# A NEW VRS1 ALLELE IDENTIFIED IN 2-ROW SPANISH LANDRACES

**Ana M Casas<sup>1</sup>,** Bruno Contreras-Moreira<sup>1,6</sup>, Shun Sakuma<sup>2</sup>, María Pilar Gracia<sup>1</sup>, Marian Moralejo<sup>3</sup>, José Luis Molina-Cano<sup>4</sup>, **Takao Komatsuda<sup>5</sup>**, **Ernesto Igartua<sup>1</sup>** 

<sup>1</sup> Estación Experimental de Aula Dei-CSIC (EEAD-CSIC), Avda Montañana 1005, 50059 Zaragoza, España; <sup>2</sup> Kihara Institute, Yokohama City University, Japan;

<sup>3</sup> Universitat de Lleida, Lleida, Spain, <sup>4</sup> Institut de Recerca i Tecnología Agroalimentàries (UdL-IRTA), Lleida, Spain,

<sup>5</sup> National Institute of Agrobiological Sciences (NIAS), Tsukuba, Japan, <sup>6</sup> Fundación ARAID, Zaragoza, Spain



brought to you by  $\widetilde{\mathbb{I}}$  CORE

# INTRODUCTION

*Vrs1*, the gene determining the type of spike in barley, has been extensively studied. The wild dominant allele encodes a homeodomain-leucine zipper transcription factor whose activity results in a two-rowed spike, whereas the recessive allele produces a six-row phenotype (Komatsuda et al. 2007). At least three alleles in two-rowed types and four different alleles in six-rowed barleys have been described.

Previous results using MWG699, a marker closely linked to *Vrs1*, suggested different geographic origins for six-row alleles. Among them, the *vrs1.a2* allele originated in the Western Mediterranean. A large proportion of Spanish and Moroccan six-row landraces, some two-row landraces as well as three wild barleys, all shared the same haplotype in MWG699. We analyze *Vrs1* sequence variation in those materials.

araid

## METHODOLOGY

**PLANT MATERIAL.** This study involved 215 accessions: 177 Spanish landraces (51 two-rowed and 126 six-rowed); 7 Moroccan landraces (3 two-rowed, 4 six-rowed), 3 wild, *Hordeum spontaneum* from Morocco and 28 barley cultivars (15 two-rowed and 13 six-rowed).

**GENOTYPING.** The haplotype for MWG699/Taql (A, D or K) was analyzed in all the samples as reported (Tanno et al. 2002; Casas et al. 2005). Also, the accessions were genotyped with a panel of 322 polymorphic SNP markers, including two within the *Vrs1* sequence. To confirm the obtained results, *Vrs1* was sequenced in selected genotypes.

**SEQUENCE COMPARISONS.** Multiple alignment of protein sequences was performed with Clustal Omega. The haplotype network was computed with TCS.

**POPULATION STRUCTURE.** The software STRUCTURE was used to infer the population structure under an admixture model. Principal Coordinate Analysis was conducted, on the same dataset, using the software DARwin, and the Simple Matching coefficient.

# RESULTS

Nine polymorphisms identified within the *Vrs1* locus allow differentiating seven alleles (4 two-rowed and 3 six-rowed), including novel allele *Vrs1.b5*.



**Figure 1.** Partial nucleotide sequences showing polymorphisms scored in three alleles of *Vrs1*. Arrows point to indels affecting the open reading frame.

440 440 440 440 440 440 550 5310 550 550 550 550 550 550 550 550 550 5
Vrs1.b2
Millellellellellellellellellellellellelle
- Eat Options Help
pen Save Export Print Next Find a Save Section   410 440 450 460 470 480 480 320
2 El V B - 1 A Sample 1416418
420 440 450 440 470 400 490 500 510
Vrs1.b5
↑

Most Spanish landraces are six-rowed and carry the *vrs1.a2* allele (ex. SBCC037), characterized by insertion of a 'T' in exon 2, causing a frame shift (Figure 1, 2<sup>nd</sup> arrow). These lines have the D haplotype in MWG699. This same haplotype was shared by wild barleys and two-rowed landraces from Morocco, all *Vrs1.b2*. Seven two-rowed Spanish landraces (ex. SBCC153), all with the D haplotype, had a new *Vrs1.b5* allele. Interestingly, this sequence contains the same 'T' found in the six-rowed *vrs1.a2*, but has an additional upstream deletion (Figure 1, 1<sup>st</sup> arrow) that results in the change of 15 amino acids and a potentially functional protein (Figure 2).

10 20 30 40 50 60 70 80 Vrs1.b OUH630 WHQLFGSSNVDTTFFAANGTAQGETSKQRARKKKKKSAKCC WHQLFGSSNVDTTFFAANGTAQGETSKQRARRRRRSARCC WHQLFGSSNVDTTFFAANGTAQGETSKQRARRRRRSARCC Vrs1.t Vrs1.b2 GDPKKRRLTDEQAEILELSFREDRK GDPKKRRLTDEQAEILELSFREDRK vrs1.a2 MDKHQLFGSSNVDTTFFAANGT/ GETSKQRARRRRRRSAEVRRRGW\* GDPKKRRLTDEQAEILELSFREDR Vrs1.b5 SBCC153 MDKHQLFGSSNVDTTFFAANGTAQGRPASRGRGAG GEMDO GGRURC vrs1.a4 MDKHQLFGSSNVDTTFFAANGTAQGETSKQRARRRRRSARCGGGDGDGGEMD GDPKKRRLTDEQAEILELSFREDRK vrs1.a1 MDKHOLFGSSNVDTTFFAANGTAOGETSKORARRRRRRSARC GDPKKRRLTDEOAEILELSFREDRK DGDGGEMD Vrs1.b3\_HarunaNijo MDKHOL HDSSNVDTTFFAANGTAOGDTSKORARRRRRSARCGGGDGDGGEMDG GDPKKRRLTDEOAEILELSFREDRK MDKHQLFDSSNVDTTFFAANGTAQGDTSKQRARRRRRRRRARCG DPKKRRLTDEOAETLELSEREDRI DPKKRRLTDEQAEILELS SNVDTTFFAANGTAQGETSKQRARRRRRRSARC

A molecular phylogeny positioned the new *Vrs1.b5* allele as sister to the six-rowed *vrs1.a2* allele (Fig. 3A), with both sharing a common ancestor with *Vrs1.b2*. The haplotype network on Fig. 3B places *Vrs1.b5* and *vrs1.a2* one step apart (indel in Figure 1).



**Figure 3.** A) Parsimony tree with bootstrap values. B) Haplotype network. Two-row barley accessions are displayed as thick circles.

PCoA separates two-rowed (left) from six-rowed (right) accessions. Further subdivision within those groups is apparent (Figure 4). A distinct group of two-row Spanish landraces stands out.



Vrs1.b_00H743	MUKHQLF655WVU	TIFFAANGIAG	or i skylivani		GDGDGGENDC	•••••	HD bein	SFREDR
	90	100	110	120	130	140	150	16
		• [ • • • • [ • • • •	1	1		1		
Vrs1.b_0UH630	LETARKVYLAAEL	GLDPKQVAVWF	QNRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLF	RLKERLGATEQ	EVRRLR
Vrs1.t	LETARKVYLAAEL	GLDPKQVAVWF	QNRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLE	RLKERLGATEQ	EVRRLR
Vrs1.b2	LETARKVYLAAEL	GLDPKQVAVWF	QNRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLF	RLKERLGATEQ	EVRRLR
vrs1.a2								
Vrs1.b5_SBCC153	LETARKVYLAAEL	GLDPKQVAVWF	QNRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLF	RLKERLGATEQ	EVRRLR
vrs1.a4	LETARKVYLAAEL	GLDPKQVAVWF	ONRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLE	RLKERLGATEQ	EVRRLR
vrs1.a1	LETARKVYLAAEL	GLDPKOVAVW	ONRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLE	LKERLGAT	RCGASG
Vrs1.b3 HarunaNijo	LETARKVYLAAEL	GLDPKOVAVW	ONRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLE	LKERLGATEO	EVRRLR
vrs1.a3	LETARKVYLAAEL	GLDPKQVAVW	QNRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLE	RLKERLGATEQ	EVRRLR
Vrs1.b OUH602	LETARKVYLAAEL	GLDPKOVAVW	ONRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLE	RLKERLGATEO	EVRRLR
Vrs1.b OUH743	LETARKVYLAAEL	GLOPKOVWW	ONRRARHKNK	TLEEEFARLK	HAHDAAILHK	CHLENELLE	LKERLGATEO	EVRRLR
	HD helix 2			HD helix 3			Leu zipper helix	2000
	170	180	190	200	210	220		
		.	1	1		1		
Vest b OUHERA	AAGSHGASVDGGH	AAGAVGVCGGS	PSSSFSTGT	OOOPGFSGAD	VLGRDDDLMM	CVPEWFLA		
AL 21'D 001020								
Vrs1.t	AAGSHGASVDGGH.	AAGAVGVCGG	PSSSFSTGT	OOOPGFSGAD	VLGRDDDLMM	<b>ICVPEWFLA</b>		
Vrs1.t Vrs1.b2	AAGSHGASVDGGH		PSSSFSTGTC PSSSFSTGTC	QQQPGFSGAD	VLGRDDDLMM VLGRDDDLMM	CVPEWFLA		
Vrs1.t Vrs1.b2 vrs1.a2	AAGSHGASVDGGH AAGSHGASVDGGH	AAGAVGVCGQ AAGAVGVCGGS	PSSSFSTGT	QQQPGFSGAD		ICVPEWFLA		
Vrs1.t Vrs1.b2 vrs1.a2 Vrs1.b5_SBCC153	AAGSHGASVDGGH AAGSHGASVDGGH AAGSHGASVDGGH	AAGAVGVCG@ AAGAVGVCGGS	PSSSFSTGTC PSSSFSTGTC	QQQPGFSGAD QQQPGFSGAD	VLGRDDDLMM VLGRDDDLMM VLGRDDDLMM	ICVPEWFLA ICVPEWFLA		
Vrs1.t Vrs1.t Vrs1.b2 Vrs1.a2 Vrs1.b5_SBCC153 Vrs1.a4	AAGSHGASVDGGH AAGSHGASVDGGH AAGSHGASVDGGH AAGSHGASVDGGH	AAGAVGVCG@ AAGAVGVCGGS AAGAVGVCGGS AAGAVGVCGGS	PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC	QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD	VLGRDDDLMA VLGRDDDLMA VLGRDDDLMA VLGRDDDLMA	ICVPEWFLA ICVPEWFLA ICVPEWFLA		
Vrs1.tb_001050 Vrs1.b2 vrs1.a2 Vrs1.b5_SBCC153 vrs1.a4 vrs1.a1	AAGSHGASVDGGH, AAGSHGASVDGGH, AAGSHGASVDGGH, AAGSHGASVDGGH, DLGATGHLWMADT	AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG PLAPLACAAG	PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC RARPSREP	QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD SSSRVSAGOT	VLGRDDDLMA VLGRDDDLMA VLGRDDDLMA VLGRDDDLMA CWGGTMT+-	ICVPEWFLA ICVPEWFLA ICVPEWFLA		
Vrs1.b_000650 Vrs1.b2 Vrs1.b2 Vrs1.a2 Vrs1.b5_SBCC153 Vrs1.a4 Vrs1.b3 HarunaNijo	AAGSHGASVDGGH AAGSHGASVDGGH AAGSHGASVDGGH AAGSHGASVDGGH DLGATGHLWMADT AAGSHGASVDGGH	AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG PLAPLACAAG AAGAVGVCGG	PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC RARPSRREP PSSSFSTGTC	QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD SSSRVSAGQT	VLGRDDDLMA VLGRDDDLMA VLGRDDDLMA VLGRDDDLMA CWGGTMT <del>X</del>	ICVPEWFLA ICVPEWFLA ICVPEWFLA ICVPEWFLA		
Vrs1.t Vrs1.b2 Vrs1.b2 Vrs1.b5_SBCC153 Vrs1.a4 Vrs1.a1 Vrs1.b3_HarunaNijo Vrs1.a3	AAGSHGASVDGGH, AAGSHGASVDGGH, AAGSHGASVDGGH, AAGSHGASVDGGH, (DLGATGHLWMADT) AAGSHGASVDGGH, AAGSHGASVDGGH,	AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG PLAPLACAAGA AAGAVGVCGG AAGAVGVCGG	PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC RARPSRREPA PSSSFSTGTC	QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD SSSRVSAGQT QQQPGFSGAD QQQPGFSGAD	VLGRDDDLMM VLGRDDDLMM VLGRDDDLMM VLGRDDDLMM CWGGTMT <del>X</del> VLGRDDDLMM VLGRDDDLMM	ICVPEWFLA ICVPEWFLA ICVPEWFLA ICVPEWFLA ICVPEWFLA		
Vrs1.t Vrs1.b2 Vrs1.b2 Vrs1.b5_SBCC153 Vrs1.a4 Vrs1.a1 Vrs1.b3_HarunaNijo Vrs1.b3_HarunaNijo Vrs1.b0UH662	AAGSHGASVDGGH, AAGSHGASVDGGH, AAGSHGASVDGGH, AAGSHGASVDGGH, QLGATGHLWMADT AAGSHGASVDGGH, AAGSHGASVDGGH, AAGTHGASVDGGH,	AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG AAGAVGVCGG PLAPLACAAGA AAGAVGVCGG AAGAVGVCGG	PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC PSSSFSTGTC RARPSRREPA PSSSFSTGTC PSSSFSTGTC	QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD SSSRVSAGQT QQQPGFSGAD QQQPGFSGAD QQQPGFSGAD	VLGRDDDLMM VLGRDDDLMM VLGRDDDLMM VLGRDDDLMM CWGGTMT <del>x</del> VLGRDDDLMM VLGRDDDLMM	ICVPEWFLA ICVPEWFLA ICVPEWFLA ICVPEWFLA ICVPEWFLA		

**Figure 2.** Multiple alignment of protein sequences of different *Vrs1* alleles. Boxed sequences highlight differences among alleles and asterisks show premature stop codons. Solid circles are protein-DNA interface residues.

### CONCLUSIONS

- A new *Vrs1.b5* allele was identified in two-row Spanish landraces.
- Phylogenies suggest that loss-of-function allele *vrs1.a2* is derived from *Vrs1.b2*. Some time later, a deletion in *vrs1.a2* gave rise to *Vrs1.b5*, which restored the ORF and reverted the two-rowed phenotype.
- PCoA separates the landraces in subpopulations also differing in the *Vrs1* allele (*.b3/.b3+.b5/.b2+.a1+.a2/.a1/.a2*). Similar results were obtained with microsatellites (Yahiaoui et al. 2008).

#### REFERENCES

Casas et al. 2005 Genome 48:41-45 Komatsuda et al. 2007 PNAS 104: 1424-1429 Tanno et al. 2002 Theor Appl Genet 104:54-60 Yahiaoui et al 2008 Theor Appl Genet 116: 271-282 ACKNOWLEDGEMENTS: work funded by the Spanish Ministry of Science and Innovation (Projects AGL2013-48756R, RFP2012-00015-00-00).