



Article Genome-Wide and Transcriptome Analysis of Jacalin-Related Lectin Genes in Barley and the Functional Characterization of *HvHorcH* in Low-Nitrogen Tolerance in Arabidopsis

Xiaoyan Quan *, Chen Meng, Chunjuan Xie, Huifang Sun, Boyang Xu, Ramon Santos Bermudez 💿 and Wenxing He *

School of Biological Science and Technology, University of Jinan, Jinan 250022, China * Correspondence: bio_quanxy@ujn.edu.cn (X.Q.); chm_hewx@ujn.edu.cn (W.H.)

Abstract: The jacalin-related lectins (JRLs) are widely distributed in plants and are involved in plant development and multiple stress responses. However, the characteristics of the HvJRL gene family at the genome-wide level and the roles of JRLs in barley's response to low-nitrogen (LN) stress have been rarely reported. In this study, 32 HvJRL genes were identified and unevenly distributed at both ends of the seven chromosomes in barley. HvJRL proteins generally exhibited low sequence similarity but shared conserved jacalin domains by multiple sequence analysis. These proteins were classified into seven subfamilies based on phylogenetic analysis, with a similar gene structure and conserved motifs in the same subfamily. The HvJRL promoters contained a large number of diverse cis-elements associated with hormonal response and stress regulation. Based on the phylogenetic relationships and functionally known JRL homologs, it was predicted that some HvIRLs have the potential to serve functions in multiple stress responses but not nutrition deficiency stress. Subsequently, nine differentially expressed genes (DEGs) encoding eight HvJRL proteins were identified in two barley genotypes with different LN tolerance by transcriptome analysis. Furthermore, 35S:HvHorcH transgenic Arabidopsis seedlings did enhance LN tolerance, which indicated that HvHorcH may be an important regulator of LN stress response (LNSR). The HvJRL DEGs identified herein could provide new candidate genes for LN tolerance studies.

Keywords: genome-wide analysis; JRL; barley; low-nitrogen stress; transcriptome

1. Introduction

Plant lectins are a kind of carbohydrate-binding protein with diverse functional roles in plants [1,2]. Jacalin-related lectins (JRLs) are a new subfamily of plant lectins with at least one jacalin domain [3]. According to the specificity of glycosylation binding, JRLs can be divided into two types, i.e., galactose-specific JRLs (gJRLs) and mannose-specific JRLs (mJRLs) [4,5]. So far, the number of known mJRLs is significantly higher than that of gJRLs. Based on their domain type, JRLs can be grouped into three categories: merojacalins with only one single jacalin domain; holojacalins with multiple jacalin domains; and chimeric jacalins with jacalin domain(s) and other domain(s), such as dirigent, NB-ARC, and Pkinase domains [2,6]. The expansion of the jacalin domain and their fusion with other functional domains in JRLs may help plants to adapt to the highly variable growth environment. The structural composition of JRL proteins varies greatly, but all of them contain a relatively conserved jacalin domain, which is generally composed of about 125 conserved amino acid residues, mainly including non-polar amino acids such as glycine and phenylalanine [7]. This sequence heterogeneity of JRL proteins strongly indicates their divergence in functional roles [6].

Since the first identification in the seeds of jackfruit (*Artocarpus integrifolia*) [8], JRLs have successively been found in more than ten plant species. At present, genome-wide identification of JRL proteins has been carried out in some plant species. There are at least



Citation: Quan, X.; Meng, C.; Xie, C.; Sun, H.; Xu, B.; Santos Bermudez, R.; He, W. Genome-Wide and Transcriptome Analysis of Jacalin-Related Lectin Genes in Barley and the Functional Characterization of *HvHorcH* in Low-Nitrogen Tolerance in Arabidopsis. *Int. J. Mol. Sci.* 2023, 24, 16641. https://doi.org/10.3390/ iims242316641

Academic Editor: Michael Moustakas

Received: 7 October 2023 Revised: 15 November 2023 Accepted: 21 November 2023 Published: 23 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). 48 AtJRLs in Arabidopsis thaliana [9], 30 OsJRLs in rice (Oryza sativa) [2], 19 ZmJRLs in maize (Zea mays) [10], 65 TaJRLs in wheat (Triticum aestivum) [3], and 25 PeJRLs in moso bamboo (Phyllostachys edulis) [10]. However, genome-wide analysis of HvJRLs in barley is scarce, although barley is a major food and feed crop worldwide.

Several plant JRL genes have been functionally characterized in recent years. Some JRLs are associated with plant growth and development. In wheat, TaVER2 regulates flowering and spikelet development, as it is induced during vernalization [11,12]. Overexpression of OsJAC1, a mannose-binding JRL, suppresses the coleoptile and stem elongation in rice [13]. In Arabidopsis, four antagonistic JRLs, namely, JAL31, ATJAL23, PBP1/AtJAL30, and AtJAL22, regulate the size of ER body-type b-glucosidase complexes [9], and At-JAC1 functions in flowering time control [14]. Some JRLs were related to biotic stress responses [15–18]. The expression of OsJRL genes was induced by Magnaporthe grisea infection [19]. Orysata (it was also called OsSalT or OsJRL) is up-regulated under pathogen and insect attacks [20]. The above mentioned OsJAC1 is involved in defense against pathogen attacks [21]. In addition, some JRLs play roles in response to abiotic stresses, such as salt [22,23], drought [24,25], and hormones [12,26,27]. In recent years, many studies have showed that the JRL family genes play important roles in salt stress response, including Orysata and OsJRL40 in rice [28–30], HvHorcH in barley [31], a PeJRL in poplar (Populus *euphratica*) [32], and several *TaJRLs* in wheat [3]. Nitrogen deficiency is a common problem in agricultural production worldwide. However, no evidence of a role of JRLs in LNSR has been reported.

Despite the importance of plant JRLs, current knowledge about molecular characterization of *HvJRLs* in barley is far from sufficient. At present, only four *HvJRL* genes, namely, *HvJPR*, *Horcolin*, *HvLem2*, and *HvHorcH*, have been reported. They are mainly related to beta-glucosidase aggregation, jasmonic acid-inducible, salicylate-inducible, and salt stress responses, respectively [27,31,33–35]. So far, no studies of *HvJRLs* on nutritional stress responses have been reported. In the current study, we conducted a genome-wide identification and analysis of the *HvJRL* gene family in barley. Meanwhile, transcriptomic data, obtained from the two barley genotypes (XZ149, LN-tolerant, and XZ56, LN-sensitive) [36] treated with LN stress, were re-analyzed to identify DEGs encoding HvJRLs [37,38]. Finally, we examined the function of *HvHorcH* in LNSR. The major objectives were to characterize the *HvJRL* gene family, identify the *HvJRL* genes in response to LN stress, and explore the roles of one *HvJRL* gene in LNSR, thus providing valuable information for further research on the functions and regulatory mechanisms of HvJRL proteins, especially those proteins that may be useful for LN tolerance breeding.

2. Results

2.1. Identification of JRL Genes in Barley

A total of 147 HvJRL protein sequences were identified by searching against the barley genome with Hidden Markov Model (HMM) profiles of the jacalin domain (PF01419) using the HMMER3 program [39,40]. Meanwhile, 150 HvJRL protein sequences were obtained using 30 OsJRL protein sequences in rice [2] through the BlastP program [41]. After removing redundant and incomplete sequences, the jacalin domain was confirmed with NCBI CDD [42] and SMART [43]. Finally, 32 *HvJRL* genes were identified in barley (Table 1 and Table S1). These genes included the published barley JRL genes, namely, *HvJPR* (HORVU2Hr1G014020) [35], *HvLem2* (HORVU0Hr1G031210) [27], *Horcolin* (HORVU1Hr1G 000160) [34], and *HvHorcH* (HORVU7Hr1G059330) [31]. The representative protein sequences of the 32 *HvJRLs* (Table S2) were analyzed by EXPASY [44]. HvJRL proteins were composed of 62 (HORVU5Hr1G119370) to 1376 (HORVU5Hr1G063690) amino acids, and their corresponding molecular weight (MW) ranged from 6.65 kDa (HORVU5Hr1G119370) to 153.70 kDa (HORVU5Hr1G063690) (Table 1). The theoretical isoelectric point (pI) was highly variable, ranging from 4.31 (HORVU3Hr1G031550) to 9.64 (HORVU4Hr1G023680) (Table 1). Most HvJRLs were predicted to localize to the cytoplasm (30/32) and nucleus

(28/32), while some were localized to other locations such as chloroplasts, extracell, and mitochondria (Table S1) using WoLF PSORT [45].

Gene ID	Chromosome Location	Amino Acids	Molecular Weight (kDa)	pI
HORVU0Hr1G031210/LEM2	chrUn:186390238-186406458	302	32.10	5.78
HORVU1Hr1G000160/Horcolin	chr1H:475026-476135	146	15.05	7.95
HORVU1Hr1G002450	chr1H:4960525-4962190	314	33.16	6.94
HORVU1Hr1G004660	chr1H:9609641-9610836	142	14.94	4.95
HORVU1Hr1G015190	chr1H:43686713-43688076	301	32.66	5.57
HORVU1Hr1G075520	chr1H:511072663-511076165	586	63.44	6.64
HORVU1Hr1G079360	chr1H:522823378-522824959	208	22.34	9.22
HORVU2Hr1G006090	chr2H:12993110-12995463	319	35.34	6
HORVU2Hr1G006530	chr2H:13676587-13677516	156	16.89	9.34
HORVU2Hr1G012010	chr2H:25563354-25569794	1082	122.10	6.55
HORVU2Hr1G013920	chr2H:30286228-30287832	104	11.10	5.23
HORVU2Hr1G013940	chr2H:30372662-30472925	304	32.69	5.64
HORVU2Hr1G014020/HvJPR	chr2H:30577014-30745532	304	32.66	6.17
HORVU3Hr1G031550	chr3H:157387209-157388376	175	18.77	4.31
HORVU3Hr1G075960	chr3H:567163924-567167067	449	48.41	9.18
HORVU3Hr1G082370	chr3H:598849424-598851100	314	33.72	5.98
HORVU4Hr1G023680	chr4H:131046735-131048211	325	35.11	9.64
HORVU4Hr1G0907204	chr4H:645950575-645970388	385	41.87	8.25
HORVU4Hr1G090780	chr4H:645995092-646003566	374	40.25	9
HORVU5Hr1G009040	chr5H:20488797-20637689	147	15.55	6.42
HORVU5Hr1G009100	chr5H:20927441-21001231	147	15.64	6.21
HORVU5Hr1G012420	chr5H:35699858-35716432	986	109.11	6.15
HORVU5Hr1G062640	chr5H:487199094-487200722	304	33.06	5.24
HORVU5Hr1G063690	chr5H:493882556-493895687	1376	153.70	6.46
HORVU5Hr1G067710	chr5H:514246154-514246870	125	13.21	6.21
HORVU5Hr1G119370	chr5H:655562953-655564277	62	6.65	6.96
HORVU6Hr1G011350	chr6H:20611501-20640921	1254	134.57	6.75
HORVU7Hr1G030210	chr7H:59140290-59141279	160	17.25	5.03
HORVU7Hr1G052850	chr7H:203881126-203882208	155	16.63	6.5
HORVU7Hr1G059330/HvHorcH	chr7H:265851060-265852199	147	15.58	6.02
HORVU7Hr1G109850	chr7H:629167675-629168735	160	17.45	7.72
HORVU7Hr1G121120	chr7H:653151090-653152795	302	32.80	7.18

Table 1. Information on 32 HvJRL genes and their encoded proteins in barley.

2.2. Structural Conservation Analysis of the Jacalin Domains in Barley

To identify the conserved features of the jacalin domains in barley, a multiple sequence alignment was carried out based on their sequences in the HvJRL proteins. A total of 25 highly conserved amino acid residues (more than 75% identity) were identified in the jacalin domains of HvJRLs (Figure 1). The jacalin domain structure was well-established in a mannose-binding JRL protein Heltuba from *Helianthus tuberosus* [46], thus the identified conserved amino acid residues in HvJRLs were analyzed with Heltuba as a reference. There were 12 key residues necessary for the integrity of the β -prism fold in Heltuba, of which 9 are highly conserved in HvJRLs, while the Asn residue on β 8 and Gly and Pro residues on β 9 are less conserved in HvJRLs (Figure 1). In addition, 13 highly conserved residues, which were marked with green stars, were not present in Heltuba (Figure 1). And 18 conserved residues in the HvJRLs, which were marked with black rhombuses, were conserved in PeJRLs (Figure 1). Gly residue in β 1 is invariant in all jacalin domains of HvJRLs, which was also observed in moso bamboo [10], indicating its importance for the biological function of plant JRLs.

H H H н н H н н н H н H н н H H H H H н H н

	β-1	β-2	β-3	β-4	β-5	β-6	
	\rightarrow -		<u> </u>	\longrightarrow			
ORVU2Hr1G013940	LLTKLGPWCGNGGI	DKDILEAVP	RRLESITVSSG	SIVDSIKFSYVDQT	GOKHTAG-PWG-GSGG	NQNTFVLG	70
ORVU2HF1G013920		SDEDILEAVP	RRLESITVSSG		GONNAG-PWG-GSGG	NONTEVIG	70
ORVU7Hr1G121120	LLTKLGPWCGDRGE	AVDILSERP	LKIESITILHE	GIIGSIEYSYIDON	SUNRTAG-PWG-SRNP	NRAEIGLG	70
ORVU2Hr1G006090	AVTK SEAWCSNGNE	VHDIINEKP	LRLRSVTVRTS	DLFDGIRFTYIDOT	OTRTEN-LWGFSODG	SEHTVTLG	71
ORVU5Hr1G062640	LPAKIGPWCGDGGT	AQNIMIKAS	RRLESITISSG	DVVDSIAFSYIDEA	QNCTTD-RLG-GPGG	SSSTIQLA	70
ORVU4Hr1G090780	- PMVLTKYGPWGPENGF	YKDMQNELP	Q <mark>R</mark> LVRVTFFHG	LYIYIMS <mark>F</mark> S <mark>Y</mark> IDHS	VTKSCQLGGRPVSGS	MPKEMNLE	74
ORVU4Hr1G090720	-PMMYRKYGPW <mark>C</mark> SEKGF	YFNMKEATC	DRITKVSVIYT	SFIHTVQFSYIDRT	JINKSAQLGGQALPG S	VVREMSLE	74
ORVU7Hr1G030210	LLTKLGPWGGDDGV	AHDI ITVAP	QRLESITIRWG	KVLDSVAFTYRDKDI	NQLHTAG-PWG-GAGG	-EKEDPDTITLG	73
ORVU3Hr1G082370	LLDKEGPLGGKGGE	EAFDIILGHP	QRLESVTIRSG	DIIDSLAFSYIDQA	SKKQTAG-PWG-GNGG	-LEENQNTITFA	73
ORVUIHF1G015190		FICDIITEAP	QRLESVTIQSG	DAIDSIMFSYTDQAC	SKKQMAG-PWG-GNDG		73
ORV01H11G002450	-MAUVMEVCAWCADDGG	PODITE DP.	NDLEGITIKCG	CUCCIECESEE VUDOO	SOKKIVG-PWG-GNGA		79
ORVU3Hr1G031550	-MADTVKVGPWCALDGO	ARDIVNIG-SRP	ORLETITISWGWDI	FENARIIGISFVYVDOEI	OPVIVG-PWGKIDPO	LSOTIEMT	80
ORVU2Hr1G006530	MSTVKVGMFCGSNGI	VCDITGLSSFPP	SSLRSMEIWSTPG	-HEGIINAIRFTFVD SKI	NNVI PVG-TWGTPLRG	QKSQQTIHLT	81
ORVU2Hr1G012010	-EIILE <mark>KIGPWC</mark> GNIGF	RNRDIIKVAP	QYLRSVKICCG	DVIDAFAFSYLDGN	SNLHQTP-LWG-GTGG	TIHTII <mark>L</mark> E	72
ORVU5Hr1G063690	-EAHLAK I GPW <mark>C</mark> GCTGI	JAKDIIKVAP	CRLSSVIICAG	EVINSIEFSYPDHK	GKMHTFG-PWG-GEGG	-HSGHHHKIM <mark>L</mark> G	75
ORVU1Hr1G000160	-MSKPVKIGPWGGNGGS	SERDIVQPKP	IRMVSMTVSSG	AIVDAIAFTYVGTDI	NVQHSSGIKWGGTGGG	TEDTINLD	74
ORVU1Hr1G004660	-MTKLTKIGTWGGEGGS	SECDIIDTKS	SVMLSVTIRSG	DTINAISFVYLGSD	BRVYSVG-PWGGTGGG	KQVQIFFD	73
ORVU5Hr1G012420	-QDLVTKIGPWCGDSG1	PNDILDVLP	RRLMSVVVHNS	TVINSLTFAYIDWDC	SQQQTAG-PWG-WTPE	PFDGKIDTIILG	76
ORVUEH-1G011350					DONUEG-UNGARSCC		73
ORVU7Hr1G059330	-MSKPVKIGLWCGVGGC	PRDIVRHAP	HRLARVEISGA	DAIHSIKFTYEDHAG	DOHVEG-VWGAPSGG	NKHTLDLE	73
ORVU5Hr1G009040	-MSRPVRIGLWGGVGGC	PRDIVRHAP	HRLARVEISGA	DAIHSIRFTYEDHAG	DQHVEG-AWGAPSGG	NKHTLDLE	73
ORVU5Hr1G067710	-MSKPVKIGLWCGVGGQ	PRDIVRHAP	HRLARVEISGA	DAIHSIKFTYEDHAC	DQHVEG-AWGAPSGG	NKHTLDLE	73
ORVU5Hr1G119370	-MSKPVKIGLWCGVGGQ	PR <mark>DI</mark> VRHAP	H <mark>R</mark> LARVE I SGA	DAIHSIKFTYEDHAC	DQHVEG-AWGAPSGG	nqntfv <mark>l</mark> g	73
ORVU1Hr1G079360	-SKKLMKVGPWGGTGGN	IPWDDGGHTG	I <mark>R</mark> GITLSVYDHRC·	SIVVSIAVE <mark>Y</mark> DRSG	3QLTVPGERH <mark>G</mark> GAG <mark>G</mark> N	httqfklcf	77
ORVU4Hr1G023680	-GRGVIKVGPWCGSGGF	AFYMRRGRG	VSAPRVRSITLQY	SITDAIHSFYQSSD	SDEAVVEQTLDRAEGG	DQDRVRLTPYEHINFA	84
ORVU7Hr1G052850	MQGSVLRMGPCCGGGGG	ICRDMDMRGV	NRVVKLVVRHGDT-	SIVDAISVLYERNG	SQREEWTDLWGGQGGG	TLSEICLR	77
ORV07H11G109850		INROMOVRAV	UWOINT. TVDAVC-		SOVOEOTEDWGGTGGG		76
ORVU3Hr1G075960	-ISTPVRFGPWCGTGGT	TIFDDGIFTG	VROINITVRGLG-	SIISSMEVLYDRNG	JOOAIWGDKRGSSGAA	RPEKIIFDF	76
	* * ** *	*	**	***	**	*	
	0.7	• 6-8	6.0	¢ 10	R-11	0.12	
		11-0					
	p-7	p-0		p-10	>	p-12	
ORVU2Hr1G013940	ASEFVKEVSGTFGIYDK	DLHNIITSLK		EAKG-TPFTIAVC	>	IYLDALGAYVRPL	147
ORVU2Hr1G013940 ORVU2Hr1G013920	ASEFVKEVSGTFGIYDK	DLHNIITSLKI	P-9 FITNVKTYGPFO FITNVKTYGPFO	EAKG-TPFTIAVC	2KN-SSIVGFFARSG-	IYLDALGAYVRPL	147 147
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020	ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK	DLHNIITSLKI DLHNIITSLKI DRHNIITSLKI	P-7 FITNVKTYGPFO FITNVKTYGPFO FITNVKTYGPFO	EAKG-TPFTIAV EAKG-TPFTIAV EAKG-TPFTIAV	QKN-SSIVGFFARSG- QKN-SSIVGFFARSG- QKN-SSIVGFFARSG-	I YLDALGAYVRPL	147 147 147
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU2Hr1G121120	ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK HGEMVTEVSGTFGTHSN	DLHNIITSLKI DLHNIITSLKI DRHNIITSLKI GLHEAISTLKI	FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF IVTNKKTYGPY	EAKG-TPFTIAV EAKG-TPFTIAV EAKG-TPFTIAV RIPGPKPFRAVA	2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- 2HG-KSIVGFFARSG-	J-12 IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGVYRPL IFLNALGVYIAPL	147 147 147 149
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU7Hr1G121120 ORVU7Hr1G006090	ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK HGEMVTEVSGTFGTHSN PSEFLKEVSGVIDINSN	DLHNIITSLKI DLHNIITSLKI DRHNIITSLKI GLHEAISTLKI EPNYFVRSIK	P-9 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF IVTNKKTYGPY LVTNQREFGPF	P-10 EAKG-TPFTIAV EAKG-TPFTIAV RIPGPRPFRAVAI GAQGGTEFKSRVI	2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- 2HG-KSIVGFFGRTDN- 2FD-SMIVAFHGSSGD	D-12 IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGYVRPL IFLNALGVYIAPL NWPGHYRMISIGVYSIPR	147 147 147 149 155
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU7Hr1G121120 ORVU2Hr1G006090 ORVU2Hr1G002640	ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK HOEMVTEVSGTFGTEN PSEFLKEVSGVIDINSN PSEFLTEVSGVIDINSN PSEFLTEVSGVIDINSN	DLHNIITSLKI DLHNIITSLKI DRHNIITSLKI GLHEAISTLK EPNYFVRSIK DLCEVVESIQ	P-7 FITNVKTYGPF FITNVKTYGPF IVTNKKTYGPF UVTNQREFGFF VVTNI-QTYGPF WVTNI-QYGPF	P-10 EAKG-TPFTIAV EAKG-TPFTIAV EAKG-TPFTIAV GAQGGTEFKSRVI GAQGGTEFKSRVI ETQDGTPFTFSVI	2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- PHG-KSIVGFFGRTDN- PHG-KSIVGFFGRTDN- PHG-KSUVGFFGRGG-	D-12 IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGVYRPL IFLMALGVYIAPL NWFG-HYRMHSIGVYSIPR LFLDALGULQPV	147 147 147 149 155 148
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU7Hr1G121120 ORVU2Hr1G006090 ORVU5Hr1G062640 ORVU4Hr1G090720	ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK ASEFVKEVSGTFGIYDK HOEMVTEVSGTFGIYDK PSEFLKEVSGVIDINSN PSEFLKEVSGVIDINSN PSEFLTEVSGTISIDGE DGEFVKSVSGTGLEDG	DLHNIITSLKI DLHNIITSLKI DRHNIITSLKI GLHEAISTLKI GLHEAISTLKI DLCEVVESIQI DAR-HNILTSVSI	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF UVTNQREFGPF FVTNIRSYGPF FVTNLRSYGPF FVTNL-RSYGPF	EAKG-TPFTIAV EAKG-TPFTIAV EAKG-TPFTIAV EAGPFPFRAVAI GAQGGTEFSSVI ETQDGTPFTFSVI REGKHASFSTFEI	2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- 2KN-SSIVGFFARSG- PED-SMIVAFHGSSGD PRY-KKVVGFFGRGG- UNNDSSIVGFFGRHVSI SIVGFFGHHVSI	IYIDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLMALGYYIAPL NWPGHYRMISIGVYSIR EFLDAIGIYLQPY NFAIDSTNYLSSLGYYASP	147 147 147 155 148 162
ORVU2Hr10013940 ORVU2Hr10013920 ORVU2Hr10014020 ORVU7Hr10121120 ORVU7Hr10006090 ORVU5Hr1062540 ORVU4Hr1090780 ORVU4Hr1090720 ORVU4Hr10030210	ASE FYREVSGTFGIVDR ASE FYREVSGTFGIVDR ASE FYREVSGTFGIVDR HGEMVTEVSGTFGIVDR PSEFLREVSGVTFGTHSN PSEFLREVSGVTGLFDG DGEFVRSVSGTTGLFDG DGEFVRSVSGTTGLFDG DGEILRIFISGSTELIDG PSEVITOVDRSV@FFRL	DLHNIITSLKI DL - HNIITSLKI DR - HNIITSLKI GGL - HEAISTLKI EEP - NYFVRSIKI DL - CEVVESIQI DAR - HNIITSVSI HAR -HNIITSLSI KE - LEHCITSLKI	P-7 PITNVKTYGPF PITNVKTYGPF PITNVKTYGPF UTNKKTYGPY UTNCREFGPF PUTNIQTYGPF PGTNLRSYGPF PGTNLRSYGPF	P-10 EAKG-TPFTIAV EAKG-TPFTIAV RIPGPKPFRAVAI GAQGTEFKSRVI ETQGTPFTSVI REGKHASFSTPEI VEGKHASFSTPEI VEGKHASFSTAE	CRN-SSIVOFFARSG- CRN-SSIVOFFARSG- SRN-SSIVOFFARSG- SRG-RSIVOFFARSG- SRG-RSIVOFFARSG- SRUVAFFARSG- LONDSSIVOFFARSG- LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSIVOFFARSSI LONDSSI LONDSSIVOFFARSSI LONDSSI	IP-12 IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGYVRPL IYLDALGYVRPL IYLDALGYVRPL IYLDALGYVRPL IYLDALGYVRPL IYLDALGYVRPL IYLDALGYVYLDP IYLDYNSVGVYLDP IYLAIGYVLPE	147 147 149 155 148 162 162
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU2Hr1G006090 ORVU5Hr1G062640 ORVU4Hr1G090780 ORVU4Hr1G090720 ORVU4Hr1G030210 ORVU4Hr1G082370	D-7 ASEFVKEVSGTFGIVDK ASEFVKEVSGTFGIVDK ASEFVKEVSGTFGIVDK HGEMVTEVSGTFGIVDK PSEFLKEVSGVIDINSN PSEFLKEVSGVIDINSD DGEVKRVSGVTSIDG DGEVLRTFSGSTELIDG DGEVLRTFSGSTELIDG PSEVLTQVDWSVOPFKL FTETLKKVGVVTIG	DL - HNIITSLKI DL - HNIITSLKI GL - HAIISIKI GL - HEAISTIK GL - CEVVESIQI DDR - HNIITSVSI HAR - HNIITSVSI KE - LEHCITSLKI DL - HTVVTSLTI	P-7 PITNVKTYGPP PITNVKTYGPY IVTNKKTYGPY IVTNKKTYGPY IVTNVREFGPT PVTNIRSYGPF PVTNQASYGPP PVTNQASYGPP PVTNQSIYGPP	P-10 2A KG - TP F T I AV EA KG - TP F T I AV RI PG PF F T AV GA	CKN-SSIVGFFARSG- CKN-SSIVGFFARSG- PHG-KSIVGFFARSG- PHG-KSIVGFFGRTDN- PHG-KSIVGFFGRGS- NNIVAFHGSSGDI PRY-KKVVGFFGRGS- NNIDSSIVGFFGRHES: LNN-SSVVGHFGRAG- SLNN-SSVVGFFGRAG-	IP-12 IYLDALGAYVRPL IYLDALGAYVRPL IFLDALGYYRPL IFLDALGYYRPL NWPGHYRMISIGYYSIPR LFLDAIGIYLOPE ITGINEINYVNSVGYLLGP DYLHAIGFYVLPE SYVDAIGYYIANL	147 147 149 155 148 162 162 152
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU7Hr1G12120 ORVU7Hr1G06090 ORVU5Hr1G062640 ORVU4Hr1G090720 ORVU4Hr1G030210 ORVU4Hr1G030210 ORVU3Hr1G032370 ORVU3Hr1G035370	ASEYVKEVSGTFGIVDK ASEYVKEVSGTFGIVDK ASEYVKEVSGTFGIVDK HGEWTEVSGTFGIVDK PSEFLKEVSGTFGIND PSEFLTEVSGTISIDGR DGEVKAVSGTISIDGR DGEVKAVSGTISIDGPSE DGEVLAVSGTIGFFL PTELKKVFGTTGFEGG	DL - HNIITSLKI DL - HNIITSLKI GL - HANITSLKI GL - HEAISTLK: EP - NYFVRSIK DL - CEVVESIQ DAR - HNIITSVSI HAR-HNIITSLSI KE - HEHCITSLKI DL - HTVVTSLTI	P-7 FITNVKTYGPF FITNVKTYGPF IVTNKKTYGPY IVTNKKTYGPY IVTNIQTYGPF FVTNIRSYGPF FGTNLRSYGPF FVTNIRSYGPF FVTNIKIYPFF	BKG-TPFTIAVÇ EAKG-TPFTIAVÇ EAKG-TPFTIAVQ EAKG-TPFTIAVQ FIQDGTPFTFSVI ETQDGTPFTFSVI REQGTFFSVI REQHASPFTFSVI FAQBGTPFTFSVI REQBGTPFTFSVI REQBGTFFSVI REQBGTFSVI XAVDS'THSLEVI XGTGGTFFSIPVI XGTGGTFFSIPVI XATGGTFFSIPVI	PII ZKN - SS IVOFFARSG- ZKN - SS IVOFFARSG- ZKN - SS IVOFFARSG- PHG - RS IVOFFARSG- DHG - SS IVOFFARSG- UNNDSS IVOFFGRGG- UNNDSS IVOFFGRGG- UNN - GSVVGMFGRAG- ZKN - ASVVGFFGRAG- ZKN - ASVVGFFGRAG-	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLMALGYYRPL IFLMAIGIYILGP NFADJSTNYLSSLGYVASF DYLHAIGYYLLGP SYVDAIGYYLPL SYVDAIGYYRPL	147 147 149 155 148 162 162 152 151
ORVU2Hr10013940 ORVU2Hr10013920 ORVU7Hr10114020 ORVU7Hr10121120 ORVU8Hr1006090 ORVU8Hr10090780 ORVU4Hr10090720 ORVU9Hr1003210 ORVU9Hr10082370 ORVU1Hr10015190 ORVU1Hr1002450	ASEYVKEVSGTFGIVDK ASEYVKEVSGTFGIVDK ASEYVKEVSGTFGIVDK HGEMVTEVSGTFGIVDK PSEFLKEVSGTFGINDS PSEFLKEVSGTFGIDFO GEYKAVSGTTGLFDG DGEYLRTFSGSTELDG FSEITGVDBVSGFFL FTETLKKVFGVGTGEFGG PTENVTKVFGTGEFGG	DLHNIITSLKI DLHNIITSLKI DRHNIITSLKI GLHEAISTLKI FDVFVRSIKI DLCEVVESIQI DAR-HNIITSUSI KAR-HNIITSLSI KE-IEHCITSLKI DLHTVVTSLTI DLTVVVTSLTI	P-7 FITNVKTYGPF FITNVKTYGPY ITNVKTYGPY UVTNGKTYGPY UVTNQREFGPT FYTNIQTYGPF FGTNLRSYGPF FGTNLRSYGPF VVTNQ-ASYGPF FVTNVATYGPF	P-10 EAKG-TPFTIAVÇ EAKG-TPFTIAVÇ EAKG-TPFTIAVQ GAQGTPFTSAVAI GAQGTPFTSSVI ETQLGTPFTTSVI REGRHASFSTPEI VEGRHSNFSITAS VAVDSTHYSLPVI RATRGTPFSIPKE RARGPFFSIPKE	PII ZRN - SS IVGFFARSG- ZRN - SS IVGFFARSG- PHG - RS IVGFFGRYDN PED - SMIVAFHGSSGDI PRY - KKVVGFFGRGG- UNNDSS IVGFFGRHVSI SNTNESIAYFGWHFS INN - GSVVGHFGRAG- ZRN - ASVVGFFGRAG- EDG - DNIVGFFGRAG-	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLMALGYVRPL IFLMALGYVRPL NWPGHYRMHSIGVYSP IFGINEINYNSVGVILGP SYVDAIGYVILGP SYVDAIGYVILGP IYLBAIGFYVLPE SYVDAIGYVIANL ESVNALGVVACPI	147 147 149 155 148 162 162 152 151 151
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU7Hr1G121120 ORVU5Hr1G062640 ORVU4Hr1G090780 ORVU4Hr1G090780 ORVU4Hr1G030210 ORVU3Hr1G03210 ORVU3Hr1G032370 ORVU1Hr1G015190 ORVU1Hr1G02450 ORVU1Hr1G02450	ASEFVKEVSGTFGI1DK ASEFVKEVSGTFGI1DK ASEFVKEVSGTFGI1DK BEVTEVSGTFGIDK HGEMVTEVSGTFGIDK BSEFILEVSGVIDINSN PSEFILEVSGVIDINSN DGEILRITSGSTELIDG DGEILRITSGSTELIDG PSEVIGUNSVGPFL PTETLKKVFGVVGTIGG FTENVTKVFGTTGEFGG ESEFIKGVLGTFGTVGE	DLHNIITSLKI DLHNIITSLKI DLHNIITSLKI GLHEAISTLKI JDLCEVVESIQ DDA-HNIITSUSI HAR-HNIITSUSI HAR-HNIITSUSI DLHTVVTSLTI DLDTVVTSLTI DLDTVVTSLTI DLHEGIGSIT	P-7 FITNVKTYGPF FITNVKTYGPF ITNVKTYGPF UVTNGREFGFT VVTNIQTGPF FVTNIQTGPF FVTNIRSYGPF GTNLRSYGPF FVTNISYGPF FVTNISYGPF FVTNISYGPF FVTNISYGPF ETTT-TKHKTY	P-10 EAKG-TPFTIAVC EAKG-TPFTIAVC EAKG-TPFTIAVC GAQGGTEFKSRVI GAQGGTEFKSRVI ETQGPFTTSVF REGRHANFSTTEI VEGRHANFSTTEI VAVSTHSLPVI KATRGTFSIPFI KATGGFTSIPFI YAVGGSFSSPIPI YAVGGSFSSPIPI	ZKN - SSIVGFFARSG- ZKN - SSIVGFFARSG- ZKN - SSIVGFFARSG- HG - KSIVGFFARSG- HG - KSIVGFFARSG- LNNDSSIVGFFGRG- LNNDSSIVGFFGRG- LNN - SSVVGFFGRAG- EXN - ASVVGFFGRAG- EXG - DNVVGFFGRAG- EDG - DNVGFFGRAG- POG-NNIGFYARSG- POG-NNIGFYARSG-	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGYVRPL IFLMALGYYIRPL IYLDAIGYYIRPL IYLDAIGYYIRPL IYLDAIGYYIPL SYVDAIGYYIANL IYLDAIGAYVRPL IYLDALGAYVRPL IYLALGYYVRL	147 147 149 155 148 162 162 152 151 151 151
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU2Hr1G014020 ORVU5Hr1G062640 ORVU5Hr1G090780 ORVU5Hr1G090780 ORVU5Hr1G030210 ORVU7Hr1G03210 ORVU1Hr1G01210 ORVU1Hr1G031210 ORVU3Hr1G031550	ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK BEFIKEVSGTFGIVDK PSEFIKEVSGTISIDGR DGE FYKSVGTISIDGR DGE FYKSVGTISIDGR DGE TYKSVGTTGEFIG PSE TITQUDSVGTFFIDT FTENTKKVFGTTGEFIG PSE IKQVLGTTGTVGG GSE FIKQVLGTTGTVGG	DLHNIITSLR DLHNIITSLR DRHNIITSLR GLHEAISTLR EPNIFVRSIR DDA-HNILTSVS HAR-HNIITSVS HAR-HNIITSLS KE-LEHCITSLR DLHEVVTSLT DLETVVTSLT DLETVVTSLT DL-HAGISSLT DL-HAGISSLT	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FUTNKKTYGPF FUTNLRSYGPF FYTNLRSYGPF FYTNLRSYGPF FYTNQASYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF FYTNVATYGPF	BAKG-TPFTIAVC EAKG-TPFTIAVC EAKG-TPFTIAVC EA	PII ZKN - SSIVGFFARSG- ZKN - SSIVGFFARSG- 2KN - SSIVGFFARSG- PHG - KSIVGFFARSG- DHG - KSIVGFFGRGG- LNNDSSIVGFFGRGG- LNNDSSIVGFFGRAG- SKN - SSVVGHFGRAG- SKN - SSVVGFFGRAG- DG - DNVVGFFGRAG- DG - DNVVGFGFGRAG- DG - DNVVGFGFGRAG- DG - DNVVGFGGRAG- DG - DNVVGFGGRAG- D	IYIDALGAYVRPL IYIDALGAYVRPL IYIDALGAYVRPL IFINALGYYAPL IFIDAIGIYLQYY NFAIDSTNYLSSLGYYAPS NFAIDSTNYLSSLGYYAPS SYUDAIGYYINL SYUDAIGYYAPL SYUNALGYYAPL SYUNALGYYAPL SYUNALGYYAPL SYUNALGYYAPL SYUNALGYYAPL	147 147 149 155 148 162 162 152 151 151 151 158 159
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU7Hr1G12120 ORVU7Hr1G06090 ORVU5Hr1G062640 ORVU4Hr1G090720 ORVU7Hr1G030210 ORVU4Hr1G030210 ORVU3Hr1G082370 ORVU3Hr1G02450 ORVU3Hr1G031550 ORVU3Hr1G031550 ORVU3Hr1G031550	ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGINDK FSE FIKEVSGTFGINDK PSE FIKEVSGTFGIND DGE VIKTSGTFGIND DGE VIKTSGSTELIDG DGE VIKTSGSTELIDG PSE YITQVDWSVGPFKL FTENTKKVFGVVGTIGG PSE FIKQVLGTTGPVGG ESE QLGLVGGTFDVDSK ANVKVIELSGTTDVDK ANVKVIELSGTTDVDK	DL HNIITSLK DL HNIITSLK DL HNIITSLK DL HSISTLK DL CEVVESIQ DAR-HNITSUS KE-IEHCITSLK DL HTVTSLT DL HTVTSLT DL ETVTSLT DL HEGISST DL HEGISST DL HEGISST DL HRITSLQ DL HRITSLQ	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FUTNKKTYGPF FVTNIQTYGPF FVTNIQTYGPF FVTNQASYGPF FVTNQASYGPF FVTNQASYGPF FVTNQASYGPF FVTNVTTGPF LETNTTKHKTY LITNQKQSYAF FRTTDAPKHGPF	P-10 EAKG-TPFTIAVÇ EAKG-TPFTIAVÇ EAKG-TPFTIAVQ EAQGTPFTFVI GAQGTFFSVI REQGTFFSVI REQGTFFSVI REQGTFFSVI REQGTFFSVI REQGTFFSVI REQGTFSVI YAVDSTHYSLPVI KG	PII ZKN - SS IVOF FARSG - ZKN - SS IVOF FARSG - ZKN - SS IVOF FARSG - PHG - RS IVOF FARSG - PHG - RS IVOF FARSG - UNNDSS IVOF FORGG - UNNDSS IVOF FORGA UNN - SSVVGMFGRAG - ENN - SSVVGMFGRAG - ENG - UNVGFFGRSG - PDG - UNIGGFYARG - PDG - UNIGGFYARG - POGCUVFFGRSG - POGCUVFFGRSD - OGCDUVFFGRSD - UHG - SS IVAFCAQAD - HIG - SS IVAFCAQAD -	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLMALGYYNPL IFLMALGYYNPL IFLMALGYYNPL IFLMALGYYNPL IYLDALGAYVYSU IFGINEINYNSUGYYLLGP 	147 147 149 155 148 162 152 151 151 151 158 159 165
ORVU2Hr10013940 ORVU2Hr10013920 ORVU7Hr1012120 ORVU7Hr10121120 ORVU7Hr1002640 ORVU4Hr10090720 ORVU4Hr10090720 ORVU4Hr10030210 ORVU3Hr10082370 ORVU1Hr10015190 ORVU1Hr1001210 ORVU3Hr10031550 ORVU3Hr1001530 ORVU2Hr10012010 ORVU3Hr1003210	ASE FYREVSGTFGIVDK ASE FYREVSGTFGIVDK ASE FYREVSGTFGIVDK HGEMTEVSGTFGINDS PSEFLEVSGTFGINDS DGEYKSYSGTFGIDD GGEYLSYSGTFGLFDG DGEYLSYSGTFGLFDG PSEVIGVDSVGFFL FTETLKKVFGVGTFGEGG ESECLKLVGGTFDNDK QEDQUNCVSGFFDDSDK ANVKVIELSGVTGFDS PSEVIELSGTWFBYGG PSEVIELSGTWFBYGG	DLHNIITSLKI DLHNIITSLKI DRHNIITSLKI GLHEAISTLK: EPNYFVRSIKI DDLCEVVESIQI DDA-HNIITSVSI HAR-HNIITSVSI HAR-HNIITSLSI HOLHTVVTSLTI DDLHTVVTSLTI DDLHEGIGSLTI DDLHEGIGSLTI DDLHEGISSL DDLHRITSLSI FPHSIIKSLMI	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF UTNVKTYGPF UTNCREFGFF FYTNIQTYGPF FGTNLRSYGPF FGTNLRSYGPF VTNV-ASYGPF UTNV-ASYGPF UTNV-ATYGPF LETNTTTYGPF LIINQKQSYAF WTNDKQSYAF WTNDSQSPF	P-10 EA KG - TP F TI AVÇ EA KG - TP F TI AVÇ EA KG - TP F TI AVÇ GA QGTP F TI SVQ GA QGTP F TI SVQ FT QGTP F TT SVI RE GRHSN F SI TA VA VD S'H YSL PVI KA TGGTT F SI PVI KA TGGTT F SI PVI RA TGGTT F SI PVI SA	PII ZRN - SS IVGFFARSG- ZRN - SS IVGFFARSG- ZRN - SS IVGFFARSG- PHG - RS IVGFFGRODN PED - SMIVAFHGSSGDI PRY - KKVVGFFGRGG- UNNDSS IVGFFGRGG- UNN - SSVVGHFGRAG- ZRN - ASVVGHFGRAG- ZRN - ASVVGFFGRGG- POGCANIGFFARG- POGCALVAFFGRSD- POGCALVAFFGRSD- POGCASJVAFFGRSD- CKN - NS IVGFFGCSG- SN - GS IVAFFARG-	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLMALGAYVRPL IFLMALGAYVRPL IFLMALGAYVRPL IFLMALGAYVRPL IFLMALGAYVRPL IFLMALGAYVRPL IFLMALGAYVRPL ITGINEINYNSVGVILGP ITGINEINYNSVGVILGP ITYLALGAYVRPL IYLALGAYVRPL IYLALGAYVRPL UTLKALGYVRPL	147 147 149 155 148 162 152 151 151 151 158 159 165 150
ORVU2Hr10013940 ORVU2Hr10013920 ORVU2Hr10014020 ORVU7Hr10121120 ORVU5Hr10062640 ORVU4Hr10090780 ORVU4Hr10090720 ORVU4Hr10030210 ORVU3Hr1003210 ORVU3Hr1002450 ORVU1Hr1002450 ORVU1Hr1001250 ORVU2Hr100530 ORVU2Hr100530 ORVU2Hr100530 ORVU2Hr100530 ORVU2Hr100530 ORVU2Hr100530 ORVU2Hr100530 ORVU2Hr100530 ORVU2Hr100530	ASE FYKEVSGTFGI 1DK ASE FYKEVSGTFGI 1DK ASE FYKEVSGTFGI 1DK ASE FYKEVSGTFGI 1DK BE FILEVSGTFGI 1DK PSEFLTEVSGTFGI 1DK DGE FYKSVSGTTGLFDG DGE FYKSVSGTTGLFDG DGE FYKSVSGTTGLFDG DGE FYKSVSGTFGLFDG DGE FYKSVSGTFGFG DGE FYKSVSGTFG DGE FYKS	DLHNIITSLR DLHNIITSLR GLHRIITSLR GLHEAISTLK: GLHEAISTLK: GLCEVVESIQ DDA-HNIITSLS KE-IEHCITSLS DLCEVVTSLT DLHEGIGSLT DLHEGIGSLT DLHAGITSLQ DLHRYITSLS DLHRYITSLS MAPG-HYVITSLT	P-7 FITNVKTYGPP FITNVKTYGPP VUTNKKTYGPP VUTNKREGPT VUTNIRESGPT VUTNLRSYGPP FVTNLRSYGPP VUTNYRSYGPP VUTNVATYGPP VUTNVATYGPP VUTNVATYGPP RTTDAPRHHGPT VUTNIHTYGPP RTTDAPRHHGPT VUTNFGYGPT	P-10 EAKG - TPFTIAVÇ EAKG - TPFTIAVÇ EAKG - TPFTIAVÇ GAQGGTE FKSRVI GAQGOTPFTSVI FEQGOTPFTSVI REGRHSNFSTFEI VEQGOTPFTSVI VGGRHSNFSTFEI VEQGOTPTTSVIP KA	2KN - SS IVGFFARSG- 2KN - SS IVGFFARSG- 2KN - SS IVGFFARSG- HG - KS IVGFFARSG- HG - SK IVGFFARSG- HG - SK IVGFFARSG- HG - SS IVGFFARSG- - SNTNESIAYFGHHSS NN - SSVVGFFARG- EGG - DNVVGFFGRSG- - PQGKGELVAFFGRSD- - QGGGUVUFFGRSD- - CKN - NS IVGFFARG- SN - SS IVGFFARG- 2SN - GS IVGFFARG- SN - SS IVGFFARG- - SN - SS IVGFFARG- - - SN - SS IVGFFARG- - - - SN - SS IVGFFARG- - - - - - - - - - - - - -	IP12 IYLDALGAYVRPI IYLDALGAYVRPI IFLKALGVYIAPI IFLKALGVYIAPI IFLKALGVYIAPI IFLKALGVYIAPI IFLDAIGIYLQVIAPI IFGINEINYUKSUGVYLLGP IFGINEINYUKSUGVYLLGP IFGINEINYUKSUGVYLLGP IFGINEINYUKSUGVYLLGP IFLKALGYYVRPI	147 147 147 155 148 162 162 151 151 151 151 155 150 165 150
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU7Hr1G121120 ORVU7Hr1G121120 ORVU5Hr1G062640 ORVU4Hr1G090720 ORVU4Hr1G090720 ORVU4Hr1G030210 ORVU3Hr1G032370 ORVU3Hr1G03250 ORVU3Hr1G031210 ORVU3Hr1G031550 ORVU3Hr1G031550 ORVU3Hr1G03650 ORVU3Hr1G012010 ORVU3Hr1G03600 ORVU3Hr1G063690 ORVU3Hr1G004660	ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK HGEWYTEVSGTFGIVDK PSE FIKEVSGTFGIDD DGE VIKSVGTTGIDG DGE VIKSVGTTGIDG DGE VIKSVGTTGEFG PSE FIK QVLGTTGFF CEG QLKLVGGTFDNDK QEDQUNCVSGTFDDSDK ANVAVIELSGTTMODK PSE SVIEISTFGFUG PSE FILGVSGTMNRYRH ATNVTEISGTVGFFG	DLHNIITSLR DLHNIITSLR DRHNIITSLR GLHEAISTLR DRHNIITSLR DLCEVVESIQ DAR-HNIITSUS HAR-HNIITSUS DLHTVVTSLT DLETVVTSLT DLETVVTSLT DLETVVTSLT DLHAGITSLQ DLHRYITSLS DLHRYITSLS DLHRYITSLS MAPG-HYVTSLT	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FUTNIRSYGPF FUTNIRSYGPF FVTNIRSYGPF FVTNIRSYGPF FVTNIRIGPF FVTNIRIGPF FVTNIKISTY LIINQKQSYAF FRTTDAPRHHGPF FVTNDCSYGPF FVTNDCSYGPF FVTNDCSYGPF	B-10 EAKG-TPFTIAVC EAKG-TPFTIAVC EAKG-TPFTIAVC EAQDGTPFTFNVI GAQGGTFKSRVI ETQDGTPFTFSVI REGKHASFSTPEI VBQBGTFKSVI VAVDSTHYSLPVI KATKGTPFSIPK KATKGTPFSIPK SPAGGFKSLPLI SPAGGFKSLPLI SPAGGFKSVPLI QALGGFFSVPK QALGGFFSVPK QA	PII ZKN - SS IVOFFARSG- ZKN - SS IVOFFARSG- ZKN - SS IVOFFARSG- PHG - RS IVOFFARTDN- PHG - RS IVOFFARTDN- PHG - RS IVOFFGRGG- UNNDSS IVOFFGRGG- UNNDSS IVOFFGRGG- UNNDSS IVOFFGRGG- EKN - ASVVOFFGRGG- DG - DNIVOFFGRGG- PDG - DNIVOFFGRGS- POGKGELVAFFGRSG- POGKGELVAFFGRSG- 28N - GS IVOFFARTG- 28N - GS IVOFFARTG- DG-GK IVOFFGRGG-	IYLDALGAYVRPI IYLDALGAYVRPI IYLDALGAYVRPI IFLMALGYYRPI IFLMALGYYRPI IFLMALGYYRPI IFLMALGYYRPI IFLMALGYYRPI IFLMALGYYRPI IFLMALGYYRPI IFLMALGYYRPI IFLDAIGIYLSSUGYYASP IFGINEINYNSUGYYLLGP IFLDAIGAYURPI IFLAIGAYURPI IFLAIGUNAGAYURPI IFLAIGUNAGAYURPI IFLAIGUNAGAYURPI IFLAIGUNAGAYURPI IFLAIGUNAGAYURPI IFLAIGUNAGAYURPI	147 147 147 155 148 162 162 151 151 151 151 155 150 154 152 152
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU7Hr1G12120 ORVU7Hr1G121120 ORVU7Hr1G06090 ORVU5Hr1G062640 ORVU4Hr1G090720 ORVU7Hr1G030210 ORVU7Hr1G030210 ORVU1Hr1G02450 ORVU1Hr1G031550 ORVU2Hr1G031550 ORVU2Hr1G031550 ORVU2Hr1G03150 ORVU2Hr1G063690 ORVU2Hr1G063690 ORVU5Hr1G06360 ORVU5Hr1G00160 ORVU1Hr1G004660 ORVU1Hr1G00160	ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK HGEMYTEVSGTFGIVDK PSEPLTEVSGTFGIDD DGEVERTSGTDIDG DGEVERTSGTDIDG DGEVERTSGTLIDG DGEVERTSGSTLIDG PSEPITQUDWSVGPFKL FTENTKKVFGVGTGFG PSEPIKQVLGTTGTVGG GEQUKUVGTTGTVGF SESEVIENGTTGFPOG PSEPIGVLGTGTVGFG PSEVIENGTMETMEN ANVRVIELSGTMGTMF PTNVVEISGTMGTMG FTNVVEENGTMGTMG	DL HNIITSLK DL HNIITSLK DL HNIITSLK DL HNIITSLK DL CEVVESIQ DL CEVVESIQ DL CEVVESIQ DL TISLK KE-IEHCITSLK DL HTVVTSLT DL HTVVTSLT DL HEGIGSLT DL HEGIGSLT DL HSIKSLK SPG-HYITSLT TD HDIVTSLK (SGN HDVVFSLK (SS HDVVFSLK (SS HDVVFSLK (SS HDVVFSLK (SS HDVVFSLK (SS HDVVFSLK (SS HDVVFSLK)		P-10 EAKG - TP FTIAVÇ EAKG - TP FTIAVÇ EAKG - TP FTIAVQ EAQGGTE FKSRVI ETQGGTE FKSRVI ETQGGTE FKSRVI ETQGGTE FKSRVI ETQGGTE FKSRVI FEGKHASF STFEI YAVQSTHYSLPVI KGQGGTF SI PVF XA	PII ZKN - SS IVGFFARSG- ZKN - SS IVGFFARSG- ZKN - SS IVGFFARSG- PHG - KS IVGFFGRODN- PHG - KS IVGFFGRODN- PED - SM IVA HGS SGDI ZKN - SS VVGFFGRGG- LNN - GS VVGFFGRGG- LNN - GS VVGFFGRGG- DG - NN VGFFGRGG- PDG - NN IGGFYARGG- POGKELVAFFGRGD- PGGELVAFFGRGD- CGGELVAFFGRGD- LGG - KIAGFFGRGG- LGG - KIAGFFGRGG- LGG - KIAGFFGRGS- LGG - KIVGFFGRGS- LGG - KIVGFFGRGS-	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLMALGYVRPL IFLMALGYVRPL IFLMALGYYRPL IFLMALGYYRPL IGLNEINYUSYGYLLGP 	147 147 149 155 148 162 152 151 151 151 151 155 150 154 152 154
ORVU2Hr10013940 ORVU2Hr10013920 ORVU7Hr1012120 ORVU7Hr10121120 ORVU7Hr1002640 ORVU4Hr10090720 ORVU4Hr10090720 ORVU4Hr10030210 ORVU4Hr10015190 ORVU1Hr10012450 ORVU1Hr10012530 ORVU3Hr1001550 ORVU2Hr10016530 ORVU2Hr1001650 ORVU5Hr1001660 ORVU5Hr100160 ORVU5Hr100160 ORVU1Hr1000460 ORVU1Hr10012420 ORVU5Hr10012420	ASE FYREVSGTFGIVDR ASE FYREVSGTFGIVDR ASE FYREVSGTFGIVDR ASE FYREVSGTFGIVDR BE FILEVSGTFGIVDR DE FILEVSGTFGIDD DE FILEVSGTFGIDFO DE FILEVSGTFGIDFO DE FILEVSGTFGIDFO DE FILEVSGTFGIDFO PTENVTRVFGVEGFE PTENVTRVFGVEGFE DESE FIRGVLGTTGFVG DE DEVCVSGTFDNDR DESE FILGVLGTTGFVG PSE FILGVSGTMINYRR ATNVTVE ISGTMGTYGR PTE VITUE SGTMGTYGR	DL HNIITSLKI DL HNIITSLKI DR HNIITSLKI DR HNIITSLKI DL CEVVESIQI DDA HNIITSLSI HAR-HNIITSLSI HAR-HNIITSLSI HDL DEVVESIQI DDL HTVVTSLTI DL HEGIGSLTI DL HEGIGSLTI DL HEGITSLGI DL HEGITSLGI SIFP HSIIKSLMI TD HDIVTSLKI :GN HDVVTSLKI 'YS HDVVTSLI	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF UTNVKTYGPF UTNCREFGFF FYTNIQTYGPF FFTRSYGPF FYTNIRSYGPF FYTNV-ASYGPF UTNV-ASYGPF UTNV-ASYGPF UTNVKQYAF FYTNO-CSYGPF UTNIGYTKIY UTSKGXFGYF UTNTGXGQPY	P-10 EA KG - TP F TI AVÇ EA KG - TP F TI AVÇ EA KG - TP F TI AVÇ GA QG TP F TI AVÇ GA QG GTP F TR SVI RE GRHAS F ST PEI VE GRHAS F ST PEI VE GG TF F SI PV KA TGGTT F SI PVI KA TGGTT F SI PVI XA VOS ST PSI SI PVI SP QGGE F SS.VPII SP QGGE F SI SVI SP QGGE F SI SVI SP QGGE P SI SVII SP LHGTP F NEWN DA KGGTP F RI PVI IA	PII ZNN - SS IVGFFARSG - ZNN - SS IVGFFARSG - ZNN - SS IVGFFARSG - PHG - KS IVGFFGRODN PED - SMIVAFHGSSGDI PRY - KKVVGFFGROG INN - GSVVGHFGRAG - INN - GSVVGHFGRAG - INN - SSVVGHFGRAG - ZNN - SSVVGHFGRAG - POGCGLVAFFGRAG - POGCGLVAFFGRAG - POGCGLVAFFGRAG - 280 - GSIVAFFARG - LDG - GKIVGFFARG - LDG - GKIVGFFARSS - LDG - GKIVGFFVARA - ZNN - GSIVGFFVARA - ZNN - GSIVGFFVARA - ZNN - GSIVGFFVARA - ZNN - GSIVGFFVARA -	IYLDALGAYVRPI IYLDALGAYVRPI IYLDALGAYVRPI IYLDALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI ITGINEINYNSVGVILGP SYVDAIGYVILGP IYLALGAYVRPI ITGINEINYNSVGVILGP IYLALGAYVRPI IYLALGAYVRPI ITYLALGAYVRPI ITYLALGAYVRPI IYLALGAYVRPI	147 147 149 155 148 162 151 151 151 151 159 165 150 154 152 151
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU2Hr1G014020 ORVU2Hr1G06200 ORVU5Hr1G06200 ORVU5Hr1G090780 ORVU4Hr1G090780 ORVU1Hr1G03210 ORVU1Hr1G015190 ORVU1Hr1G031210 ORVU3Hr1G031550 ORVU2Hr1G031550 ORVU2Hr1G031550 ORVU2Hr1G01350 ORVU2Hr1G00160 ORVU5Hr1G01600 ORVU5Hr1G01420 ORVU5Hr1G01400160 ORVU5Hr1G01420 ORVU5Hr1G01400160 ORVU5Hr1G01350 ORVU5Hr1G01350	ASE FYREVSGTFGIVDE ASE FYREVSGTFGIVDE ASE FYREVSGTFGIVDE ASE FYREVSGTFGIVDE HGEMYTEVSGTFGIVDE PSETLEVSGTISIDGE DGE FYREVSGTISIDGE DGE FYREVSGTISIDGE PSE FILTEVSGTIGE FYETTERVFGTGEFG SE FILTEVSGTFGIVGG QEDQUNCUSGTFDDSDE ANVAVIELSGTFGPYGG PSE SVIELSTFGPYGG PSE SVIELSGTFGPYGG SE FILGVSGTMMEYRE ATNYVELSGTVGKFGK RSEFLIEVSGTLGRNSG DFE FILSTVGSTGRNSG	DL HNIITSLRI DL HNIITSLRI DR HNIITSLRI DR HNIITSLRI EP NIFVRSIRI DL CEVVESIQU DAR -HNILTSUSI MAR-HNIITSUSI DL DTVVTSLT DL DTVVTSLT DL HRYITSLSI DL HRYITSLSI DL HRYITSLSI DL HRYITSLSI DL HRYITSLSI SGN HDVVPSLRI VIS HDVVTSLI VIS HDVVTSLI VIS HDVVTSLI		B-10 EAKG - TP FT IAVÇ EAKG - TP FT IAVÇ EAKG - TP FT IAVÇ EAQDGPP FT FAVAI GAQGGPT FS ST REQDGPP FT SVI PA	PII ZKN - SS IVGFFARSG - ZKN - SS IVGFFARSG - ZKN - SS IVGFFARSG - PHG - KS IVGFFARSG - PHG - KS IVGFFARSG - ZKN - SS IVGFFGRGG - UNNDSS IVGFFGRGG - UNNDSS IVGFFGRGG - EGG - DNVVGFFGRGG - EGG - DNVVGFFGRGG - DG - DNVVGFFGRGG - DG - DNVVGFFGRGG - PGG CELVAFFGRGG - SSN - GS IVGFFARGG - DG - GK IVGFFGRGG - ZKN - GS IVGFFGRGG - ZKN - GS IVGFFGRGG - ZKN - GS IVGFFARGG - ZKN - GS IVGFARGG - ZKN - GS ZKN - ZKN	IYIDALGAYVRPI IYIDALGAYVRPI IYIDALGAYVRPI IYIDALGAYVRPI IFIDAIGUYLQYIAPI LFIDAIGUYLQY NFAIDSTNYLSSLGYVAS SYUDAIGYVLGP SYUDAIGYVIAPI SYUALGAYVRPI DYLSAIGAYVRPI DYLSAIGAYLKI DYLSAIGAYLKI 	147 147 149 155 148 162 152 151 151 151 151 158 159 165 150 154 152 151
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU7Hr1G121120 ORVU7Hr1G02640 ORVU5Hr1G062640 ORVU5Hr1G090720 ORVU4Hr1G090720 ORVU4Hr1G030210 ORVU4Hr1G03210 ORVU1Hr1G015190 ORVU1Hr1G01210 ORVU2Hr1G031210 ORVU2Hr1G031250 ORVU2Hr1G03680 ORVU5Hr1G03680 ORVU5Hr1G0160 ORVU5Hr1G01420 ORVU5Hr1G01450 ORVU5Hr1G01350 ORVU5Hr1G01350 ORVU5Hr1G01350 ORVU5Hr1G0330	ASEYKEVSGTFGIIDE ASEYKEVSGTFGIDE ASEYKEVSGTFGIDE ASEYKEVSGTFGIDE BEFLTEVSGTISIDGR DGEVKAVSGTFGIDE DGEVKAVSGTTGLEDG DGEVLATTGEFG PSEVIGVUSGTFDE SEVIGUUSGTFDNDE GEQLELVGGTFDNDE ANVAVIELSGTTMODE PSEVIGTFDDSE ANVAVIELSGTFDOSE ANVAVIELSGTFORSE ANVAVIELSGTGFGAFG PSEFLIGVSGTMARINE ATNVTEISGTVGKFG RSEFLIEVSGTLGRNSG DFFILMEISGTVGSED DFEVTEISGTVGKGSED	DL HNIITSLR DL HNIITSLR DR HNIITSLR DR HNIITSLR DR HEAISTLR DL CEVVESIQ DAR -HNILTSVS DL CEVVESIQ DDL HTVVTSLT DL ETVVTSLT DL ETVVTSLT DL HEGIGSLT DL HRYITSLS NG HOVVTSLT TD HDIVTSLT TD HDIVTSLT TS HDVVTSLT YS HDVVTSLY VS HSIVTSLR VS HSIVTSLR		B-10 EAKG - TP FT IAVÇ EAKG - TP FT IAVÇ EAKG - TP FT IAVÇ EA	JEN-SSIVGFFARG- ZKN-SSIVGFFARGG- ZKN-SSIVGFFARGG- ZKN-SSIVGFFARGDN- PHG-RSIVGFFGRGG- UNNDSIVGFFGRGG- UNNDSIVGFFGRGG- UNNDSIVGFFGRGG- UNNTSSIAFFGRAG- ZKN-ASVVGFFGRG- PCGKGELVAFFGRG- PCGKGELVAFFGRG- CGGGGUVDFFGRSD- QGDGDVVDFFGRSD- QGDGDVVDFFGRSD- QGDGUVDFFGRSD- QGDGSVVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- QRD-GAIVGFFGRSU- QRD-GAIVGFFGRSU- NG-GQITGFFGRVG- NG-GQITGFFGRVG-	IYIDALGAYVRPI IYIDALGAYVRPI IYIDALGAYVRPI IFINALGYYRPI IFINALGYYRPI IFINALGYYRPI IFINALGYYRPI IFICAL IYIDALGAYVRPI IFICAL IY	147 147 149 155 148 162 152 151 151 151 151 158 159 165 150 154 152 151 154 153 151
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU7Hr1G121120 ORVU7Hr1G121120 ORVU7Hr1G06090 ORVU5Hr1G062640 ORVU4Hr1G090720 ORVU7Hr1G030210 ORVU7Hr1G030210 ORVU1Hr1G015190 ORVU1Hr1G015190 ORVU2Hr1G01550 ORVU2Hr1G013550 ORVU2Hr1G01663690 ORVU5Hr1G01660 ORVU5Hr1G012420 ORVU5Hr1G012420 ORVU5Hr1G01350 ORVU5Hr1G01300 ORVU5Hr1G01300 ORVU5Hr1G05300 ORVU5Hr1G05300 ORVU5Hr1G05300 ORVU5Hr1G05300	ASEYVKEVSGTFGIVDK ASEYVKEVSGTFGIVDK ASEYVKEVSGTFGIVDK ASEYVKEVSGTFGIVDK BETLEVSGTFGIVDK DEEYLATVSGTFGIDD DEEYLATVSGTVGIDG PSEFLTEVSGTVGTGE PSEFLTEVSGTVGTVG PSEFLTEVSGTVGTVG PSEFLTEVSGTVGTVG PSEFLIGVSGTWGYDK ANVRVIELSGTVGVDK PSEFLIGVSGTWGYDK SEFLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDK DFSFTLIGVSGTWGYDC DFSFTLIGVSGTWGYDC DFSFTLIGVGTWGWGS DFSFTUFLIGTYGRWGS DFSFTUFLIGTYGRWGS	DL HNIITSLK DL HNIITSLK DL HNIITSLK DL HNIITSLK DL CEVVESIQ DL CEVVESIQ DL CEVVESIQ DL THITSLS KE - IEHCITSLK DL HTVVTSLT DL HTVVTSLT DL HEGIGSIT DL HEGIGSIT DL HSIKSLM SPG HSIKSLM SGN HDVVTSLT TD HDVVTSLT TS HDVVTSLT VS HSIVTSLK VS HSIVTSLK VS HSIVTSLK VS HSIVTSLK SGN HSIVT		P-10 EAKG-TPFTIAVÇ EAKG-TPFTIAVÇ EAKG-TPFTIAVQ EARG-TPFTIAVQ GA	PII ZNN-SSIVGFFARSG- ZNN-SSIVGFFARSG- ZNN-SSIVGFFARSG- HG-RSIVGFFGRGDN- PHG-RSIVGFFGRGDN- PHG-RSIVGFFGRGG- UNNDSSIVGFFGRGG- UNN-GSVVGMFGRAG- ENN-GSVVGMFGRAG- ENN-GSVVGMFGRAG- DGC-NNIGGFYARAG- PGG-NNIGGFYARAG- PGG-NNIGGFYARAG- CGCDUVFFGRSD- QGCDUVFFFGRD- LGG-SSIVAFCAQAD- KNN-SSIVGFFARGG- LGG-GKIAGFFGRAG- LGG-GKIAGFFGRAG- LGG-GKIVGFFNARG- PHG-GQITGFFGRVG- PHG-GQITGFFGRVG- PHG-GQITGFFGRVG- PHG-GQITGFFGRVG-	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLNALGYVRPL IFLNALGYVRPL IFLNALGYYRPL IFLNALGYYRPL IFLNALGYYRPL IFLNALGYYRPL IFGINEINYNSGGYYLGP IFGINEINYNSGGYYLGP IFGINEINYNSGGYYLGP IFGINEINYNSGGYYLGP IFGINEINYNSGGYYLGP IFGINEINYNSGGYYLGP IFGINEINYNSGGYYLGP IFFLAIGYYRPL IFLALGYYRPL IFLAIGYYRPL IFLAIGYYRPL IFLAIGYYRPL IFLAIGYYFSPE IFLAIGYYFSPE IFLAIGYYRPL IFLAIGYYFPL IFLAIGYYRPL IFLAIGYYFPL IFLAIGYYFPL IFLAIGYYFPL IFLAIGYYPL IFLAIGYYPPL IFLAIGYYPPL IFLAIGYYPL IFLAIGYYPPL IFLAIGYYPPL IFLAIGYYPL	147 147 149 155 148 162 162 152 151 151 151 158 159 1650 154 154 152 151 154
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU2Hr1G014020 ORVU2Hr1G006090 ORVU5Hr1G090780 ORVU4Hr1G090780 ORVU4Hr1G090780 ORVU1Hr1G03210 ORVU1Hr1G01210 ORVU3Hr1G031210 ORVU3Hr1G031210 ORVU3Hr1G03150 ORVU2Hr1G01500 ORVU5Hr1G00160 ORVU5Hr1G01420 ORVU5Hr1G01350 ORVU5Hr1G0141350 ORVU5Hr1G0141350 ORVU5Hr1G0141350 ORVU5Hr1G014140 ORVU5Hr1G014140 ORVU5Hr1G014	ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK ASE FYKEVSGTFGIVDK PSE FIKEVSGTFGIVDG DGE FVKSVSGTFGIVDG DGE FVKSVSGTFGIVDG DGE FVKSVSGTGIDDG PSE FIKEVGVGTGG PTE TVKKVFGTGEFQG PSE FIKQULGTTGTVGG GEGQLKUVGTFDNTKK ANVRVIELSGTVGGTGN SES FIKGVTEISGTVGKGS DTE FVKSGTGAVKGS DTE FVKSISGTYGKVGS DTE VVTEISGTYGKVGS DTE VVTEISGTYGKVGS DTE VVTEISGTYGKVGS	DL HNIITSLR DL HNIITSLR DR HNIITSLR GL HEAISTLR IGL HEAISTLR IGL HEAISTLR IGL CEVVESIQ DDAR-HNILTSVS HAR-HNILTSVS HAR-HNILTSVS HAR-HNITSLS IGL HEGTSL IGL HEGTSL IGL HAGITSLQ IGL HAGITSLQ IGL HAGITSLS IGN HAVITSLR ISS HOVVFSLR ISS HOVVFSLR ISS HSIVTSLR IVS HS		B-10 EA KG - TP F TI AVÇ EA KG - TP F TI AVÇ EA KG - TP F TI AVÇ GA QG FP F TI AVÇ GA QG FP F TI SVQ FT QD GPP F TT SVI RE GRHAS F ST PEI VE GRHAS F ST PEI VE GG TF F SI PVI KA TG GT F F SI PVI KA ANGSP F SS QVI YP VQ GEF F SI PVI SG TK GTP F SI PVI GA ANGSP F SS QVI YP VQ GEF F SI PVI SG KG GP F SV PI QA LH GPP F SR VVI YP CG GG F F NTPMG SG KK GPP F RI PVI TA GG GG F F NTPMG SG	JENN-SSIVGFFARSG- JENN-SSIVGFFARSG- JENN-SSIVGFFARSG- JENSIVGFFARSG- JENSIVGFFARSG- JENSIVGFFARSG- JENSIVGFFGRGG- JENSIVGFFGRGS- JENSIVGFFGRGS- JENSIVAFFGRGS- JEG-DNVVGFFGRSG- JEG-DNVVGFFGRSG- JEG-DNVVGFFGRSG- JEG-DNVVGFFGRSG- JEG-SIVAFFGRSG- JEG-SIVGFFARSG- JEG-SIVGFFARSG- JEG-GRIVGFFGRGS- JEG-GRIVGFFGRGS- JEG-GRIVGFFGRGS- JEG-GRIVGFFGRGS- JEG-GRIVGFFGRGS- JEG-GRIVGFFGRGS- JEG-GRIVGFFGRGS- JEG-GRIVGFFARSG- JEG-GGITGFFGRVG- FHG-GGITGFFGRVG- FHG-GGITGFFGRGS- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG- JEG-GRIVGFFARSG-	IYID ALGAYVRPI 	147 147 149 155 148 162 152 152 151 151 151 158 159 165 150 154 151 151 151 151
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU7Hr1G121120 ORVU7Hr1G121120 ORVU7Hr1G062640 ORVU5Hr1G062640 ORVU4Hr1G090720 ORVU4Hr1G090720 ORVU3Hr1G03210 ORVU3Hr1G03210 ORVU3Hr1G031210 ORVU3Hr1G031210 ORVU3Hr1G031550 ORVU2Hr1G063690 ORVU3Hr1G063690 ORVU3Hr1G063690 ORVU5Hr1G063690 ORVU5Hr1G01420 ORVU5Hr1G01500 ORVU5Hr1G01350 ORVU5Hr1G07300	ASE FYKEVSGTFGIVE ASE FYKEVSGTFGIVE ASE FYKEVSGTFGIVE ASE FYKEVSGTFGIVE BEFIEVSGTISIDGR PSEFIEVSGTISIDGR DGE ILRTFSSTELIG DGE ILRTFSSTELIG DFE ILRVFGTGEFG PSEFIEVGVGTIGG FYENVTKVFGTTGEFG SECILLVGGTFDNTNE QEDQUNCVSGTFDDSDK ANVAVIELSGTTGVGG PSEFIEVGTFDNTNE QEDQUNCVSGTFDDSDK ANVAVIELSGTTGVGFG PSEFIEVSGTLGNDS DFEFIEVSGTLGNDS DFEFIEVSGTLGNDS DFEFIEVSGTLGNDS DFEFIEVSGTLGNDS DFEFIEVSGTLGNDS DTEVTTEISGTYGRWGS DTEVTTEISGTYGRWGS DTEVTVEISGTYGRWGS DTEVTVEISGTYGRWGS	DL HNIITSLR DL HNIITSLR DR HNIITSLR DR HNIITSLR DR HNIITSLR DL CEVVESIQ DAR -HNITSVS HAR -HNITSVS HAR -HNITSVS DL DTVVTSLT DL DTVVTSLT DL ETVVTSLT DL HAGITSLQ DL HAGITSLQ DL HRYITSLS MACHIVITSLT TD HDIVTSLR SGN HDVVPSLR VS HSIVTSLR VS HSIVTSLR VS HSIVTSLR VS HSIVTSLR VS HSIVTSLR VS HSIVTSLR DL HNITSLR	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FYTNIRSYGPF FYTNIRSYGPF FYTNIRSYGPF FYTNIRSYGPF FYTNIRSYGPF FYTNIKIYFT LIINQKQSYAF FYTNOSYGPF FYTNOSYGPF FYTNOSYGPF FYTNAGYIEC FYTNA-GQTIEC FYTNNGQTIEC FYTNNGQTIEC FYTNNGQTIEC FYTNNGQTIEC FYTNNGQTIEC FYTNNGQTIEC FYTNNGQTIEC FYTNNGQTIEC FYTNNSTYGPF	B-10 EA	JENN-SSIVGFFARSG- JENN-SSIVGFFARSG- JENN-SSIVGFFARSG- DENSIVGFFARSG- DENSIVGFFARSG- DENSIVGFFGRGG- LNNDSSIVGFFGRGG- LNNDSSIVGFFGRGG- LNNDSSIVGFFGRGG- DEG-DNVVGFFGRGG- DEG-DNVVGFFGRGG- DEG-DNVVGFFGRSG- DEG-DNVGFFGRSG- DEG-DNVGFFGRSG- DEG-DNVGFFGRSG- DEG-DNIGFFGRGAG- DEG-DNIGFFGRGAG- DEG-DNIGFFGRGG- LGG-GGIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFFARGG- DG-GRIVGFCARGG-	IYIDALGAYVRPI IYIDALGAYVRPI IYIDALGAYVRPI IFINALGYYRPI IYDALGAYYRPI IYDALGAYYAPI IYDALGAYYAPI	147 1477 1479 1555 1488 1662 152 151 151 151 1551 150 154 153 151 151 151 151 151 151 151 151
ORVU2Hr10013940 ORVU2Hr10013920 ORVU7Hr10121120 ORVU7Hr10121120 ORVU7Hr1006090 ORVU5Hr1006090 ORVU5Hr10090720 ORVU3Hr10090720 ORVU3Hr10090720 ORVU3Hr10015190 ORVU3Hr10015190 ORVU3Hr1001520 ORVU2Hr1001550 ORVU2Hr1001550 ORVU2Hr1001600 ORVU3Hr1001600 ORVU3Hr1001600 ORVU5Hr100160 ORVU5Hr1001530 ORVU5Hr1005330 ORVU5Hr1005710 ORVU5Hr1005710 ORVU5Hr10079360 ORVU5Hr10079360	ASEYKEVSGTFGIVDE ASEYKEVSGTFGIVDE ASEYKEVSGTFGIVDE ASEYKEVSGTFGIVDE BEFLTEVSGTFGIDDE DGEVERTSGTDINS PSEFLEVSGTDISIDGE DGEVERTSGTELDG DGEVERTSGTELDG DGEVERTSGTELDG PSEVIQUDWSVGPFL FTENTKVFGVVGTIGG PSEVIQUGTGTVGG SEGULKLVGGTFDNDE QEDQVNCVSGTFDDSE ANV&IELSGTTGVGG PSEVIELSGTFGVGG DFEVTELSGTVGKFG FNVVELSGTVGKFGS DTEVTELSGTVGKVGS DTEVTELSGTVGKVGS DTEVTELSGTVGKVGS SEFVEVSGTFGIVDE	DL HNIITSLK DL HNIITSLK DR HNIITSLK DR HNIITSLK DL CEVVESIQ DAR -HNITSLS DL CEVVESIQ DAR -HNITSLS KE-IEHCITSLS NE-IEHCITSLS DL HTVVTSLT DL ETVVTSLT DL HEGIGSIT DL HAGITSLQ DL HRIITSLS NG HDVVTSLT SG HDVVTSLK VS HSIVTSLK VS HSIVTSLK VS HSIVTSLK US HSIVTSLK SG SPVIRSLA		P-10 EAKG - TP FT IAV; EAKG - TP FT IAV; EAKG - TP FT IAV; EA	JEN-SSIVGFFARSG- ZKN-SSIVGFFARSG- ZKN-SSIVGFFARSG- ZKN-SSIVGFFARSG- ZKN-SSIVGFFARSG- DEG-SKIVGFFGRGG- LNNDSSIVGFFGRGG- LNNDSSIVGFFGRGG- LNN-GSVVGMFGRAG- ENN-SSVVGFFGRGG- DG-DNVUGFFGRGG- DG-DNVUGFFGRGG- DG-DNVUGFFGRSG- ZGKGELVAFFGRSG- ZGKGELVAFFGRSG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- DG-GKIVGFFARGG- ZKN-SSIVGFFARSG- ZKG-GVIVGFFARSG- ZKG-GVIVGFFARSG- ZKG-GVIVGFFARSG- ZKG-GVIVGFCGRSG- XKN-CCIVGFWGRSG-	IYLDALGAYVRPL IYLDALGAYVRYL IYLDALGAYVRPL	147 147 147 149 155 148 162 162 152 151 151 151 151 154 152 151 151 151 151 151 151 151 151
ORVU2Hr10013940 ORVU2Hr10013920 ORVU7Hr1012120 ORVU7Hr1021210 ORVU7Hr1002640 ORVU3Hr10002640 ORVU3Hr1000210 ORVU3Hr1000210 ORVU3Hr1003210 ORVU1Hr1002450 ORVU1Hr10015190 ORVU3Hr10031550 ORVU2Hr10013550 ORVU2Hr100160 ORVU3Hr10031550 ORVU3Hr1001320 ORVU3Hr1001320 ORVU3Hr1001320 ORVU3Hr1001320 ORVU3Hr1001320 ORVU3Hr1001320 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr1003930 ORVU3Hr10032850 ORVU3Hr10032850	ASEYVKEVSGTFGIVK ASEYVKEVSGTFGIVK ASEYVKEVSGTFGIVK HGEMYTEVSGTFGIVK HGEMYTEVSGTFGINS PSEPLEVSGTISIDG DGEVKTSGSTLIDG DGEVKTSGSTLIDG DGEVKTSGSTLIDG DGEVKTFGSTLIDG DGEVKTFGSTLIDG PSEPLKQVLGTGTVGG PSEPLKQVLGTGTVGG SEGELKLVGGTFDNDK ANVRVIELSGTMGVK ANVRVIELSGTMGVK ANVRVIELSGTMGVKFG PSEVLIGVSGTMRVKH ATNYVEISGTVGKGSED DTEVTEISGTVGKWGS	DL - HNIITSLK DL - HNIITSLK DL - HNIITSLK DL - HNIITSLK DL - CEVVESIQ DDA - HNITSLK DL - CEVVESIQ DDA - HNITSVS KE - IEHCITSLK ME - IEHCITSLK ME - IEHCITSLK DL - HTVVTSLT DL - HTVVTSLT DL - HEGIGSLT DL - HEGIGSLT TD - HGINTSLS SUB - HSITSLS SUB - HSITSLS VS - HSIVTSLK VS - HSIVTSLK SG - SPVIRSLA CG - SPVIRSLA CG - SPVIRSLA	P-7 FITNVKTYGPF FITNVKTYGPF FITNVKTYGPF FUTNKKTYGPF FVTNIQTYGPF FVTNIQTYGPF FVTNISYGPF FVTNQASYGPF FVTNQASYGPF FVTNQASYGPF FVTNQSYGPF FVTNVGYTAF FRTTDAPRHHGPF JVSSITTYGPF FRTTDAPRHHGPF IVTSKGVTKIY CVPKNGQTIEC FVTNN	P-10 EA	PII ZNN-SSIVGFFARSG- ZNN-SSIVGFFARSG- ZNN-SSIVGFFARSG- PHG-RSIVGFFGRGDN- PED-SMIVAFHGSSGDI PSY-KKVVGFFGRGG- UNNDSSIVGFFGRGG- UNN-GSVVGMFGRAG- ENN-GSVVGMFGRAG- ENN-GSVVGMFGRAG- DGO-NNIGGFYARAG- PGG-NNIGGFYARAG- PGG-NNIGGFYARAG- PGG-NNIGGFYARAG- CGGUSVFFGRSD- QGDGDVVFFGRSD- QGDGDVVFFGRSD- DG-GKIVGFFARGG- LDG-GKIVGFFARGG- LDG-GKIVGFFARGG- LDG-GKIVGFFARGG- RHG-GGITGFFGRVG- PHG-GGITGFFGRVG- PHG-GGITGFFGRVG- HGG-GGITGFFARSG- HGG-GGITGFFAR	IYLDALGAYVRPI IYLDALGAYVRPI IYLDALGAYVRPI IYLDALGAYVRPI IYLDALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IFLMALGAYVRPI IYLDALGAYVRPI	147 147 149 155 148 162 152 151 151 151 151 154 159 165 153 151 151 151 151 151 151 151 151 15
ORVU2Hr1G013940 ORVU2Hr1G013920 ORVU2Hr1G014020 ORVU2Hr1G014020 ORVU2Hr1G0160609 ORVU5Hr1G062640 ORVU4Hr1G090780 ORVU4Hr1G090780 ORVU4Hr1G030210 ORVU3Hr1G03210 ORVU3Hr1G031210 ORVU3Hr1G031210 ORVU3Hr1G031550 ORVU2Hr1G013550 ORVU2Hr1G01350 ORVU5Hr1G0160 ORVU5Hr1G01420 ORVU5Hr1G01460 ORVU5Hr1G01350 ORVU5Hr1G01350 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0310 ORVU5Hr1G0330 ORVU5Hr1G0330 ORVU5Hr1G032850 ORVU5Hr1G032850	ASE FYREVSGTFGIVDK ASE FYREVSGTFGIVDK ASE FYREVSGTFGIVDK ASE FYREVSGTFGIVDK FSEFIREVSGTIGIDK FSEFIREVSGTIGIDG DGE FVRSVSGTIGIDG FSEFIREVSGTIGG FSEFIREVGVGTIGG FSEFIREVGVGTIGG FSEFIREVGGTFDNTNK QEDQVNCVSGTFDDSDK ANVAVIELSGTFGVGG PSESVIEJSGTFGVGG PSESVIEJSGTFGANGG FNVVEEISGTVGKFGA FTNVVEEISGTVGKFGA FTNVVEISGTVGKFGA SEFILIEVSGTLGANGG DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVTEISGTVGRWGS DTEVTEISGTVGRWGS DTEVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DTEVVTEISGTVGRWGS DELTAVSGRVGHGIISF	DL - HNIITSLR DL - HNIITSLR DR - HNIITSLR DR - HNIITSLR DR - HRISTR EF - NYFYRSIR DL - CEVVESIQ DAR - HNIITSLS KE - LEHCITSLR DL - HVVTSLT DL - HEVTSLT DL - HEVTSLT DL - HRITSLS NE - LEHCITSLR SGN - HDVVTSLT TD - HAGITSLQ SGN - HDVVTSLT VS - HSVTSLR VS - HSVTSLR VS - HSIVTSLR VS - HSIVTSLR US - HNITSLR MG - SVUTSLR MG - SVUTSLR MF - VTCLKSLT MF - VTCLKSLT		B-10 EA	PII ZKN - SS I VOF FARSG - ZKN - SS I VOF FARSG - ZKN - SS I VOF FARSG - PHG - RS I VOF FARSG - PHG - RS I VOF FGRTDN. PED - SM I VH HG SSGO XKN - SS VVGF FGRGG - JKN - SS VVGF FGRGG - JKN - SS VVGF FGRGG - DG - NN I GF YARAG - PCG C SI VAF FGRSG - PCG C SI VAF FGRSG - ZKN - SS I VAF FGRGG - JKN - SG I VOF FGRGG - JKD - GR I VOF FGRGG - ZKN - SS I VOF FGRSG - ZKN - SS I VOF FGRSG - ZKN - SS I VOF FGRSG - ZKN - SS I VOF FFGRSG - ZKN - SS I VOF FFRSG - ZKN - SG I VOF FGRSG - XKG - GV I VOF CGRSG - XKN - GC I VOF CGRSG - XKN - GC I VOF CGRSG - XKN - GC I VOF FGRSG - ZKN - SG I VOF FGRSG - ZKN -	IYIDALGAYYRPI SYUDALGAYYRPI IYIDALGAYYRPI IYIDALGAYYRPI IYIDALGAYYRPI IYIDALGAYYRPI IYIDALGAYYRPI VIDALGAYYRPI IYIDALGAYYRPI IYIDALGAYYRPI <	147 147 147 147 155 162 152 151 151 151 155 150 154 151 151 151 151 151 151 151 151 151 151 151 151 151 151 151 151 151 151 152 151 151 151 152 151 152 151 151 151 152 151 152 151 151 151 152 151 151 151 151 152 151 151 151 151 152 151 151 151 151 151 151 151 151 152 151 151 151 151 151 152 151 1551 1551 1551 1551 1551 1551 1551 1551 1551 1551 1552 1554 1552 1554 1552 1554 1555 1555 1554 1555 1555 1554 15555 1555 15555 15555555555
ORVU2Hr10013940 ORVU2Hr10013920 ORVU7Hr10112120 ORVU7Hr1012120 ORVU5Hr1006090 ORVU5Hr10062640 ORVU4Hr10090720 ORVU4Hr10090720 ORVU4Hr10030210 ORVU4Hr1003210 ORVU1Hr1001210 ORVU1Hr1001210 ORVU5Hr1001210 ORVU5Hr1001200 ORVU5Hr100160 ORVU5Hr100160 ORVU5Hr1001200 ORVU5Hr1001200 ORVU5Hr1001200 ORVU5Hr1001200 ORVU5Hr100700 ORVU5Hr100700 ORVU5Hr1007300 ORVU5Hr1007360 ORVU5Hr1007360 ORVU5Hr1007520	ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK PSEFLEVSGTISIDGR DGEVKAVSGTFGIVDF DGEVKAVSGTTGIFD PSEFLEVSGTISIDGR DGEVLAVTSGTTGEFG PSEFLEVSGTGTGG FTENTKNFGTTGEFG PSEFLKQVGTTGEFG PSEFLKQVGTFDDSDK ANVAVIELSGTTGVGG PSEFLIGVSGTMNRYRH ATNVTEISGTGKFGA FTNVTEISGTGKFGA SEFLIEVSGTLGNSG DFFLTMISGTVGKFGA DFFLTMISGTVGKFGA DFFLTMISGTVGKFGA DFFVTEISGTYGRWGS DTEVTTEISGTYGRWGS DTEVTEISGTYGRWGS DTEVTEISGTYGRWGS DTEVTEISGTYGRWGS DTEVTEISGTYGRWGS DTEVTEISGTYGRWGS DTEVTEISGTYGRWGS SAEFVKEVSGTFGIVDA PDEHLAVSGRYGAVBFI DELTAVGKHGIISF PDEVLTMISGRYGAVBFI	DL - HNIITSLR DL - HNIITSLR DL - HNIITSLR DR - HNIITSLR DR - HNIITSLR DL - CEVVESIQ DAR - HNIITSLS DL - CEVVESIQ DDAR - HNIITSLS RE - LEHCITSLR DL - HTVVTSLT DL - ETVVTSLT DL - ETVVTSLT DL - HRYITSLS DL - HRYITSLS NG - HDVVTSLT VS - HSIVTSLR VS - HSIVTSLR NG - SFAVRSLT MG - FTAVRSLT		B-10 EA	JENN - SSIVGFFARSG- ZKN - SSIVGFFARSG- ZKN - SSIVGFFARSG- ZKN - SSIVGFFARSG- ZKN - SSIVGFFARSG- ZKN - SSIVGFFGRGG- UNNDSIVGFFGRGG- UNNDSIVGFFGRGG- UNNDSIVGFFGRGG- ZKN - SSVVGFFGRGG- ZKN - SSVVGFFGRGS- ZKN - SSIVGFFARG- ZKN - SSIVGFARG- ZKN -	IYIDALGAYVRPI IYIDALGAYVRPI IYIDALGAYVRPI IFINALGYYRPI IFINALGYYRPI IFINALGYYRPI IFINALGYYRPI IFIDALGYYRPI IFIDALGYYRPI IFIDALGYYRPI IFIDALGYYRPI IFIDALGYYRPI IFIDALGYYRPI IFIGINEINYNSUGYILGP IFIDALGAYURPI IFIDALGYINPI IFIDALGYINPI IFIDALGYINPI IFIDALGYINPI IFIDALGYINPI IYIDALGAYVRPI	147 147 147 147 145 152 152 151 151 151 151 155 150 154 151 151 151 154 151 151 151 151 155 151 151 151 151 151 151 151 152 151 151 151 152 151 151 151 152 151 151 152 151 151 151 152 151 151 151 152 151 151 151 152 151 151 151 152 151 151 152 151 151 151 152 151 151 151 151 152 151 151 151 151 151 151 151 152 151 151 151 151 152 151 151 151 151 151 152 151 152 154 152 154 152 154 152 154 152 154 152 154 152 154 152 154 154 152 154
ORVU2Hr10013940 ORVU2Hr10013920 ORVU2Hr10014020 ORVU7Hr10121120 ORVU7Hr1005090 ORVU5Hr10062640 ORVU4Hr10090720 ORVU7Hr10090720 ORVU7Hr10090720 ORVU7Hr10082370 ORVU1Hr10082370 ORVU1Hr1008250 ORVU1Hr1001510 ORVU2Hr1001550 ORVU2Hr10012010 ORVU2Hr100160390 ORVU1Hr100160390 ORVU1Hr1001420 ORVU5Hr1001420 ORVU5Hr1001350 ORVU5Hr1009300 ORVU5Hr1009300 ORVU5Hr1007530 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012360 ORVU5Hr1012520	ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK ASEYKEVSGTFGIVDK HGEMMTEVSGTFGINDK PSEFLÆVSGTFGIND DGEVKAVSGTFGIDDE DGEVKAVSGTGIDE PSEFLÆVSGTISIDGR PSEFLÆVSGTFGIDE PSEFLÆVSGTFGIDE PSEFLÆVSGTFGIDE SEFLÆVSGTFGIDE SEFLÆVSGTFGIDE SEFLEVSGTGE PSEFLIGVSGTMERKE ANVENIELSGTTGGVGS DFETMILISGTYGRWGS DTEVTELISGTYGRWGS DTEVTELISGTYGRWGS ASEFVÆVSGTFGIDE SESLIGTYGRWGS DTEVTELISGTYGRWGS ASEFVÆVSGTFGIDE SESLIGTYGRWGS DEFVÆVSGTFGIDE SESLIGTYGRWGS DEFVÆVSGTFGIDE SESLIGTYGRWGS DEFVÆVSGTFGIDE SESLIGTYGRWGS DEFVÆVSGTFGIDE SESLIGTYGRWGS DEFVÆVSGTFGIDE SESLIGTYGRWGS DEFVÆVSGTFGIDE SESLIGTYGRWGS DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG DESLIGTYG	DL HNIITSLK DL HNIITSLK DR HNIITSLK DR HNIITSLK DL CEVVESIQ DAR-HNITSLS DL CEVVESIQ DAR-HNITSLS KE-IEHCITSLK IDL HTVVTSLT DL ETVVTSLT DL HEGIGSLT DL HUTTSLK GG SPURSLK MG FTAVRSLT MG FTAVRSLT		B-10 EAKG - TP FT IAV; EAKG - TP FT IAV; EAKG - TP FT IAV; EA	JEN-SSIVGFFARSG- ZKN-SSIVGFFARSG- ZKN-SSIVGFFARSG- ZKN-SSIVGFFARSG- HG-KSIVGFFGRGDN- PHG-KSIVGFFGRGG- UNNDSSIVGFFGRGG- UNNDSSIVGFFGRGG- UNN-SSVVGMFGRAG- EXN-SSVVGMFGRAG- DG-UNIGFFGRG- DG-UNIGFFGRG- DG-UNIGFFARG- DG-UNIGFFARG- DG-SIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- DG-GKIVGFFARG- UGFGQITGFFGRVG- PHG-GQITGFFGRVG- PHG-SSIVGFFARG- ZKN-SSIVGFFARG- ZKN-SSIVGFFARG- ZKN-SSIVGFFARG- ZKN-SSIVGFFARG- ZKN-SSIVGFFARG- ZKN-SSIVGFFARG- ZKN-SSIVGFFARG- ZKN-SGIVGFACG- ZKN-SGIVGFACG	IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IYLDALGAYVRPL IFLNALGYYRPL IFLNALGYYRPL IFLNALGYYRPL IFLNALGYYRPL IYLDALGAYV	147 147 147 149 155 148 162 152 151 151 151 151 154 153 155 155 155 155 155 155 154 154

Figure 1. Multiple sequence alignment of the jacalin domains of HvJRLs. Amino acids with 100% identity are shown in black, and those with greater than 75% identity are shown in light blue. Dotted lines indicate gaps. Regions corresponding to β -strands are labeled by horizontal arrows above the sequences, and red stars indicate key residues in Heltuba used to package the three Greek key motifs. Green stars indicate residues that are highly conserved in HvJRLs but not present in Heltuba. Black rhombuses indicate highly conserved residues in both HvJRLs and PeJRLs.

2.3. Gene Structure, Conserved Motif, and Domain Analysis of HvJRLs

A phylogenetic tree was constructed on the 32 HvJRL proteins using the maximum likelihood (ML) method with 1000 bootstrap replicates. These proteins were classified into seven subfamilies (A–G) based on the clades with over 50% bootstrap support (Figure S1). To characterize the features of HvJRL proteins, the domain compositions based on the 32 HvJRL protein sequences were analyzed. The structural domains of HvJRL proteins are various, with different numbers of jacalin domains and different types of other domains (Figure S1). Of the 32 HvJRL proteins, 14 (43.75%) were merojacalins with only one jacalin domain, only 3 (9.38%) were holojacalins with multiple jacalin domains, and

15 (46.88%) were chimeric jacalins with jacalin domain(s) plus other domain(s) (Figure S1). For chimeric jacalins, 11 (73.33%) had an N-terminal dirigent domain plus one jacalin domain (Figure S1), and the remaining 4 had jacalin domain(s) plus other domain(s), including the PKc_like superfamily, B3, Motile_Sperm superfamily, NB-ARC superfamily, Rx_CC_like, LRR superfamily, and Rx_N (Figure S1). The HvJRL proteins in the same subfamily generally exhibited a similar domain composition. For instance, all the members in subfamily A, except HORVU7Hr1G030210, were composed of one dirigent domain and one jacalin domain, and all of those in subfamily F contained only one jacalin domain (Figure S1).

In order to support the phylogenetic relationship, the gene structure analysis of *HvJRLs* was displayed. The result showed that most *HvJRLs* have 2–5 exons, with the exceptions of three chimeric lectins, including HORVU5Hr1G063690 (9), HORVU5Hr1G012420 (12), and HORVU6Hr1G011350 (21) (Figure 2B). In general, most members in the same subfamily exhibited a similar gene structure (Figure 2B). For instance, 11 of 12 genes in subfamily A had four exons.



Figure 2. Phylogenetic tree, conserved motifs, and gene structure of the *HvJRL* genes from barley. (**A**) The motif compositions of 32 HvJRL proteins. The motifs, numbers 1–10, are displayed in different colored boxes. The sequence information for each motif is shown in Figure S2. The length of protein can be estimated using the scale at the bottom. The capital letters on the left of gene ID were the subfamilies to which the corresponding genes belong. (**B**) Exon/intron structure of *HvJRL* genes. Yellow boxes represent exons, and black lines represent introns. The upstream/downstream regions of *HvJRL* genes are indicated in green boxes. The length of exons can be inferred from the scale at the bottom.

In addition, ten conserved motifs were predicted in the HvJRL proteins (Figure S2). Almost all members of the *HvJRL* family contain motif 3. Motifs 1, 3, 4, and 5 together make up the jacalin domain (Figure 2A). HvJRL proteins consist of 2–32 motifs, and their motif compositions in the same subfamily were generally similar (Figure 2A). For example, motifs 1, 3, 4, 5, 6, 7, 8, and 9 existed in almost all members of subfamily A. Interestingly, some motifs were unique to one or more subfamilies. For instance, motifs 6, 7, 8, and 10 are specific to the members of subfamily A.

2.4. Chromosomal Distribution and Gene Duplication of HvJRL Genes

The 32 *HvJRL* genes were unevenly distributed on the seven barley chromosomes, with most located at both ends of the chromosomes (Figure 3). More than 59% of the *HvJRL* genes were found on Chr5H (7), Chr1H (6), and Chr2H (6), and only one was located on Chr6H (Figure 3). In addition, HORVU0Hr1G031210 was distributed on scaffolds (Chr0H) (Table 1).



Figure 3. Chromosomal distribution and duplication events for JRL genes in barley. The tandem duplicated genes are highlighted in yellow, and the segmental duplicated genes are linked by red lines. The scale bar on the left indicates the length (Mb) of barley chromosomes.

The duplication of *HvJRL* genes was determined using a Multiple Collinearity Scan toolkit (MCScanX) [47]. At least two tandem duplicated regions on Chr2H and Chr5H and nine segmental duplicated gene pairs on Chr5H and Chr7H were found in the *HvJRL* family (Figure 3 and Table S3). This result indicated that tandem and segmental duplication are important drivers for the generation of new *HvJRL* genes [10]. Moreover, each pair of duplicated genes belongs to the same subfamily, sharing the same domain composition and a similar motif composition and gene structure (Figures 2 and S1).

2.5. Analysis of cis-Acting Elements in HvJRL Promoters

cis-Acting elements are of importance for the regulation of gene expression. To analyze the potential regulatory mechanisms of JRL genes, the 1500 bp upstream promoter sequences from the start codon of *HvJRLs* were used to predict *cis*-acting elements through PlantCARE [48]. A large number of *cis*-elements were found in *HvJRL* promoters, which mainly contained hormone-responsive elements and stress-responsive elements (Figure 4, Table S4). For the hormone-responsive *cis*-acting elements, there were the MeJA response-related TGACG motif and the CGTCA motif; abscisic acid response-related ABRE; auxin response-related AuxRR-core and TGA-box; gibberellins response-related GARE motif, TATC-box, and P-box; coupled with the salicylic acid response-related TCA-element. The abiotic and biotic stress-responsive *cis*-acting elements included elements associated with low-temperature response TC-rich repeats (Figure 4, Table S4). The hormone- and abiotic stress-responsive *cis*-acting elements included stress response to *cis*-acting elements and stress response to *cis*-acting elements included elements associated with low-temperature response TC-rich repeats (Figure 4, Table S4). The hormone- and abiotic stress-responsive *cis*-acting elements in almost all the *HvJRL* promoters, suggesting that *HvJRL* genes may play extensive roles in abiotic stresses and be involved in multiple plant hormone signaling pathways.



Figure 4. Predicted *cis*-elements in HvJRL promoters. Promoter sequences (-1500 bp) of 32 HvJRL genes were analyzed by PlantCARE. The upstream length to the translation starting site can be inferred according to the scale at the bottom. The capital letters on the left of gene ID were the subfamilies to which the corresponding genes belong.

2.6. Evolutionary Analysis of JRL Gene Family

To clarify the evolutionary relationships of JRL proteins between barley and other species, protein sequences from counterparts in barley, Arabidopsis, and rice, coupled with the verified functional JRLs in other plants, were used to construct the phylogenetic tree. A total of 48 *AtJRLs* (Table S5), which were also reported by Nagano et al. (2008) [9], were found by re-examining the Arabidopsis genome. All the 30 *OsJRLs* (Table S5) discovered by re-examining the rice genome were previously identified by Song et al. (2014) [3] and Han et al. (2018) [2], but LOC_Os01g25150 reported by Song et al. (2014) [3] has no JRL domain. Finally, an ML phylogenetic tree was constructed using protein sequences including 48 AtJRLs, 30 OsJRLs, 32 HvJRLs, and 12 JRLs in other plants (Figure 5). The results showed that the 122 genes could be grouped into 22 subgroups (subgroup 1~22). Notably, there were no HvJRLs in subgroups 9 and 12–22 but only JRLs from dicotyledonous plants, while subgroups 1–8 and 10 only contained JRLs from monocots, indicating that plant JRLs may have been functionally differentiated between monocots and dicots.

The potential functions of HvJRLs were predicted based on their functionally known homologs in the same subgroup (Figure 5). Numerous members in subgroups 1, 6, 7, 10, and 11 were involved in the response to biotic and abiotic stresses, such as pathogen resistance [16,21,49,50], response to hormones [26,27], salt tolerance [30,31,51], and drought tolerance [25]. The members of subgroup 1 might also be associated with beta-glucosidase aggregation [35,52], vernalization signaling, and spike development regulation [11,12]. The detailed functions of HvJRLs still need further study. The analysis has important reference



value, although the evolutionary relationships could not be accurately deciphered for the functions of all the HvJRLs.

Figure 5. Phylogenetic analysis of JRL proteins among different species. A phylogenetic tree was constructed using MEGA v.10.1.8 with full-length amino acid sequences of JRLs from barley, Arabidopsis, rice, and those functionally known in other species, including TaWC1 (TAU32427), Moringa G (AY048576), PALb (AB099933), Heltuba (AF064029), Banlec-1 (AF001527), Calsepa (U56820), Ipomoelin (D89823), TaJRLL1 (HQ317136), TaJRLL1 (HQ317136), TaJA1 (AY372111), SbBGAF1 (ABI24164), and TaVER2 (BAA32786), and the number of bootstrap test replicates was set as 1000. Sequence accessions of functionally known proteins have been highlighted in red and green. The different-colored arcs indicate different subgroups of the JRLs. The numbers on the outside of the arc represent the subgroups to which the corresponding proteins belong.

2.7. Expression Profiles of the HvJRLs in Response to Low-Nitrogen Stress

Many plant JRLs were confirmed to be involved in the biotic and abiotic stress responses. Combined with the phylogenetic and *cis*-element analyses, almost all the *HvJRLs* were potentially related to stress responses. However, it is unclear whether any *HvJRL* responds to nutritional stress. Thus, transcriptome data of barley roots under LN stress in our previous studies were employed and re-analyzed (Tables S6 and S7). Two barley genotypes (XZ149, LN-tolerant, and XZ56, LN-sensitive) were used for the RNA-seq analysis at 6 h, 48 h, and 12 d after LN stress [37,38]. Differentially expressed genes (DEGs) were screened with the threshold of FDR < 0.05 and FPKM \geq 1 in at least one replicate. Finally, nine DEGs which encode eight HvJRL proteins were identified in the two genotypes (Figure 6A and Table S8). The RNA-Seq data were validated using nine *HvJRL* DEGs through real-time PCR analysis. The results showed that expressions of the identified *HvJRL* DEGs from real time PCR were highly consistent with that from RNA-Seq analysis (Figure S3).



Figure 6. Expression profiles of the *HvJRL* DEGs. (**A**) Expression profiles of the *HvJRL* genes from the transcriptome in XZ149 and XZ56 at 6 h, 48 h, and 12 d under LN stress. Expression data were the ratio of FPKM of LN stress to that of the control at each time point. (**B**) Expression profiles (FPKM) of the *HvJRL* genes in different tissues of barley cultivar Morex. The color scale represents relative expression levels from high (red) to low (blue). The numbers on the left of each heat map are the subgroups to which the corresponding DEGs belong.

Of the nine *HvJRL* DEGs, three were down-regulated and six up-regulated in their expression under LN stress. All the down-regulated DEGs were only changed in XZ149 but not changed in XZ56. All up-regulated DEGs except one were changed both at 6 h and 48 h in XZ149, but they only changed at 6 h or were not changed in XZ56. Interestingly, the fold change of these up-regulated DEGs was larger in the tolerant genotype XZ149 compared with the sensitive genotype XZ56 (Figure 6A). Noteworthily, only one DEG (HORVU7Hr1G121120, MLOC_34131) was found at 12 d after LN stress; it was up-regulated at 6 h and 48 h in XZ149 and at 6 h and 12 d in XZ56.

To further understand the tissue expression patterns of these *HvJRL* DEGs, the transcriptomic data in eight plant tissues of Morex were downloaded from the Expression Atlas database (https://www.ebi.ac.uk/gxa/plant/experiments (accessed on 21 September 2023)). A total of 30 *HvJRLs*, containing all the DEGs, were found expressed in the examined tissues from the available data (Table S9). And no DEG was expressed in all the eight tissues (Figure 6B), that is to say their expression was tissue-specific. Remarkably, all DEGs were expressed in roots, and four up-regulated DEGs showed the highest expression in roots among the eight tissues.

Generally, the DEGs in the same subgroup showed a similar expression pattern in their response to LN stress and in tissues, such as HORVU5Hr1G009040 and HORVU7Hr1G059330

in subgroup 7. However, there were also distinct expression patterns among the members of the same subgroup, for instance, HORVU2Hr1G013940 and HORVU3Hr1G082370 in subgroup 1, indicating that they may function in LNSR in different ways.

2.8. HvHorcH Enhanced Low-Nitrogen Tolerance in Transgenic Arabidopsis

To confirm the RNA-seq results and to determine the function of *HvJRLs* in LNSR, *HvHorcH* (HORVU7Hr1G059330) was taken as a candidate gene, as two DEGs (MLOC_76780 and MLOC_1563) were both aligned to this gene, and their fold changes of the up-expression were the highest in LN-tolerant genotype XZ149 under LN stress (Figure 6A). Firstly, the tissue and LN responsive expression patterns of *HvHorcH* were analyzed in XZ149 and XZ56. The real-time PCR result showed that the expression of *HvHorcH* was the highest in the root among the examined tissues for both the two genotypes (Figure S4), which was in accordance with that in Morex (Figure 6B). To investigate the dynamic expression of *HvHorcH* in response to LN stress, seven time points (1 h, 3 h, 6 h, 24 h, 2 d, 6 d, and 12 d) were set after LN treatment (Figure S5). For leaves, *HvHorcH* was up-regulated at all the time points in XZ149, while it was down-regulated at 1 h, 3 h, 6 h, 24 h, and 12 d in XZ56. For roots, *HvHorcH* was up-regulated from 3 h to 2 d in XZ149, while it was only up-regulated at 3 h and 6 h in XZ56. In general, *HvHorcH* was mainly expressed in roots, and its expression pattern under LN stress was different in XZ149 and XZ56.

To further validate the function of *HvHorcH* in imparting LN tolerance, transgenic Arabidopsis plants overexpressing this gene were generated. Three independent transgenic lines, L1, L2, and L3, with high expression levels of *HvHorcH*, were used to evaluate the LN stress tolerance. The four-day-old of wild-type (WT) seedlings and three transgenic lines were treated for 7 d with a medium containing 0.1 mM NO₃⁻ (10 mM NO₃⁻ as a control) [53]. The growth performance of the three overexpressed lines was better than that of the WT, when exposed to LN stress, while there was no significant difference in the growth phenotype between the WT and transgenic lines in the control (Figure 7). In other words, all three lines displayed a significantly improved LN tolerance compared with the WT plants.



Figure 7. *HvHorcH* transgenic Arabidopsis showed enhanced low-nitrogen tolerance. The WT and three independent transgenic lines (L1, L2, and L3) were subjected to LN stress for 7 d and analyzed. (**A**,**B**) Morphology, (**C**) shoot fresh weight, (**D**) root length, (**E**) lateral root number. Four biological replications were performed. Different lowercase letters indicate a significant difference between treatments and genotypes at *p* < 0.05.

3. Discussion

JRL proteins are widely distributed in plants and are found in many species, including Arabidopsis, rice, wheat, maize, barley, *Sorghum bicolor*, and *Musa acuminata* [2]. Several of them have been demonstrated to be involved in plant growth regulation and stress responses [3,10,28,30,31,54]. Owing to their essential roles, the identification and characterization of JRL gene families and their expression profiles in response to abiotic stresses have been performed at the genome-wide level in several plant species, such as rice, wheat, and moso bamboo [2,3,10]. However, until now, no systematic studies have been reported on *HvJRLs* in barley nor on their roles in response to nutrition deficiency. Crops often suffer LN stress, as N deficiency is a common issue in agricultural production worldwide. Because JRL genes could take roles in so many environmental stress responses, such as salt, drought, and low-temperature stresses, we wondered if any of them respond to LN stress. In the present study, we performed a comprehensive investigation of *HvJRLs* and their expression under LN stress.

Here, a total of 32 HvJRLs were identified in barley combined from a BlastP and HMMER search (Table 1). Moreover, 48 AtJRLs and 30 OsJRLs were identified through a re-examined Arabidopsis and rice genome (Table S5). It has been reported that there were 46 to 123 JRL genes in cruciferous plants and 20 to 41 in graminaceous plants [2], which was accordance with our results. The accuracy of the *HvJRL* gene screening was further confirmed by multiple sequence alignment analysis, as all the 32 HvJRL proteins had jacalin domains (Figure 1). The HvJRL proteins have low sequence similarity, but the key sequences in their jacalin domain are relatively conserved, similar to those of Heltuba in Helianthus *tuberosus* and PeJRLs in moso bamboo [10,46]. According to the phylogenetic analysis, HvJRLs could be categorized into seven subfamilies, which was favored by the similar conserved domain and exon/intron structure in the same subfamily (Figure 2). Tandem duplication and segmental duplication were identified in the HvJRL family, and each duplicated pair existed in the same subfamily (Figures 2 and 3), further confirming their similar evolutionary origins and biological functions. On the other hand, gene duplication acts as a primary driver allowing for sub-functionalization and neofunctionalization [55]. Herein, the duplicate genes may show different response patterns under LN stress, for example, HORVU2Hr1G013940 and HORVU2Hr1G014020, where one responds to LN stress and the other does not (Figure 6), indicating that functional divergence occurred as the *HvJRL* gene family expanded.

Phylogenetic relationships showed that most JRL genes from monocots and dicots have been grouped in some individual subgroups, indicating the basic characteristics' divergency in this gene family between monocot and dicot plants. Interestingly, HvJRLs and OsJRLs are often clustered together, suggesting a closer relationship. The orthologous genes in the same subgroup often share similar functional roles. In this research, many HvJRL proteins were classified into some functional subgroups of functionally known JRLs, which provided a valuable reference for the study of the function of HvJRLs. Most functionally known JRLs in subgroup 1 are involved in agglutinating activity [35,50,52,56], which demonstrated that proteins from the same subgroup may display similar roles. Meanwhile, two members in subgroup 7, Orysata and HvHorcH, both confer a tolerance to salt stress [31,51], and the member OsJRL40 in subgroup 10 also positively regulated salt stress tolerance [30]. Both OsJAC1 and TaJA1 in subgroup 1 were associated with pathogen resistance [21,49,50], and TaJRLL1 in subgroup 6 was also involved in pathogen resistance [16]. These results suggest that there is functional redundancy in JRLs in different subfamilies. Therefore, the HvJRLs in subgroups 7 and 10 may play important roles in salt stress responses, while the HvJRLs in subgroups 1 and 6 participate in pathogen resistance. Notably, HvJRL proteins may serve similar functions in different ways. For instance, some *HvJRL* DEGs displayed diverse expression patterns under LN stress (Figure 6).

Members of the same subgroup also show multiple functions in their response to stress. For instance, in subgroup 8, Banlec-1 from *Musa acuminata* could prevent viral infection [57]; Calsepa in *Calystegia sepium* displayed agglutinating activity [56]; and Ipomoelin in *Ipomoea*

batatas showed JA-mediated wound induction and antioxidation [58,59]. Additionally, some individual JRLs could display functions in multiple stress responses. For example, Orysata is induced by various stresses, including salt, drought, hormones, senescence, pathogen attack, and insect attack [20,23,28,60–62]. The diversity of multiple *cis*-elements in each *HvJRL* promoter may also indicate that one *HvJRL* has multiple functions in the complex environment. For example, Horcolin was related with agglutinating activity and jasmonic acid induction, and it also enabled the specific inhibition of cellular HIV infection [33,34,63].

Consequently, most *HvJRLs* were supposed to be involved in stress responses, such as salt tolerance in subgroup 7. However, no studies on the role of JRLs in nutrition deficiency has been published. To clarify whether JRLs participate in the LNSR, the transcript profiles of *HvJRLs* were analyzed employing two barley genotypes with differences in low-nitrogen tolerance. As a result, nine DEGs encoding eight HvJRL proteins were identified under LN stress (Figure 6A). In general, the fold change of six up-regulated DEGs was higher in XZ149 than that in XZ56, and the increase in the abundance of their transcripts lasted longer in XZ149 than in XZ56 (Figure 6A). Notably, MLOC_4840 (HORVU7Hr1G030210) and MLOC_1563 (*HvHorcH*) were up-regulated only in XZ149 (Figure 6A). Moreover, both MLOC_1563 and MLOC_76780 were sequenced to the *HvHorcH* gene. Thus, *HvHorcH* might play key roles in LNSR. To further reveal the function of *HvHorcH* in LNSR, the 35S:*HvHorcH* transgenic Arabidopsis was obtained. The transgenic Arabidopsis seedlings exhibited enhanced LN tolerance (Figure 7). The specific function requires further verification and discussion in the future.

4. Materials and Methods

4.1. Identification of JRL Genes in Barley

Genome data of Morex barley were downloaded from the Phytozome v.13 database (https://phytozome-next.jgi.doe.gov/ (accessed on 6 August 2023)). The Hidden Markov Model (HMM) profile of the jacalin-like lectin domain (PF01419) was obtained from the PFAM database (https://www.ebi.ac.uk/interpro/entry/pfam/#table (accessed on 10 August 2023)) [39]. The HvJRL proteins were identified from the barley protein database using the HMMER3 program with a threshold of e-value $< 1 \times 10^{-20}$ [40]. Meanwhile, 30 OsJRL proteins reported by Han et al. (2018) were used as query sequences to search for predicted HvJRL proteins with a threshold of e-value $< 1 \times 10^{-10}$ [41]. After eliminating redundant and incomplete protein sequences, the predicted HvJRL sequences were confirmed by the presence of the jacalin domain using NCBI-CDD (https://www.ncbi.nlm.nih.gov/ cdd/ (accessed on 10 August 2023)) [42] and SMART (http://smart.embl-heidelberg.de/ (accessed on 10 August 2023)) [43]. Representative sequences of HvJRL proteins obtained by Tbtools v.1.120 [64] were taken for subsequent analysis. The molecular weight (MW), number of amino acids, and isoelectric point (pI) of HvJRL proteins were analyzed using ExPASy v.7 (https://web.expasy.org/protparam/ (accessed on 12 August 2023)) [44]. Subcellular localization predictions were generated through WoLF PSORT (https://wolfpsort.hgc.jp/ (accessed on 12 August 2023)) [45].

4.2. Sequence Alignment and Phylogenetic Analysis

The alignment of the core jacalin domain sequences of HvJRLs were carried out with ClustralW (http://www.genome.jp/tools/clustalw/ (accessed on 16 August 2023)) [65], and the conserved sequences were visualized with DNAMAN (v.2). Genome data of Arabidopsis and rice were retrieved from the Arabidopsis Information Resource database (http://www.arabidopsis.org (accessed on 6 August 2023)) and the Rice Genome Annotation Project (http://rice.uga.edu/ (accessed on 6 August 2023)), respectively. The AtJRLs in Arabidopsis and OsJRLs in rice were re-examined from the corresponding protein database employing the methods described above. Finally, 50 AtJRL and 30 OsJRL protein sequences were identified and used for phylogenetic analysis. An ML phylogenetic tree

with 32 HvJRLs and a tree containing 112 JRL proteins from barley, Arabidopsis, rice, and other plants were constructed with 1000 bootstrap replicates in MEGA v.10.1.8, respectively.

4.3. Gene Structure, Conserved Motif, and Domain Analysis

The alignment of the cDNAs and their corresponding genomic DNA sequences of *HvJRLs* was performed to determine their exon/intron structure. The conserved motifs in HvJRLs were obtained using the MEME 5.5.1 online software (https://meme-suite.org/meme/ (accessed on 12 August 2023)) [66]. A visual map of gene structures and motif compositions was displayed through TBtools v.1.120 [64]. The conserved structural domains identified with the NCBI CDD and SMART were also visualized using TBtools v.1.120 [64].

4.4. Chromosomal Distribution and Duplication of HvJRL Genes

The chromosome location of *HvJRLs* was mapped based on their specific positions in the barley genome by TBtools v.1.120 [64]. The gene duplications of *HvJRLs* were determined by MCScanX (http://chibba.pgml.uga.edu/mcscan2/#tm (accessed on 17 August 2023)) [47]. If the alignment sequence covers more than 70% of the longer gene and the sequence identity is greater than 70%, it was defined as a gene duplication event [67]. The collinearity relationships were visualized using TBtools v.1.120 [64].

4.5. cis-Elements in the Promoter Regions of HvJRLs

To gain a preliminary understanding of the potential regulatory mechanism of the *HvJRL* genes, the 1500 bp upstream sequences from the start codon of the *HvJRLs* were used to determine the *cis*-elements' distribution in PlantCare (http://bioinformatics.psb. ugent.be/webtools/plantcare/html/ (accessed on 16 August 2023)) [48]. The stress-related *cis*-elements in each promoter region were visualized using TBtools v.1.120 software [64].

4.6. Vector Construction and Arabidopsis Transformation

The full-length coding sequence of *HvHorcH* was cloned in LN-tolerant genotype XZ149. In brief, total RNA was extracted from the LN stress treated roots of XZ149. After quality monitoring, first-strand cDNA was synthesized with oligo dT primer and Random 6 mers (Takara, Kyoto, Japan). And then *HvHorcH* was cloned using its gene specific primer. It was then recombined into a pCAMBIA1300 vector with a flag tag. The recombinant 35S:*HvHorcH*-flag vector was used to transform *Agrobacterium tumefaciens* GV3101. And then the MES suspensions containing transformed Agrobacterium were transferred into the inflorescence of Arabidopsis (Col-0). The positive transgenic lines were screened through genomic DNA PCR with a *HvHorcH*-specific primer. The expression of the *HvHorcH* was analyzed by semiquantitative RT-PCR. The primers are listed in Table S10.

4.7. Plant Materials and Treatments

The transcriptome analysis was carried out using roots of XZ149 and XZ56. The barley seedlings were hydroponic with the basic culture solution containing 2 mM NaNO₃ in a plant growth chamber (22/18 °C, day/night). Three-leaf-stage seedlings were treated with the nutrient solution containing 0.2 mM N (2 mM N as control) according to Yang et al. (2014) [68]. The culture medium was continuously aerated with an air pump and renewed every five days. The details were the same as those reported by Quan et al. (2016) [37]. The samples were taken at 6 h, 48 h, and 12 d under treatment. Finally, 24 samples (2 treatments (control and LN stress) × 3 time points (6 h, 48 h, and 12 d after LN stress) × 2 genotypes (XZ149 and XZ56) × 2 biological replications) were prepared for RNA-seq analysis. For real-time PCR analysis, we performed the same experiment as above, and then leaves and roots were taken 1 h, 3 h, 6 h, 24 h, 48 h, 6 d, and 12 d after LN stress.

The Arabidopsis genotypes used for phenotypic analysis under LN stress were the Col-0 ecotype (WT) and homozygous 35S:*HvHorcH* T3 transgenic Arabidopsis lines. The experiment was carried out in a plant growth chamber (25 °C) with a photoperiod of

16 h light/8 h dark and a light intensity of 100 μ mol m⁻² s⁻¹. The Arabidopsis seeds were vernalized for three days and then cultivated for four days with basic medium in 13 cm \times 13 cm vertical agar plates. The seedlings with uniform size were treated with the medium of 1 mM NO₃⁻ (10 mM NO₃⁻ as control). The basic medium was prepared according to Remans et al. (2006) [53] with a small modification, containing 10 mM KNO₃, 2.5 mM MES, 1 mM KH₂PO₄, 0.5 mM CaSO₄, 0.5 mM MgCl₂, 50 μ M NaFe(III)-EDTA, 50 μ M H₃BO₃, 12 μ M MnCl₂, 1 μ M ZnCl₂, 1 μ M CuCl₂, and 0.03 μ M (NH₄)₆Mo₇O₂₄. The K concentration was supplemented to 10 mM by adding KCl in the media for the LN treatment [69]. The growth phenotype was determined after 7 days under LN stress.

4.8. Transcriptome Analysis of HvJRL Genes in Response to LN Stress

A total of 24 cDNA libraries were constructed. Briefly, poly-A containing mRNA was obtained from the total RNA with poly-T oligo-attached magnetic beads. Following purification, the random fragmentation of mRNA was generated using divalent cations under elevated temperature and then reversely transcribed into cDNA. After adenylation of the 3' ends of cDNA fragments, sequence adapters were ligated on both ends. The purified cDNA library fragments were selectively amplified and enriched. Then, the purified and quantified PCR products were applied for the cluster generation processing. Subsequently, each final library was sequenced on an Illumina NextSeq 500 platform to generate 2×75 bp pair-end reads.

After removing the adapter sequences, empty sequences, and low-quality bases from raw reads, quality analysis of the clean data was performed. And then, the clean reads were aligned to the Morex barley genomes using the ultra-high-throughput short read aligner on TopHat 2.1.1 (https://ccb.jhu.edu/software/tophat/index.shtml (accessed on 20 August 2023)), followed by the identification of splice junctions.

The abundance of the gene expression was calculated using FPKM (fragments per kilobase of exon per million mapped fragments) [70]. The differential expression of the transcripts under LN stress was re-analyzed employing DESeq2 [71]. The thresholds of FDR < 0.05 and FPKM \geq 1 in at least one replicate were used to identify DEGs [72]. The DEGs were annotated using the genomes of Arabidopsis and Morex barely as reference. The fold change of *HvJRL* DEGs was calculated by the ratio of FPKM of LN stress/control and displayed in the heatmaps produced by TBtools v.1.120 software [64]. The datasets used are available on the following websites: https://doi.org/10.1186/s12870-016-0721-8 and https://doi.org/10.1186/s12870-019-1668-3 [37,38].

4.9. Tissue Expression Profile of JRL Gene Family in Barley

Transcriptomic data of Morex tissue were downloaded from the EMBL-EBI database (https://www.ebi.ac.uk/gxa/plant/experiments (accessed on 21 September 2023)) [73]. Eight different tissues included 4-day embryos dissected from germinating grains, shoots from seedlings (10 cm shoot stage), roots, developing inflorescence (5 mm and 1–1.5 cm), the third internode at the six-leaf visible stage, and 5 and 15 days post-anthesis developing grain. The tissue expression levels of *HvJRL* genes were obtained from the data and presented by heatmap.

4.10. Real-Time PCR

RNA of each sample was extracted with the FastPure Plant Total RNA Isolation Kit (Vazyme, Nanjing, China), followed by the first-strand cDNA synthesis using Hiscript III Reverse Transcriptase (Vazyme, Nanjing, China). The gene-specific primers were designed by primer-blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/ (accessed on 12 September 2023)), and *HvGAPDH* was used as the internal control gene to normalize all the data (Table S10). The real-time PCR was conducted in three biological replicates and three technical replicates on a CFX96 system (Bio-Rad, Hercules, CA, USA). The comparative CT method was used to calculate the relative expression which means that the fold change referred to the expression in the control [74].

15 of 18

4.11. Statistical Analysis

Significant differences for physiological traits were tested with Duncan's Multiple Range Test on SPSS 16.0 software, and the difference at p < 0.05 was considered significant. The bar charts were generated in SigmaPlot 10.0.

5. Conclusions

In the present study, the barley JRL family genes were systematically characterized. The comparative transcriptome data indicated that at least eight *HvJRL* genes were associated with LNSR in barley. The expression pattern of *HvHorcH* under LN stress was further analyzed, and its potential functions in LNSR were also confirmed using the transgenic method. The *HvJRL* DEGs in XZ149 may be of significance for further functional study of the JRL genes in LNSR. This work may provide new candidate genes for breeding barley cultivars with LN tolerance.

Supplementary Materials: The supporting information can be downloaded at https://www.mdpi. com/article/10.3390/ijms242316641/s1.

Author Contributions: X.Q. and W.H. designed the research. X.Q., C.M., C.X., H.S., B.X. and R.S.B. performed research. X.Q. analyzed data and wrote the manuscript. R.S.B. and W.H. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of Shandong Province (ZR2023MC114), National Natural Science Foundation of China (31901437), and Shandong Province Double Hundred Talent Plan (WSG2020001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the first author on reasonable request, and her email address is bio_quanxy@ujn.edu.cn.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

DEG: differentially expressed gene; DPA, day post-anthesis; FPKM, fragments per kilobase of exon per million fragments mapped reads; LN, low nitrogen; LNSR, LN stress response; MCScanX, Multiple Collinearity Scan toolkit; ML, maximum likelihood; N, nitrogen.

References

- 1. Jung, I.J.; Ahn, J.W.; Jung, S.; Hwang, J.E.; Hong, M.J.; Choi, H.I.; Kim, J.B. Overexpression of rice jacalin-related mannose-binding lectin (OsJAC1) enhances resistance to ionizing radiation in Arabidopsis. *BMC Plant Biol.* **2019**, *19*, 561. [CrossRef]
- Han, Y.J.; Zhong, Z.H.; Song, L.L.; Stefan, O.; Wang, Z.H.; Lu, G.L. Evolutionary analysis of plant jacalin-related lectins (JRLs) family and expression of rice JRLs in response to *Magnaporthe oryzae*. J. Integr. Agric. 2018, 17, 1252–1266. [CrossRef]
- Song, M.; Xu, W.Q.; Xiang, Y.; Jia, H.Y.; Zhang, L.X.; Ma, Z.Q. Association of jacalin-related lectins with wheat responses to stresses revealed by transcriptional profiling. *Plant Mol. Biol.* 2014, 84, 95–110. [CrossRef]
- Sharon, N. Lectins: Carbohydrate-specific reagents and biological recognition molecules. J. Biol. Chem. 2007, 282, 2753–2764. [CrossRef] [PubMed]
- Peumans, W.J.; Hause, B.; Van Damme, E.J. The galactose-binding and mannose-binding jacalin-related lectins are located in different sub-cellular compartments. *FEBS Lett.* 2000, 477, 186–192. [CrossRef] [PubMed]
- 6. Van Damme, E.J.; Lannoo, N.; Peumans, W.J. Plant lectins. Adv. Bot. Res. 2008, 48, 107–209.
- Raval, S.; Gowda, S.B.; Singh, D.D.; Chandra, N.R. A database analysis of jacalin-like lectins: Sequence-structure-function relationships. *Glycobiology* 2004, 14, 1247–1263. [CrossRef]
- 8. Bunn-Moreno, M.M.; Campos-Neto, A. Lectin(s) extracted from seeds of *Artocarpus integrifolia* (jackfruit): Potent and selective stimulator(s) of distinct human T and B cell function. *J. Immunol.* **1981**, 127, 427–429. [CrossRef]
- Nagano, A.J.; Fukao, Y.; Fujiwara, M.; Nishimura, M.; Hara-Nishimura, I. Antagonistic jacalin-related lectins regulate the size of ER body-type b-glucosidase complexes in *Arabidopsis thaliana*. *Plant Cell Physiol.* 2008, 49, 969–980. [CrossRef]

- Zhang, Z.; Huang, B.; Chen, J.; Jiao, Y.; Guo, H.; Liu, S.; Ramakrishnan, M.; Qi, G. Genome-wide identification of JRL genes in Moso Bamboo and their expression profiles in response to multiple hormones and abiotic stresses. *Front. Plant Sci.* 2022, 12, 809666. [CrossRef]
- 11. Yong, W.D.; Xu, Y.Y.; Xu, W.Z.; Wang, X.; Li, N.; Wu, J.S.; Liang, T.B.; Chong, K.; Xu, Z.H.; Tan, K.H.; et al. Vernalization induced flowering in wheat is mediated by a lectin-like gene VER2. *Planta* **2003**, *217*, 261–270. [CrossRef]
- 12. Xing, L.; Li, J.; Xu, Y.; Xu, Z.; Chong, K. Phosphorylation modification of wheat lectin VER2 is associated with vernalizationinduced O-GlcNAc signaling and intracellular motility. *PLoS ONE* 2009, *3*, e4854. [CrossRef]
- 13. Jiang, J.; Xu, Y.; Chong, K. Overexpression of *OsJAC1*, a lectin gene, suppresses the coleoptile and stem elongation in rice. *J. Integr. Plant Biol.* **2007**, 49, 230–237. [CrossRef]
- Xiao, J.; Li, C.; Xu, S.; Xing, L.; Xu, Y.; Chong, K. JACALIN-LECTIN LIKE1 regulates the nuclear accumulation of GLYCINE-RICH RNA-BINDING PROTEIN7, influencing the RNA processing of FLOWERING LOCUS C antisense transcripts and flowering time in Arabidopsis. *Plant Physiol.* 2015, *3*, 2102–2117.
- 15. Yamaji, Y.; Maejima, K.; Komatsu, K.; Shiraishi, T.; Okano, Y.; Himeno, M.; Sugawara, K.; Neriya, Y.; Minato, N.; Miura, C.; et al. Lectin-mediated resistance impairs plant virus infection at the cellular level. *Plant Cell*. **2012**, *24*, 778–793. [CrossRef]
- Xiang, Y.; Song, M.; Wei, Z.Y.; Tong, J.; Zhang, L.X.; Xiao, L.; Ma, Z.Q.; Wang, Y. A jacalin-related lectin-like gene in wheat is a component of the plant defence system. *J. Exp. Bot.* 2011, 62, 5471–5483. [CrossRef] [PubMed]
- Subramanyam, S.; Smith, D.F.; Clemens, J.C.; Webb, M.A.; Sardesai, N.; Williams, C.E. Functional characterization of HFR-1, a high mannose *N*-glycan-specific wheat lectin induced by Hessian fly larvae. *Plant Physiol.* 2008, 147, 1412–1426. [CrossRef] [PubMed]
- Subramanyam, S.; Sardesai, N.; Puthoff, D.P.; Meyer, J.M.; Nemacheck, J.A.; Gonzalo, M.; Williams, C.E. Expression of two wheat defense-response genes, *Hfr-1* and *Wci-1*, under biotic and abiotic stresses. *Plant Sci.* 2006, 170, 90–103. [CrossRef]
- 19. Qin, Q.M.; Zhang, Q.; Zhao, W.S.; Wang, Y.Y.; Peng, Y.L. Identification of a lectin gene induced in rice in response to *Magnaporthe* grisea infection. *Acta Bot. Sin.* **2003**, 45, 76–81.
- Atalah, B.A.; Smagghe, G.; Damme, E.V.J.P.S. Orysata, a jacalin-related lectin from rice, could protect plants against biting-chewing and piercing-sucking insects. *Plant Sci.* 2014, 221–222, 21–28. [CrossRef]
- Weidenbach, D.; Esch, L.; Moller, C.; Hensel, G.; Kumlehn, J.; Hofle, C.; Huckelhoven, R.; Schaffrath, U. Polarized defense against fungal pathogens is mediated by the jacalin-related lectin domain of modular Poaceae-specific proteins. *Mol. Plant* 2016, *9*, 514–527. [CrossRef] [PubMed]
- 22. Claes, B.; Dekeyser, R.; Villarroel, R.; Van den Bulcke, M.; Bauw, G.; Van Montagu, M.; Caplanet, A. Characterization of a rice gene showing organ-specific expression in response to salt stress and drought. *Plant Cell* **1990**, *2*, 19–27. [PubMed]
- Filho, G.A.S.; Ferreira, B.S.; Dias, J.M.; Queiroz, K.S.; Branco, A.T.; Bressan-Smith, R.E.; Oliveira, J.G.; Garcia, A.B. Accumulation of SALT protein in rice plants as a response to environmental stresses. *Plant Sci.* 2003, 164, 623–628.
- Paul, S.; Gayen, D.; Datta, S.K.; Datta, S.K. Dissecting root proteome of transgenic rice cultivars unravels metabolic alterations and accumulation of novel stress responsive proteins under drought stress. *Plant Sci.* 2015, 234, 133–143. [CrossRef] [PubMed]
- Sahid, S.; Roy, C.; Shee, D.; Datta, R.; Paul, S. Jacalin domain-containing protein OsSalT interacts with OsDREB2A and OsNAC1 to impart drought stress tolerance in planta. *Environ. Exp. Bot.* 2021, 183, 104362. [CrossRef]
- 26. Feng, H.; Xu, W.Z.; Lin, H.H.; Chong, K. Transcriptional regulation of wheat VER2 promoter in rice in response to abscisic acid, jasmonate, and light. *J. Genet. Genom.* 2009, *6*, 371–377. [CrossRef]
- Abebe, T.; Skadsen, R.W.; Kaeppler, H.F. A proximal upstream sequence controls tissue-specific expression of *Lem2*, a salicylateinducible barley lectin-like gene. *Planta* 2005, 221, 170–183. [CrossRef]
- Zhang, W.; Peumans, W.J.; Barre, A.; Astoul, C.H.; Rovira, P.; Rougé, P.; Proost, P.; Truffa-Bachi, P.; Jalali, A.A.; Van Damme, E.J.M. Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. *Planta* 2000, 210, 970–978.
- 29. He, X.; Li, L.; Xu, H.; Xi, J.; Cao, X.; Xu, H. A rice jacalin-related mannose-binding lectin gene, *OsJRL*, enhances *Escherichia coli viability* under high salinity stress and improves salinity tolerance of rice. *Plant Biol.* **2017**, *19*, 257–261. [CrossRef]
- Gao, Q.; Yin, X.; Wang, F.; Hu, S.; Liu, W.; Chen, L.; Dai, X.; Liang, M. OsJRL40, a jacalin-related lectin gene, promotes salt stress tolerance in rice. Int. J. Mol. Sci. 2023, 24, 7441. [CrossRef]
- 31. Witzel, K.; Matros, A.; Bertsch, U.; Aftab, T.; Rutten, T.; Ramireddy, E.; Melzer, M.; Kunze, G.; Mock, H. The jacalin-related lectin HvHorcH is involved in the physiological response of barley roots to salt stress. *Int. J. Mol. Sci.* **2021**, 22, 10248. [CrossRef]
- Zhang, H.; Deng, C.; Yao, J.; Zhang, Y.L.; Zhang, Y.N.; Deng, S.; Zhao, N.; Sa, G.; Zhou, X.; Lu, C.; et al. *Populus euphratica* JRL mediates ABA response, ionic and ROS homeostasis in Arabidopsis under salt stress. *Int. J. Mol. Sci.* 2019, 20, 815. [CrossRef] [PubMed]
- 33. Van Damme, E.J.; Zhang, W.; Peumans, W.J. Induction of cytoplasmic mannose-binding jacalin-related lectins is a common phenomenon in cereals treated with jasmonate methylester. *Commun. Agric. Appl. Biol. Sci.* 2004, *69*, 23–31. [PubMed]
- Grunwald, I.; Heinig, I.; Thole, H.H.; Neumann, D.; Kahmann, U.; Kloppstech, K.; Gau, A.E. Purification and characterization of a jacalin-related, coleoptile specific lectin from *Hordeum vulgare*. *Planta* 2007, 226, 225–234. [CrossRef]
- 35. Wang, X.M.; Ma, Q.H. Characterization of a jasmonate- regulated wheat protein related to a beta-glucosidase aggregating factor. *Plant Physiol. Biochem.* **2005**, *2*, 185–192. [CrossRef]

- 36. Quan, X.; Zeng, J.; Han, Z.; Zhang, G. Ionomic and physiological responses to low nitrogen stress in Tibetan wild and cultivated barley. *Plant Physiol. Biochem.* **2017**, *111*, 257–265. [CrossRef]
- 37. Quan, X.; Zeng, J.; Ye, L.; Chen, G.; Han, Z.; Shah, J.M.; Zhang, G. Transcriptome profiling analysis for two Tibetan wild barley genotypes in responses to low nitrogen. *BMC Plant Biol.* **2016**, *16*, 30. [CrossRef]
- 38. Quan, X.; Zeng, J.; Chen, G.; Zhang, G. Transcriptomic analysis reveals adaptive strategies to chronic low nitrogen in Tibetan wild barley. *BMC Plant Biol.* **2019**, *19*, 68. [CrossRef]
- 39. 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. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49*, 412–419. [CrossRef]
- Finn, R.D.; Clements, J.; Eddy, S.R. HMMER web server: Interactive sequence similarity searching. Nucleic Acids Res. 2011, 39, 29–37. [CrossRef] [PubMed]
- He, Z.; Li, L.; Luan, S. Immunophilins and parvulins. Superfamily of peptidyl prolyl isomerases in Arabidopsis. *Plant Physiol.* 2004, 134, 1248–1267. [CrossRef]
- 42. Marchler, B.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. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **2015**, *43*, 222–226. [CrossRef]
- 43. Letunic, I.; Khedkar, S.; Bork, P. SMART: Recent updates, new developments and status in 2020. *Nucleic Acids Res.* 2021, 49, 458–460. [CrossRef]
- 44. Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; Castro, E.D.; Flegel, V.; Fortier, A.; Gasteiger, E.; Ioannidis, V. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **2012**, *40*, 597–603. [CrossRef] [PubMed]
- 45. Horton, P.; Park, K.J.; Obayashi, T.; Fujita, N.; Harada, H.; Adams, C.C.J.; Nakai, K. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* 2007, *35*, 585–587. [CrossRef]
- 46. Bourne, Y.; Zamboni, V.; Barre, A.; Peumans, W.J.; Van Damme, E.J.; Rougé, P. *Helianthus tuberosus* lectin reveals a widespread scaffold for mannose-binding lectins. *Structure* **1999**, *7*, 1473–1482. [CrossRef] [PubMed]
- 47. Wang, Y.; Tang, H.; DeBarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.; Jin, H.; Marler, B.; Guo, H. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [CrossRef]
- Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 2002, 30, 325–327. [CrossRef] [PubMed]
- Ma, Q.H.; Tian, B.; Li, Y.L. Overexpression of a wheat jasmonate-regulated lectin increases pathogen resistance. *Biochimie* 2010, 92, 187–193. [CrossRef]
- 50. Ma, Q.H.; Zhen, W.B.; Liu, Y.C. Jacalin domain in wheat jasmonate-regulated protein Ta-JA1 confers agglutinating activity and pathogen resistance. *Biochimie* 2013, *2*, 359–365. [CrossRef] [PubMed]
- Hirano, K.; Teraoka, T.; Yamanaka, H.; Harashima, A.; Kunisaki, A.; Takahashi, H.; Hosolawa, D. Novel mannose-binding rice lectin composed of some isolectins and its relation to a stress-inducible *salT* gene. *Plant Cell Physiol.* 2000, 41, 258–267. [CrossRef]
- Kittur, F.S.; Lalgondar, M.; Yu, H.Y.; Bevan, D.R.; Esen, A. Maize beta-glucosidase-aggregating factor is a polyspecific jacalinrelated chimeric lectin, and its lectin domain is responsible for beta-glucosidase aggregation. J. Biol. Chem. 2007, 10, 7299–7311. [CrossRef]
- Remans, T.; Nacry, P.; Pervent, M.; Girin, T.; Tillard, P.; Lepetit, M.; Gojon, A. A central role for the nitrate transporter NRT2.1 in the integrated morphological and physiological responses of the root system to nitrogen limitation in Arabidopsis. *Plant Physiol.* 2006, 140, 909–921. [CrossRef] [PubMed]
- 54. Ray, S.; Kapoor, S.; Tyagi, A.K. Analysis of transcriptional and upstream regulatory sequence activity of two environmental stress-inducible genes, *NBS-Str1* and *BLEC-Str8*, of rice. *Transgenic Res.* **2012**, *21*, 351–366. [CrossRef] [PubMed]
- 55. Cannon, S.B.; Mitra, A.; Baumgarten, A.; Young, N.D.; May, G. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *BMC Plant Biol.* **2004**, *4*, 10. [CrossRef] [PubMed]
- Bourne, Y.; Roig-Zamboni, V.; Barre, A.; Peumans, W.J.; Astoul, C.H.; Van Damme, E.J.; Rougé, P. The crystal structure of the *Calystegia sepium* agglutinin reveals a novel quaternary arrangement of lectin subunits with a beta-prism fold. *J. Biol. Chem.* 2004, 279, 527–533. [CrossRef]
- 57. Liu, X.Y.; Li, H.; Zhang, W. The lectin from *Musa paradisiaca* binds with the capsid protein of tobacco mosaic virus and prevents viral infection. *Biotechnol. Biotec Eq.* **2014**, *3*, 408–416. [CrossRef] [PubMed]
- 58. Imanishi, S.; Kito-Nakamura, K.; Matsuoka, K.; Morikami, A.; Nakamura, K. A major jasmonate-inducible protein of sweet potato, ipomoelin, is an ABA-independent wound-inducible protein. *Plant Cell Physiol.* **1997**, *38*, 643–652. [CrossRef]
- 59. Chang, W.C.; Liu, K.L.; Hsu, F.C.; Jeng, S.T.; Cheng, Y.S. Ipomoelin, a jacalin-related lectin with a compact tetrameric association and versatile carbohydrate binding properties regulated by its N terminus. *PLoS ONE* **2012**, *7*, e40618. [CrossRef]
- Chen, P.Y.; De Schutter, K.; Pauwels, J.; Gevaert, K.; Van Damme, E.J.M.; Smagghe, G. The lectin Orysata induces phosphatasemediated and carbohydrate-independent aggregation of insect cells. *J. Insect Physiol.* 2021, 131, 104241. [CrossRef]
- 61. Lee, R.H.; Wang, C.H.; Huang, L.T.; Chen, S.C.G. Leaf senescence in rice plants: Cloning and characterization of senescence up-regulated genes. *J. Exp. Bot.* **2001**, *52*, 1117–1122. [CrossRef]
- 62. Sofie, V.H.; Kristof, D.S.; Lore, E.; Mariya, T.; Dang, L.; Els, V.D. Comparative study of lectin domains in model species: New insights into evolutionary dynamics. *Int. J. Mol. Sci.* 2017, *18*, 1136–1164.

- Jayaprakash, N.G.; Singh, A.; Vivek, R.; Yadav, S.; Pathak, S.; Trivedi, J.; Jayaraman, N.; Nandi, D.; Mitra, D.; Surolia, A. The barley lectin, horcolin, binds high-mannose glycans in a multivalent fashion, enabling high-affinity, specific inhibition of cellular HIV infection. J. Biol. Chem. 2020, 295, 12111–12129. [CrossRef]
- 64. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant.* **2020**, *13*, 1194–1202. [CrossRef]
- 65. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, 22, 4673–4680. [CrossRef]
- 66. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res.* 2009, *37*, W202–W208. [CrossRef]
- 67. Gu, Z.; Cavalcanti, A.; Chen, F.C.; Bouman, P.; Li, W.H. Extent of gene duplication in the genomes of Drosophila, nematode, and yeast. *Mol. Biol. Evol.* 2002, *19*, 256–262. [CrossRef] [PubMed]
- Yang, L.; Hu, H.; Zhu, B.; Jin, X.; Wu, F.; Zhang, G. Genotypic variations of N use efficiency in Tibetan wild and cultivated barleys. J. Zhejiang Univ. Agric. Life Sci. 2014, 40, 155–164.
- Kiba, T.; Feria-Bourrellier, A.B.; Lafouge, F.; Lezhneva, L.; Boutet-Mercey, S.; Orsel, M.; Brehaut, V.; Miller, A.; Daniel-Vedele, F.; Sakakibara, H.; et al. The Arabidopsis nitrate transporter NRT2.4 plays a double role in roots and shoots of nitrogen-starved plants. *Plant Cell* 2012, 24, 245–258. [CrossRef] [PubMed]
- Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26, 139–140. [CrossRef]
- 71. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
- 72. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; Baren, M.J.V.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 2010, 28, 511–515. [CrossRef] [PubMed]
- Mayer, K.F.X.; Waugh, R.; Brown, J.W.S.; Schulman, A.; Langridge, P.; Platzer, M.; Fincher, G.B.; Muehlbauer, G.J.; Sato, K.; Close, T.J.; et al. A physical, genetic and functional sequence assembly of the barley genome. *Nature* 2012, 491, 711–716. [PubMed]
- 74. Schmittgen, T.D.; Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 2008, *3*, 1101–1108. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.