, *CHS3*, and γ*-RpVPE,* elevated RE of *RpMC-IIb* occurred in the outer TZ before the rise in the inner TZ. In contrast to the pattern observed in the differentiating xylem, elevated RE of *RpMC-IIa* and *-b* appeared later in the year in the TZ compared to the other up-regulated genes. The results calls for experimental repeat for confirmations. Moreover, the three MC4-like *RpMC-IIs* showed a high co-expression in the phloem in June 2019, suggesting their possible involvement in the differentiation of phloem cells.

6. CONCLUSIONS

Although these findings need to be substantiated at the phenotypic level before concrete conclusions can be drawn, the present study offers a preliminary view of the expression patterns of the marker genes that define the two essential aspects of HWF, HWS biosynthesis and dPCD, over an annual cycle. It appears not redundant to emphasize that gene expression studies at the transcription level do not tell the story of the activated final gene products. Each candidate gene investigated is a PCD regulatory protease synthesized as inactive zymogens, and they may be accumulated and stored in cellular compartments until their activation, which is post-translationally regulated. The timing and level of up-regulated gene expression reflect the relatively higher amount of transcripts present in the cell, which indicates that the use of such gene products is likely to be imminent. The candidate genes had not been investigated in the context of HWF or in xylem parenchyma cells until now; hence, further investigations are needed to clarify their roles in the process.

With the overview of PCD regulatory protease expression patterns across the seasons and regions, one follow-up question is about their transcriptional regulation and initiation triggering signals. Gene expression profiling on HWF in gymnosperm trees reported the upregulation of MYB-like and NAC domain TFs (Lim *et al.*, 2016; Yeh *et al.*, 2020). Identifying the CREs and TFs that mediate the transcription of PCD regulatory proteases in angiosperms would present useful connecting dots. Furthermore, previous works proposed the possible involvement of ethylene signaling; the mechanism has remained a knowledge gap, which makes the mediation of *VPE* on ethylene production as a TF particularly intriguing.

The mechanism behind HWF has remained largely unknown, and a consensus on the definition of the TZ at the cellular and molecular level is also in dispute (Nakada & Fukatsu, 2012). With advancing technology and commercially available laboratory aids, methods in molecular biology are reliable for unveiling highly complex cellular processes. As high-throughput methods, such as next-generation sequencing and multi-omics studies, become more readily available, setting a reference point on the timeline for the various cellular processes of HWF would greatly assist our further

understanding. HWF is defined as converting and depositing reserve materials into bioactive substances and the cessation of containing living cells. However, pinpointing the exact timing for the biosynthesis of HWS appears to be a confusing pursuit; there are remarkable variations in HWS depending on species and between individual trees within a species (Taylor *et al.* 2002, Nakada 2007, Bito *et al.* 2011, Bush *et al.* 2011). Moreover, among the composition of HWS in each species, there is no consensus on the method or criteria for assigning the signatures molecules. Regarding the deposition of HWS, there is even less understanding of its mechanism and the diversity of quantification methods renders comparisons irrelevant. On the contrary, dPCD in plants has a defined timeline, with vacuolar collapse marking the moment of cell death. Furthermore, it has been well-established in the extant literature that γVPE play a protagonist role in tonoplast rupture. Despite the limited data available in the present study, the results showed the co-expression patterns of *γVPE* with *PAL1* and *CHS3* that fit into the HWF timeline and regions, reflecting the two essential aspects that define the process. Hence, we propose the elevated expression of *γVPE* as an additional molecular marker to set a reference point on the timeline of HWF for comparing the various cellular events during the differentiation process.

Terminal differentiation is an intricate yet dramatic event for the cell; multiple major biochemical processes and structural changes are at play simultaneously. Its regulation occurs on multiple levels, and transcriptional expression of regulatory proteases is among the many. Similar to xylogenesis, HWF involves the orchestrated activation of over a thousand genes and various molecular signaling pathways. While it is generally agreed upon that weather conditions contribute to the seasonal initiation of HWF, the exact correlation is still a conundrum, as data on the phenological effects has remained scarce. Therefore, the accompanying weather data in future studies will be highly beneficial for identifying their correlation patterns. The mechanism behind HWF implies enormous values beyond durable timers. Under the influence of global climate change, there is an increasing urgency to understand how forests adapt to environmental stress, and HWF is a sophisticated defense strategy in higher plants that enhances tree performance and sustains longevity.

7. REFERENCES

Abeles, F. B., Morgan, P. W., & Saltveit, M. E. (Mikal E. (1992). *Ethylene in plant biology*. Academic Press. http://www.sciencedirect.com:5070/book/9780080916286/ethylene-in-plant-biology

Amborella Genome Project. (2013). The Amborella genome and the evolution of flowering plants. *Science*, *342*(6165). https://doi.org/10.1126/science.1241089

Améglio, T., Decourteix, M., Alves, G., Valentin, V., Sakr, S., Julien, J. L., Petel, G., Guilliot, A., & Lacointe, A. (2004). Temperature effects on xylem sap osmolarity in walnut trees: evidence for a vitalistic model of winter embolism repair. *Tree Physiology*, *24*(7), 785–793. https://doi.org/10.1093/TREEPHYS/24.7.785

Andersson-Gunnerås, S., Mellerowicz, E. J., Love, J., Segerman, B., Ohmiya, Y., Coutinho, P. M., Nilsson, P., Henrissat, B., Moritz, T., & Sundberg, B. (2006). Biosynthesis of cellulose-enriched tension wood in Populus: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *The Plant Journal*, *45*(2), 144–165. https://doi.org/10.1111/J.1365-313X.2005.02584.X

Bamber, R. K. (1976). Heartwood, its function and formation. *Wood Science and Technology*. https://doi.org/10.1007/BF00376379

Barros, J., Serk, H., Granlund, I., & Pesquet, E. (2015). The cell biology of lignification in higher plants. *Annals of Botany*, *115*(7), 1053–1074. https://doi.org/10.1093/aob/mcv046

Bartels, S., & Boller, T. (2015). Quo vadis, Pep? Plant elicitor peptides at the crossroads of immunity, stress, and development. *Journal of Experimental Botany*, *66*(17), 5183–5193. https://doi.org/10.1093/JXB/ERV180

Baskett, J. A. (2012). *A Type II Metacaspase Interacts with RPS1-K-2 in Soybean and Analysis of the Soybean Metacaspase Gene Family*.

Batistič, O., & Kudla, J. (2012). Analysis of calcium signaling pathways in plants. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1820*(8), 1283–1293. https://doi.org/10.1016/J.BBAGEN.2011.10.012

Becker, C., Shutov, A. D., Nong, V. H., Senyuk, V. I., Jung, R., Horstmann, C., Fischer, J., Nielsen, N. C., & Müntz, K. (1995). Purification, cDNA Cloning and Characterization of Proteinase B, an Asparagine-specific Endopeptidase from Germinating Vetch (Vicia sativa L.) Seeds. *European Journal of Biochemistry*, *228*(2), 456–462. https://doi.org/10.1111/J.1432-1033.1995.0456N.X

Bennett, R. N., & Wallsgrove, R. M. (1994). Secondary metabolites in plant defence mechanisms. *New Phytologist*, *127*(4), 617–633. https://doi.org/10.1111/j.1469-8137.1994.tb02968.x

Berghe, T. Vanden, Vanlangenakker, N., Parthoens, E., Deckers, W., Devos, M., Festjens, N., Guerin, C. J., Brunk, U. T., Declercq, W., & Vandenabeele, P. (2009). Necroptosis, necrosis and secondary necrosis converge on similar cellular disintegration features. *Cell Death & Differentiation 2010 17:6*, *17*(6), 922–930. https://doi.org/10.1038/cdd.2009.184

Bergström, B. (2003). Chemical and structural changes during heartwood formation in Pinus sylvestris. *Forestry: An International Journal of Forest Research*, *76*(1), 45–53. https://doi.org/10.1093/FORESTRY/76.1.45

Bergström, B., Gustafsson, G., Gref, R., & Ericsson, A. (1999). Seasonal changes of pinosylvin distribution in the sapwood/heartwood boundary of Pinus sylvestris. *Trees 1999 14:2*, *14*(2), 65–71. https://doi.org/10.1007/PL00009754

Beritognolo, I., Magel, E., Abdel-Latif, A., Charpentier, J.-P. P., Jay-Allemand, C., & Breton, C. (2002). Expression of genes encoding chalcone synthase, flavanone 3-hydroxylase and dihydroflavonol 4-reductase correlates with flavanol accumulation during heartwood formation in Juglans nigra. *Tree Physiology*, *22*(5), 291–300. https://doi.org/10.1093/treephys/22.5.291

Biggs, A. R. (1987). Occurrence and Location of Suberin in Wound Reaction Zones in Xylem of 17 Tree Species. *Phytopathology*, *77*(5), 718. https://doi.org/10.1094/PHYTO-77-718

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin Biosynthesis. *Annual Review of Plant Biology*, *54*, 519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Boller, T. (1991). Ethylene in Pathogenesis and Disease Resistance. In *The Plant Hormone Ethylene* (1st ed., p. 22). CRC Press. https://doi.org/10.1201/9781351075763-16

Bollhöner, B. (2013). *Significance of hydrolytic enzymes expressed during xylem cell death*. http://umu.diva-portal.org/

Bollhöner, B., Jokipii-Lukkari, S., Bygdell, J., Stael, S., Adriasola, M., Muñiz, L., Van Breusegem, F., Ezcurra, I., Wingsle, G., & Tuominen, H. (2018). The function of two type II metacaspases in woody tissues of Populus trees. *New Phytologist*, *217*(4), 1551–1565. https://doi.org/10.1111/nph.14945

Bollhöner, B., Prestele, J., & Tuominen, H. (2012). Xylem cell death: Emerging understanding of regulation and function. *Journal of Experimental Botany*, *63*(3), 1081–1094. https://doi.org/10.1093/jxb/err438

Bollhöner, B., Zhang, B., Stael, S., Denancé, N., Overmyer, K., Goffner, D., Van Breusegem, F., & Tuominen, H. (2013). Post mortem function of AtMC9 in xylem vessel elements. *New Phytologist*. https://doi.org/10.1111/nph.12387

Bozhkov, P. V., Suarez, M. F., Filonova, L. H., Daniel, G., Zamyatnin, A. A., Rodriguez-Nieto, S., Zhivotovsky, B., & Smertenko, A. (2005). Cysteine protease mcII-Pa executes programmed cell death during plant embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(40), 14463–14468. https://doi.org/10.1073/pnas.0506948102

Brodersen, C. R. (2013). Visualizing wood anatomy in three dimensions with high-resolution X-ray micro-tomography (MCT) - A review. *IAWA Journal*, *34*(4), 408–424. https://doi.org/10.1163/22941932-00000033

Brodersen, C. R., & McElrone, A. J. (2013). Maintenance of xylem network transport capacity: A review of embolism repair in vascular plants. *Frontiers in Plant Science*, *4*(APR), 1–11. https://doi.org/10.3389/fpls.2013.00108

Carlquist, S. (2018). Living Cells in Wood 3. Overview; Functional Anatomy of the Parenchyma Network. *Botanical Review*, *84*(3), 242–294. https://doi.org/10.1007/S12229-018-9198-5/TABLES/1

Carmona-Gutierrez, D., Fröhlich, K. U., Kroemer, G., & Madeo, F. (2010). Editorial: Metacaspases are caspases. Doubt no more. In *Cell Death and Differentiation* (Vol. 17, Issue 3, pp. 377–378). https://doi.org/10.1038/cdd.2009.198

Celedon, J. M., & Bohlmann, J. (2018). An extended model of heartwood secondary metabolism informed by functional genomics. In *Tree Physiology* (Vol. 38, Issue 3). https://doi.org/10.1093/treephys/tpx070

Celedon, J. M., Chiang, A., Yuen, M. M. S., Diaz-Chavez, M. L., Madilao, L. L., Finnegan, P. M., Barbour, E. L., & Org Bohlmann, J. (2016). *Heartwood-specific transcriptome and metabolite signatures of tropical sandalwood (Santalum album) reveal the final step of (Z)-santalol fragrance biosynthesis*. https://doi.org/10.1111/tpj.13162

Chaffey, N., & Barlow, P. (2001). The cytoskeleton facilitates a three-dimensional symplasmic continuum in the long-lived ray and axial parenchyma cells of angiosperm trees. *Planta 2001 213:5*, *213*(5), 811–823. https://doi.org/10.1007/S004250100560

Chappell, J., Hahlbrock, K., & Boller, T. (1984). Rapid induction of ethylene biosynthesis in cultured parsley cells by fungal elicitor and its relationship to the induction of phenylalanine ammonia-lyase. *Planta*, *161*, 475–480.

Chattaway, M. M. (1949). The development of tyloses and secretion of gum in heartwood formation. *Australian Journal of Biological Sciences*, *2*(3), 227–240. https://doi.org/10.1071/BI9490227

Cheng, Z., Zhang, J., Yin, B., Liu, Y., Wang, B., Li, H., & Lu, H. (2019). γVPE plays an important role in programmed cell death for xylem fiber cells by activating protease CEP1 maturation in Arabidopsis thaliana. *International Journal of Biological Macromolecules*, *137*, 703–711. https://doi.org/10.1016/j.ijbiomac.2019.07.017

Choat, B., & Pittermann, J. (2009). New insights into bordered pit structure and cavitation resistance in angiosperms and conifers. *New Phytologist*, *182*(3), 557–560. https://doi.org/10.1111/J.1469-8137.2009.02847.X

Cierjacks, A., Kowarik, I., Joshi, J., Hempel, S., Ristow, M., von der Lippe, M., & Weber, E. (2013). Biological flora of the british isles: Robinia pseudoacacia. *Journal of Ecology*, *101*(6), 1623–1640. https://doi.org/10.1111/1365-2745.12162

Clearwater, M. J., & Goldstein, G. (2005). Embolism Repair and Long Distance Water Transport. *Vascular Transport in Plants*, 375–399. https://doi.org/10.1016/B978-012088457-5/50020-4

Cohen, G. M. (1997). Caspases: The executioners of apoptosis. In *Biochemical Journal* (Vol. 326, Issue 1, pp. 1–16). Portland Press Ltd. https://doi.org/10.1042/bj3260001

Coll, N. S., Epple, P., & Dangl, J. L. (2011). Programmed cell death in the plant immune system. *Cell Death and Differentiation*, *18*(8), 1247–1256. https://doi.org/10.1038/cdd.2011.37

Conaway, R. C., & Conaway, J. W. (1997). General Transcription Factors for RNA Polymerase II. *Progress in Nucleic Acid Research and Molecular Biology*, *56*(C), 327–346. https://doi.org/10.1016/S0079-6603(08)61009-0

Conlan, B. F., & Anderson, M. A. (2012). Circular Micro-Proteins and Mechanisms of Cyclization. *Current Pharmaceutical Design*, *17*(38), 4318–4328. https://doi.org/10.2174/138161211798999410

Corbin, D. R., Sauer, N., & Lamb, C. J. (1987). Differential regulation of a hydroxyproline-rich glycoprotein gene family in wounded and infected plants. *Molecular and Cellular Biology*, *7*(12), 4337–4344. https://doi.org/10.1128/MCB.7.12.4337-4344.1987

Courtois-Moreau, C. L., Pesquet, E., Sjödin, A., Muñiz, L., Bollhöner, B., Kaneda, M., Samuels, L., Jansson, S., & Tuominen, H. (2009). A unique program for cell death in xylem fibers of Populus stem. *Plant Journal*, *58*(2), 260–274. https://doi.org/10.1111/j.1365-313X.2008.03777.x

Craik, D. J. (2012). Host-defense activities of cyclotides. *Toxins*, *4*(2), 139–156. https://doi.org/10.3390/toxins4020139

Creux, N. M., Ranik, M., Berger, D. K., & Myburg, A. A. (2008). Comparative analysis of orthologous cellulose synthase promoters from Arabidopsis, Populus and Eucalyptus: evidence of conserved regulatory elements in angiosperms. *The New Phytologist*, *179*(3), 722–737. https://doi.org/10.1111/J.1469-8137.2008.02517.X

D'Mello, S. R. (2007). Plant Programmed Cell Death. In *Medical Cell Biology: Third Edition*. https://doi.org/10.1016/B978-0-12-370458-0.50015-3

Dall, E., & Brandstetter, H. (2013). Mechanistic and structural studies on legumain explain its zymogenicity, distinct activation pathways, and regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 10940–10945. https://doi.org/10.1073/pnas.1300686110

Dall, E., & Brandstetter, H. (2016). Structure and function of legumain in health and disease. In *Biochimie* (Vol. 122, pp. 126–150). Elsevier B.V. https://doi.org/10.1016/j.biochi.2015.09.022

De Filippis, L., & Magel, E. (1998). Differences in genomic DNA extracted from bark and from wood of different zones in Robinia trees using RAPD-PCR. *Trees 1998 12:6*, *12*(6), 377–384. https://doi.org/10.1007/PL00009723

Dharmawardhana, P., Brunner, A. M., & Strauss, S. H. (2010). Genome-wide transcriptome analysis of the transition from primary to secondary stem development in Populus trichocarpa. *BMC Genomics*, *11*(1), 1–19. https://doi.org/10.1186/1471-2164-11-150

Dixon, H. H., & Joly, J. (1895). On the Ascent of Sap on JSTOR. *Philosophical Transactions of the Royal Society of London. B*, *186*, 563–576. https://www.jstor.org/stable/91804

Domec, J. C., Lachenbruch, B., Meinzer, F. C., Woodruff, D. R., Warren, J. M., & McCulloh, K. A. (2008). Maximum height in a conifer is associated with conflicting requirements for xylem design. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(33), 12069–12074. https://doi.org/10.1073/PNAS.0710418105/SUPPL_FILE/0710418105SI.PDF

Du, J., Yap, K., Chan, L. Y., Rehm, F. B. H., Looi, F. Y., Poth, A. G., Gilding, E. K., Kaas, Q., Durek, T., & Craik, D. J. (2020). A bifunctional asparaginyl endopeptidase efficiently catalyzes both cleavage and cyclization of cyclic trypsin inhibitors. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-020-15418-2

Dünisch, O., Richter, H. G., & Koch, G. (2010). Wood properties of juvenile and mature heartwood in Robinia pseudoacacia L. *Wood Science and Technology*, *44*(2), 301–313. https://doi.org/10.1007/s00226-009-0275-0

Dutton, J. L., Renda, R. F., Waine, C., Clark, R. J., Daly, N. L., Jennings, C. V., Anderson, M. A., & Craik, D. J. (2004). Conserved structural and sequence elements implicated in the processing of gene-encoded circular proteine. *Journal of Biological Chemistry*, *279*(45), 46858–46867. https://doi.org/10.1074/jbc.M407421200

Eckhart, L., Lippens, S., Tschachler, E., & Declercq, W. (2013). Cell death by cornification. *Biochimica et Biophysica Acta*, *1833*(12), 3471–3480. https://doi.org/10.1016/J.BBAMCR.2013.06.010

El Meanawy, M. A., Aji, T., Phillips, N. F. B., Davis, R. E., Salata, R. A., Malhotra, I., McClain, D., Aikawa, M., & Davis, A. H. (1990). Definition of the Complete Schistosoma mansoni Hemoglobinase mRNA Sequence and Gene Expression in Developing Parasites. *The American Journal of Tropical Medicine and Hygiene*, *43*(1), 67–78. https://doi.org/10.4269/AJTMH.1990.43.67

Escamez, S., & Tuominen, H. (2014). Programmes of cell death and autolysis in tracheary elements: When a suicidal cell arranges its own corpse removal. *Journal of Experimental Botany*, *65*(5), 1313–1321. https://doi.org/10.1093/jxb/eru057

Essiamah, S., & Eschrich, W. (1985). Changes of Starch Content in the Storage Tissues of Deciduous Trees During Winter and Spring. *IAWA Journal*, *6*(2), 97–106. https://doi.org/10.1163/22941932-90000921

Evert, R. F. (2006). Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function, and Development. In *Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function, and Development: Third Edition* (3rd ed.). John Wiley and Sons. https://doi.org/10.1002/0470047380

Ferguson, A. R., Eiseman, J. A., & Leonard, J. A. (1983). Xylem Sap from Actinidia chinensis: seasonal Changes in Composition. *Annals of Botany*, *51*(6), 823–833. https://doi.org/10.1093/OXFORDJOURNALS.AOB.A086533

Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research*, *39*(SUPPL. 2), 28–37. https://doi.org/10.1093/nar/gkr367

Fischer, J., Becker, C., Hillmer, S., Horstmann, C., Neubohn, B., Schlereth, A., Senyuk, V., Shutov, A., & Müntz, K. (2000). The families of papain- and legumain-like cysteine proteinases from embryonic axes and cotyledons of Vicia seeds: developmental patterns, intracellular localization and functions in globulin proteolysis. *Plant Molecular Biology*, *43*(1), 83–101. https://doi.org/10.1023/A:1006456615373

Floyd, B. E., Pu, Y., Soto-Burgos, J., & Bassham, D. C. (2015). To Live or Die: Autophagy in Plants. In *Plant Programmed Cell Death* (pp. 369–300). Springer, Cham. https://doi.org/10.1007/978-3-319-21033-9_11

Fortin, J., & Lam, E. (2018). Domain swap between two type-II metacaspases defines key elements for their biochemical properties. *Plant Journal*, *96*(5), 921–936. https://doi.org/10.1111/tpj.14079

Frey-Wyssling, A., & Bosshard, H. H. (1959). Cytology of the Ray Cells in Sapwood and Heartwood. *Holzforschung*, *13*(November), 129–137. https://doi.org/10.1515/HFSG.1959.13.5.129

Friedman, W. E. (1994). The evolution of embryogeny in seed plants and the developmental origin and early history of endosperm. *American Journal of Botany*, *81*(11), 1468–1486. https://doi.org/10.1002/J.1537-2197.1994.TB15633.X

Fromard, L., Babin, V., Fleurat-Lessard, P., Fromont, J. C., Serrano, R., & Bonnemain, J. L. (1995). *Control of Vascular Sap pH by the Vessel-Associated Cells in Woody Species (Physiological and Immunological Studies)*. *108*(3), 913–918. https://doi.org/10.1104/PP.108.3.913

Fujita, M., Nakagawa, K., Mori, N., & Harada, H. (1978). The season of tylosis development and changes in parenchyma cell structure in Robinia pseudoacacia L. *BULLETIN OF THE KYOTO UNIVERSITY FORESTS*, 6. http://hdl.handle.net/2433/191652

Fukuda, H. (1996). *XYLOGENESIS: INITIATION, PROGRESSION, AND CELL DEATH*. 299–325.

Funada, R., Yamagishi, Y., Begum, S., Kudo, K., Nabeshima, E., Nugroho, W. D., Hasnat, R., Oribe, Y., & Nakaba, S. (2016). Xylogenesis in Trees: From Cambial Cell Division to Cell Death. In *Secondary Xylem Biology: Origins, Functions, and Applications* (pp. 25–43). Elsevier Inc. https://doi.org/10.1016/B978-0-12-802185-9.00002-4

Galluzzi, L., Maiuri, M. C., Vitale, I., Zischka, H., Castedo, M., Zitvogel, L., & Kroemer, G. (2007). Cell death modalities: Classification and pathophysiological implications. *Cell Death and Differentiation*, *14*(7), 1237–1243. https://doi.org/10.1038/SJ.CDD.4402148

Galluzzi, L., Vitale, I., Aaronson, S. A., Abrams, J. M., Adam, D., Agostinis, P., Alnemri, E. S., Altucci, L., Amelio, I., Andrews, D. W., Annicchiarico-Petruzzelli, M., Antonov, A. V., Arama, E., Baehrecke, E. H., Barlev, N. A., Bazan, N. G., Bernassola, F., Bertrand, M. J. M., Bianchi, K., … Kroemer, G. (2018). Molecular mechanisms of cell death: Recommendations of the Nomenclature Committee on Cell Death 2018. In *Cell Death and Differentiation*. https://doi.org/10.1038/s41418-017-0012-4

Gatehouse, J. A., Bown, D., Evans, I. M., Gatehouse, L. N., Jobes, D., Preston, P., & Croy, R. R. D. (1987). Sequence of the seed lectin gene from pea (Pisum sativum L.). *Nucleic Acids Research*, *15*(18), 7642. https://doi.org/10.1093/NAR/15.18.7642

Gerry, E. (1914). *TYLOSES: THEIR OCCURRENCE AND PRACTICAL SIGNIFICANCE IN SOME AMERICAN WOODS*.

Gruis, D., Selinger, D. A., Curran, J. M., & Jung, R. (2002). Redundant Proteolytic Mechanisms Process Seed Storage Proteins in the Absence of Seed-Type Members of the Vacuolar Processing Enzyme Family of Cysteine Proteases. *The Plant Cell*, *14*(11), 2863–2882. https://doi.org/10.1105/TPC.005009

Guiboileau, A., Sormani, R., Meyer, C., & Masclaux-Daubresse, C. (2010). Senescence and death of plant organs: nutrient recycling and developmental regulation. *Comptes Rendus Biologies*, *333*(4), 382–391. https://doi.org/10.1016/J.CRVI.2010.01.016

Han, J., Li, H., Yin, B., Zhang, Y., Liu, Y., Cheng, Z., Liu, D., & Lu, H. (2019). The papain-like cysteine protease CEP1 is involved in programmed cell death and secondary wall thickening during xylem development in Arabidopsis. *Journal of Experimental Botany*, *70*(1), 205–215. https://doi.org/10.1093/jxb/ery356

Hara-Nishimura, I. (2013). Plant Legumain, Asparaginyl Endopeptidase, Vacuolar Processing Enzyme. *Handbook of Proteolytic Enzymes*, *2*, 2314–2320. https://doi.org/10.1016/B978-0-12-382219-2.00517-2

Hara-Nishimura, I., & Hatsugai, N. (2011). The role of vacuole in plant cell death. *Cell Death and Differentiation*, *18*, 1298–1304. https://doi.org/10.1038/cdd.2011.70

Hara-Nishimura, I., Hatsugai, N., Nakaune, S., Kuroyanagi, M., & Nishimura, M. (2005). Vacuolar processing enzyme: an executor of plant cell death. *Curr Opin Plant Biol*, *8*(4), 404–408. https://doi.org/10.1016/j.pbi.2005.05.016

Hara-Nishimura, I., Inoue, K., & Nishimura, M. (1991). A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. *FEBS Lett*, *294*(1–2), 89–93. https://doi.org/10.1016/0014-5793(91)81349-d

Hara-Nishimura, I., Kinoshita, T., Hiraiwa, N., & Nishimura, M. (1998). Vacuolar processing enzymes in protein-storage vacuoles and lytic vacuoles. *Journal of Plant Physiology*, *152*(6), 668–674. https://doi.org/10.1016/S0176-1617(98)80028-X

Hara-Nishimura, I., Takeuchi, Y., & Nishimura, M. (1993). Molecular characterization of a vacuolar processing enzyme related to a putative cysteine proteinase of Schistosoma mansoni. *The Plant Cell*, *5*(11), 1651–1659. https://doi.org/10.1105/tpc.5.11.1651

Hartmann, U., Sagasser, M., Mehrtens, F., Stracke, R., & Weisshaar, B. (2005). Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. *Plant Molecular Biology*, *57*(2), 155–171. https://doi.org/10.1007/s11103-004-6910-0

Hatsugai, N., & Hara-Nishimura, I. (2018). Measurement of the Caspase-1-Like Activity of Vacuolar Processing Enzyme in Plants. In *Methods Mol Biol* (2018/01/15, pp. 163–171). https://doi.org/10.1007/978-1-4939-7668-3_15

Hatsugai, N., Kuroyanagi, M., Yamada, K., Meshi, T., Tsuda, S., Kondo, M., Nishimura, M., & Hara-Nishimura, I. (2004). A plant vacuolar protease, VPE, mediates, virus-induced hypersensitive cell death. *Science*, *305*(5685), 855–858. https://doi.org/10.1126/science.1099859

Hatsugai, N., Yamada, K., Goto-Yamada, S., Hara-Nishimura, I., Noriyuki Hatsugai, Kenji Yamada, ShinoGoto-Yamada, & IkukoHara-Nishimura. (2015). Vacuolar processing enzyme in plant programmed cell death. *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2015.00234

Hauch, S., & Magel, E. (1998). Extractable activities and protein content of sucrose-phosphate synthase, sucrose synthase and neutral invertase in trunk tissues of Robinia pseudoacacia L. are related to cambial wood production and heartwood formation. *Planta*, *207*(2), 266–274. https://doi.org/10.1007/s004250050482

Haywood, J., James, A. M., & Mylne, J. S. (2018). Macrocyclization by asparaginyl endopeptidases. *New Phytologist*, *218*(3), 923–928. https://doi.org/10.1111/nph.14511

He, R., Drury, G. E., Rotari, V. I., Gordon, A., Willer, M., Farzaneh, T., Woltering, E. J., & Gallois, P. (2008). Metacaspase-8 modulates programmed cell death induced by ultraviolet light and H2O2 in Arabidopsis. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M704185200

Helm, M., Schmid, M., Hierl, G., Terneus, K., Tan, L., Lottspeich, F., Kieliszewski, M. J., & Gietl, C. (2008). KDEL-tailed cysteine endopeptidases involved in programmed cell death, intercalation of new cells, and dismantling of extensin scaffolds. *American Journal of Botany*, *95*(9), 1049–1062. https://doi.org/10.3732/ajb.2007404

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlén, M., Teeri, T. T., Lundeberg, J., Sundberg, B., Nilsson, P., & Sandberg, G. (2001). A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(25), 14732–14737. https://doi.org/10.1073/pnas.261293398

Higuchi, T. (1997). *Biochemistry and Molecular Biology of Wood*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-60469-0

Higuchi, T., Fukazawa, K., & Shimada, M. (1967). Biochemical Studies on the Heartwood Formation. *Hokkaido University Collection of Scholarly and Academic Papers*.

Hillis, W. E. (1968). Chemical aspects of heartwood formation. *Wood Science and Technology*, *2*(4), 241–259. https://doi.org/10.1007/BF00350271

Hillis, W. E. (1977). Secondary Changes in Wood. *The Structure, Biosynthesis, and Degradation of Wood*, 247–309. https://doi.org/10.1007/978-1-4615-8873-3_7

Hillis, W. E. (1987). *Heartwood and Tree Exudates*. *4*. https://doi.org/10.1007/978-3-642-72534-0

Hiraiwa, N., Nishimura, M., & Hara-Nishimura, I. (1997). Expression and activation of the vacuolar processing enzyme in Saccharomyces cerevisiae. *Plant Journal*. https://doi.org/10.1046/j.1365-313X.1997.12040819.x

Hofius, D., Munch, D., Bressendorff, S., Mundy, J., & Petersen, M. (2011). Role of autophagy in disease resistance and hypersensitive response-associated cell death. *Cell Death and Differentiation*, *18*(8), 1257–1262. https://doi.org/10.1038/cdd.2011.43

Holbrook, N. M., Zwieniecki, M. A., & Melcher, P. J. (2002). The Dynamics of “Dead Wood”: Maintenance of Water Transport Through Plant Stems. *Integrative and Comparative Biology*, *42*(3), 492–496. https://doi.org/10.1093/ICB/42.3.492

Hoson, T. (1991). Structure and Function of Plant Cell Walls: Immunological Approaches. *International Review of Cytology*, *130*(C), 233–268. https://doi.org/10.1016/S0074-7696(08)61505-3

Huffaker, A., Pearce, G., & Ryan, C. A. (2006). An endogenous peptide signal in Arabidopsis activates components of the innate immune response. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(26), 10098–10103. https://doi.org/10.1073/PNAS.0603727103/ASSET/0459E184-8E57-43EE-B993-7A75C44053F2/ASSETS/GRAPHIC/ZPQ0250626070005.JPEG

Iakimova, E. T., & Woltering, E. J. (2017). Xylogenesis in zinnia (Zinnia elegans) cell cultures: unravelling the regulatory steps in a complex developmental programmed cell death event. In *Planta* (Vol. 245, Issue 4). https://doi.org/10.1007/s00425-017-2656-1

IAWA. (1964). *Multilingual Glossary Of Terms Used In Wood Anatomy* (Vol. 40).

Imaseki, H. (1991). The Biochemistry of Ethylene Biosynthesis. In *The Plant Hormone Ethylene* (1st ed.). CRC Press. https://doi.org/10.1201/9781351075763-1/BIOCHEMISTRY-ETHYLENE-BIOSYNTHESIS-HIDEMASA-IMASEKI

Jackson, M. A., Gilding, E. K., Shafee, T., Harris, K. S., Kaas, Q., Poon, S., Yap, K., Jia, H., Guarino, R., Chan, L. Y., Durek, T., Anderson, M. A., & Craik, D. J. (2018). Molecular basis for the production of cyclic peptides by plant asparaginyl endopeptidases. *Nature Communications 2018 9:1*, *9*(1), 1–12. https://doi.org/10.1038/s41467-018-04669-9

Jackson, M. B. (1991). Ethylene in Root Growth and Development. In *The Plant Hormone Ethylene*. CRC Press. https://doi.org/10.1201/9781351075763-9

Jokipii-Lukkari, S., Delhomme, N., Schiffthaler, B., Mannapperuma, C., Prestele, J., Nilsson, O., Street, N. R., & Tuominen, H. (2018). Transcriptional roadmap to seasonal variation in wood formation of Norway spruce. *Plant Physiology*, *176*(4), 2851–2870. https://doi.org/10.1104/pp.17.01590

Kampe, A., & Magel, E. (2013). New Insights into Heartwood and Heartwood Formation. *Plant Cell Monographs*, *20*, 71–95. https://doi.org/10.1007/978-3-642-36491-4_3

Kinoshita, T., Nishimura, M., & Hara-Nishimura, I. (1995). Homologues of a vacuolar processing enzyme that are expressed in different organs in Arabidopsis thaliana. *Plant Mol Biol*, *29*(1), 81–89. https://doi.org/10.1007/bf00019120

Kinoshita, T., Yamada, K., Hiraiwa, N., Kondo, M., Nishimura, M., & Hara-Nishimura, I. (1999). Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions. *Plant J*, *19*(1), 43–53. https://doi.org/10.1046/j.1365-313x.1999.00497.x

Klauser, D., Desurmont, G. A., Glauser, G. D. S., Vallat, A., Flury, P., Boller, T., Turlings, T. C. J., & Bartels, S. (2015). The Arabidopsis Pep-PEPR system is induced by herbivore feeding and contributes to JA-mediated plant defence against herbivory. *Journal of Experimental Botany*, *66*(17), 5327–5336. https://doi.org/10.1093/JXB/ERV250

Klemenčič, M., & Funk, C. (2018). Type III metacaspases: calcium-dependent activity proposes new function for the p10 domain. *New Phytologist*. https://doi.org/10.1111/nph.14660

Kornberg, R. D. (2007). The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(32), 12955–12961. https://doi.org/10.1073/PNAS.0704138104/ASSET/573DB0B8-41E5-4EFE-A261-EFFE5771389F/ASSETS/GRAPHIC/ZPQ0260768180018.JPEG

Kroemer, G., Mariño, G., & Levine, B. (2010). Autophagy and the integrated stress response. *Molecular Cell*, *40*(2), 280–293. https://doi.org/10.1016/J.MOLCEL.2010.09.023

Kuroyanagi, M., Nishimura, M., & Hara-Nishimura, I. (2002). Activation of Arabidopsis vacuolar processing enzyme by self-catalytic removal of an auto-inhibitory domain of the C-terminal propeptide. *Plant Cell Physiology*, *43*(2), 143–151. https://doi.org/10.1093/pcp/pcf035

Kuroyanagi, M., Yamada, K., Hatsugai, N., Kondo, M., Nishimura, M., & Hara-Nishimura, I. (2005). Vacuolar processing enzyme is essential for mycotoxin-induced cell death in Arabidopsis thaliana. *Journal of Biological Chemistry*, *280*(38), 32914–32920. https://doi.org/10.1074/jbc.M504476200

Landi, M., Tattini, M., & Gould, K. S. (2015). Multiple functional roles of anthocyanins in plant-environment interactions. *Environmental and Experimental Botany*, *119*, 4–17. https://doi.org/10.1016/J.ENVEXPBOT.2015.05.012

Lange, H. (2009). *Molekulare Grundlagen von Verfärbungsprozessen im Holz der Robinie (Robinia pseudoacacia L.): Genexpressionsstudien an Schlüsselgenen der Flavonoidsynthese*. Universität Hamburg.

Larson, P. R. (1994). *The Vascular Cambium*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-78466-8

Larsson, C., & Møller, I. M. (1990). The Plant Plasma Membrane: Structure, Function and Molecular Biology. In Christer Larsson & Ian M. Møller (Eds.), *The Plant Plasma Membrane*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-74522-5

Li, C. P., Qi, Y. P., Zhang, J., Yang, L. T., Wang, D. H., Ye, X., Lai, N. W., Tan, L. L., Lin, D., & Chen, L. S. (2017). Magnesium-deficiency-induced alterations of gas exchange, major metabolites and key enzymes differ among roots, and lower and upper leaves of Citrus sinensis seedlings. *Tree Physiology*, *37*(11), 1564–1581. https://doi.org/10.1093/TREEPHYS/TPX067

Li, Z., Yue, H., & Xing, D. (2012). MAP Kinase 6-mediated activation of vacuolar processing enzyme modulates heat shock-induced programmed cell death in Arabidopsis. *New Phytologist*. https://doi.org/10.1111/j.1469-8137.2012.04131.x

Lim, K. J., Paasela, T., Harju, A., Venäläinen, M., Paulin, L., Auvinen, P., Kärkkäinen, K., & Teeri, T. H. (2016). Developmental changes in scots pine transcriptome during heartwood formation. *Plant Physiology*, *172*(3), 1403–1417. https://doi.org/10.1104/pp.16.01082

Linnestad, C., Doan, D. N. P., Brown, R. C., Lemmon, B. E., Meyer, D. J., Jung, R., & Olsen, O. A. (1998). Nucellain, a barley homolog of the dicot vacuolar-processing protease, is localized in nucellar cell walls. *Plant Physiology*, *118*(4), 1169–1180. https://doi.org/10.1104/pp.118.4.1169

Liu, J., Osbourn, A., & Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Molecular Plant*, *8*(5), 689–708. https://doi.org/10.1016/j.molp.2015.03.012

Liu, Y., Schiff, M., Czymmek, K., Tallóczy, Z., Levine, B., & Dinesh-Kumar, S. P. (2005). Autophagy regulates programmed cell death during the plant innate immune response. *Cell*, *121*(4), 567–577. https://doi.org/10.1016/j.cell.2005.03.007

Lund, S. T., Stall, R. E., & Klee, H. J. (1998). Ethylene Regulates the Susceptible Response to Pathogen Infection in Tomato. *The Plant Cell*, *10*(3), 371–382. https://doi.org/10.1105/TPC.10.3.371

Magel, E. (2001). Physiology of Cambial Growth, Storage of Reserves and Heartwood Formation. In *Trends in European Forest Tree Physiology Research* (pp. 19–32). https://doi.org/10.1007/978-94-015-9803-3_2

Magel, E., Einig, W., & Hampp, R. (2000). Carbohydrates in trees. *Developments in Crop Science*. https://doi.org/10.1016/S0378-519X(00)80016-1

Magel, E., Hillinger, C., Wagner, T., & Höll, W. (2001). Oxidative pentose phosphate pathway and pyridine nucleotides in relation to heartwood formation in Robinia pseudoacacia L. *Phytochemistry*, *57*(7), 1061–1068. https://doi.org/10.1016/S0031-9422(01)00091-7

Magel, E., Jay-Allemand, C., & Ziegler, H. (1994). Formation of heartwood substances in the stemwood of Robinia pseudoacacia L. II. Distribution of nonstructural carbohydrates and wood extractives across the trunk. *Trees*, *8*(4), 165–171. https://doi.org/10.1007/BF00196843

Makoto, K., Susloparova, E., Tsuyama, I., Shimase, T., Nakaba, S., Takahashi, N., & Yoshida, T. (2021). Influence of soil properties on the heartwood colour of Juglans mandshurica var. sachalinensis in a cool temperate forest. *Journal of Wood Science*, *67*(1), 1–5. https://doi.org/10.1186/s10086-021-01981-9

Marttila, S., Porali, I., Ho, T. H. D., & Mikkonen, A. (1993). Expression of the 30 kD cysteine endoprotease B in germinating barley seeds. *Cell Biology International*, *17*(2), 205–212. https://doi.org/10.1006/CBIR.1993.1056

Matarasso, N., Schuster, S., & Avni, A. (2005). A novel plant cysteine protease has a dual function as a regulator of 1-aminocyclopropane-1-carboxylic acid synthase gene expression. *The Plant Cell*, *17*(4), 1205–1216. https://doi.org/10.1105/tpc.105.030775

Mattoo, A. K., & Handa, A. K. (2008). Higher polyamines restore and enhance metabolic memory in ripening fruit. *Plant Science*, *174*, 386–393. https://doi.org/10.1016/j.plantsci.2008.01.011

Minina, E. A., Filonova, L. H., Fukada, K., Savenkov, E. I., Gogvadze, V., Clapham, D., Sanchez-Vera, V., Suarez, M. F., Zhivotovsky, B., Daniel, G., Smertenko, A., & Bozhkov, P. V. (2013). Autophagy and metacaspase determine the mode of cell death in plants. *Journal of Cell Biology*, *203*(6), 917–927. https://doi.org/10.1083/jcb.201307082

Morris, H. (2016). *The structure and function of ray and axial parenchyma in woody seed plants* (Issue October) [Universität Ulm]. https://doi.org/10.18725/OPARU-4087

Morris, H., Brodersen, C., Schwarze, F. W. M. R., & Jansen, S. (2016). The parenchyma of secondary xylem and its critical role in tree defense against fungal decay in relation to the CODIT model. *Frontiers in Plant Science*, *7*(November 2016), 1–18. https://doi.org/10.3389/fpls.2016.01665

Morris, H., Plavcová, L., Cvecko, P., Fichtler, E., Gillingham, M. A. F., Martínez-Cabrera, H. I., Mcglinn, D. J., Wheeler, E., Zheng, J., Ziemińska, K., & Jansen, S. (2016). A global analysis of parenchyma tissue fractions in secondary xylem of seed plants. *New Phytologist*, *209*(4), 1553–1565. https://doi.org/10.1111/nph.13737

Moya, R., & Marín, J. D. (2011). Grouping of Tectona grandis (L.f.) clones using wood color and stiffness. *New Forests*, *42*(3), 329–345. https://doi.org/10.1007/S11056-011-9255-Y/FIGURES/6

Mylne, J. S., Chan, L. Y., Chanson, A. H., Daly, N. L., Schaefer, H., Bailey, T. L., Nguyencong, P., Cascales, L., & Craik, D. J. (2012). Cyclic Peptides Arising by Evolutionary Parallelism via Asparaginyl-Endopeptidase–Mediated Biosynthesis. *The Plant Cell*, *24*(7), 2765–2778. https://doi.org/10.1105/TPC.112.099085

Mylne, J. S., Colgrave, M. L., Daly, N. L., Chanson, A. H., Elliott, A. G., McCallum, E. J., Jones, A., & Craik, D. J. (2011). Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nature Chemical Biology*, *7*(5), 257–259. https://doi.org/10.1038/NCHEMBIO.542

Nakaba, S., Yamagishi, Y., Sano, Y., & Funada, R. (2012). Temporally and spatially controlled death of parenchyma cells is involved in heartwood formation in pith regions of branches of Robinia pseudoacacia var. inermis. *Journal of Wood Science*, *58*(1), 69–76. https://doi.org/10.1007/s10086-011-1221-y

Nakada, R., & Fukatsu, E. (2012). Seasonal variation of heartwood formation in Larix kaempferi. *Tree Physiology*, *32*(12), 1497–1508. https://doi.org/10.1093/treephys/tps108

Nakaune, S., Yamada, K., Kondo, M., Kato, T., Tabata, S., Nishimura, M., & Hara-Nishimura, I. (2005). A vacuolar processing enzyme, δVPE, is involved in seed coat formation at the early stage of seed development. *Plant Cell*, *17*(3), 876–887. https://doi.org/10.1105/tpc.104.026872

Nardini, A., Lo Gullo, M. A., & Salleo, S. (2011). Refilling embolized xylem conduits: Is it a matter of phloem unloading? *Plant Science*, *180*(4), 604–611. https://doi.org/10.1016/J.PLANTSCI.2010.12.011

Nelson, N. D., Maeglin, R. R., & Wahlgren, H. E. (1969). Relationship of Black Walnut Wood Color to Soil Properties and Site. *Wood and Fiber Science*, 29–37. https://wfs.swst.org/index.php/wfs/article/view/1952

Neuhaus, J. M., & Rogers, J. C. (1998). Integral membrane protein sorting to vacuoles in plant cells: Evidence for two pathways. *Journal of Cell Biology*, *143*(5), 1183–1189. https://doi.org/10.1083/jcb.143.5.1183

Nguyen, G. K. T., Wang, S., Qiu, Y., Hemu, X., Lian, Y., & Tam, J. P. (2014). Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nature Chemical Biology*, *10*(9), 732–738. https://doi.org/10.1038/nchembio.1586

Nobuchi, T., Kuroda, K., Iwata, R., & Harada, H. (1982). Cytological study of the seasonal features of heartwood formation of sugi (Cryptomeria japonica D. Don). *Mokuzai Gakkaishi*, *28*(11), 669–676. https://agris.fao.org/agris-search/search.do?recordID=JP19830920325

Obst, J. R. (1998). Special (Secondary) Metabolites from Wood. *Forest Products Biotechnology*, 161–176.

Ogorodnikov, A., Kargapolova, Y., & Danckwardt, S. (2016). Processing and transcriptome expansion at the mRNA 3′ end in health and disease: finding the right end. *Pflugers Archiv*, *468*(6), 993. https://doi.org/10.1007/S00424-016-1828-3

Olvera-Carrillo, Y., Van Bel, M., Van Hautegem, T., Fendrych, M., Huysmans, M., Simaskova, M., van Durme, M., Buscaill, P., Rivas, S., Coll, N. S., Coppens, F., Maere, S., & Nowack, M. K. (2015). A conserved core of programmed cell death indicator genes discriminates developmentally and environmentally induced programmed cell death in plants. *Plant Physiology*, *169*(4), 2684–2699. https://doi.org/10.1104/pp.15.00769

Panshin, A. J., & Zeeuw, C. de. (1980). *Textbook of wood technology. Vol. 1, Structure, identification, uses, and properties of the commercial woods of the United States*.

Parmentier, Y., Durr, A., Marbach, J., Hirsinger, C., Criqui, M. C., Fleck, J., & Jamet, E. (1995). A novel wound-inducible extensin gene is expressed early in newly isolated protoplasts of Nicotiana sylvestris. *Plant Molecular Biology*, *29*(2), 279–292. https://doi.org/10.1007/BF00043652

Patten, A. M., Vassão, D. G., Wolcott, M. P., Davin, L. B., & Lewis, N. G. (2010). Trees: A remarkable biochemical bounty. *Comprehensive Natural Products II: Chemistry and Biology*, *3*, 1173–1296. https://doi.org/10.1016/b978-008045382-8.00083-6

Pennington, R. T., Lavin, M., & Oliveira-Filho, A. (2009). Woody plant diversity, evolution, and ecology in the tropics: Perspectives from seasonally dry tropical forests. *Annual Review of Ecology, Evolution, and Systematics*, *40*(February), 437–457. https://doi.org/10.1146/annurev.ecolsys.110308.120327

Pesquet, E., Korolev, A. V., Calder, G., & Lloyd, C. W. (2010). The microtubule-associated protein AtMAP70-5 regulates secondary wall patterning in Arabidopsis wood cells. *Current Biology : CB*, *20*(8), 744–749. https://doi.org/10.1016/J.CUB.2010.02.057

Pesquet, E., Korolev, A. V., Calder, G., & Lloyd, C. W. (2011). Mechanisms for shaping, orienting, positioning and patterning plant secondary cell walls. *Plant Signaling & Behavior*, *6*(6), 843. https://doi.org/10.4161/PSB.6.6.15202

Pesquet, E., & Tuominen, H. (2011). *Ethylene stimulates tracheary element differentiation in Zinnia elegans cell cultures*. *190*(1), 138–149. https://doi.org/10.1111/J.1469-8137.2010.03600.X

Poncet, V., Scutt, C., Tournebize, R., Villegente, M., Cueff, G., Rajjou, L., Balliau, T., Zivy, M., Fogliani, B., Job, C., de Kochko, A., Sarramegna-Burtet, V., & Job, D. (2015). The Amborella vacuolar processing enzyme family. *Front Plant Sci*, *6*, 618. https://doi.org/10.3389/fpls.2015.00618

Ponstein, A. S., Bres-Vloemans, S. A., Sela-Buurlage, M. B., Van Den Elzen, P. J. M., Melchers, L. S., & Cornelissen, B. J. C. (1994). A Novel Pathogen- and Wound-Inducible Tobacco (Nicotiana tabacum) Protein with Antifungal Activity. *Plant Physiology*, *104*(1), 109–118. https://doi.org/10.1104/PP.104.1.109

Poth, A. G., Colgrave, M. L., Lyons, R. E., Dalya, N. L., & Craik, D. J. (2011). Discovery of an unusual biosynthetic origin for circular proteins in legumes. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(25), 10127–10132. https://doi.org/10.1073/PNAS.1103660108/SUPPL_FILE/PNAS.1103660108_SI.PDF

Prassinos, C., Ko, J. H., Yang, J., & Han, K. H. (2005). Transcriptome Profiling of Vertical Stem Segments Provides Insights into the Genetic Regulation of Secondary Growth in Hybrid Aspen Trees. *Plant and Cell Physiology*, *46*(8), 1213–1225. https://doi.org/10.1093/PCP/PCI130

Raes, J., Rohde, A., Christensen, J. H., Van De Peer, Y., & Boerjan, W. (2003). Genome-Wide Characterization of the Lignification Toolbox in Arabidopsis. *Plant Physiology*, *133*(3), 1051–1071. https://doi.org/10.1104/pp.103.026484

Rocha, A. Ô. J., Soares, E. L., Costa, J. H., Costa, W. L. G., Soares, A. A., Nogueira, F. C. S., Domont, G. B., & Campos, F. A. P. (2013). Differential expression of cysteine peptidase genes in the inner integument and endosperm of developing seeds of Jatropha curcas L. (Euphorbiaceae). *Plant Science*, *213*, 30–37. https://doi.org/10.1016/j.plantsci.2013.08.009

Salleo, S., Lo Gullo, M. A., Trifilò, P., & Nardini, A. (2004). New evidence for a role of vessel-associated cells and phloem in the rapid xylem refilling of cavitated stems of Laurus nobilis L. *Plant, Cell & Environment*, *27*(8), 1065–1076. https://doi.org/10.1111/J.1365-3040.2004.01211.X

Salleo, Sebastiano, LastNameLastNameTrifilò, P., Esposito, S., Nardini, A., & Gullo, M. A. Lo. (2009). Starch-to-sugar conversion in wood parenchyma of field-growing Laurus nobilis plants: a component of the signal pathway for embolism repair? *Functional Plant Biology*, *36*(9), 815–825. https://doi.org/10.1071/FP09103

Saranpää, P., & Nyberg, H. (1987). Lipids and sterols of Pinus sylvestris L. sapwood and heartwood. *Trees*, *1*(2), 82–87. https://doi.org/10.1007/BF00203575

Sauter, J. J. (1988). Seasonal changes in the efflux of sugars from parenchyma cells into the apoplast in poplar stems (Populus × canademis “robusta”). *Trees*, *2*(4), 242–249. https://doi.org/10.1007/BF00202379

Schmitt, U., & Liese, W. (1993). Response of xylem parenchyma by suberization in some hardwoods after mechanical injury. *Trees 1993 8:1*, *8*(1), 23–30. https://doi.org/10.1007/BF00240978

Schweichel, J. -U, & Merker, H. -J. (1973). The morphology of various types of cell death in prenatal tissues. *Teratology*, *7*(3), 253–266. https://doi.org/10.1002/TERA.1420070306

Secchi, F., Pagliarani, C., & Zwieniecki, M. A. (2017). The functional role of xylem parenchyma cells and aquaporins during recovery from severe water stress. *Plant Cell and Environment*, *40*(6), 858–871. https://doi.org/10.1111/pce.12831

Secchi, F., & Zwieniecki, M. A. (2011). Sensing embolism in xylem vessels: the role of sucrose as a trigger for refilling. *Plant, Cell & Environment*, *34*(3), 514–524. https://doi.org/10.1111/J.1365-3040.2010.02259.X

Secondary Growth: The Vascular Cambium. (n.d.). Retrieved December 13, 2022, from https://www.whfreeman.com/BrainHoney/Resource/6716/SitebuilderUploads/Hillis2e/Student Resources/Animated Tutorials/pol2e_at_2401_Secondary_Growth_The_Vascular_Cambium/pol2e_at_2401_Secondary_Growth_The_Vascular_Cambium.html

Shafee, T., Harris, K., & Anderson, M. (2015). Biosynthesis of Cyclotides. In *Advances in Botanical Research* (Vol. 76). Elsevier Ltd. https://doi.org/10.1016/bs.abr.2015.08.005

Shain, L. (1995). Stem defense against pathogens. In B. L. Gartner (Ed.), *Plant stems : physiology and functional morphology* (pp. 383–406). Academic Press.

Sharples, A. ., & Gunnery, H. . (1933). Callus Formation in Hibiscus Rosa-sinensis L. and Hevea brasiliensis Müll. Arg on JSTOR. *Annals of Botany* , *47*(188), 827–839.

Shi, Y. (2004). Caspase activation: Revisiting the induced proximity model. *Cell*, *117*(7), 855–858. https://doi.org/10.1016/j.cell.2004.06.007

Shigo, A. L. (1984). Compartmentalization: A Conceptual Framework for Understanding How Trees Grow and Defend Themselves. *Annual Review of Phytopathology*, *22*(1), 189–214. https://doi.org/10.1146/annurev.py.22.090184.001201

Shigo, A. L., & Tippett, J. T. (1981). Compartmentalization of American Elm Tissues Infected byCeratocystis ulmi. *Plant Disease*, *65*(9), 715. https://doi.org/10.1094/PD-65-715

Shimada, T., Takagi, J., Ichino, T., Shirakawa, M., & Hara-Nishimura, I. (2018). Plant Vacuoles. *Annual Review of Plant Biology*, *69*(April), 123–145. https://doi.org/10.1146/annurev-arplant-042817-040508

Shimada, T., Yamada, K., Kataoka, M., Nakaune, S., Koumoto, Y., Kuroyanagi, M., Tabata, S., Kato, T., Shinozaki, K., Seki, M., Kobayashi, M., Kondo, M., Nishimura, M., & Hara-Nishimura, I. (2003). Vacuolar processing enzymes are essential for proper processing of seed storage proteins in Arabidopsis thaliana. *Journal of Biological Chemistry*, *278*(34), 32292–32299. https://doi.org/10.1074/jbc.M305740200

Shirahama-Noda, K., Yamamoto, A., Sugihara, K., Hashimoto, N., Asano, M., Nishimura, M., & Hara-Nishimura, I. (2003). Biosynthetic processing of cathepsins and lysosomal degradation are abolished in asparaginyl endopeptidase-deficient mice. *Journal of Biological Chemistry*, *278*(35). https://doi.org/10.1074/jbc.M302742200

Simpson, M. G. (2019). Plant Systematics, Third Edition. In *Plant Systematics, Third Edition* (3rd ed.). Elsevier. https://doi.org/10.1016/C2015-0-04664-0

Singh, T., & Singh, A. P. (2012). A review on natural products as wood protectant. *Wood Science and Technology*, *46*(5), 851–870. https://doi.org/10.1007/S00226-011-0448-5/TABLES/1

Sokolowska, K., & Zagórska-Marek, B. (2012). Symplasmic, long-distance transport in xylem and cambial regions in branches of Acer pseudoplatanus (Aceraceae) and Populus tremula x P. tremuloides (Salicaceae). *American Journal of Botany*, *99*(11), 1745–1755. https://doi.org/10.3732/AJB.1200349

Song, K., Yin, Y., Salmén, L., Xiao, F., & Jiang, X. (2014). Changes in the properties of wood cell walls during the transformation from sapwood to heartwood. *Journal of Materials Science*, *49*(4), 1734–1742. https://doi.org/10.1007/s10853-013-7860-1

Spicer, R. (2005). Senescence in Secondary Xylem: Heartwood Formation as an Active Developmental Program. In *Vascular Transport in Plants*. https://doi.org/10.1016/B978-012088457-5/50024-1

Spicer, R. (2014). Symplasmic networks in secondary vascular tissues: Parenchyma distribution and activity supporting long-distance transport. *Journal of Experimental Botany*, *65*(7), 1829–1848. https://doi.org/10.1093/jxb/ert459

Spicer, R. (2016). *Variation in Angiosperm Wood Structure and Its Physiological and Evolutionary Significance*. *June*, 19–60. https://doi.org/10.1007/7397_2016_28

Spicer, R., & Groover, A. (2010). Evolution of development of vascular cambia and secondary growth. *New Phytologist*, *186*(3), 577–592. https://doi.org/10.1111/j.1469-8137.2010.03236.x

Stewart, C. M. (1966). Excretion and heartwood formation in living trees. *Science*. https://doi.org/10.1126/science.153.3740.1068

Stewart, W. N., & Rothwell, G. W. (2009). *Paleobotany and the evolution of plants* (2nd ed.). Cambridge University Press.

Suarez, M. F., Filonova, L. H., Smertenko, A., Savenkov, E. I., Clapham, D. H., Von Arnold, S., Zhivotovsky, B., & Bozhkov, P. V. (2004). Metacaspase-dependent programmed cell death is essential for plant embryogenesis. *Current Biology*, *14*(9), R339–R340. https://doi.org/10.1016/j.cub.2004.04.019

Sundström, J. F., Vaculova, A., Smertenko, A. P., Savenkov, E. I., Golovko, A., Minina, E., Tiwari, B. S., Rodriguez-Nieto, S., Zamyatnin, A. A., Välineva, T., Saarikettu, J., Frilander, M. J., Suarez, M. F., Zavialov, A., Ståhl, U., Hussey, P. J., Silvennoinen, O., Sundberg, E., Zhivotovsky, B., & Bozhkov, P. V. (2009). Tudor staphylococcal nuclease is an evolutionarily conserved component of the programmed cell death degradome. *Nature Cell Biology*, *11*(11), 1347–1354. https://doi.org/10.1038/ncb1979

Szallies, A., Kubata, B. K., & Duszenko, M. (2002). A metacaspase of Trypanosoma brucei causes loss of respiration competence and clonal death in the yeast Saccharomyces cerevisiae. *FEBS Letters*. https://doi.org/10.1016/S0014-5793(02)02608-X

Takabe, K., & Kim, J. S. (2016). Rapid Freezing and Immunocytochemistry Provide New Information on Cell Wall Formation in Woody Plants. In *Secondary Xylem Biology: Origins, Functions, and Applications*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-802185-9.00016-4

Tanaka, T., Inazawa, J., & Nakamura, Y. (1996). Molecular cloning of a human cDNA encoding putative cysteine protease (PRSC1) and its chromosome assignment to 14q32.1. *Cytogenetic and Genome Research*, *74*(1–2), 120–123. https://doi.org/10.1159/000134397

Tang, D., Kang, R., Berghe, T. Vanden, Vandenabeele, P., & Kroemer, G. (2019). The molecular machinery of regulated cell death. *Cell Research*, *29*(5), 347–364. https://doi.org/10.1038/s41422-019-0164-5

Taylor, A. M., Gartner, B. L., & Morrell, J. J. (2002). Heartwood formation and natural durability - A review. *Wood and Fiber Science*, *34*(4), 587–611.

Thomas, H. (2013). Senescence, ageing and death of the whole plant. *New Phytologist*, *197*(3), 696–711. https://doi.org/10.1111/NPH.12047

Tohge, T., Ramos, M. S., Nunes-Nesi, A., Mutwil, M., Giavalisco, P., Steinhauser, D., Schellenberg, M., Willmitzer, L., Persson, S., Martinoia, E., & Fernie, A. R. (2011). Toward the Storage Metabolome: Profiling the Barley Vacuole. *Plant Physiology*, *157*(3), 1469–1482. https://doi.org/10.1104/PP.111.185710

Tsiatsiani, L., Timmerman, E., De Bock, P. J., Vercammen, D., Stael, S., van de Cotte, B., Staes, A., Goethals, M., Beunens, T., Van Damme, P., Gevaert, K., & Van Breusegem, F. (2013). The arabidopsis METACASPASE9 degradome. *Plant Cell*. https://doi.org/10.1105/tpc.113.115287

Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E., & Bozhkov, P. V. (2011). Metacaspases. *Cell Death and Differentiation*, *18*(8), 1279–1288. https://doi.org/10.1038/cdd.2011.66

Turner, S., Gallois, P., & Brown, D. (2007). Tracheary element differentiation. *Annual Review of Plant Biology*, *58*, 407–433. https://doi.org/10.1146/ANNUREV.ARPLANT.57.032905.105236

Tuteja, N., & Mahajan, S. (2007). Calcium Signaling Network in Plants. *Plant Signaling & Behavior*, *2*(2), 79–85. https://doi.org/10.4161/psb.2.2.4176

Uren, A. G., O'Rourke, K., Aravind, L., Pisabarro, M. T., Seshagiri, S., Koonin, E. V., & Dixit, V. M. (2000). Identification of paracaspases and metacaspases: Two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. *Molecular Cell*. https://doi.org/10.1016/s1097-2765(05)00086-9

van der Schoot, C., & van Bel, A. J. E. (1990). Mapping membrane potential differences and dye-coupling in internodal tissues of tomato (Solanum lycopersicum L.). *Planta*, *182*(1), 9–21. https://doi.org/10.1007/BF00239977

Van Doorn, W. G., Beers, E. P., Dangl, J. L., Franklin-Tong, V. E., Gallois, P., Hara-Nishimura, I., Jones, A. M., Kawai-Yamada, M., Lam, E., Mundy, J., Mur, L. A. J., Petersen, M., Smertenko, A., Taliansky, M., Van Breusegem, F., Wolpert, T., Woltering, E., Zhivotovsky, B., & Bozhkov, P. (2011). Morphological classification of plant cell deaths. *Cell Death and Differentiation*, *18*(8), 1241–1246. https://doi.org/10.1038/cdd.2011.36

Van Doorn, W. G., & Papini, A. (2013). Ultrastructure of autophagy in plant cells: A review. *Autophagy*, *9*(12), 1922–1936. https://doi.org/10.4161/auto.26275

Vercammen, D., Belenghi, B., Romero-Puertas, M. C., Brackenier, A., Inzé, D., Delledonne, M., & Van Breusegem, F. (2007). Metacaspase activity of Arabidopsis thaliana is regulated by S-nitrosylation of a critical cysteine residue. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M608931200

Vercammen, D., De Cotte, B. Van, De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inzé, D., & Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana cleave substrates after arginine and lysine. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M406329200

von Arx, G., Arzac, A., Olano, J. M., & Fonti, P. (2015). Assessing conifer ray parenchyma for ecological studies: Pitfalls and guidelines. *Frontiers in Plant Science*, *6*(NOVEMBER), 1–10. https://doi.org/10.3389/fpls.2015.01016

Wang, M., Qi, X., Zhao, S., Zhang, S., & Lu, M. Z. (2009). Dynamic changes in transcripts during regeneration of the secondary vascular system in Populus tomentosa Carr. revealed by cDNA microarrays. *BMC Genomics*, *10*(1), 1–8. https://doi.org/10.1186/1471-2164-10-215/FIGURES/3

Watanabe, N., & Lam, E. (2011). Calcium-dependent activation and autolysis of arabidopsis metacaspase 2d. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M110.194340

Weston, L. A., & Mathesius, U. (2013). Flavonoids: Their Structure, Biosynthesis and Role in the Rhizosphere, Including Allelopathy. *Journal of Chemical Ecology*, *39*(2), 283–297. https://doi.org/10.1007/s10886-013-0248-5

Wilkins, O., Nahal, H., Foong, J., Provart, N. J., & Campbell, M. M. (2009). Expansion and Diversification of the Populus R2R3-MYB Family of Transcription Factors. *Plant Physiology*, *149*(2), 981–993. https://doi.org/10.1104/PP.108.132795

Wilkins, T. A., & Raikhel, N. V. (1989). Expression of rice lectin is governed by two temporally and spatially regulated mRNAs in developing embryos. *The Plant Cell*, *1*(5), 541–549. https://doi.org/10.1105/TPC.1.5.541

Wycoff, K. L., Powell, P. A., Gonzales, R. A., Corbin, D. R., Lamb, C., & Dixon, R. A. (1995). Stress Activation of a Bean Hydroxyproline-Rich Glycoprotein Promoter Is Superimposed on a Pattern of Tissue-Specific Developmental Expression. *Plant Physiology*, *109*(1), 41–52. https://doi.org/10.1104/PP.109.1.41

Yamada, K., Basak, A. K., Goto-Yamada, S., Tarnawska-Glatt, K., & Hara-Nishimura, I. (2019). Vacuolar processing enzymes in the plant life cycle. *New Phytologist*, *1593*(October 2019), 21–31. https://doi.org/10.1111/nph.16306

Yamada, K., Shimada, T., Nishimura, M., & Hara-Nishimura, I. (2005). A VPE family supporting various vacuolar functions in plants. *Physiologia Plantarum*, *123*(4), 369–375. https://doi.org/10.1111/j.1399-3054.2005.00464.x

Yamaguchi, M., Kubo, M., Fukuda, H., & Demura, T. (2008). VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. *The Plant Journal*, *55*(4), 652–664. https://doi.org/10.1111/J.1365-313X.2008.03533.X

Yang, J., Kamdem, D. P., Keathley, D. E., & Han, K. H. (2004). Seasonal changes in gene expression at the sapwood-heartwood transition zone of black locust (Robinia pseudoacacia) revealed by cDNA microarray analysis. *Tree Physiology*, *24*(4), 461–474. https://doi.org/10.1093/treephys/24.4.461

Yang, J., Park, S., Kamdem, D. P., Keathley, D. E., Retzel, E., Paule, C., Kapur, V., & Han, K. H. (2003). Novel gene expression profiles define the metabolic and physiological processes characteristic of wood and its extractive formation in a hardwood tree species, Robinia pseudoacacia. *Plant Molecular Biology*. https://doi.org/10.1023/A:1025445427284

Yeh, T. F., Chu, J. H., Liu, L. Y., & Chen, S. Y. (2020). Differential gene profiling of the heartwood formation process in Taiwania cryptomerioides hayata xylem tissues. *International Journal of Molecular Sciences*, *21*(3). https://doi.org/10.3390/ijms21030960

Zauner, F. B., Dall, E., Regl, C., Grassi, L., Huber, C. G., Cabrele, C., & Brandstetter, H. (2018). Crystal structure of plant legumain reveals a unique two-chain state with pH-dependent activity regulation. *Plant Cell*, *30*(3), 686–699. https://doi.org/10.1105/tpc.17.00963

Zeng, S., Shen, W. H., & Liu, L. (2018). Senescence and Cancer. *Cancer Translational Medicine*, *4*(3), 70. https://doi.org/10.4103/CTM.CTM_22_18

Zhong, R., Lee, C., Haghighat, M., & Ye, Z. H. (2021). Xylem vessel-specific SND5 and its homologs regulate secondary wall biosynthesis through activating secondary wall NAC binding elements. *New Phytologist*, *231*(4), 1496–1509. https://doi.org/10.1111/nph.17425

Zhong, R., Lee, C., & Ye, Z. H. (2010). Global Analysis of Direct Targets of Secondary Wall NAC Master Switches in Arabidopsis. *Molecular Plant*, *3*(6), 1087–1103. https://doi.org/10.1093/MP/SSQ062

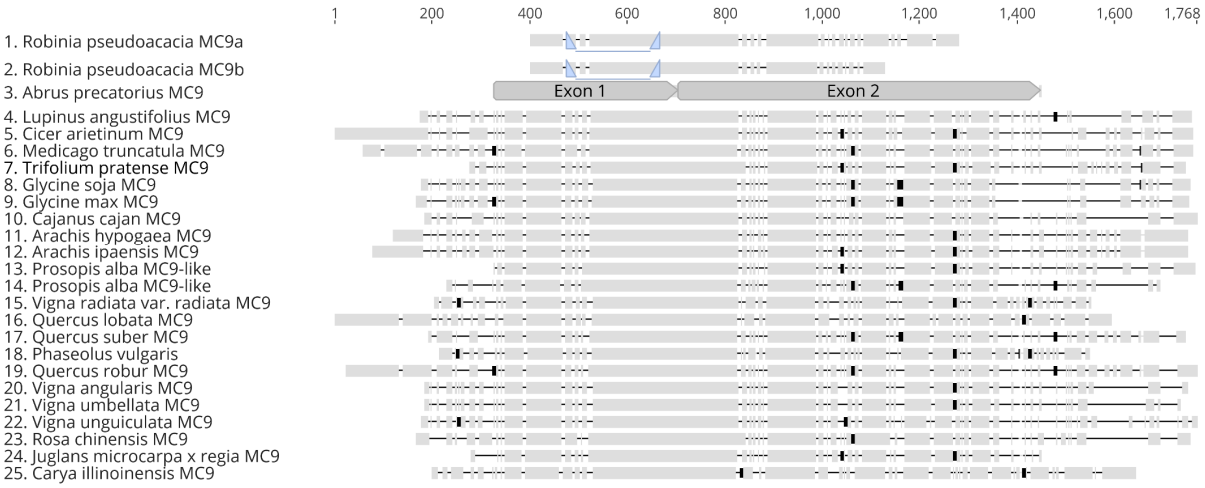
Zhong, R., & Ye, Z. H. (2007). Regulation of cell wall biosynthesis. *Current Opinion in Plant Biology*, *10*(6), 564–572. https://doi.org/10.1016/J.PBI.2007.09.001

Zhong, R., & Ye, Z. H. (2013). Transcriptional Regulation of Wood Formation in Tree Species. *Plant Cell Monographs*, *20*, 141–158. https://doi.org/10.1007/978-3-642-36491-4_5

Zhong, R., & Ye, Z. H. (2015). Secondary cell walls: Biosynthesis, patterned deposition and transcriptional regulation. *Plant and Cell Physiology*, *56*(2), 195–214. https://doi.org/10.1093/pcp/pcu140

Zhu, P., Yu, X., Wang, C., Zhang, Q., Liu, W., McSweeney, S., Shanklin, J., Lam, E., & Liu, Q. (2020). Structural basis for Ca2+-dependent activation of a plant metacaspase. *Submitted*.

8. Appendix



Lane	Accession	Description	Query coverage	% Pairwise Identity	Bit-Score
1	--	Robinia pseudoacacia MC9a	--	--	--
2	--	Robinia pseudoacacia MC9b	--	--	--
3	XM_027495318.1	PREDICTED: Abrus precatorius metacaspase-9 (LOC113862209), mRNA	99.00%	86.0%	924.611
4	XM_019601616.1	PREDICTED: Lupinus angustifolius metacaspase-9 (LOC109357640), mRNA	99.00%	84.1%	856.083
5	XM_004493225.3	PREDICTED: Cicer arietinum metacaspase-9 (LOC101488845), mRNA	98.75%	82.5%	806.49
6	XM_013594075.3	PREDICTED: Medicago truncatula metacaspase-9 (LOC25498997), mRNA	98.87%	81.4%	762.308
7	XM_045940070.1	PREDICTED: Trifolium pratense metacaspase-9 (LOC123890464), mRNA	98.37%	81.4%	757.8
8	XM_028361387.1	PREDICTED: Glycine soja metacaspase-9 (LOC114399243), mRNA	99.00%	81.0%	738.864
9	XM_003554030.5	PREDICTED: Glycine max metacaspase-9 (LOC100781510), mRNA	99.00%	81.0%	738.864
10	XM_020364252.2	PREDICTED: Cajanus cajan metacaspase-9 (LOC109802849), mRNA	99.00%	79.7%	691.075
11	XM_025749698.2	PREDICTED: Arachis hypogaea metacaspase-9 (LOC112696814), mRNA	97.62%	75.9%	554.019
12	XM_016306384.2	PREDICTED: Arachis ipaensis metacaspase-9 (LOC107604728), mRNA	97.62%	75.7%	549.51
13	XM_028916909.1	PREDICTED: Prosopis alba metacaspase-9-like (LOC114729879), mRNA	99.00%	75.1%	526.968
14	XM_028910894.1	PREDICTED: Prosopis alba metacaspase-9-like (LOC114724535), mRNA	99.00%	75.1%	526.968
15	XM_014635660.2	PREDICTED: Vigna radiata var. radiata metacaspase-9 (LOC106753803), mRNA	99.00%	74.9%	526.067
16	XM_031068375.1	PREDICTED: Quercus lobata metacaspase-9 (LOC115951191), mRNA	99.25%	74.6%	513.443
17	XM_024073957.1	PREDICTED: Quercus suber metacaspase-9-like (LOC112041046), mRNA	99.25%	74.4%	506.23
18	XM_007161767.1	Phaseolus vulgaris hypothetical protein (PHAVU_001G101400g) mRNA, complete cds	99.00%	74.2%	497.213
19	XM_050389100.1	PREDICTED: Quercus robur metacaspase-9 (LOC126693194), mRNA	99.25%	74.2%	499.918
20	XM_017562410.1	PREDICTED: Vigna angularis metacaspase-9 (LOC108328531), mRNA	99.00%	74.0%	492.704
21	XM_047291339.1	PREDICTED: Vigna umbellata metacaspase-9 (LOC124819764), mRNA	99.00%	73.8%	488.196
22	XM_028061471.1	PREDICTED: Vigna unguiculata metacaspase-9 (LOC114176425), mRNA	99.00%	73.7%	477.376
23	XM_024303475.2	PREDICTED: Rosa chinensis metacaspase-9 (LOC112166596), mRNA	96.99%	72.7%	441.308
24	XM_041169463.1	PREDICTED: Juglans microcarpa x Juglans regia metacaspase-9 (LOC121265797), mRNA	98.49%	72.0%	410.651
25	XM_043087374.1	PREDICTED: Carya illinoinensis metacaspase-9 (LOC122277404), mRNA	98.49%	71.5%	397.126

-- Indicates not applicable

Figure 1: *RpMC9* and homologs DNA sequence alignment overview and sequence information table on BLASTn result with *RpMC9* as query.

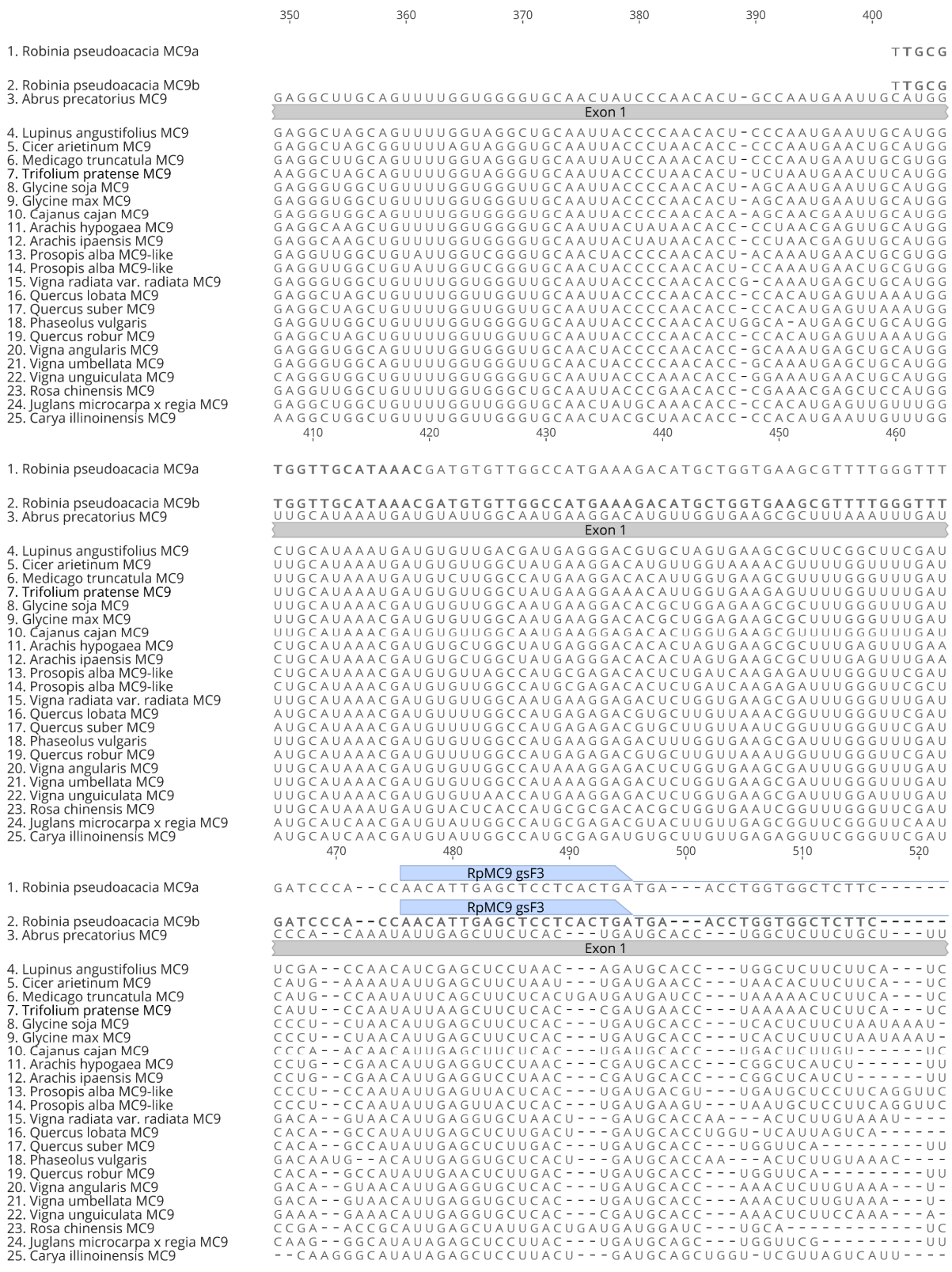


Figure 2: DNA sequence alignment with RpMC9 and homologs (page 1).

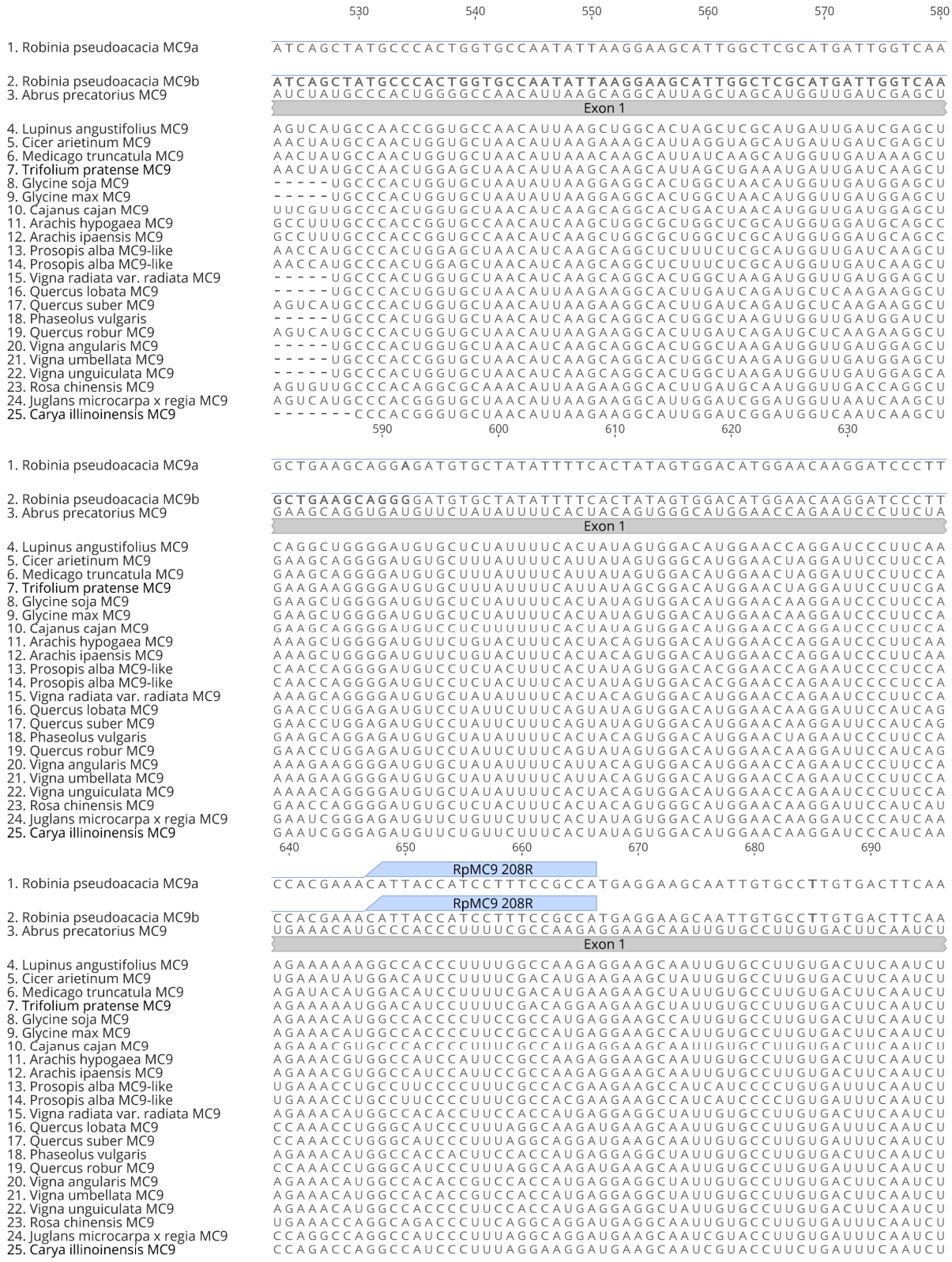


Figure 2: DNA sequence alignment with RpMC9 and homologs (page 2).

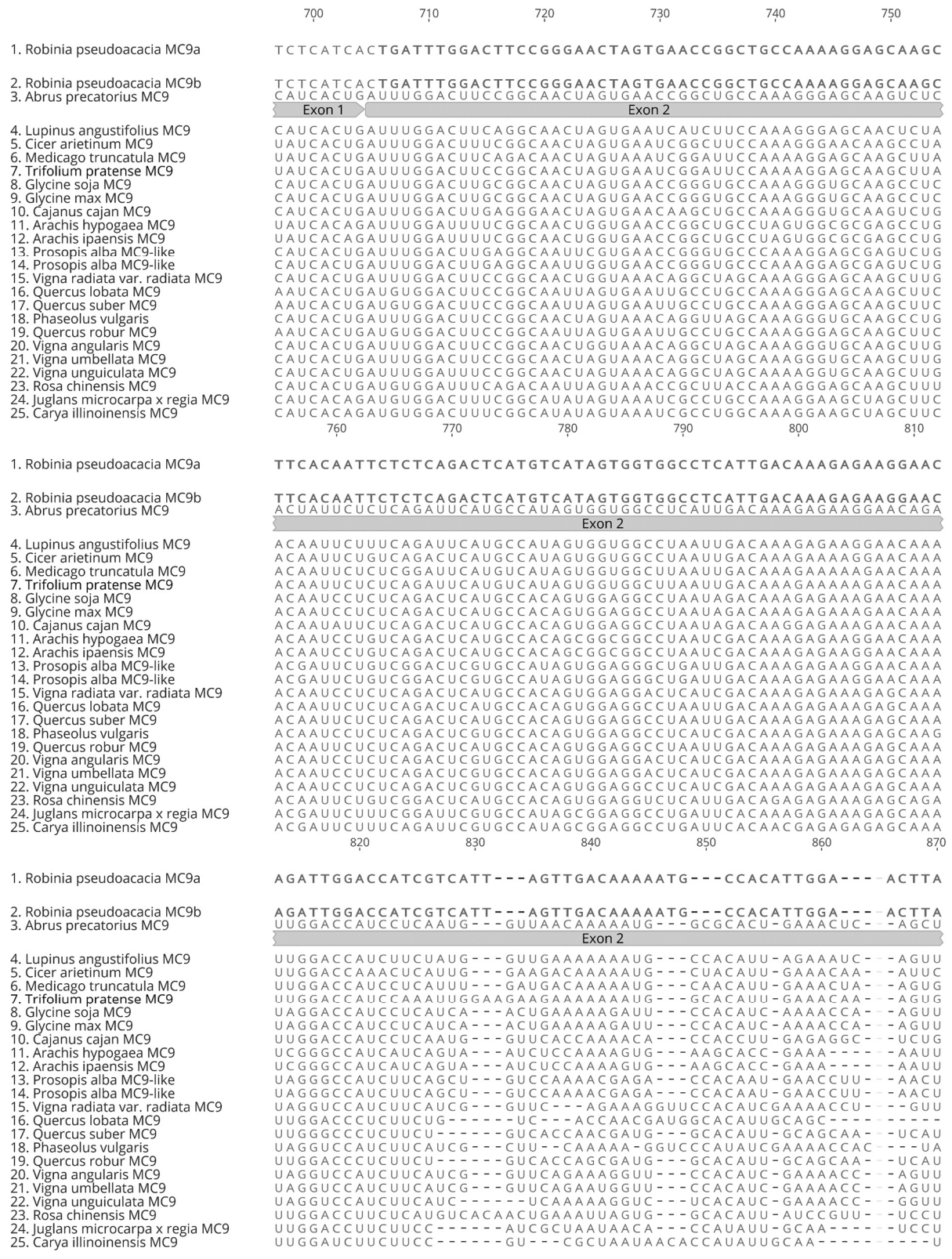


Figure 2: DNA sequence alignment with RpMC9 and homologs (page 3).

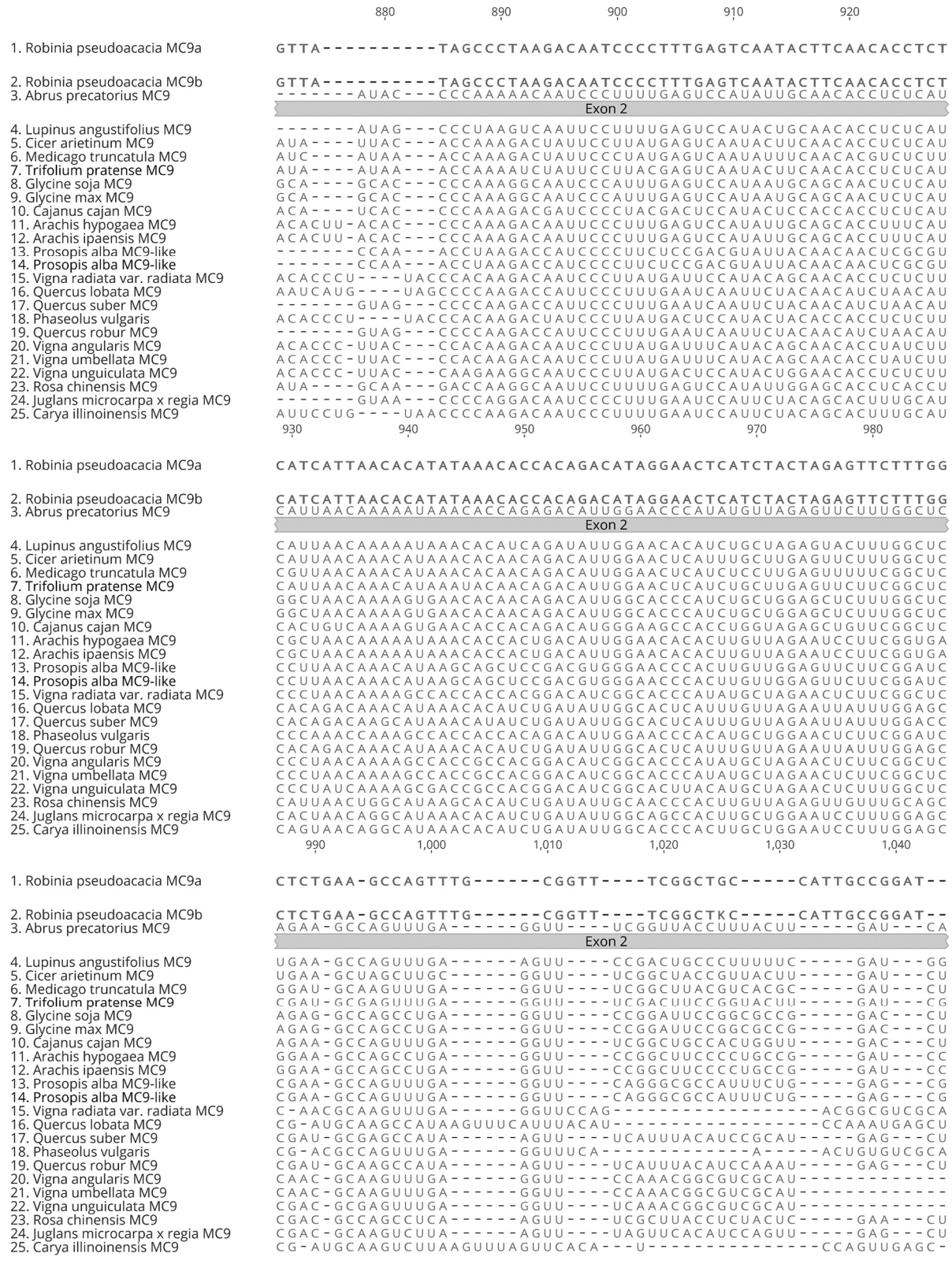


Figure 2: DNA sequence alignment with RpMC9 and homologs (page 4).

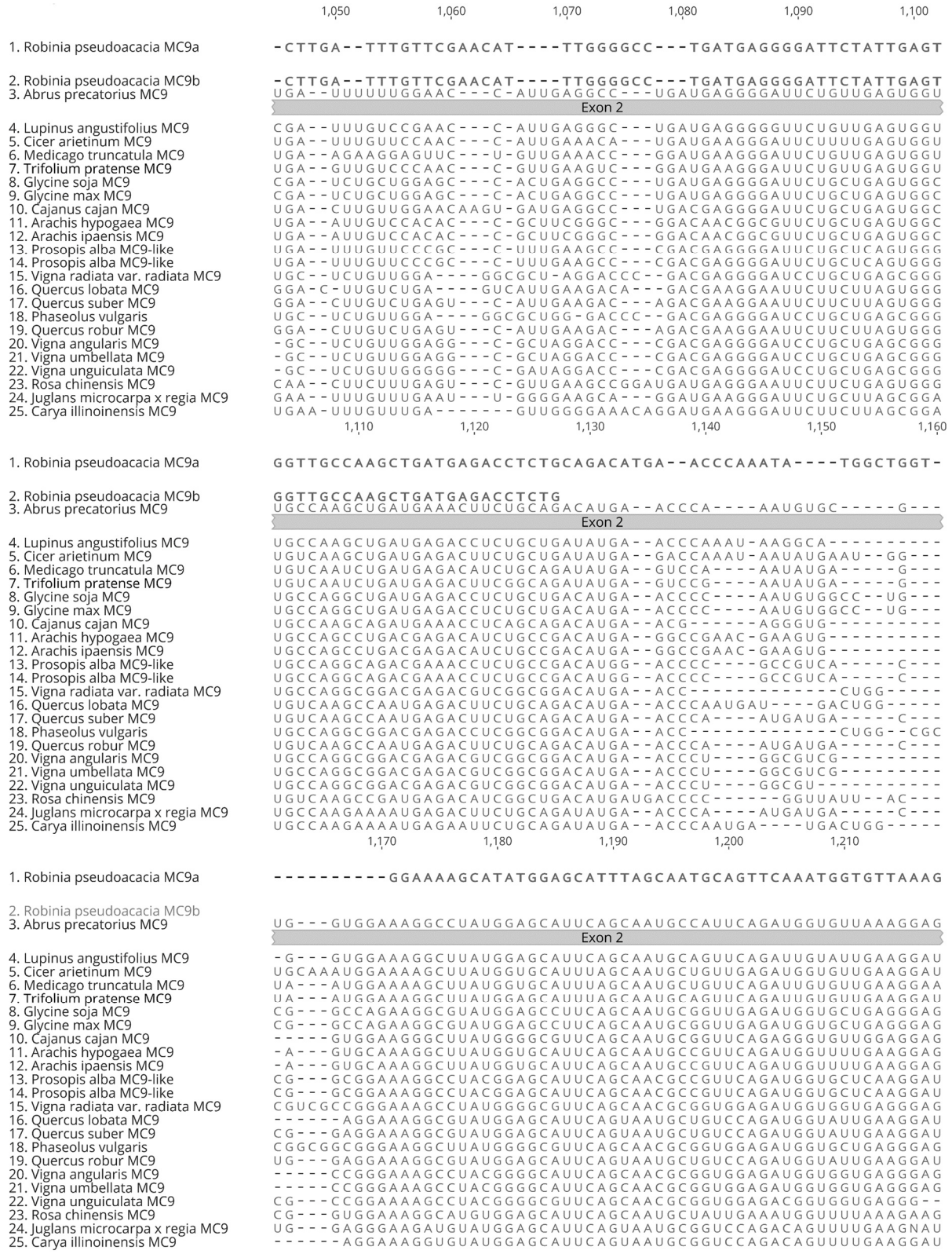


Figure 2: DNA sequence alignment with RpMC9 and homologs (page 5).

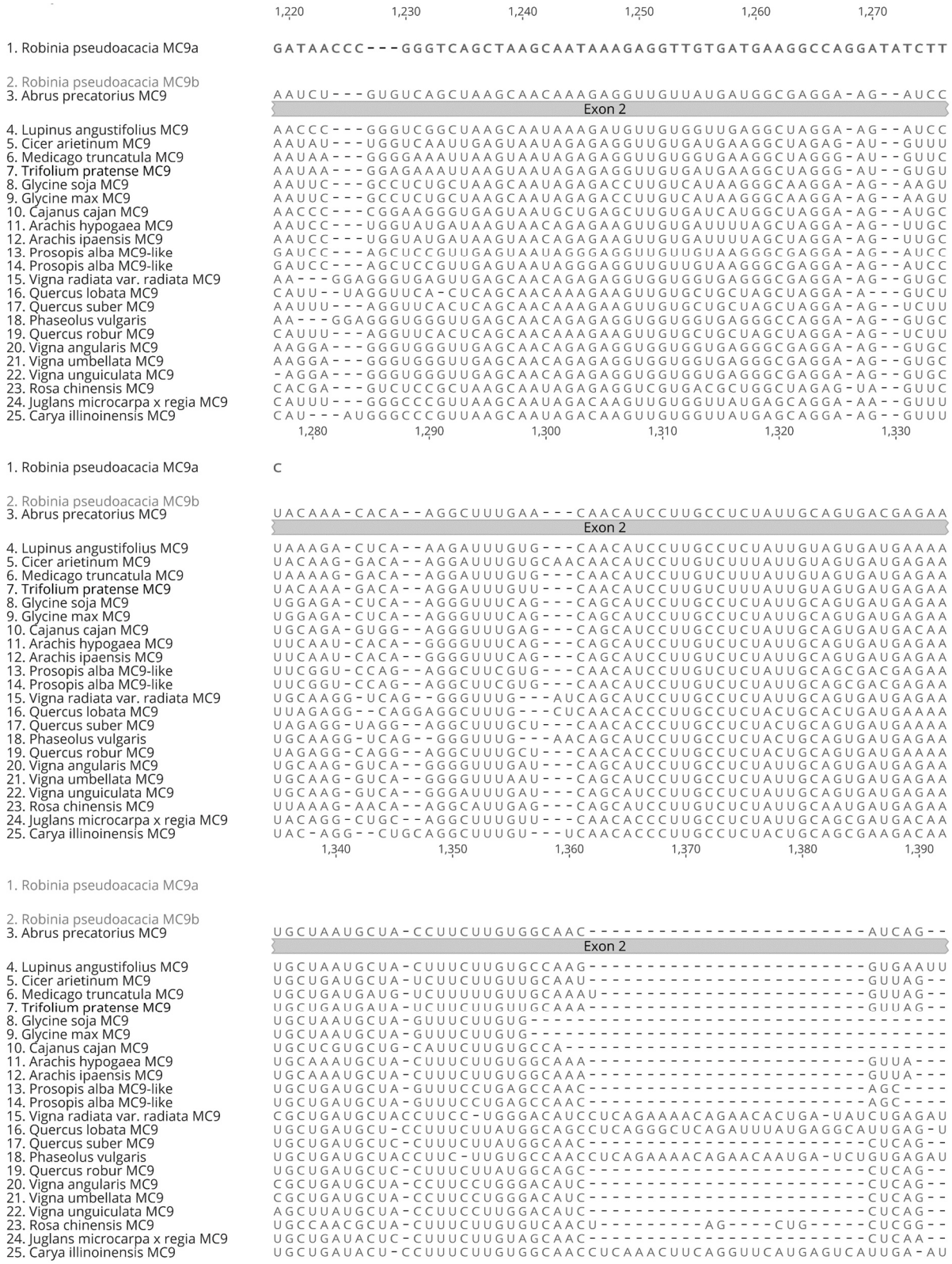
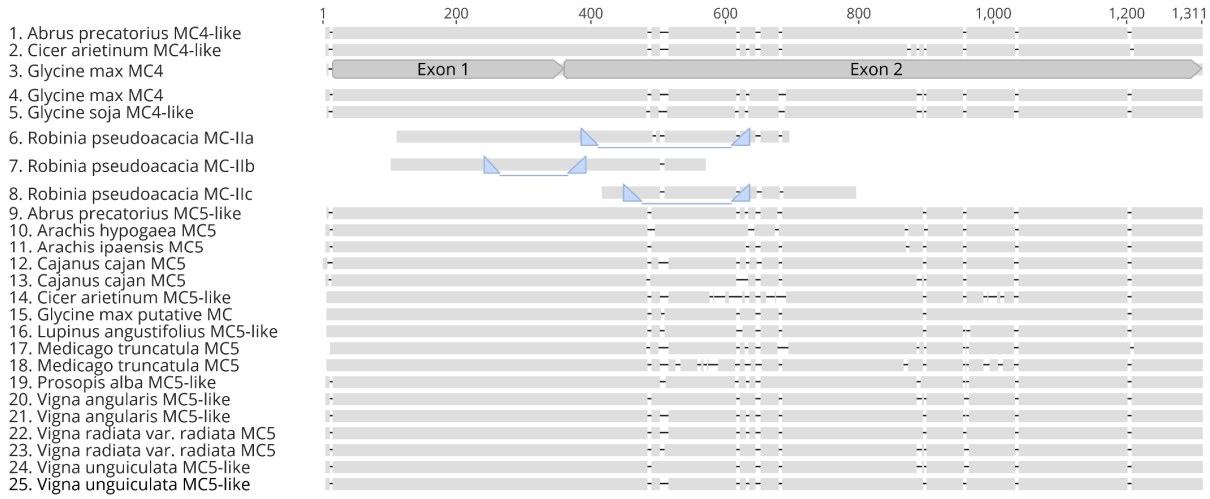


Figure 2: DNA sequence alignment with RpMC9 and homologs (page 6).



-- Indicates not applicable

Figure 3: *RpMC-II* and homologs DNA sequence alignment overview and sequence information table.

Lane	Sequence label	Description	Accession
1	Abrus precatorius MC4-like	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X1, mRNA	XM_027499192.1
2	Cicer arietinum MC4-like	PREDICTED: Cicer arietinum metacaspase-4-like (LOC101509880), mRNA	XM_004509944.3
3	Glycine max MC4	PREDICTED: Glycine max metacaspase-4 (LOC100795043), mRNA	XM_003531762.5
4	Glycine max MC4	PREDICTED: Glycine max metacaspase-4 (LOC100802658), mRNA	XM_003546612.5
5	Glycine soja MC4-like	PREDICTED: Glycine soja metacaspase-4-like (LOC114422611), mRNA	XM_028389058.1
6	Robinia pseudoacacia MC-IIa	Robinia pseudoacacia Type-II matecaspase RpMC-IIa	
7	Robinia pseudoacacia MC-IIb	Robinia pseudoacacia Type-II matecaspase RpMC-IIb	
8	Robinia pseudoacacia MC-IIc	Robinia pseudoacacia Type-II matecaspaseRpMC-IIc	
9	Arachis hypogaea MC5	PREDICTED: Arachis hypogaea metacaspase-5 (LOC112707436), mRNA	XM_025759180.2
10	Arachis ipaensis MC5	PREDICTED: Arachis ipaensis metacaspase-5 (LOC107635285), mRNA	XM_016338719.2
11	Abrus precatorius MC5-like	PREDICTED: Abrus precatorius metacaspase-5-like (LOC113865563), transcript variant X1, mRNA	XM_027500219.1
12	Cajanus cajan MC5	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789350), mRNA	XM_020348274.2
13	Cajanus cajan MC5	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789349), transcript variant X1, mRNA	XM_020348272.2
14	Cicer arietinum MC5-like	PREDICTED: Cicer arietinum metacaspase-5-like (LOC101509553), mRNA	XM_004488501.3
15	Glycine max putative MC	Glycine max putative metacaspase (LOC100527809), mRNA	NM_001250818.2
16	Lupinus angustifolius MC5-like	PREDICTED: Lupinus angustifolius metacaspase-5-like (LOC109352001), mRNA	XM_019593760.1
17	Medicago truncatula MC5	PREDICTED: Medicago truncatula metacaspase-5 (LOC25502442), mRNA	XM_013589850.3
18	Medicago truncatula MC5	PREDICTED: Medicago truncatula metacaspase-5 (LOC11419585), mRNA	XM_003595881.4
19	Prosopis alba MC5-like	PREDICTED: Prosopis alba metacaspase-5-like (LOC114754970), mRNA	XM_028943832.1
20	Vigna angularis MC5-like	PREDICTED: Vigna angularis metacaspase-5-like (LOC108338807), mRNA	XM_017575836.1
21	Vigna angularis MC5-like	PREDICTED: Vigna angularis metacaspase-5-like (LOC108340250), mRNA	XM_017577495.1
22	Vigna radiata var. radiata MC5	PREDICTED: Vigna radiata var. radiata metacaspase-5 (LOC106752949), mRNA	XM_014634726.2
23	Vigna radiata var. radiata MC5	PREDICTED: Vigna radiata var. radiata metacaspase-5 (LOC106752959), mRNA	XM_014634744.2
24	Vigna unguiculata MC5-like	PREDICTED: Vigna unguiculata metacaspase-5-like (LOC114169553), mRNA	XM_028054754.1
25	Vigna unguiculata MC5-like	PREDICTED: Vigna unguiculata metacaspase-5-like (LOC114170591), mRNA	XM_028056101.1

Table 1: BLASTn result with *RpMC-IIa* as query.

Accession	Description	Query coverage	% Pairwise Identity	Bit-Score
NM_001250818	Glycine max putative metacaspase (LOC100527809), mRNA	99.46%	80.0%	409.23
XM_028389055	PREDICTED: Glycine soja metacaspase-5-like (LOC114422609), mRNA	99.46%	80.0%	409.23
XM_020348272	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789349), transcript variant X1, mRNA	100.00%	80.3%	409.23
XM_020348273	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789349), transcript variant X2, mRNA	100.00%	80.3%	409.23
XM_014634726	PREDICTED: Vigna radiata var. radiata metacaspase-5 (LOC106752949), mRNA	99.11%	80.3%	411.07
XM_019593760	PREDICTED: Lupinus angustifolius metacaspase-5-like (LOC109352001), mRNA	96.43%	80.5%	409.23
XM_020348274	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789350), mRNA	100.00%	80.7%	431.39
XM_017577495	PREDICTED: Vigna angularis metacaspase-5-like (LOC108340250), mRNA	99.11%	80.8%	427.69
XM_047312313	PREDICTED: Vigna umbellata metacaspase-5-like (LOC124837030), mRNA	99.11%	80.8%	427.69
XM_007133394	Phaseolus vulgaris hypothetical protein (PHAVU_011G180200g) mRNA, complete cds	99.11%	81.0%	433.23
XM_045944455	PREDICTED: Trifolium pratense metacaspase-5-like (LOC123894447), mRNA	97.86%	81.2%	431.39
XM_004509944	PREDICTED: Cicer arietinum metacaspase-4-like (LOC101509880), mRNA	100.00%	81.4%	448.01
XM_027500219	PREDICTED: Abrus precatorius metacaspase-5-like (LOC113865563), transcript variant X1, mRNA	90.54%	81.6%	416.61
XM_003531762	PREDICTED: Glycine max metacaspase-4 (LOC100795043), mRNA	98.57%	81.7%	455.39
XM_028389058	PREDICTED: Glycine soja metacaspase-4-like (LOC114422611), mRNA	98.57%	81.7%	455.39
XM_003546612	PREDICTED: Glycine max metacaspase-4 (LOC100802658), mRNA	99.11%	82.0%	466.47
XM_028348014	PREDICTED: Glycine soja metacaspase-4-like (LOC114387810), mRNA	99.11%	82.0%	466.47
XM_007133396	Phaseolus vulgaris hypothetical protein (PHAVU_011G180300g) mRNA, complete cds	99.64%	82.6%	484.94
XM_027499192	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X1, mRNA	97.50%	83.8%	512.64

Table 2: BLASTn result with *RpMC-IIb* as query.

Accession	Description	Query coverage	% Pairwise Identity	Bit-Score
XM_007133394	Phaseolus vulgaris hypothetical protein (PHAVU_011G180200g) mRNA, complete cds	99.14%	81.1%	355.677
XM_027500219	PREDICTED: Abrus precatorius metacaspase-5-like (LOC113865563), transcript variant X1, mRNA	99.14%	81.5%	372.297
XM_020348272	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789349), transcript variant X1, mRNA	96.98%	81.8%	370.45
XM_020348273	PREDICTED: Cajanus cajan metacaspase-5 (LOC109789349), transcript variant X2, mRNA	96.98%	81.8%	370.45
NM_001250818	Glycine max putative metacaspase (LOC100527809), mRNA	93.09%	82.0%	363.064
XM_028389055	PREDICTED: Glycine soja metacaspase-5-like (LOC114422609), mRNA	93.09%	82.0%	363.064
XM_047312313	PREDICTED: Vigna umbellata metacaspase-5-like (LOC124837030), mRNA	90.71%	82.0%	350.137
XM_045944455	PREDICTED: Trifolium pratense metacaspase-5-like (LOC123894447), mRNA	90.50%	82.1%	351.984
XM_027500220	PREDICTED: Abrus precatorius metacaspase-5-like (LOC113865563), transcript variant X2, mRNA	90.28%	82.1%	351.984
XM_019593760	PREDICTED: Lupinus angustifolius metacaspase-5-like (LOC109352001), mRNA	91.58%	82.2%	359.37
XM_004509944	PREDICTED: Cicer arietinum metacaspase-4-like (LOC101509880), mRNA	95.03%	82.7%	379.683
XM_028389058	PREDICTED: Glycine soja metacaspase-4-like (LOC114422611), mRNA	96.33%	82.7%	387.07
XM_003531762	PREDICTED: Glycine max metacaspase-4 (LOC100795043), mRNA	96.33%	83.0%	392.61
XM_003546612	PREDICTED: Glycine max metacaspase-4 (LOC100802658), mRNA	97.84%	83.0%	399.997
XM_028348014	PREDICTED: Glycine soja metacaspase-4-like (LOC114387810), mRNA	97.84%	83.0%	399.997
XM_007133396	Phaseolus vulgaris hypothetical protein (PHAVU_011G180300g) mRNA, complete cds	91.36%	83.1%	379.683
XM_027499192	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X1, mRNA	99.14%	83.4%	416.616
XM_027499193	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X2, mRNA	84.67%	84.4%	385.223

Table 3: BLASTn result with RpMC-IIc as query.

Accession	Description	Query coverage	% Pairwise Identity	Bit-Score
XM_013589850	PREDICTED: Medicago truncatula metacaspase-5 (LOC25502442), mRNA	97.20%	79.3%	222.718
XM_027500219	PREDICTED: Abrus precatorius metacaspase-5-like (LOC113865563), transcript variant X1, mRNA	96.64%	80.9%	267.038
NM_001250818	Glycine max putative metacaspase (LOC100527809), mRNA	96.64%	81.1%	274.424
XM_028389055	PREDICTED: Glycine soja metacaspase-5-like (LOC114422609), mRNA	96.64%	81.1%	274.424
XM_003531762	PREDICTED: Glycine max metacaspase-4 (LOC100795043), mRNA	99.16%	82.2%	291.044
XM_028389058	PREDICTED: Glycine soja metacaspase-4-like (LOC114422611), mRNA	99.16%	82.2%	291.044
XM_004509944	PREDICTED: Cicer arietinum metacaspase-4-like (LOC101509880), mRNA	97.76%	83.2%	309.511
XM_003546612	PREDICTED: Glycine max metacaspase-4 (LOC100802658), mRNA	100.00%	84.3%	335.364
XM_028348014	PREDICTED: Glycine soja metacaspase-4-like (LOC114387810), mRNA	100.00%	84.3%	335.364
XM_014634744	PREDICTED: Vigna radiata var. radiata metacaspase-5 (LOC106752959), mRNA	99.16%	84.6%	350.137
XM_017575836	PREDICTED: Vigna angularis metacaspase-5-like (LOC108338807), mRNA	99.16%	84.7%	355.677
XM_047312309	PREDICTED: Vigna umbellata metacaspase-5-like (LOC124837025), mRNA	99.16%	84.7%	355.677
XM_027499192	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X1, mRNA	99.16%	84.9%	350.137
XM_027500220	PREDICTED: Abrus precatorius metacaspase-5-like (LOC113865563), transcript variant X2, mRNA	27.17%	87.6%	113.766
XM_027499193	PREDICTED: Abrus precatorius metacaspase-4-like (LOC113864938), transcript variant X2, mRNA	21.85%	88.5%	952.994

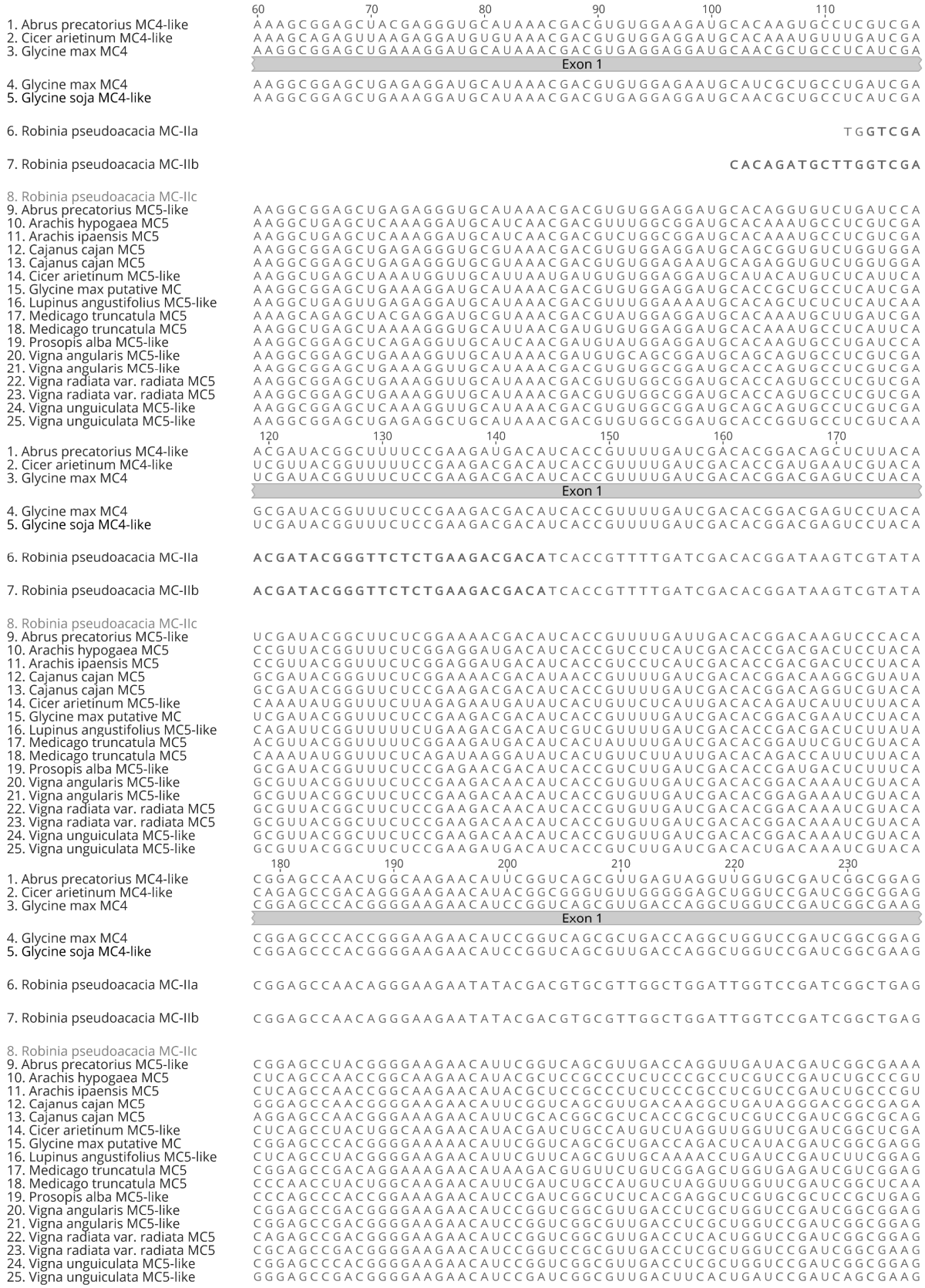


Figure 4: DNA sequence alignment with RpMC-II, and homologs of MC4 and MC5 in Fabales (page 1).

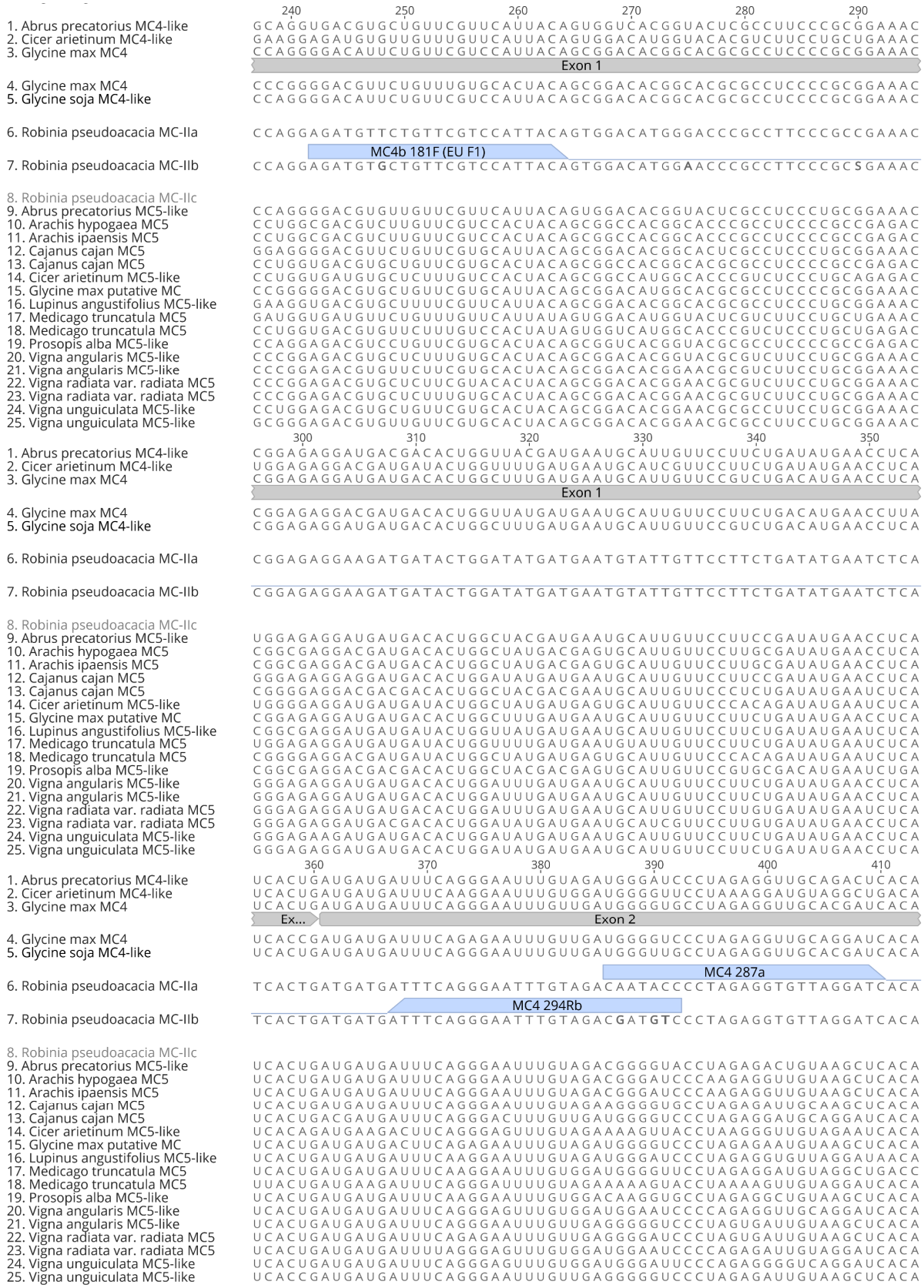


Figure 4: DNA sequence alignment with RpMC-II, and homologs of MC4 and MC5 in Fabales (page 2).

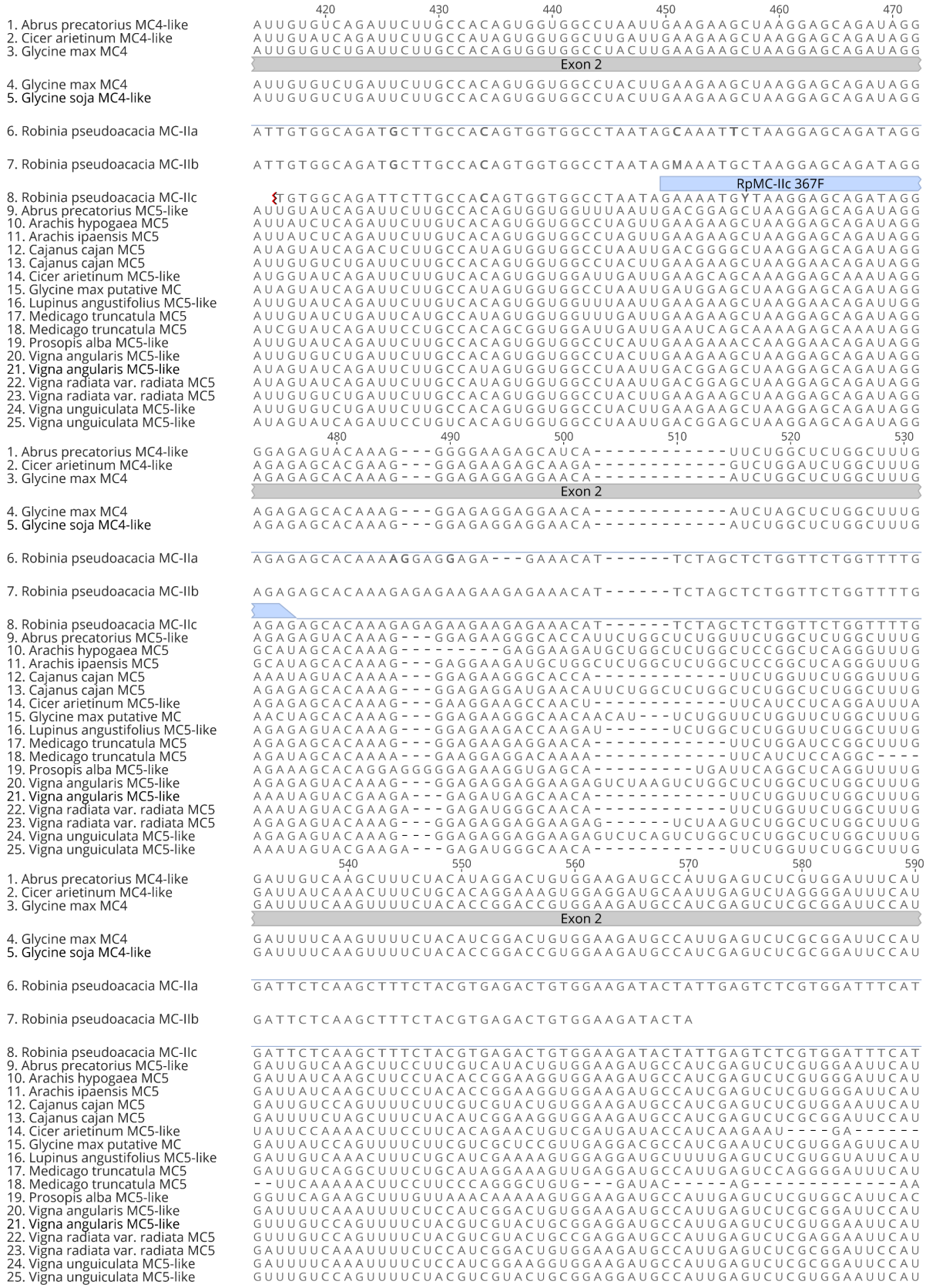


Figure 4: DNA sequence alignment with RpMC-II, and homologs of MC4 and MC5 in Fabales (page 3).

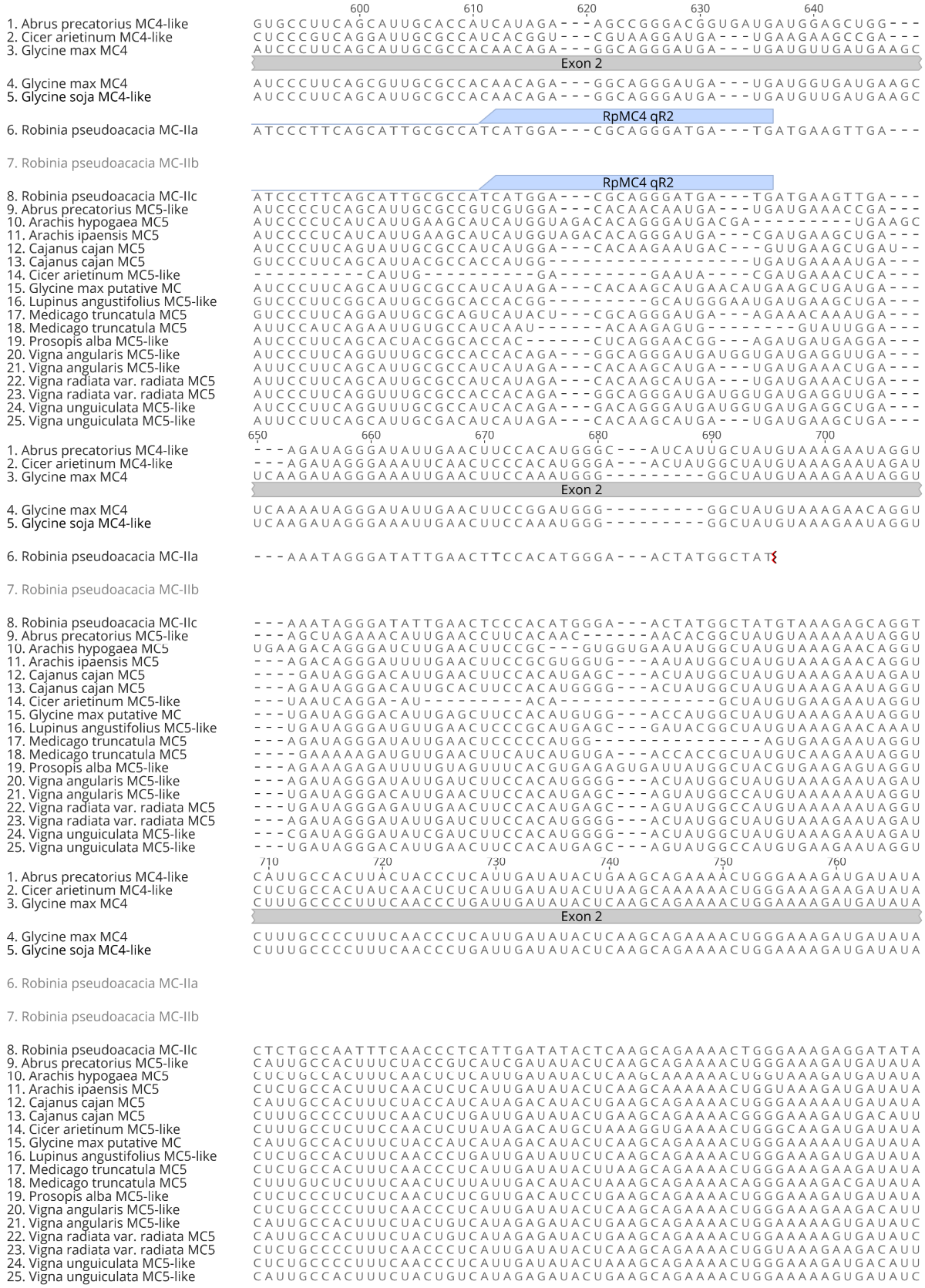


Figure 4: DNA sequence alignment with RpMC-II, and homologs of MC4 and MC5 in Fabales (page 4).

Figure 4: DNA sequence alignment with RpMC-II, and homologs of MC4 and MC5 in Fabales (page 5).

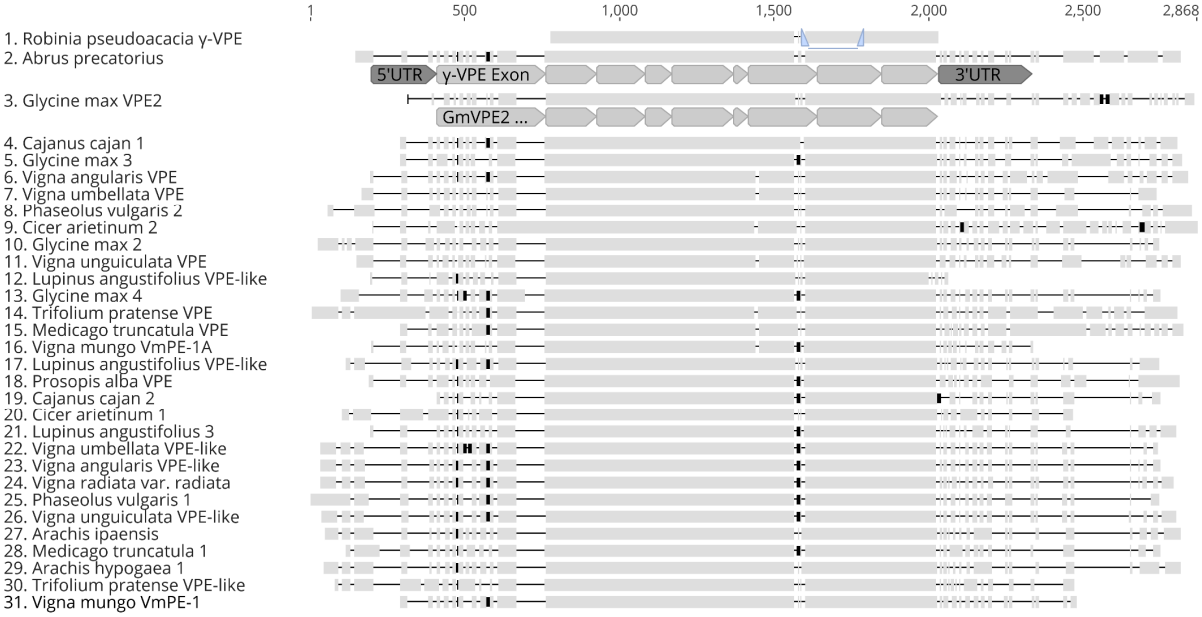


Figure 5: γ-RpVPE and homologs DNA sequence alignment overview, sequence information table on BLASTn results is listed in Table 4

Table 4: Information on the sequences in Figure 5

Lane	Sequence label	Accession	Description	Query coverage	% Pairwise Identity	Bit-Score
1	Robinia pseudoacacia	--	Robinia pseudoacacia γ-VPE	--	--	--
2	Abrus precatorius	XM_027513192.1	PREDICTED: Abrus precatorius vacuolar-processing enzyme-like (LOC113874911), transcript variant X1, mRNA	92.05%	90.6%	1716.66
3	Glycine max VPE2	NM_001249635.	Glycine max vacuolar processing enzyme 2 (VPE2), mRNA	96.05%	90.5%	1703.67
4	Cajanus cajan 1	XM_020366661.2	PREDICTED: Cajanus cajan vacuolar-processing enzyme (LOC109804800), mRNA	92.05%	89.0%	1598.47
5	Glycine max 3	XM_003550235.5	PREDICTED: Glycine max vacuolar-processing enzyme (LOC100784318), mRNA	97.16%	88.8%	1659.41
6	Vigna angularis VPE	XM_017574898.1	PREDICTED: Vigna angularis vacuolar-processing enzyme (LOC108338162), transcript variant X1, mRNA	91.77%	88.1%	1526.45
7	Vigna umbellata VPE	XM_047306939.1	PREDICTED: Vigna umbellata vacuolar-processing enzyme (LOC124832678), mRNA	91.77%	88.1%	1526.45
8	Phaseolus vulgaris 2	XM_007160923.1	Phaseolus vulgaris hypothetical protein (PHAVU_001G033500g) mRNA, complete cds	91.70%	87.8%	1506.14
9	Cicer arietinum 2	XM_027333646.1	PREDICTED: Cicer arietinum vacuolar-processing enzyme-like (LOC101514688), transcript variant X1, mRNA	92.05%	87.8%	1511.68
10	Glycine max 2	XM_003525931.5	PREDICTED: Glycine max vacuolar-processing enzyme (LOC100791675), mRNA	88.79%	87.8%	1459.97
11	Vigna unguiculata VPE	XM_028084964.1	PREDICTED: Vigna unguiculata vacuolar-processing enzyme (LOC114194621), mRNA	91.77%	87.6%	1491.37
12	Lupinus angustifolius VPE-like	XM_019573985.1	PREDICTED: Lupinus angustifolius vacuolar-processing enzyme-like (LOC109337103), mRNA	84.74%	87.1%	1345.48
13	Glycine max 4	XM_014774474.3	PREDICTED: Glycine max vacuolar-processing enzyme (LOC100807010), transcript variant X1, mRNA	88.79%	87.0%	1406.42
14	Trifolium pratense VPE	XM_045968705.1	PREDICTED: Trifolium pratense vacuolar-processing enzyme (LOC123917080), mRNA	88.86%	86.8%	1387.95
15	Medicago truncatula VPE	XM_013610451.3	PREDICTED: Medicago truncatula vacuolar-processing enzyme (LOC25482000), mRNA	91.34%	86.8%	1417.5
16	Vigna mungo VmPE-1A	D89972.1	Vigna mungo mRNA for asparaginyl endopeptidase (VmPE-1A), complete cds	91.77%	86.7%	1426.73
17	Lupinus angustifolius VPE-like	XM_019561080.1	PREDICTED: Lupinus angustifolius vacuolar-processing enzyme-like (LOC109327911), mRNA	89.99%	86.5%	1389.8
18	Prosopis alba VPE	XM_028940230.1	PREDICTED: Prosopis alba vacuolar-processing enzyme (LOC114751576), mRNA	87.01%	86.5%	1341.79
19	Cajanus cajan 2	XM_020368305.2	PREDICTED: Cajanus cajan vacuolar-processing enzyme (LOC109805986), mRNA	88.22%	86.1%	1336.25
20	Cicer arietinum 1	XM_004501420.3	PREDICTED: Cicer arietinum vacuolar-processing enzyme-like (LOC101515346), mRNA	88.72%	85.7%	1308.55
21	Lupinus angustifolius 3	XM_019582316.1	PREDICTED: Lupinus angustifolius vacuolar-processing enzyme-like (LOC109343833), mRNA	89.07%	85.7%	1315.93
22	Vigna umbellata VPE-like	XM_047316214.1	PREDICTED: Vigna umbellata vacuolar-processing enzyme-like (LOC124840200), mRNA	88.72%	85.6%	1310.39
23	Vigna angularis VPE-like	XM_017566611.1	PREDICTED: Vigna angularis vacuolar-processing enzyme-like (LOC108331721), mRNA	88.72%	85.6%	1304.85
24	Vigna radiata var. radiata	XM_014645928.2	PREDICTED: Vigna radiata var. radiata vacuolar-processing enzyme (LOC106762174), mRNA	88.72%	85.5%	1299.31
25	Phaseolus vulgaris 1	XM_007136747.1	Phaseolus vulgaris hypothetical protein (PHAVU_009G076100g) mRNA, complete cds	88.72%	85.4%	1293.78
26	Vigna unguiculata VPE-like	XM_028047286.1	PREDICTED: Vigna unguiculata vacuolar-processing enzyme-like (LOC114163155), mRNA	88.72%	85.2%	1279.0
27	Arachis ipaensis	XM_016308381.2	PREDICTED: Arachis ipaensis vacuolar-processing enzyme (LOC107606332), mRNA	90.70%	85.2%	1304.85
28	Medicago truncatula 1	XM_003603073.4	PREDICTED: Medicago truncatula vacuolar-processing enzyme gamma-isozyme (LOC11432331), mRNA	87.51%	84.8%	1234.68
29	Arachis hypogaea 1	XM_025757656.2	PREDICTED: Arachis hypogaea vacuolar-processing enzyme (LOC112706381), mRNA	90.70%	84.7%	1271.62
30	Trifolium pratense VPE-like	XM_045944193.1	PREDICTED: Trifolium pratense vacuolar-processing enzyme-like (LOC123894249), mRNA	90.70%	84.7%	1271.62
31	Vigna mungo VmPE-1	D89971.1	Vigna mungo mRNA for asparaginyl endopeptidase (VmPE-1), complete cds	88.29%	84.6%	1232.84

-- Indicates not applicable

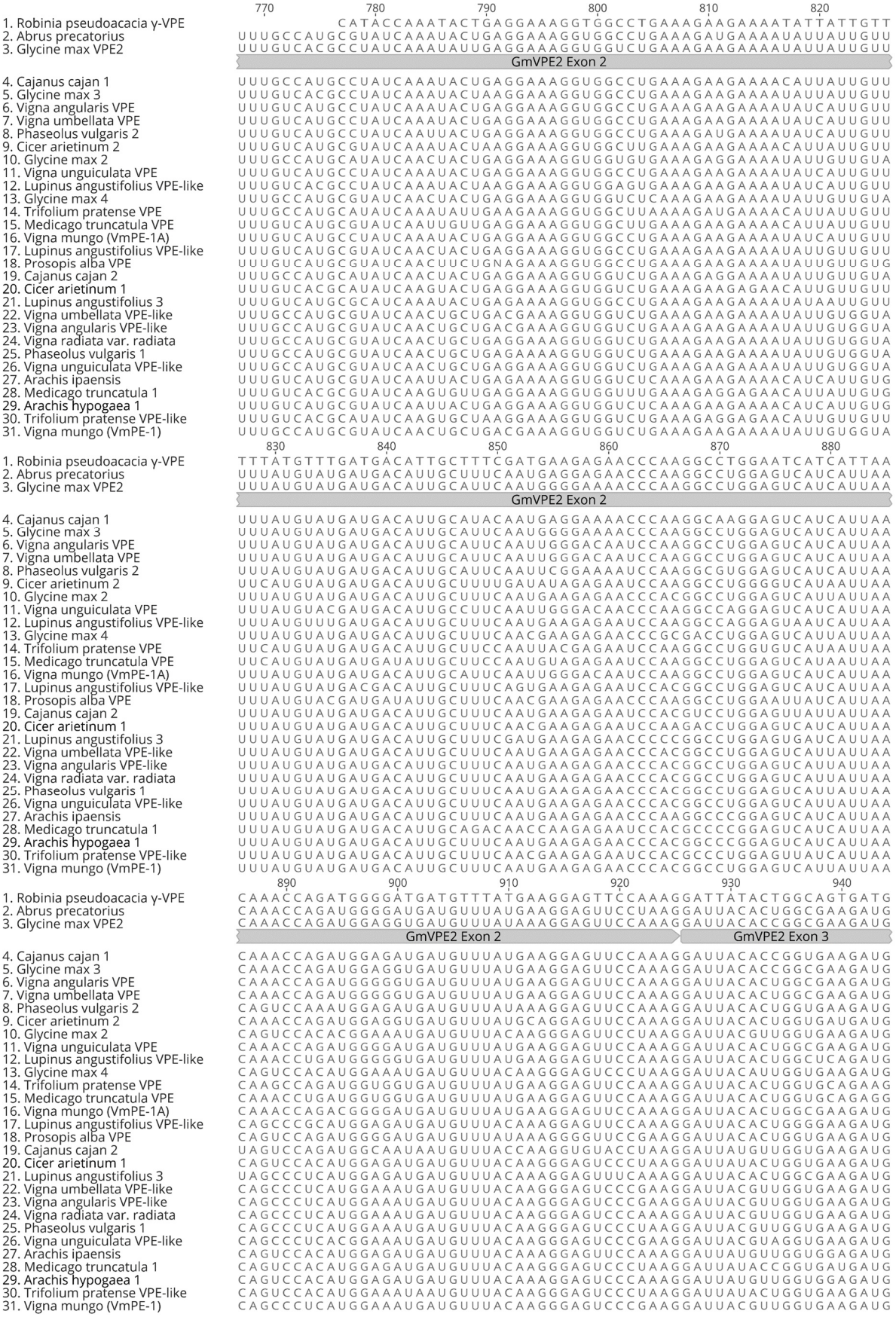


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 1).

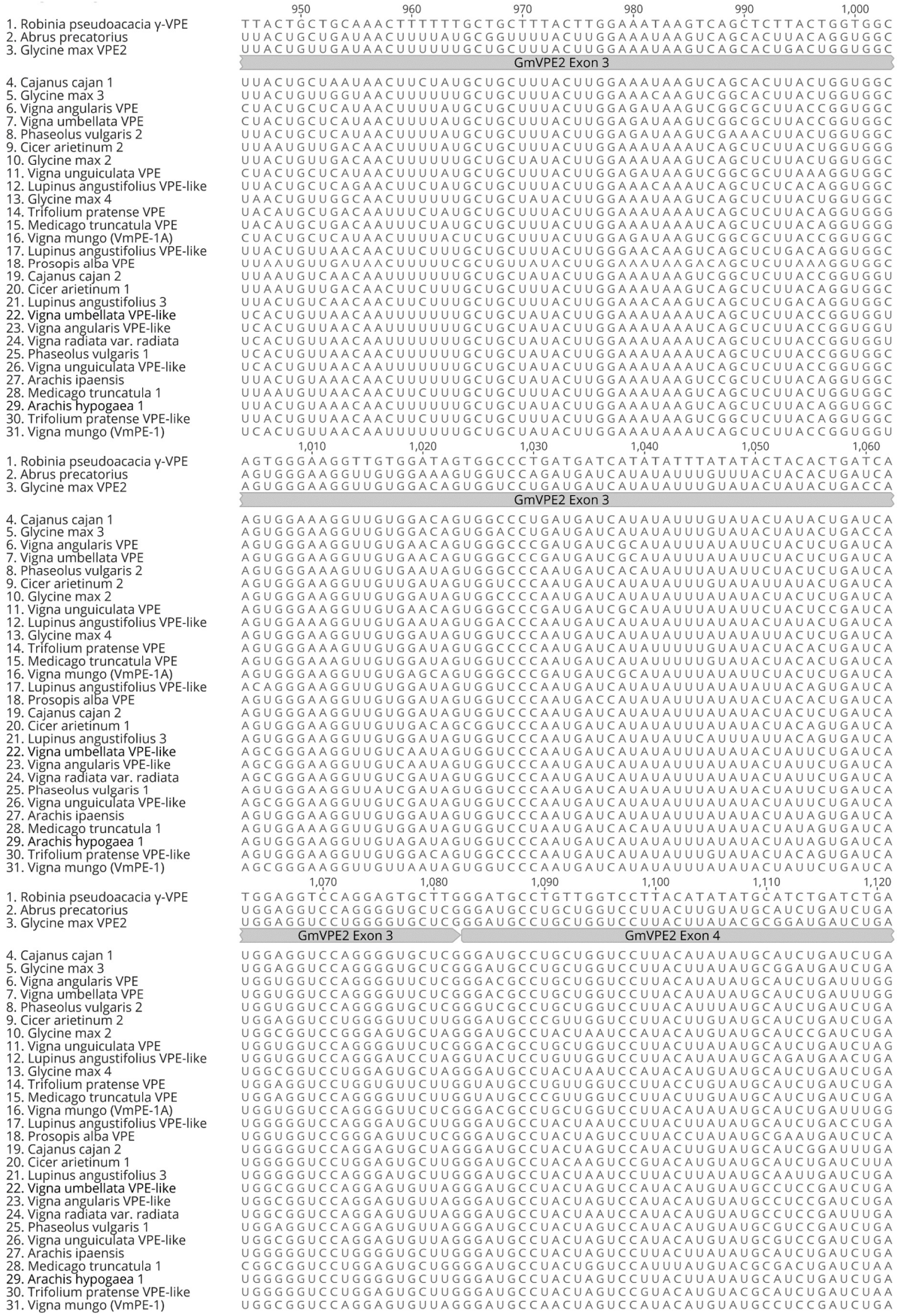


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 2).

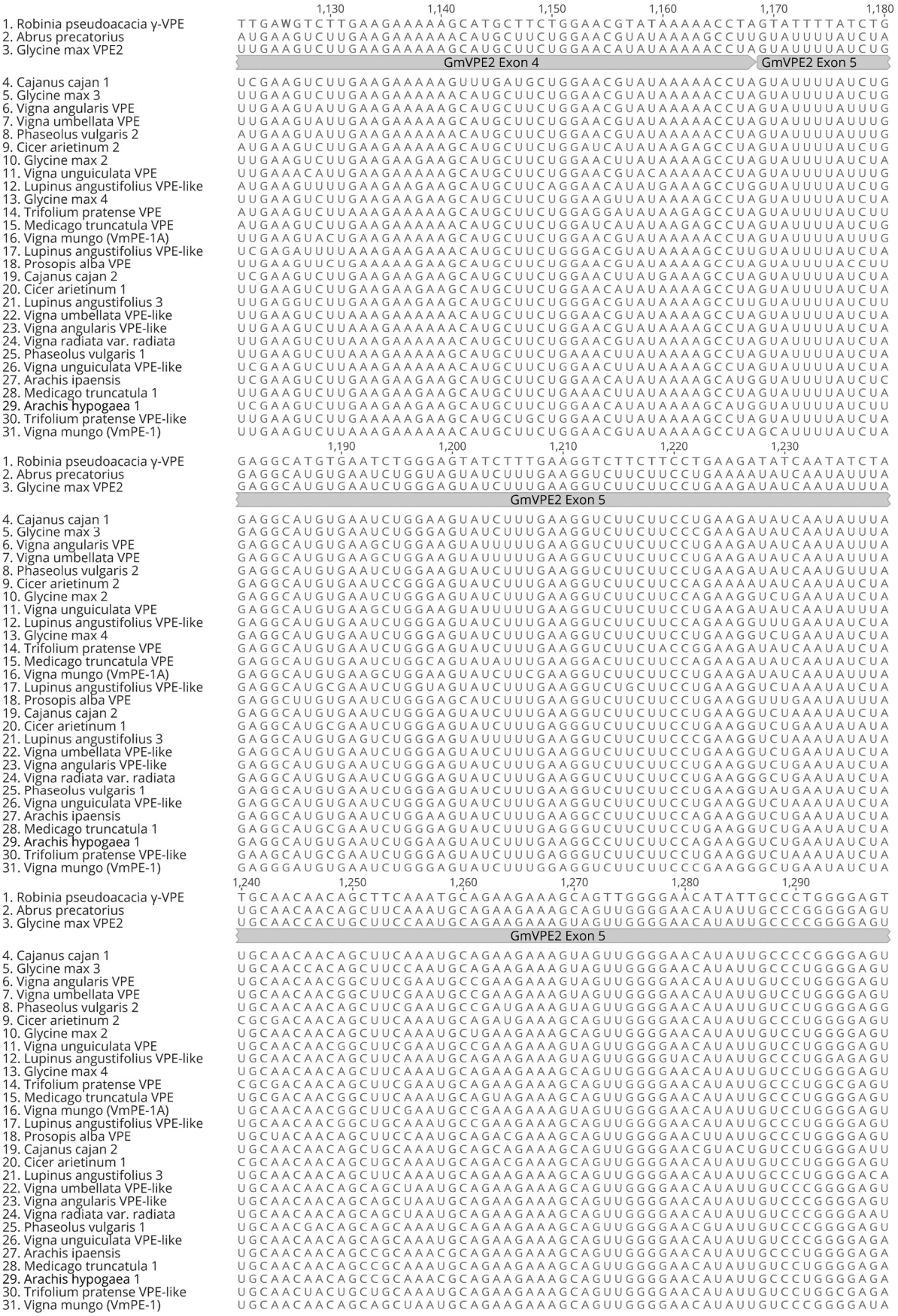


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 3).

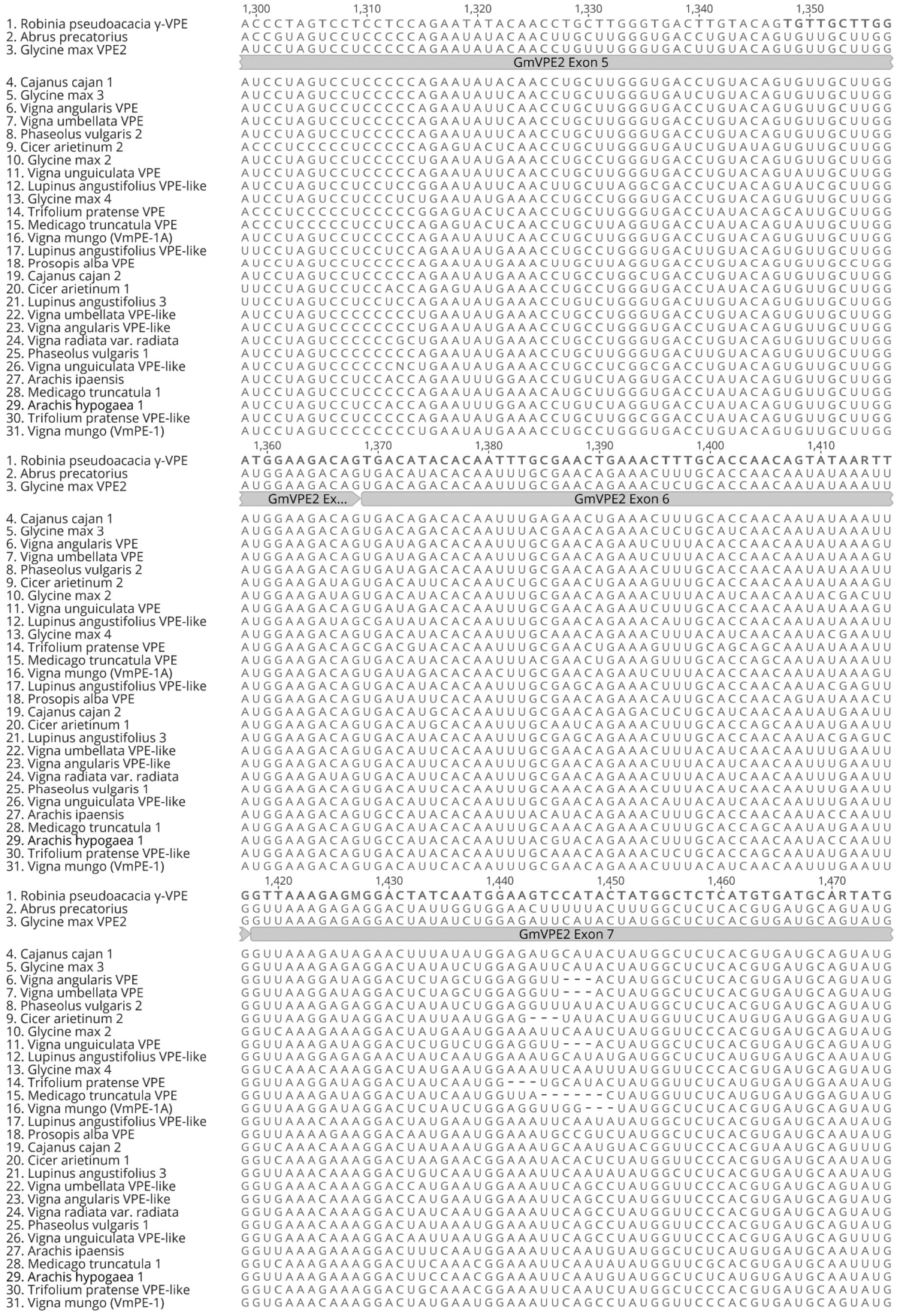


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 4).

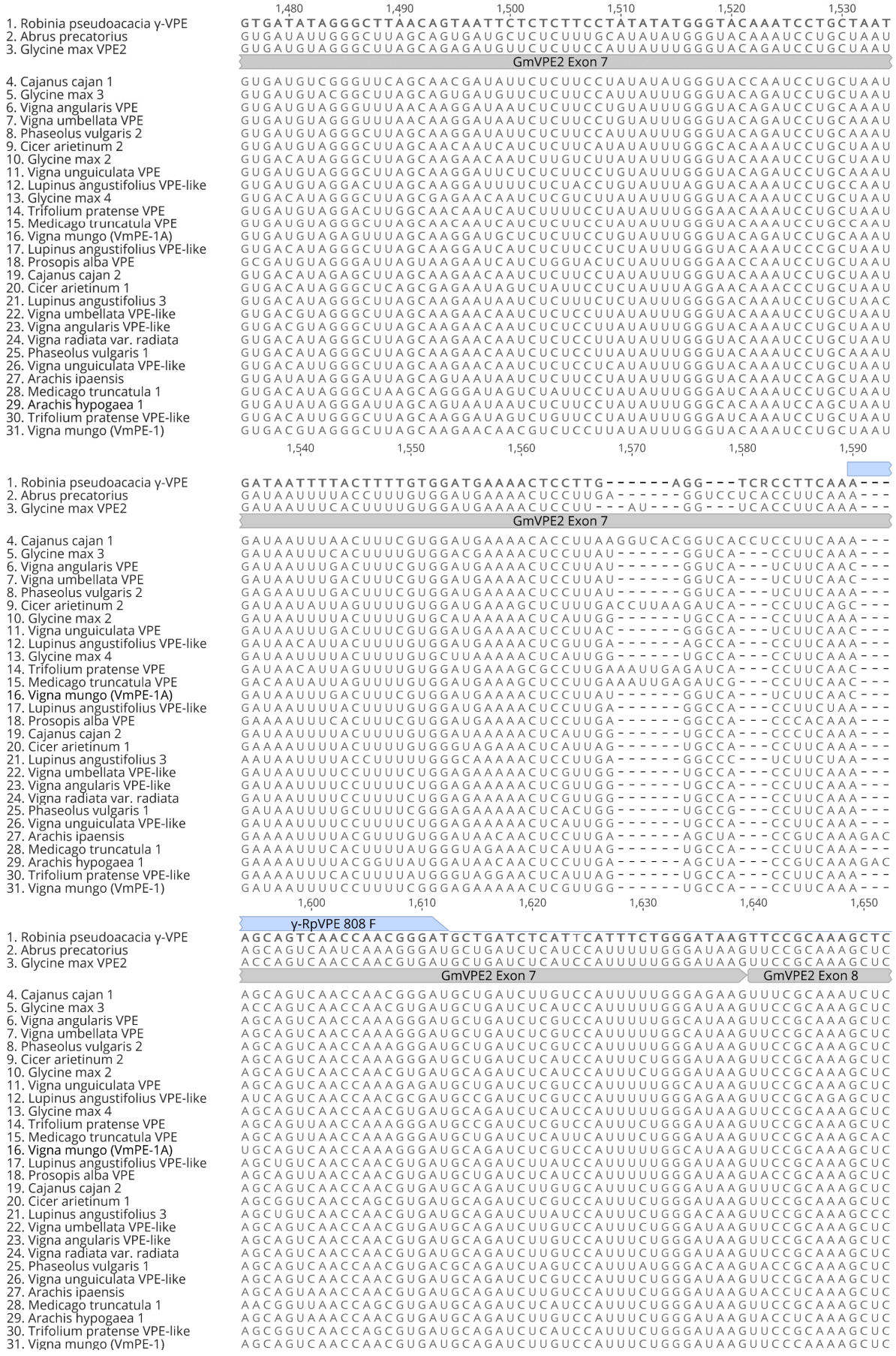


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 5).

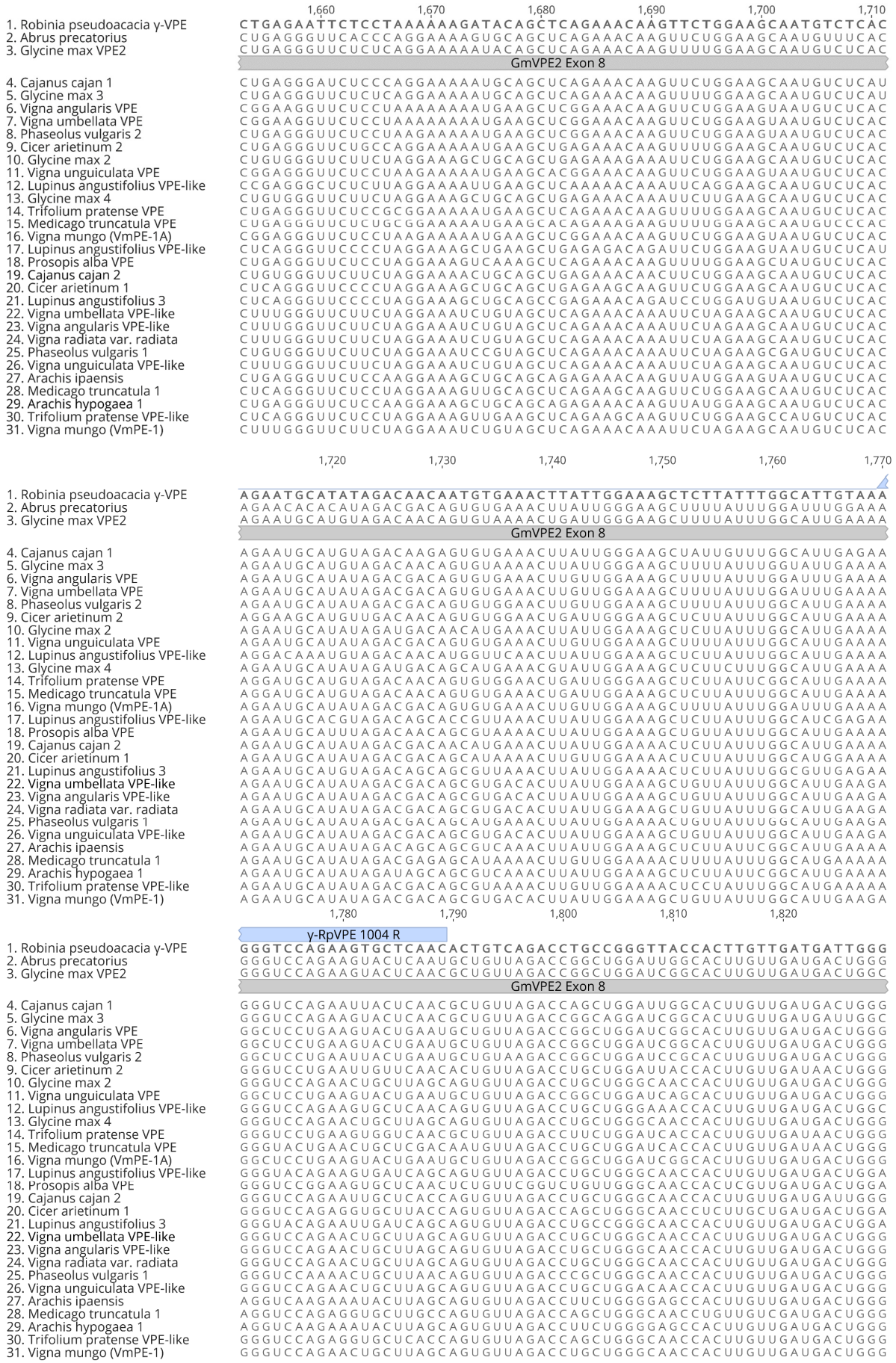


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 6).

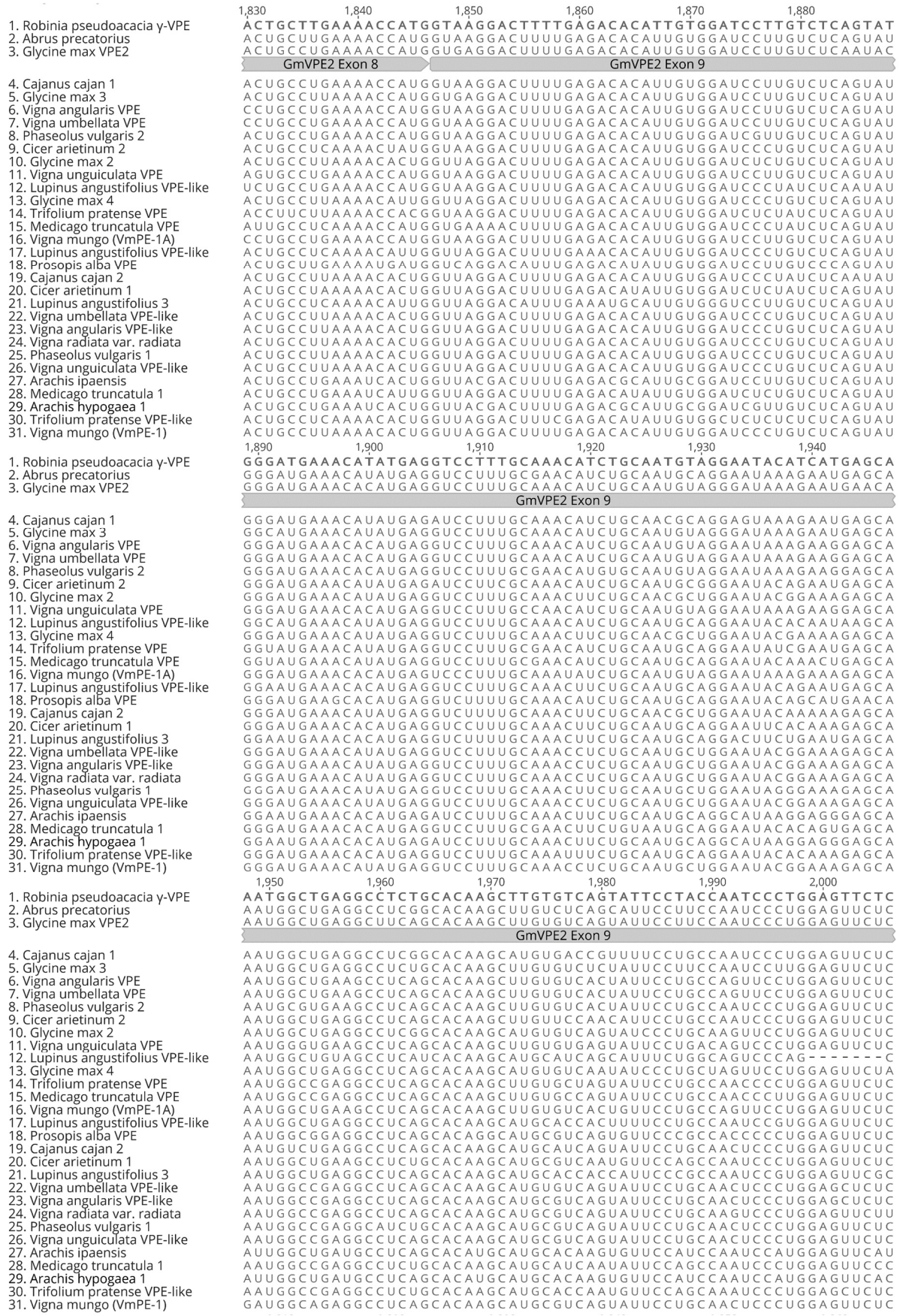


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 7).

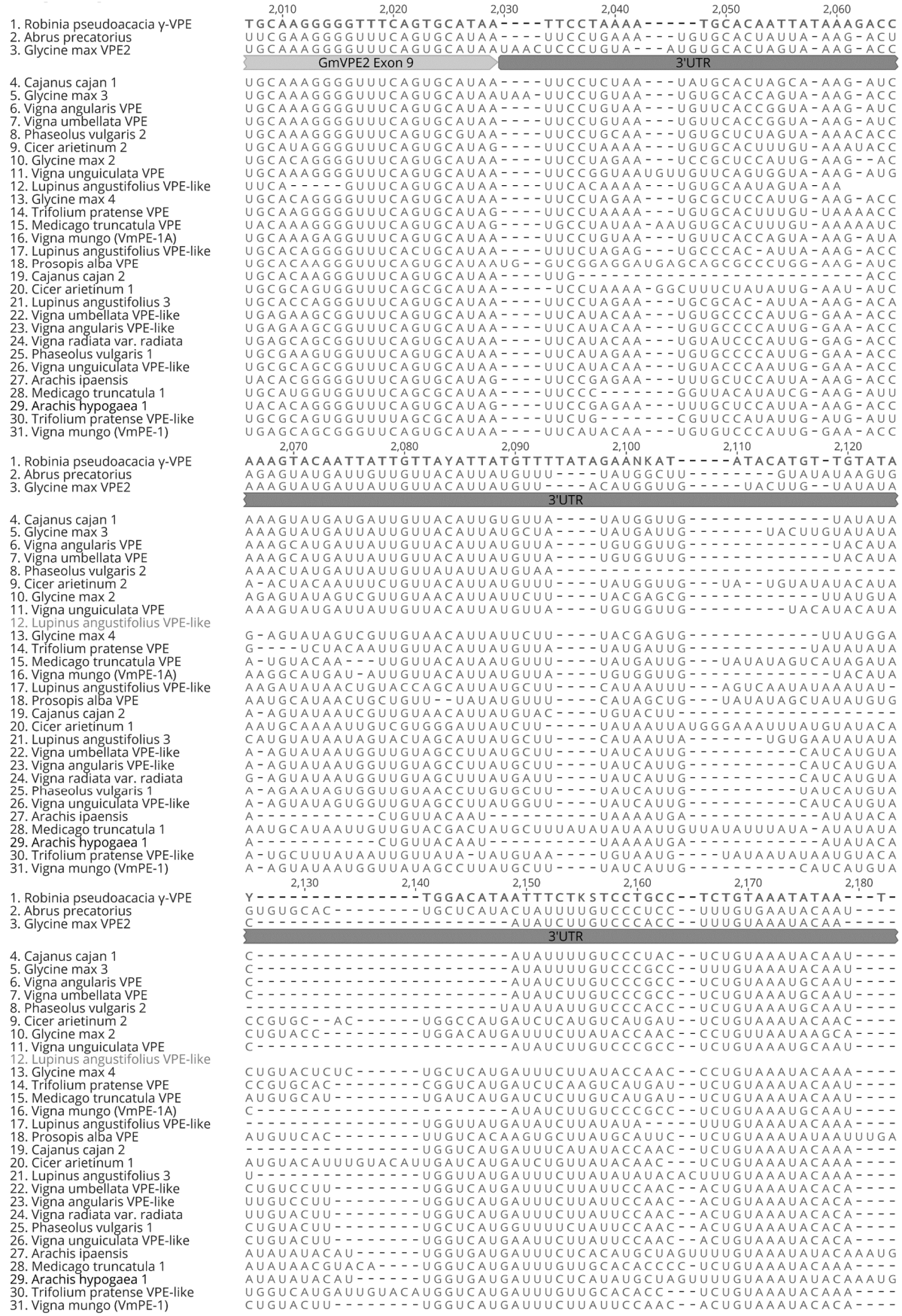
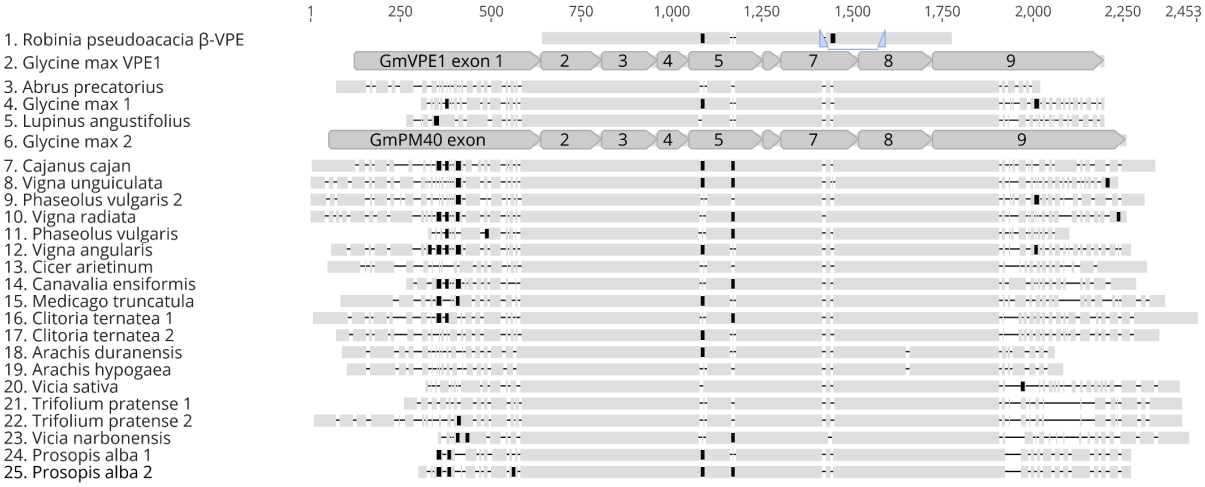


Figure 6: DNA sequence alignment with γ-RpVPE and homologs (page 8).

Figure 6: DNA sequence alignment with *γ-RpVPE* and homologs (page 9).



Lane	Sequence label	Accession	Description	Query coverage	% Pairwise Identity	Bit-Score
1	Robinia pseudoacacia β-VPE	--	Robinia pseudoacacia β-VPE	--	--	--
2	Glycine max VPE	NM_001249749.2	Glycine max vacuolar-processing enzyme (VPE), mRNA	100.00%	89.6%	1421.19
3	Abrus precatorius	XM_027479555.1	PREDICTED: Abrus precatorius legumain (LOC113849561), mRNA	100.00%	89.5%	1410.11
4	Glycine max 1	D28876.1	Glycine max mRNA for cysteine proteinase, complete cds	100.00%	89.5%	1410.11
5	Lupinus angustifolius	XM_019591087.1	PREDICTED: Lupinus angustifolius legumain (LOC109349992), mRNA	100.00%	89.3%	1399.03
6	Glycine max 2	NM_001349197.1	Glycine max seed maturation protein PM40 (PM40), mRNA	100.00%	89.1%	1387.95
7	Cajanus cajan	XM_020370788.2	PREDICTED: Cajanus cajan vacuolar-processing enzyme (LOC109808005), mRNA	100.00%	88.0%	1321.47
8	Vigna unguiculata	XM_028063059.1	PREDICTED: Vigna unguiculata vacuolar-processing enzyme (LOC114177634), mRNA	100.00%	88.0%	1319.63
9	Vigna radiata	XM_014654608.2	PREDICTED: Vigna radiata var. radiata vacuolar-processing enzyme (LOC106769123), mRNA	100.00%	87.8%	1308.55
10	Phaseolus vulgaris 2	XM_007155562.1	Phaseolus vulgaris hypothetical protein (PHAVU_003G217500g) mRNA, complete cds	100.00%	87.8%	1304.85
11	Phaseolus vulgaris	Z99957.1	Phaseolus vulgaris Moldavian encoding legumain-like proteinase precursor (clone p21b)	100.00%	87.8%	1304.85
12	Vigna angularis	XM_017554660.1	PREDICTED: Vigna angularis vacuolar-processing enzyme (LOC108322532), mRNA	100.00%	87.3%	1277.16
13	Cicer arietinum	XM_004515685.3	PREDICTED: Cicer arietinum legumain (LOC101491855), mRNA	100.00%	87.0%	1260.54
14	Canavalia ensiformis	D31787.1	Canavalia ensiformis mRNA for asparaginyl endopeptidase, complete cds	100.00%	87.0%	1255.0
15	Medicago truncatula	XM_003608718.4	PREDICTED: Medicago truncatula legumain (LOC11405547), mRNA	100.00%	86.8%	1243.92
16	Clitoria ternatea 1	MT468738.1	Clitoria ternatea clone ctr29014_c2_g1_i2 asparaginyl endopeptidase 2 mRNA, complete cds	100.00%	86.8%	1243.92
17	Clitoria ternatea 2	KR912009.1	Clitoria ternatea clone Ct_flwr_1565_g1_i1 asparaginyl endopeptidase mRNA, complete cds	100.00%	86.7%	1238.38
18	Arachis hypogaea	XM_025846712.2	PREDICTED: Arachis hypogaea legumain (LOC112803196), mRNA	100.00%	85.6%	1170.05
19	Arachis duranensis	XM_016107728.2	PREDICTED: Arachis duranensis legumain (LOC107487130), mRNA	100.00%	85.6%	1170.05
20	Vicia sativa	AJ007743.1	Vicia sativa mRNA for cysteine proteinase precursor	100.00%	85.2%	1142.35
21	Trifolium pratense 2	XM_045932441.1	PREDICTED: Trifolium pratense legumain (LOC123883594), transcript variant X1, mRNA	100.00%	84.8%	1122.04
22	Trifolium pratense 1	XM_045932442.1	PREDICTED: Trifolium pratense legumain (LOC123883594), transcript variant X2, mRNA	100.00%	84.8%	1122.04
23	Vicia narbonensis	Z99174.1	Vicia narbonensis mRNA for cysteine proteinase precursor	100.00%	84.7%	1116.5
24	Prosopis alba 2	XM_028906873.1	PREDICTED: Prosopis alba legumain-like (LOC114721067), mRNA	100.00%	84.1%	1077.72
25	Prosopis alba 1	XM_028931650.1	PREDICTED: Prosopis alba legumain-like (LOC114743439), mRNA	100.00%	84.1%	1077.72

-- Indicates not applicable

Figure 7: *β-RpVPE* and homologs DNA sequence alignment overview and sequence information table on BLASTn result with *β-RpVPE* as query.

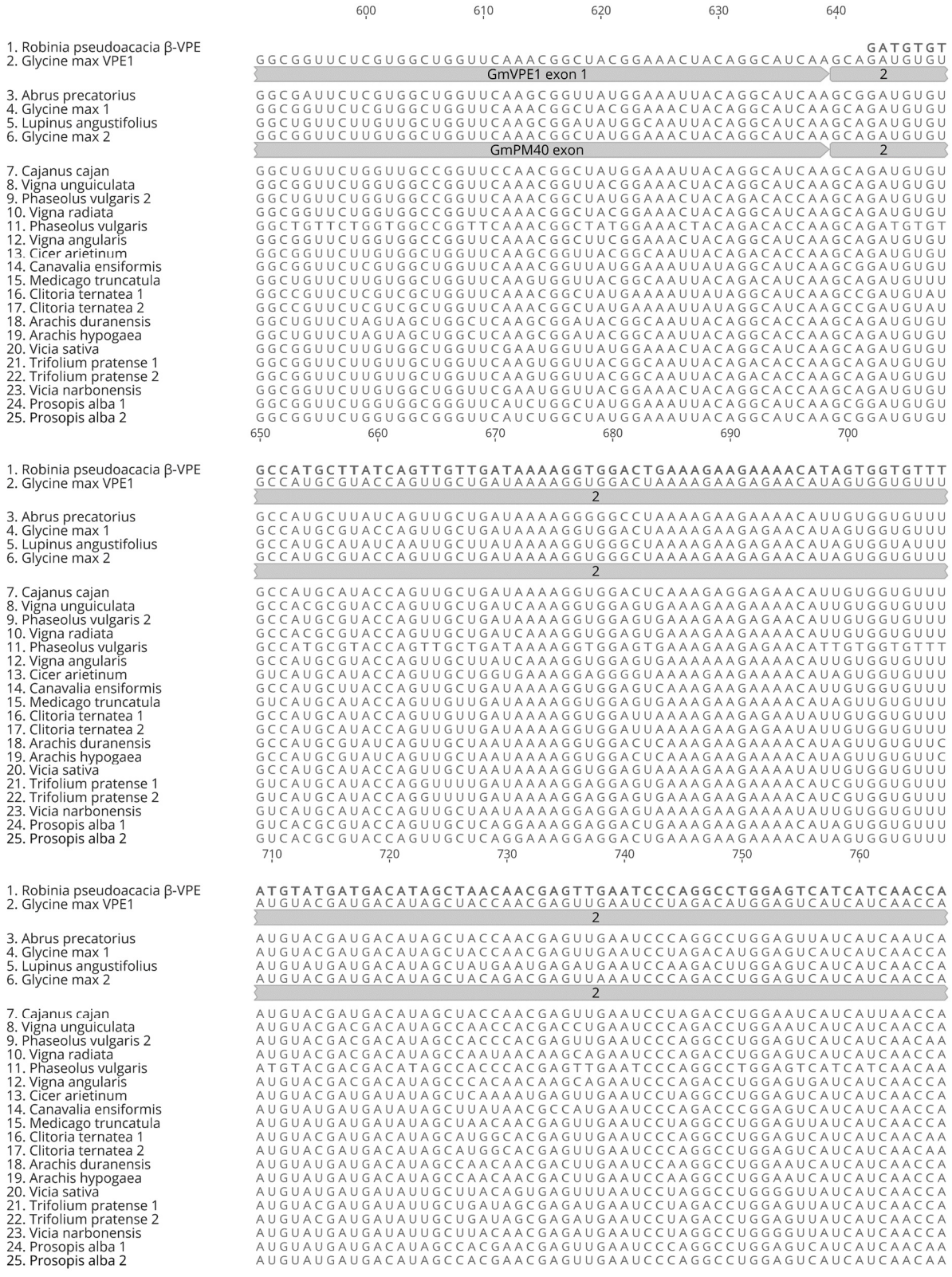


Figure 8: DNA sequence alignment with β-RpVPE and homologs (page 1).

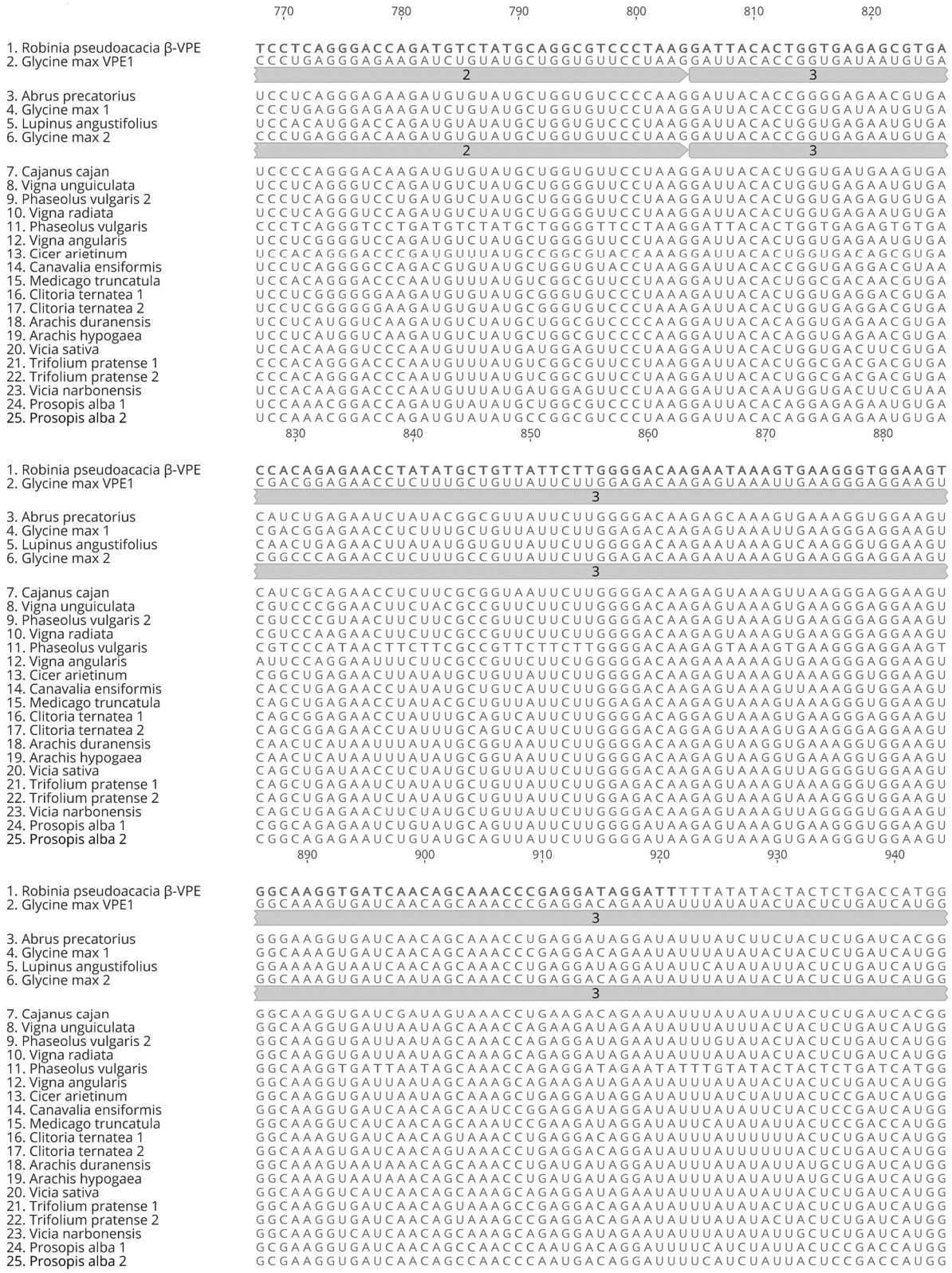


Figure 8: DNA sequence alignment with β-RpVPE and homologs (page 2).

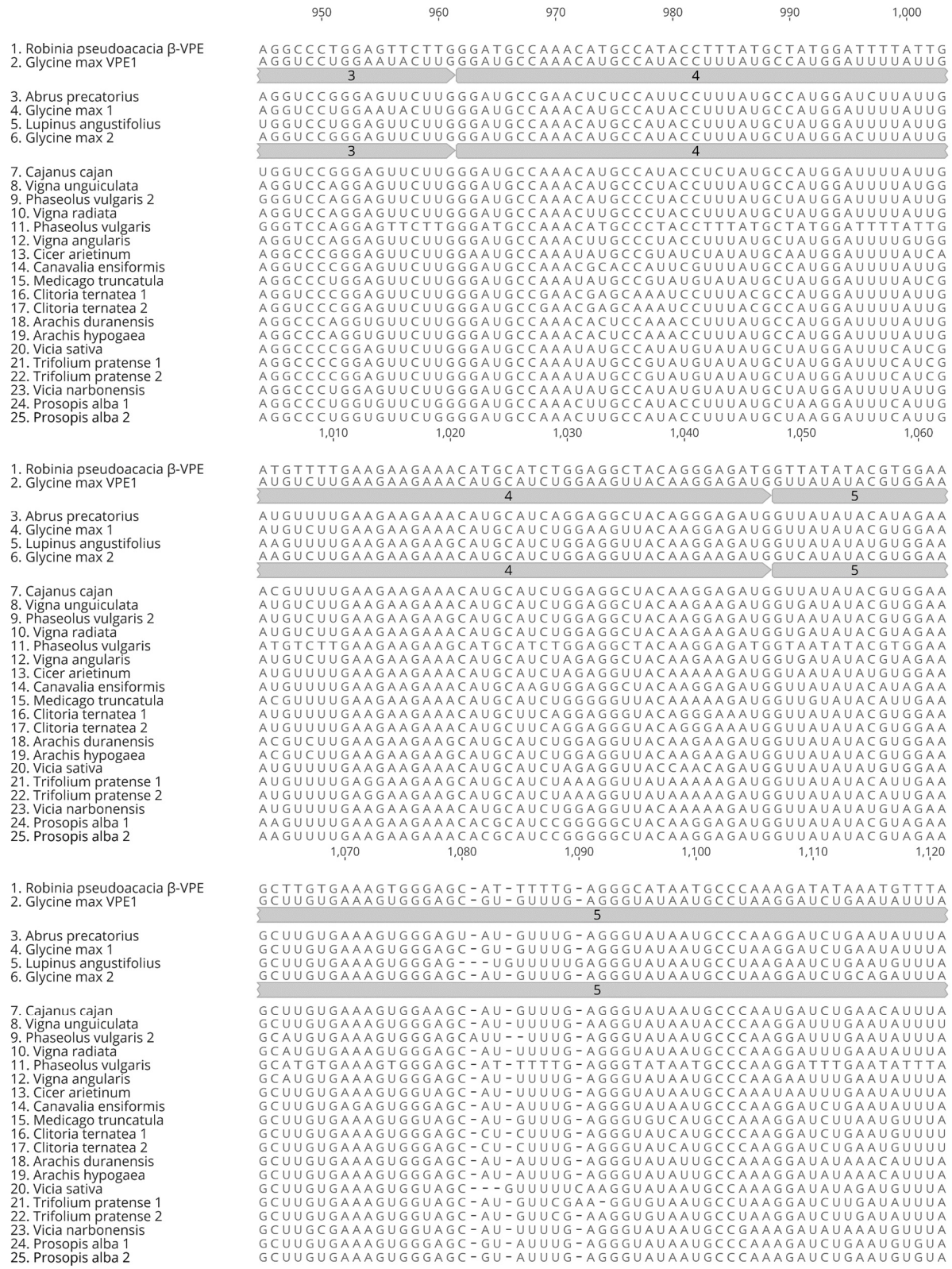


Figure 8: DNA sequence alignment with β-RpVPE and homologs (page 3).

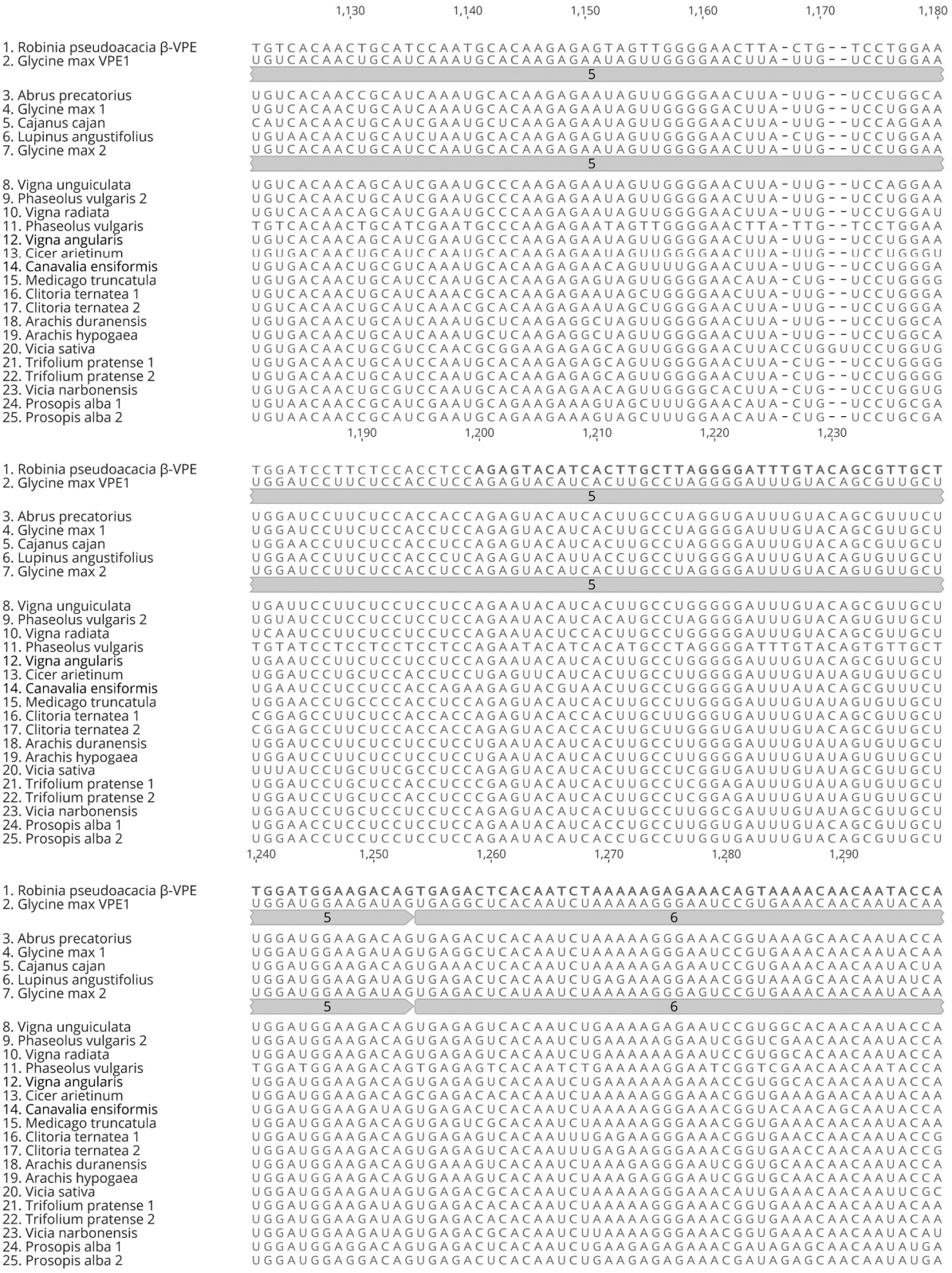


Figure 8: DNA sequence alignment with β-RpVPE and homologs (page 4).

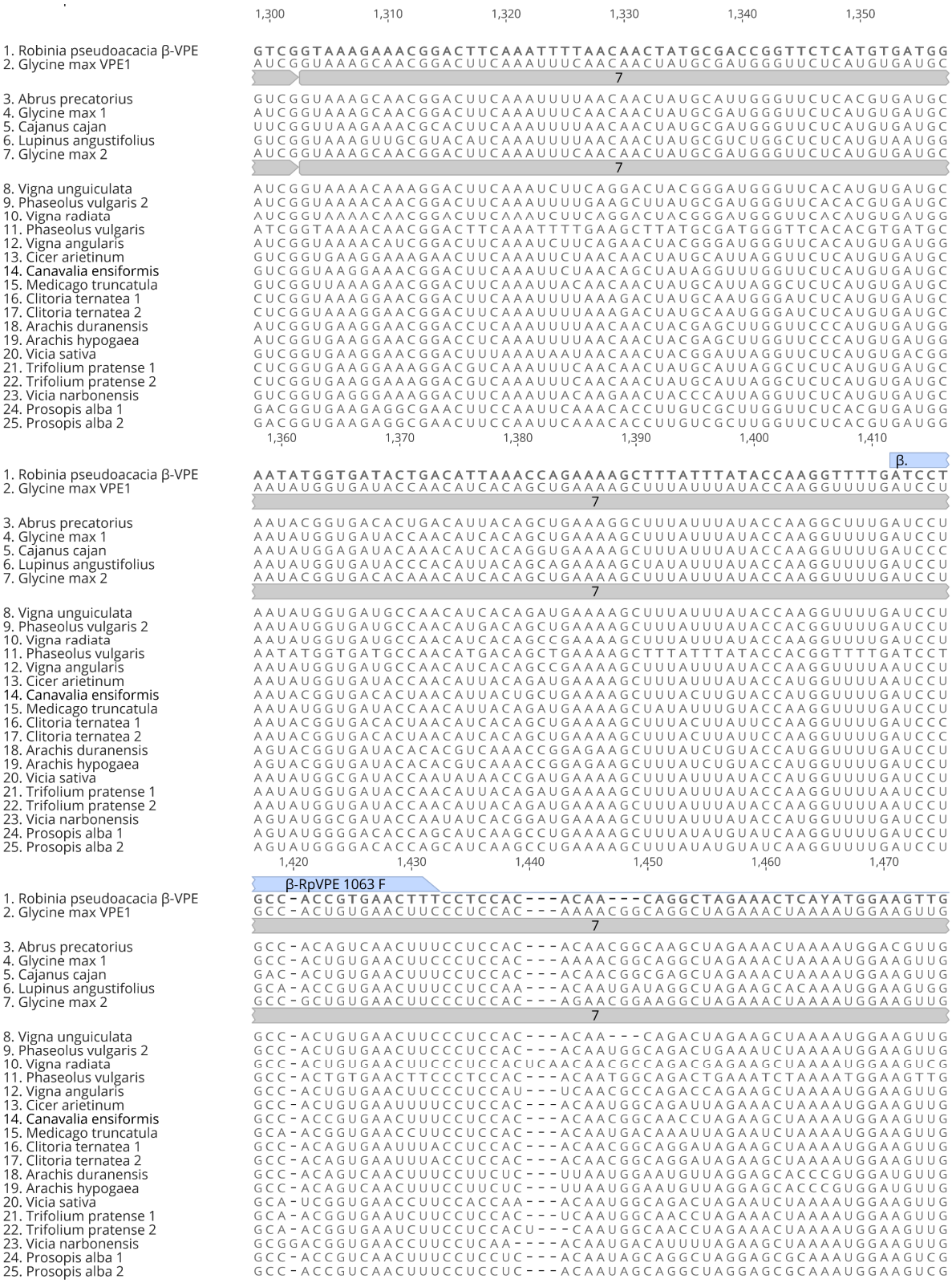


Figure 8: sequence alignment with β-RpVPE and homologs (page 5).

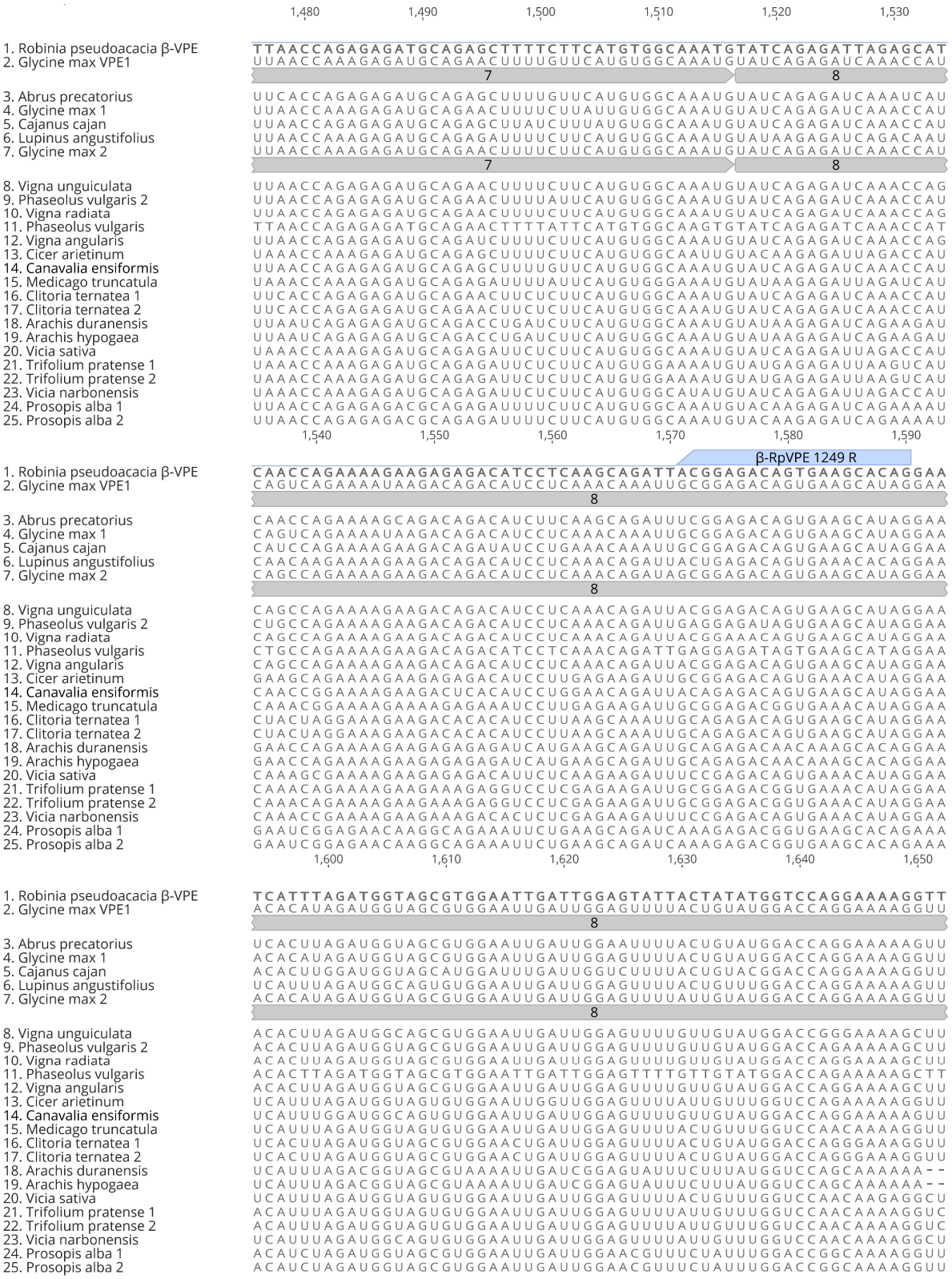


Figure 8: DNA sequence alignment with β-RpVPE and homologs (page 6).

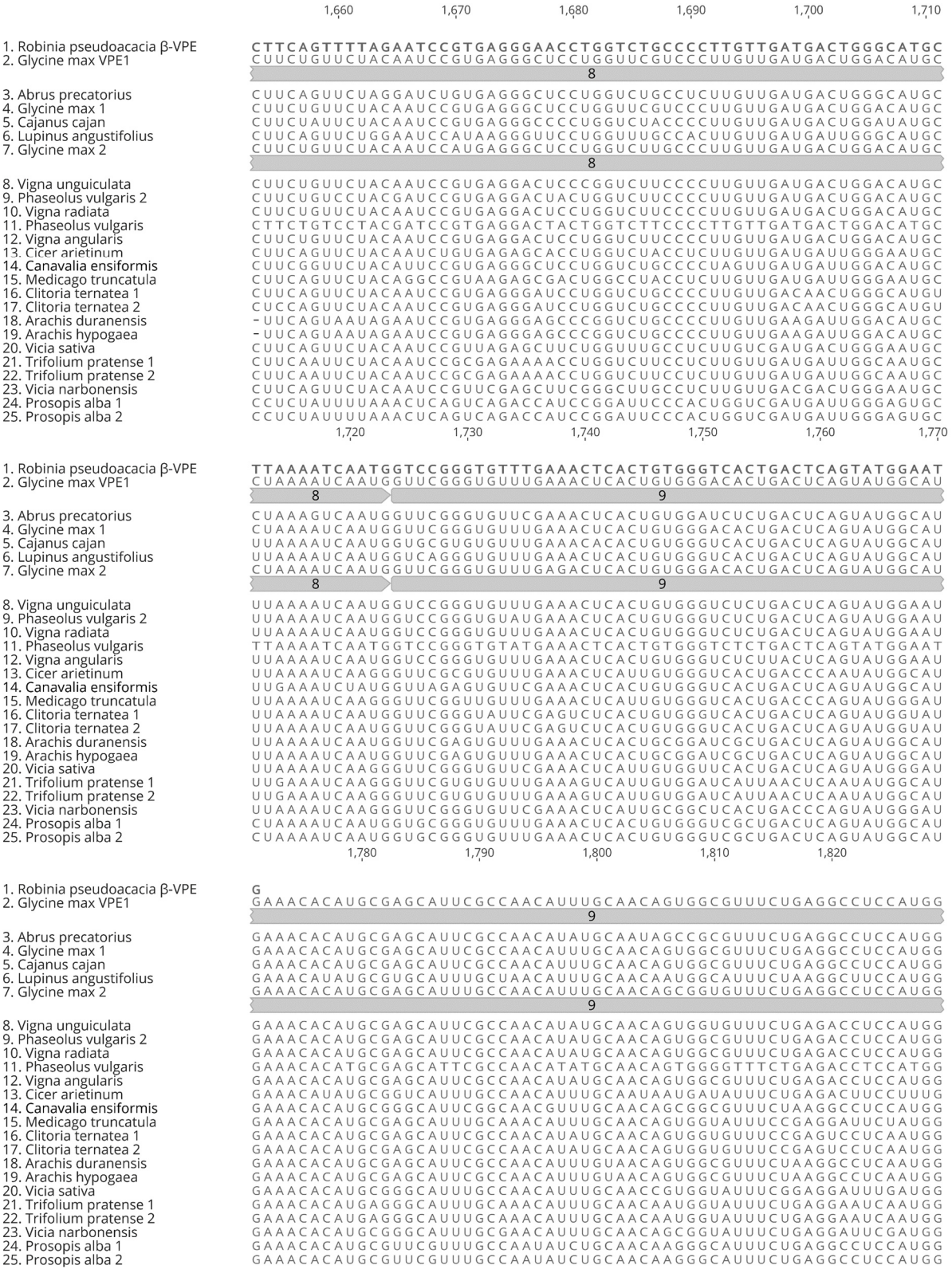


Figure 8: DNA sequence alignment with β-RpVPE and homologs (page 7).