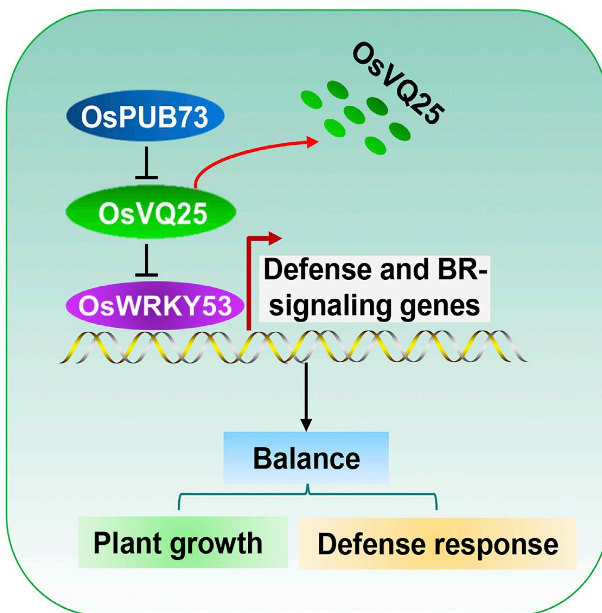
LIÈGE université
Gembloux
Agro-Bio Tech

Characterization of *VQ* gene family and associated plant defense pathways in rice



Author: Jinfu Tian

Supervisor: Prof. Frédéric Francis

Co-supervisor: Prof. Lanqin Xia

Year: 2024

COMMUNAUTÉ FRANÇAISE DE BELGIQUE

UNIVERSITÉ DE LIÈGE – GEMBLOUX AGRO-BIO TECH

Characterization of *VQ* gene family and associated plant defense pathways in rice

Jinfu Tian

Dissertation originale présentée en vue de l'obtention du grade de docteur en sciences agronomiques et ingénierie biologique

Promoteurs: Prof. Frédéric Francis
Prof. Lanqin Xia

Année civile: 2024

Copyright. Cette œuvre est sous licence Creative Commons. Vous êtes libre de reproduire, de modifier, de distribuer et de communiquer cette création au public selon les conditions suivantes:

- paternité (BY): vous devez citer le nom de l'auteur original de la manière indiquée par l'auteur de l'œuvre ou le titulaire des droits qui vous confère cette autorisation (mais pas d'une manière qui suggérerait qu'ils vous soutiennent ou approuvent votre utilisation de l'œuvre);
- pas d'utilisation commerciale (NC): vous n'avez pas le droit d'utiliser cette création à des fins commerciales;
- partage des conditions initiales à l'identique (SA): si vous modifiez, transformez ou adaptez cette création, vous n'avez le droit de distribuer la création qui en résulte que sous un contrat identique à celui-ci. À chaque réutilisation ou distribution de cette création, vous devez faire apparaitre clairement au public les conditions contractuelles de sa mise à disposition. Chacune de ces conditions peut être levée si vous obtenez l'autorisation du titulaire des droits sur cette œuvre. Rien dans ce contrat ne diminue ou ne restreint le droit moral de l'auteur.

© Jinfu Tian 2024

Abstract

Jinfu Tian (2024). "Characterization of *VQ* gene family and associated plant defense pathways in rice" (PhD Dissertation in English).

Gembloux, Belgium, Gembloux Agro-Bio Tech, University of Liege.

143 pages, 19 figures, 3 tables.

Plant pathogens cause significant crop damages and yield losses each year. Rice bacterial blight and rice blast have the most dominant and destructive effects on rice production and quality. In plants, genes encoding valine glutamine (VQ)-motif containing proteins play an important role in plant defense responses. The objectives of this thesis are (1) to summarize the sequence and evolutionary characteristics, as well as roles and working mechanisms of *VQ* genes in plants, (2) to explore new *VQ* genes for rice resistance to diseases or pests, (3) generate novel germplasm resistant to rice bacterial blight and rice blast through CRISPR/Cas9, and (4) analyze the potential underlying mechanisms.

We reviewed the research progress of plant *VQ* family genes. The *VQ* family genes were successively identified in various plant species and found to be up- or down-regulated when exposed to environmental stress, pathogen invasion, and phytohormone treatment. They participate in various biological processes and responses to biotic or abiotic stress. VQ proteins function as important transcription regulators via interacting with WRKY transcription factors (WRKY), Mitogen-activated protein kinases (MAPK), and other proteins. Here, we summarized the detailed molecular characteristics, biological functions, and working mechanisms of published VQ proteins in plants.

In order to provide more detailed information on *VQ* family genes in plants, a total of 2469 *VQ* genes were identified from 56 plant species, including bryophytes, gymnosperms, and angiosperms. A comprehensive bioinformatics analysis, including conserved motifs, basic molecular characterization, and systemic clustering was carried out. Importantly, the research on the *VQ* family genes in gymnosperms has not been reported yet. So, we selected *Ginkgo biloba*, *Taxus chinensis* and *Pinus tabuliformis* as target species, identified their *VQ* gene members, and analyzed their molecular features, gene structure, subcellular location, chromosome distribution, duplication events, expression levels, synteny blocks, and evolutionary comparisons.

The CRISPR/Cas9 technology has been widely applied in plant genome editing and crop improvement. Here, we report that a rice valine-glutamine (VQ) motif-containing protein, OsVQ25, balances broad-spectrum disease resistance and plant growth by interacting with a U-Box E3 ligase, OsPUB73, and a transcription factor, OsWRKY53. OsPUB73 was found to positively regulate rice resistance against *Magnaporthe oryzae* and *Xanthomonas oryzae* by interacting with and promoting OsVQ25 degradation via the 26S proteasome pathway. Knockout mutants of *OsVQ25* exhibited enhanced resistance to both pathogens without growth and yield penalty. Furthermore, OsVQ25 interacted with and suppressed the

transcriptional activity of OsWRKY53, a positive regulator of plant immunity. OsWRKY53 downstream defense-related genes and brassinosteroid signaling genes were upregulated in *osvq25* mutants. Our findings revealed a OsPUB73-OsVQ25-OsWRKY53 module that finetunes plant immunity and growth at the transcriptional and posttranslational levels.

In conclusion, we have reviewed and summarized the characteristics, functions, and working mechanisms of *VQ* family genes in plants, and engineered novel resistant rice germplasms through CRISPR/Cas9-mediated genome editing technology and further revealed their disease resistance mechanisms.

Keywords: rice bacterial blight, rice blast, CRISPR/Cas9, *VQ* genes, sequence and evolutionary characteristics, roles and working mechanisms, OsVQ25, resistant rice germplasms, resistance mechanisms

Résumé

Jinfu Tian (2024). "Caractérisation de la famille des gènes VQ et des voies de défense des plantes associées dans le riz" (thèse de doctorat en anglais).

Gembloux, Belgique, Gembloux Agro-Bio Tech, Université de Liège.

143 pages, 19 figures, 3 tableaux.

Résumé:

De nombreux phytopathogènes induisent d'importants dommages causant des pertes de rendement chaque année. Les effets du flétrissement bactérien et de la pyriculariose sur la production et la qualité du riz sont les plus marqués et les plus dévastateurs. Chez les plantes, les gènes codant pour les protéines du motif Valine - glutamine (VQ) jouent un rôle important dans la réponse végétale. Les objectifs de cette thèse sont les suivants: (1) Étudier les séquences et les caractéristiques évolutives, le rôle et les mécanismes d'action du gène *VQ* chez les plantes; (2) Explorer de nouveaux gènes *VQ* pour la résistance du riz aux maladies et aux ravageurs; (3) Produire un nouveau germoplasme résistant au flétrissement des feuilles du riz et à la pyriculariose du riz par CRISPR/Cas9; (4) Analyser les mécanismes potentiels.

Nous avons passé en revue les progrès de la recherche sur les gènes de la famille VQ des plantes. Les protéines *VQ* sont une classe spécifique de cofacteurs transcriptionnels largement présents dans les plantes. De nombreux gènes *VQ* sont régulés à la hausse ou à la baisse lorsqu'ils sont exposés à des stress environnementaux, et à des attaques d'agents pathogènes. Ils sont impliqués dans divers processus physiologiques en réaction au stress biologique ou abiotique. Les protéines VQ interagissent principalement avec le facteur de transcription WRKY (WRKY), la protéine kinase activée par mitogène (MAPK) et d'autres protéines, jouant un rôle important de régulateur de la transcription. Malgré quelques études sur les gènes de la famille *VQ*, les perspectives d'application n'ont pas été développées. Les caractéristiques moléculaires, les fonctions biologiques et les mécanismes de travail des protéines VQ chez les plantes ont été détaillées ainsi que leur importance dans la régulation de l'activité transcriptionnelle. Aussi, des perspectives d'application sont évoquées.

Bien que les gènes *VQ* aient été identifiés structurellement et fonctionnellement chez certaines plantes, il manque une analyse systématique pour montrer clairement leurs caractéristiques dans une plus grande variété de plantes. Un total de 2469 gènes *VQ* provenant de 56 espèces de plantes telles que les Bryophytes, les nudistes et les Angiospermes ont été identifiés. Une analyse bioinformatique complète comprenant des motifs conservés a été réalisée ainsi que la caractérisation moléculaire et le regroupement systématique. Les *Ginkgo biloba*, *Taxus chinensis* et *Pinus tabuliformis* ont été sélectionnés comme modèles d'étude. Des informations fondamentales sur la caractérisation et l'évolution des gènes *VQ* chez les plantes ont été obtenues, avec des implications importantes pour la

poursuite de la recherche sur leurs fonctions biologiques et leurs mécanismes d'action.

La technologie CRISPR/Cas9 a été largement utilisée dans l'édition du génome des plantes et l'amélioration des cultures. Ici, une protéine OsVQ25 contenant un motif VQ induit la résistance à large spectre aux maladies et la croissance des plantes en interagissant avec la ligase U - box E3 OsPUB73 et le facteur de transcription OsWRKY53. OsPUB73 interagit avec OsVQ25 et favorise sa dégradation par la voie du protéasome 26s, régulant ainsi positivement la résistance du riz à *Aspergillus oryzae* et *Xanthomonas oryzae*. Les mutants Knockout d'*OsVQ25* ont montré une résistance accrue aux deux pathogènes sans inhibition de la croissance. De plus, OsVQ25 interagit avec OsWRKY53, un régulateur positif de l'immunité des plantes, et inhibe son activité transcriptionnelle. Les gènes liés à la défense en aval d'OsWRKY53 et les gènes de signalisation des stéroïdes ont été régulés à la hausse dans le Mutant *osvq25*. Nos résultats révèlent un module de facteur de transcription de la protéine VQ de l'Ubiquitine E3 ligase, qui ajuste finement l'immunité et la croissance des plantes au niveau de la transcription et de la post - traduction.

En conclusion, nous avons passé en revue les caractéristiques, les fonctions et les mécanismes d'action des gènes de la famille *VQ* chez les plantes et avons construit un nouveau germoplasme de riz résistant grâce à la technologie d'édition génomique médiée par CRISPR/Cas9, révélant davantage ses mécanismes de résistance aux maladies.

Mots clés: flétrissement bactérien et de la pyriculariose, CRISPR/Cas9, gène *VQ*, séquences et les caractéristiques évolutives, rôle et les mécanismes, OsVQ25, germoplasme de riz résistant, mécanisme de résistance

Acknowledgment

I would like to express my sincere appreciation and thanks to all of the members who helped and encouraged me during my PhD study. I would like to thank the Functional & evolutionary Entomology, Gembloux Agro-Bio Tech, University of Liège, and Institute of Crop Sciences, Chinese Academy of Agricultural Sciences. Thanks for the China Scholarship Council project provided me the financial support in Belgium.

I wish to express my deepest gratitude to my supervisors, Prof. Frédéric Francis and Prof. Lanqin Xia. Thanks for the opportunity they offered me to study in Beijing and Gembloux. Their professional guidance and patience helped me to gain more scientific knowledge and develop logical thinking during my study. I would like to thank my thesis committee and defense committee: the chairman, Prof. Blecker Christophe, Prof. Delaplace Pierre, Prof. Massart Sébastien, Dr Caparros Megido Rudy, and Prof. Julian Chen, for their insightful comments and encouragement. Thanks to Mr Zhang Mingjun for his kind help and efficient collaboration.

A large part of my work would have been impossible to complete without efficient collaboration with Taifei Yu and Lei Yan during my study at the Institute of Crop Science of CAAS, and Jiahui Liu, Jiahui Zhang, Xin Jiang, Lallie Glacet, Lisa Iannello and Noël Grégoire during my study at the lab of Functional & evolutionary Entomology. They provide me with all kinds of suggestions and conveniences whenever I need help. Thanks to their selfless help, and I was able to spend a wonderful and unforgettable time in Beijing and Gembloux.

I would like to convey my heartfelt gratitude to my family, my parents, my wife and children for their persistent support, understanding and encouragement, which helps and supports me complete my doctoral study.

In the end, I would like to express my gratitude and best wishes to all the teachers, friends, colleagues, and relatives who have helped and cared for me. Sincerely wish everyone forever happiness and well-being.

This study was funded by the China Scholarship Council (No. 202103250061) and GSCAAS-ULg Joint PhD Program. Thanks for their strong support.

Jinfu Tian
12/21/2023 in Gembloux, Belgium

Table of Contents

Abstract I
Résumé III
Acknowledgment V
Table of Contents VII
List of Figures IX
List of Tables X
List of Abbreviations XI
Chapter I General introduction 1
I.1. Context and objectives 3
I.2. Research roadmap and outline 5
Chapter II Literature review on CRISPR/Cas9-mediated genome editing in plants 7
II.1. Introduction 10
II.1.1. The action mechanism of CRISPR/Cas9 system 10
II.1.2. Development and application of CRISPR/Cas9 system 11
II.2. Conclusions and perspectives 14
Chapter III Literature review on the role and pathway of VQ family in plants 17
III.1. Introduction 19
III.1.1. Structural features of VQ genes and proteins 20
III.1.2. The role of the VQ protein in plant growth and development 23
III.1.3. The role of VQ protein in response to biotic stress 25
III.1.4. The role of VQ protein in response to abiotic stress 27
III.1.5. Interactions between VQ and WRKY 28
III.1.6. Interactions between VQ and MAPK 31
III.1.7. Interactions between VQ and other proteins 33
III.1.8. The importance of VQ domain and the molecular breeding strategies based on CRISPR/Cas9 system and VQ genes 37
III.2. Conclusions and perspectives 40
Chapter IV Large-scale identification and characterization analysis of *VQ* family genes in plants, especially gymnosperms 43
IV.1. Introduction 45
IV.2. Materials and methods 48
IV.2.1. Identification of VQ genes in multiple plants 48
IV.2.2. Molecular features and chromosomal localization analysis 48
IV.2.3. Conserved motifs, gene structure, and phylogenetic tree analysis 49
IV.2.4. Gene duplication and collinearity analysis 49
IV.2.5. Gene expression patterns analysis 50
IV.3. Results 50
IV.3.1. Members identification and conserved motif analysis of VQ genes in multiple plants 50
IV.3.2. Sequence similarities, length, introns, and GC content analysis of VQ genes in multiple plants 53
IV.3.3. Molecular features and gene structure analysis of VQ genes in

gymnosperms 54
IV.3.4. Gene duplication and collinearity analysis of VQ genes in gymnosperms 58
IV.3.5. Expression patterns of VQ genes in different tissues of gymnosperms 61
IV.3.6. Phylogenetic analysis of VQ genes 63
IV.4. Discussion 66
Chapter V A VQ-motif-containing protein fine-tunes rice immunity and growth by a hierarchical regulatory mechanism 71
V.1. Introduction 73
V.2. Materials and methods 75
V.2.1. Plants and materials 75
V.2.2. Expression pattern analysis 76
V.2.3. Magnaporthe oryzae (M. oryzae) inoculation and disease symptom evaluation 76
V.2.4. Xanthomonas oryzae pv. oryzae (Xoo) inoculation and disease symptom evaluation 76
V.2.5. Luciferase complementation imaging (LCI) assay in N. benthamiana . 76
V.2.6. In vivo co-immunoprecipitation assay 77
V.2.7. Bimolecular fluorescence complementation assay in N. benthamiana . 77
V.2.8. Protein degradation assay in rice protoplasts 78
V.2.9. Transcriptional activity assay in Arabidopsis protoplasts 78
V.2.10. Electrophoretic mobility shift assay 78
V.2.11. Dual-luciferase assay in rice protoplasts 78
V.2.12. RNA extraction and quantitative reverse transcriptase PCR 79
V.2.13. Quantitative and statistical analysis 79
V.3. Results 79
V.3.1. Regulation of rice resistance to Magnaporthe oryzae and Xanthomonas oryzae by OsPUB73 79
V.3.2. VQ-motif protein OsVQ25 interactions with OsPUB73 82
V.3.3. OsVQ25 degradation by OsPUB73 via the ubiquitin 26S proteasome . 84
V.3.4. Enhanced resistance of rice osvq25 mutant against Magnaporthe oryzae and Xanthomonas oryzae 84
V.3.5. Agronomic traits of resistant osvq25 mutant 87
V.3.6. OsVQ25 interactions with OsWRKY53 transcription factor 88
V.4. Discussion 93
Chapter VI General discussion, conclusions and perspectives 99
VI.1. CRISPR/CaS9-mediated genome editing 101
VI.2. The applications and limitations of *VQ* family genes in plants 102
VI.3. Limitations and the prospects of the research on *OsVQ25* in rice 103
References 107
Appendix-Publications 127

List of Figures

Figure 1- 1 The overview contents and research roadmap of thesis 5
Figure 2- 1 The working mechanism of basic CRISPR/Cas9 system 11
Figure 3- 1 The top 5 conserved motifs of VQ proteins. 21
Figure 3- 2 The functional model and interaction network of *VQ* family genes 37
Figure 3- 3 Molecular design breeding strategy based on *VQ* gene 40
Figure 4- 1 Members identification and sequence characteristics of *VQ* genes. 51
Figure 4- 2 The ratio of single exon (intron-free) *VQ* genes 53
Figure 4- 3 The average GC content of genes in coding sequence 54
Figure 4- 4 Chromosomal distribution of *VQ* genes in three gymnosperms 56
Figure 4- 5 Phylogenetic tree, conserved motifs, and gene structure of VQ protein in three gymnosperms 57
Figure 4- 6 Collinear analysis between different gymnosperms 61
Figure 4- 7 Expression level of *VQ* genes in different tissues of three gymnosperms62
Figure 4- 8 Unrooted phylogenetic tree of VQ proteins in 56 plant species 63
Figure 5- 1 Expression pattern of *OsPUB73* and resistance phenotypes of the *ospub73* knockout mutant to *Magnaporthe oryzae* and *Xanthomonas oryzae*. 80
Figure 5- 2 OsPUB73 interactions with and promoting the degradation of OsVQ2582
Figure 5- 3 Expression pattern of *OsVQ25* and enhanced resistance of the *osvq25* mutant to *Magnaporthe oryzae* and *Xanthomonas oryzae* 85
Figure 5- 4 Major agronomic traits measured in *osvq25* mutant lines 87
Figure 5- 5 Interaction between OsVQ25 and OsWRKY53 and suppression of the transcriptional activity of OsWRKY53 by OsVQ25 90
Figure 5- 6 Working model of the role of OsVQ25 in balancing plant defense and growth 96

List of Tables

Table 3- 1 Identification of *VQ* family genes in different plants. 22
Table 3- 2 Biological functions and interacting factors of VQ proteins in different plants. 35
Table 4- 1 Summary of molecular characterization and subcellular localization of VQ proteins in three gymnosperms 55

List of Abbreviations

Chapter I

CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DSB	Double strand DNA breaks
HDR	The homology-directed DNA repair
NHEJ	Non-homologous end joining
VQ	VQ motif containing protein

Chapter II

crRNA	CRISPR-derived RNA
tracrRNA	Trans-activating crRNA
PAM	Protospacer Adjacent Motif
Indel	Insertion or deletion
sgRNA	Single guiding RNA
SBE	Starch Branching Enzyme
T-DNA	Transfer DNA

Chapter III

MAPK	Mitogen-activated protein kinase
CaM	Calmodulin
miRNA	Micro RNA
RNAi	RNA interference
Ka	Nonsynonymous substitutions ratio
Ks	Synonymous substitutions ratio
ETI	Effector-triggered immunity
PTI	Pattern-triggered immunity
SA	Salicylic acid
JA	Jasmonic acid
ABA	Abscisic acid
Xoo	*Xanthomonas oryzae* pv. *oryzae*
Xoc	*Xanthomonas oryzae* pv. *oryzicola*

Chapter IV

HMM	Hidden Markov Model
MW	molecular weight
pI	isoelectric point
WGD	whole-genome duplication
TPM	Transcripts Per Million

Chapter V

SNS	Species-non-specific

BSR	Broad-spectrum resistance
UPS	Ubiquitin-proteasome system
PCD	Programmed cell death
qRT-PCR	Quantitative real-time polymerase chain reaction
BR	Brassinosteroid

1

Chapter I General introduction

I.1. Context and objectives

I.1.1. Context

Agricultural pests and pathogens can destroy many important crops and cause significant damage and yield losses each year both by direct feeding or infection. Rice bacterial blight and rice blast have the most dominant and destructive effects on rice production and quality.

At present, the prevention and control of pests and pathogens usually depend on pesticide or biocide, however the application of such chemicals is not desirable in the long term, because of the development of insecticide resistance, the potential negative effects on non-target organisms, and especially environment pollution. Many attempts have been used to improve resistance under the biotic stress in rice through conventional breeding. They are being confronted with slow progress due to lack of resistant germplasm, unsteady inheritance and demanding long period of time. Therefore, exploring new type of resistant genes and applying biotechnology mediated by the CRISPR/Cas9 system to obtain transgene-free resistant germplasms will be an efficient alternative strategy.

CRISPR/Cas9 system is a fast developing and wide used tool for plant genome editing because of its specificity, simplicity, versatility, robust and cost-effective (Shan et al., 2013; Wang et al., 2014; Sun et al., 2016; Zhang et al., 2017; Hu et al., 2018; Hua et al., 2019; Walton et al., 2020). The Cas9 endonuclease generates double strand DNA breaks (DSBs) in target genes in a guide RNA dependent manner (Cong et al., 2013). DSBs are repaired either through the error-prone non-homologous end joining (NHEJ) pathway or the homology-directed DNA repair (HDR) pathway. NHEJ often causes small indels and has been widely used in generating knock-out mutants (Ma et al., 2015; Zhang et al., 2018). HDR or homologous recombination-based gene targeting (GT) plays essential roles in targeted gene replacement or precise modifications (Li et al., 2019). Therefore, CRISPR/Cas9 system can serve as a powerful tool for creating gene mutants and studying gene function.

VQ proteins are a class of plant-specific transcription regulatory cofactors that contain a conservative VQ-motif: FxxhVQxhTG, where x is any amino acid and h represents hydrophobic amino acid. Research on the function of the *VQ* gene found that it is not only involved in regulating multiple life processes of plants (Li et al., 2014a; Li et al., 2014b), but also involved in plant response to biotic and abiotic stresses (Hu et al., 2013a; Wang et al., 2015a). Studies have shown that the expression of *VQ* gene is induced or inhibited by SA, JA or pathogens (Kim et al., 2013; Wang et al., 2015b; Zhang et al., 2015), indicating that *VQ* gene may have a important role in responding to plant defense. Given the important biological role of *VQ* genes, *VQ* gene resources are worth further research and development in plants.

Jasmonates, including jasmonic acid and its cyclopentanone derivatives, are well-known to act as a regulator in plant growth and development as well as a vital

defense signal to regulate plant response against herbivorous insects and necrotrophic pathogens. AtVQ22/JAV1 is a VQ domain gene in the *Arabidopsis* jasmonate pathway, reported as a negative regulator controlling plant defense. When encountering insect attack and pathogen infection, plants accumulate jasmonates that trigger JAV1 degradation via the 26S proteasome to activate defensive gene expression and elevate resistances against both insects and pathogens. Knocking out JAV1 gene significantly enhances jasmonate-regulated defense responses against the necrotrophic pathogen and herbivorous insect without growth penalty (Hu et al., 2013b).

Although some *VQ* genes have been studied in *Arabidopsis thaliana* already, there are few reports and studies in crops, especially in crops. The purpose of this study is to research and summarize the characteristics of the *VQ* gene family in plants, and explore disease resistance related *VQ* genes in rice that are functionally similar to JAV1. After determining the candidate *VQ* genes, we will select suitable CRISPR/Cas9 tools for targeted editing and modification of *VQ* genes based on actual needs. After gene editing using the CRISPR/Cas9 system, new germplasm with disease resistance but no loss of plant agronomic traits can be obtained. In addition, analyzing their disease resistance mechanisms is also of great significance for the research and utilization of *VQ* gene resources in crops.

I.1.2. Objectives

Cereal plants are often attacked by different pathogens, resulting in reduced yield and decreased quality. In order to ensure food security, it is necessary to search for effective resistance genes and create new resistant germplasms. According to reports, *VQ* family genes play important roles in regulating plant defense and various life activities. Therefore, the *VQ* genes deserve in-depth attention and research. However, although there have been some reports on the *VQ* family genes, they are scattered, not systematic, and still lack clearer feature displays and deeper summaries to further promote the application of *VQ* gene resources. In addition, the role and mechanism of *VQ* genes in plant disease resistance in major food crops such as rice also urgently need to be developed and studied. The CRISPR/Cas9 system is currently playing an increasingly important role in plant mutant creation, gene function research, and crop improvement. Therefore, fully understanding the principle, application, advantages and disadvantages of the CRISPR/Cas9 system, and using the appropriate tools of the CRISPR/Cas9 system to edit the rice genome, will greatly promote the research and utilization of rice *VQ* genes.

Given the importance of CRISPR/Cas9 system and *VQ* family genes, our specific research objectives are as follows:

(1) Reviewing the principles, applications, problems, and prospects of CRISPR/Cas9 genome editing tools;

(2) Summarizing and studying the characteristics, functions, mechanisms, current problems, and application prospects of the *VQ* gene family in a wide range of plants;

(3) Developing *VQ* genes significantly associated with disease resistance in rice, and obtaining transgene-free disease resistant rice germplasm through CRISPR/Cas9-mediated genome editing technology and genetic isolation;

(4) Investigating the related plant defense pathway and mechanism of these *VQ* genes in rice.

I.2. Research roadmap and outline

I.2.1. Research roadmap

The research roadmap is presented in Figure 1-1.

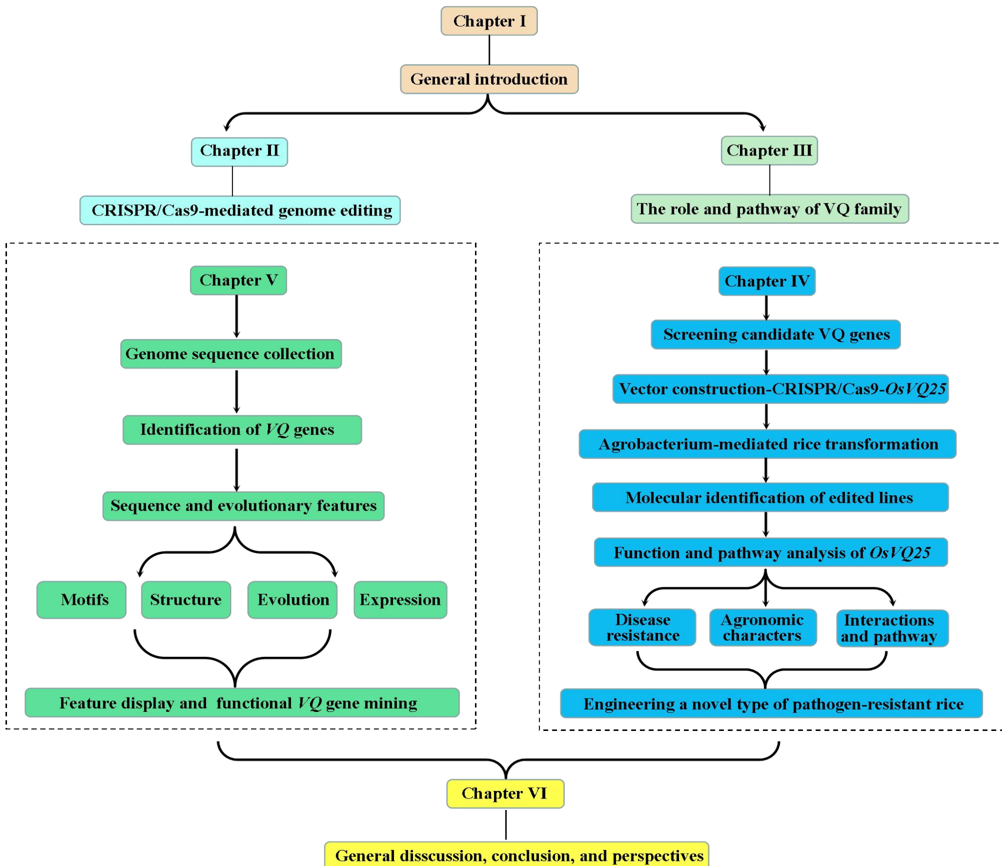


Figure 1- 1. The overview contents and research roadmap of thesis.

Firstly, review the working mechanism, current research status, and applications of CRISPR/Cas9 system. We plan to select suitable CRISPR/Cas9 tool for subsequent mutant creation and functional research of candidate genes. Secondly, review the structure, function, and working mechanism of the *VQ* gene family.

Through literature investigation, it was found that the VQ family have important functions in plants, and the *VQ* genes was identified sas the main research target. Thirdly, we will identify the *VQ* gene family in various plants, study their molecular and evolutionary characteristics, and these results will lay the foundation for further candidate *VQ* gene selection. Fourthly, based on literature reports and sequence alignment, candidate disease or pest resistance related *VQ* genes with negative regulatory effects were screened in rice. Furthermore, we will construct CRISPR/Cas9 vector to knock out target *VQ* gene and use *Agrobacterium*-mediated transgenic technology for genetic transformation. In addition, we will perform phenotype identification on the mutant of the *VQ* gene. For example, testing for disease or insect resistance and investigating their agronomic traits. Importantly, we will also further analyze the disease resistance pathways and mechanisms involved in this *VQ* gene. Finally, we will obtain new rice materials and germplasms with breeding potential for disease or insect resistance.

I.2.2. Outline

In Chapter II, we reviewed and summarized the basic working principle, research progress, applications in plants, existing problems, and future research prospects of CRISPR/Cas9 system.

In Chapter III, we reviewed the research progress of the *VQ* gene family in the plant kingdom, including structural characteristics, functions, mechanisms, current issues, and application prospects.

Reference: Tian, J., Zhang, J. Francis, F. (2023). The role and pathway of VQ family in plant growth, immunity, and stress response. Planta 259, 16.

In Chapter IV, we conducted large-scale identification and molecular and evolutionary characteristics analysis of *VQ* family genes in 56 plant species, especially for three gymnosperms.

Reference: Tian, J., Zhang, J., Francis, F. (2023). Large-scale identification and characterization analysis of *VQ* family genes in plants, especially gymnosperms. Int. J. Mol. Sci. 2023, 24, 14968.

In Chapter V, we developed pathogen resistant rice through knockout of a key VQ-motif containing gene *OsVQ25*, and studied the involved disease resistance pathways.

Reference: Hao, Z.[†], Tian, J.[†], Fang, H.[†], Fang, L., Xu, X., He, F., et al. (2022). A VQ-motif-containing protein fine-tunes rice immunity and growth by a hierarchical regulatory mechanism. Cell Rep. 40(7), 111235.

2

Chapter II Literature review on CRISPR/Cas9-mediated genome editing in plants

Abstract

With the advent of the genome era, a massive amount of gene data has emerged, and the functions of a large number of genes need further exploration. At present, there are many ways to study gene function. The traditional methods such as RNAi (RNA interference) cannot meet the needs of a large number of gene function analyses, and there are also cases of incomplete interference. In addition, physical or chemical mutagenesis techniques, as well as the random mutations of exogenous gene insertion (T-DNA), have extremely low efficiency. With the continuous evolution of bioengineering technologies, CRISPR/Cas9-mediated genome editing technology has emerged, providing effective means to create gene loss-of-function mutants. There are four main types of genome editing in crops using the CRISRP/Cas9 system: gene targeted knockout, target site replacement, base editing, and fragment insertion. The CRISPR/Cas9 system can simultaneously edit different loci of the same gene or edit multiple genes. Using the CRISPR/Cas9 system for gene knockout is simple and fast, requiring only 20 bp nucleotide sequences to be replaced on the original vector for each target, making it more suitable for large-scale operations. The CRISPR/Cas9 system has quickly become an important tool for studying the gene function and creating new materials for plants, animals, and microorganisms, relying on its simple operation, high editing efficiency, and wide applicability.

Keywords: CRISPR/Cas9, gene targeted knockout, target site replacement, base editing, fragment insertion, studying the gene function, creating new materials

II.1. Introduction

In 1987, a Japanese research team discovered a special tandem spacer repeat structure in Escherichia coli (Ishino et al., 1987). In 2002, scientists named this special repetitive sequence 'CRISPR' (Clustered Regularly Interspaced Short Palindromic Repeats), which refers to regular clusters of short palindromic repeats (Jansen et al., 2002). In 2005, three research teams reported that the CRISPR interval repeat sequences present in bacteria and archaea were highly homologous to the sequences of certain viruses or plasmids outside the host, providing evidence that the CRISPR system may play a role in the immune response of microorganisms to invasion of exogenous genetic material (Bolotin et al., 2005; Mojica et al., 2005; Poourcel et al., 2005). In 2006, a research team in the United States predicted through computer bioinformatics analysis that the CRISPR/Cas system may cleave exogenous DNA by binding them to form specific structures, similar to the RNAi of eukaryotic cells (Makarova et al., 2006). In 2007, the Barrangou team first experimentally confirmed that the Type II CRISPR/Cas system is directly involved in the adaptive immune response of bacteria to bacteriophages (Barrangou et al., 2007). In 2013, the CRISPR/Cas9 system, which was modified based on Type II CRISPR/Cas, was first applied to mammals and successfully achieved simultaneous editing of multiple genes (Cong et al., 2013). Subsequently, the CRISPR/Cas9 system rapidly became a powerful tool for genome editing research in organisms such as humans, animals, and plants.

II.1.1. The action mechanism of CRISPR/Cas9 system

When bacteria are invaded by exogenous DNA, the corresponding region of CRISPR is transcribed as pre-crRNA (crRNA: CRISPR-derived RNA), and tracrRNA complementary to the repetitive regions in CRISPR is also transcribed. tracrRNA forms a complex structure with Cas9 protein and works with RNase III to process pre-crRNA into mature crRNA. Subsequently, crRNA, tracrRNA, and Cas9 form a secondary complex structure. The secondary complex cleaves exogenous DNA by combining crRNA sequences with complementary DNA sequences, ultimately leading to the production of double strand breaks (DSB). The repair of DSB usually leads to random insertion or deletion of bases, resulting in the loss of gene function. The cleavage site of CRISPR/Cas9 is usually located at the third nucleotide upstream of NGG PAM (Protospacer Adjacent Motif) in the 5 '-N20-NGG-3' sequence.

The CRISPR/Cas9 system can cleave specific target sites and trigger the generation of DSB, while cells are often able to repair the DSB through their own repair mechanism (Figure 2-1). There are two main repair mechanisms for DSB in cells: non homologous end joining (NHEJ) and homologous directed repair (HDR) (Wyman et al., 2006). NHEJ is the main repair mechanism for DNA double strand breakage damage in cells. Due to the high probability of errors in this repair mechanism, random insertion or deletion (Indel) of bases can occur, resulting in frameshift or functional region deletion mutations, disrupting the normal structure

of genes, causing them to lose their original functions, and achieving gene knockout. In contrast, the repair mechanism of HDR homologous recombination has a lower error rate, but the probability of occurrence is very low. The HDR repair mechanism mediated by CRISPR/Cas9 can be used to achieve targeted gene editing, such as targeted gene insertion and precise gene replacement (Yin et al., 2014). It is worth noting that the frequency of HDR is significantly lower than that of NHEJ.

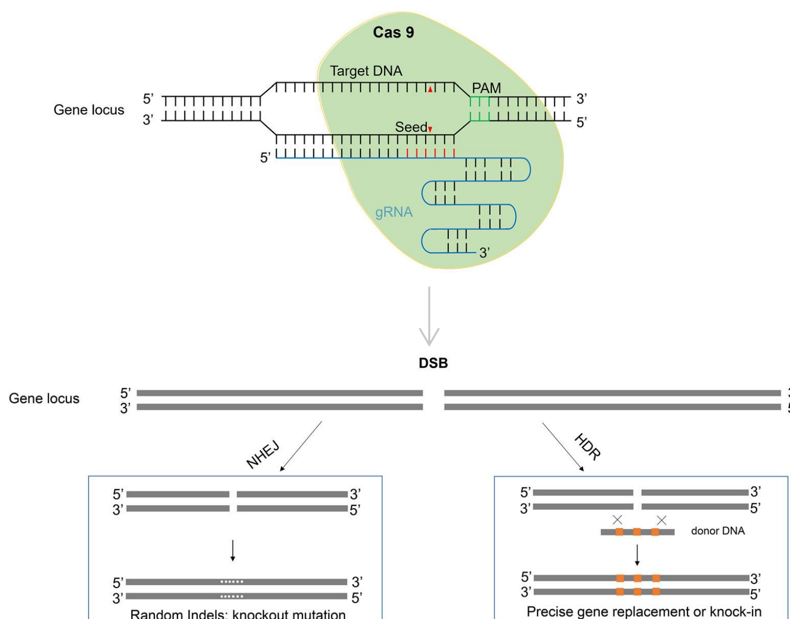


Figure 2-1. The working mechanism of basic CRISPR/Cas9 system. Two major pathways underlying the repair of DSB induced by CRISPR/Cas9 system (This figure is sourced from Li et al., 2019).

II.1.2. Development and application of CRISPR/Cas9 system

Since scientists discovered the existence of CRISPR in 1987, in less than 30 years, CRISPR technology has undergone unprecedented development. Especially in recent years, as an emerging genome editing tool, CRISPR/Cas9 system has the advantages of simple construction, low cost, fast speed, high efficiency, and being able to be used for multi-gene and multi-site editing simultaneously. With the continuous development of technology and in-depth research, the CRISPR/Cas9 system has been extended and applied to multiple species such as humans, rats, mice, zebrafish, bacteria, *Arabidopsis*, rice, wheat, etc., specifically involving single base substitution, targeted integration of foreign fragment, gene knockout, precise gene replacement, and regulation of gene transcription levels.

In 2012, scientists artificially connected crRNA and tracrRNA into a guiding RNA (sgRNA) in vitro and confirmed that sgRNA also had a specific guiding effect on the endonuclease activity of Cas9. This study simplified the construction process of the CRISPR/Cas9 system and laid an important foundation for the subsequent modification and application of the Type II CRISPR/Cas system (Jinek et al., 2012). In 2013, Zhang Feng's research team first reported that the CRISPR/Cas9 system can mediate gene editing in eukaryotes (Ran et al., 2013). In the same year, CRISPR/Cas9 system was used to achieve site-specific mutations in human and mouse cells in vitro (Cong et al., 2013); CRISPR/Cas9 system was successfully mediated gene knockout in zebrafish embryos (Hwang et al., 2013); Gao Caixia's research team conducted the first rice gene knockout experiment using the CRISPR/Cas9 system in plants and obtained stable rice gene editing plants, thus establishing the CRISPR/Cas9 genome editing system in plants (Shan et al., 2013). *TaMLO* was successfully targeted and knocked out in wheat using CRISPR technology, obtaining wheat new materials resistant to powdery mildew (Wang et al., 2014). CRISPR/Cas9 system was used to edit a total of 46 target sites in rice, with an average mutation rate of 85.4%, most of which were biallelic mutations (54.9%) and homozygous mutations (24.7%), and could be inherited to offspring, indicating that the CRISPR/Cas9 system can achieve efficient point mutations for specific genes in crop (Ma et al., 2015). In 2016, Zhu Jiankang's research team used the CRISPR/Cas9 system to simultaneously edit multiple genes in *Arabidopsis* (Zhang et al., 2016). Similarly, through CRISPR/Cas9-mediated multiplex gene editing, six important agronomic genes were simultaneously edited in wheat (Luo et al., 2021). Li et al. used the CRISPR-Cas9 system to selectively knock out the *ZmTMS5* gene in maize, and the tms5 mutant was stably transmitted to the next generation, producing temperature sensitive male sterile maize (Li et al., 2017b). In rice, after mutating the starch branching enzyme related genes *SBE* I and *SBE* II, it was found that there was no significant difference between the *SBE* I mutant and the wild-type, while the *SBE* II mutant had a higher proportion of long chains in the branching chain, with a significant increase of 25.0% and 9.8% in high amylose content (AC) and resistant starch content (RS), respectively, thereby altering the fine structure and nutritional characteristics of starch (Sun et al., 2017). Abe et al. utilized CRISPR/Cas9 gene editing technology to edit the *Qsd1* gene related to seed dormancy duration in wheat, and successfully developed wheat materials that are resistant to rainwater and less prone to wheat ear germination (Abe el al., 2019). Sun et al. used homologous recombination mediated by the CRISPR/Cas9 system to perform site-specific substitution of OsALS, endowing it with herbicide resistance. Finally, through genetic isolation, herbicide resistant plants without screening markers were obtained (Sun et al., 2016).

CRISPR/Cas9 composite system BE3 (APOBEC1-XTEN-nCas9-UGI) was used for single base editing of rice genes (Li et al., 2017). In 2020, for the first time, Ronald's team utilized CRISPR/Cas9 to insert a 5.2kb carotenoid synthesis expression component at a specific location in rice genome, obtaining golden rice rich in carotenoids (Dong et al., 2020). Maize genetically modified with Cas9/guide RNA (gRNA) was utilized for inter generic pollination of wheat to

promote site-specific mutations in any wheat germplasm. The authors generated new allelic variants of wheat genes *BRASSINOSTEROID INSENSITIVE 1* (*BRI1*) and *SEMIDWARF 1* (*SD1*) involved in plant height regulation (Budhagatapalli et al., 2020). Xie Chuanxiao's team utilized CRISPR/Cas9 genome editing technology to create a one-step technical system for both recessive genic male sterile lines (GMS) and controlled genic male sterile maintainer lines (MGM maintainer lines) in crops, improving the development efficiency of crop hybrid breeding (Qi et al., 2020). In addition to being used for genome editing, CRISPR/Cas9 system can also be used for gene expression regulation and epigenetic modification. Qi et al. modified the Cas9 system into a transcription regulatory factor dCas9 that does not have cleavage activity but has binding function, and designed sgRNA in the promoter region. dCas9 can bind to the corresponding site on the promoter to inhibit the transcription of the target gene, thereby achieving interference with the expression of the target gene. The researchers referred to it as CRISPRi (CRISPR interference) (Qi et al., 2013).

In recent years, the CRISPR/Cas9 genome editing system has also been widely applied and improved. There are many variants of Cas9 designed to continuously expand the recognition scope of PAM sites, to improve editing efficiency and accuracy (Hu et al., 2017; Hu et al., 2018; Hua et al., 2019; Kulcsár et al., 2020; Walton et al., 2020). For example, Cas9 orthologs with shorter PAM have been reported such as SpCas9-NG for a 5'-NG-3' PAM (Nishimasu et al., 2018), and SpRY for 5'-NR/YN-3' PAM (Walton et al., 2020). In addition, after fusing Cas9 with other proteins, the CRISPR/Cas9 toolbox has also added new technologies such as base editing and primer editing tools to achieve base replacement and targeted insertion or deletion (Gao, 2021; Molla et al., 2021; Sun et al., 2023). For example, CRISPR/Cas9-based cytosine base editors (CBEs) and adenine base editors (ABEs) can efficiently mediate C-to-T/G-to-A and A-to-G/T-to-C substitutions, respectively (Zeng et al., 2022). They are formed by the fusion of nCas9 (Cas9 nickase) with cytosine deaminase and adenine deaminase, respectively. PE (Primer Editing) is an effector protein composed of nCas9 fused with engineered reverse transcriptase (RT). Compared to CBE (Cytodine Base Editor) and ABE (Adenine Base Editor), the biggest advantage of PE is its diverse editing types, including multiple sites editing, DNA short segment insertion and deletion. Compared to traditional HDR (Homologous Directed Recognition), base editors and PE can achieve editing without DSB (Double Strand Break), making it more efficient and secure. However, both base editors and PE tools still have low editing efficiency in plants, and further optimization and improvement are needed in the future. These new CRISPR/Cas9-based technologies have become powerful tools in plant research.

With the continuous optimization and improvement of basic technology, the CRISPR/Cas9 system has undergone rapid changes in recent years and has become an important means of modern molecular biology research. Its efficiency, specificity, precision, and safety have brought revolutionary impacts to the development of modern agriculture, molecular design breeding. Moreover, simultaneous mutation of multiple genes and high-throughput batch detection have

also become possible. CRISPR/Cas9 has become more widely used in animal modeling, drug screening, single base editing technology, cell lineage tracing, basic disease research, disease diagnosis, in vivo editing, and genetic disease correction, becoming increasingly widely used tools in scientific research.

II.2. Conclusions and perspectives

The CRISPR/Cas9 system has the characteristics of simple principle, low cost, easy operation, and high mutation efficiency. It has been rapidly developed and applied since its birth in 2013, and is currently the main means of plant gene editing, achieving precise editing and regulation of target genes. Compared with traditional gene editing methods such as T-DNA insertion and RNAi, this system can obtain plant materials without transgenic fragments through genetic separation, bringing good prospects for the commercial production of gene editing crops. The new breeding technology based on genome editing technology allows for the simultaneous modification of multiple genetic loci in excellent varieties, which will accelerate the gene functional research and strengthen global crop improvement. The CRISPR/Cas9 system has been used to improve various traits of crops, including stress resistance, disease resistance, insect resistance, herbicide resistance, yield levels, and nutritional value. Multiple gene editing can quickly improve multiple traits in the context of excellent varieties or valuable wild varieties, which is of great significance for improving complex agronomic traits in crops.

With the development of high-throughput sequencing technology, more than 100 crops have been sequenced, such as rice, corn, wheat, millet, peanut, soybean, chickpea, potato, cassava, banana, apple, pear, peach, tomato, cabbage, cucumber, and so on (Purugganan and Jackson, 2021). The large amount of genome data obtained by researchers will provide convenience for gene function research and crop improvement through genome editing. Nevertheless, plant genome editing still faces enormous challenges, mainly in improving the efficiency of homologous recombination, base editing, and large fragment insertion/deletion, further breaking through the limitations of PAM recognition sites, and quickly obtaining mutant materials that can be stably inherited. This requires more optimization and innovation for CRISPR/Cas9 system. For example, using artificial intelligence (AI) to design and develop smaller, more efficient, or more adaptable Cas9 proteins. Alternatively, discovering new Cas9 proteins that have even stronger functions in living organisms, or scientifically combining Cas9 with other proteins forming a more powerful fusion protein to further expand the functionality of Cas9 and improve editing efficiency. In addition, efficient transformation and regeneration methods are another necessary condition limiting the application of gene editing in many crops. At present, many excellent varieties and other important commodity crops have low conversion efficiency, which reduces the efficiency of CRISPR/Cas9-mediated genome editing events. The safety of CRISPR/Cas9 gene editing materials also requires continuous observation. Therefore, we should further improve and optimize the technologies and methods matched with CRISPR/Cas9 to achieve efficient and precise utilization of CRISPR/Cas9 system.

With more practical applications and technological breakthroughs of CRISPR/Cas9, this system will become a more powerful tool for gene and genome research, as well as for the creation of new materials.

3

Chapter III Literature review on the role and pathway of VQ family in plants

Reference: Tian, J., Zhang, J. Francis, F. (2023). The role and pathway of VQ family in plant growth, immunity, and stress response. Planta 259, 16. https://doi.org/10.1007/s00425-023-04292-z

Abstract

Valine-glutamine (VQ) motif-containing proteins are a large class of transcriptional regulatory cofactors. VQ proteins have their own unique molecular characteristics. Amino acids are highly conserved only in the VQ domain, while other positions vary greatly. Most *VQ* genes do not contain introns and the length of their proteins is less than 300 amino acids. A majority of VQ proteins are predicted to be localized in the nucleus. The promoter of many *VQ* genes contains stress or growth related elements. Segment duplication and tandem duplication are the main amplification mechanisms of the *VQ* gene family in angiosperms and gymnosperms, respectively. Purification selection plays a crucial role in the evolution of many *VQ* genes. By interacting with WRKY, MAPK, and other proteins, VQ proteins participate in the multiple signaling pathways to regulate plant growth and development, as well as defense responses to biotic and abiotic stresses. Although there have been some reports on the *VQ* gene family in plants, most of them only identify family members, with little functional verification, and there is also a lack of complete, detailed, and up-to-date review of research progress. Here, we comprehensively summarized the research progress of *VQ* genes that have been published so far, mainly including their molecular characteristics, biological functions, importance of VQ motif, and working mechanisms. Finally, the regulatory network and model of *VQ* genes were drawn, a precise molecular breeding strategy based on *VQ* genes was proposed, and the current problems and future prospects were pointed out, providing a powerful reference for further research and utilization of *VQ* genes in plant improvement.

Keywords: VQ proteins, molecular characteristics, growth and development, biotic and abiotic stresses, importance of VQ-motif, working mechanisms, molecular precision breeding

III.1. Introduction

Gene expression is usually regulated by the interactions between cis-acting elements and transcription factors, and the latter often requiring some co-factors to function in regulation of gene expression. VQ motif-containing (VQ) protein is a specific class of transcriptional co-factor widely found in plants, containing the highly conserved VQ motif FxxhVQxhTG (x represents arbitrary amino acid, h represents hydrophobic amino acid) (Jiang et al., 2018). In 2002, *AtVQ23*/*AtSIB1* was first identified in *Arabidopsis* (Morikawa et al., 2002). Subsequently, the *VQ* genes were successively identified in various plant species (Yuan et al., 2021). Many *VQ* genes were up- or down-regulated when exposed to environmental stress, pathogen invasion, and phytohormone treatment, involved in various life processes

and responses to biotic and abiotic stresses (Kim et al., 2013; Zhang et al., 2015; Song et al., 2016; Jiang et al., 2018). By mainly interacting with WRKY transcription factors, MAPK kinases and other proteins, VQ proteins function as important transcription regulators (Cheng et al., 2012; Li et al., 2014a; Ali et al., 2019; Yuan et al., 2021). Although there have been some studies on the *VQ* family genes, they are scattered, and the overall regulatory context and application prospects are not clear enough. In this study, we reviewed the detailed molecular characteristics, biological functions, and working mechanisms of VQ proteins in plants and highlight its importance in regulation of transcriptional activity.

III.1.1. Structural features of VQ genes and proteins

The *VQ* family genes have been individually identified and systematically analyzed from different plants by bioinformatic and experimental methods, for example, *Arabidopsis*, soybean, apple, tomato, cucumber, potato, tobacco, cotton, rice, maize, wheat, etc. (Table 3-1). Some studies have also performed large-scale analysis and comparison of *VQ* gene families in multiple plant species, including angiosperms and gymnosperms, to elucidate their characteristics and patterns (Jiang et al., 2018; Cai et al., 2019; Ma et al., 2023b; Tian et al., 2023). VQ proteins have their own specific molecular features. The conserved VQ-motif is FxxhVQxhTG with three main terminal amino acids, namely LTG, FTG, and VTG, but occasionally there are other types in some plants, such as ITG, YTG, LTR, and LTS (Zhang et al., 2022a) (Figure 3-1). The core element 'VQ' in FxxhVQxhTG, is slightly changed in some plants, such as FxxhVHxhTG (Wang et al., 2017; Liu et al., 2022a; Zhang et al., 2022a). Except VQ-motif, the amino acid sequences at other positions are very variable (Weyhe et al., 2014). Top 10 conserved motif of VQ proteins were pointed out in angiosperms (Cai et al., 2019), which maybe related to protein localization and interaction (Figure 3-1). Studies have shown that many VQ proteins contain single or dual component nuclear localization signals, and some also contain chloroplast targeting signals. Most VQ proteins are expected to be located in the nucleus, with a few located in the cytoplasm, chloroplast, and mitochondria (Cheng et al., 2012; Kim et al., 2013; Jing and Lin, 2015; Guo et al., 2018). Moreover, the N-terminal of *Arabidopsis* AtVQ15 and AtVQ22 contain a calmodulin (CaM) binding domain, which is essential for their interaction with CaM (Perruc et al., 2004; Yan et al., 2018). Some VQ proteins also contain the predicted MAPK docking domains, which are necessary for their interaction with MAPK (Pecher et al., 2014). Most *VQ* genes in higher plants without introns and encode relatively small proteins less than 300 amino acids (Jiang et al., 2018; Cai et al., 2019).

According to structural characteristics, phylogenetic tree was systematically built and VQ proteins from different species have been clustered into 7 groups (Kim et al., 2013; Wang et al., 2017), 8 groups (Zhang et al., 2022a), 9 groups (Dong et al., 2018), and 10 groups (Pecher et al., 2014; Cai et al., 2019). Among

them, gymnosperms and angiosperms, as well as monocotyledons and dicotyledons, were found to have their own independent and intersecting branches, which provides evidence for the evolutionary history of *VQ* genes (Jiang et al., 2018; Cai et al., 2019; Tian et al., 2023). Segmental duplication and tandem duplication are considered to be the main mechanisms for the expansion of the *VQ* gene family in angiosperms and gymnosperms, respectively, and there is no necessary relationship between *VQ* numbers and genome size (Jiang et al., 2018; Xu and Wang, 2022; Zhang et al., 2022a; Tian et al., 2023). The estimated number of *VQ* genes based on the times of genome replication events is inconsistent with the actual number, indicating that there may be gene loss events after genome replication (Wang et al., 2019). The substitution rates of Ka and Ks are the basis of analyzing the selection pressure in gene duplication events (Wang et al., 2010a). The Ka/Ks values of most duplicated *VQ* gene pairs are < 1, indicating that they mainly evolved under purification selection (Wang et al., 2017; Cao et al., 2018; Jiang et al., 2018; Zhang et al., 2019; Zhang et al., 2022a). These provide an important reference for evolutionary comparison and biological function of *VQ* genes in different plants.

In addition, cis-acting elements such as W-box (WRKY binding site), and SA-, JA-, and ABA- related elements are enriched in the promoter region of *VQ* genes in different plants. These cis-acting elements are mainly classified into four categories: hormone, stress regulatory, growth and development, and photo-reactive related elements respectively (Wang et al., 2015a; Song et al., 2016; Wang et al., 2017; Zhang et al., 2019; Zhang et al., 2022a). These findings suggest that *VQ* genes may be associated with response to biotic and abiotic stresses as well as growth and development. Additionally, *VQ* genes were reported to be regulated by microRNA (Guo et al., 2018; Zhang et al., 2022a). For example, 38 of 113 wheat *VQ* genes were predicted to be targeted by 15 putative miRNA. The latter belong to different miRNA families, such as miR160, miR395, miR1130 and miR9657, which play key roles in various biotic and abiotic stresses (Zhang et al., 2022a). These studies have expanded our understanding of the structure, function, and regulation of *VQ* genes, indicating that there may be more complex regulatory networks in plants.

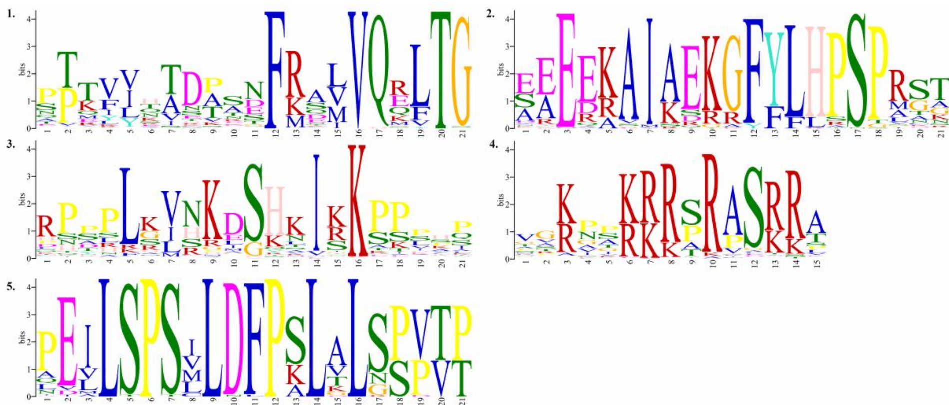


Figure 3-1. The top 5 conserved motifs of VQ proteins.
MEME online program (https://meme-suite.org) (Bailey et al., 2015) was used to identify motifs of VQ proteins from *Arabidopsis*, soybean, grape, tomato, Chinese cabbage, cotton, rice, maize, and wheat. Here, the top 5 motifs were listed. Motif analysis can help us to better understand protein structure and determine homologous sequences, and these motifs may play important roles in subcellular localization and interactions with different proteins

Table 3-1. Identification of *VQ* family genes in different plants

Species	No. of total *VQ* genes	Reference
Arabidopsis thaliana	34	Cheng et al., 2012
Arabidopsis thaliana	34	Jing et al., 2015
Brassica juncea	120	Zheng et al., 2022
Brassica napus	118	Zou et al., 2021
Brassica oleracea	64	Yang et al., 2023
Brassica rapa	57	Zhang et al., 2015
Camellia sinensis	25	Guo et al., 2018
Cicer arietinum	19	Ling et al., 2020
Coix lacryma-jobi	31	Wang et al., 2023b
Cucumis melo	30	Zhang et al., 2019
Cucumis sativus	32	Shan et al., 2021
Cucurbita pepo	44	Xu et al., 2022
Eucalyptus grandis	27	Yan et al., 2019
Fragaria	19 (*F. nipponica*)	Zhong et al., 2018
	21 (*F. iinumae*)	
	23 (*F. orientalis*)	
	23 (*F. vesca*)	
	23 (*F. nubicola*)	
	25 (*F. x ananassa*)	
Fragaria	25	Garrido-Gala et al., 2019
Glycine max	74	Wang et al., 2014
Glycine max	74	Zhou et al., 2016
Glycine max	75	Wang et al., 2019
Gossypium	89 (*G. hirsutum*)	Chen et al., 2020
	89 (*G. barbadense*)	
	45 (*G. raimondii*)	
	45 (*G. arboretum*)	
Helianthus annuus	20	Ma et al., 2021

Ipomoea	55 (*I. batatas*)	Si et al., 2023
	58 (*I. triflida*)	
	50 (*I. triloba*)	
	47 (*I. nil*)	
Malus domestica	49	Dong et al., 2018
Medicago truncatula	32	Ling et al., 2020
Nicotiana tabacum	59	Liu et al., 2020
Nicotiana tobacum	61	Yan et al., 2023
Oryza sativa	39	Kim et al., 2013
Oryza sativa	40	Li et al., 2014b
Phyllostachys edulis	29	Wang et al., 2017
Populus trichocarpa	51	Chu et al., 2016
Prunus	55 (*P. yedoensis*)	Zhong et al., 2021
	70 (*P. domestica*)	
	25 (*P. avium*)	
	23 (*P. dulcis*)	
	26 (*P. persica*)	
	23 (*P. yedoensis* var. *nudiflora*)	
Pyrus	41 (*Pyrus bretschneideri*)	Cao et al., 2018
	28 (*Pyrus communis*)	
Saccharum spontaneum	78	Liu et al., 2022a
Setaria italica	32	Liu et al., 2023
Solanum lycopersicum	26	Ding et al., 2019
Triticum aestivum	113	Zhang et al., 2022a
Triticum aestivum	65	Cheng et al., 2022a
Vitis vinifera	18	Wang et al., 2015a
Zea mays	61	Song et al., 2016

III.1.2. The role of the VQ protein in plant growth and development

The diverse spatiotemporal expression levels of *VQ* genes in different tissues and stages indicate their widespread involvement in regulating plant growth and development, suggesting that VQ proteins may balance, promote, or inhibit plant growth (Wang et al., 2014; Zhou et al., 2016; Wang et al., 2017; Cao et al., 2018; Guo et al., 2018; Cai et al., 2019; Ling et al., 2020; Zhang et al., 2022a).

VQ proteins play important roles in regulating seed size, plant fertility, and

growth. For example, the *AtVQ14/IKU1* is strongly expressed in early endosperm development, and the *vq14/iku1* mutant produces only small seeds, indicating that this gene may regulate endosperm development and thus affect seed size (Wang et al., 2010b). The loss-of-function mutants of *AtVQ8* showed yellow-green leaves and delayed growth, while the overexpressing plants of *AtVQ17*, *AtVQ18* and *AtVQ22* respectively showed a stunted phenotype with severely inhibited growth (Cheng et al., 2012). *AtVQ18* and *AtVQ26* are identified as two key VQ members involved in seed germination and early seedling establishment in *Arabidopsis*. Overexpression of *AtVQ18* or *AtVQ26* reduces the ABA response during seed germination, and simultaneously reducing the expression of *AtVQ18* and *AtVQ26* can make the germinated seeds more sensitive to ABA, indicating *AtVQ18* and *AtVQ26* are functionally redundant (Pan et al., 2018). Growth and development of *AtVQ21/MKS1* overexpressing plants is inhibited, but the phenotype of RNAi plants coincides with the wild type (Andreasson et al., 2005). Heterologous overexpression of *AtVQ21* in long-lived flowers (*Kalanchoë blossfeldiana*) and petunia (*Petunia hybrida*) promotes plant dwarfing and delayed flowering (Gargul et al., 2015). Compared with wild-type plants, the overexpression lines of *AtVQ23/SIB1* showed *Pseudomonas syringae* increased resistance but accompanied by varying degrees of growth retardation (Xie et al., 2010). *AtVQ29* plays an important role in the photomorphogenesis of *Arabidopsis* seedlings. The expression level of *AtVQ29* is relatively higher in stem, and the length of the hypocotyl in overexpressed plants is significantly longer than that in wild-type, while it is opposite in *vq29* mutants, indicating that *AtVQ29* regulates the elongation of the hypocotyl of seedlings under far red or weak light conditions (Li et al., 2014a). Meanwhile, *AtVQ29* overexpressing plants also exhibited a delayed flowering phenotype without altering vegetative growth (Cheng et al., 2012). Leaf senescence is another developmental process that may be regulated by the VQ proteins. AtVQ23/SIB1 and AtVQ16/SIB2 can interact with WRKY75 transcription factors to negatively regulate ABA-mediated leaf senescence and seed germination (Zhang et al., 2022b). In addition to *Arabidopsis*, the *VQ* genes in other plants were also found to affect plant growth and development. Overexpression of soybean *GmVQ43* and *GmVQ62* in *Arabidopsis* promotes flowering, and *GmVQ37* is associated with plant fertility (Zhou et al., 2016). Furthermore, respective overexpression of apple *MdVQ37*, *MdVQ25*, and *MdVQ15* in *Arabidopsis* and tobacco affected their vegetative and reproductive growth (Dong et al., 2018). Overexpression of *OsVQ13* can increase grain size, and OsVQ13 can associate with OsMPK6 to influence grain development in rice (Uji et al., 2019). *ZmVQ52* is mainly expressed in maize leaves, and the overexpression of *ZmVQ52* in *Arabidopsis* can accelerate leaf senescence (Yu et al., 2019). Overexpression of wheat *TaVQ25* accelerates leaf senescence in *Arabidopsis* and leads to hypersensitivity reactions in ABA-induced leaf senescence, which is a positive regulatory factor for ABA-related leaf senescence (Meng et al., 2023).

III.1.3. The role of VQ protein in response to biotic stress

In general, plant resists to pathogen infections through two different immune systems: ETI (effector-triggered immunity) and PTI (pattern-triggered immunity) (Jones and Dangl, 2006). ETI is a late defense response induced by the interactions between plant resistance genes and pathogen effectors, associated with programmed cell death known as hypersensitivity. Whereas PTI is an early defense response in which plant pattern recognition receptors (PRRs) at the plasma membrane recognize the invasive PAMP (pathogen-associated molecular patterns) (Coll et al., 2011). The recognition of PAMP activates a series of cell signal events, leading to gene expression, reprogramming of cell biochemistry and metabolism, and ultimately resulting in plant resistance to attempted attacks by pathogens (Senthil-Kumar and Mysore 2013). This is commonly referred to as PAMP-triggered immunity, including surface immune responses such as plasma membrane-localized receptor recognition of PAMPs, production of reactive oxygen and different intracellular molecular events such as the changes in cytosolic calcium concentrations, second messengers, phytohormones, and various phosphorylation cascades (Pecher et al., 2014).

Salicylic acid (SA), and jasmonic acid (JA) are two important hormone signaling molecules mediating plant disease resistance. They are key mediators of plant immunity, participating in plant growth, development and stress resistance. They are accumulated in response to pathogen infection to activate different defense-related gene sets (Glazebrook et al., 2005; Bari and Jones 2009; Caarls et al., 2017; Deenamo et al., 2018). SA signaling pathways generally play a role in the defense of living trophic (biotroph) pathogens; JA-dependent signaling pathways usually exert effects in the defense of dead trophic (necrotroph) pathogens and insects; the two signaling pathways sometimes antagonize and sometimes synergize (Halim et al., 2006; Bari and Jones, 2009). Defense response (DR) related genes, including signal transduction genes and ultimately effective defense genes, often have a crosstalk (overlapping) in plant defense against two nutritional pathogens, regardless of the lifestyle of the pathogen in the host tissue (Lai et al., 2013; Ke et al., 2017). This reflects the complexity and sophistication of plant defense mechanisms when facing various pathogens. Studies have shown that in different plants, the expression of many *VQ* genes is induced or inhibited by JA and SA hormones or different pathogen treatment, and the expression patterns of some genes significantly depend on pathogen types and inoculation time, which may be closely related to different defense mechanisms adopted by plants against different pathogens. Those results indicate that *VQ* genes play an important role in responding to complex signaling pathways mediated by hormones such as JA and SA (Cheng et al., 2012; Kim et al., 2013; Zhou et al., 2016; Jiang et al., 2018; Zhang et al., 2022a).

Overexpression of *AtVQ4/MVQ1* in *Arabidopsis* reduces Flg22-induced resistance and negatively regulates pathogen resistance induced by PAMP (pathogen-associated molecular patterns) (Pecher et al., 2014). *AtVQ10* as a positive regulator, involved in the JA-mediated signaling pathway to resist the

infection of *Botrytis cinerea* (Chen et al., 2018). *AtVQ16/SIB2* is strongly induced by pathogenic *B. cinerea*, and the ability of disease resistance is reduced in its loss-of-function mutants (Lai et al., 2011). Overexpressing *AtVQ20* in *Arabidopsis* reduces the resistance to *Pseudomonas syringae*, indicating that *AtVQ20* negatively regulates the defense response (Cheng et al., 2012). However, *AtVQ21/MKS1* overexpressed plants showed significantly increased resistance to pathogenic *P. syringae* by participating in the SA signal transduction pathway (Andreasson et al., 2005), but with decreased resistance to *B. cinerea* pathogen through negative regulation of the JA signal transduction pathway (Petersen et al., 2010; Fiil and Petersen, 2011), which suggests that *AtVQ21* plays a positive role in SA-regulated defense against biotrophic pathogens, but plays a negative role in JA-mediated immunity against necrotrophic pathogens. The expression level of *AtVQ22* significantly increase after JA treatment, and rapidly accumulate after mechanical injury. Its functional deficient mutants not only enhance its resistance to necrotic pathogen *B. cinerea*, but also enhance resistance to herbivorous insect. It is a negative regulatory factor controlling plant defense without detected adverse effects on plant growth and development (Hu et al., 2013a). *AtVQ23/SIB1* can be significantly induced by SA, JA, as well as pathogens *P. syringae* and *B. cinerea*, and its overexpression plants observably enhance the disease resistance. *AtVQ23* mediates the cross interaction of SA and JA disease resistance signaling pathways (Narusaka et al., 2008; Xie et al., 2010). Overexpression of *AtVQ28* in *Arabidopsis* reduces resistance to *Phytophthora sojae* and *Phytophthora infestans*, promotes *Phytophthora parasitica* infection. *AtVQ28* negatively regulates non-host resistance (NHR) of plants to *Phytophthora* (Lan et al., 2022). Both *AtVQ12* and *AtVQ29* are strongly induced by JA and *B. cinerea* treatment, and its overexpressed plants showed significant susceptible phenotypes, whereas single mutants and double mutants showed significantly improved resistance to *B. cinerea* (Wang et al., 2015b).

Moreover, overexpression of soybean *GmVQ35* or *GmVQ47* in *Arabidopsis* makes plants more susceptible to *B. cinerea*, suggesting that these two *VQ* genes act as negative regulators in the response to necrotrophic pathogens (Zhou et al., 2016). Silencing *GmVQ58* can improve the resistance of soybean to common cutworm (Li et al., 2020). Overexpressing *BnVQ7/BnMKS1* in rapeseed showed enhanced resistance against *Leptosphaeria maculans* in adulthood stage (Zou et al., 2021). Tomato *SlVQ15* mutants showed reduced resistance to *B. cinerea*, while its overexpression plants showed enhanced pathogen resistance (Huang et al., 2022a). Overexpression of *OsVQ14* or *OsVQ32* in rice both enhance plant resistance to *Xoo* (Li et al., 2021). Recently, it was found that rice *OsVQ25* knockout mutants showed significantly enhanced resistance to the pathogens of rice blast and bacterial blight, without penalty of plant growth and development. The explanation of function and mechanism of *OsVQ25* is of great significance for the study of *VQ* gene in food crops (Hao et al., 2022). These results reflect the functional importance and diversity of VQ proteins in response to biotic stress.

III.1.4. The role of VQ protein in response to abiotic stress

Studies have found that the transcriptional expression of a large number of *VQ* genes in many plants is induced or inhibited by salt, drought, temperature, low nitrogen stress and ABA, indicating that *VQ* genes also play an important role in regulating plant responses to abiotic stress (Kim et al., 2013; Wang et al., 2014; Wang et al., 2015a; Zhang et al., 2015; Song et al., 2016; Wang et al., 2017; Dong et al., 2018; Zhang et al., 2019; Liu et al., 2020; Liu et al., 2022; Zhang et al., 2022a). Low temperature, salinity, and drought are common abiotic stresses that impair the growth and development of plants and increase endogenous ABA levels, which is a key hormone helping plants to adapt and survive under these extreme environments (Tuteja 2007; Roychoudhury et al., 2013).

For example, the expression of *AtVQ9* is strongly induced after NaCl treatment, and its mutants showed higher seed germination rate and better seedling growth under NaCl treatment, while the overexpressing lines are opposite, indicating that *AtVQ9* negatively regulates the resistance of *Arabidopsis* to NaCl stress (Hu et al., 2013b). AtVQ15 belongs to the calmodulin binding protein (AtCaMBP25), and its overexpression lines are highly sensitive to NaCl and osmotic stress during seed germination and seedling growth, while the mutants exhibit significant resistance, indicating *AtVQ15* negatively regulates *Arabidopsis* tolerance to NaCl and osmotic stress (Perruc et al., 2004). Overexpression of poplar *VQ1* in *Arabidopsis* enhances its resistance to salt stress and *P. syringae*, and confers a variety of biotic and abiotic stress resistance by mediating ABA and SA signaling pathways (Liu et al., 2022b). The expression of *PeVQ28* in Moso bamboo is induced by salt and ABA; overexpressing *PeVQ28* in *Arabidopsis* showed increased resistance to salt stress and showed more sensitive to ABA; under salt stress, *PeVQ28* overexpressing plants have low malondialdehyde, high proline, and increased expression levels of ABA signaling and ABA synthesis related genes; these results indicate that *PeVQ28* can mediate the positive regulation of salt tolerance through ABA-dependent signaling pathways (Cheng et al., 2020). Overexpressing sweet potato *IbWRKY2* in *Arabidopsis* showed strong drought and salt tolerance. IbWRKY2 can interact with IbVQ4, and *IbVQ4* can be induced by PEG or NaCl treatments, indicating that *IbVQ4* may play an important role in drought and salt tolerance of sweet potatoes (Zhu et al., 2020). The latest research showed that *SlVQ16* positively regulates tomato resistance to salt stress, while *SlVQ21* negatively regulates tomato resistance to salt stress (Ma et al., 2023a). In addition, another latest study found that overexpression of *TAVQ4*-D in *Arabidopsis* and wheat plants increases their tolerance to drought stress. Under drought stress conditions, compared with the wild-type, transgenic wheat plants overexpressing *TAVQ4*-D have increased levels of superoxide dismutase and proline, while decreased levels of malondialdehyde, and at the same time, the expression of genes related to reactive oxygen clearance and stress-related genes are upregulated (Zhang et al., 2023a).

Overexpressing tomato *SlVQ6* in *Arabidopsis* showed high sensitivity to high-temperature stress, illustrating that *SlVQ6* negatively regulates plant

thermotolerance (Ding et al., 2019). Expression of soybean *GmVQ47* in *Arabidopsis* enhances thermotolerance in transgenic plants (Zhou et al., 2016). Under high temperature stress, the relative transcription level of apple *MdVQ37* is significantly down-regulated; the apple plants overexpressing *MdVQ37* exhibited a heat sensitive phenotype, leading to a significant decrease of endogenous SA content and disruption of SA-dependent signaling pathways; and external SA can partially improve the survival rate of transgenic lines. These results show that the regulation of apple *MdVQ37* in response to high temperature stress is related to the changes of transcription factor activity and SA homeostasis (Dong et al., 2021). RsVQ4, as a positive regulator mediating plant thermotolerance in radish, its overexpression can improve heat tolerance, while its RNAi plants have reduced heat tolerance (He et al., 2023).

In addition, the analysis of the dynamic expression patterns of rice *VQ* genes under NO treatment found that 45% (14/31) of *VQ* genes showed significantly differential expression, indicating that VQ genes may play an important role in the physiological processes mediated by NO signals (Peng et al., 2020). Transcriptome data of oxygen-related processes in *Arabidopsis* revealed that 56% (19/34) of *VQ* genes and 64% (48/75) of WRKY genes were up-regulated by ozone treatment (Leon et al., 2021). *COX6B-3* and *COA6-L* can be induced by osmotic stress, and their encoded proteins can interact with AtVQ27 in the presence of NO. The mitochondrial biogenesis was impaired in *vq27* mutant (Kumari et al., 2023). These data indicate that some VQ proteins have potential regulatory roles in response to NO and O_2.

III.1.5. Interactions between VQ and WRKY

Plants have evolved many transcription factor families, which bind to the promoter regions of gene to regulate their expression (Shiu et al., 2005; Qu and Zhu, 2006). The precise regulations of gene expression often require transcription factors and transcriptional co-factors (Buscaill et al., 2014). As one of the largest transcription factor families in plant kingdom, WRKY family is involved in a variety of biological processes, including plant signal transduction, growth and development, and response to biotic and abiotic stresses (Jiang et al., 2017; Li et al., 2020; Wani et al., 2021; Goyal et al., 2022; Khoso et al., 2022). Many VQ proteins interact with WRKY transcription factors by the conserved 'VQ' residues to regulate various biological activities, and WRKY proteins are also the main interacting protein of VQ family (Lai et al., 2011; Cheng et al., 2012; Weyhe et al., 2014). VQ proteins were found to only interacts with class I and class IIc WRKY proteins, the interaction between class I WRKY protein and VQ protein depends on the C-terminal WRKY domain, and the interaction between group IIc WRKY protein and VQ protein depends on its single WRKY domain (Cheng et al., 2012). WRKY domain is mainly associated with DNA binding, so the interaction between VQ protein and WRKY protein affects its DNA binding activity (Bakshi and Oelmüller, 2014).

WRKY33 has been reported to positively regulate plant defense response to *B. cinerea* (Zheng et al., 2006). AtVQ23/SIB1 and AtVQ16/SIB2 can bind to the C-terminal WRKY domain of WRKY33, stimulate the DNA binding activity of WRKY33, enhance its binding ability with W-box, and positively regulate the plant defense against necrotic pathogen *B. cinerea* (Lai et al., 2011; Cheng et al., 2012). Further research has found that AtVQ23/SIB1 and AtVQ16/SIB2 can interact with WRKY33 and WRKY57. The presence of AtVQ23 and AtVQ16 further enhances the competition between WRKY57 and WRKY33, and their fine regulation of *B. cinerea* resistance will be more conducive to plant defense (Jiang and Yu, 2016). In addition, AtVQ23/SIB1 and AtVQ16/SIB2 can interact with WRKY75, inhibit its transcriptional repression function, and regulate the expression of downstream GLKs genes, thus regulating ABA-mediated leaf senescence and seed germination (Zhang et al., 2022b). By interacting with WRKY8 and inhibiting the binding activity between WRKY8 and W-box, AtVQ9 regulates the expression of the salt stress resistance gene *AtRD29A*, thereby negatively regulating plant salt resistance (Hu et al., 2013b). The interaction between AtVQ10 and WRKY8 in the nucleus activates the transcriptional activity of WRKY8, increases the expression of downstream defense related gene *PDF1.2*, and positively regulates the basic defense of *Arabidopsis* against *B. cinerea* (Chen et al., 2018). AtVQ10 also interacts with WRKY25 and WRKY33, and F1 hybrids that overexpressed AtVQ10 and WRKY25 or WRKY33 showed significantly slow and weakened growth, indicating that AtVQ10 and WRKY protein synergistically inhibit plant growth and development (Cheng et al., 2012). AtVQ14/IKU1 can interact with WRKY10/MINI3 to promote the expression of IKU2, which synergistically regulates the development of endosperm, and then affect the size of seeds (Luo et al., 2005; Wang et al., 2010b). AtVQ15 can interact with CaM, WRKY25, and WRKY51, and may link Ca^{2+} signals with transcriptional regulation of downstream targets of osmotic stress signaling pathways, regulating plant tolerance to osmotic stress (Perruc et al., 2004; Cheng et al., 2012). Pollen-specific expressed AtVQ20 can interact with WRKY2 and WRKY34 to form complexes respectively, enhance the transcriptional repression activity of WRKY2 and WRKY34 to regulate the expression of downstream genes related to pollen development and pollen tube germination (such as *AtMYB97*, *AtMYB101* and *AtMYB120*), thereby affecting male fertility (Lei et al., 2017; Lei et al., 2018). AtVQ22/JAV1 interacts with CaMs, JAZ8, and WRKY51 proteins to form a JAV1-JAZ8-WRKY51 complex, which jointly regulates the synthesis of JA. Normally, the JAV1-JAZ8-WRKY51 complex can inhibit the expression of JA biosynthetic gene; when plants are subjected to insect feeding or mechanical damage, extracellular Ca^{2+} flows in and activates intracellular CaMs. CaMs bind to JAV1 in a Ca^{2+} dependent manner, inducing JAV1 phosphorylation and degradation through the ubiquitin-26S proteasome pathway, thereby breaking down the JAV1-JAZ8-WRKY51 complex, relieving the inhibition of JA signals, and promoting the expression of JA synthesis genes (such as *AtAOS*), leading to rapid JA burst and activating plant defense (Hu et al., 2013a; Yan et al., 2018; Ali et al., 2019).

In addition, rice OsVQ8 interacts with OsWRKY10 and inhibits its DNA binding and transcriptional activity; OsWRKY10 negatively regulates rice thermotolerance by regulating ROS balance and allergic reaction, and its interacting protein OsVQ8 plays a antagonistic role. This functional module provides a safe and effective regulatory mechanism for rice responses to heat stress (Chen et al., 2022). Recently, studies have found that rice OsVQ25 balances plant broad-spectrum disease resistance and growth by interacting with the U-Box E3 ubiquitin ligase OsPUB73 and the transcription factor OsWRKY53. OsVQ25 interacts with OsPUB73 and promotes the degradation of OsVQ25 through the 26S proteasome pathway, negatively regulating rice resistance to *M. oryzae* and *Xoo.* Moreover, OsVQ25 interacts with OsWRKY53, a positive regulator of plant immunity, and inhibits its transcriptional activity, and the downstream defense-related genes of OsWRKY53 are up-regulated in OsVQ25 mutants. These results reveal a E3-VQ-WRKY module that can control plant immunity and growth at the transcriptional and post-translational levels (Hao et al., 2022). Wheat TaVQ25 has recently been reported to interact with TaWRKY133 to regulate ABA-mediated leaf senescence (Meng et al., 2023). Soybean GmVQ58 can interact with GmWRKY32, and silencing *GmVQ58* can significantly increase the expression levels of GmWRKY32-downstream defense related genes (*GmVSPβ* and *GmN:IFR*), thereby improving soybean resistance to common cutworm (Li et al., 2020). Banana *MaWRKY26* is induced by cold stress or methyl jasmonate (MeJA), which enhances the cold tolerance of banana fruit. MaVQ5 can interact with MaWRKY26, which weakens the transcriptional activation of JA biosynthesis genes (*MaLOX2*, *MaAOS3*, and *MaOPR3*), indicating that MaVQ5 may act as a suppressor of MaWRKY26 and participate in the JA-mediated responses to cold stress (Ye et al., 2016). The interaction between RsVQ4 and RsWRKY26 promotes the expression of *RsHSP70-20*, thereby positively regulating the heat stress response of radish (He et al., 2023). Apple MdVQ10 and MdVQ15 were found to interact with MdWRKY52 to regulate apple pathogen defense and development (Dong et al., 2018). Apple MdVQ10 can also interact with MdWRKY75 to promote the expression of downstream senescence-related genes, thereby accelerating leaf senescence induced by injury, and is regulated by MdCML15 and MdJAZs (Zhang et al., 2023b). Moso bamboo PeVQ28 and PeWRKY83 interact in the nucleus, and the overexpression of *PeVQ28* promotes the expression of ABA-related genes downstream of PeWRKY83, and positively regulates salt tolerance in plants (Wu et al., 2017; Cheng et al., 2020). Tomato SlWRKY37 interacts with SlVQ7 to promote the expression of downstream *SlWRKY5*3 and *SlSGR1*, positively regulating jasmonic acid- and dark-induced leaf senescence (Wang et al., 2022a). Tomato SlVQ15 interacts with SlJAZ and SlWRKY31 to participate in JA-mediated plant defense responses against *B. cinerea* (Huang et al., 2022a). The latest study found that tomato SlWRKY57 plays a negative regulatory role in salt stress response. SlVQ16 and SlVQ21 competitively interact with SlWRKY57 and antagonize the transcriptional inhibitory activity of SlWRKY57, thereby regulating the expression of downstream salt stress-related genes *SlRD29B*,

SlDREB2, and *SlSOS1*. In addition, the SlWRKY57-SlVQ21/SlVQ16 module also interacts with the SlJAZ proteins and regulates the expression of JA-induced salt stress-related genes, suggesting that this module may be involved in the JA pathway to regulate tomato resistance to salt stress (Ma et al., 2023a).

Based on the studies above, the WRKY transcription factors seem to play a critical role in the function of the VQ protein. VQ proteins regulate the expression levels of downstream related gene by regulating the transcriptional activation or suppression activity of different WRKY transcription factors, and meanwhile the interaction between a WRKY protein and multiple VQ proteins may endow WRKY protein with extensive biological functions (Zheng et al., 2006; Jiang and Deyholos, 2009; Li et al., 2011; Chi et al., 2013).

III.1.6. Interactions between VQ and MAPK

Plant immunity is a precise mechanism by which plants perceive and respond to pathogen attacks (Jones and Dangl, 2006). The immune response triggered by PAMP includes the changes in various phosphorylation cascades, among which one of the earliest and key phosphorylation-mediated signaling events is the activation of MAPK (mitogen-activated protein kinase) (Pecher et al., 2014). The MAPK cascade reaction is composed of three kinases, which play a crucial role in plant immunity. Pathogen attacks trigger phosphorylation of MAPK kinase kinases (MAPKKKs); MAPKKKs then phosphorylate MAPK kinases (MAPKKs), which in turn phosphorylate and activate MAPKs (Meng and Zhang, 2013; Yamada et al. 2016; Bi et al., 2018). Finally, activated MAPKs phosphorylate downstream substrates to post-translationally regulate the function of many proteins, including changes in protein-protein interactions, protein activity, protein stability, etc., therefore promoting signal transduction in various environmental stresses and development processes (Taj et al., 2010; Pecher et al., 2014; Bigeard and Hirt, 2018).

Phosphorylation usually results in the targeted degradation of protein by the ubiquitin-proteasome pathway (Henriques et al., 2009). Some VQ proteins can be targeted phosphorylated by MPK3, MPK4, and MPK6, leading to its degradation, thereby affecting proteins that interact with VQ protein, such as WRKY, and regulating plant immune responses (Pecher et al., 2014; Weyhe et al., 2014). A total of 10 VQ proteins (MPK3/6 targeted VQ proteins, MVQ1-10) in *Arabidopsis* are identified as phosphorylation substrates of MPK3 and MPK6, and these VQ proteins can also interact with a specific subset of WRKYs (Pecher et al., 2014). Normally, AtVQ4/MVQ1 can interact with WRKYs such as WRKY33, and suppress the transcriptional activity of WRKYs, thereby inhibiting the expression of downstream defense genes such as *NHL10*. However, when invaded by bacterial pathogens, AtVQ4/MVQ1 is phosphorylated by MPK3/6, which promotes its instability and degradation, leading to a weakened even ineffective activity repression on WRKYs; the expression levels of downstream disease resistance genes increase along with the phosphorylation level. MAPK3/6 participates in the interaction between AtVQ4/MVQ1 and multiple WRKY transcription factors to

finely regulate the immune response in *Arabidopsis* (Pecher et al., 2014). AtVQ21/MKS1 is identified as a downstream substrate of MPK4. In the absence of pathogen infection, AtVQ21 can bridge the MPK4 and WRKY33 to form a ternary complex, and physical interactions constrain the activity of WRKY33. When infected with pathogens, the activated MPK4 phosphorylates AtVQ21, leading to the denaturation and disassociation of the MPK4-AtVQ21-WRKY33 complex, fully releasing WRKY33 and targeting the promoter of antitoxin gene (*PAD3*) to regulate its expression to enhance plant resistance to pathogens (Andreasson et al., 2005; Qiu et al., 2008; Fiil and Petersen, 2011).

In addition, tomato SlVQ6 has been identified as a phosphorylation substrate of SIMPK1 and plays an important role in the response to some abiotic stresses such as drought, high temperature and salt stress; the overexpression of *SlVQ6* can reduce high temperature tolerance of plants (Ding et al., 2019). It was reported that OsMPK6 positively regulates rice resistance to *Xoc* and grain size/weight (Liu et al., 2015; Ma et al., 2017; Xu et al., 2018). OsVQ1 interacts with OsMPK6 in rice, and *OsVQ1* knockout mutants exhibit stronger resistance to *Xoo*, accumulate high levels of hydrogen peroxide, and show a delayed flowering phenotype under natural long-term conditions (Wang et al., 2021). OsVQ13 can interact with OsMPK6 and activate the OsMPK6-OsWRKY45 component, positively regulating the JA signaling pathway and mediating rice resistance to bacterial blight. Overexpression of *OsVQ13* can also increase grain size, which is associated with OsMPK6 to affect rice grain development (Uji et al., 2019). Both OsVQ14 and OsVQ32 can positively regulate rice resistance to *Xoo*. OsVQ14 and OsVQ32 can interact with and be phosphorylated by OsMPK4. In *OsMPKK6* transgenic plants, OsMPK4 is highly phosphorylated after pathogen infection, enhancing its resistance to *Xoo*, and meantime, phosphorylated OsVQ14 and OsVQ32 also accumulate significantly before and after infection. The results indicate that OsVQ14 and OsVQ32 serve as the substrate of the OsMPKK6-OsMPK4 signaling cascades, enhancing rice resistance to *Xoo*, thus defining a more complete signal transduction pathway in plant induced defense (Li et al., 2021). Recent studies have found that wheat TaVQ4 interacts with MPK3 and MPK6 and plays a role as a phosphorylation substrate for MPK3 and MPK6 in plant drought stress resistance (Zhang et al., 2023a).

Moreover, MAPK can also directly interact with WRKY (Chi et al., 2013). It was shown that WRKY34 is phosphorylated by MPK3 and MPK6 at the early stages of male gametogenesis, and the MAPK-WRKY signaling module plays a crucial role in early pollen development (Guan et al., 2014). When inoculating *Arabidopsis* with *B. cinerea*, the MPK3/6 signaling pathway is activated, which in turn phosphorylates and activates WRKY33 and promotes the expression of *PAD3* to synthesize a large amount of phytoalexin, ultimately leading to a defense response in plant (Zhou et al., 2020). Rice OsWRKY45 has also been identified as a downstream target of OsMPK6 to regulate rice defense responses (Shimono et al., 2007; Ueno et al., 2015). OsWRKY53 can interact with and be phosphorylated by

the OsMAPK6 to regulate BR signaling and plant architecture (Tian et al., 2017), and herbivore-induced defense responses (Hu et al., 2015). A recent study showed that MPK3/MPK6 interacts with and phosphorylates WRKY18 to regulate the expression of two protein phosphatases AP2C1 and PP2C5, and through the MPK3/6-WRKYs-PP2Cs module, PTI triggers and suppresses the ETI responses to balance plant growth and defense (Wang et al., 2023a). Moreover, both barley MPK4 and WRKY1 negatively regulate powdery mildew resistance, and MPK4 phosphorylates WRKY1, enhancing its DNA binding ability and transcriptional inhibitory activity. MKK1-MPK4-WRKY1 module can regulate powdery mildew resistance (Xue et al., 2023). Therefore, it can be seen that VQ protein plays an important role in balancing the interaction between MAPK and WRKY to regulate the function of WRKY transcription factors.

III.1.7. Interactions between VQ and other proteins

Except WRKY and MAPK, VQ protein can also interact with other proteins, such as E3 ubiquitin ligases, CaMs, JAZs, ABI5, and PIF1, participating in the regulation of various life processes in plants. AtVQ15/AtCaMBP25 binds to typical CaM in a calcium-dependent manner, functions as a negative effector of osmotic stress tolerance, and may participate in stress signal transduction pathways (Perruc et al., 2004). AtVQ18 and AtVQ26 can interact with ABI5 transcription factors to inhibit the transcriptional activation ability of ABI5, reducing the expression levels of *AtEM1* and *AtEM6*, and negatively regulate ABA responses during seed germination and seedling establishment (Pan et al., 2018). AtVQ29 interacts with phytochrome binding factor PIF1 during early seedling development to enhance its transcriptional activation activity on downstream genes such as *PIL1* and *XTR7*, thus regulating the elongation of hypocotyls in *Arabidopsis* under different spectra (Li et al., 2014a). AtVQ22/JAV1 can interact with CaMs, JAZ8, and WRKY51 to regulate JA-mediated plant defense (Yan et al., 2018). Meanwhile, AtVQ22/JAV1 also interacts with Ring type E3 ubiquitin ligase JUL1, and JUL1 ubiquitinates JAV1, causing JAV1 degradation through ubiquitin-26S proteasome system, triggering up-regulation expression of *PDF1.2* and activation of JA resistance pathway, thereby conferring *Arabidopsis* biotic stress resistance. This study provides insights into the mechanisms by which the JAV1/JUL1 system specifically coordinates plant defense responses without interfering with plant development or growth (Ali et al., 2019). Similarly, rice OsVQ25 balances plant broad-spectrum disease resistance and growth by interacting with the U-Box E3 ubiquitin ligase OsPUB73 and the transcription factor OsWRKY53 (Hao et al., 2022). Tomato SlVQ15 can interact with SlJAZ2, SlJAZ5, SlJAZ6, SlJAZ7, and SlJAZ11 to regulate JA-mediated plant defense against *B. cinerea* (Huang et al., 2022a). Besides, VQ proteins can also form homodimers or heterodimers. For example, AtVQ12 can strongly interact with AtVQ3, AtVQ8, AtVQ10, AtVQ12,

AtVQ17, AtVQ18, AtVQ29, and AtVQ32. AtVQ12 and AtVQ29 physically interact to form homodimers and heterodimers, negatively regulating the plant defense against *B. cinerea* (Wang et al., 2015b). Apple MdVQ1, MdVQ10, MdVQ15 and MdVQ36 can interact with multiple MdVQ proteins to form heterodimers, while MdVQ15 can form homodimers with itself through C-terminal fragments (Dong et al., 2018). These results provide more important information for studying the interactions and functional pathways of VQ family.

In conclusion, there is a complex and subtle triangular relationship between MAPK, VQ and WRKY, which jointly realize the precise regulation of various physiological processes in a complex environment. Meanwhile, the interaction between VQ protein and other types of proteins may also affect the interaction between VQ protein and transcription factors such as WRKY, thereby altering the gene expression. Based on the results above, we summarized the functions and interacting proteins of the *VQ* genes in different plants (Table 3-2), and finally plotted its intracellular regulatory patterns (Figure 3-2). When subjected to external environmental pressures, phosphorylation signals, calcium signals, hormone signals and photoelectric signals transduction occur in plant cell. This subsequently alters the interactions and modifications between VQ protein and one or more of MAPK/CaM/E3/VQ/WRKY/JAZ8/ABI5/SIG1/PIF1 proteins to activate or suppress related transcription factors to further regulate the expression level of different downstream genes, including hormone, nutrient and metabolite related genes. Finally, plants will respond to external biotic or abiotic stress, such as pathogens, insects, drought, salt, high/low temperature stresses. Here, we proposed a typical signal transduction model, the MAPK-VQ-E3-WRKY cascade, based on these previous studies. When plants encounter external stimuli, MAPK is strongly activated, interacts with VQ proteins, and phosphorylates them, leading to their ubiquitination by E3 ubiquitin ligase and subsequent degradation by 26S proteasome pathway, ultimately triggering the disintegration of VQ-WRKY complex, releasing WRKY to regulate downstream genes (Figure 3-2). Overexpression or mutation of some *VQ* genes, in addition to altering the plant's response to external stimuli, may also be accompanied by the changes of plant growth and development, which is likely due to the sustained expression or suppression of certain downstream genes. These results not only provide a complete molecular perspective on the regulatory network of *VQ* genes, but also provide a reference for cultivating new materials by manipulating these related genes.

Table 3-2. Biological functions and interacting factors of VQ proteins in different plants

Species	Gene ID	Name	Other name	Interacting proteins	Functions	References
Arabidopsis thaliana	AT1G28280	AtVQ4	MVQ1	MPK3/6/10; WRKY33/68	Negatively regulate disease resistance	Pecher et al. 2014; Weyhe et al. 2014
	AT1G32585	AtVQ5			Negatively regulate disease resistance	Cheng et al. 2012
	AT1G68450	AtVQ8		WRKY20/24/34	Regulate plant growth and development	Cheng et al. 2012
	AT1G78310	AtVQ9	MVQ10	WRKY8/20; MPK3/4/6/10/11	Negatively regulate salt stress responses	Cheng et al. 2012; Hu et al. 2013b; Pecher et al. 2014
	AT1G78410	AtVQ10		WRKY8; WRKY25/26/33	Positively regulate disease resistance; Regulate vegetative growth; Regulate meristem development and tolerance to oxidative stress and NO	Cheng et al. 2012; Chen et al. 2018; Gayubas et al. 2023
	AT2G22880	AtVQ12		VQ proteins; WRKY20/23/24	Negatively regulate disease resistance	Cheng et al. 2012; Wang et al. 2015b
	AT2G35230	AtVQ14	IKU1/MVQ9	WRKY10 (MINI3);	Influence endosperm growth and seed size	Wang et al. 2010b; Cheng et al. 2012
	AT2G41010	AtVQ15	CAMBP25	CaM; WRKY25/51	Negatively regulate osmotic stress tolerance	Perruc et al. 2004; Cheng et al. 2012
	AT2G41180	AtVQ16	SIB2	WRKY25/33; WRKY57/75; MPK1-18/20	Positively regulate disease resistance; Regulate leaf senescence and seed germination	Lai et al. 2011; Cheng et al. 2012; Pecher et al. 2014; Jiang and Yu 2016; Zhang et al. 2022b
	AT2G42140	AtVQ17			Regulate growth and development	Cheng et al. 2012
	AT3G44340	AtVQ18		ABI5	Regulate seed germination; Regulate growth and development	Cheng et al. 2012; Pan et al. 2018
	AT3G18360	AtVQ20		WRKY2/34; WRKY20/75	Negatively regulate disease resistance; Modulate pollen development and function	Cheng et al. 2012; Lei et al. 2017
	AT3G18690	AtVQ21	MKS1	MPK4/11; WRKY25/33	Regulate plant growth and disease resistance	Andreasson et al. 2005; Qiu et al. 2008; Petersen et al. 2010; Cheng et al. 2012; Pecher et al. 2014; Gargul et al. 2015
	AT3G22160	AtVQ22	JAV1	WRKY28/51; JAZ8; JUL1	Negatively regulate disease resistance; Coordinate growth and defense	Cheng et al. 2012; Hu et al. 2013; Ali et al. 2019
	AT3G56710	AtVQ23	SIB1	WRKY33/57/75; SIG1; MPK6/14/16; WRKY3/4/20/25	Positively regulate disease resistance; Regulate growth and development; Regulate leaf senescence and seed germination	Morikawa et al. 2002; Narusaka et al. 2008; Xie et al. 2010; Lai et al. 2011; Pecher et al. 2014; Jiang and Yu 2016; Zhang et al. 2022b
	AT3G58000	AtVQ25			Regulate plant disease resistance	Cheng et al. 2012
	AT3G60090	AtVQ26		ABI5	Regulate seed germination	Pan et al. 2018
	AT4G15120	AtVQ27		COX6b-3 and COA6-L	Osmotic stress and mitochondrial biogenesis	Kumari et al. 2023
	AT4G20000	AtVQ28			Negatively regulate disease resistance	Lan et al. 2022
	AT4G37710	AtVQ29		VQ proteins; PIF1	Negatively regulate disease resistance; Regulate photomorphogenesis and flowering	Wang et al. 2015b; Li et al. 2014a

Oryza sativa	Os01g17050	OsVQ1	OsMPK6	Negatively regulate disease resistance; Regulate flowering time	Wang et al. 2021
	Os02g33600	OsVQ8	WRKY10	Regulates plant thermotolerance	Chen et al. 2022
	Os03g47280	OsVQ13	OsMPK6	Positively regulate disease resistance; Positively regulate seed size	Uji et al. 2019
	Os03g57520	OsVQ14	OsMPK4	Positively regulate disease resistance	Li et al. 2021
	Os06g45570	OsVQ25	OsPUB73; OsWRKY53	Negatively regulate disease resistance; Coordinate growth and defense	Hao et al. 2022
	Os08g01260	OsVQ32	OsMPK4	Positively regulate disease resistance	Li et al. 2021
Solanum lycopersicum	Solyc02g078030	SlVQ6	SlMPK1	Negatively regulate thermotolerance	Ding et al. 2019
	Solyc04g074520	SlVQ7	SlWRKY37	Regulate leaf senescence	Wang et al. 2022
	Solyc07g043250	SlVQ15	SlWRKY31; SlJAZ	Positively regulate disease resistance	Huang et al. 2022
	Solyc07g056600	SlVQ16	SlWRKY57; SlJAZ	Positively regulate salt stress	Ma et al. 2023
	Solyc10g007580	SlVQ21	SlWRKY57; SlJAZ	Negatively regulate salt stress	Ma et al. 2023
Malus domestica	MDP0000193206	MdVQ10	VQ proteins; MdWRKY52; MdWRKY75	Regulate pathogen defense and development; Regulating leaf senescence caused by wound	Dong et al. 2018; Zhang et al. 2023b
	MDP0000182830	MdVQ15	VQ proteins; MdWRKY52	Regulate pathogen defense and development	Dong et al. 2018
	MDP0000312336	MdVQ25		Regulate growth and development	Dong et al. 2018
	MDP0000248043	MdVQ37		Regulate growth and development; Negatively regulate thermotolerance	Dong et al. 2018; Dong et al. 2021
Musa acuminata	-	MaVQ5	MaWRKY26	Regulate JA biosynthesis and cold tolerance	Ye et al. 2016
Glycine max	Glyma08g15620	GmVQ35		Regulate disease resistance	Zhou et al. 2016
	Glyma08g18820	GmVQ37		Regulate development and seed setting	Zhou et al. 2016
	Glyma09g05700	GmVQ43		Regulate flowering time	Zhou et al. 2016
	Glyma11g04970	GmVQ47		Regulate disease resistance; Regulate plant thermotolerance	Zhou et al. 2016
	Glyma.14g002800	GmVQ58	GmWRKY32	Negatively regulate pest resistance	Li et al. 2020
	Glyma15g16990	GmVQ62		Regulate flowering time	Zhou et al. 2016
Ipomoea batatas	-	IbVQ4	IbWRKY2	Regulate drought and salt tolerance	Zhu et al. 2020
Phyllostachys edulis	PH01007611G0010	PeVQ28	WRKY83	Positively regulate plant salt tolerance	Cheng et al. 2020
Raphanus sativus	LOC108856901	RsVQ4	RsWRKY26	Regulate plant thermotolerance	He et al. 2023
Zea mays	GRMZM2G122447	ZmVQ52	ZmWRKY20/36/50/71	Regulate leaf senescence	Yu et al. 2019
Brassica napus	BnaA01g36880D	BnVQ7	BnMKS1	Positively regulate disease resistance	Zou et al. 2021
Populus trichocarpa	Potri.001G029700	PtVQ1		Positively regulate salt tolerance/disease resistance	Liu et al. 2022b
Triticum aestivum	TraesCS1D02G340900	TaVQ4	TaMPK3/6	Positively regulate drought tolerance	Zhang et al. 2023a
Triticum aestivum	TraesCS4A02G290800	TaVQ25	TaWRKY133	Regulate leaf senescence	Meng et al. 2023

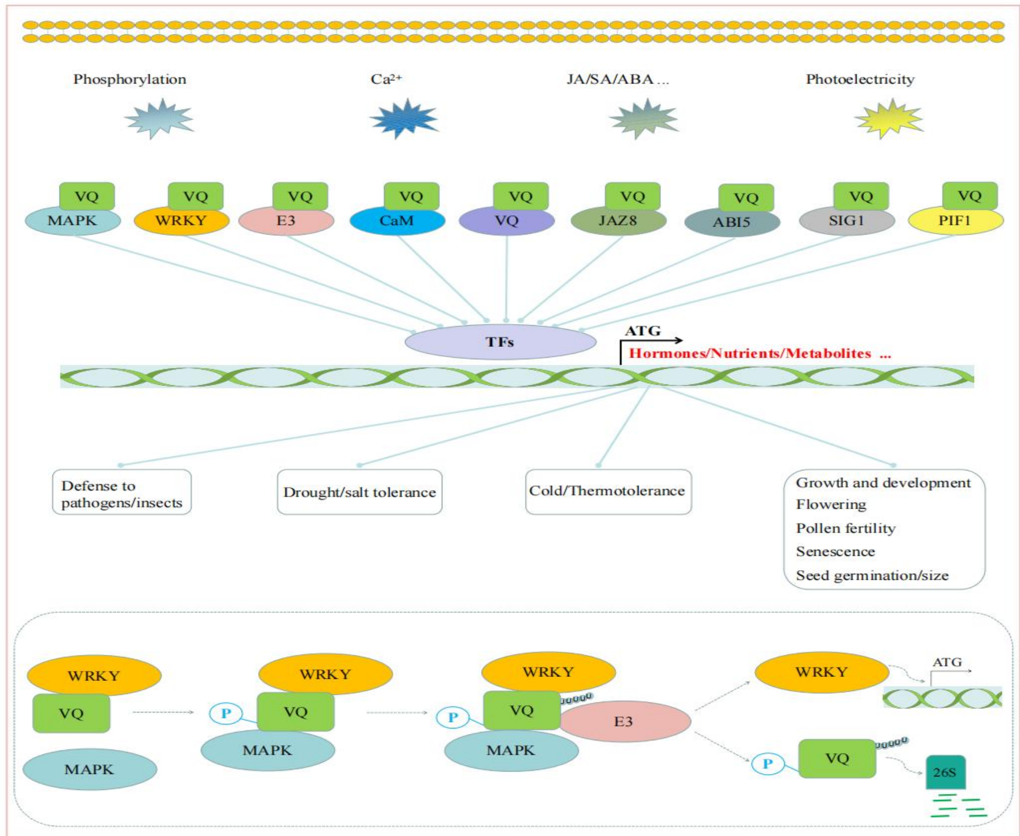


Figure 3-2. The functional model and interaction network of *VQ* family genes.
When faced with pathogen invasion or adversity stress, plants intracellular signal transduction will occur. Subsequently, the interaction between VQ protein and different proteins such as MAPK, CaM, E3, and transcription factors will be strengthened, which in turn affects the transcriptional activation or inhibition activity of these transcription factors, causing changes in downstream gene expression, ultimately affecting plant response and phenotype. The bottom frame displays a typical signal transduction mode, namely the 'MAPK-VQ-E3-WRKY' cascade. When exposed to external stimuli, the MAPK phosphorylation cascades are activated, MAPK phosphorylates VQ protein. Then, E3 ubiquitin ligase recognizes and ubiquitinates the phosphorylated VQ protein, leading to its degradation through the 26S proteasome pathway, thereby disassembling the VQ-WRKY complex and completely releasing WRKY transcription factor to regulate the expression of downstream genes

III.1.8. The importance of VQ domain and the molecular breeding strategies based on CRISPR/Cas9 system and VQ genes

The VQ motif is the most conserved and important feature of VQ proteins, and several studies have shown that VQ motif has a significant impact on the function of VQ proteins. For example, mutation of IVQQ to EDLE in the VQ domain of AtVQ14 results in smaller seeds, while mutation at other positions do not have this

phenotype (Wang et al., 2010b). Further study found that the role of the VQ domain may be achieved through the following two mechanisms. First, VQ domain can affect the subcellular localization of VQ protein. For example, the fusion protein of AtVQ9-GFP is localized in the nucleus, but when VVQK mutated into EDLE in the VQ domain of AtVQ9, it is distributed in the nucleus and cytoplasm (Hu et al., 2013b). The VQ motif of AtVQ21/MKS1 is not only associated with the interactions of WRKY33/WRKY25, but also involved in the nuclear localization of AtVQ21 (Petersen et al., 2010). However, not all VQ domain are associated with subcellular localization, for instance, the mutation of VQ to AA at the VQ domain of AtVQ23 does not affect its localization, possibly because AtVQ23 contains a nucleus and chloroplast-targeting signal peptides in addition to the VQ motif (Lai et al., 2011). Second, the VQ domain can affect the interactions between VQ protein and other proteins. For example, a mutation in the VQ domain can cause the interactions of AtVQ9-WRKY8 (Hu et al., 2013b), AtVQ14-WRKY10 (Wang et al., 2010b), AtVQ20-WRKR2/34 (Lei et al., 2017), and AtVQ23-WRKY33 (Lai et al., 2011) failed respectively, suggesting that the VQ motif is the core element for interaction with WRKY. In addition, an amino acid substitution in AtVQ29 (V70A or V70D) abolishes its repression activity, indicating that the VQ motif is essential to mediate AtVQ29-regulated transcriptional activity, and interestingly, the double substitution in the VQ motif (V70D and Q71L) causes significant induction of the reporter *LUCIFERASE* (Li et al., 2014a), indicating that transcription regulatory activity can be shifted from repression to activation through modification of the VQ motif.

However, the VQ motif is not the only critical domain for VQ protein to interact with other proteins, and mutation in other sites can also lead to the loss of original function of VQ protein. For example, the VQ motif is not essential for the interaction or phosphorylation with MPK3/MPK6 (Pecher et al., 2014). Mutations in the VQ domain affect the transcription regulatory activity of AtVQ29, but do not alter its ability to interact with PIF1 (Li et al., 2014a). Studies have shown that the amino acid sequences on both sides of the VQ motif are highly diverse in VQ protein. The variable regions and sub-motifs of VQ protein can also regulate the interplay with WRKY, and affect the binding affinity and specificity of WRKY-VQ complex by interacting with other proteins (Cheng et al., 2012; Chi et al., 2013). When studying the interaction between AtVQ12 and AtVQ29, it was found that mutating the amino acid residues LVQR into EDLE does not affect their interaction, but the C-terminal mutation abolishes the interaction between the two VQ proteins, speculating that the interaction between VQ-VQ proteins requires a C-terminal fragment rather than a VQ motif. Therefore, AtVQ12 and AtVQ29 may interact with each other through C-terminal, and interact with WRKY33 through the VQ motif to form a large protein complex mediating the plant defense against *B. cinerea* (Wang et al., 2015b). In addition, when two lysine residues K52 and K179

in AtVQ22/JAV1 mutate into two threonine T, the interaction ability between AtVQ22/JAV1 and E3 ubiquitin ligase JUL1 is significantly decreased (Ali et al., 2019). The interaction between apple MdVQ10/MdVQ15 and MdWRKY52 requires not only the VQ motif, but also several amino acid residues flanking the VQ motif (Dong et al., 2018).

With the advancement of technology, genome editing, including gene knockout, gene replacement, base editing, fragment insertion and deletion, has been widely applied in plants based on CRISPR/Cas system (Gao 2021; Molla et al., 2021). Therefore, molecular design breeding based on the genes related to certain important agronomic traits is a promising crop breeding strategy (Zhang et al., 2018; Xing et al., 2020; Xu et al., 2021; Huang et al., 2022b; Song et al., 2022). Given the regulatory functions of VQ proteins in plant disease resistance, stress resistance, and growth and development, as well as its multiple functional modules in structure, *VQ* genes showed the potential for precise manipulation to achieve modern molecular breeding. Therefore, we proposed a molecular design breeding strategy based on *VQ* genes (Figure 3-3). Firstly, through big-data analysis, including genomics, transcriptomics, metabolomics, proteomics, and phenomics, candidate *VQ* genes associated with important traits in plants can be identified. These *VQ* genes may be related to plant responses to biotic or abiotic stress, growth and development, yield and quality, and on the other hand, may be related to plant hormones, metabolism, and nutrition. Secondly, based on the structure, interaction proteins, and action pathways of the *VQ* gene, key active sites and domains on the *VQ* gene, as well as core regulatory elements on its promoter sequence, can be identified. Using these information, we can knock out one or more *VQ* gene, replace key loci, delete specific fragment, insert specific fragment, and edit promoter. Thirdly, strategies such as manual design, AI (artificial intelligence) design, genome editing, transgene, and mutant library screening can be utilized to achieve the modification of *VQ* gene. Especially in AI design, it has begun to be recognized and increasingly applied in protein structure prediction and new functional protein development (Baek et al., 2021; Tunyasuvunakool et al., 2021; Bryant et al., 2022; Dauparas et al., 2022; Wang et al., 2022b; Huang et al., 2023; Lutz et al., 2023; Madani et al., 2023; Watson et al., 2023). Fourthly, the characteristics of *VQ* gene molecular design materials are clarified through molecular level evaluation, phenotype identification, and functional verification. Finally, the action mechanism of the *VQ* gene can be revealed, and new materials with specific or comprehensive trait improvements can also be obtained. This provides potential promising research ideas and implementation approaches for further utilizing gene resources to achieve molecular design breeding, intelligent breeding, and precision breeding.

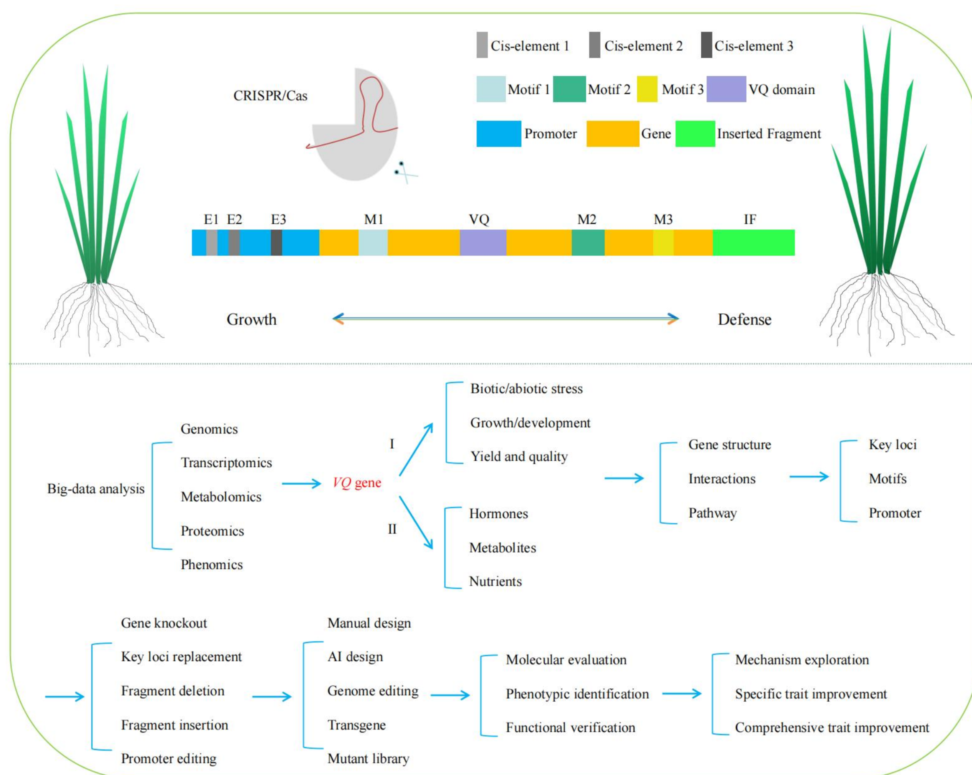


Figure 3-3. Molecular design breeding strategy based on *VQ* gene.
In the structure of *VQ* genes, blocks with different colors represent different functional units. Three different elements are exemplified on the promoter and gene sequences, respectively. E1 represents cis-element 1, E2 represents cis-element 2, and E3 represents cis-element 3; M1 represents motif 1, M2 represents motif 2, and M3 represents motif 3; VQ represents VQ domain; IF represents the inserted fragment. Mining candidate *VQ* genes through big-data analysis, identifying key loci, and applying genome editing tools for modification. The modification of cis-elements in the promoter and motifs in the coding region will further promote the working mechanism analysis of VQ proteins and create a series of new materials with different phenotypes

III.2. Conclusions and perspectives

VQ protein is a kind of conserved VQ motif protein, which widely exists in various plants in the form of multi-gene family. Early studies suggested that VQ family is a class of transcription regulatory factors unique to plants (Jing and Lin, 2015), but afterwards, single or several *VQ* genes were identified in some fungi, lower animals and bacteria, proving that it is not unique to plants but an ancient gene family (Jiang et al., 2018). In recent years, a series of studies on the function and mechanism of *VQ* genes have found that they are not only involved in the response to biotic and abiotic stresses, but also involved in the regulation of various life processes of plants by interacting with MAPK, WRKY and other proteins (Jing and Lin, 2015). In this paper, we comprehensively summarized the published research papers about *VQ* genes, and further depicted the functional model and

regulation network of VQ proteins, which help us to better understand the interactions between plant and environment. In view of the results above, by the way of interaction with many proteins such as MAPK and WRKY, VQ proteins play an very important role in balancing plant stress resistance, immunity, and growth and development. Therefore, the functional diversity and structural importance of VQ proteins endow them with operability significance, which has great potential value for modern molecular breeding based on *VQ* genes.

However, there are still many limitations on *VQ* gene in the current study. Firstly, although the number, structure, and expression patterns of *VQ* genes have been preliminarily analyzed in some plants, published researches about its function and mechanism are mainly focused on *Arabidopsis*, and there is very little research on other higher plants, especially crops. Additionally, the *VQ* family genes are rarely systematically identified and analyzed in gymnosperms and other plants, therefore, further research is needed on their molecular features, expression levels and evolutionary comparisons with angiosperms. Secondly, VQ family contains a large number of members with significant sequence variations, therefore it remains to be determined whether VQ motif only plays an important role only in certain VQ proteins, which means although some proteins contain VQ motif, this domain or even the entire protein actually do not play any roles. Therefore, identifying VQ proteins with key regulatory capabilities and discovering their structural and functional patterns is highly valuable. Moreover, mutations of the VQ domain can change its transcriptional regulation activity from inhibition to activation (Li et al., 2014a), reminding us to manipulate its activity through gene replacement or base editing via CRISPR/Cas9 system, which may produce diverse plant phenotypes. Thirdly, from the perspective of sequence similarity, interacting proteins, and gene expression patterns, VQ proteins exhibit functional diversity or redundancy, which still need further experimental verification. For example, both AtVQ12 and AtVQ29 can negatively regulate plant defense against *B. cinerea*, and *vq12/vq29* double mutant plants displayed greater resistance than *vq29* single mutant and wild type (Wang et al., 2015b). But once these genes are mutated simultaneously, although it may greatly improve plant resistance, it is also likely to have a significant impact on plant development. So, it will be of great significance for crop breeding if knocking out some *VQ* genes can observably improve plant resistance or agronomic traits without affecting growth and development, such as *OsVQ25* (Hao et al., 2022). Fourthly, the working mechanisms of VQ protein are mainly focused on some known interacting patterns, such as MAPK-VQ, WRKY-VQ, CaM-VQ, and E3-VQ. However, the multiple roles of VQ protein in plant growth, development and defense may indicate that signal transduction of these processes requires strict regulation and fine-tuning. Therefore, more interaction combinations (such as MYB-VQ, MADS-VQ, AP2/ERF-VQ, and bHLH-VQ), more specific pathway or mechanisms (such as the regulatory mechanisms of plant development), and even more upstream and downstream networks need to be further uncovered.

In summary, although many achievements have been made, research on the VQ family is still limited now. Therefore, exploring the specific functions and

mechanisms of *VQ* genes in more plants and practicing molecular breeding strategies based on *VQ* genes will further promote the development and application of genetic resources.

4

Chapter IV Large-scale identification and characterization analysis of *VQ* family genes in plants, especially gymnosperms

Reference: Tian, J., Zhang, J., Francis, F. (2023). Large-scale identification and characterization analysis of *VQ* family genes in plants, especially gymnosperms. Int. J. Mol. Sci. 2023, 24, 14968. https://doi.org/10.3390/ijms241914968

Abstract

VQ motif-containing (VQ) proteins are a class of transcription regulatory cofactors widely present in plants, playing crucial roles in growth and development, stress response, and defense. Although there have been some reports on the member identification and functional research of *VQ* genes in some plants, there is still a lack of large-scale identification and clear graphical presentation of their basic characterization information to help us better understand this family. Especially in gymnosperms, the *VQ* family genes and their evolutionary relationships have not been reported yet. In this study, we systematically identified 2469 *VQ* genes from 56 plant species, including bryophytes, gymnosperms, and angiosperms, and analyzed their molecular and evolutionary features. We found that amino acids are highly conserved only in the VQ domain, while other positions are relatively variable; most *VQ* genes encode relatively small proteins and do not have introns; the GC content in Poaceae plants is the highest (up to 70%); these VQ proteins can be divided into 9 subgroups. In particular, we analyzed molecular characteristics, chromosome distribution, duplication events, and expression levels of *VQ* genes in three gymnosperms: *Ginkgo biloba*, *Taxus chinensis*, and *Pinus tabuliformis*. In gymnosperms, *VQ* genes are classified into 11 groups, with highly similar motifs in each group; most VQ proteins have less than 300 amino acids and are predicted to be located in nucleus; tandem duplication is an important driving force for the expansion of the *VQ* gene family, and the evolutionary process of most *VQ* genes and duplication events is relatively independent; some candidate *VQ* genes are preliminarily screened, and they are likely to be involved in plant growth, stress and defense responses. These results provide detailed information and powerful references for further understanding and utilizing the *VQ* family genes in various plants.

Keywords: large-scale identification, clear graphical presentation, *VQ* family genes, molecular and evolutionary features, gymnosperms

IV.1. Introduction

VQ motif-containing (VQ) protein is known as a transcription regulatory cofactor for interacting with transcription factors (TFs) to regulate gene expression. In 2002, *AtVQ23*/*AtSIB1* was first identified in *Arabidopsis* (Morikawa et al., 2002). Subsequently, the *VQ* genes were successively identified by bioinformatics and experimental methods in various plants, such as rice, soybean, maize, grape, and wheat (Yuan et al., 2021; Zhang et al., 2022).

VQ proteins contain the highly conserved VQ motif (FxxxVQxhTG, x represents arbitrary amino acid, h represents hydrophobic amino acid) (Jiang et al.,

2018; Yuan et al., 2021). The three terminal amino acids in the VQ motif may have different types in different plants. For example, six types (LTG, FTG, VTG, YTG, LTS, and LTD) were identified in *Arabidopsis*, four types (LTG, FTG, VTG, and ITG) were identified in rice, and five types (LTG, FTG, VTG, ITG, and VMA) were identified in wheat (Zhang et al., 2022). Most *VQ* genes in higher plants have no introns and encode relatively small proteins with less than 300 amino acids (Jiang et al., 2018; Yuan et al., 2021). The lack of introns in these gene sequences makes transcription and translation more efficient, resulting in the production of these small proteins. Most VQ proteins are located in the nucleus, with a few in the chloroplast and cytoplasm (Cheng et al., 2012; Kim et al., 2013; Garrido-Gala et al., 2019). Moreover, it has been reported that VQ motif is related to the interaction with WRKY (Yuan et al., 2021). These basic characteristics provide a certain reference for us to understand and identify *VQ* family genes.

Jasmonic acid (JA), salicylic acid (SA), and abscisic acid (ABA) are important hormone signaling molecules in plants, which are widely involved in plant growth and development, disease resistance, and stress responses (Bari et al., 2009; Glazebrook et al., 2003; Tuteja, 2007; Roychoudhury et al., 2013). Research has shown that the expression of many *VQ* genes in different plants is induced or inhibited by JA, SA, or ABA hormones, as well as treatments such as pathogens, drought, and salt (Zhang et al., 2022; Cheng et al., 2012; Kim et al., 2013). This indicates that the *VQ* gene plays an important role in complex signaling pathways in response to JA, SA, and ABA hormones and various stresses. In addition, WRKY, MAPK, and calmodulin (CaM) have been reported to play very important roles in plant life processes and responses to external stimuli (Khoso et al., 2022; Wani et al., 2021; Pitzschke et al., 2009; Taj et al., 2010; Perochon et al., 2011). By interacting with proteins such as CaM, WRKY, and MAPK, VQ proteins can mediate plant growth and development, as well as defense responses to biotic and abiotic stresses (Yuan et al., 2021; Cheng et al., 2012; Perruc et al., 2004; Yan et al., 2018; Pecher et al., 2014; Weyhe et al., 2014).

Studies have shown that VQ proteins are involved in growth and development. For example, in *Arabidopsis*, the loss-of-function mutants of *AtVQ8* showed yellow-green leaves and delayed growth, while plants overexpressing *AtVQ17*, *AtVQ18*, or *AtVQ22* showed a stunted phenotype with severely inhibited growth (Cheng et al., 2012). AtVQ14/IKU1 interacts with WRKY10/MINI3 to regulate the development of seed (Luo et al., 2005; Wang et al., 2010). AtVQ20 interacts with WRKY2 and WRKY34 to regulate the expression of downstream *MYB* genes, thereby affecting male fertility (Lei et al., 2018). AtVQ18 or AtVQ26 can interact with ABI5 and inhibit its transcriptional activation ability to negatively regulate ABA responses during seed germination and seedling establishment (Pan et al., 2018). AtVQ29 interacts with PIF1 to enhance its transcriptional activation activity, regulating the elongation of hypocotyls under different spectra (Li et al., 2014).

VQ proteins are involved in the response to abiotic stresses. For example, AtVQ9 interacts with WRKY8 to regulate the expression of *AtRD29A*, negatively regulating plant salt tolerance (Hu et al., 2013). AtVQ15 interacts with CaM, WRKY25, and WRKY51 to regulate plant tolerance to osmotic stress (Cheng et al., 2012; Perruc et al., 2004). Bamboo PeVQ28 interacts with PeWRKY83 to promote the expression of ABA-related genes, and positively regulates salt tolerance in plants (Cheng et al., 2020). Ectopic overexpression of tomato *SlVQ6* in *Arabidopsis* showed decreased high temperature tolerance (Ding et al., 2019). The hypermorphic mutant of *AtVQ10* exhibited enhanced meristem development, increased tolerance to oxidative stress, and reduced sensitivity to NO (Gayubas et al., 2023).

VQ proteins are involved in the response to biotic stresses. AtVQ4/MVQ1 is phosphorylated by MPK3/6, which in turn affects the interaction of VQ-WRKY complex to finely regulate the immune response in *Arabidopsis* (Pecher et al., 2014). Both AtVQ16/SIB2 and AtVQ23/SIB1 can bind to WRKY33 to regulate plant defense against *Botrytis cinerea* (Cheng et al., 2012; Lai et al., 2011). AtVQ21/MKS1 interacts with MPK4 and WRKY25/33 to regulate the expression of downstream genes. Overexpression of *AtVQ21*/*MKS1* significantly increases resistance to *Pseudomonas syringae* by participating in the SA pathway (Andreasson et al., 2005), but with decreased resistance to *B. cinerea* through negative regulation of the JA pathway (Petersen et al., 2010; Fiil et al., 2011). AtVQ22/JAV1 interacts with Ca^{2+}/CaM, JUL1, JAZ8, and WRKY51 to jointly regulate JA synthesis, leading to rapid JA burst and activating plant defense (Yan et al., 2018; Ali et al., 2019). OsVQ25 interacts with OsPUB73 and OsWRKY53 to balance broad-spectrum disease resistance and growth of rice (Hao et al., 2022).

Some gymnosperms have high ornamental, medicinal, and economic values. They are also important materials for studying plant evolution. However, the research on the *VQ* family genes in gymnosperms has not been reported yet. Many gymnosperms have large genomes and numerous repetitive sequences, making it difficult to assemble a complete genome. But in recent years, with the advancement of technology, some gymnosperms have gradually completed genome sequencing and assembly (Liu et al., 2021; Xiong et al., 2021; Liu et al., 2022; Niu et al., 2022; Fu et al., 2023). Gymnosperms can also encounter various diseases and pests (such as leaf spot disease, rust, aphids, and pine caterpillars), as well as environmental stresses such as drought, low temperature, high temperature, and saline alkali, which endanger their growth and survival. In view of the prominent role of *VQ* genes in the growth and development and the responses to environmental stress in angiosperms such as *Arabidopsis* and rice, *VQ* genes may also participate in various life processes in gymnosperms. Therefore, studying the morphology, evolution, and expression of *VQ* genes in gymnosperms will be very interesting and meaningful.

Although *VQ* genes have been structurally and functionally identified in some plants, there is a lack of systematic analysis and intuitive display to clearly show their characteristics in a wider range of plants. Here, in order to know the detailed information on *VQ* family genes in the plant kingdom, we identified a total of 2469 *VQ* genes from 56 plant species, including bryophytes, gymnosperms, and angiosperms. We carried out a comprehensive bioinformatics analysis, including conserved motifs, basic molecular characterization, and systemic clustering. Importantly, for gymnosperms, we selected *Ginkgo biloba*, *Taxus chinensis* and *Pinus tabuliformis* as the research objects, identified their *VQ* gene members, and analyzed their molecular features, gene structure, subcellular location, chromosome distribution, duplication events, expression levels, synteny blocks, and evolutionary comparisons. Our results provide fundamental information for the characterization and evolution of *VQ* genes in gymnosperms and angiosperms, which will be valuable for further research on their biological function and working mechanism.

IV.2. Materials and methods

IV.2.1. Identification of VQ genes in multiple plants

In the present study, genome sequence and annotation files of the 56 plant species were mainly collected from Ensembl Plant database (http://plants.ensembl.org/index.html/, accessed on 2 March 2023) (Bolser et al., 2017), NCBI plant genome database (ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/plant/, accessed on 2 March 2023), phytozome database (https://phytozome-next.jgi.doe.gov/, accessed on 2 March 2023) (Goodstein et al., 2012), and National Genomics Data Center (https://ngdc.cncb.ac.cn/, accessed on 2 March 2023). The Hidden Markov model (HMM) profile of the VQ domain (PF05678) was obtained from the Pfam database (http://pfam.xfam.org/, accessed on 20 March 2023) (Mistry et al., 2021). The VQ family members were retrieved from plant protein sequences using the HMMSEARCH program of the HMMER (v3.0) software (Potter et al., 2018). The online program SMART tool (http://smart.embl-heidelberg.de/, accessed on 25 March 2023) (Letunic et al., 2021), the NCBI Conserved Domains Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/, accessed on 25 March 2023) (Marchler-Bauer et al., 2015), and NCBI-BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi/, accessed on 25 March 2023) were used to ultimately determine the *VQ* genes. Default parameters were adopted for all the softwares mentioned above.

IV.2.2. Molecular features and chromosomal localization analysis

The genome information, such as genome size, chromosome ploidy, and

numbers of total coding genes were obtained from Ensembl Plant database (http://plants.ensembl.org/index.html/, accessed on 2 March 2023) and other related published papers. GC content in coding region, protein length, and sequence similarities of aligned sites of *VQ* genes were calculated by Perl program (https://www.perl.org/, accessed on 2 March 2023). The relevant scripts and their input data can be downloaded from https://github.com/ywxkjtsd123/code (accessed on 15 September 2023). After alignment of 2469 VQ protein sequences, the largest proportion of the same amino acids in each alignment site was calculated. The biophysical properties of the VQ proteins, including peptide length, isoelectric point (pI), and molecular weight (MW), were estimated using the online program ExPasy (https://web.expasy.org/protparam/, accessed on 7 April 2023) (Gasteiger et al., 2003). The physical locations of the *VQ* genes on the chromosomes were visualized using the MapChart (v2.3) software (Voorrips, 2002). The subcellular localization prediction tool DeepLoc-2.0 (https://services.healthtech.dtu.dk/services/DeepLoc-2.0/, accessed on 7 April 2023) (Thumuluri et al., 2022) was used to predict the likely location of the *VQ* genes. Online tools NLStradamus (http://www.moseslab.csb.utoronto.ca/NLStradamus/, accessed on 15 September 2023) (Nguyen et al., 2009) and PSORT (https://www.genscript.com/psort.html/, accessed on 15 September 2023) (Horton et al., 2007) were used to predict nuclear localization signals of protein.

IV.2.3. Conserved motifs, gene structure, and phylogenetic tree analysis

The conserved motifs of the VQ proteins were detected using the online program MEME (https://meme-suite.org/, accessed on 15 April 2023) with default parameters (Bailey et al., 2015). Conserved motifs were drawn using the TBtools (v1.123) software (Chen et al., 2020). The exon and intron structures were determined by GFF file. After extracting their location information from the annotation file, the gene structures of the *VQ* genes were visualized using TBtools software. Based on the protein sequences, all multiple sequence alignments were carried out using the MAFFT (v7.511) software (Katoh et al., 2013) with default parameters to study the evolutionary relationships and classification of the *VQ* genes. Depending on the alignment results, the phylogenetic tree was built using the FastTree2 software (Price et al., 2010) with default parameters. The online tool iTOL (https://itol.embl.de/, accessed on 15 April 2023) (Letunic et al., 2021) was used to draw and adjust the phylogenetic tree. TBtools software was used to integrate phylogenetic trees, conserved motifs results, and gene structure results in three gymnosperms.

IV.2.4. Gene duplication and collinearity analysis

Gene duplication mainly includes segmental duplication, tandem duplication, dispersed duplication, and proximal duplication (Wang et al., 2012). Segmental duplicates exist in collinear blocks. Tandem duplicates are defined as closely

adjacent to each other on the same chromosome. Proximal duplicates are on the same chromosome and are close to each other, but separated by several other genes. Dispersed duplicates occur on the same or different chromosomes, and they are neither close to each other nor within conserved collinearity blocks (Wang et al., 2012). For the definition of tandem duplicates, we added the following concepts based on literature: two or more *VQ* genes adjacent to each other within 200 kb can be defined as tandem duplication events (Xu et al., 2022). To identify duplication events of *VQ* gene family in the gymnosperms, the coding sequence (CDS) of all *VQ* genes were aligned using BLASTN with E-value below 1×10^{-15}. For *VQ* gene with different isoforms, we selected the longest one for analysis. We used the following criteria to search for duplicate *VQ* gene pairs: both identity and coverage > 75% at the nucleotide level (Zhang et al., 2022; Kovach et al., 2010). In addition, BLASTP (E-value $< 1 \times 10^{-5}$, top 5 matches) was used for seqnence alignment, MCScanX (Wang et al., 2012) with the default parameters was used to identify collinear blocks within or between species, and TBtools software and Circos program (Krzywinski et al., 2009) were used to visualize the collinearity maps and exhibit segmentally duplicated *VQ* gene pairs. The values of nonsynonymous substitution rate (Ka) and synonymous substitution rate (Ks) of duplicated *VQ* gene pairs were calculated to evaluate the selection pressure using the KaKs_Calculator 2.0 (Wang et al., 2010) with the NG method. The relevant scripts and their input data can be obtained from https://github.com/ywxkjtsd123/code (accessed on 15 September 2023).

IV.2.5. Gene expression patterns analysis

The RNA-seq data of *VQ* genes in different tissues were obtained from previous research (Liu et al., 2021; Xiong et al., 2021; Niu et al., 2019; Zhang et al., 2021; Li et al., 2023). The sample name, accession number, and data source were listed in Table 4-S5. Fastqc (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, accessed on 20 April 2023) was used for quality control and Trimmomatic (v0.39) (Bolger et al., 2014) was used for data filtering to obtain the clean reads. HISAT2 (Kim et al., 2019) was used for genome library construction and alignment. Transcript abundance was measured using TPM value. TPM values were calculated with StringTie2 (Pertea et al., 2015) for *Ginkgo biloba* and *Taxus chinensis.* Due to the larger genome size and the limitations in computing resources, Bowtie2 (Langmead et al., 2012) was used to build the index library and conduct sequence alignment, and RSEM (v1.3.3) (Li et al., 2011) was used to obtained the TPM values for *Pinus tabuliformis*. Heatmap was generated using log2 (TPM + 1) values by TBtools software. Default parameters were used for all the softwares.

IV.3. Results

IV.3.1. Members identification and conserved motif analysis of VQ genes in multiple plants

The Hidden Markov Model (HMM) of the VQ motif (PF05678) was used to

search the coding proteins for putative *VQ* genes in each species. After manually removing invalid entries and validating the search results, we systematically identified a total of 2469 *VQ* family genes from 56 plant species, including 3 bryophytes, 3 gymnosperms, and 50 angiosperms. Moreover, their basic information, such as VQ protein sequences (Table 4-S1), chromosome ploidy, genome size, and numbers of total coding genes or *VQ* genes (Table 4-S2) were summarized and listed for each species. We found that there is not any necessary relationship between genome size and *VQ* gene numbers (Figure 4-1a), consistent with previous studies (Jiang et al., 2018; Xu et al., 2022).

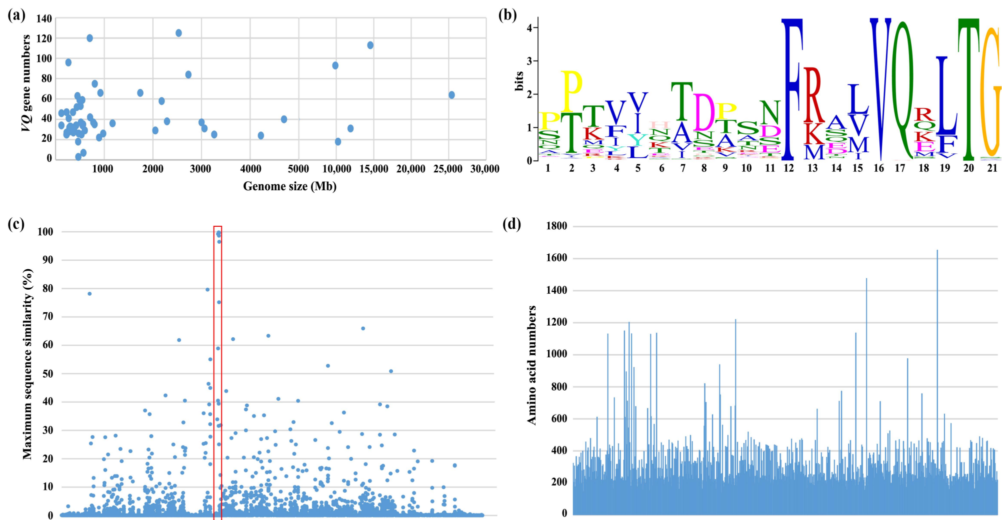


Figure 4-1. Members identification and sequence characteristics of *VQ* genes. (a) The relationship between genome size and *VQ* gene numbers in 56 plant species. (b) The conserved VQ motif. (c) Maximum sequence similarity in each alignment site. The red frame is region of VQ motif and X-axis represents alignment site. (d) The length of VQ proteins. X-axis represents 2469 VQ proteins and Y-axis shows amino acid numbers of VQ protein.

We scanned 2469 VQ proteins, and found the most conserved motif is FxxxVQxhTG, where three main terminal amino acids, namely LTG, FTG, and VTG (Figure 4-1b). The core element 'VQ' in FxxxVQxhTG, is slightly changed in some plants, such as FxxxVHxhTG (Table 4-S1), which is in agreement with previous study (Zhang et al., 2022). Studies have shown that many VQ proteins contain single or dual component nuclear localization signals, and some also contain chloroplast targeting signals (Cheng et al., 2012; Kim et al., 2013; Garrido-Gala et al., 2019). Moreover, some VQ proteins were reported containing calmodulin (CaM) binding domain, such as AtVQ15 and AtVQ22 (Perruc et al., 2004; Yan et al., 2018), or containing MAPK phosphorylation sites (Pecher et al., 2014). We predicted conservative motifs using MEME software. The top 20 conserved motifs were listed in Figure 4-S1, and these motifs may be related to

protein localization and protein interactions. We searched the database for these motifs using the Tomtom program within MEME software and did not find any clear functional annotations. Given that most VQ proteins were predicted to be localized in the nucleus, we speculated that there should be motifs associated with nuclear localization. It is known that proteins such as JAV1 and OsVQ25 are located in the nucleus (Yan et al., 2018; Hao et al., 2022). Therefore, we used NLStradamus and PSORT online tools to predict their nuclear localization signals. We found that Motif 4 is directly related to the nuclear localization of these proteins (Figures 4-S1 and 4-S2). Interestingly, we also found a significant overlap between the CaM binding domain (Perruc et al., 2004; Yan et al., 2018) and Motif 4, suggesting that Motif 4 may both guide nuclear localization and participate in interactions with CaM in these proteins.

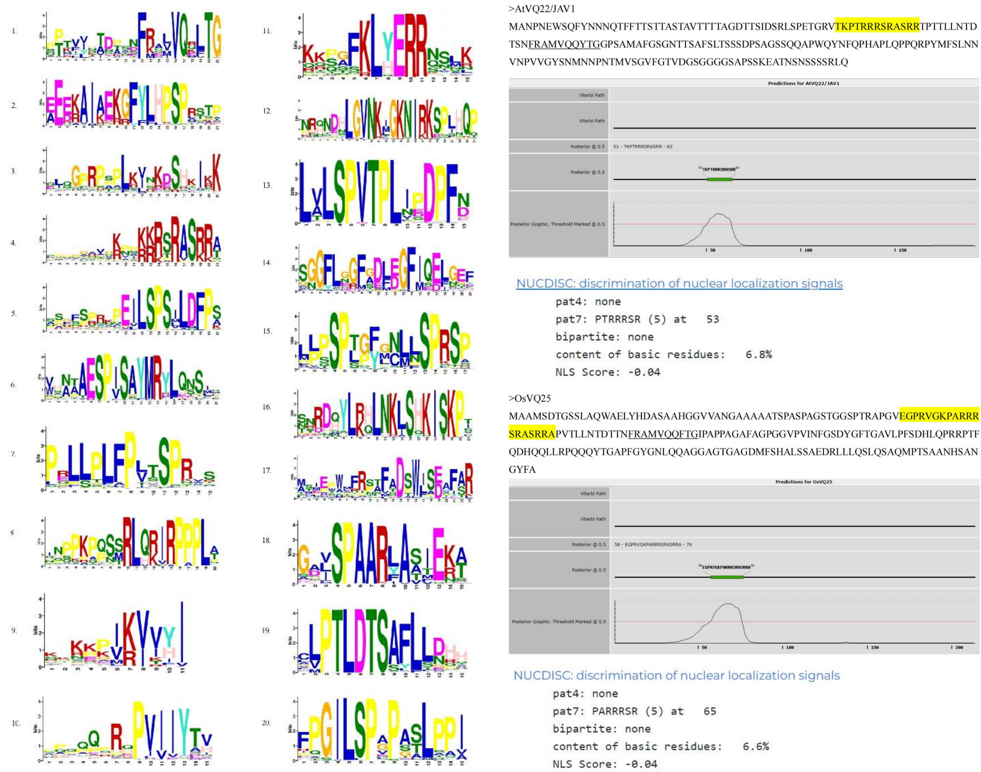


Figure 4-S1. Top 20 conserved motifs of VQ protein in 56 plant species
Figure 4-S2. Nuclear localization signal prediction

IV.3.2. Sequence similarities, length, introns, and GC content analysis of VQ genes in multiple plants

After performing multiple sequence alignment using MAFFT software on 2469 VQ proteins, all VQ proteins have an equal alignment length, and the proportion of identical amino acids at each aligned site was calculated. The highest ratio is the maximum sequence similarity of this matching point. This can help us clearly see the overall sequence similarity of *VQ* family genes in plants. We found that the proportion of alignment sites with maximum sequence similarity ≤ 20% is 99.2%, and except the VQ motif, the amino acid sequences at other positions are very variable (Figure 4-1c). Overall, the similarity of VQ proteins is low, but there are also some amino acid sites showing slightly higher similarity. Therefore, for the certain *VQ* gene, it is still possible to find relatively homologous genes in different species according to their sequence similarity.

We conducted the basic statistics on these 2469 *VQ* genes. Most *VQ* genes encode relatively small protein, with 83% and 93% of them less than 300 amino acids and than 400 amino acids respectively (Figure 4-1d), almost consistent with previous studies (Yuan et al., 2021; Jiang et al., 2018). In species with multiple *VQ* genes, moss (*Physcomitrella patens*) keeps high proportion (more than 70%) of intron-containing *VQ* genes, while most *VQ* genes in higher plants do not have introns, no matter angiosperms or gymnosperms (Figure 4-2). It has previously been reported that most *VQ* genes in moss (*Physcomitrella patens*) have introns (Ding et al., 2019), but interestingly, we inadvertently found that in different tea varieties (*Camellia sinensis*), such as 'Tieguanyin' (used in this study), 'Longjing 43′, 'Shuchazao', 'Yunkang 10′, and 'Biyun', the ratio of intron-free *VQ* genes are mostly high, except for 'Longjing 43′, whose proportion is actually as low as 15% (Table 4-S3).

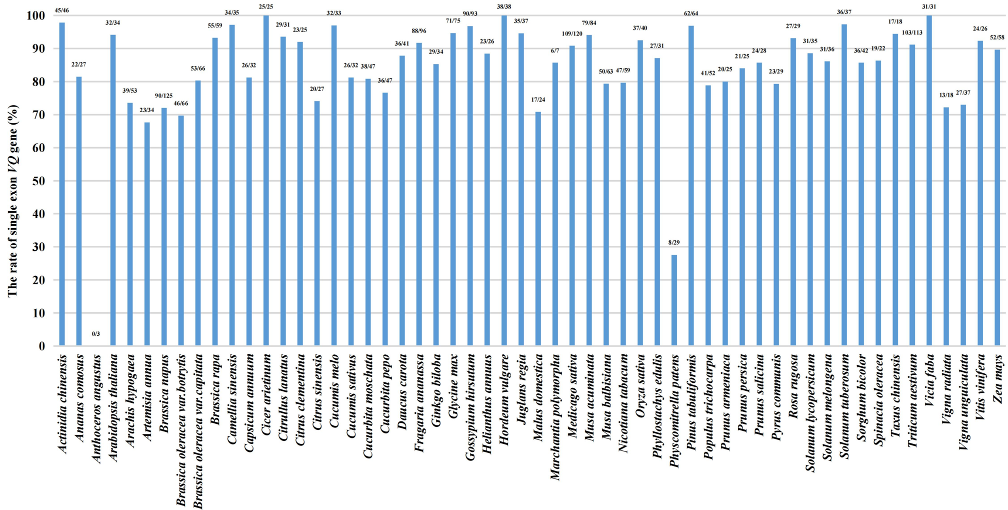


Figure 4-2. The ratio of single exon (intron-free) *VQ* genes. The labels above the column mean 'numbers of single exon *VQ* genes/numbers of all *VQ* genes'.

Furthermore, we calculated GC content in the coding region of *VQ* gene. Interestingly, we found that the average GC content in coding sequence of *VQ* genes among all species is greater than 40%; in commelinids, including Poales and Musa plants, GC content is more than 60%; especially, in Poaceae plants such as bamboo, sorghum, rice, maize, barley, and wheat, their GC content is up to 70%, which is much higher than the average level of species (55%), making it a relatively special occurence in plants (Figure 4-3). This suggests that *VQ* genes are more stable in Poaceae plants and may play more prominent roles in evolution and function.

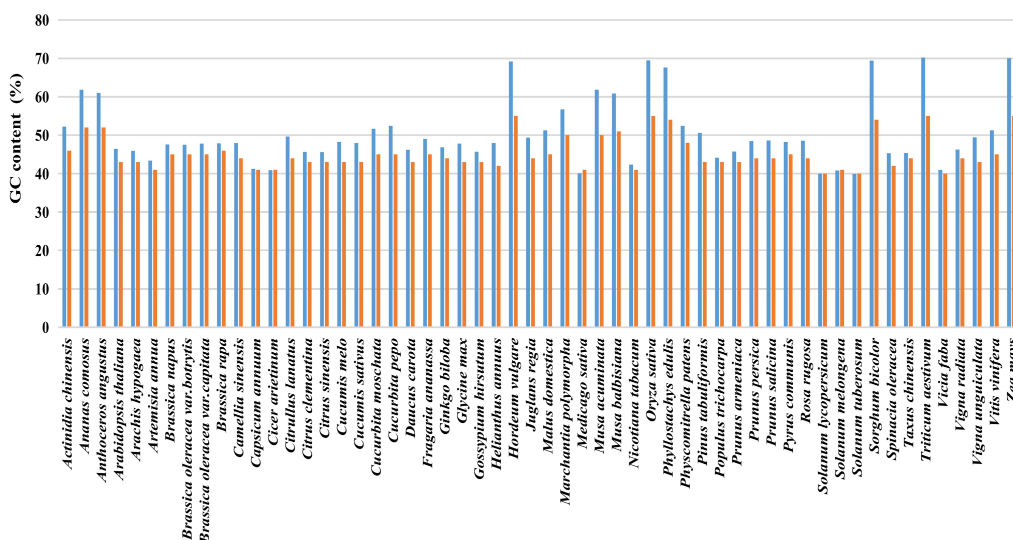


Figure 4-3. The average GC content of genes in coding sequence. Blue column is just for *VQ* genes, and the red column is for all coding genes in each species.

IV.3.3. Molecular features and gene structure analysis of VQ genes in gymnosperms

In total, 34, 18, and 64 *VQ* genes were identified respectively in *Ginkgo biloba*, *Taxus chinensis*, and *Pinus tabuliformis*, designated *GbVQ1* to *GbVQ34*, *TcVQ1* to *TcVQ18*, and *PtVQ1* to *PtVQ64* based on their chromosomal physical location (Table 4-1 and Table 4-S4). In order to clearly see the distribution of *VQ* genes on chromosomes, MapChart software was used to produce the chromosome map on chromosomes according to detailed location information of *VQ* genes in GFF file. These *VQ* genes are not evenly distributed across different chromosomes, and on

certain chromosomes, there is no *VQ* gene distribution (Figure 4-4). Among them, chromosome 3 and chromosome 10 have the largest number of *VQ* genes in *Ginkgo biloba* (n = 9, n = 9), chromosome 1 with the largest number of *VQ* genes in *Taxus chinensis* (n = 5), and chromosome 5 and chromosome 8 with the largest number of *VQ* genes in *Pinus tabuliformis* (n = 14, n = 18). In addition, we noticed that there are multiple *VQ* gene stacking within specific chromosomal regions, which may be due to the huge repetitive sequence or related to genome replication events of gymnosperms.

Table 4-1. Summary of molecular characterization and subcellular localization of VQ proteins in three gymnosperms.

Species	Number	Motif Type						Length (aa)	MW (kDa)	pI	Subcellular localization	
		LTG	FTG	VTG	MTG	YTG	FTA				Nucleus	Cytoplasm l Nucleus
Ginkgo biloba	34	23	8	2	-	1	-	283 ± 98	30.86 ± 10.25	7.67 ± 1.73	26	8
Taxus chinensis	18	13	3	1	-	-	1	165 ± 43	18.37 ± 4.81	8.55 ± 1.73	17	1
Pinus tabuliformis	64	36	22	2	4	-	-	260 ± 88	28.09 ± 9.10	8.50 ± 1.60	49	15

Protein length, MW, and pI: mean ± SD.

In gymnosperms, we analyzed the basic molecular features of *VQ* genes, the detailed data for each *VQ* genes, including gene ID, chromosome position, protein length, motif type, isoelectric point, molecular weight, and subcellular localization, were shown in Table 4-S4. The length of their encoded VQ protein ranges from 126 amino acids (GbVQ19) to 519 amino acids (GbVQ32) with average 283 amino acids in *Ginkgo biloba*, ranges from 110 amino acids (TcVQ12) to 243 amino acids (TcVQ9) with average 165 amino acids in *Taxus chinensis*, and ranges from 121 amino acids (PtVQ11) to 709 amino acids (PtVQ13) with average 260 amino acids in *Pinus tabuliformis*. The analysis of physiochemical properties further revealed that VQ proteins are widely varied in molecular weight (MW), ranging from 14.43 (GbVQ19) to 54.51 kDa (GbVQ32) with average 30.86 kDa in *Ginkgo biloba*, ranging from 12.26 (TcVQ12) to 54.51 kDa (TcVQ9) with average 18.37 kDa in *Taxus chinensis*, and ranging from 13.08 (PtVQ11) to 76.10 kDa (PtVQ13) with average 28.09 kDa in *Pinus tabuliformis*. The isoelectric point (pI) of these *VQ* genes varies between 4.83 (GbVQ21) and 10.43 (GbVQ3) in *Ginkgo biloba* (average 7.67), between 5.06 (TcVQ5) and 10.27 (TcVQ15) in *Taxus chinensis* (average 8.55), and between 5.07 (PtVQ24) and 11.03 (PtVQ31) in *Pinus tabuliformis* (average 8.50) (Table 4-1 and Table 4-S4).

In addition, we identified four VQ motif types (LTG, FTG, VTG, and YTG) in *Ginkgo biloba*, four VQ motif types (LTG, FTG, VTG, and FTA) in *Taxus chinensis*, and four VQ motif types (LTG, FTG, MTG, and VTG) in *Pinus tabuliformis*. Most VQ proteins belong to the LTG type (23/34), with eight FTG

type (8/34), two VTG type (2/34), and one YTG type (1/34) in *Ginkgo biloba*; most LTG type (13/18), three FTG type (3/18), one VTG type (1/18), and one FTA type (1/18) in *Taxus chinensis*; most LTG type (36/64), twenty-two FTG type (22/64), four MTG type (4/64), and two VTG type (2/64) in *Pinus tabuliformis* (Table 4-1 and Table 4-S4). We also observed that in *Pinus tabuliformis*, these are slightly changes in the core VQ motif from FxxxVQxhTG to FxxxVExhTG (PtVQ37, PtVQ38, PtVQ39) and FxxxVHxhTG (PtVQ28) (Table 4-S1). The subcellular localization analysis revealed that most *VQ* genes were predicted as nuclear-localized proteins, some were predicted as cytoplasm/nucleus-localized proteins (Table 4-1 and Table 4-S4).

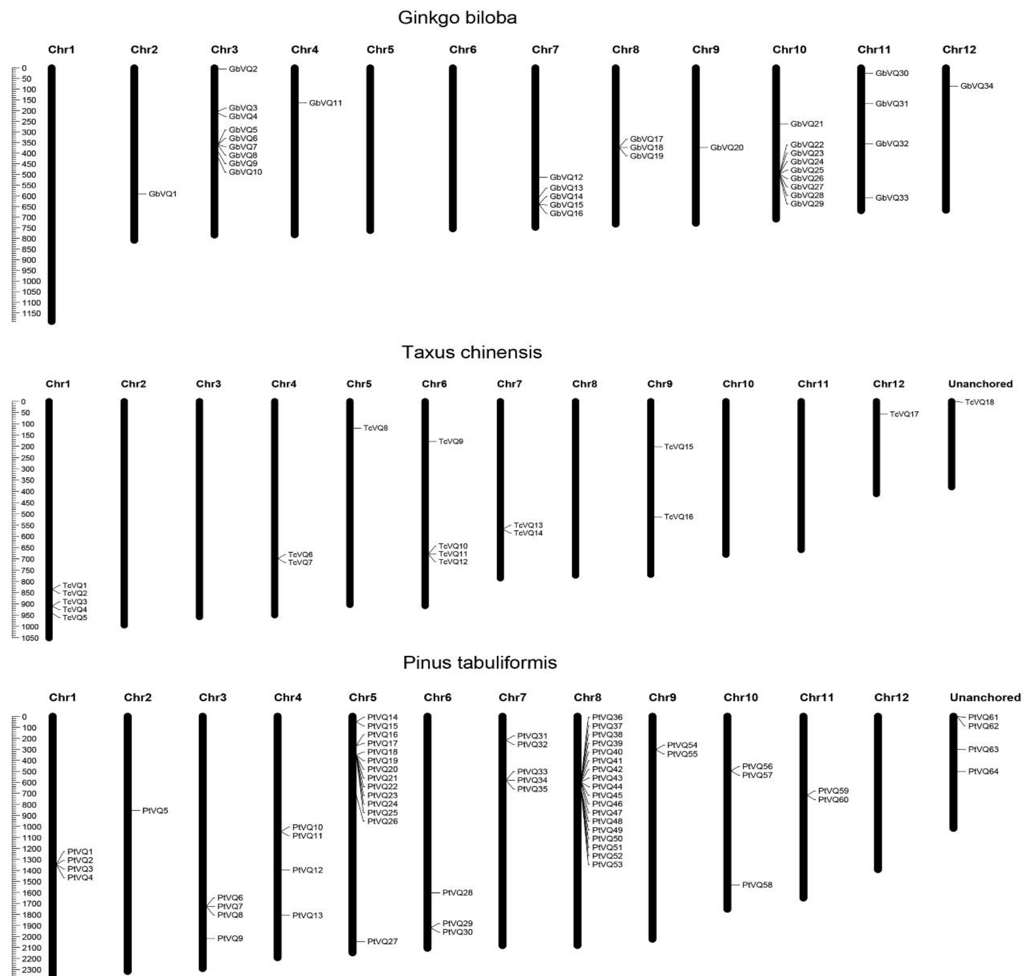


Figure 4-4. Chromosomal distribution of *VQ* genes in three gymnosperms. Chromosome numbers are listed on the top. The length of the chromosome is displayed in megabase (Mb) scale.

Gene structure can provide more information about the evolutionary relationship in a gene family. We conducted a systematic clustering and gene structure analysis on the *VQ* genes in *Ginkgo biloba*, *Taxus chinensis*, and *Pinus tabuliformis*. We detected 41 conservative motifs using MEME software, and we can clearly see the differences between these *VQ* genes. Motif 1 was identified as the core motif that composes the VQ domain, which was included in all VQ proteins. It is noteworthy that *VQ* genes with closer clustering relationships have almost similar conserved motifs, indicating that the phylogenetic classification is relatively reliable and proteins in the same group maybe perform similar functions (Figure 4-5a,b). We created an exon/intron structure map based on the location information of each *VQ* gene in the GFF file. The structure of *VQ* genes was analyzed and we found 85%, 94%, and 96% of *VQ* genes have no introns respectively in *Ginkgo biloba*, *Taxus chinensis*, and *Pinus tabuliformis* (Figure 4-5c). These indicates that most *VQ* genes do not contain introns in gymnosperms, which is similar to angiosperms (Yuan et al., 2021).

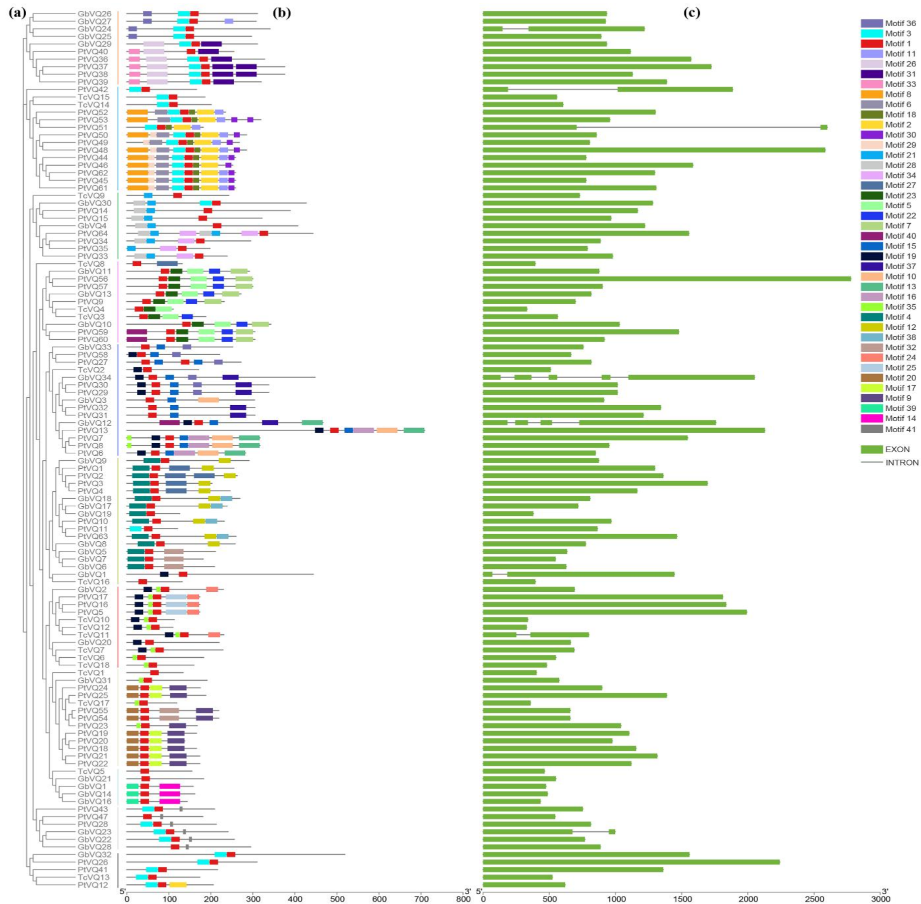


Figure 4-5. Phylogenetic tree, conserved motifs, and gene structure of VQ protein in three gymnosperms. (**a**) phylogenetic tree of 116 VQ proteins in three gymnosperms based on the results of sequence alignment. (**b**) 41 conserved motifs of the VQ protein. Each specific motif is indicated by a different colored box. (**c**) Gene structure (exon-intron) of the *VQ* genes. Exons are indicated by green rectangles, and lines connecting two exons represent introns.

IV.3.4. Gene duplication and collinearity analysis of VQ genes in gymnosperms

Gene duplication is an important event contributing to genome evolution, and also an important factor in the expansion of gene family. It is primarily divided into tandem duplication, segmental duplication, dispersed duplication, and proximal duplication (Wang et al., 2012). To better understand the evolutionary mechanism of *VQ* genes in gymnosperms, the duplication events of the *VQ* family gene were evaluated. Following the BLAST and MCScanX results, among the 34 *GbVQs*, a total of 10 members and 8 gene pairs participated in duplication events in *Ginkgo biloba*; among the 18 *TcVQs*, a total of 3 members and 2 gene pairs participated in duplication events in *Taxus chinensis*; and among the 64 *PtVQs*, a total of 18 members and 33 gene pairs participated in duplication events in *Pinus tabuliformis* (Table 4-2). *Ginkgo biloba* exhibited 5 tandem duplication events (5/8), 2 proximal duplication events (2/8), and 1 segmental duplication events (1/8); *Taxus chinensis* exhibited 2 proximal duplication events (2/2); and *Pinus tabuliformis* exhibited 14 dispersed duplication events (14/33), ten tandem duplication events (10/33), and 9 proximal duplication events (9/33). In these three gymnosperms, there is a interesting phenomenon: although the distance between two or more *VQ* gene pairs is a little bigger than 200 kb (here marked as proximal duplication), they are continuously arranged coding genes in one genome block, therefore, this type of gene pair may also be named as tandem duplication (Figure 4-4, Table 4-2). These results suggest that tandem duplication is important to expand the *VQ* family gene in gymnosperms. While whole-genome duplication (WGD) played a critical role in adaptive evolution in angiosperms (Jiao et al., 2011), but few recent WGD events were found in extant gymnosperms (Liu et al., 2021; Xiong et al., 2021; Liu et al., 2022; Niu et al., 2022; Fu et al., 2023; Li et al., 2015), which indicates that these duplicate *VQ* gene pairs were evolved from independent duplication events or derived from older ancestors.

The selective evolutionary pressure on all *VQ* gene pairs was investigated by calculating the Ka, Ks, and Ka/Ks ratios of the duplication events. The Ka/Ks values of most duplicated gene pairs (5/8) are less than 1.0, 2 gene pairs (2/8) are slightly greater than 1.0, and 1 gene pair (3/8) (segmental duplication events) is unable to calculate a valid value using KaKs_Calculator 2.0 software in *Ginkgo biloba*; 1 duplicated gene pair (1/2) is slightly greater 1.0, and 1 gene pair (1/2) is unable to calculate a valid value in *Taxus chinensis*; most duplicated gene pairs (21/33) are less than 1.0, 4 gene pairs (4/33) are more than 1.0, 1 gene pair (1/33) is unable to calculate a valid value, and 7 duplicated gene pairs (7/33) with Ka = Ks =

0, which means these two genes have no difference in their coding region in *Pinus tabuliformis* (Table 4-2). The Ka/Ks values of the majority of *VQ* gene pairs in gymnosperms are less than 1.0, indicating they have mainly undergone purifying selection pressures during evolution process. The Ka/Ks values of some duplicated gene pairs are greater than 1.0, which shows the presence of positive selection pressure.

Table 4-2. Duplication events of *VQ* genes in three gymnosperms.

Species	Gene 1	Gene 2	Duplication Type	Ka	Ks	Ka/Ks
Ginkgo biloba	GbVQ5	GbVQ6	Tandem	0.103073	0.129092	0.798447
	GbVQ5	GbVQ7	Droximal	0.095308	0.094741	1.005980
	GbVQ6	GbVQ7	Droximal	0.057674	0.059349	0.971771
	GbVQ14	GbVQ15	Tandem	0.044280	0.052344	0.845953
	GbVQ14	GbVQ16	Tandem	0.045420	0.033470	1.357020
	GbVQ15	GbVQ16	Tandem	0.062036	0.067561	0.918212
	GbVQ26	GbVQ27	Tandem	0.040319	0.041262	0.977136
	GbVQ2	GbVQ20	Segmental	NA	NA	NA
Taxus chinensis	TcVQ10	TcVQ12	Droximal	NA	NA	NA
	TcVQ11	TcVQ12	Droximal	0.098532	0.094428	1.043470
Pinus tabuliformis	PtVQ3	PtVQ4	Droximal	0.054239	0.179528	0.302121
	PtVQ5	PtVQ17	Dispersed	0.035409	0.034928	1.013750
	PtVQ6	PtVQ13	Dispersed	NA	NA	NA
	PtVQ6	PtVQ7	Droximal	0.004765	0.028118	0.169469
	PtVQ7	PtVQ13	Dispersed	0.142177	0.190426	0.746623
	PtVQ7	PtVQ8	Tandem	0	0	NA
	PtVQ8	PtVQ6	Droximal	0.004765	0.028118	0.169469
	PtVQ8	PtVQ13	Dispersed	0.179357	0.193909	0.924953
	PtVQ16	PtVQ5	Dispersed	0.009938	0.008609	1.154340
	PtVQ16	PtVQ17	Droximal	0.025090	0.044106	0.568853
	PtVQ18	PtVQ19	Droximal	0.032103	0.017421	1.842760
	PtVQ18	PtVQ20	Dispersed	0.035765	0.065227	0.548313
	PtVQ19	PtVQ20	Tandem	0.045874	0.053754	0.853421
	PtVQ21	PtVQ22	Tandem	0.098428	0.225648	0.436201
	PtVQ29	PtVQ30	Droximal	0.001346	0	NA
	PtVQ31	PtVQ32	Dispersed	0.001485	0.025303	0.058707
	PtVQ37	PtVQ38	Tandem	0	0	NA
	PtVQ44	PtVQ61	Dispersed	0.010209	0.005417	1.884470
	PtVQ44	PtVQ62	Dispersed	0.028367	0.030457	0.931388
	PtVQ44	PtVQ45	Tandem	0.025740	0.068154	0.377675
	PtVQ44	PtVQ46	Droximal	0.045012	0.063465	0.709246
	PtVQ45	PtVQ46	Tandem	0.059890	0.101127	0.592221
	PtVQ48	PtVQ49	Tandem	0	0	NA
	PtVQ48	PtVQ50	Droximal	0.004649	0.024470	0.189988
	PtVQ49	PtVQ50	Droximal	0.004947	0.026018	0.190129
	PtVQ54	PtVQ55	Dispersed	0	0	NA
	PtVQ56	PtVQ57	Tandem	0	0	NA

PtVQ59	PtVQ60	Tandem	0	0	NA
PtVQ61	PtVQ62	Tandem	0.031880	0.036130	0.882364
PtVQ61	PtVQ45	Dispersed	0.026613	0.071133	0.374138
PtVQ61	PtVQ46	Dispersed	0.048733	0.069510	0.701090
PtVQ62	PtVQ45	Dispersed	0.021352	0.065554	0.325719
PtVQ62	PtVQ46	Dispersed	0.058954	0.085332	0.690883

To gain insight into the evolution of *VQ* genes in gymnosperms, the collinear blocks were searched in the chromosomes using MCScanX software. *Ginkgo biloba* and *Pinus tabulaeformis* have the most collinear blocks at the whole genome level, followed by *Ginkgo biloba* and *Taxus chinensis*, *Taxus chinensis* and *Pinus tabuliformis.* This indicates that compared to *Taxus chinensis*, *Ginkgo biloba* has a higher similarity and closer evolutionary relationship with *Pinus tabuliformis.* Similarly, the *VQ* gene pairs in these collinear blocks are the most between *Ginkgo biloba* and *Pinus tabuliformis* (*n* = 12), followed by *Ginkgo biloba* and *Taxus chinensis* (*n* = 5), however, *Taxus chinensis* and *Pinus tabuliformis* do not have orthologous *VQ* gene pairs (*n* = 0) in fewer collinear blocks (Figure 4-6). Among all *VQ* gene pairs above (*n* = 17), only one *VQ* gene (*GbVQ29*, *TcVQ13*, and *PtVQ38*) shared by these three gymnosperms. In order to better see the conservatism of these *VQ* genes in different gymnosperms, we introduced the collinearity comparison with another two gymnosperms *Cycas panzhihuaensis* (Liu et al., 2022) and *Metasequoia glyptostroboides* (Fu et al., 2023). We found this *VQ* gene and its related collinear block is also existed in *Metasequoia glyptostroboides* and *Cycas panzhihuaensis*. Moreover, in all collinear blocks, we also found five other *VQ* genes (*GbVQ2*, *GbVQ13*, *GbVQ17*, *GbVQ20*, and *GbVQ34*), most of which contain collinear *VQ* genes in these five gymnosperms (Figure 4-6). This indicates that the genome blocks containing these six *VQ* genes, may have been preserved from ancestors of these gymnosperms. Additionally, through intragenomic collinearity analysis, we found that there are relatively more collinear blocks in *Ginkgo biloba*, including one *VQ* gene pair, which belongs to the segmental duplication events. Meanwhile, there are very few collinear blocks in other two gymnosperms, and no *VQ* gene pairs were found (Figure 4-S3, Table 4-2). All these results suggest that most *VQ* genes and its duplication events in every gymnosperm mentioned in this study may have evolved independently in the later stages. Furthermore, almost no collinear blocks and *VQ* gene pairs were observed between these three gymnosperms and *Arabidopsis* or rice.

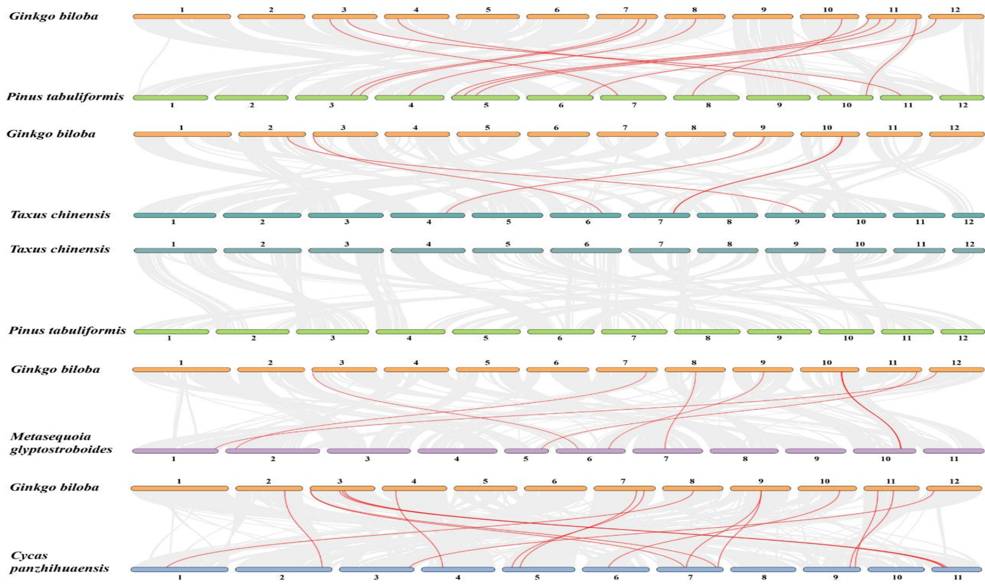


Figure 4-6. Collinear analysis between different gymnosperms. The grey lines in the background represent the collinear regions between different plant genomes, while the red lines highlight the collinear *VQ* gene pairs. The numbers represents chromosomes.

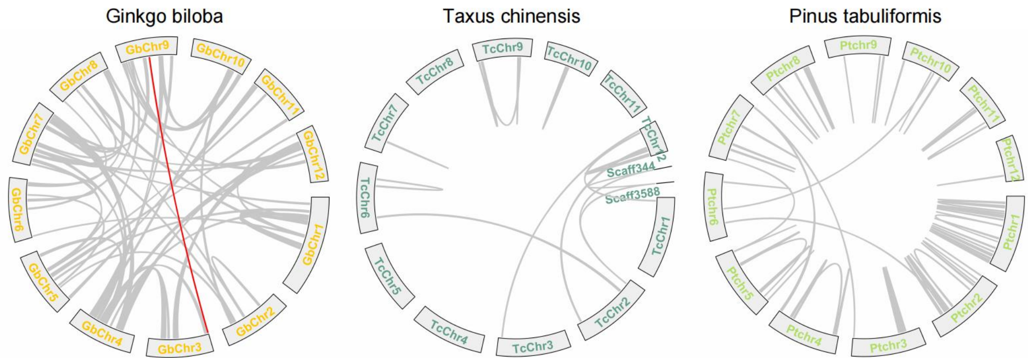


Figure 4-S3. Collinearity analysis in gymnosperms.

IV.3.5. Expression patterns of VQ genes in different tissues of gymnosperms

To characterize the expression patterns of *VQ* genes, the expression levels were analyzed in different tissues (8 for *Ginkgo biloba*, 4 for *Taxus chinensis*, and 4 for *Pinus tabulaeformis*) based on the published data in previous research to gain a preliminary insight into their potential functions (Figure 4-7, Table 4-S5). Based on the TPM values, we found 12 *VQ* genes (*GbVQ3*, *GbVQ10*, *GbVQ11*, *GbVQ12*, *GbVQ13*, *GbVQ26*, *GbVQ27*, *GbVQ29*, *GbVQ30*, *GbVQ32*, *GbVQ33*, and

GbVQ34), 6 *VQ* genes (*TcVQ2*, *TcVQ3*, *TcVQ4*, *TcVQ9*, *TcVQ13*, and *TcVQ14*), and 15 *VQ* genes (*PtVQ1*, *PtVQ9, PtVQ14, PtVQ15, PtVQ29*, *PtVQ30*, *PtVQ34, PtVQ37*, *PtVQ38*, *PtVQ56*, *PtVQ57*, *PtVQ58*, *PtVQ59, PtVQ60*, and *PtVQ64*) were broadly and prominently expressed in different tissues respectively in *Ginkgo biloba*, *Taxus chinensis* and *Pinus tabulaeformis* (Figure 4-7). By contrast, some *VQ* genes were expressed only in a few tissues, and some even with low or zero expression levels in different tissues (Figure 4-7, Tables 4-S6–S8). These *VQ* genes are suggested to play different roles in regulating plant growth and development of gymnosperms. These *VQ* genes expressed in multiple tissues are more easily focused and will receive further attention in this study. They may play a certain role in plant development and defense response (Zhang et al., 2022; Pedrosa et al., 2015; Eulgem et al., 2000). Of course, in addition to these *VQ* genes, many *VQ* genes may also be induced to express and play a certain role under certain conditions, and further research can be conducted in the future.

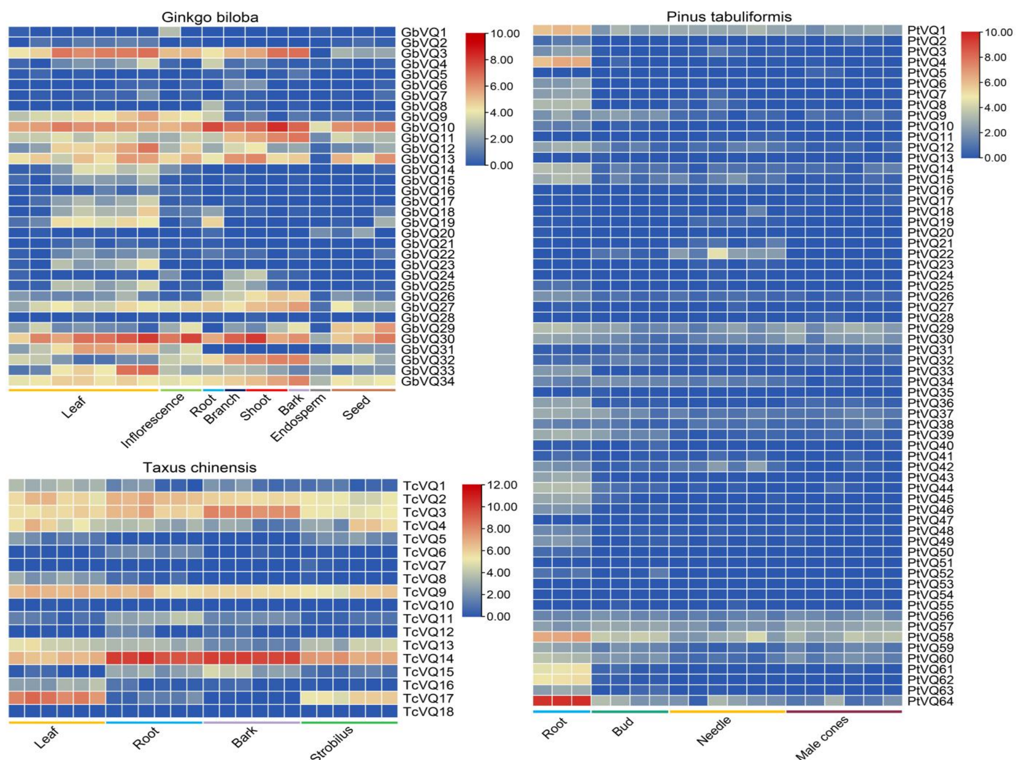


Figure 4-7. Expression level of *VQ* genes in different tissues of three gymnosperms. Expression levels of these *VQ* genes were obtained using RNA-seq data and measured with TPM values. Heatmap was generated using log2 (TPM + 1) values.

IV.3.6. Phylogenetic analysis of VQ genes

To detect the evolutionary relationships and classification of the *VQ* family genes in 56 plant species, circular and unrooted phylogenetic trees were constructed with 2469 VQ proteins. In previous studies, according to structural characteristics, VQ proteins from different species have been clustered into 7 groups (Kim et al., 2013), 8 groups (Zhang et al., 2022), 9 groups (Dong et al., 2018), and 10 groups (Pecher et al., 2014). In this study, we also built the phylogenetic tree for these 2469 VQ proteins from 56 plant species to explore their phylogenetic relationship, in which these proteins could be divided into 9 groups (Figures 4-8 and Figure 4-S4). *VQ* genes with unknown functions and pathways can often be inferred from species with close genetic relationships. Therefore, we have also constructed a species evolution tree based the taxonomy database of NCBI (Figure 4-S5). Systemic clustering and biological classification of these 56 plant species can help us to better understand the function and evolution of *VQ* genes in different plants. Moreover, *VQ* genes were clustered into 11 groups in three gymnosperms, which is basically consistent with the results of conservative motifs (Figure 4-5).

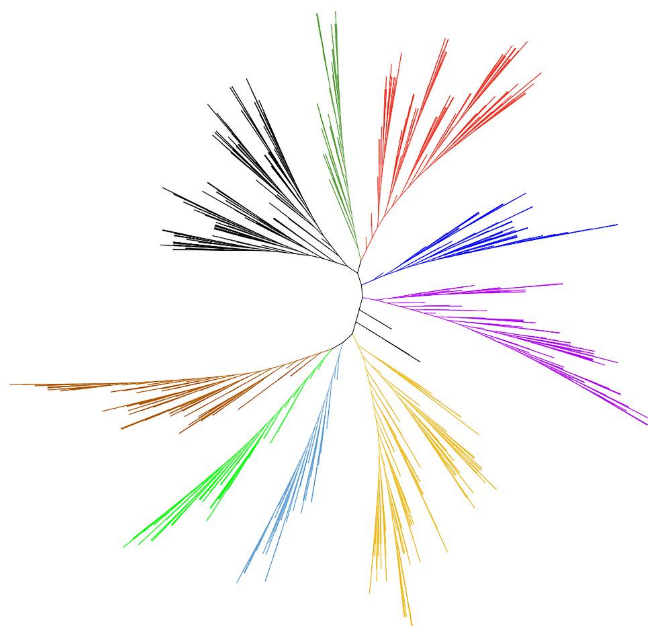


Figure 4-8. Unrooted phylogenetic tree of VQ proteins in 56 plant species. Multiple sequences alignment was performed with MAFFT software, and the phylogenetic tree was constructed by FastTree. Different colors indicate different groups, and these 2469 VQ proteins were clustered into nine subgroups.

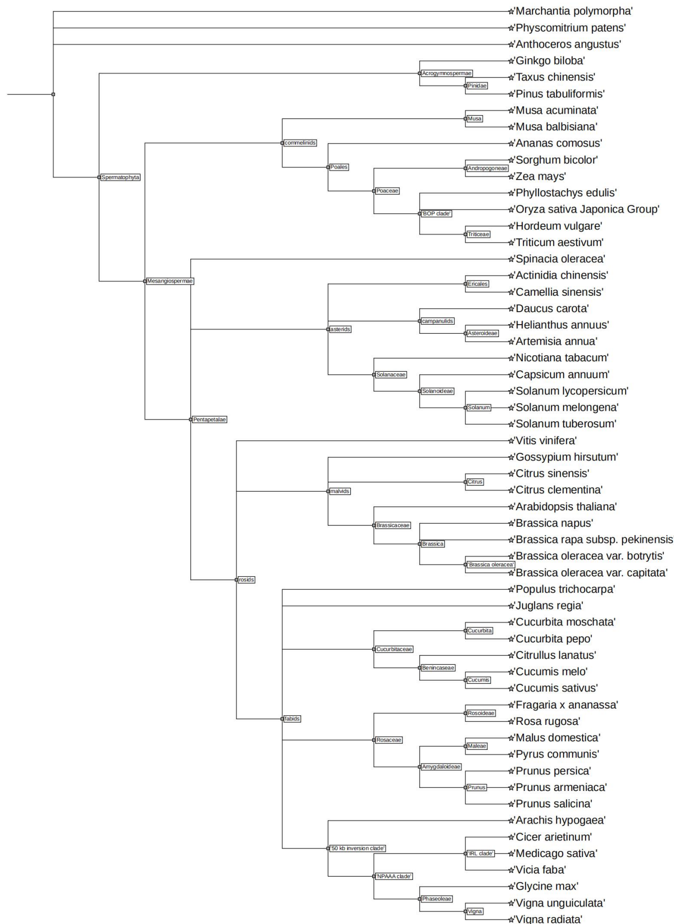


Figure 4-S5. Taxonomy classification for 56 plant species.

To further explore the evolutionary relationship of *VQ* gene between gymnosperms and angiosperms, and between monocotyledons and dicotyledons, nine species including three gymnosperms (ginkgo, taxus, Chinese pine) and six angiosperms (monocot: rice, wheat, maize; dicotyledon: *Arabidopsis*, soybean, tomato) were selected for further evolutionary analysis. Phylogenetic tree of VQ proteins from these species was drawn and it showed that some branches are unique to gymnosperms, some are unique to angiosperms, and some have intersections, indicating that some VQ proteins may appear after the differentiation of angiosperms and gymnosperms (Figure 4-S6). Similarly, monocotyledonous and dicotyledonous plants also have their own independent and intersecting branches (Figure 4-S6), which is consistent with previous research (Zhang et al., 2022). Interestingly, among widely expressed *VQ* genes, we found four genes (*GbVQ3*,

GbVQ12, GbVQ33, and *GbVQ34*) in *Ginkgo biloba*, one gene (*TcVQ2*) in *Taxus chinensis*, and three genes (*PtVQ29*, *PtVQ30*, and *PtVQ58*) in *Pinus tabulaeformis* were clustered with *AtVQ14, AtVQ9*, and *AtVQ5* (Figure 4-S6), which was reported influencing seed development, mediating salinity stress responses, and regulating plant defense, respectively (Cheng et al., 2012; Wang et al., 2010; Hu et al., 2013); three genes (*GbVQ26, GbVQ27,* and *GbVQ29*) in *Ginkgo biloba*, three genes (*TcVQ13*, *TcVQ14*, and *TcVQ15*) in *Taxus chinensis*, and three genes (*PtVQ37*, *PtVQ38*, and *PtVQ39*) in *Pinus tabulaeformis* were clustered with *AtVQ22*, *AtVQ27*, and *AtVQ28* (Figure 4-S6), which are related to plant defense and growth (Yan et al., 2018; Lan et al., 2022; Kumari et al., 2023). Among them, *GbVQ29*, *TcVQ13* and *PtVQ38* are the only *VQ* gene in the collinear block shared by these gymnosperms, which indicates that these genes have a very conservative, important role and status in seed plants. Additionally, *GbVQ10*, *TcVQ3*, *PtVQ9*, *PtVQ59*, and *PtVQ60* were clustered with *AtVQ4*, which regulates disease resistance (Pecher et al., 2014); *GbVQ30*, *TcVQ9*, *PtVQ14*, *PtVQ15*, *PtVQ34*, and *PtVQ64* were clustered with *AtVQ15*, which negatively regulates osmotic stress tolerance (Cheng et al., 2012; Perruc et al., 2004) (Figure 4-S6). These clusters can serve as references for the potential functions of these *VQ* genes in some degree.

The different characteristics of *VQ* genes in different types of plants can provide a molecular perspective for us to better understand plant evolution. All these results provide important references for the research and utilization of *VQ* genes in plants, as well as for the evolutionary comparisons between gymnosperms and angiosperms.

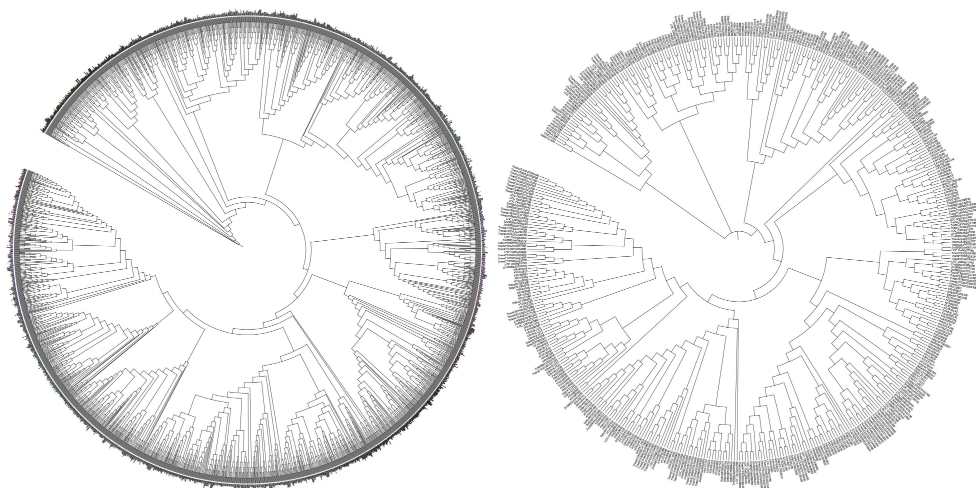


Figure 4-S4. Circular phylogenetic tree of VQ proteins in 56 plant species.
Figure 4-S6. Circular phylogenetic tree of VQ proteins in three gymnosperms and six angiosperms.

IV.4. Discussion

Plants are often affected by various environmental stresses throughout their lifetime (Suzuki et al., 2014). *VQ* family genes have been proven to play important roles in growth and development, and the responses to various abiotic and biotic stresses (Yuan et al., 2021; Cheng et al., 2012). Therefore, it is of great significance to study the characteristics and functions of *VQ* genes in a wide range of plants. Up to now, *VQ* family genes have been identified and analyzed in multiple plant species. However, although there have been some reports on the structure and function of *VQ* genes, clear and intuitive feature information and how to utilize these features are still limited in the plant kingdom, including bryophytes, gymnosperms, and angiosperms. Importantly, the *VQ* gene family of gymnosperms has not been reported yet. Thus, a comprehensive bioinformatics analysis of *VQ* genes in multiple plants, especially in gymnosperms, can provide a overall basis for evolutionary and functional studies on *VQ* genes.

4.1. Molecular characteristics and phylogeny of VQ family genes

In this study, we systematically identified 2469 *VQ* genes from 56 plant species. Gene numbers of VQ family are various among different plants. Compared to other plants, plants such as wheat and soybean contain abundant *VQ* genes (Table 4-S2), which have undergone significant polyploidization and WGD events, respectively (Zhang et al., 2022; Kim et al., 2015). *Ginkgo biloba*, *Taxus chinensis* and *Pinus tabulaeformis* were detected 34, 18, and 64 *VQ* gene members respectively, but compared with other species, such as *Arabidopsis* with 34 members (Cheng et al., 2012), or rice with 40 members (Li et al., 2014), the *VQ* gene numbers in these three gymnosperms are far lower than expected considering their big genome size. Therefore, it can be concluded that the number of *VQ* genes has no necessary connection with the genome size, but is related to species polyploidization and whole genome replication events.

Except for the conserved VQ domain, VQ protein exhibits significant variability in other regions (Figure 4-1c). However, through MEME domain scanning, we also found some other motifs in different subgroups (Figures 4-5 and Figure 4-S1), which may be related to the interaction, modification, and subcellular localization of these VQ proteins, affecting protein function. Although we have enriched multiple motifs in these VQ proteins, their specific functions still need further analysis. Many *VQ* family genes play important roles in growth and development, biotic and abiotic stress, and interact with various proteins such as MAPK and WRKY. By conducting targeted editing and modification of these enriched candidate loci on the *VQ* gene, the functions and mechanisms of these motifs can be further revealed. Moreover, given the extensive biological functions of the *VQ*

gene, this also has great potential in plant molecular precision breeding. Our study provides a method and idea for searching and utilizing conserved motifs and loci in gene families. The average length of VQ proteins is fewer than 300 amino acid residues (Figure 4-1d). This is one of the characteristics of the *VQ* family genes in plants, which helps us identify them. Noteworthy, some VQ proteins were found to be longer than 600 amino acids (Figure 4-1d), and these VQ proteins may contain other domains and participate in more diverse regulatory pathways.

Although the average GC content in the coding region of the *VQ* gene varies among different plants, the GC content in commelinids plants, here including Poales and Musa plants, has significantly increased and is most prominent in Poaceae (GC content > 70%) (Figure 4-3). The *VQ* gene, as an ancient transcription regulatory cofactor in plants, plays an important role in the interaction between plants and environment. High GC content means better alkaline and high-temperature tolerance, allowing these *VQ* genes to remain relatively stable in environmental changes. Poaceae plants are the main food source for humans and many animals, and high GC content of their *VQ* genes is likely possessed by their ancestors and has been selected and fixed during evolution, which has become one of the characteristics of these plants. As for the deeper reasons why the GC content of *VQ* genes in Poaceae plants is prominent, further research is needed, which is a very interesting and meaningful topic.

In the higher eukaryotes, intron-free genes are very common in their genomes (Mourier et al., 2003). In our study, based on the gene structure analysis, we found that most *VQ* genes in higher plants are intronless, no matter in angiosperms or gymnosperms, which is consistent with previous researches, including *Arabidopsis* (Cheng et al., 2012), rice (Kim et al., 2013), tomato (Ding et al., 2019), apple (Dong et al., 2018), and wheat (Zhang et al., 2022), and so on. In contrast, lower plant moss is exactly the opposite (Ding et al., 2019) (Figure 4-2). Therefore, many studies speculated that the *VQ* genes in higher plants lost its intron during evolution (Yuan et al., 2021; Zhang et al., 2022; Xu et al., 2022). However, we found that most *VQ* genes in tea variety ‘Longjing 43′ also contain introns, and at the same time, *Marchantia polymorpha*, known as a lower bryophyte plant, although there are only seven *VQ* genes in its genome, six of them (6/7) have no intron. So, the explanation of intron loss from the perspective of evolution history may need further investigation. Moreover, our results revealed that intron-containing *VQ* genes of gymnosperms are located in different subgroups, suggesting that these introns appear relatively independent, and this is also common in angiosperms (Zhang et al., 2022; Xu et al., 2022). Taken together, identification of gene structure of *VQ* genes enriched our understanding for the evolution of introns in the plant kingdom.

Phylogenetic trees represent the genetic relationships among gene families from different species and reflect the similarity of protein-coding genes. To further understand the evolutionary relationships among these 2469 *VQ* genes from 56

plant species, we established a phylogenetic tree based on their protein sequences. They were classified into 9 groups from our phylogenetic analysis (Figures 4-8 and Figure 4-S4) and showed obvious evolutionary characteristics, such as the differentiation between angiosperms and gymnosperms, monocotyledons and dicotyledons. Additionally, in three gymnosperms, the *VQ* genes were classified into 11 groups, and in each group, they harbored similar type of motifs (Figure 4-5), suggesting a potential functional similarity. These results highlight the conservatism and diversity among *VQ* gene family of different plants. Based on the phylogenetic tree results, we can also use *VQ* genes with known functions to quickly search for homologous genes in specific plants. In theory, gene structure determines its function, and the more similar motifs shared between VQ proteins, the higher probability of their functional similarity. Therefore, systematic clustering combined with motifs analysis will help us quickly identify homologous genes in different plants and make preliminary judgments on their functions. At the same time, further experiments are needed to verify their detailed functionality.

4.2. Expansion and duplication mechanism of VQ gene family in gymnosperms

Genome replication events play an important role in expanding the size of the genome (Adams et al., 2005) and diversifying gene functions (Rensing et al., 2014). Chromosome fragment replication is considered to be the main expansion mechanism of gene family (Kaltenegger et al., 2018), and thus, the evolutionary process can explain the number of specific *VQ* genes in a species, not the genome size. Previous research indicated that segmental duplication is the major mechanism contributing to the expansion of the *VQ* gene family in many angiosperms (Zhang et al., 2022; Jiang et al., 2018; Xu et al., 2022). In the present study, we found both tandem duplication and segmental duplication events of *VQ* genes exist simultaneously in *Ginkgo biloba*; only proximal duplication events exist in *Taxus chinensis*; dispersed duplication, tandem duplication, and proximal duplication events exist in *Pinus tabulaeformis*; but segmental duplication events only appear in *Ginkgo biloba* and account for a small proportion (Table 4-2). Segmental duplication is an important way of expanding the *VQ* gene family in angiosperms, while tandem duplication dominates the expansion of the *VQ* gene family in gymnosperms, which may be caused by different genetic and evolutionary mechanisms between angiosperms and gymnosperms. In addition, tandem duplication can reduce genetic instability caused by single gene mutations or deletions, enhance plant resistance to environmental stress, such as stress resistance and disease resistance. Given the importance of *VQ* genes, this is of great significance for ensuring the normal life of gymnosperms.

The Ka, Ks and Ka/Ks ratios of all *VQ* gene pairs were calculated to investigate whether *VQ* genes underwent selection pressure. The Ka/Ks values of most *VQ* gene pairs are less than 1.0, which demonstrates that purifying selection (Ka/Ks < 1)

play an important role in the evolution of *VQ* gene family in gymnosperms. Besides, positive selection (Ka/Ks > 1) is also existed in some *VQ* genes (Table 4-2). Most *VQ* gene pairs are undergoing purification selection, while a few are undergoing positive selection, indicating that the *VQ* gene family in gymnosperms is in a stable dynamic evolution process.

Furthermore, only a few *VQ* gene pairs in collinear blocks were detected within and between *Ginkgo biloba, Taxus chinensis*, and *Pinus tabulaeformis*. For *Taxus chinensis* or *Pinus tabulaeformis* themself, there were even no *VQ* gene pairs detected in very few collinear blocks (Figures 4-6 and Figure 4-S3). The results of synteny analysis indicate that the conservation degree of *VQ* genes among these three gymnosperms is low and their evolution process is relatively independent, which is similar to the difference between dicotyledons and monocotyledons in angiosperms (Xu et al., 2022), and this may be related to the fact that these three gymnosperms belong to different phytoclasses.

4.3. Expression patterns of VQ members in gymnosperms

Previous studies have demonstrated that *VQ* genes are involved in regulating plant responses to biotic stresses, abiotic stresses, and growth and development (Yuan et al., 2021; Cheng et al., 2012). The expression of many *VQ* genes showed various levels between different tissues and significant changes under pathogen, stress or hormone treatments (Zhang et al., 2022; Cheng et al., 2012; Kim et al., 2013). In this study, we detected the expression level of gymnosperm *VQ* genes in different tissues by analyzing RNA-seq data. The results showed that some *VQ* genes are not expressed in any tissues, some are only expressed in certain tissues, while some are widely expressed in different tissues. These widely expressed genes are usually associated with growth and development, hormone response, and plant defense (Zhang et al., 2022; Pedrosa et al., 2015; Eulgem et al., 2000), which requires further experimental verification in gymnosperms.

In *Arabidopsis*, *AtVQ14* is mainly associated with seed development, and its mutation produces small seeds (Luo et al., 2005; Wang et al., 2010). Transgenic plants overexpressing *AtVQ5* displayed increased susceptibility to *B. cinerea* (Cheng et al., 2012). *AtVQ9* is strongly induced by NaCl treatment and negatively regulates the resistance to NaCl stress (Hu et al., 2013). Compared with wild type plants, overexpression of *AtVQ17*, *AtVQ18*, or *AtVQ22* causes highly stunted growth of the transgenic plans (Cheng et al., 2012). AtVQ22/JAV1 can regulate JA-mediated plant defense and coordinate growth and defense (Yan et al., 2018). Among these widely expressed *VQ* genes in gymnosperms, we totally found eight *VQ* genes (*GbVQ3, GbVQ12, GbVQ33, GbVQ34, TcVQ2, PtVQ29, PtVQ30*, and *PtVQ58*) as the candidate homologs of *AtVQ14*/*AtVQ9*/*AtVQ5,* nine *VQ* genes (*GbVQ26, GbVQ27, GbVQ29*, *TcVQ13*, *TcVQ14*, *TcVQ15*, *PtVQ37*, *PtVQ38*, and *PtVQ39*) as candidate homologs of *AtVQ22*/*AtVQ27*/*AtVQ28*, five *VQ* genes (*GbVQ10*, *TcVQ3*, *PtVQ9*, *PtVQ59*, and *PtVQ60*) as candidate homologs of *AtVQ4*,

and six *VQ* genes (*GbVQ30*, *TcVQ9*, *PtVQ14*, *PtVQ15*, *PtVQ34*, and *PtVQ64*) as candidate homologs of *AtVQ15* (Figures 4-7 and Figure 4-S6). This suggests that these candidate *VQ* genes may have similar functions and play an important role in growth, development, and response to external environmental stimuli. Among these genes, we found that TcVQ14 is almost identical to TcJAV*3*, which was reported to be a VQ motif-containing protein and homologous to AtVQ22/JAV1. TcJAV3-TcWRKY26 complex can regulate the expression of the downstream paclitaxel biosynthesis gene *DBAT* and participate in JA-mediated plant defense (Chen et al., 2022). This indicates that our research results are trustworthy. In addition, gene duplication can produce gene function redundancy, and most of these repeated or collinear *VQ* genes showed almost the same expression patterns (Figure 4-7, Table 4-2). Taken together, our study suggested that some *VQ* genes are involved in growth and development, and participate in multiple life processes of gymnosperms.

5

Chapter V A VQ-motif-containing protein fine-tunes rice immunity and growth by a hierarchical regulatory mechanism

Reference: Hao, Z.†, Tian, J.†, Fang, H.†, Fang, L., Xu, X., He, F., Li, S., Xie, W., Du, Q., You, X., Wang, D., Chen, Q., Wang, R., Zuo, S., Yuan, M., Wang, G. L., Xia, L., and Ning, Y. (2022). A VQ-motif-containing protein fine-tunes rice immunity and growth by a hierarchical regulatory mechanism. Cell Rep. 40(7), 111235. https://doi.org/10.1016/j.celrep.2022.111235
Author Contributions:
Y.N. and L.X. conceived the project. Z.H., J.T., H.F., L.F., X.X., F.H., S.L., W.X., Q.D., X.Y., D.W., Q.C., S.Z., and R.W. carried out the experiments. Z.H, J.T, H.F. S.L., and M.Y. performed the transgenic plant generation and analysis. Z.H., J.T., and H.F. wrote the paper. Y.N., L.X., and G.L.W. analyzed the data and revised the manuscript.

Abstract

Rice blast and bacterial blight, caused by the fungus *Magnaporthe oryzae* and the bacterium *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), respectively, are devastating diseases affecting rice. Here, we report that a rice valine-glutamine (VQ) motif-containing protein, OsVQ25, balances broad-spectrum disease resistance and plant growth by interacting with a U-Box E3 ligase, OsPUB73, and a transcription factor, OsWRKY53. We show that OsPUB73 positively regulates rice resistance against *M. oryzae* and *Xoo* by interacting with and promoting OsVQ25 degradation via the 26S proteasome pathway. Knockout mutants of OsVQ25 exhibit enhanced resistance to both pathogens without a growth penalty. Furthermore, OsVQ25 interacts with and suppresses the transcriptional activity of OsWRKY53, a positive regulator of plant immunity. OsWRKY53 downstream defense-related genes and brassinosteroid signaling genes are upregulated in *osvq25* mutants. Our findings reveal a ubiquitin E3 ligase-VQ protein-transcription factor module that finetunes plant immunity and growth at the transcriptional and posttranslational levels.

Keywords: OsVQ25, OsPUB73, OsWRKY53, rice resistance against *M. oryzae* and *Xoo*, finetunes plant immunity and growth

V.1. Introduction

Plant diseases are caused by many different pathogens and can result in devastating yield losses in crop production (Nelson et al., 2018; Ning et al., 2017). Developing crop varieties with durable, broad-spectrum resistance (BSR) is the most economical and sustainable way to control diseases (Kou and Wang, 2010; Nelson et al., 2018). Among BSR types, species-non-specific (SNS) BSR confers resistance against two or more pathogens (Ke et al., 2017; Li et al., 2020). Because crops are often successively attacked by multiple pathogens during the growing season, SNS BSR provides better disease control than species-specific BSR (Ke et al., 2017). In rice, over 40 SNS BSR genes have been identified; they encode membrane-associated pattern recognition receptors, defense-signaling (DS) proteins, pathogenesis-related proteins, and susceptibility (S) proteins (Li et al., 2020). For example, the DS protein (IPA1) positively regulates SNS BSR to the

blight bacterium *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) and the rice blast fungus *Magnaporthe oryzae* (Liu et al., 2019; Wang et al., 2018). By contrast, the S gene Bsr-k1 encoding an RNA-binding protein negatively regulates SNS BSR to *Xoo* and *M. oryzae* (Zhou et al., 2018). However, the functions of these proteins in SNS BSR and their potential underlying mechanisms in plants remain elusive.

The ubiquitin-proteasome system (UPS) plays critical roles in plant-microbe interactions and in immune responses to pathogens (Ning et al., 2016; Zeng et al., 2006). In the UPS, E3 ligases (E3s) are key factors that ubiquitinate target proteins and promote the degradation of some ubiquitinated proteins (Sadanandom et al., 2012). Depending on their structures, E3s are assigned to three categories: RING or U-box ligases, HECT ligases, and Cullin-RING ligases (Vierstra, 2009). Functions for E3s in rice SNS BSR have been reported. For instance, the U-box E3 ligase SPL11 negatively regulates programmed cell death (PCD) and defense against rice pathogens by degrading the Rho GTPase-activating protein SPIN6 and the S-domain receptor-like kinase SDS2 (Fan et al., 2018; Liu et al., 2015). The spl11 mutant generates PCD symptoms in leaves and exhibits non-specific resistance to *Xoo* and *M. oryzae* (Zeng et al., 2004). The RING-type E3 ligase EBR1 negatively regulates PCD and defense against pathogens in rice by degrading OsBAG4, and the ebr1 mutant demonstrates enhanced resistance to *Xoo* and *M. oryzae* (You et al., 2016). In addition, the Cullin-RING E3 ligase OsCUL3a negatively regulates PCD and immunity by degrading OsNPR1 in rice (Liu et al., 2017). OsCUL3a loss-of-function mutation promotes H_2O_2 accumulation in rice and shows enhanced resistance to *Xoo* and *M. oryzae* (Liu et al., 2017). All these rice E3s negatively regulate SNS BSR, while rice E3s that positively regulate SNS BSR have not been documented.

Valine-glutamine (VQ) proteins are an ancient protein family with the conserved VQ-motif structure FxxhVQxhTG (Jiang et al., 2018). VQ proteins are involved in plant defense. In *Arabidopsis* (*Arabidopsis thaliana*), *MKS1* (also named *VQ21*) overexpression confers increased resistance to *Pseudomonas syringae* pv. tomato DC3000 but greater susceptibility to *Botrytis cinerea* with severe growth penalties (Andreasson et al., 2005; Fiil and Petersen, 2011; Petersen et al., 2010). Differently, *SIB1* (also named *VQ23*) overexpression increases resistance to *Pseudomonas syringae* and *Botrytis cinerea* compared with wild-type (WT) plants. Importantly, *SIB1*-overexpression plants also display growth retardation (Lai et al., 2011; Xie et al., 2010). In addition, plants with reduced *JAV1* (also named *VQ22*) expression are more resistant to necrotrophic pathogens and herbivorous insects without affecting growth and development in *Arabidopsis* (Hu et al., 2013). These studies indicate that VQ proteins have various functions in plant immunity. However, only a few VQ proteins confer SNS BSR without a growth penalty.

In terms of signal transduction, VQ proteins always interact with WRKY transcription factors (TFs), which then regulate downstream responses (Chi et al., 2013; Jing and Lin, 2015). For example, JAV1 interacts with WRKY51 to negatively regulate the expression of jasmonic acid (JA) biosynthesis genes (Yan et al., 2018). The rice genome encodes around 40 VQ-motif-containing proteins (Jiang et al., 2018). Among them, OsVQ13 positively regulates rice resistance to

Xoo and affects rice grain size (Uji et al., 2019), while OsVQ14 and OsVQ32 function redundantly to positively regulate rice resistance to *Xoo* (Li et al., 2021). However, whether these or other *VQ* genes are involved in SNS BSR and associated signal transduction cascades needs further investigation.

Here, we report a VQ motif-containing protein, OsVQ25, which shares the same branch as JAV1 on the phylogenetic tree although with low similarity (Zhang et al., 2022a). We found that OsVQ25 interacts with and is degraded by the U-box E3 ligase OsPUB73 via the 26S proteasome. Genetic analysis showed that OsPUB73 positively regulates SNS BSR against rice pathogens. OsVQ25 interacts with the positive immune regulator OsWRKY53 and suppresses its transcriptional activity to negatively regulate SNS BSR in rice. Importantly, we demonstrate that loss of OsVQ25 function in rice results in SNS BSR against *M. oryzae* and *Xoo* without a growth penalty. Our study reveals that the OsPUB73-OsVQ25-OsWRKY53 module balances SNS BSR and plant growth at the transcriptional and posttranslational levels under a hierarchical regulatory mechanism. We obtained a new rice material with improved broad-spectrum disease resistance due to the mutation of *OsVQ25*. Meanwhile, we found that these *osvq25* mutants do not contain transgenic components (vector sequences) and have breeding potential.

V.2. Materials and methods

V.2.1. Plants and materials

The rice (*Oryza sativa*) cultivars Nipponbare (NPB) and Zhonghua11 (ZH11) were used for disease evaluation in this study. Rice seeds were surface-sterilized by immersion in 75% (v/v) ethanol for 5 min, followed by immersion in 40% (v/v) sodium hypochlorite for 25 min. After washing 5 times with sterile water, the seeds were germinated on 1/2 Murashige Skoog (MS) medium for 1 week. The seedlings were then maintained in an incubator with a 12-h light/12-h dark photoperiod, a 28/26℃ light/dark temperature regime, and 65% relative humidity. After 7 days in the incubator, the seedlings were transferred to soil and maintained in a growth chamber at 26℃ with a 12-h light/12-h dark photoperiod and 70% relative humidity.

The CRISPR/Cas9 system and rice variety NPB were used to generate the *osvq25* and *ospub73* knockout mutants. The single guide RNA (sgRNA) sequence was designed to specifically target the genomic loci of *OsVQ25* and *OsPUB73*. The sgRNA expression cassette was then inserted into the pYLCRISPR/Cas9-MTmono binary vector (Ma et al., 2015). The resulting constructs were introduced into *Agrobacterium* (*Agrobacterium tumefaciens*) strain EHA105 for rice transformation. *oswrky53* mutant plants and ZH11 were described previously (Xie et al., 2021). All primer sequences used for the constructs are listed in Table 5-S1.

N. Benthamiana plants were cultivated in soil under a 12-h light/12-h dark photoperiod at 25℃. Five-week-old *N. benthamiana* leaves were used in luciferase complementation imaging and bimolecular fluorescence complementation assays.

Arabidopsis thaliana accession Columbia-0 (Col-0) were cultivated in soil at 25°C with a 16/8 h light/dark cycle. Fiveweek-old *Arabidopsis* leaves were used for the transcriptional activity assay.

V.2.2. Expression pattern analysis

Rice leaves were collected at different time points after spray inoculation with Magnaporthe oryzae isolates. Water with 0.05% (v/v) Tween 20 was used as a mock inoculation control (Mock) (Fang et al., 2021). Total RNA was isolated with Trizol reagent (Invitrogen) according to the manufacturer's instructions. First-strand cDNA was synthesized with reverse transcriptase (Promega) after digestion of total RNA with DNase I (TransGen). qRT-PCR was performed with 23SYBR Green Mix (GeneStar) on an ABI Prism 7500 PCR instrument. Gene expression levels were calculated with the data from three technical replicates. The primer sequences used for the qRT-PCR assay are listed in Table 5-S1.

V.2.3. Magnaporthe oryzae (M. oryzae) inoculation and disease symptom evaluation

M. oryzae isolates were cultivated on oatmeal agar plates under weak light for 14 days to generate spores. Six-week-old rice plants were used for punch inoculation (Fang et al., 2021) with a suspension of *M. oryzae* spores (about 53105 spores/mL). After a mouse ear clip was used to lightly punch rice leaves, the punched sites were treated with one drop (10 mL) of the spore suspension. The spores were held in place by sealing both sides of the treated sites with scotch tape. Two weeks after inoculation, the inoculated leaves were photographed. A 4-cm-long segment of rice leaf with lesion was then cut and subjected to DNA extraction with the cetyltrimethyl ammonium bromide (CTAB) protocol (Clarke, 2009). Relative fungal biomass was measured as previously described with DNA-based quantitative PCR (qPCR) using the threshold cycle value (Ct) of *M. oryzae* MoPot2 and rice genomic UBIQUITIN (OsUbq) according to the formula $2^{[CT(OsUbq)-CT(MoPot2)]}$ (Shi et al., 2018). qPCR was performed with 23SYBR Green Mix (GeneStar) on an ABI Prism 7500 PCR instrument.

V.2.4. Xanthomonas oryzae pv. oryzae (Xoo) inoculation and disease symptom evaluation

Xoo isolates were cultured on potato dextrose liquid medium (30 C, 200 rpm) until the optical density (OD600) of the culture was 1.0 (Liu et al., 2017); the resulting suspension was used to inoculate rice leaves. Leaves of 6-week-old rice plants were cut with a scissors that had been dipped into the bacterial suspension (Liu et al., 2017). Two weeks after inoculation, the inoculated leaves were photographed and lesion lengths were measured.

V.2.5. Luciferase complementation imaging (LCI) assay in N. benthamiana

The LCI assay was performed in *N. benthamiana* as previously described (Zhang

et al., 2020). The coding sequences of *OsPUB73* and *OsWRKY53* were cloned into the pCAMBIA-NLuc vector (OsPUB73-NLuc, OsWRKY53-NLuc), and the coding sequence of *OsVQ25* was cloned into the pCAMBIA-CLuc vector (CLuc-OsVQ25). Agrobacterium cultures (strain EHA105) individually containing the respective constructs, were adjusted to an OD600 of 0.5 with MES buffer (10 mM $MgCl_2$, 10 mM MES and 0.2mM acetosyringone, pH 5.6) and used for co-infiltration of N. benthamiana leaves. At 48 h after co-infiltration, infiltrated leaves were incubated with 150 ng/mL D-luciferin potassium salt and photographed with NightSHADE LB 985 in vivo plant imager to qualitatively measure LUC activity. Leaf discs were then taken and incubated with 150 ng/mL D-luciferin potassium salt in a 96-well plate, and the relative LUC activity was quantified with a GLOMAX 96 microplate luminometer (Promega). All primer sequences used for the constructs are listed in Table 5-S1.

V.2.6. In vivo co-immunoprecipitation assay

Co-immunoprecipitation (Co-IP) assays were performed as previously described (Wang et al., 2016). The desired constructs were co-transfected into rice protoplasts, 20 mM MG132 was added to the treated rice protoplasts at 12 h after co-transfection, and 12 h later, total protein was extracted in native buffer (50 mM Tris-MES, pH 8.0, 0.5 M sucrose, 1 mM $MgCl_2$, 10 mM EDTA, 5 mM DTT, 50 mM MG132, and protease inhibitor cocktail). The proteins were then incubated with anti-HA antibody with gentle shaking for 6 h at 4°C before 25 mL of pre-rinsed Protein G beads (Millipore) was added to the protein-antibody samples, after which the preparation was incubated for another 3 h. The beads were then washed 3-5 times with 13phosphate buffered saline with 1% Tween-20 (PBST) buffer before 1xSDS loading buffer was added to each sample, which was boiled for 8 min. About 10 mL of each sample was separated by SDS-PAGE for immunoblot analysis. All primer sequences used for the constructs are listed in Table 5-S1.

V.2.7. Bimolecular fluorescence complementation assay in N. benthamiana

For the bimolecular fluorescence complementation (BiFC) assays, full-length coding sequences of *OsVQ25*, *OsPUB73*, *OsWRKY53*, and *WRKY45* were individually cloned into the p2YN (nYFP) or p2YC (cYFP) vectors to produce the fusion to the N- or C-terminal half of YFP (You et al., 2019), including p2YN-OsVQ25, p2YC-OsPUB73, p2YC-OsWRKY53, and p2YC-WRKY45. The resulting plasmids were separately transformed into *Agrobacterium* (strain EHA105) and then transiently infiltrated in *N. benthamiana* leaves. Fluorescent signals were observed using a laser scanning confocal microscope (Zeiss LSM880) between 48 and 72 h after infiltration. All primer sequences used for the constructs are listed in Table 5-S1.

V.2.8. Protein degradation assay in rice protoplasts

Protein degradation assays in rice protoplasts were performed as previously described (Wang et al., 2021). Briefly, 12 h after the desired constructs were co-transfected into rice protoplasts, 20 mM MG132 was added to the transfected protoplasts, and 12 h later, the rice protoplasts were harvested for protein extraction. Total protein was extracted with denaturation buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% [v/v] NP40, 4 M urea, and 1 mM PMSF). Protein abundance was detected by immunoblotting and normalized to rice ACTIN levels in each sample. In addition, total RNA was isolated, and the relative transcript levels of each gene were determined by RT-PCR. All primer sequences used for the constructs are listed in Table 5-S1.

V.2.9. Transcriptional activity assay in Arabidopsis protoplasts

The transcriptional activity assay in *Arabidopsis* protoplasts was performed as previously described (Wang et al., 2021). The GUS reporter gene was cloned downstream of four GAL4 DNA-binding sites and four tandem copies of the constitutive D1-3 element (GAL4(4X)–D1-3(4X)). The OsWRKY53 full-length coding sequence was cloned in-frame with the sequence of GAL4-DB under the control of the cauliflower mosaic virus (CaMV) 35S promoter to obtain the GAL4DB-OsWRKY53 construct. The 35S promoter was also used to drive the firefly LUC gene as the internal control. The OsVQ25 coding sequence was cloned into the pGreenII 62-SK vector and placed under the control of the 35S promoter as an effector. All primer sequences used for the constructs are listed in Table 5-S1.

Arabidopsis protoplasts were isolated from the leaves of approximately 5-week-old WT (Columbia-0) plants, and polyethylene glycol (PEG)-mediated transformation was used for transfection as previously described (Yoo et al., 2007). LUC and GUS activities were measured with a GLOMAX 96 microplate luminometer (Promega) (Promega) and a FlexStation 3 (Molecular Devices), respectively. The GUS/LUC ratios were used to evaluate the transcriptional activity of OsWRKY53.

V.2.10. Electrophoretic mobility shift assay

Recombinant His-OsWRKY53 protein was prepared as described previously (Xie et al., 2021). The plasmids harboring the OsVQ25 coding sequences were introduced into Escherichia coli BL21(DE3) cells, and then 0.2 mM isopropylthio-b-galactoside (IPTG) was added to induce protein production overnight at 16°C, and the proteins were purified using Amylose Resin (BioLabs, #E8021V). The electrophoretic mobility shift assay (EMSA) and 50’-FAM-modified oligonucleotide probes containing W-box elements from *OsMYB63* promoter regions were described previously (Xie et al., 2021).

V.2.11. Dual-luciferase assay in rice protoplasts

The dual-luciferase assay in rice protoplasts and reporter constructs used in this

assay were described previously (Xie et al., 2021). The OsVQ25 effector construct was co-transfected with effector and reporter constructs into rice protoplasts to test the effect of OsVQ25 on OsWRKY53 transcription. The protoplasts were collected 24 h after transfection with the indicated combinations. The LUC and REN activities were measured according to the manufacturer's instructions (Promega, USA). The relative reporter gene expression level was calculated as the ratio between LUC and REN activity.

V.2.12. RNA extraction and quantitative reverse transcriptase PCR

Total RNA isolation and quantitative reverse transcriptase PCR (qRT-PCR) were performed as previously described (Fang et al., 2021). Total RNA was extracted from rice tissues using a plant RNA extraction kit (Sangon Biotech, Shanghai, China). For RT-qPCR, 2 mg of total RNA was reverse transcribed into first-strand cDNA with a one-step gDNA removal and cDNA synthesis supermix (TransGen Biotech, Beijing, China). qPCR was performed with 2xSYBR Green Mix (GeneStar, Beijing, China) on an ABI Prism 7500 PCR instrument (Applied Biosystems, Waltham, USA). The primer sequences used for qPCR are listed in Table 5-S1.

V.2.13. Quantitative and statistical analysis

Data for quantification analyses are presented as mean ± standard error of mean (SEM). At least two independent biological replicates were performed for each experiment. The asterisks indicate significant differences from the controls by two-tailed Student's t-test (*p < 0.05, **p < 0.01). The details are included in the figure legends.

V.3. Results

V.3.1. Regulation of rice resistance to Magnaporthe oryzae and Xanthomonas oryzae by OsPUB73

OsPUB73 is a functional U-box E3 ligase, the expression of whose encoding gene is induced by *M. oryzae* infection in rice cultivar TP309 (Zeng et al., 2008). To further explore the contribution of *OsPUB73* to blast resistance, we confirmed its expression pattern in Nipponbare (NPB) rice plants infected with the compatible *M. oryzae* isolate RB22. Quantitative reverse transcriptase PCR (qRT-PCR) analysis showed that *OsPUB73* is highly induced at 96 h post inoculation (Figure 5-1A). To determine *OsPUB73* function in rice immunity, we generated *ospub73* knockout mutants by clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein 9 (Cas9)-mediated genome editing. We selected two independent homozygous lines, 3-1 and 4-1 (in which a 1-bp insertion or a 20-bp deletion leads to early termination and a frameshift mutation, respectively) (Figures 5-S1A and S1B). We used the punch inoculation method to assess the resistance of

the *ospub73* mutants to the compatible *M. oryzae* isolate RB22. At 2 weeks after inoculation, the *ospub73* mutants had developed larger disease lesions than the WT (Figures 5-1B and 1C). The relative fungal biomass, as measured by quantitative PCR (qPCR), was higher in *ospub73* mutants than in the WT (Figure 5-1D), suggesting that *OsPUB73* positively regulates *M. oryzae* resistance. To assess if the loss of *OsPUB73* function affects resistance to other pathogens, we inoculated WT and the ospub73 mutants with the *Xoo* compatible isolate PXO99A. At 2 weeks after inoculation, the *Xoo* disease lesions were longer on the *ospub73* mutants than on the WT (Figures 5-1E and 1F), indicating that *OsPUB73* also positively regulates *Xoo* resistance. We also measured the relative transcript levels of three defense-related genes, *OsPBZ1*, *WRKY45*, and *OsPR1*, in WT and the *ospub73* mutants by qRT-PCR. All three genes were significantly downregulated in the *ospub73* mutants compared with the WT (Figures 5-1G–1I). Together, these results demonstrated that *OsPUB73* positively regulates SNS BSR in rice.

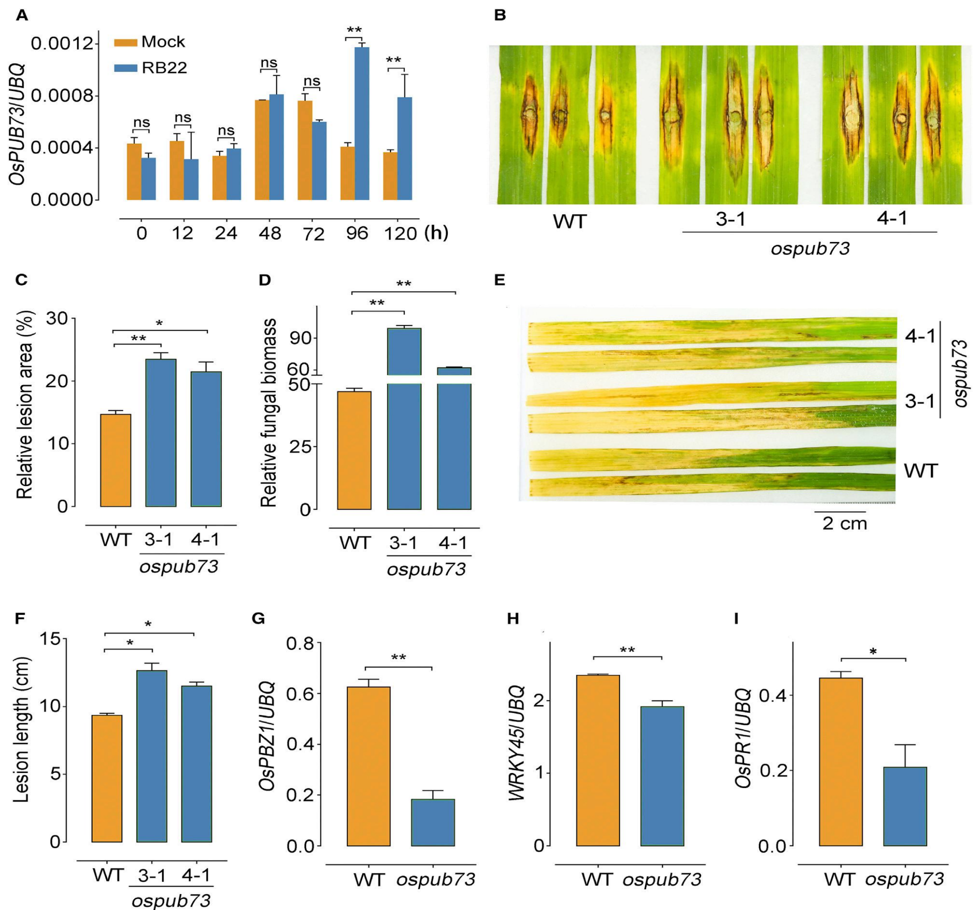


Figure 5-1. Expression pattern of *OsPUB73* and resistance phenotypes of the *ospub73* knockout mutant to *Magnaporthe oryzae* and *Xanthomonas oryzae.*

(A) *OsPUB73* expression in Nipponbare (NPB) plants inoculated with the compatible *M. oryzae* isolate RB22, as determined by qRT-PCR. We used ddH_2O containing 0.1% (v/v) Tween 20 as the mock-inoculation control and rice *UBIQUITIN* (*UBQ*) as the reference gene to normalize gene expression. Values are means ± SEM of two biological replicates.
(B–D) Phenotypes of leaves from 8-week-old *ospub73* mutant plants inoculated with the compatible *M. oryzae* isolate RB22 (B), the percentage of leaf area with lesions, as measured by ImageJ (C), and relative fungal biomass, as determined by qPCR [$2^{[CT(OsUbq)\ CT(MoPot2)]}$] (D). Values are means ± SEM (n = 3 biological replicates).
(E and F) Phenotypes of leaves from 8-week-old *ospub73* mutant plants inoculated with the *Xoo* isolate PXO99A (E), and lesion length (F). Values are means ± SEM (n = 3 biological replicates).
(G–I) Relative expression levels of the defense-related genes *OsPBZ1* (G), *WRKY45* (H), and *OsPR1* (I) in *ospub73* mutants and WT plants, as determined by qRTPCR. Values are means ± SEM (n = 3 biological replicates). For (A), (C)–(D), and (F)–(I), asterisks indicate statistical significance ("ns" indicates no statistical significance at p > 0.05, *p % 0.05, **p % 0.01, Student's t test). See also Figure 5-S1.

A

Target-base editing types of *OsPUB73* CRSPR/Cas9 lines	
5'—CCACGAGGTCTGACTACGCCGCG—3'	WT
5'—CCACAGAGGTCTGACTACGCCGCG—3'	3-1 (+A, homo)
5'—CC -------------------- GCG—3'	4-1 (-20, homo)

B

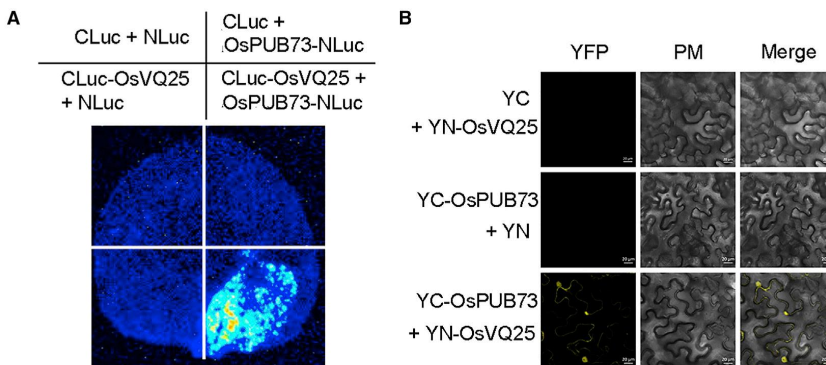
```
OsPUB73 WT     MDPEAEEAQLRLEMELAKKAKADMSGLQRSSSLGLDHAGLYPLPLPPGWRSAPTSPLRTPSSPPPLQFPP   70
ospub73 3-1    MDPEAEEAQLRLEMELAKKAKADMSGLQRSSSLGLDHAGLYPLPLPPGWRSAPTSPLRTPSSPPPLQFPP   70
ospub73 4-1    MDPEAEEAQLRLEMELAKKAKADMSGLQRSSSLGLDHAGLYPLPLPPGWRSAPTSPLRTPSSPPPLQFPP   70

OsPUB73 WT     AWAADVAGTSGSAAPEDDGPARNAGADEATAGSAPKNEDPARAAGADDGPTRSDYAAMMRMALAKFQDDD  140
ospub73 3-1    AWAADVAGTSGSAAPEDDGPARNAGADEATAGSAPKNEDPARAAGADDGPTEV.................  123
ospub73 4-1    AWAADVAGTSGSAAPEDDGPARNAGADEATAGSAPKNEDPARAAGADDGRDDADGIGQVPRRRCCCRRRG  140

OsPUB73 WT     AAADDEEAASAVMEQAMTGLMDLTYRKAKPPELPYEFATRWPIPIAHDGTLQAEVMRDPVILPSGYSVDQ  210
ospub73 3-1    ......................................................................  123
ospub73 4-1    GGVRGDGAGDDRPHGPHLPQSEASRAALRVRHKMAYSYCS..............................  180

OsPUB73 WT     TYQNNQKRQNPWTNTSTFTDHSLPYSLSVPNHLLRDMISAWCLDHSDLSPSTTSDTPSTPLEPSEEEQIQ  280
ospub73 3-1    ......................................................................  123
ospub73 4-1    ......................................................................  180

OsPUB73 WT     RILKLFSGNSASQREALKLIQLLTKTTKGVQPCLAKYADIIPVLINLRRKYKSSWTQDLEEERLTIILNL  350
ospub73 3-1    ......................................................................  123
ospub73 4-1    ......................................................................  180

OsPUB73 WT     TMHRQNREILAGQNELAGAIKKIVKKAGNRGKRTSSLAKVASIVAVLSEFDMFRKRMLDAGGMKMLRGML  420
ospub73 3-1    ......................................................................  123
ospub73 4-1    ......................................................................  180

OsPUB73 WT     KIKDTEVITEAATAILALYADGEGEQPARFHEVPQMLLECHMFTDGILLLLDRLPKSPRVFRKICDQALQ  490
ospub73 3-1    ......................................................................  123
ospub73 4-1    ......................................................................  180

OsPUB73 WT     LVNIVMAEDASGPVTRKGILSAISLIYEIVERDVGKMNAVKNMEDFIERLRQLSSDRLPMQKMLQVERII  560
ospub73 3-1    ......................................................................  123
ospub73 4-1    ......................................................................  180

OsPUB73 WT     RTLSDAFPAPTVRGRCQEPSGSRLL                                              585
ospub73 3-1    .........................                                              123
ospub73 4-1    .........................                                              180
```

Figure 5-S1. Analysis of *OsPUB73* CRISPR/Cas9 lines. Related to Figure 5-1.

(A) Sequence of the WT and *ospub73* mutant alleles, based on PCR amplification and Sanger sequencing of *OsPUB73* genomic DNA.
(B) Predicted protein sequence for OsPUB73 in WT and *ospub73* mutants. Black shading indicates shared amino acids, red boxes indicate the U-box domain of OsPUB73, and black dots indicate missing amino acids.

V.3.2. VQ-motif protein OsVQ25 interactions with OsPUB73

To investigate the molecular mechanism of OsPUB73 in rice resistance, we identified OsPUB73-interacting proteins using immunoprecipitation (IP) followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) assays in rice protoplasts. Among the candidate proteins obtained from the LC-MS data, a VQ-motif-containing protein (encoded by LOC_Os06g45570, annotated and designated as *OsVQ25*) caught our attention, as it is mainly expressed in rice leaves and seedlings and is induced by *Xoo* and *M. oryzae* (Kim et al., 2013; Li et al., 2014). We tested the interaction between OsVQ25 and OsPUB73 by a luciferase complementation imaging (LCI) assay in *Nicotiana benthamiana* leaves. Co-infiltration of OsPUB73-Nluc and Cluc-OsVQ25 constructs led to a stronger luciferase reporter signal and higher luciferase activity than in the control combinations (Figures 5-2A and 5-S2). We also performed a bimolecular fluorescence complementation (BiFC) assay in *N. benthamiana* leaves. Pairwise expression of constructs encoding the N-terminal part of the yellow fluorescent protein (YFP) fused to OsVQ25 (YN-OsVQ25) and of the C-terminal part of YFP fused to OsPUB73 (YC-OsPUB73) resulted in a YFP fluorescence signal in the cytoplasm and the nucleus at 72 h post infiltration, but not with the control combinations encoding YN-OsVQ25/YC and YN/YC-OsPUB73 (Figure 5-2B), indicating that OsVQ25 interacts with OsPUB73 in the cytoplasm and the nucleus. To explore the OsVQ25-OsPUB73 interaction in vivo, we performed a co-immunoprecipitation (Co-IP) assay by transiently expressing OsVQ25-GFP with OsPUB73-HA in rice protoplasts, using NLuc-HA and 23GFP (two tandemly repeated Green fluorescent protein [GFP] genes) as negative controls. Immunoblot analysis with an anti-GFP antibody showed that OsVQ25-GFP co-precipitates with OsPUB73-HA, but not with the control NLuc-HA; in addition, 23GFP did not co-precipitate with OsPUB73-HA (Figure 5-2C). Together, these results demonstrate that OsVQ25 interacts with the U-box E3 ligase OsPUB73 in planta.



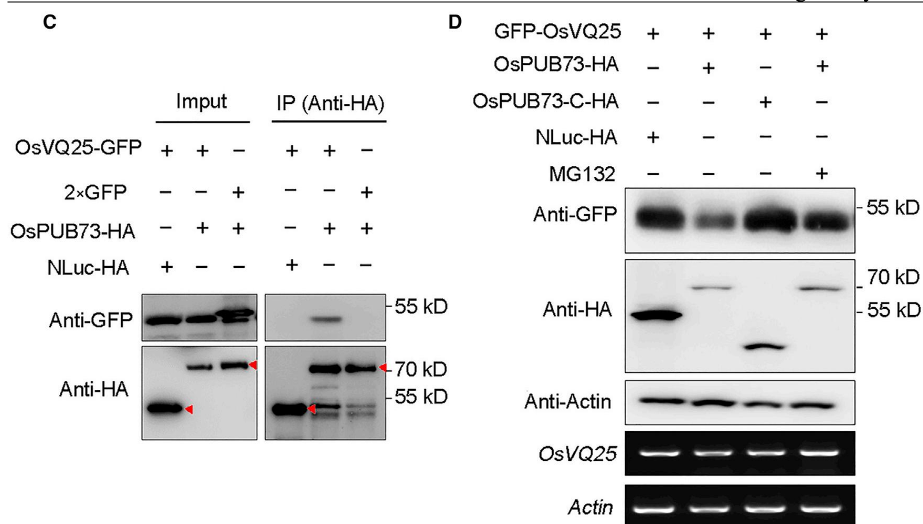


Figure 5-2. OsPUB73 interactions with and promoting the degradation of OsVQ25.
(A) LCI assay showing the interaction between OsVQ25 and OsPUB73 in N. benthamiana leaves. Luminescence level was determined at 48 h after infiltration with the indicated constructs.
(B) BiFC assay to test the interaction between OsVQ25 and OsPUB73 in *N. benthamiana* leaves. Fluorescence from *N. benthamiana* leaf cells was acquired on a confocal microscope at 48 h after infiltration with the indicated constructs. Scale bar, 20 mm.
(C) Co-IP assay to test the interaction between OsVQ25 and OsPUB73 in transfected rice protoplasts. Total protein from rice protoplasts transfected with the indicated plasmid combinations was extracted and subjected to immunoprecipitation with anti-HA antibody. Red arrowheads indicate the expected proteins. There were three biological replicates with similar results.
(D) Degradation of OsVQ25 by OsPUB73 via the 26S proteasome pathway. *OsVQ25-GFP* was co-transfected with OsPUB73-HA, *OsPUB73-C-HA*, or *NLuc-HA* in rice protoplasts, followed by immunoblotting. ACTIN serves as an internal control. We added 20 mM MG132 or an equivalent volume of DMSO (as a control) at 12 h before sampling. The relative transcript levels of *OsVQ25* and *ACTIN* were detected by RT-PCR. There were three biological replicates with similar results.

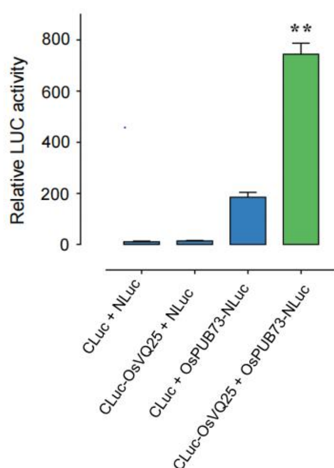


Figure 5-S2. OsPUB73 interacts with OsVQ25 *in planta.* Related to Figure 5-2.

Quantitative luciferase complementation imaging assay to test the interaction between OsVQ25 and OsPUB73 in *N. benthamiana* leaves; data are shown as means ± SEM of three biological replicates. Asterisks represent statistical significance (** $P \leq 0.01$, Student's t-test).

V.3.3. OsVQ25 degradation by OsPUB73 via the ubiquitin 26S proteasome

OsPUB73 possesses E3 ubiquitin ligase activity (Zeng et al., 2008). The interaction between OsVQ25 and OsPUB73 suggested that OsPUB73 might promote OsVQ25 degradation via ubiquitination. To test this idea, we generated a construct encoding OsPUB73-C, which is a truncated OsPUB73 variant with the C terminus of OsPUB73 but lacking the U-box domain, for a degradation assay. We co-transfected OsVQ25-GFP and OsPUB73-HA or OsPUB73-C-HA in rice protoplasts, using NLuc-HA as the control. OsVQ25-GFP abundance was comparable when OsVQ25-GFP was co-expressed with NLuc-HA or OsPUB73-C-HA. However, the intensity of the OsVQ25-GFP band was clearly weaker when OsVQ25-GFP was co-expressed with full-length OsPUB73-HA (Figure 5-2D, lane 2). We then asked if OsVQ25 degradation is affected by the proteasome inhibitor MG132. Accordingly, we treated rice protoplasts transfected with the OsPUB73-HA and OsVQ25-GFP plasmids with MG132, which revealed that MG132 inhibits OsVQ25-GFP degradation (Figure 5-2D, lane 4). ACTIN, the internal control, displayed a similar accumulation in all tested combinations, and the relative OsVQ25 and ACTIN transcript levels were also similar (Figure 5-2D). These results suggest that OsPUB73 specifically promotes OsVQ25 degradation via the 26S proteasome-dependent pathway in planta.

V.3.4. Enhanced resistance of rice osvq25 mutant against Magnaporthe oryzae and Xanthomonas oryzae

OsVQ25 expression is induced by *M. oryzae* in rice cultivar CO39 (Li et al., 2014). To determine if *OsVQ25* participates in disease resistance, we confirmed its expression pattern in NPB plants infected with the compatible *M. oryzae* isolate RB22. *OsVQ25* was rapidly induced at 12 h and was highly expressed at 96 and 120 h post inoculation compared with mock-inoculated plants (Figure 5-3A). To explore the function of *OsVQ25* in resistance, we generated *osvq25* knockout mutants by CRISPR-Cas9-mediated gene editing. After genotyping, we selected three independent homozygous mutant lines (28-1, 29-1, and 48-1; with a 1-bp insertion of an A, a C, and a T, respectively, all causing frameshift mutations) for punch inoculation assays (Figures 5-S3A and 5-S3B). Two weeks after inoculation with the compatible *M. oryzae* isolate RB22, *osvq25* mutants developed smaller disease lesions and accumulated less fungal biomass than the WT (Figures 5-3B–3D). Because *OsVQ25* is also induced by *Xoo* (Kim et al., 2013; Li et al., 2014), we inoculated *osvq25* mutants and the WT with the *Xoo* isolate PXO99A and observed that all three mutant lines are also more resistant to *Xoo* (Figures 5-3E and 3F). To verify that the *osvq25* mutants have SNS BSR, we inoculated the mutants with another compatible *M. oryzae* strain, RO1-1, and another compatible

Xoo strain, PXO86. The *osvq25* mutants also showed enhanced resistance to these pathogen strains (Figures 5-S3C-S3F). Consistent with the disease-resistant phenotypes of the mutants, expression of the defense-related genes *OsPBZ1*, *WRKY45*, and *OsPR1* was significantly higher in *osvq25* mutant plants than in WT plants (Figures 5-3G-3I). Together, these results demonstrate that *OsVQ25* negatively regulates SNS BSR in rice.

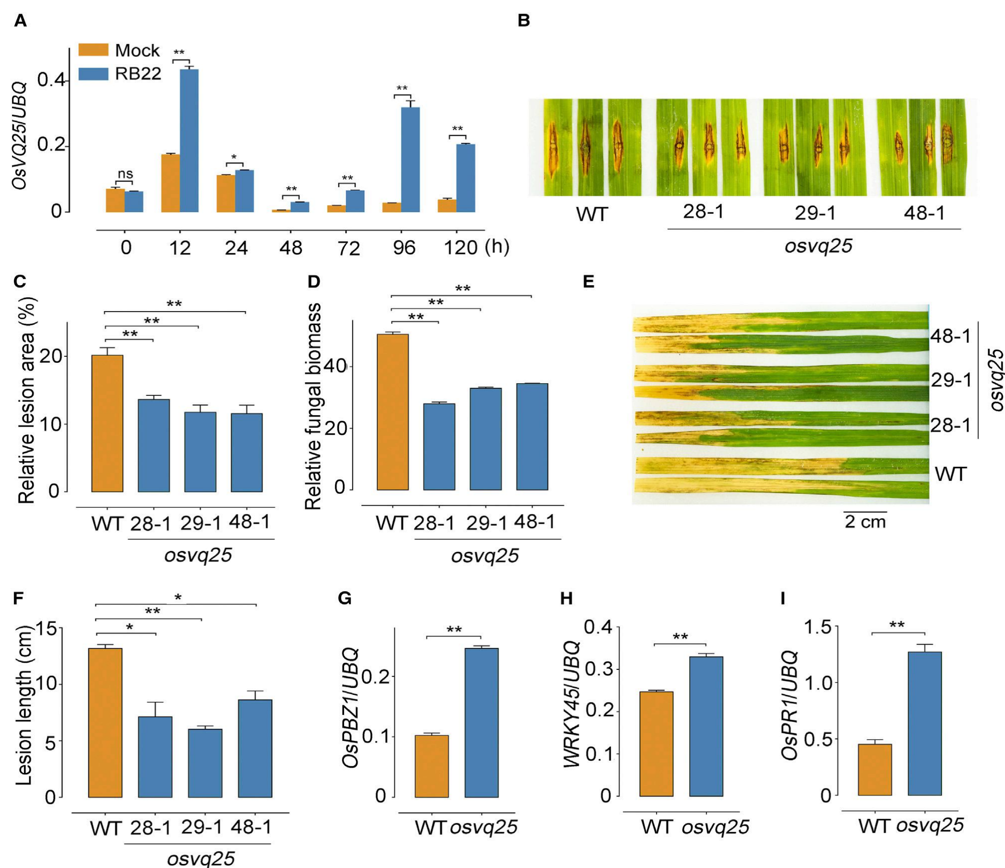


Figure 5-3. Expression pattern of *OsVQ25* and enhanced resistance of the *osvq25* mutant to *Magnaporthe oryzae* and *Xanthomonas oryzae.*

(A) *OsVQ25* expression in NPB plants inoculated with the compatible *M. oryzae* isolate RB22, as determined by qRT-PCR. We used ddH_2O containing 0.1% (v/v) Tween 20 as the mock-inoculation control and rice *UBIQUITIN* (*UBQ*) as the reference gene to normalize gene expression. Values are means ± SEM (n = 2 biological replicates).

(B–D) Phenotypes of the leaves from 8-week-old *osvq25* mutant plants inoculated with the compatible *M. oryzae* isolate RB22 (B), percentage of leaf area with lesions, as measured by ImageJ (C), and relative fungal biomass, as determined by qPCR [$2^{[CT(OsUbq)\ CT(MoPot2)]}$] (D). Values are means ± SEM (n = 3 biological replicates). Phenotypes of *osvq25* mutants inoculated with another compatible *M. oryzae* isolate, RO1-1.

(E and F) Phenotypes of the leaves of 8-week-old *osvq25* mutant plants inoculated with the *Xoo* isolate PXO99A (E), and lesion length (F). Values are means ± SEM (n = 3 biological replicates).

(G–I) Relative transcript levels of the defense-related genes *OsPBZ1* (G), *WRKY45* (H), and *OsPR1* (I) in *osvq25* mutants and WT plants as determined by qRT PCR. Values are means ± SEM (n = 3 biological replicates). For (A),

(C)–(D), and (F)–(I), asterisks indicate statistical significance ("ns" indicates no statistical significance at $p > 0.05$, *p % 0.05, **p % 0.01, Student's t test). See also Figure 5-S3.

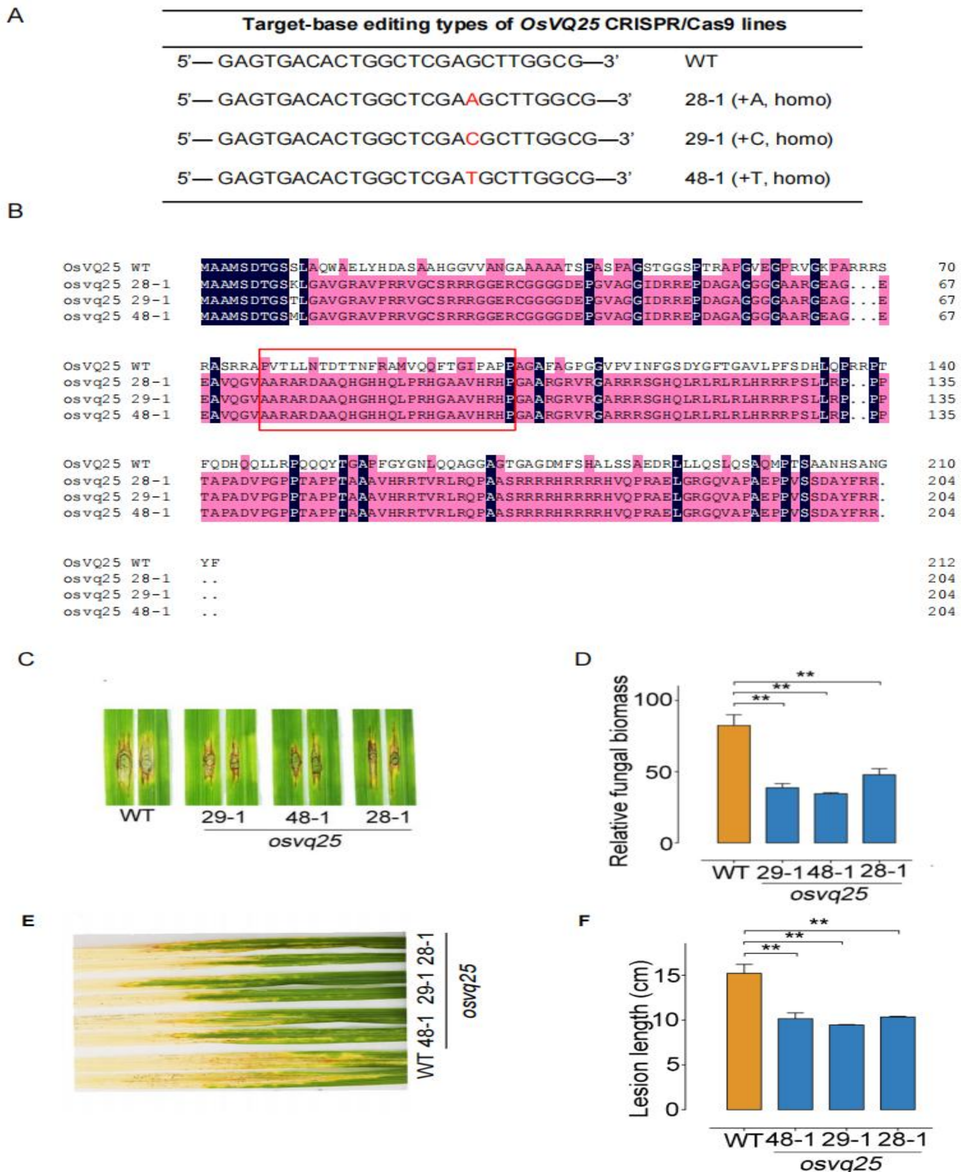


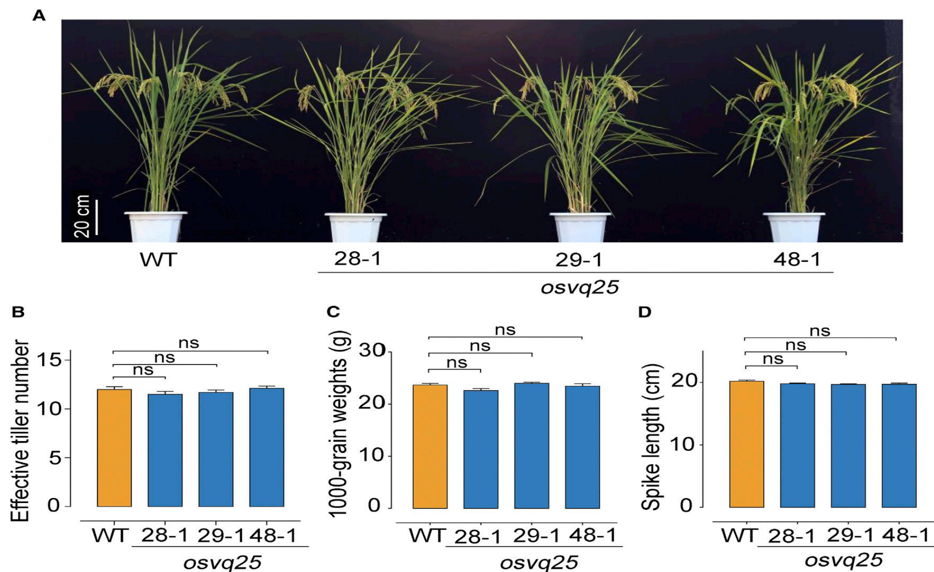
Figure 5-S3. Loss of OsVQ25 function enhances resistance against *Magnaporthe oryzae* and *Xanthomonas oryzae*. Related to Figure 5-3.

(A) Sequence of the WT and *osvq25* mutant alleles, based on PCR amplification and Sanger sequencing of *OsVQ25* genomic DNA.

(B) Predicted protein sequence of OsVQ25 in WT and *osvq25* mutants. Black shading indicates shared amino acids, pink shading indicates amino acids not present in the WT sequence, red box indicates the VQ domain of OsVQ25, and black dots indicate missing amino acids.
(C) Phenotypes of the leaves from eight-week-old *osvq25* mutant plants inoculated with the compatible *M. oryzae* isolate RO1-1.
(D) Relative fungal biomass, as determined by qPCR [$2^{[CT(OsUbq)-CT(MoPot2)]}$]; data are shown as means ± SEM (n = 3 biological replicates).
(E) Phenotypes of the leaves from eight-week-old *osvq25* mutant plants inoculated with the *Xoo* isolate PXO86 .
(F) Lesion length in WT and *osvq25* mutants inoculated with the *Xoo* isolate PXO86; data are shown as means ± SEM (n = 3 biological replicates). Asterisks in (D) and (F) represent statistical significance (**$P \leq 0.01$, Student's t-test).

V.3.5. Agronomic traits of resistant osvq25 mutant

The *osvq25* mutants used in this article are all materials without transgenic elements. We have detected *osvq25* mutants without vector sequences (such as *Cas9*, *sgRNA*, *Hpt*, etc.) through genetic segregation and PCR detection for further resistance and agronomic trait testing (Figure 5-S4). To test if the enhanced resistance of the *osvq25* mutants affects rice growth, we assessed key agronomic traits in the WT and the mutants under field conditions. Both plant architecture and panicle type were similar in the *osvq25* mutants and WT plants (Figure 5-4A). The yield-related traits (effective tiller number, thousand-grain weight, spike length, spikelet number, plant height, and kernel number per spike) were also similar across the genotypes (Figures 5-4B–4G). In addition, grain length and grain width were almost the same in *osvq25* and WT plants (Figure 5-4H). These results indicate that the loss of *OsVQ25* function does not incur a growth penalty in rice, and that *OsVQ25* is a valuable candidate gene for breeding SNS BSR rice varieties by genome editing.



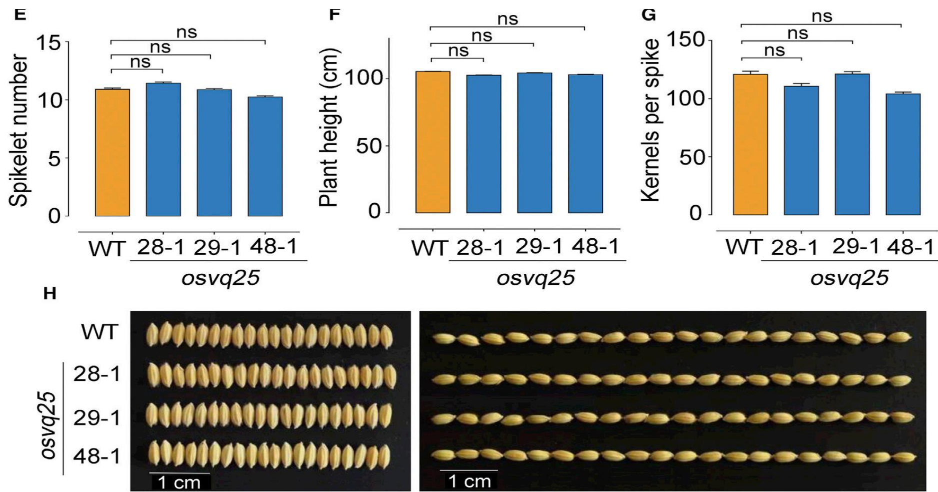


Figure 5-4. Major agronomic traits measured in *osvq25* mutant lines.
(A) Gross morphology of *osvq25* mutants and WT plants at the heading stage.
(B–G) Major agronomic traits measured: effective tiller number (B), thousand-grain weight (C), spike length (D), spikelet number (E), plant height (F), kernels per spike (G). For (B)–(G), values are means ± SEM (n ≥ 30, ≥ 10, ≥ 10, ≥ 30, ≥ 30, and ≥ 10, respectively). 'ns' indicates no statistical significance at $p > 0.05$ according to Student's t test.
(H) Grain length (left) and grain width (right) of *osvq25* mutants and the WT.

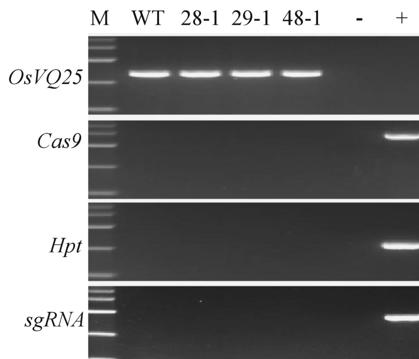


Figure 5-S4. Vector sequence detection in *osvq25* mutant plants.
M: Marker; '-' represents negative control; '+' represents positive plasmid control. Hpt: Hygromycin resistance gene

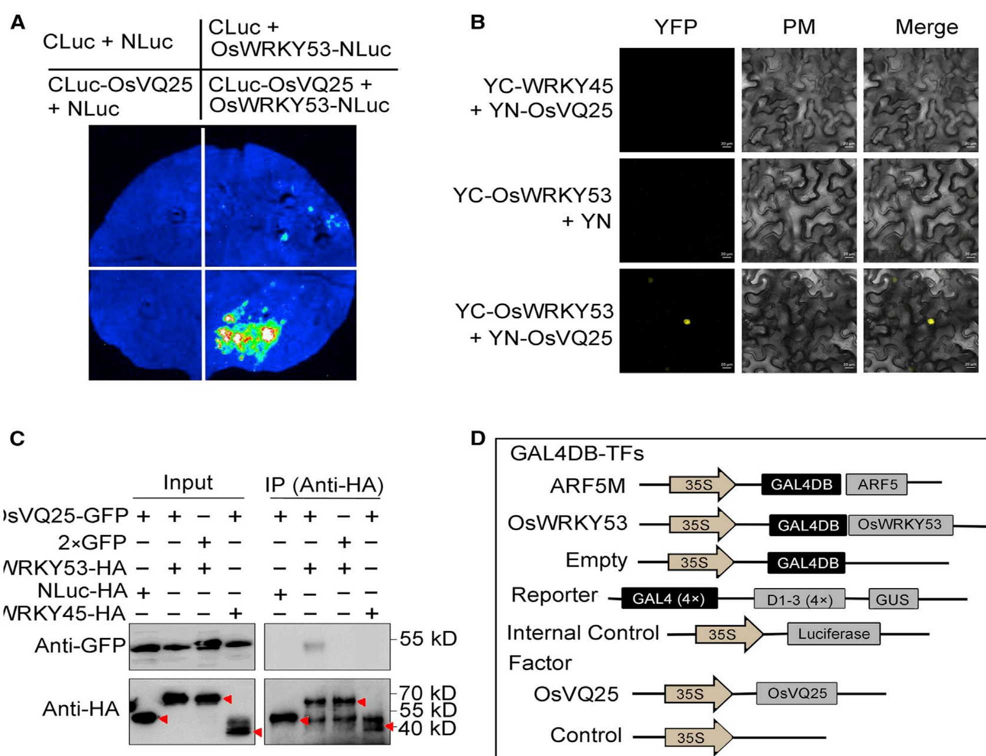
V.3.6. OsVQ25 interactions with OsWRKY53 transcription factor

To further dissect the underlying mechanism of OsVQ25 function in rice disease resistance, we queried the STRING database (https://string-db.org) for potential OsVQ25-interacting proteins. We focused on one such protein, the transcription factor OsWRKY53. We used LCI assays to investigate their potential physical

interaction in *N. benthamiana* leaves infiltrated with CLuc-OsVQ25 and OsWRKY53-NLuc constructs. Indeed, we detected strong luminescence signal when CLuc-OsVQ25 and OsWRKY53-NLuc were co-infiltrated, but not with any of the control combinations (Figure 5-5A). We also established that the interaction between OsVQ25 and OsWRKY53 is specific, as LCI assays with an unrelated WRKY family member, WRKY45, yielded luminescence signals as low as those of the negative controls (Figure 5-S5A). We also performed BiFC in *N. benthamiana* leaves by co-infiltrating the YN-OsVQ25 and YC-OsWRKY53 constructs, which reconstituted YFP fluorescence in the nucleus, unlike the co-infiltration of YN-OsVQ25 and YC-WRKY45, which produced no detectable fluorescence (Figure 5-5B), indicating that OsVQ25 specifically interacts with OsWRKY53 in the nucleus. We also performed a Co-IP assay by transiently transfecting rice protoplasts with the OsVQ25-GFP and OsWRKY53-HA constructs. OsVQ25-GFP co-precipitated with OsWRKY53-HA when OsVQ25-GFP was co-transfected with OsWRKY53-HA but not with the negative controls NLuc-HA or WRKY45-HA. The 2xGFP protein also did not co-precipitate with OsWRKY53-HA (Figure 5-5C). These results indicate that OsVQ25 interacts with OsWRKY53 in planta. *OsWRKY53*-overexpressing rice plants have increased resistance to *M. oryzae* (Chujo et al., 2007, 2014). We investigated if *OsWRKY53* is involved in *M. oryzae* resistance by analyzing its expression pattern during *M. oryzae* infection: we observed that *OsWRKY53* is highly induced at 12 h post inoculation (Figure 5-S6A). To explore the role of *OsWRKY53* in basal defense against *M. oryzae*, we used a previously generated *oswrky53* mutant in the Zhonghua11 (ZH11) background for punch inoculation (Xie et al., 2021). The *oswrky53* mutant exhibited decreased resistance to *M. oryzae* compared with ZH11 (Figures 5-S6B and S6C), suggesting a positive role in the *M. oryzae* response. Next, we compared *OsWRKY53* expression in *osvq25* mutants and the WT. The *OsVQ25* knockouts did not affect *OsWRKY53* transcript levels (Figure 5-S5B); furthermore, *OsWRKY53* expression was similar between *ospub73* and WT plants (Figure 5-S5C). We then asked if OsVQ25 affects OsWRKY53 transcriptional activity in *Arabidopsis* protoplasts. Transcriptional activator AUXIN RESPONSE FACTOR 5 MIDDLE REGION (ARF5M) was used as a positive effector control (Wang et al., 2021). To this end, we cloned *OsWRKY53* coding sequences in-frame and downstream of the sequence of GAL4 DNA-binding domain, and we used a reporter construct consisting of the b-glucuronidase (GUS) reporter gene under the control of the GAL4(4X)-D1-3(4X) (Four copies of the GAL4 DNA-binding site fused immediately upstream of four tandem copies of the constitutive D1-3 element). The OsWRKY53 effector construct showed a significantly higher relative GUS activity (normalized to firefly luciferase [LUC] activity) than that obtained with the ARF5M effector or the empty effector vector. However, relative GUS activity decreased significantly when we co-transfected the *OsWRKY53* effector construct with a construct expressing *OsVQ25*, but not with the control vector (Figures 5-5D and 5E). These results revealed that OsVQ25 reduces the transcriptional activity of OsWRKY53.

Previous study reported that OsWRKY53 directly binds to the *OsMYB63* promoter in vitro and in vivo, and it represses *OsMYB63* transcription (Xie et al., 2021). To test if OsVQ25 inhibits OsWRKY53 DNA binding, we performed an electrophoretic mobility shift assay (EMSA) and a dual-luciferase assay, which both demonstrated that OsVQ25 impairs OsWRKY53 binding to the *OsMYB63* promoter (Figures 5-S5D–S5F). We measured the expression levels of the defense genes *Chitinase* (Os01g0687400) and *PR-5* (Os12g0629700) in the *osvq25* mutants, as they are downstream of OsWRKY53 and are upregulated in *OsWRKY53*-overexpression plants (Chujo et al., 2014). The expression of these genes was significantly induced in *osvq25* mutants (Figures 5-5F and 5G).

OsWRKY53 increases brassinosteroid (BR) signaling, and two BR-responsive genes, *OsBU1* and *OsXTR1*, are upregulated in *OsWRKY53*-overexpression plants (Tian et al., 2017). We therefore also assessed *OsBU1* and *OsXTR1* expression in *osvq25* mutants. The expression of both genes was induced in *osvq25* mutants compared with the WT (Figures 5-5H and 5I). By contrast, *PR-5*, *OsBU1*, and *OsXTR1* were significantly suppressed in *ospub73* mutants (Figures 5-S5G–S5J). Taken together, our results demonstrate that OsVQ25 suppresses the transcriptional activity of OsWRKY53, which impairs the downstream defense and growth-related BR-signaling responses.



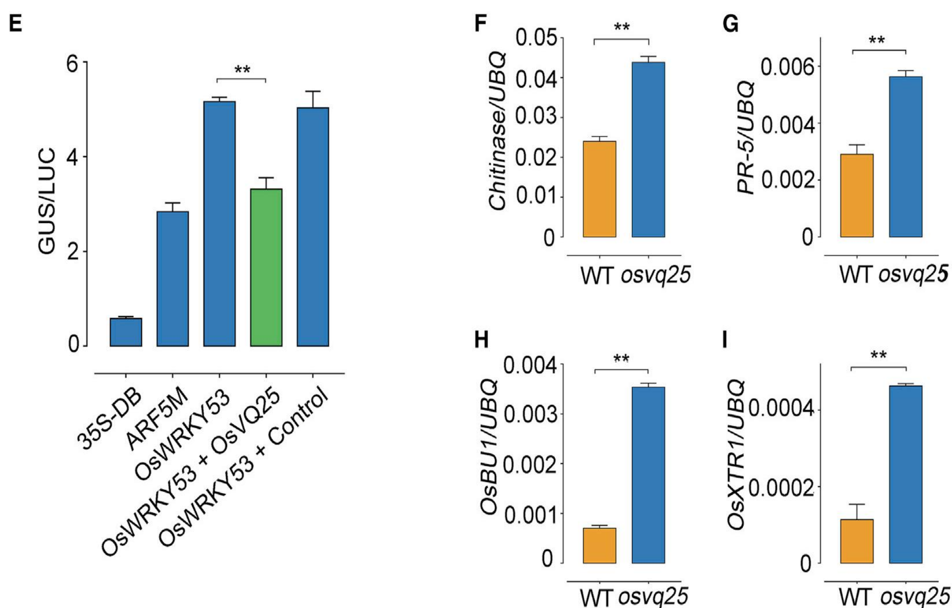


Figure 5-5. Interaction between OsVQ25 and OsWRKY53 and suppression of the transcriptional activity of OsWRKY53 by OsVQ25.

(A) LCI assay of the interaction between OsVQ25 and OsWRKY53 in N. benthamiana leaves. Luminescence level was determined at 48 h after infiltration with the indicated constructs.

(B) BiFC assay to test the interaction between OsVQ25 and OsWRKY53 in *N. benthamiana* leaves. YFP fluorescence was acquired from *N. benthamiana* leaf cells on a confocal microscope at 48 h after infiltration of the indicated constructs. Scale bar, 20 mm.

(C) Co-IP assay to test the interaction between OsVQ25 and OsWRKY53 in rice protoplasts. Total protein was extracted from rice protoplasts transfected with the indicated constructs, followed by immunoprecipitation with anti-HA antibody. Red arrowheads indicate the expected proteins. There were three biological replicates with similar results.

(D) Constructs used in the transcriptional activity assay in (E).

(E) OsVQ25 suppresses the transcriptional activity of WRKY53. We used ARF5M, a transcription activator, as a control. Asterisks indicate a significant difference between OsWRKY53 and the OsWRKY53 + OsVQ25 combination. Values are means ± SEM (n = 3 biological replicates).

(F and G) Relative transcript levels of the OsWRKY53-downstream defense-related genes *Chitinase* (F) and *PR-5* (G) in *osvq25* mutants and WT plants, as determined by qRT-PCR. We used rice *UBIQUITIN* (*UBQ*) as the reference gene to normalize gene expression. Values are means ± SEM (n = 3 biological replicates).

(H and I) Relative transcript levels of the OsWRKY53-downstream BR-signaling genes *OsBU1* (H) and *OsXTR1* (I) in *osvq25* mutants and WT plants, as determined by qRT-PCR. Values are means ± SEM (n = 3 biological replicates). Asterisks in (E)–(I) indicate statistical significance (**p % 0.01, Student's t test).

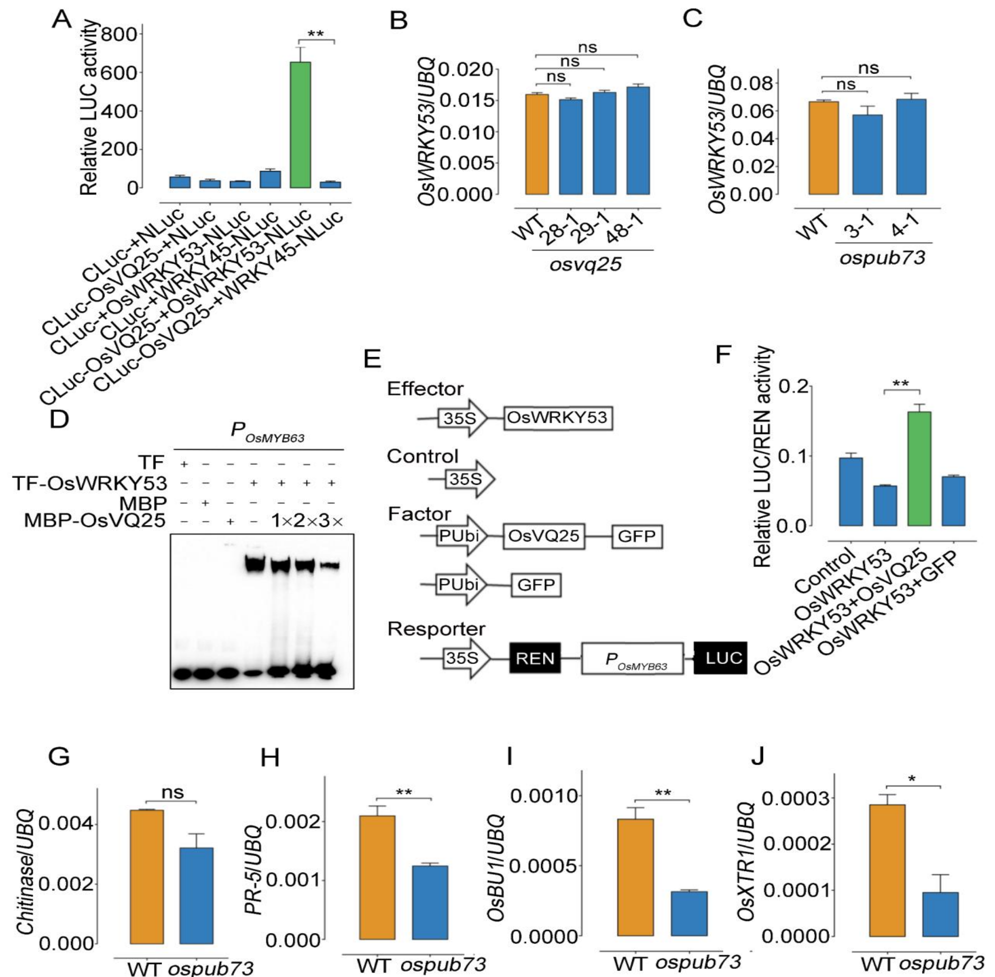


Figure 5-S5. OsVQ25 interacts with OsWRKY53 and suppresses the transcriptional activity and DNA binding of OsWRKY53. Related to Figure 5-5.

(A) Quantitative luciferase complementation imagine assay to test the interaction between OsVQ25 and OsWRKY53 in *N. benthamiana* leaves. Data are shown as means ± SEM (n = 3 biological replicates).

(B) Relative *OsWRKY53* transcript levels in *osvq25* mutants and WT plants, as determined by qRTPCR. Rice *UBIQUITIN* (*UBQ*) was used as the reference gene to normalize gene expression. Data are shown as means ± SEM (n = 3 biological replicates).

(C) Relative *OsWRKY53* transcript levels in *ospub73* mutants and WT plants, as determined by qRTPCR. Rice *UBIQUITIN* (*UBQ*) was used as the reference gene to normalize gene expression.

(D) OsVQ25 inhibits OsWRKY53 binding to the *OsMYB63* promoter in EMSA. '1×', '2×', and '3×' indicate the addition of 10 μg, 20 μg, or 30 μg recombinant purified MBP-OsVQ25 in the EMSA, respectively.

(E) Constructs used in dual-luciferase assay of OsWRKY53 in regulating *OsMYB63* expression.

(F) OsVQ25 increases *OsMYB63* expression in the dual-luciferase assay. The *ProOsMYB63:LUC* reporter construct was co-transfected with control, or with the constructs *35S:OsWRKY53*, *35S:OsWRKY53* + *Ubi:OsVQ25*, or *35S:OsWRKY53* + *Ubi:GFP* in rice protoplasts respectively. LUC and REN activity was determined at 24 h after transformation. Values are means ± SEM (n = 3 biological replicates).

(G-H) Expression of the OsWRKY53 downstream defense-related genes *Chitinase* and *PR-5* in *ospub73* mutants and WT plants, as determined by qRT-PCR. Rice *UBIQUITIN* (*UBQ*) was used as the reference gene to normalize gene expression. Data are shown as means ± SEM (n = 3 biological replicates).
(I-J) Relative transcript levels of the OsWRKY53 downstream BR-signaling genes *OsBU1* (C) and *OsXTR1* (D) in *ospub73* mutants and WT plants, as determined by qRT-PCR. Data are shown as means ± SEM (n = 3 biological replicates).
For (A)-(C), (F)-(J), 'ns' indicates no statistical significance at p > 0.05, asterisks represent statistical significance, *P ≤ 0.05, **P ≤ 0.01, Student's t-test.

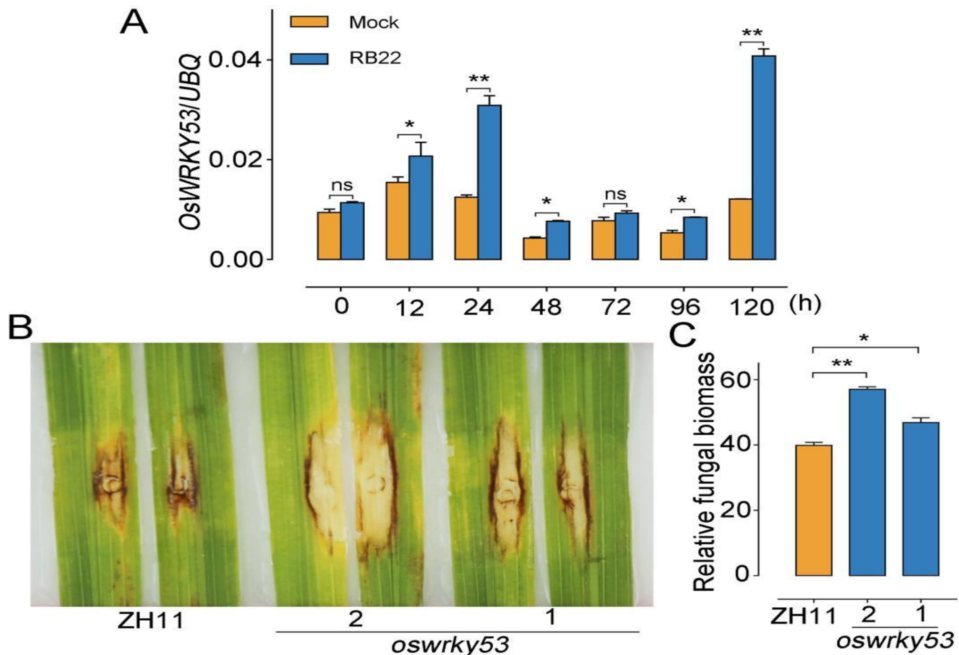


Figure 5-S6. Loss of OsWRKY53 function suppresses *Magnaporthe oryzae* resistance. Related to Figure 5-5.
(A) Relative *OsWRKY53* transcript levels in NPB plants inoculated with the compatible *M. oryzae* isolate RB22, as determined by qRT-PCR. We used ddH2O containing 0.1% (v/v) Tween 20 as the mock-inoculation control and rice *UBIQUITIN* (*UBQ*) as the reference gene to normalize gene expression. Values are means ± SEM of two biological replicates.
(B) Phenotypes of the leaves from eight-week-old *oswrky53* mutant plants inoculated with the compatible *M. oryzae* isolate RB22.
(C) Relative fungal biomass as determined by qPCR [$2^{[CT(OsUbq)-CT(MoPot2)]}$]; data are shown as means ± SEM (n = 3 biological replicates).
For (A) and (C), 'ns' indicates no statistical significance at p > 0.05, asterisks represent statistical significance, *P ≤ 0.05, **P ≤ 0.01, Student's t-test.

V.4. Discussion

Because crop plants growing in a region are often attacked by more than one pathogen, SNS BSR is a highly desirable trait in a crop breeding program (Li et al., 2020; Nelson et al., 2018). However, strong immune responses usually come with growth penalties (Ning et al., 2017). Therefore, SNS BSR genes that provide immunity without a growth penalty are favored by plant breeders. SNS BSR genes

with potential applications in crop breeding have been identified in rice. The tetratricopeptide repeat domain RNA-binding protein Bsr-k1 negatively regulates SNS BSR (Zhou et al., 2018). *Bsr-k1* knockout in rice leads to enhanced resistance against *M. oryzae* and *Xoo* without obvious growth penalties via moderately elevating the expression of phenylalanine ammonia lyase genes (Zhou et al., 2018). In another example, *IPA1* encodes a transcription factor that reduces the number of unproductive tillers, increases the number of grains per panicle, and positively regulates SNS BSR. Upregulating *IPA1* leads to enhanced resistance against *M. oryzae* and *Xoo*, as well as improved yield (Liu et al., 2019; Wang et al., 2018). Furthermore, VQ proteins contribute to disease resistance in *Arabidopsis* (Jing and Lin, 2015), but the role of VQ protein in SNS BSR in crop plants has not been documented. In this study, we demonstrate that knockouts in the *OsVQ25* gene encoding a VQ protein increases resistance to diverse isolates of *M. oryzae* and *Xoo* in rice. Compared with the three previously reported VQ proteins (OsVQ13, OsVQ14, and OsVQ32) related to rice immunity, OsVQ25 clustered on a different branch of a phylogenetic tree (Figure 5-S7), indicating that OsVQ25 has different functions in rice growth and disease responses. Importantly, the *osvq25* mutant presented SNS BSR with no obvious growth penalty in the major agronomic traits evaluated here. The moderately elevated activation of defense-related genes in *osvq25* may be sufficient to defend against pathogen attacks, making it a good candidate for rice breeding.

Plants use the UPS to regulate protein turnover for growth, development, and responses to abiotic and biotic stresses (Duplan and Rivas, 2014; Vierstra, 2009). In *Arabidopsis*, methyljasmonate treatment and wounding induce the degradation of the VQ protein JAV1 via the 26S proteasome pathway in a COI1-dependent manner, but COI1 does not directly recruit JAV1 for degradation (Hu et al., 2013). The RING-type E3 ligase JUL1 interacts with and ubiquitinates JAV1, leading to its proteasomal degradation (Ali et al., 2019). *jul1* mutants have impaired resistance to *B. cinerea* and herbivorous insects, which is opposite the phenotypes of *JAV1* RNA interference lines (Ali et al., 2019; Hu et al., 2013). In this study, we established that the U-box-type E3 ligase OsPUB73 interacts with OsVQ25 and promotes OsVQ25 degradation via the UPS in rice. The *ospub73* mutants had reduced resistance to *M. oryzae* and *Xoo*, which was opposite the *osvq25* phenotype. Our results suggest that OsVQ25 is a substrate of OsPUB73 in rice, and it comprises an example of E3 ligase-VQ protein module in crop plants. While the E3-type proteins SPL11, EBR1, and OsCUL3a negatively regulate SNS BSR (Liu et al., 2017; You et al., 2016; Zeng et al., 2004), we demonstrated here that OsPUB73 positively regulates SNS BSR. Furthermore, OsPUB73 is required for anther development, and *ospub73* knockouts in rice display low pollen fertility (Chen et al., 2019), suggesting that OsPUB73 positively regulates rice development. These results indicate that OsPUB73 modulates OsVQ25 at the posttranslational level to regulate the balance between immunity and growth.

VQ proteins interact with WRKY TFs and affect downstream processes in *Arabidopsis* (Jing and Lin, 2015). For instance, JAV1 interacts with WRKY28 and WRKY51, functioning as a positive regulator of WRKY28 and as a negative

regulator of WRKY51 (Hu et al., 2013). JAV1 and JAZ8 interact with WRKY51 to form a so-called JJW (JAV1-JAZ8-WRKY51) complex that represses the expression of JA biosynthesis genes. JAV1 phosphorylation, which is triggered by injury, causes the JJW complex to disintegrate, thus alleviating the prior repression and leading to the activation of JA biosynthesis (Yan et al., 2018). In rice, expression pattern analysis in different tissues and developmental stages indicated that 12 out of 40 *VQ* genes are co-expressed with 20 *WRKY* genes (Li et al., 2014), but no studies have reported that VQs and WRKYs interact with each other. Here, we showed that OsVQ25 interacts with OsWRKY53 and suppresses OsWRKY53 transcriptional activity. OsWRKY53 is a typical transcription factor (Chujo et al., 2007). *OsWRKY53* expression is induced by treatment with the oligosaccharide elicitor chitin and by *M. oryzae* infection. *OsWRKY53* overexpression increases resistance to *M. oryzae* (Chujo et al., 2007, 2014). We determined that loss of *OsWRKY53* function decreases the resistance to *M. oryzae*. Furthermore, the expression of defenserelated genes *PR-5* and *Chitinase*, which are positively regulated by OsWRKY53, increased in *osvq25* mutants. In addition, the BR-responsive genes *OsBU1* and *OsXTR1*, also positively regulated by OsWRKY53 (Tian et al., 2017), were also upregulated in *osvq25*. *oswrky53* mutant plants were slightly dwarf and had more erect leaves, but *OsWRKY53* overexpression results in enlarged leaf angles and significantly decreased plant height, which did not occur in *osvq25* mutant lines (Xie et al., 2021). It is possible that the moderately increased expression levels of BR-responsive genes in *osvq25* mutants maintain, but are not enough to alter, the key agronomic traits of *osvq25* mutants in the field. Thus, our results demonstrate that OsVQ25 modulates the transcriptional activity of OsWRKY53 to balance immunity and growth in rice. Remarkably, OsWRKY53 negatively regulates rice resistance to *Xoo* (Xie et al., 2021). *OsWRKY53* overexpression reduced resistance to *Xoo*, while its knockout enhanced the strength of resistance. However, defense-related genes, including *Chitinase*, are induced in *OsWRKY53*-overexpression plants (Xie et al., 2021), which is consistent with the report that OsWRKY53 positively regulates *M. oryzae* resistance (Chujo et al., 2007, 2014), indicating that the resistance mechanisms mediated by OsWRKY53 differ for *M. oryzae* and *Xoo*. Therefore, there might be another pathway regulated by OsVQ25 that positively regulates resistance to *Xoo*.

We propose the following regulatory model of how OsPUB73, OsVQ25, and OsWRKY53 interact to balance immunity and growth in rice (Figure 5-6). OsVQ25 negatively regulates SNS BSR to *M. oryzae*. The U-box-type E3 ligase OsPUB73 interacts with and promotes OsVQ25 degradation via the UPS to positively regulate SNS BSR. By reducing the transcriptional activity of OsWRKY53, OsVQ25 suppresses downstream defense signaling and BR signaling, balancing plant defense responses and growth in rice. Our results highlight the hierarchical regulatory mechanism of an OsPUB73-OsVQ25-OsWRKY53 module that balances BSR and plant growth in rice, proposing a new insight for breeding SNS BSR rice varieties as well as other agriculturally important crop plants for improved food security and sustainable agriculture.

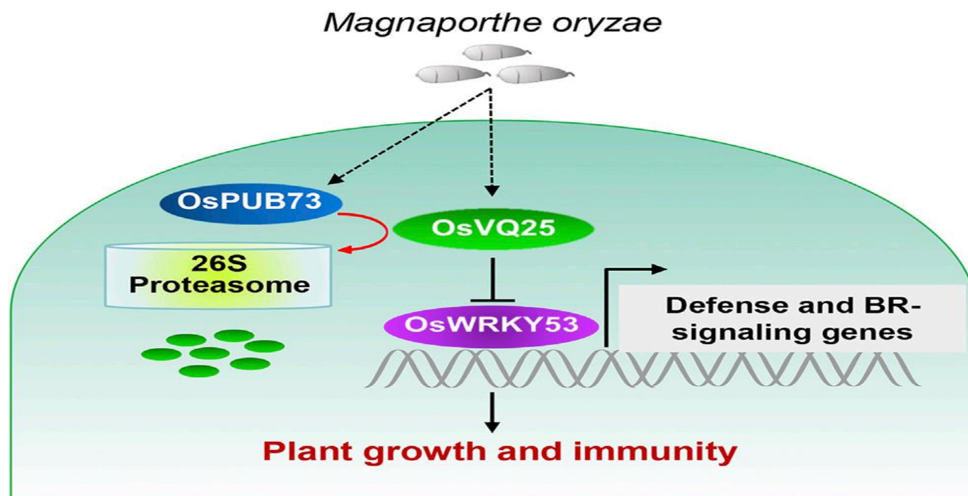


Figure 5-6. Working model of the role of OsVQ25 in balancing plant defense and growth.
OsVQ25 negatively regulates SNS BSR. OsPUB73 promotes OsVQ25 degradation via the 26S proteasome-dependent pathway to positively regulate SNS BSR at the posttranslational level. *OsVQ25* and *OsPUB73* expression is induced by *M. oryzae* at the transcriptional level. OsVQ25 then suppresses the transcriptional activity of OsWRKY53, which consequently reduces the expression of OsWRKY53-downstream defenseand BR-signaling-related genes, to balance plant growth and defense responses.

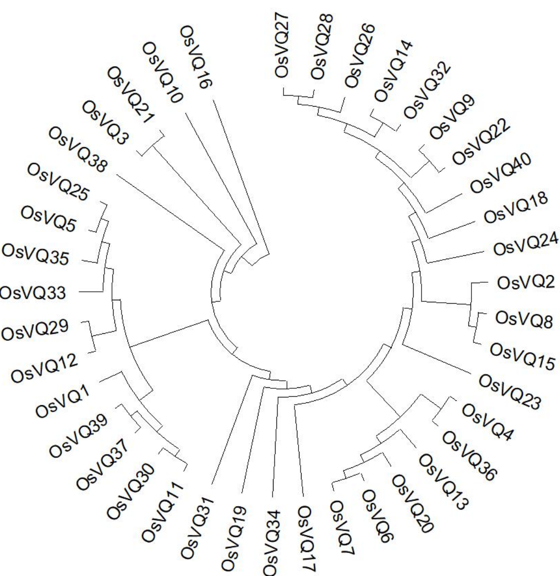


Figure 5-S7. Phylogenic tree of VQ domain proteins in rice. Related to Figure 5-6.
The full-length amino acid sequences of the 40 VQ proteins in rice were aligned by ClustalW, from which a phylogenic tree of VQ proteins was constructed, using the neighbor-joining method with 1,000 bootstrap replicates.

Table 5-S1. Primers used in the paper.

Primer Name	Primer Sequence （5’-3’）
qRT-PCR	
Q-OsPUB73-Forward	ATTTCATCGAGCGGTTGCGT
Q-OsPUB73-Rreverse	CTCCAGAAAGTTAAGCGAGCA
Q-OsVQ25-Forward	CCTCCAGTCAGCTCAGAT
Q-OsVQ25-Reverse	AACACAACTGGTCCATGAAT
Q-OsWRKY53-Forward	GAGCGACATCGACATCCT
Q-OsWRKY53-Reverse	TTGTGCTTGCCCTCGTAG
Q-OsPBZ1-Forward	TGGCATGCTCAAGATGATCGAGGA
Q-OsPBZ1-Reverse	TTACTCTCACGGACTCAAACGCCA
Q-OsPR1-Forward	CTGGTGGAGGGCGGCGGCAT
Q-OsPR1-Reverse	CCGGCTCGCCGGCGACCAT
Q-WRKY45-Forward	CGGGTAAAACGATCGAAAGA
Q-WRKY45-Reverse	GACCCCCAGCTCATAATCAA
Q-Chitinase-Forward	GCTACGCCTACGAACCATTC
Q-Chitinase-Reverse	GTCCGGTCGGTGTACATTCT
Q-PR-5-Forward	GCATTAGCTGGCTGCTATAGAT
Q-PR-5-Reverse	CCATGGACGATTATTATCTTATTATTT
Q-OsBU1-Forward	ATCTCCAAGCTCCAGTCCCT
Q-OsBU1-Reverse	GCTCTTGATGTAGCTGCACG
Q-OsXTR1-Forward	GGAGCCGTACATCCTGCAGA
Q-OsXTR1-Reverse	AGGCTGGAGTAGAGCTTCATCG
Q-UBQ-Forward	CGCAAGAAGAAGTGTGGTCA
Q-UBQ-Reverse	GGGAGATAACAACGGAAGCA
Mutant identification	
OsVQ25-Identify-Forward	GATGTCATGATCGGACGGCT
OsVQ25-Identify-Reverse	ACACGGACACCACCAACTT
OsPUB73-Identify-Forward	GCTTGTATCCGCTGCCATTGC
OsPUB73-Identify-Reverse	CTAAAACAAACAGGAACAAAG
Vector detection	
Cas9-F	TCGACAAGAAGTACTCCATCGGC
Cas9-R	CAAGAGAGAGGGCGATCAGGTTG
sgRNA-F	AAGGAATCTTTAAACATACGAACAGATC
sgRNA-R	ACTTTTTCAAGTTGATAACGG
Hpt-F	GAGGGCGTGGATATGTCCTG
Hpt-R	ATTGACCGATTCCTTGCGGT

6

Chapter VI General discussion, conclusions and perspectives

VI.1. CRISPR/CaS9-mediated genome editing

The CRISPR/Cas9 system is currently the most widely used genome editing tool. We have summarized the working principle, current research and application status, existing problems, and research prospects of CRISPR/Cas9 system. Gene function has been studied by creating series of mutants via CRISPR/Cas9 system in animals such as humans, mice, and pigs, as well as in plants such as *Arabidopsis*, rice, maize, and wheat. By utilizing the characteristic of Cas9 protein targeted cleavage of target gene sequences under the guidance of gRNA, multiple strategies can be applied to cause insertion, deletion, and replacement of bases or fragments at the target site, thereby triggering gene mutations and changes in biological traits. CRISPR/Cas9 system has the advantages of high accuracy, simple operation, low cost, and the ability to edit multiple target sites simultaneously.

In the potential applications of plant genetics and breeding, due to the separation of transgenic vector elements and Cas9 protein cleavage targets in the CRISPR/Cas9 system, transgenic plants can be backcrossed or hybridized in the later stage to cultivate plants with target genes that have been edited but do not contain any transgenic elements, which can further improve the application safety of CRISPR/Cas9 genome editing technology. On the other hand, homologous recombination strategy of CRISPR/Cas9 system can also integrate exogenous fragments or genes in a targeted manner, avoiding the mutations caused by random insertion. At present, in the research of plant functional genes, the CRISPR/Cas9 system has become very mature in the genome editing of *Arabidopsis*, and its application in other plants has also shown explosive development, gradually becoming a powerful tool for creating gene functional mutants and studying target sites, such as promoter regions.

Although A technology has achieved many results, there are still some areas to improve. For example, although the CRISPR/Cas9 system is increasingly widely and efficiently applied in some diploid plants, there are problems such as low knockout and point mutation efficiency in the field of genome editing in polyploid plants such as wheat. Although some Cas9 variants and methods have fewer or even undetectable off-target mutations (Kleinstiver et al., 2016; Manghwar et al., 2020; Guo et al., 2023), it is still necessary to detect whether there is off-target events after obtaining the target mutant in plants. In addition to the limitations of complex genomes and low transgenic efficiency, how to further improve the efficiency of base substitution, fragment targeted insertion, and large fragment deletion is also a challenge that scientists are tackling in the field of CRISPR/Cas9 genome editing. The development and research of new or optimized CRISPR/Cas9 genome editing tools, such as smaller Cas9 protein, higher efficiency and specificity, and wider PAM recognition sites, will further promote the application of CRISPR/Cas9 system in animals and plants.

In recent years, with the continuous advancement of technology, the CRISPR/Cas9 system has been continuously improved. The off-target rate of its different variants has decreased, the efficiency of targeted gene editing has been

continuously improved, and the targeting range has gradually expanded. It has gradually become the most powerful tool in the field of gene editing. CRISPR/Cas9 genome editing technology has laid a solid foundation for gene targeted knockout, replacement, insertion, and gene function research, becoming a powerful tool in the field of animal and plant variety improvement and disease treatment, bringing revolutionary breakthroughs to human development.

VI.2. The applications and limitations of *VQ* family genes in plants

VQ proteins are a class of plant specific regulatory proteins that widely exist in many species in the form of gene families. With the rapid development of science and technology, scientists have gradually conducted a series of experimental and bioinformatics studies on the function and mechanism of *VQ* genes. We reviewed the research progress of plant *VQ* family genes, including gene structure, gene function, and pathways of action. The results show that by interacting with MAPK, WRKY, and other proteins, VQ gene family not only participates in the stress response to drought, salt, temperature, the defense response to pathogens and pests, but also participates in regulating various life processes of plants. Based on the published research results, we summarized a cascade of *VQ* genes involved in the plant's response to external stimuli, called 'MAPK-E3-VQ-WRKY'. We can see that VQ protein plays an important bridging role in this cascade, ensuring normal plant life activities and responding to external stimuli. In view of the importance of *VQ* genes in growth and development, stress response, and plant defense, as well as their unique molecular characteristics, we proposed a molecular design breeding strategy based on *VQ* genes and CRISPR/Cas9 system. Through big-data analysis, *VQ* genes with biological functions were preliminarily screened, and key loci on *VQ* genes were analyzed. Subsequently, CRISPR/Cas9-mediated genome editing technology is used to create a series of knockout, replacement, insertion, or deletion mutants. Through molecular and phenotypic identification, a series of new materials with different agronomic traits are created. This is of great significance for crop breeding. In summary, the specific biological functions and working mechanisms of the *VQ* family genes are worthy of further research and utilization in the future.

We further conducted large-scale identification and analysis of *VQ* family genes in various plants, including gymnosperms, angiosperms, and mosses, to explore the molecular and evolutionary characteristics of *VQ* genes in a larger range of plants. We identified a total of 2469 *VQ* genes from 56 plant species and conducted sequence similarity, gene structure, GC content, phylogenetic tree, and gene expression analysis of these *VQ* genes. Many gymnosperms have high economic value, medicinal value, ornamental value, and evolutionary research value. They often encounter various stress and invasion of diseases and pests in the external environment at different developmental stages, resulting in a loss of their growth adaptability. Prior to this, no one had systematically studied the *VQ* gene family in

gymnosperms, and there were almost no reports on the *VQ* genes in gymnosperms. Therefore, studying the characteristics, evolution, and function of the *VQ* family genes in gymnosperms is urgent and of great significance. In this article, for the first time, we conducted detailed molecular and evolutionary analysis of the *VQ* family genes in three gymnosperms, and predicted some candidate functional *VQ* genes in these three gymnosperms according to sequence similarity with known functional *VQ* genes in *Arabidopsis*. These results has laid a solid foundation for the further research of *VQ* family genes in gymnosperms. Our study clearly visualized and displayed the structural composition and molecular characteristics of the *VQ* family genes in plants, allowing us to better understand their basic information. In conclusion, our study has provided a good reference for further research and utilization of *VQ* gene resources to improve specific traits for various plants.

However, there are still many limitations in current research on the *VQ* gene. Firstly, preliminary research on the functions and mechanisms of certain members of the VQ family is currently limited to *Arabidopsis*, and there is little research on crops, such as rice, maize, and wheat, and specific in-depth research has not been reported. As important food crops, the functions and mechanisms of *VQ* genes in these plants deserve further research. Secondly, due to the large scale of the *VQ* gene family and the fact that its proteins are only conserved in the VQ domain, there are significant differences in the similarity of *VQ* family genes among different plant species. Therefore, further extensive analysis is needed to determine whether these different *VQ* genes truly have biological functions and what functions they possess. Thirdly, there are differences and commonalities in the expression patterns of different *VQ* family genes, which to some extent indicates the diversity and redundancy of *VQ* gene functions. Sufficient experimental evidence is needed to verify this. Fourthly, currently, research on the working mechanism of VQ proteins mainly focuses on their interactions with WRKY transcription factors, while its interactions with other proteins remain to be further revealed.

In summary, we have conducted a comprehensive and detailed review and analysis of the *VQ* gene family, including its structure, function, and mechanism, providing sufficient reference for more researchers. *VQ* family genes are very important in plants, and although we have made some achievements in their research, there are still many questions and unknowns that need further research and answers. With the advancement of technology and in-depth research, we believe that the *VQ* gene will be more fully studied and utilized in more plants.

VI.3. Limitations and the prospects of the research on *OsVQ25* in rice

We identified *OsVQ25* as a candidate gene through identification and screening. Next, we cloned *OsVQ25* from rice and constructed a CRISPR/Cas9 gene knockout vector for *OsVQ25*. The CRISPR/Cas9 genome editing vector was transformed into

rice mature embryo callus using the *Agrobacterium tumefaciens* transformation method. After screening and identification, series of *OsVQ25* mutants were obtained. Functional studies have shown that compared to wild-type rice plants, *OsVQ25* homozygous mutants exhibit significant resistance improvement to rice blast (*M. oryzae*) and bacterial blight (*Xoo*). We have obtained a series of transgene-free rice materials with improved disease resistance. Further mechanism analysis indicates that OsVQ25 can interact with OsPUB73 and OsWRKY53. The ubiquitin ligase OsPUB73 can ubiquitinize and modify OsVQ25, leading to the degradation of OsVQ25, thereby releasing OsWRKY53 to regulate downstream BR-related genes. OsVQ25 plays an important role in balancing plant growth and disease resistance.

However, there are also many shortcomings and deficiencies in this study. Although we demonstrated that OsVQ25 is a substrate of the E3 ligase OsPUB73, we cannot exclude the possibility that there are other substrates of OsPUB73, or that other E3 ligases or pathways promote OsVQ25 degradation. OsVQ25 negatively regulates resistance to *M. oryzae* and *Xoo* and hinders the transcriptional activity of OsWRKY53 to suppress *M. oryzae* resistance in rice, but the molecular mechanism by which OsVQ25 suppresses OsWRKY53 transcriptional activity remains unclear. In addition, our model only presents how OsVQ25 regulates *M. oryzae* resistance in rice. It is likely that there are other target proteins or mechanisms by which OsVQ25 modulates *Xoo* resistance. Future study is needed to investigate the OsVQ25 downstream signaling cascade to better understand the molecular mechanism of OsVQ25-mediated SNS BSR. In addition, except for *OsVQ25*, rice also has 39 *VQ* genes, so it is necessary to further develop more functional *VQ* genes in rice.

Given the important role of *VQ* genes in growth and development, as well as in response to biotic or abiotic stress, the exploration and development of some functional *VQ* genes in crops is of great significance. Our study indicates that rice *OsVQ25* mutant significantly improves its broad-spectrum disease resistance without causing growth loss. This indicates that the some *VQ* genes of crops have enormous breeding potential. We can further search for *VQ* genes with similar or more powerful functions in crops such as corn and wheat, and conduct in-depth research on their specific biological functions and mechanisms of action. This can improve plant traits without compromising yield and development, which is of great significance for the potential application of *VQ* gene resources in crop breeding.

In summary, this article summarizes the working principle and practical application of CRISPR/Cas9 system. Reviewing the research progress of *VQ* family genes in plants, including molecular characteristics, roles in growth and development, responses to biotic or abiotic stresses, working mechanisms, and the importance of VQ domain. And based on these characteristics of the *VQ* gene, we proposed a molecular design breeding strategy. In addition, large-scale identification and analysis of the members, molecular and evolutionary characteristics of the *VQ* family genes were conducted in 56 plant species, especially gymnosperms. Some candidate functional *VQ* genes were preliminarily

identified for three gymnosperms through bioinformatics methods, but further experiments are needed to verify their specific functions in the future. Furthermore, we identified a VQ protein, OsVQ25, related to broad-spectrum plant disease resistance in rice, and its mutant showed significantly improved resistance to rice blast (*M. oryzae*) and bacterial blight (*Xoo*). We studied its related disease resistance mechanisms and regulatory pathways, and ultimately created a series of new rice disease resistant germplasms. These studies and findings in this article provide strong references and important directions for characterization of the *VQ* family genes in plants, as well as the exploration and utilization of *VQ* genetic resources in plant, especially crops.

References

Abe, F., Haque, E., Hisano, H., Tanaka, T., Kamiya, Y., Mikami, M., et al. (2019). Genome-Edited Triple-Recessive Mutation Alters Seed Dormancy in Wheat. Cell Rep. 28(5), 1362-1369.e4.

Adams, K.L., Wendel, J.F. (2005). Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. 8, 135-141.

Ali, M. R. M., Uemura, T., Ramadan, A., Adachi, K., Nemoto, K., Nozawa, A., et al. (2019). The Ring-Type E3 Ubiquitin Ligase JUL1 Targets the VQ-Motif Protein JAV1 to Coordinate Jasmonate Signaling. Plant Physiol. 179(4), 1273-1284.

Andreasson, E., Jenkins, T., Brodersen, P., Thorgrimsen, S., Petersen, N. H., Zhu, S., et al. (2005). The MAP kinase substrate MKS1 is a regulator of plant defense responses. EMBO J. 24(14), 2579-2589.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science 373(6557), 871-876.

Bailey, T. L., Johnson, J., Grant, C. E., Noble, W. S. (2015). The MEME Suite. Nucleic Acids Res. 43(W1), W39-W49.

Bakshi, M., and Oelmüller, R. (2014). WRKY transcription factors: Jack of many trades in plants. Plant Signal. Behav. 9(2), e27700.

Bari, R., and Jones, J. D. (2009). Role of plant hormones in plant defence responses. Plant Mol. Biol. 69(4), 473-488.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science 315(5819), 1709-1712.

Bi, G., Zhou, Z., Wang, W., Li, L., Rao, S., Wu, Y., et al. (2018). Receptor-Like Cytoplasmic Kinases Directly Link Diverse Pattern Recognition Receptors to the Activation of Mitogen-Activated Protein Kinase Cascades in *Arabidopsis*. Plant Cell 30(7), 1543-1561.

Bigeard, J., and Hirt, H. (2018). Nuclear Signaling of Plant MAPKs. Front. Plant Sci. 9, 469.

Bolger, A.M., Lohse, M., Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.

Bolotin, A., Quinquis, B., Sorokin, A., Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151(Pt 8), 2551-2561.

Bolser, D. M., Staines, D. M., Perry, E., Kersey, P. J. (2017). Ensembl Plants: Integrating Tools for Visualizing, Mining, and Analyzing Plant Genomic Data. Methods Mol. Biol. 1533, 1-31.

Bryant, P., Pozzati, G., Elofsson, A. (2022). Improved prediction of protein-protein interactions using AlphaFold2. Nat. Commun. 13(1), 1265.

Budhagatapalli, N., Halbach, T., Hiekel, S., Büchner, H., Müller, A. E., Kumlehn, J. (2020). Site-directed mutagenesis in bread and durum wheat via pollination by

cas9/guide RNA-transgenic maize used as haploidy inducer. Plant Biotechnol. J. 18(12), 2376-2378.

Buscaill, P., and Rivas, S. (2014). Transcriptional control of plant defence responses. Curr. Opin. Plant Biol. 20, 35-46.

Caarls, L., Elberse, J., Awwanah, M., Ludwig, N. R., de Vries, M., Zeilmaker, T., et al. (2017). *Arabidopsis* JASMONATE-INDUCED OXYGENASES down-regulate plant immunity by hydroxylation and inactivation of the hormone jasmonic acid. Proc. Natl. Acad. Sci. U. S. A. 114(24), 6388-6393.

Cai, H., Zhang, M., Liu, Y., He, Q., Chai, M., Liu, L., et al. (2019). Genome-Wide Classification and Evolutionary and Functional Analyses of the VQ Family. Tropical Plant Biol. 12, 117-131.

Cao, Y., Meng, D., Abdullah, M., Jin, Q., Lin, Y., Cai, Y. (2018). Genome Wide Identification, Evolutionary, and Expression Analysis of VQ Genes from Two *Pyrus* Species. Genes 9(4), 224.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., Xia, R. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol. Plant 13, 1194-1202.

Chen, J., Wang, H., Li, Y., Pan, J., Hu, Y., Yu, D. (2018). *Arabidopsis* VQ10 interacts with WRKY8 to modulate basal defense against *Botrytis cinerea*. J. Integr. Plant Biol. 60(10), 956-969.

Chen, L., Deng, R., Liu, G., Jin, J., Wu, J., and Liu, X. (2019). Cytological and transcriptome analyses reveal OsPUB73 defect affects the gene expression associated with tapetum or pollen exine abnormality in rice. BMC Plant Biol. 19, 546.

Chen, L., Wu, L., Yang, L., Yu, H., Huang, P., Wang, Y., Yao, R., Zhang, M. (2022). TcJAV3-TcWRKY26 Cascade Is a Missing Link in the Jasmonate-Activated Expression of Taxol Biosynthesis Gene *DBAT* in *Taxus chinensis*. Int. J. Mol. Sci. 23, 13194.

Chen, P., Wei, F., Cheng, S., Ma, L., Wang, H., Zhang, M., et al. (2020). A comprehensive analysis of cotton *VQ* gene superfamily reveals their potential and extensive roles in regulating cotton abiotic stress. BMC genomics 21(1), 795.

Chen, S., Cao, H., Huang, B., Zheng, X., Liang, K., Wang, G. L., et al. (2022). The WRKY10-VQ8 module safely and effectively regulates rice thermotolerance. Plant Cell Environ. 45(7), 2126-2144.

Cheng, X., Gao, C., Liu, X., Xu, D., Pan, X., Gao, W., et al. (2022). Characterization of the wheat VQ protein family and expression of candidate genes associated with seed dormancy and germination. BMC Plant Biol. 22(1), 119.

Cheng, X., Wang, Y., Xiong, R., Gao, Y., Yan, H., Xiang, Y. (2020). A Moso bamboo gene *VQ28* confers salt tolerance to transgenic *Arabidopsis* plants. Planta 251(5), 99.

Cheng, Y., Zhou, Y., Yang, Y., Chi, Y. J., Zhou, J., Chen, J. Y., et al. (2012). Structural and functional analysis of VQ motif-containing proteins in *Arabidopsis* as interacting proteins of WRKY transcription factors. Plant Physiol. 159(2), 810-825.

Chi, Y., Yang, Y., Zhou, Y., Zhou, J., Fan, B., Yu, J. Q., et al. (2013).

Protein-protein interactions in the regulation of WRKY transcription factors. Mol. Plant 6(2), 287-300.

Chu, W., Liu, B., Wang, Y., Pan, F., Chen, Z., Yan, H., et al. (2016). Genome-wide analysis of poplar *VQ* gene family and expression profiling under PEG, NaCl, and SA treatments. Tree Genetics & Genomes 12, 124.

Chujo, T., Miyamoto, K., Ogawa, S., Masuda, Y., Shimizu, T., Kishi-Kaboshi, M., Takahashi, A., Nishizawa, Y., Minami, E., Nojiri, H., et al. (2014). Overexpression of phosphomimic mutated OsWRKY53 leads to enhanced blast resistance in rice. PLoS One 9, e98737.

Chujo, T., Takai, R., Akimoto-Tomiyama, C., Ando, S., Minami, E., Nagamura, Y., Kaku, H., Shibuya, N., Yasuda, M., Nakashita, H., et al. (2007). Involvement of the elicitor-induced gene OsWRKY53 in the expression of defense-related genes in rice. Biochim. Biophys. Acta 1769, 497-505.

Clarke, J.D. (2009). Cetyltrimethyl ammonium bromide (CTAB) DNA miniprep for plant DNA isolation. Cold Spring Harb. Protoc. 2009. pdb.prot5177.

Coll, N. S., Epple, P., Dangl, J. L. (2011). Programmed cell death in the plant immune system. Cell Death Differ. 18(8), 1247-1256.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science 339:819-823.

Dauparas, J., Anishchenko, I., Bennett, N., Bai, H., Ragotte, R. J., Milles, L. F., et al. (2022). Robust deep learning-based protein sequence design using ProteinMPNN. Science 378(6615), 49-56.

Deenamo, N., Kuyyogsuy, A., Khompatara, K., Chanwun, T., Ekchaweng, K., Churngchow, N. (2018). Salicylic Acid Induces Resistance in Rubber Tree against *Phytophthora palmivora*. Int. J. Mol. Sci. 19(7), 1883.

Ding, H., Yuan, G., Mo, S., Qian, Y., Wu, Y., Chen, Q., et al. (2019). Genome-wide analysis of the plant-specific VQ motif-containing proteins in tomato (*Solanum lycopersicum*) and characterization of SlVQ6 in thermotolerance. Plant Physiol. Biochem. 143, 29-39.

Dong, O. X., Yu, S., Jain, R., Zhang, N., Duong, P. Q., Butler, C., et al. (2020). Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. Nat. Commun. 11(1), 1178.

Dong, Q., Duan, D., Zheng, W., Huang, D., Wang, Q., Li, X., et al. (2021). *MdVQ37* overexpression reduces basal thermotolerance in transgenic apple by affecting transcription factor activity and salicylic acid homeostasis. *Hortic. Res*. 8 (1), 220.

Dong, Q., Zhao, S., Duan, D., Tian, Y., Wang, Y., Mao, K., et al. (2018). Structural and functional analyses of genes encoding VQ proteins in apple. Plant Sci. 272, 208-219.

Duplan, V., and Rivas, S. (2014). E3 ubiquitin-ligases and their target proteins during the regulation of plant innate immunity. Front. Plant Sci. 5, 42.

Eulgem, T., Rushton, P. J., Robatzek, S., Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. Trends Plant Sci. 5, 199-206.

Fan, J., Bai, P., Ning, Y., Wang, J., Shi, X., Xiong, Y., Zhang, K., He, F., Zhang, C., Wang, R., et al. (2018). The monocot-specific receptor-like kinase SDS2

controls cell death and immunity in rice. Cell Host Microbe 23, 498-510.e5.

Fang, H., Shen, S., Wang, D., Zhang, F., Zhang, C., Wang, Z., Zhou, Q., Wang, R., Tao, H., He, F., et al. (2021). A monocot-specific hydroxycinnamoylputrescine gene cluster contributes to immunity and cell death in rice. Sci. Bull. 66, 2381-2393.

Fiil, B. K., Petersen, M. (2011). Constitutive expression of *MKS1* confers susceptibility to *Botrytis cinerea* infection independent of *PAD3* expression. Plant Signal. Behav. 6(10), 1425-1427.

Fu, F., Song, C., Wen, C., Yang, L., Guo, Y., Yang, X., Shu, Z., Li, X., Feng, Y., Liu, B., et al. (2023). The Metasequoia genome and evolutionary relationship among redwoods. Plant Commun. 4, 100643.

Gao C. (2021). Genome engineering for crop improvement and future agriculture. Cell 184(6), 1621-1635.

Gargul, J. M., Mibus, H., Serek, M. (2015). Manipulation of *MKS1* gene expression affects *Kalanchoë blossfeldiana* and *Petunia hybrida* phenotypes. Plant Biotechnol. J. 13(1), 51-61.

Garrido-Gala, J., Higuera, J. J., Muñoz-Blanco, J., Amil-Ruiz, F., Caballero, J. L. (2019). The VQ motif-containing proteins in the diploid and octoploid strawberry. Sci. Rep. 9(1), 4942.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784-3788.

Gayubas, B., Castillo, M. C., Ramos, S., León, J. (2023). Enhanced meristem development, tolerance to oxidative stress and hyposensitivity to nitric oxide in the hypermorphic *vq10-H* mutant in *AtVQ10* gene. Plant Cell Environ. 46, 3445-3463.

Glazebrook, J., Chen, W., Estes, B., Chang, H. S., Nawrath, C., Métraux, J. P., et al. (2003). Topology of the network integrating salicylate and jasmonate signal transduction derived from global expression phenotyping. Plant J. 34(2), 217-228.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., et al. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40, D1178-D1186.

Goyal, P., Devi, R., Verma, B., Hussain, S., Arora, P., Tabassum, R., et al. (2023). WRKY transcription factors: evolution, regulation, and functional diversity in plants. Protoplasma 260(2), 331-348.

Guan, Y., Meng, X., Khanna, R., LaMontagne, E., Liu, Y., Zhang, S. (2014). Phosphorylation of a WRKY transcription factor by MAPKs is required for pollen development and function in *Arabidopsis*. PLoS Genet. 10(5), e1004384.

Guo, C., Ma, X., Gao, F., Guo, Y. (2023). Off-target effects in CRISPR/Cas9 gene editing. Front. Bioeng. Biotechnol. 11, 1143157.

Guo, J., Chen, J., Yang, J., Yu, Y., Yang, Y., Wang, W. (2018). Identification, characterization and expression analysis of the VQ motif-containing gene family in tea plant (*Camellia sinensis*). BMC genomics 19(1), 710.

Halim, V. A., Vess, A., Scheel, D., Rosahl, S. (2006). The role of salicylic acid and jasmonic acid in pathogen defence. Plant Biol. (Stuttg.) 8(3), 307-313.

Hao, Z., Tian, J., Fang, H., Fang, L., Xu, X., He, F., et al. (2022). A VQ-motif-containing protein fine-tunes rice immunity and growth by a hierarchical regulatory mechanism. Cell Rep. 40(7), 111235.

He, Q., He, M., Zhang, X., Zhang, X., Zhang, W., Dong, J., et al. (2023). RsVQ4-RsWRKY26 module positively regulates thermotolerance by activating RsHSP70-20 transcription in radish (*Raphanus sativus* L.). Environ. Exp. Bot. 214, 105467.

Henriques, R., Jang, I. C., Chua, N. H. (2009). Regulated proteolysis in light-related signaling pathways. Curr. Opin. Plant Biol. 12(1), 49-56.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35, W585-W587.

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556(7699), 57-63.

Hu, L., Ye, M., Li, R., Zhang, T., Zhou, G., Wang, Q., et al. (2015). The Rice Transcription Factor WRKY53 Suppresses Herbivore-Induced Defenses by Acting as a Negative Feedback Modulator of Mitogen-Activated Protein Kinase Activity. Plant Physiol. 169(4), 2907-2921.

Hu, P., Zhou, W., Cheng, Z., Fan, M., Wang, L., Xie, D. (2013a). JAV1 controls jasmonate-regulated plant defense. Mol. Cell 50(4), 504-515.

Hu, X., Meng, X., Liu, Q., Li, J., Wang, K. (2018). Increasing the efficiency of CRISPR-Cas9-VQR precise genome editing in rice. Plant Biotechnol. J. 16(1), 292-297.

Hu, Y., Chen, L., Wang, H., Zhang, L., Wang, F., Yu, D. (2013b). *Arabidopsis* transcription factor WRKY8 functions antagonistically with its interacting partner VQ9 to modulate salinity stress tolerance. Plant J. 74(5), 730-745.

Hua, K., Tao, X., Han, P., Wang, R., Zhu, J. K. (2019). Genome Engineering in Rice Using Cas9 Variants that Recognize NG PAM Sequences. Mol. Plant 12(7), 1003-1014.

Huang, H., Zhao, W., Li, C., Qiao, H., Song, S., Yang, R., et al. (2022a). SlVQ15 interacts with jasmonate-ZIM domain proteins and SlWRKY31 to regulate defense response in tomato. Plant Physiol. 190(1), 828-842.

Huang, J., Lin, Q., Fei, H., He, Z., Xu, H., Li, Y., et al. (2023). Discovery of deaminase functions by structure-based protein clustering. Cell 186(15), 3182-3195.e14.

Huang, X., Huang, S., Han, B., Li, J. (2022b). The integrated genomics of crop domestication and breeding. Cell 185(15), 2828-2839.

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., et al. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. 31(3), 227-229.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. 169(12), 5429-5433.

Jansen, R., Embden, J. D., Gaastra, W., Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. 43(6), 1565-1575.

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., Zhang, J. (2017). WRKY transcription factors in plant responses to stresses. J. Integr. Plant Biol. 59(2), 86-101.

Jiang, S. Y., Sevugan, M., Ramachandran, S. (2018). Valine-glutamine (VQ) motif coding genes are ancient and non-plant-specific with comprehensive expression regulation by various biotic and abiotic stresses. BMC genomics 19(1), 342.

Jiang, Y., Deyholos, M. K. (2009). Functional characterization of *Arabidopsis* NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. Plant Mol. Biol. 69(1-2), 91-105.

Jing, Y., Lin, R. (2015). The VQ Motif-Containing Protein Family of Plant-Specific Transcriptional Regulators. Plant Physiol. 169(1), 371-378.

Jiang, Y., Yu, D. (2016). The WRKY57 Transcription Factor Affects the Expression of Jasmonate ZIM-Domain Genes Transcriptionally to Compromise *Botrytis cinerea* Resistance. Plant Physiol. 171(4), 2771-2782.

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., Tomsho, L. P., Hu, Y., Liang, H., Soltis, P. S., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature 473, 97-100.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096), 816-821.

Jones, J. D., Dangl, J. L. (2006). The plant immune system. Nature 444(7117), 323-329.

Kaltenegger, E., Leng, S., Heyl, A. (2018). The effects of repeated whole genome duplication events on the evolution of cytokinin signaling pathway. BMC Evol. Biol. 18, 76.

Katoh, K., Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772-780.

Ke, Y., Deng, H., Wang, S. (2017). Advances in understanding broad-spectrum resistance to pathogens in rice. Plant J. 90(4), 738-748.

Khoso, M. A., Hussain, A., Ritonga, F. N., Ali, Q., Channa, M. M., Alshegaihi, R. M., et al. (2022). WRKY transcription factors (TFs): Molecular switches to regulate drought, temperature, and salinity stresses in plants. Front. Plant Sci. 13, 1039329.

Kim, D., Paggi, J.M., Park, C., Bennett, C., Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37, 907-915.

Kim, D. Y., Kwon, S. I., Choi, C., Lee, H., Ahn, I., Park, S. R., et al. (2013). Expression analysis of rice *VQ* genes in response to biotic and abiotic stresses. Gene 529(2), 208-214.

Kim, K. D., El Baidouri, M., Abernathy, B., Iwata-Otsubo, A., Chavarro, C.,

Gonzales, M., Libault, M., Grimwood, J., Jackson, S. A. (2015). A Comparative Epigenomic Analysis of Polyploidy-Derived Genes in Soybean and Common Bean. Plant Physiol. 168, 1433-1447.

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., Joung, J. K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529(7587), 490–495.

Kou, Y., and Wang, S. (2010). Broad-spectrum and durability: understanding of quantitative disease resistance. Curr. Opin. Plant. Biol. 13, 181-185.

Kovach, A., Wegrzyn, J. L., Parra, G., Holt, C., Bruening, G. E., Loopstra, C. A., Hartigan, J., Yandell, M., Langley, C. H., Korf, I., et al. (2010). The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. BMC Genom. 11, 420.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., Marra, M. A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. 19, 1639-1645.

Kulcsár, P. I., Tálas, A., Tóth, E., Nyeste, A., Ligeti, Z., Welker, Z., et al. (2020). Blackjack mutations improve the on-target activities of increased fidelity variants of SpCas9 with 5'G-extended sgRNAs. Nat. Commun. 11(1), 1223.

Kumari, A., Kaladhar, V. C., Yadav, N., Singh, P., Reddy, K., Gupta, K. J. (2023). Nitric oxide regulates mitochondrial biogenesis in plants. Plant Cell Environ. 46(8), 2492-2506.

Lai, Z., Li, Y., Wang, F., Cheng, Y., Fan, B., Yu, J. Q., et al. (2011). *Arabidopsis* sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. Plant Cell 23(10), 3824-3841.

Lai, Z., Mengiste, T. (2013). Genetic and cellular mechanisms regulating plant responses to necrotrophic pathogens. Curr. Opin. Plant Biol. 16(4), 505-512.

Lan, X., Wang, X., Tao, Q., Zhang, H., Li, J., Meng, Y., et al. (2022). Activation of the VQ Motif-Containing Protein Gene *VQ28* Compromised Nonhost Resistance of *Arabidopsis thaliana* to *Phytophthora* Pathogens. Plants 11(7), 858.

Langmead, B., Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357-359.

Lei, R., Li, X., Ma, Z., Lv, Y., Hu, Y., Yu, D. (2017). *Arabidopsis* WRKY2 and WRKY34 transcription factors interact with VQ20 protein to modulate pollen development and function. Plant J. 91(6), 962-976.

Lei, R., Ma, Z., Yu, D. (2018). WRKY2/34-VQ20 Modules in *Arabidopsis thaliana* Negatively Regulate Expression of a Trio of Related MYB Transcription Factors During Pollen Development. Front. Plant Sci. 9, 331.

León, J., Gayubas, B., Castillo, M. C. (2021). Valine-Glutamine Proteins in Plant Responses to Oxygen and Nitric Oxide. Front. Plant Sci. 11, 632678.

Letunic, I., Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293-W296.

Letunic, I., Khedkar, S., Bork, P. (2021). SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. 49, D458-D460.

Li, B., Dewey, C.N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform. 12, 323.

Li, J., Han, F., Yuan, T., Li, W., Li, Y., Wu, H.X., Wei, H., Niu, S. (2023). The methylation landscape of giga-genome and the epigenetic timer of age in Chinese pine. Nat. Commun. 14, 1947.

Li, J., Sun, Y., Du, J., Zhao, Y., Xia, L. (2017). Generation of Targeted Point Mutations in Rice by a Modified CRISPR/Cas9 System. Mol. Plant. 10(3), 526-529.

Li, J., Zhang, H., Si, X., Tian, Y., Chen, K., Liu, J., et al. (2017). Generation of thermosensitive male-sterile maize by targeted knockout of the *ZmTMS5* gene. J. Genet. Genomics 44(9), 465-468.

Li, N., Yang, Z., Li, J., Xie, W., Qin, X., Kang, Y., et al. (2021). Two VQ Proteins are Substrates of the OsMPKK6-OsMPK4 Cascade in Rice Defense Against Bacterial Blight. Rice 14(1), 39.

Li, N., Li, X., Xiao, J., Wang, S. (2014b). Comprehensive analysis of VQ motif-containing gene expression in rice defense responses to three pathogens. Plant Cell Rep. 33(9), 1493-1505.

Li, S., Fu, Q., Chen, L., Huang, W., Yu, D. (2011). *Arabidopsis thaliana* WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. Planta 233(6), 1237-1252.

Li, S., Xia, L. (2019). Precise gene replacement in plants through CRISPR/Cas genome editing technology: current status and future perspectives. aBIOTECH, 1(1), 58–73.

Li, W., Deng, Y., Ning, Y., He, Z., and Wang, G.L. (2020). Exploiting broadspectrum disease resistance in crops: from molecular dissection to breeding. Annu. Rev. Plant Biol. 71, 575-603.

Li, W., Pang, S., Lu, Z., Jin, B. (2020). Function and Mechanism of WRKY Transcription Factors in Abiotic Stress Responses of Plants. Plants 9(11), 1515.

Li, X., Qin, R., Du, Q., Cai, L., Hu, D., Du, H., et al. (2020). Knockdown of *GmVQ58* encoding a VQ motif-containing protein enhances soybean resistance to the common cutworm (*Spodoptera litura* Fabricius). J. Exp. Bot. 71(10), 3198-3210.

Li, Y., Jing, Y., Li, J., Xu, G., Lin, R. (2014a). *Arabidopsis* VQ MOTIF-CONTAINING PROTEIN29 represses seedling deetiolation by interacting with PHYTOCHROME-INTERACTING FACTOR1. Plant Physiol. 164(4), 2068-2080.

Li, Z., Baniaga, A.E., Sessa, E.B., Scascitelli, M., Graham, S.W., Rieseberg, L.H., Barker, M. S. (2015). Early genome duplications in conifers and other seed plants. Sci. Adv. 1, e1501084.

Ling, L., Qu, Y., Zhu, J., Wang, D., Guo, C. (2020). Genome-wide identification and expression analysis of the *VQ* gene family in *Cicer arietinum* and *Medicago truncatula*. PeerJ 8, e8471.

Liu, C., Liu, H., Zhou, C., Timko, M. P. (2020). Genome-Wide Identification of the VQ Protein Gene Family of Tobacco (*Nicotiana tabacum* L.) and Analysis of Its Expression in Response to Phytohormones and Abiotic and Biotic Stresses. Genes 11(3), 284.

Liu, H., Wang, X., Wang, G., Cui, P., Wu, S., Ai, C., Hu, N., Li, A., He, B., Shao,

X., et al. (2021). The nearly complete genome of *Ginkgo biloba* illuminates gymnosperm evolution. Nat. Plants 7, 748-756.

Liu, J., Park, C.H., He, F., Nagano, M., Wang, M., Bellizzi, M., Zhang, K., Zeng, X., Liu, W., Ning, Y., et al. (2015). The RhoGAP SPIN6 associates with SPL11 and OsRac1 and negatively regulates programmed cell death and innate immunity in rice. PLoS Pathog. 11, e1004629.

Liu, M., Li, C., Li, Y., An, Y., Ruan, X., Guo, Y., et al. (2023). Genome-Wide Identification and Characterization of the VQ Motif-Containing Gene Family Based on Their Evolution and Expression Analysis under Abiotic Stress and Hormone Treatments in Foxtail Millet (*Setaria italica* L.). Genes 14(5), 1032.

Liu, M., Shi, Z., Zhang, X., Wang, M., Zhang, L., Zheng, K., Liu, J., Hu, X., Di, C., Qian, Q., et al. (2019). Inducible overexpression of Ideal Plant Architecture1 improves both yield and disease resistance in rice. Nat. Plants 5, 389-400.

Liu, Q., Ning, Y., Zhang, Y., Yu, N., Zhao, C., Zhan, X., Wu, W., Chen, D., Wei, X., Wang, G.L., et al. (2017). OsCUL3a negatively regulates cell death and immunity by degrading OsNPR1 in rice. Plant Cell 29, 345-359.

Liu, S., Hua, L., Dong, S., Chen, H., Zhu, X., Jiang, J., et al. (2015). OsMAPK6, a mitogen-activated protein kinase, influences rice grain size and biomass production. Plant J. 84(4), 672-681.

Liu, S., Wang, Z., Wu, J., Wu, C., Xiong, R., Xiang, Y., et al. (2022b). The poplar *VQ1* gene confers salt tolerance and pathogen resistance in transgenic *Arabidopsis* plants via changes in hormonal signaling. G3 (Bethesda, Md.) 12(4), jkac044.

Liu, Y., Liu, X., Yang, D., Yin, Z., Jiang, Y., Ling, H., et al. (2022a). A Comprehensive Identification and Expression Analysis of VQ Motif-Containing Proteins in Sugarcane (*Saccharum spontaneum* L.) under Phytohormone Treatment and Cold Stress. Int. J. Mol. Sci. 23(11), 6334.

Liu, Y., Wang, S., Li, L., Yang, T., Dong, S., Wei, T., Wu, S., Liu, Y., Gong, Y., Feng, X., et al. (2022). The *Cycas* genome and the early evolution of seed plants. Nat. Plants 8, 389-401.

Luo, J., Li, S., Xu, J., Yan, L., Ma, Y., Xia, L. (2021). Pyramiding favorable alleles in an elite wheat variety in one generation by CRISPR-Cas9-mediated multiplex gene editing. Molecular plant, 14(6), 847-850.

Luo, M., Dennis, E. S., Berger, F., Peacock, W. J., Chaudhury, A. (2005). *MINISEED3* (*MINI3*), a WRKY family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A. 102(48), 17531-17536.

Lutz, I. D., Wang, S., Norn, C., Courbet, A., Borst, A. J., Zhao, Y. T., et al. (2023). Top-down design of protein architectures with reinforcement learning. Science 380(6642), 266-273.

Ma, H., Chen, J., Zhang, Z., Ma, L., Yang, Z., Zhang, Q., et al. (2017). MAPK kinase 10.2 promotes disease resistance and drought tolerance by activating different MAPKs in rice. Plant J. 92(4), 557-570.

Ma, J., Li, C., Sun, L., Ma, X., Qiao, H., Zhao, W., et al. (2023a). The SlWRKY57-SlVQ21/SlVQ16 module regulates salt stress in tomato. J. Integr.

Plant Biol. 10.1111/jipb.13562. Advance online publication.

Ma, J., Ling, L., Huang, X., Wang, W., Wang, Y., Zhang, M., et al. (2021). Genome-wide identification and expression analysis of the *VQ* gene family in sunflower (*Helianthus annuus* L.). J. Plant Biochem. Biotechnol. 30, 56-66.

Ma, J., Wang, R., Zhao, H., Li, L., Zeng, F., Wang, Y., et al. (2023b). Genome-wide characterization of the *VQ* genes in Triticeae and their functionalization driven by polyploidization and gene duplication events in wheat. Int. J. Biol. Macromol. 243, 125264.

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., et al. (2015). A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. Mol. Plant 8(8), 1274-1284.

Madani, A., Krause, B., Greene, E. R., Subramanian, S., Mohr, B. P., Holton, J. M., et al. (2023). Large language models generate functional protein sequences across diverse families. Nat. Biotechnol. 41(8), 1099-1106.

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct 1, 7.

Manghwar, H., Li, B., Ding, X., Hussain, A., Lindsey, K., Zhang, X., Jin, S. (2020). CRISPR/Cas Systems in Genome Editing: Methodologies and Tools for sgRNA Design, Off-Target Evaluation, and Strategies to Mitigate Off-Target Effects. Adv. Sci. 7(6), 1902312.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., et al. (2015). CDD: NCBI's conserved domain database. Nucleic Acids Res. 43, D222-D226.

Meng, X., Lu, M., Xia, Z., Li, H., Liu, D., Li, K., et al. (2023). Wheat VQ Motif-Containing Protein VQ25-A Facilitates Leaf Senescence via the Abscisic Acid Pathway. Int. J. Mol. Sci. 24, 13839.

Meng, X., Zhang, S. (2013). MAPK cascades in plant disease resistance signaling. Annu. Rev. Phytopathol. 51, 245-266.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., et al. (2021). Pfam: The protein families database in 2021. Nucleic Acids Res. 49, D412–D419.

Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J., Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60(2), 174-182.

Molla, K. A., Sretenovic, S., Bansal, K. C., Qi, Y. (2021). Precise plant genome editing using base editors and prime editors. Nat. Plants 7(9), 1166-1187.

Morikawa, K., Shiina, T., Murakami, S., Toyoshima, Y. (2002). Novel nuclear-encoded proteins interacting with a plastid sigma factor, Sig1, in *Arabidopsis thaliana*. FEBS Lett. 514(2-3), 300-304.

Mourier, T., Jeffares, D. C. (2003). Eukaryotic intron loss. Science 300, 1393.

Narusaka, M., Kawai, K., Izawa, N., Seki, M., Shinozaki, K., Seo, S., et al. (2008). Gene coding for SigA-binding protein from *Arabidopsis* appears to be

transcriptionally up-regulated by salicylic acid and NPR1-dependent mechanisms. J. Gen. Plant Pathol. 74, 345-354.

Nelson, R., Wiesner-Hanks, T., Wisser, R., and Balint-Kurti, P. (2018). Navigating complexity to breed disease-resistant crops. Nat. Rev. Genet. 19, 21-33.

Nguyen Ba, A.N., Pogoutse, A., Provart, N., Moses, A.M. (2009). NLStradamus: A simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinform. 10, 202.

Ning, Y., Liu, W., and Wang, G.L. (2017). Balancing immunity and yield in crop plants. Trends. Plant. Sci. 22, 1069-1079.

Ning, Y., Wang, R., Shi, X., Zhou, X., and Wang, G.L. (2016). A layered defense strategy mediated by rice E3 ubiquitin ligases against diverse pathogens. Mol. Plant 9, 1096-1098.

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science, 361(6408), 1259-1262.

Niu, S., Li, J., Bo, W., Yang, W., Zuccolo, A., Giacomello, S., Chen, X., Han, F., Yang, J., Song, Y., et al. (2022). The Chinese pine genome and methylome unveil key features of conifer evolution. Cell 185, 204-217.e14.

Niu, S.H., Liu, S.W., Ma, J.J., Han, F.X., Li, Y., Li, W. (2019). The transcriptional activity of a temperature-sensitive transcription factor module is associated with pollen shedding time in pine. Tree Physiol. 39, 1173-1186.

Pan, J., Wang, H., Hu, Y., Yu, D. (2018). *Arabidopsis* VQ18 and VQ26 proteins interact with ABI5 transcription factor to negatively modulate ABA response during seed germination. Plant J. 95(3), 529-544.

Pecher, P., Eschen-Lippold, L., Herklotz, S., Kuhle, K., Naumann, K., Bethke, G., et al. (2014). The *Arabidopsis thaliana* mitogen-activated protein kinases MPK3 and MPK6 target a subclass of ‘VQ-motif’-containing proteins to regulate immune responses. New Phytol. 203(2), 592-606.

Pedrosa, A. M., Martins, C.deP., Gonçalves, L. P., Costa, M. G. (2015). Late Embryogenesis Abundant (LEA) Constitutes a Large and Diverse Family of Proteins Involved in Development and Abiotic Stress Responses in Sweet Orange (*Citrus sinensis* L. Osb.). PLoS ONE 10, e0145785.

Peng, X., Xiao, T. A., Meng, J. A., Tao, Z. A., Zhou, D., Tang, X., et al. (2020). Differential expression of rice valine-qlutamine gene family in response to nitric oxide and regulatory circuit of OsVQ7 and OsWRKY24. Rice Science 27(1), 10-20.

Perochon, A., Aldon, D., Galaud, J. P., Ranty, B. (2011). Calmodulin and calmodulin-like proteins in plant calcium signaling. Biochimie 93, 2048-2053.

Perruc, E., Charpenteau, M., Ramirez, B. C., Jauneau, A., Galaud, J. P., Ranjeva, R., et al. (2004). A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in *Arabidopsis thaliana* seedlings. Plant J. 38(3), 410-420.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33, 290-295.

Petersen, K., Qiu, J. L., Lütje, J., Fiil, B. K., Hansen, S., Mundy, J., et al. (2010). *Arabidopsis* MKS1 is involved in basal immunity and requires an intact N-terminal domain for proper function. PLoS One 5(12), e14364.

Pitzschke, A., Schikora, A., Hirt, H. (2009). MAPK cascade signalling networks in plant defence. Curr. Opin. Plant Biol. 12, 421-426.

Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., Finn, R.D. (2018). HMMER web server: 2018 update. Nucleic Acids Res. 46, W200-W204.

Pourcel, C., Salvignol, G., Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology 151(Pt 3), 653-663.

Price, M.N., Dehal, P.S., Arkin, A.P. (2010). FastTree 2—Approximately maximum-likelihood trees for large alignments. PLoS ONE 5, e9490.

Purugganan, M. D., Jackson, S. A. (2021). Advancing crop genomics from lab to field. Nat. Genet. 53(5), 595-601.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152(5), 1173-1183.

Qi, X., Zhang, C., Zhu, J., Liu, C., Huang, C., Li, X., et al. (2020). Genome Editing Enables Next-Generation Hybrid Seed Production Technology. Mol. Plant 13(9), 1262-1269.

Qiu, J. L., Fiil, B. K., Petersen, K., Nielsen, H. B., Botanga, C. J., Thorgrimsen, S., et al. (2008). *Arabidopsis* MAP kinase 4 regulates gene expression through transcription factor release in the nucleus. EMBO J. 27(16), 2214-2221.

Qu, L. J., Zhu, Y. X. (2006). Transcription factor families in *Arabidopsis*: major progress and outstanding issues for future research. Curr. Opin. Plant Biol. 9(5), 544-549.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8(11), 2281-2308.

Rensing, S. A. (2014). Gene duplication as a driver of plant morphogenetic evolution. Curr. Opin. Plant Biol. 17, 43-48.

Roychoudhury, A., Paul, S., Basu, S. (2013). Cross-talk between abscisic acid-dependent and abscisic acid-independent pathways during abiotic stress. Plant Cell Rep. 32(7), 985-1006.

Sadanandom, A., Bailey, M., Ewan, R., Lee, J., and Nelis, S. (2012). The ubiquitin-proteasome system: central modifier of plant signalling. New. Phytol. 196, 13-28.

Senthil-Kumar, M., Mysore, K. S. (2013). Nonhost resistance against bacterial pathogens: retrospectives and prospects. Annu. Rev. Phytopathol. 51, 407-427.

Shan, N., Xiang, Z., Sun, J., Zhu, Q., Xiao, Y., Wang, P., et al. (2021). Genome-wide analysis of valine-glutamine motif-containing proteins related to abiotic stress response in cucumber (*Cucumis sativus* L.). BMC Plant Biol. 21(1), 492.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31(8), 686-688.

Shi, X., Long, Y., He, F., Zhang, C., Wang, R., Zhang, T., Wu, W., Hao, Z., Wang, Y., Wang, G.L., and Ning, Y. (2018). The fungal pathogen *Magnaporthe oryzae* suppresses innate immunity by modulating a host potassium channel. PLoS Pathog. 14, e1006878.

Shimono, M., Sugano, S., Nakayama, A., Jiang, C. J., Ono, K., Toki, S., et al. (2007). Rice WRKY45 plays a crucial role in benzothiadiazole-inducible blast resistance. Plant cell 19(6), 2064-2076.

Shiu, S. H., Shih, M. C., Li, W. H. (2005). Transcription factor families have much higher expansion rates in plants than in animals. Plant Physiol. 139(1), 18-26.

Si, Z., Wang, L., Ji, Z., Qiao, Y., Zhang, K., Han, J. (2023). Genome-wide comparative analysis of the valine glutamine motif containing genes in four *Ipomoea* species. BMC Plant Biol. 23(1), 209.

Song, W., Zhao, H., Zhang, X., Lei, L., Lai, J. (2016). Genome-Wide Identification of VQ Motif-Containing Proteins and their Expression Profiles Under Abiotic Stresses in Maize. Front. Plant Sci. 6, 1177.

Song, X., Meng, X., Guo, H., Cheng, Q., Jing, Y., Chen, M., et al. (2022). Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. Nat. Biotechnol. 40(9), 1403-1411.

Sun, C., Lei, Y., Li, B., Gao, Q., Li, Y., Cao, W., et al. (2023). Precise integration of large DNA sequences in plant genomes using PrimeRoot editors. Nat. Biotechnol. 10.1038/s41587-023-01769-w. Advance online publication.

Sun, Y., Jiao, G., Liu, Z., Zhang, X., Li, J., Guo, X., et al. (2017). Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. Front. Plant Sci. 8, 298.

Sun, Y., Zhang, X., Wu, C., He, Y., Ma, Y., Hou, H., et al. (2016). Engineering Herbicide-Resistant Rice Plants through CRISPR/Cas9-Mediated Homologous Recombination of Acetolactate Synthase. Mol. Plant 9(4), 628-631.

Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E., Mittler, R. (2014). Abiotic and biotic stress combinations. New Phytol. 203, 32-43.

Taj, G., Agarwal, P., Grant, M., Kumar, A. (2010). MAPK machinery in plants: recognition and response to different stresses through multiple signal transduction pathways. Plant Signal. Behav. 5(11), 1370-1378.

Thumuluri, V., Almagro Armenteros, J. J., Johansen, A. R., Nielsen, H., Winther, O. (2022). DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. Nucleic Acids Res. 50, W228-W234.

Tian, J., Zhang, J., Francis, F. Large-Scale Identification and Characterization Analysis of *VQ* Family Genes in Plants, Especially Gymnosperms. Int. J. Mol. Sci. 2023, 24, 14968.

Tian, X., Li, X., Zhou, W., Ren, Y., Wang, Z., Liu, Z., et al. (2017). Transcription Factor OsWRKY53 Positively Regulates Brassinosteroid Signaling and Plant Architecture. Plant Physiol. 175(3), 1337-1349.

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., et al. (2021). Highly accurate protein structure prediction for the human proteome. Nature 596(7873), 590-596.

Tuteja N. (2007). Abscisic Acid and abiotic stress signaling. Plant Signal. Behav. 2(3), 135-138.

Ueno, Y., Yoshida, R., Kishi-Kaboshi, M., Matsushita, A., Jiang, C. J., Goto, S., et al. (2013). MAP kinases phosphorylate rice WRKY45. Plant Signal. Behav. 8(6), e24510.

Uji, Y., Kashihara, K., Kiyama, H., Mochizuki, S., Akimitsu, K., Gomi, K. (2019). Jasmonic Acid-Induced VQ-Motif-Containing Protein OsVQ13 Influences the OsWRKY45 Signaling Pathway and Grain Size by Associating with OsMPK6 in Rice. Int. J. Mol. Sci. 20(12), 2917.

Vierstra, R.D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. Nat. Rev. Mol. Cell. Biol. 10, 385-397.

Voorrips, R. E. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. J. Hered. 93, 77-78.

Walton, R. T., Christie, K. A., Whittaker, M. N., Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. Science 368(6488), 290-296.

Wang, A., Garcia, D., Zhang, H., Feng, K., Chaudhury, A., Berger, F., et al. (2010b). The VQ motif protein IKU1 regulates endosperm growth and seed size in *Arabidopsis*. Plant J. 63(4), 670-679.

Wang, D., Wei, L., Liu, T., Ma, J., Huang, K., Guo, H., et al. (2023a). Suppression of ETI by PTI priming to balance plant growth and defense through an MPK3/MPK6-WRKYs-PP2Cs module. Mol. Plant 16(5), 903-918.

Wang, D., Zhang, Y., Zhang, Z., Zhu, J., Yu, J. (2010a). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteomics Bioinformatics 8(1), 77-80.

Wang, H., Hu, Y., Pan, J., Yu, D. (2015b). *Arabidopsis* VQ motif-containing proteins VQ12 and VQ29 negatively modulate basal defense against *Botrytis cinerea*. Sci. Rep. 5, 14185.

Wang, J., Lisanza, S., Juergens, D., Tischer, D., Watson, J. L., Castro, K. M., et al. (2022b). Scaffolding protein functional sites using deep learning. Science 377(6604), 387-394.

Wang, J., Wang, R., Fang, H., Zhang, C., Zhang, F., Hao, Z., You, X., Shi, X., Park, C.H., Hua, K., et al. (2021). Two VOZ transcription factors link an E3 ligase and an NLR immune receptor to modulate immunity in rice. Mol. Plant 14, 253-266.

Wang, J., Zhou, L., Shi, H., Chern, M., Yu, H., Yi, H., He, M., Yin, J., Zhu, X., Li, Y., et al. (2018). A single transcription factor promotes both yield and immunity in rice. Science 361, 1026-1028.

Wang, M., Vannozzi, A., Wang, G., Zhong, Y., Corso, M., Cavallini, E., et al. (2015a). A comprehensive survey of the grapevine *VQ* gene family and its transcriptional correlation with WRKY proteins. Front. Plant Sci. 6, 417.

Wang, P., Li, J., Zhang, Z., Zhang, Q., Li, X., Xiao, J., et al. (2021). OsVQ1

links rice immunity and flowering via interaction with a mitogen-activated protein kinase OsMPK6. Plant Cell Rep. 40(10), 1989-1999.

Wang, R., Ning, Y., Shi, X., He, F., Zhang, C., Fan, J., Jiang, N., Zhang, Y., Zhang, T., Hu, Y., et al. (2016). Immunity to rice blast disease by suppression of effector-triggered necrosis. Curr. Biol. 26, 2399-2411.

Wang, X., Zhang, H., Sun, G., Jin, Y., Qiu, L. (2014). Identification of active VQ motif-containing genes and the expression patterns under low nitrogen treatment in soybean. Gene 543(2), 237-243.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., et al. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat. Biotechnol. 32(9), 947-951.

Wang, Y., Jiang, Z., Li, Z., Zhao, Y., Tan, W., Liu, Z., et al. (2019). Genome-wide identification and expression analysis of the *VQ* gene family in soybean (*Glycine max*). PeerJ 7, e7509.

Wang, Y., Liu, H., Zhu, D., Gao, Y., Yan, H., Xiang, Y. (2017). Genome-wide analysis of VQ motif-containing proteins in Moso bamboo (*Phyllostachys edulis*). Planta 246(1), 165-181.

Wang, Y., Lu, X., Fu, Y., Wang, H., Yu, C., Chu, J., et al. (2023b). Genome-wide identification and expression analysis of *VQ* gene family under abiotic stress in *Coix lacryma-jobi* L. BMC Plant Biol. 23(1), 327.

Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X., Lee, T. H., Jin, H., Marler, B., Guo, H., et al. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 40, e49.

Wang, Z., Gao, M., Li, Y., Zhang, J., Su, H., Cao, M., et al. (2022a). The transcription factor SlWRKY37 positively regulates jasmonic acid- and dark-induced leaf senescence in tomato. J. Exp. Bot. 73(18), 6207-6225.

Wani, S. H., Anand, S., Singh, B., Bohra, A., Joshi, R. (2021). WRKY transcription factors and plant defense responses: latest discoveries and future prospects. Plant Cell Rep. 40(7), 1071-1085.

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., et al. (2023). De novo design of protein structure and function with RFdiffusion. Nature. Advance online publication.

Weyhe, M., Eschen-Lippold, L., Pecher, P., Scheel, D., Lee, J. (2014). Ménage à trois: the complex relationships between mitogen-activated protein kinases, WRKY transcription factors, and VQ-motif-containing proteins. Plant Signal. Behav. 9(8), e29519.

Wu, M., Liu, H., Han, G., Cai, R., Pan, F., Xiang, Y. (2017). A moso bamboo WRKY gene *PeWRKY83* confers salinity tolerance in transgenic *Arabidopsis* plants. Sci. Rep. 7(1), 11721.

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

Xie, W., Ke, Y., Cao, J., Wang, S., and Yuan, M. (2021). Knock out of transcription factor WRKY53 thickens sclerenchyma cell walls, confers bacterial blight resistance. Plant Physiol. 187, 1746-1761.

Xie, Y. D., Li, W., Guo, D., Dong, J., Zhang, Q., Fu, Y., et al. (2010). The

Arabidopsis gene SIGMA FACTOR-BINDING PROTEIN 1 plays a role in the salicylate- and jasmonate-mediated defence responses. Plant Cell Environ. 33(5), 828-839.

Xing, S., Chen, K., Zhu, H., Zhang, R., Zhang, H., Li, B., et al. (2020). Fine-tuning sugar content in strawberry. Genome Biol. 21(1), 230.

Xiong, X., Gou, J., Liao, Q., Li, Y., Zhou, Q., Bi, G., Li, C., Du, R., Wang, X., Sun, T., et al. (2021). The *Taxus* genome provides insights into paclitaxel biosynthesis. Nat. Plants 7, 1026-1036.

Xu, K., Wang, P. (2022). Genome-wide identification and expression analysis of the *VQ* gene family in *Cucurbita pepo* L. PeerJ 10, e12827.

Xu, R., Duan, P., Yu, H., Zhou, Z., Zhang, B., Wang, R., et al. (2018). Control of Grain Size and Weight by the OsMKKK10-OsMKK4-OsMAPK6 Signaling Pathway in Rice. Mol. Plant 11(6), 860-873.

Xu, Y., Lin, Q., Li, X., Wang, F., Chen, Z., Wang, J., et al. (2021). Fine-tuning the amylose content of rice by precise base editing of the *Wx* gene. Plant Biotechnol. J. 19(1), 11-13.

Xue, P., Zhang, L., Fan, R., Li, Y., Han, X., Qi, T., et al. (2023). HvMPK4 phosphorylates HvWRKY1 to enhance its suppression of barley immunity to powdery mildew fungus. J. Genet. Genomics S1673-8527(23)00115-7.

Yamada, K., Yamaguchi, K., Shirakawa, T., Nakagami, H., Mine, A., Ishikawa, K., et al. (2016). The *Arabidopsis* CERK1-associated kinase PBL27 connects chitin perception to MAPK activation. EMBO J. 35(22), 2468-2483.

Yan, C., Fan, M., Yang, M., Zhao, J., Zhang, W., Su, Y., et al. (2018). Injury Activates Ca^{2+}/Calmodulin-Dependent Phosphorylation of JAV1-JAZ8-WRKY51 Complex for Jasmonate Biosynthesis. Mol. Cell 70(1), 136-149.e7.

Yan, H., Wang, Y., Hu, B., Qiu, Z., Zeng, B., Fan, C. (2019). Genome-Wide Characterization, Evolution, and Expression Profiling of *VQ* Gene Family in Response to Phytohormone Treatments and Abiotic Stress in *Eucalyptus grandis*. Int. J. Mol. Sci. 20(7), 1765.

Yan, X., Luo, R., Liu, X., Hou, Z., Pei, W., Zhu, W., et al. (2023). Characterization and the comprehensive expression analysis of tobacco valine-glutamine genes in response to trichomes development and stress tolerance. Bot. Stud. 64(1), 18.

Yang, M., Liu, Z., Yu, Y., Yang, M., Guo, L., Han, X., et al. (2023). Genome-wide identification of the valine-glutamine motif containing gene family and the role of *VQ25-1* in pollen germination in *Brassica oleracea*. Genes Genomics 45(7), 921-934.

Ye, Y. J., Xiao, Y. Y., Han, Y. C., Shan, W., Fan, Z. Q., Xu, Q. G., et al. (2016). Banana fruit VQ motif-containing protein5 represses cold-responsive transcription factor MaWRKY26 involved in the regulation of JA biosynthetic genes. Sci. Rep. 6, 23632.

Yin, H., Xue, W., Chen, S., Bogorad, R. L., Benedetti, E., Grompe, M., et al. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat. Biotechnol. 32(6), 551-553.

Yoo, S.D., Cho, Y.H., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: a

versatile cell system for transient gene expression analysis. Nat. Protoc. 2, 1565-1572.

You, Q., Zhai, K., Yang, D., Yang, W., Wu, J., Liu, J., Pan, W., Wang, J., Zhu, X., Jian, Y., et al. (2016). An E3 ubiquitin ligase-BAG protein module controls plant innate immunity and broad-spectrum disease resistance. Cell Host Microbe 20, 758-769.

You, X., Zhu, S., Zhang, W., Zhang, J., Wang, C., Jing, R., Chen, W., Wu, H., Cai, Y., Feng, Z., et al. (2019). OsPEX5 regulates rice spikelet development through modulating jasmonic acid biosynthesis. New Phytol. 224, 712-724.

Yu, T., Lu, X., Bai, Y., Mei, X., Guo, Z., Liu, C., et al. (2019). Overexpression of the maize transcription factor *ZmVQ52* accelerates leaf senescence in *Arabidopsis*. PLoS One 14(8), e0221949.

Yuan, G., Qian, Y., Ren, Y., Guan, Y., Wu, X., Ge, C., et al. (2021). The role of plant-specific VQ motif-containing proteins: An ever-thickening plot. Plant Physiol. Biochem. 159, 12-16.

Zeng, D., Zheng, Z., Liu, Y., Liu, T., Li, T., Liu, J., et al. (2022). Exploring C-to-G and A-to-Y Base Editing in Rice by Using New Vector Tools. Int. J. Mol. Sci. 23(14), 7990.

Zeng, L.R., Park, C.H., Venu, R.C., Gough, J., and Wang, G.L. (2008). Classification, expression pattern, and E3 ligase activity assay of rice U-boxcontaining proteins. Mol. Plant 1, 800-815.

Zeng, L.R., Qu, S., Bordeos, A., Yang, C., Baraoidan, M., Yan, H., Xie, Q., Nahm, B.H., Leung, H., and Wang, G.L. (2004). Spotted leaf11, a negative regulator of plant cell death and defense, encodes a U-box/armadillo repeat protein endowed with E3 ubiquitin ligase activity. Plant Cell 16, 2795-2808.

Zeng, L.R., Vega-Sa´ nchez, M.E., Zhu, T., and Wang, G.L. (2006). Ubiquitination-mediated protein degradation and modification: an emerging theme in plant-microbe interactions. Cell Res. 16, 413-426.

Zhang, C., Fang, H., Shi, X., He, F., Wang, R., Fan, J., Bai, P., Wang, J., Park, C.H., Bellizzi, M., et al. (2020). A fungal effector and a rice NLR protein have antagonistic effects on a Bowman-Birk trypsin inhibitor. Plant Biotechnol. J. 18, 2354-2363.

Zhang, G., Wang, F., Li, J., Ding, Q., Zhang, Y., Li, H., et al. (2015). Genome-Wide Identification and Analysis of the VQ Motif-Containing Protein Family in Chinese Cabbage (*Brassica rapa* L. ssp. Pekinensis). Int. J. Mol. Sci. 16(12), 28683-28704.

Zhang, G., Wei, B. (2019). Characterization of VQ motif-containing protein family and their expression patterns under phytohormones and abiotic stresses in melon (*Cucumis melo* L.). Plant Growth Regul. 89, 273-285.

Zhang, H., Si, X., Ji, X., Fan, R., Liu, J., Chen, K., et al. (2018). Genome editing of upstream open reading frames enables translational control in plants. Nat. Biotechnol. 36(9), 894-898.

Zhang, H., Zhang, L., Ji, Y., Jing, Y., Li, L., Chen, Y., et al. (2022b). *Arabidopsis* SIGMA FACTOR BINDING PROTEIN1 (SIB1) and SIB2 inhibit WRKY75 function in abscisic acid-mediated leaf senescence and seed germination. J. Exp.

Bot. 73(1), 182-196.

Zhang, J., Zhang, H., Botella, J. R., Zhu, J. K. (2018). Generation of new glutinous rice by CRISPR/Cas9-targeted mutagenesis of the *Waxy* gene in elite rice varieties. J. Integr. Plant Biol. 60(5), 369-375.

Zhang, L., Wang, K., Han, Y., Yan, L., Zheng, Y., Bi, Z., et al. (2022a). Genome-wide analysis of the VQ motif-containing gene family and expression profiles during phytohormones and abiotic stresses in wheat (*Triticum aestivum* L.). BMC genomics 23(1), 292.

Zhang, L. L., Zheng, Y., Xiong, X. X., Li, H., Zhang, X., Song, Y. L., et al. (2023). The wheat VQ motif-containing protein TaVQ4-D positively regulates drought tolerance in transgenic plants. J. Exp. Bot. erad280.

Zhang, X., Cao, H., Wang, H., Zhang, R., Jia, H., Huang, J., Zhao, J., Yao, J. (2021). Effects of graphene on morphology, microstructure and transcriptomic profiling of Pinus tabuliformis Carr. roots. PLoS ONE 16, e0253812.

Zhang, X. W., Xu, R. R., Liu, Y., You, C. X., An, J. P. (2023b). MdVQ10 promotes wound-triggered leaf senescence in association with MdWRKY75 and undergoes antagonistic modulation of MdCML15 and MdJAZs in apple. Plant J. 115(6), 1599–1618.

Zhang, Y., Bai, Y., Wu, G., Zou, S., Chen, Y., Gao, C., et al. (2017). Simultaneous modification of three homoeologs of TaEDR1 by genome editing enhances powdery mildew resistance in wheat. Plant J. 91(4), 714-724.

Zhang, Z., Mao, Y., Ha, S., Liu, W., Botella, J. R., Zhu, J. K. (2016). A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. Plant Cell Rep. 35(7), 1519-1533.

Zheng, J., Li, H., Guo, Z., Zhuang, X., Huang, W., Mao, C., et al. (2022). Comprehensive Identification and Expression Profiling of the VQ Motif-Containing Gene Family in *Brassica juncea*. Biology 11(12), 1814.

Zheng, Z., Qamar, S. A., Chen, Z., Mengiste, T. (2006). *Arabidopsis* WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J. 48(4), 592-605.

Zhong, Y., Guo, C., Chu, J., Liu, H., Cheng, Z. M. (2018). Microevolution of the *VQ* gene family in six species of *Fragaria*. Genome 61(1), 49-57.

Zhong, Y., Wang, P., Zhang, X., Cheng, Z. M. (2021). Recent Duplications Dominate VQ and WRKY Gene Expansions in Six *Prunus* Species. Int. J. Genomics 2021, 4066394.

Zhou, J., Wang, X., He, Y., Sang, T., Wang, P., Dai, S., et al. (2020). Differential Phosphorylation of the Transcription Factor WRKY33 by the Protein Kinases CPK5/CPK6 and MPK3/MPK6 Cooperatively Regulates Camalexin Biosynthesis in *Arabidopsis*. Plant Cell 32(8), 2621-2638.

Zhou, X., Liao, H., Chern, M., Yin, J., Chen, Y., Wang, J., Zhu, X., Chen, Z., Yuan, C., Zhao, W., et al. (2018). Loss of function of a rice TPR-domain RNA-binding protein confers broad-spectrum disease resistance. Proc. Natl. Acad. Sci. USA 115, 3174-3179.

Zhou, Y., Yang, Y., Zhou, X., Chi, Y., Fan, B., Chen, Z. (2016). Structural and Functional Characterization of the VQ Protein Family and VQ Protein Variants

from Soybean. Sci. Rep. 6, 34663.

Zhu, H., Zhou, Y., Zhai, H., He, S., Zhao, N., Liu, Q. (2020). A Novel Sweetpotato WRKY Transcription Factor, IbWRKY2, Positively Regulates Drought and Salt Tolerance in Transgenic *Arabidopsis*. Biomolecules 10(4), 506.

Zou, Z., Liu, F., Huang, S., Fernando, W. G. D. (2021). Genome-Wide Identification and Analysis of the Valine-Glutamine Motif-Containing Gene Family in *Brassica napus* and Functional Characterization of *BnMKS1* in Response to *Leptosphaeria maculans*. Phytopathology 111(2), 281-292.

Appendix-Publications

1. Hao, Z.[†], **Tian, J**.[†], Fang, H.[†], Fang, L., Xu, X., He, F., et al. (2022). A VQ-motif-containing protein fine-tunes rice immunity and growth by a hierarchical regulatory mechanism. Cell Rep. 40(7), 111235.

2. **Tian, J**., Zhang, J., Francis, F. (2023). Large-scale identification and characterization analysis of *VQ* family genes in plants, especially gymnosperms. Int. J. Mol. Sci. 2023, 24, 14968.

3. **Tian, J**., Zhang, J. Francis, F. (2023). The role and pathway of VQ family in plant growth, immunity, and stress response. Planta 259, 16.

4. Zhang, J., Li, H., Zhong, X., **Tian, J.**, Segers, A., Xia, L., Francis, F. (2023). Silencing an aphid-specific gene *SmDSR33* for aphid control through plant-mediated RNAi in wheat. Front. Plant Sci. 13, 1100394.

5. Zhang, J., Li, H., Zhong, X., **Tian, J.**, Segers, A., Xia, L., Francis, F. (2022). RNA-Interference-Mediated Aphid Control in Crop Plants: A Review. Agriculture 12, 2108.