

RESEARCH

Open Access



Integrating natural variation through GWAS – genetics of drought and flood tolerance in grass pea reveal independent yet interconnected mechanisms

Matilde Sanches^{1,2,3*} , Marnik Vuylsteke⁴ , Carmen Santos¹ , Amna Mhamdi^{2,3} , Susana Araújo^{5,6} , Frank Van Breusegem^{2,3}  and Maria Carlota Vaz Patto¹ 

Abstract

Grass pea (*Lathyrus sativus* L.) is a grain legume of increasing importance in the Mediterranean region due to its outstanding tolerance to abiotic stresses such as salinity, heat, drought, and flooding, outperforming many other legume species. Despite established natural phenotypic variation in response to water-related stresses, the genetic basis of this resilience remains poorly understood, hindering precision breeding for single and combined stress tolerance. A genome-wide association study was conducted here to investigate the genetic architecture of water stress responses in grass pea. Previously, phenotypic data, including gas exchange, chlorophyll *a* fluorescence, photosynthetic pigments, leaf water status, and biomass partitioning traits, were assessed under well-watered, mild drought, and partial submergence conditions across 194 representative grass pea accessions worldwide. The data were associated with 5,651 single nucleotide polymorphisms (SNPs) using linear mixed models under a restricted maximum likelihood framework, incorporating population structure and the newly assembled L0007 genome. A total of 130 unique SNPs associated with at least one trait-treatment combination or with trait variation between stress and control conditions, providing a valuable resource for precision breeding of multi-stress tolerance in grass pea. The loci associated with drought and waterlogging were largely non-overlapping, suggesting distinct genetic bases for the two stress tolerances. However, some common mechanisms, such as redox regulation and carbohydrate metabolism, emerged among the identified candidate genes, highlighting some interconnectedness of biological pathways involved in grass pea responses to water stress.

Key message

Genotypic-phenotypic data associations suggest that the documented tolerance to water deficit and to waterlogging in grass pea is achieved through independent molecular mechanisms that might nevertheless overlap.

Keywords *Lathyrus sativus*, GWAS, Water deficit, Waterlogging, Genotype–phenotype association

*Correspondence:
Matilde Sanches
matsanches@itqb.unl.pt

Full list of author information is available at the end of the article



Introduction

Lathyrus sativus L., commonly known as grass pea, is an annual, primarily self-pollinating grain legume recognized for its outstanding tolerance to adverse conditions, such as water stress (defined here as water deficit and waterlogging), salinity, and temperature extremes [5, 19, 55, 98]. Not only is grass pea an insurance crop in some of the world's most vulnerable regions, it also holds huge potential as a genetic resource for breeding climate-resilient varieties [30].

Grass pea's tolerance to drought is partially explained by its morphology, including narrow leaves, winged stem margins, and a deep, extensive root system [48], as well as with a high photosynthetic capacity and efficient water use [15, 35, 89]. Additionally, biochemical changes can also occur, such as enhanced performance of antioxidant activity, reactive oxygen species (ROS) scavenging, and effective osmotic adjustment [37, 67, 95]. However, knowledge on the molecular basis of drought resilience in grass pea remains limited, with existing studies focusing on broad changes in proteomic [20, 75] and transcriptomic profiles [37], and the identification of a few drought-responsive microRNAs [14].

Short-term waterlogging (14 days with the water table 10 mm below the soil surface) had no significant effect on shoot and root dry biomass in grass pea [55]. This resilience has been attributed to increased antioxidant activity as well as the formation of aerenchyma and lateral roots under partial submergence [110]. Under combined stress conditions, such as heat with drought [4] or salinity with drought [95], grass pea accessions exhibited significant differences in yield-related and physiological traits, including osmotic adjustment and β -N-oxalyl-L- α , β -diamino propionic acid (β -ODAP) accumulation.

In addition to the aforementioned variation among individual accessions, evidence suggests that grass pea exhibits distinct responses to water availability based on geographical origin and seed type—the main clustering parameters in this species [68]. For example, seed germination in waterlogged soils varies by geographical origin, with accessions from Bangladesh and Ethiopia being the most well adapted [105]. Under Turkish rainfed field conditions, subtropical accessions consistently outperformed East European accessions across multiple yield parameters, regardless of the water availability [39].

We recently conducted an extensive phenotyping study on a worldwide collection of 194 grass pea accessions, which were subjected to three distinct water regimes: well-watered (WW), water deficit (WD), and waterlogging (WL). Several physiological and morphological traits were quantified to characterize the diverse responses to contrasting water stresses [82]. Under WD, increased root growth and stomatal closure were often

observed, contributing to improved water use efficiency. In contrast, enhanced shoot growth and carotenoid accumulation appeared to be the key responses to the WL treatment. Interestingly, the physiological and relative performance ranking of accessions varied across water treatments, indicating a strong genotype-treatment interaction [82]. We also demonstrated variability across ecotypes. For instance, light and large seeds (typically the Mediterranean ecotype) produced enlarged plants with increased total dry biomass overall, whereas dark and small-seeded accessions (mainly of African and Asian origin) seemed to be physiologically less affected by contrasting water stresses, exhibiting significantly augmented relative water content, stomatal conductance, and water use efficiency under WD, as well as improved stability or even amplified pigment content under WL [82].

Nevertheless, the molecular and genetic architecture of water stress resilience in grass pea remains largely unexplored, resulting in limited precision breeding for multi-stress tolerance. The availability of cost-effective, high-throughput platforms for single-nucleotide polymorphism (SNP) genotyping are important in prompting genetic research in underused crops such as grass pea [38]. Thus far, only a few studies have investigated the genetics of interesting traits in this species. They identified candidate genes for disease resistance [58, 59, 80, 84] and discovered a key gene in the β -ODAP biosynthesis pathway [25], but none directly assessed the genetic basis of abiotic stress tolerance. Given the complexity of these traits, influenced by multiple genes and with strong environmental interactions, genome-wide association studies (GWAS) offer a powerful and adequate approach to clarify the genetic architecture underlying such complexity [42, 64].

The identification of quantitative trait loci (QTLs) for water stress tolerance in other legumes has predominantly been achieved through GWAS, although breeding for complex traits in legumes has historically lagged behind that of cereals [78]. With regard to drought tolerance, GWAS has facilitated the characterisation of marker-trait associations and candidate genes in soybean (*Glycine max*) [2, 40, 77, 79], common bean (*Phaseolus vulgaris*) [34, 50], faba bean (*Vicia faba*) [32], and cowpea (*Vigna unguiculata*) [83, 106]. Waterlogging tolerance has also been investigated through GWAS in soybean [88], common bean [90] and mung bean (*Vigna radiata*) [45]. The wide array of gene-trait associations reported in these studies (with the exception of the mung bean study, in which a single SNP was found to be associated with five out of six traits), highlights the complex nature of drought and waterlogging stress tolerance in legumes.

Here, we utilized the natural variation of grass pea through a GWAS approach, by means of the newly

released genome assembly [100], to elucidate the molecular mechanisms and genetic basis of its water stress tolerance. We combined a genotypic dataset comprising 5,651 SNP markers mapped along the seven chromosomes of this species [81], with phenotyping data from 194 accessions that were evaluated under the WW, WD, and WL treatments, as previously described [82]. The marker-trait associations identified across the different water regimes provide the first steps towards developing molecular tools to support precision breeding for multi-stress tolerance, contributing to closing the gap between fundamental genetic research and its application in sustainable crop improvement [24, 41].

Materials and methods

Plant material

A set of 194 *Lathyrus sativus* L. accessions representative of the worldwide diversity of grass pea was used that had previously been genotyped using high-throughput SNP markers [81] and phenotyped for disease response, adaptation to the Mediterranean environment, and tolerance to different water stresses [30, 58, 59, 80, 82]. Within this collection, 92 and 102 accessions had light-coloured and dark-coloured seeds, respectively (visually scored). Additionally, 58 and 136 accessions were classified as large and small, based on 100-seed weight, using 18 g as the threshold [80]. Geographically, the set included 87 Mediterranean, 49 South Asian, 21 East European, 15 Sub-Saharan African, 13 North and Central Asian, 2 North American, and 2 South American accessions, and 5 of unknown origin (Supplementary Table S1).

The studied accessions were either provided by plant germplasm banks or donors, as in the case of cultivars, breeding lines or traditional landraces seeds. In this last case, the formal identification was done by Dr Susana Neves (PhD, botanist). In detail, two of the studied accessions, SITNICA and STUDENICA, are registered Serbian grass pea cultivars (obtained and provided by IFVCNS: Institute of Field and Vegetable Crops, Serbia); the accession LISA is a breeding line (obtained and provided by Dr Diego Rubiales, IAS-CSIC, Spain), as well as the accessions W6-39,226/Raipur and LS87124, provided by Prof. Fernand Lambein (IPBO, Belgium). Accessions named PTLS#### and SNVP-64 are traditional grass pea landraces provided by Portuguese farmers, which names can be found in Supplementary Table S1. Regarding these traditional landraces, vouchers have not been deposited in any publicly available herbarium, their seeds were voluntarily provided by the farmers who own and cultivate them with the support of Ana Moral, Diego Rubiales, Leticia Gonçalves, Nuno Almeida, Susana Neves, and Carlota Vaz Patto. All the other studied accessions were obtained from three different plant germplasm banks (CRF-INIA: Plant Genetic Resources Centre—National

Center of CSIC (Spain), ICARDA: International Center for Agricultural Research in the Dry Areas (Lebanon) and USDA: United States Department of Agriculture (USA)). All these germplasm bank accessions, as well as accession W6-39,226/Raipur are indexed at Genesys online platform (<https://www.genesys-pgr.org>). There, detailed information relative to its origin, identification and available vouchers can be retrieved using the accession IDs indicated in Supplementary Table S1.

Phenotypic data

The phenotypic dataset used in the present association analysis has been previously described [82]. Briefly, 20 different physiological traits including pigment content, gas exchange rates, chlorophyll *a* fluorescence, and dry biomass-related traits (listed in Supplementary Table S2.1), were measured in 194 grass pea accessions grown under one of three water treatments for 21 days. The treatments were well-watered (WW), which maintained approximately 95% field capacity (FC); water deficit (WD), which consisted of regular watering for 16 days followed by 5 days without irrigation, reaching approximately 55% FC; and waterlogging (WL), involving well-watered conditions for 7 days, followed by 14 days of partial submergence (approximately 1 cm above soil level). For additional methodological details, we refer to Sanches et al. [82] (Supplementary Table S2).

After filtering outliers, best linear unbiased estimates (BLUES) were calculated for each accession, water treatment, and trait through a restricted maximum likelihood procedure [82]. These BLUES datasets, one per trait, constituted the first phenotypic input data used for the present association study. The second phenotypic input data were the calculated deltas (Δ) for each trait and accession, specifically $\Delta_{WD} = BLUE_{WD} - BLUE_{WW}$ and $\Delta_{WL} = BLUE_{WL} - BLUE_{WW}$. Finally, the eigenvalues of each accession-treatment combination in the first five principal components (PCs) of the multivariate analysis [82], integrating all the 20 measured traits, were used as the third set of phenotypic input data.

Genotypic data

DNA had previously been isolated from one representative individual of each grass pea accession in the studied collection at the PlantX Lab and been sent to the relevant providers for genotyping using DArTseq™ and genotyping-by-sequencing (GBS) analysis [81]. A genotypic dataset consisting of 5,651 high-quality SNPs was retrieved from this previous analysis. Importantly, genotypic variation was guaranteed in the dataset, with all included markers exhibiting a minor allele frequency greater than 5%.

Of the 5,651 genotyped high-quality SNPs, 4,871 had a unique location on the seven grass pea chromosomes

[100], determined by alignment with the BLASTn tool (e-value < 1e-5) with the OmicsBox v2.0 software. The remaining 780 SNPs were considered unmapped, either because no BLAST hits passed the e-value threshold, or because multiple hits occurred of identical alignment probability to different genomic regions (based on bit score, percentage of similarity, and e-values).

Association-mapping analysis

Genome-wide association studies (GWAS) were conducted with Genstat® software [102] in the mixed model framework to detect SNPs associated with the variation of the measured traits in the grass pea collection. Markers and accessions were fitted as fixed and random terms respectively using restricted maximum likelihood [56]. A model accounting for familial relatedness (K) was employed using a kinship matrix K [Phenotype = SNP + Accession + Error], computed using the Dice similarity coefficient (DICE method) in Genstat® with structured accession random effects [56]. The effect of the minor-frequency SNP allele was estimated in relation to the most frequent reference allele.

Marker–trait associations were filtered using a false discovery rate (FDR) procedure based on the Benjamini–Hochberg method, adapted as described by Brzyski et al. [18]. We controlled the FDR at $Q = 0.20$, using the number of linkage disequilibrium (LD) blocks ($k = 5,433$) as the effective number of independent tests rather than the total number of markers. Such a threshold was chosen to allow the detection of a larger set of marker–trait associations compared with more stringent FDR approaches, which was considered preferable given the exploratory nature of this study, and particularly the limited genomic resources available for orphan crops such as grass pea, even at the cost of following up a limited number of false-positive associations.

Favourable allele frequency analysis

Previously, differences in response to the water treatments depending on grass pea ecotype clustering parameters (seed colour, seed size, and geographical origin of the accessions) have been reported [82]. To deepen our understanding of the differences among ecotypes, we searched for differences between the frequency of the favourable allele (i.e. maximizing the trait value) of each associated marker in light- vs. dark-seeded accessions, in large- vs. light-seeded accessions, and among the seven geographical origins in the collection.

Favourable allele frequencies were calculated for each seed morphology category (dark or light, small or large) and geographical origin, by dividing the number of accessions with the favourable allele by the total number of accessions in the same category. Differences between favourable allele frequencies depending on seed colour

and size were tested with penalized generalized linear mixed models for binomial data. Because accessions share ancestry and population structure, allele frequencies are correlated within seed morphology and origin groups. Population structure therefore needs to be accounted for in post-GWAS analyses. However, for binomial allele-frequency data, kinship matrices cannot be incorporated as straightforwardly as in Gaussian GWAS LMMs, since non-Gaussian GLMMs lack closed-form likelihoods and become computationally slow and difficult to diagnose. We therefore accounted for population structure in the post-GWAS GLMMs by including the first 15 principal components from a previous eigen analysis of a subset of markers as a low-rank approximation of kinship. Seed colour and seed size were therefore set as fixed factors and population structure data as random factors. A significance level of 0.05 was applied with Bonferroni correction for multiple testing within each set of same trait- and treatment-associated markers.

The genetic variance contributed by a QTL (V_{QTL}) to the total variance of each phenotypic trait (V_g) was calculated as $\frac{2 \times Freq \times (1 - Freq) \times Effect^2}{s^2}$ [103], where Freq is the frequency of the favourable QTL-associated allele in the population, Effect is the estimated effect of the same allele over the trait, and s^2 is the trait variance (the variance of the BLUEs of all accessions in all treatments; Supplementary Table S2). All statistics, visual representations, and follow-up analysis were performed with R software [73].

Candidate gene identification

Following GWAS, a gene was considered a candidate contributing to the phenotypic trait if it contained an associated SNP with that trait. Putative candidate genes were identified with the MapMan and Mercator4 v2.0 web tools [86] to locate the sequences of associated SNPs in relation to the grass pea genome. Swiss-Prot was the preferred source for annotation, although the InterPro-UniProt database was consulted as well.

Results

Marker-trait associations

A total of 114 significant marker-trait associations were detected across all GWAS using BLUEs from each trait-treatment combination: 15 under well-watered (WW), 50 under water deficit (WD), and 49 under waterlogging (WL) conditions (Table 1, Fig. 1a-c). Some SNPs were associated with multiple trait-treatment combinations, resulting in a total of 104 SNPs being significantly associated with at least one combination, of which 94 were uniquely mapped to one of the seven chromosomes of the reference grass pea genome [100]. Circular Manhattan plots (Fig. 1) illustrate the distribution of SNPs along

Table 1 Single nucleotide polymorphism (SNP, Marker) significant associations (adj.p-BH: adjusted p-values < 0.01 with an adaptation of Benjamini–Hochberg method, $\alpha = 0.2$, $k = 3007$) across 115 independent GWAS, their position within grass pea chromosomes (chr), allelic variants (Ref.Allele/Var.Allele), frequency and effect of the allelic variant, and the proportion of genotypic variance explained by each SNP-trait association (V_{QTL}/V_g) using 194 grass pea accessions subject to well-watered (WW), water deficit (WD) and waterlogging (WL) treatments

GWAS phenotypic input: BLUEs from all the accessions in each trait and treatment combination										
Marker	Trait	Treat	adj.p-BH ^a	Grass Pea genome location (chr ^b position (Mpb))		Ref.Allele ^c	Var.Allele ^d	Frequency ^e	Effect ^f	V_{QTL}/V_g (%) ^g
S0363	A ₃₄₇	WW	3.89E-05	1	22.92	G	A	0.216	0.658	0.06
	\sqrt{WUE}_{869}	WD	6.22E-04						0.206	0.05
S0294	A ₈₆₉	WW	1.94E-04	7	431.67	T	A	0.093	-1.118	0.04
S0644	A ₈₆₉	WW	1.56E-04	2	547.76	T	A	0.110	1.066	0.04
S1100	A ₈₆₉	WW	7.78E-05	4	608.98	T	C	0.136	1.016	0.04
S1556	A ₈₆₉	WW	1.17E-04	5	645.58	C	T	0.098	1.079	0.04
S2404	A ₈₆₉	WW	3.89E-05	6	537.36	C	T	0.068	1.643	0.06
	E ₈₆₉	WW	3.89E-05						0.671	0.07
	\sqrt{gS}_{347}	WW	3.89E-05						0.060	0.06
	TDB	WD	3.89E-05						-0.064	0.05
S2616	A ₈₆₉	WD	3.89E-05	6	653.14	T	A	0.079	-1.420	0.07
S3107	E ₈₆₉	WL	7.78E-05	5	7.08	C	T	0.143	-0.710	0.06
S2636	\sqrt{WUE}_{347}	WL	3.89E-05	6	669.65	A	G	0.119	0.120	0.06
S2672	\sqrt{WUE}_{347}	WL	7.78E-05	6	682.29	C	T	0.128	0.108	0.05
S4328	\sqrt{WUE}_{347}	WL	1.17E-04	3	679.93	G	A	0.340	0.078	0.06
S3074	\sqrt{WUE}_{869}	WW	3.89E-05	2	447.99	A	G	0.157	-0.118	0.08
S3774	\sqrt{WUE}_{869}	WW	4.28E-04	6	24.51	G	A	0.110	-0.088	0.03
S0258	\sqrt{WUE}_{869}	WD	5.45E-04	1	149.60	T	G	0.108	0.255	0.04
S0303	\sqrt{WUE}_{869}	WD	8.95E-04	1	85.90	G	A	0.051	0.365	0.04
S0511	\sqrt{WUE}_{869}	WD	9.72E-04	2	459.29	G	A	0.053	0.335	0.04
S0787	\sqrt{WUE}_{869}	WD	7.00E-04	3	664.13	G	A	0.069	0.283	0.03
S1051	\sqrt{WUE}_{869}	WD	3.89E-05	1	871.77	A	C	0.096	0.368	0.08
S1389	\sqrt{WUE}_{869}	WD	5.83E-04	7	142.65	G	C	0.061	0.349	0.05
S1429	\sqrt{WUE}_{869}	WD	6.61E-04	4	17.24	C	T	0.092	0.330	0.06
S1513	\sqrt{WUE}_{869}	WD	7.78E-05	Unk	23.00	C	G	0.102	0.326	0.06
S1668	\sqrt{WUE}_{869}	WD	1.01E-03	5	538.18	G	A	0.069	0.264	0.03
S1839	\sqrt{WUE}_{869}	WD	2.72E-04	5	410.28	G	A	0.089	0.315	0.05
S1840	\sqrt{WUE}_{869}	WD	1.17E-04	5	410.28	G	C	0.083	0.349	0.06
S1930	\sqrt{WUE}_{869}	WD	1.94E-04	5	48.53	G	A	0.052	0.413	0.05
S2133	\sqrt{WUE}_{869}	WD	8.17E-04	6	9.43	C	T	0.063	0.326	0.04
S2269	\sqrt{WUE}_{869}	WD	2.33E-04	7	22.70	T	G	0.071	0.339	0.05
S2316	\sqrt{WUE}_{869}	WD	4.28E-04	6	449.83	G	A	0.109	0.271	0.05
S2318	\sqrt{WUE}_{869}	WD	4.67E-04	6	449.83	A	C	0.108	0.271	0.05
S2319	\sqrt{WUE}_{869}	WD	4.67E-04	6	449.83	G	T	0.108	0.271	0.05
S2740	\sqrt{WUE}_{869}	WD	3.11E-04	7	684.22	C	T	0.067	0.336	0.05
S2742	\sqrt{WUE}_{869}	WD	1.56E-04	7	684.22	C	T	0.084	0.344	0.06
S3020	\sqrt{WUE}_{869}	WD	1.09E-03	7	537.70	C	G	0.075	0.273	0.03
S3240	\sqrt{WUE}_{869}	WD	7.39E-04	1	508.25	A	G	0.093	0.247	0.03
S3395	\sqrt{WUE}_{869}	WD	9.33E-04	4	391.86	T	C	0.080	0.272	0.04
S3657	\sqrt{WUE}_{869}	WD	8.56E-04	4	644.24	T	A	0.073	0.284	0.04
S3852	\sqrt{WUE}_{869}	WD	7.78E-04	1	760.05	C	A	0.190	0.180	0.03
S5038	\sqrt{WUE}_{869}	WD	1.05E-03	Unk	198.50	C	T	0.105	0.227	0.03
S5235	\sqrt{WUE}_{869}	WD	3.50E-04	4	681.16	T	C	0.061	0.364	0.05
S5478	\sqrt{WUE}_{869}	WD	3.89E-04	6	190.05	T	G	0.158	0.242	0.05
S1323	RWC	WW	3.89E-05	4	97.69	T	C	0.444	0.013	0.07
	TDB	WD	7.78E-05	4	159.16	T	G	0.110	-0.035	0.06
	TDB	WL	1.56E-04						0.046	0.04
								0.048	0.04	

Table 1 (continued)

GWAS phenotypic input: BLUEs from all the accessions in each trait and treatment combination										
S5474	RWC	WD	3.89E-05	5	94.45	C	A	0.152	-0.036	0.08
S0310	RWC	WL	9.72E-04	1	101.32	C	T	0.061	0.017	0.04
S0339	RWC	WL	1.01E-03	1	58.40	T	C	0.063	0.018	0.04
S0354	RWC	WL	1.17E-04	1	35.50	C	T	0.073	0.018	0.05
S0546	RWC	WL	8.56E-04	7	307.92	A	C	0.186	-0.018	0.12
S0603	RWC	WL	5.45E-04	2	4.57	G	A	0.139	-0.013	0.05
S0932	RWC	WL	3.11E-04	3	478.79	A	G	0.073	0.020	0.06
S1222	RWC	WL	1.56E-04	4	493.01	C	T	0.119	0.013	0.05
S1452	RWC	WL	1.24E-03	4	12.81	T	C	0.287	-0.010	0.05
S1477	RWC	WL	8.95E-04	2	505.94	T	C	0.087	-0.018	0.06
S1708	RWC	WL	8.17E-04	5	509.01	C	T	0.135	-0.013	0.05
S2026	RWC	WL	2.33E-04	3	20.73	C	T	0.152	-0.017	0.09
S2271	RWC	WL	3.89E-04	6	518.08	C	T	0.093	-0.016	0.05
S2507	RWC	WL	9.33E-04	Unk	229.00	A	G	0.148	-0.017	0.09
S2565	RWC	WL	7.00E-04	Unk	230.00	C	T	0.276	0.011	0.06
S2566	RWC	WL	2.72E-04	7	659.91	A	G	0.386	-0.015	0.14
S2681	RWC	WL	3.89E-05	3	707.38	T	A	0.099	-0.017	0.06
S2718	RWC	WL	6.61E-04	7	697.94	T	C	0.102	-0.013	0.04
S2719	RWC	WL	5.83E-04	7	697.94	A	G	0.101	-0.013	0.04
S2872	RWC	WL	6.22E-04	7	612.47	G	T	0.167	-0.011	0.04
S2874	RWC	WL	3.50E-04	7	612.47	G	A	0.166	-0.012	0.05
S2983	RWC	WL	4.28E-04	2	28.92	G	T	0.061	-0.017	0.04
S2996	RWC	WL	1.05E-03	5	454.11	C	A	0.153	0.012	0.05
S3386	RWC	WL	1.09E-03	1	709.44	A	G	0.089	-0.014	0.04
S3670	RWC	WL	1.21E-03	1	243.07	C	T	0.137	0.011	0.04
S3984	RWC	WL	1.13E-03	6	572.88	A	T	0.199	-0.012	0.06
S4424	RWC	WL	4.67E-04	7	111.95	G	A	0.410	0.014	0.12
S4696	RWC	WL	5.06E-04	Unk	163.50	A	G	0.206	-0.017	0.11
S4766	RWC	WL	7.78E-04	Unk	170.00	A	G	0.193	0.010	0.04
S4883	RWC	WL	1.94E-04	4	665.72	T	C	0.156	-0.016	0.09
S4903	RWC	WL	7.78E-05	5	343.93	A	G	0.127	0.015	0.06
S5567	RWC	WL	7.39E-04	7	152.76	G	A	0.257	-0.022	0.2341
S5574	RWC	WL	1.17E-03	7	152.76	G	A	0.296	-0.019	0.1933
S0870	Fv/Fo	WW	3.89E-05	3	570.74	C	G	0.122	-0.147	0.06
	(Fv/Fm) ²	WW	3.89E-05						-0.009	0.06
S1252	Fv/Fo	WD	3.89E-05	4	441.37	G	C	0.273	0.129	0.06
	(Fv/Fm) ²	WD	3.89E-05						0.008	0.06
S2661	C _{ab}	WW	3.89E-05	6	680.06	A	T	0.077	3.595	0.06
S2102	C _{ab}	WD	7.78E-05	1	877.17	G	A	0.201	2.801	0.06
	√C _{cx}	WD	3.89E-05						0.109	0.07
S4194	C _{ab}	WD	2.72E-04	6	434.21	T	C	0.303	1.971	0.04
S0225	√C _{ab} C _{cx}	WW	7.78E-05	6	404.15	C	T	0.104	-0.065	0.05
S0837	SPAD	WD	3.89E-05	3	611.22	T	A	0.117	-1.828	0.06
S2857	SPAD	WD	7.78E-05	7	619.24	A	T	0.098	1.741	0.04
S1286	√R/S	WD	1.56E-04	4	377.21	C	T	0.119	-0.066	0.05
S3553	√R/S	WD	3.89E-05	1	898.67	G	A	0.077	0.085	0.05
S5269	√R/S	WD	1.17E-04	1	307.23	G	A	0.050	-0.112	0.06
S5270	√R/S	WD	7.78E-05	1	307.23	C	T	0.050	-0.113	0.06
S0154	√R/S	WL	1.94E-04	1	627.92	C	T	0.052	0.068	0.04
S2977	√R/S	WL	1.17E-04	7	500.86	A	G	0.082	0.060	0.05
S3623	√R/S	WL	7.78E-05	Unk	266.00	A	T	0.057	0.070	0.04
S0018	TDB	WD	1.17E-04	1	772.04	G	A	0.072	0.059	0.04
S0020	TDB	WD	7.78E-05	1	772.04	A	G	0.072	0.059	0.04

Table 1 (continued)

GWAS phenotypic input: BLUEs from all the accessions in each trait and treatment combination										
S0021	TDB	WD	1.17E-04	1	772.04	C	T	0.072	0.059	0.04
S1291	TDB	WD	3.50E-04	1	15.13	C	T	0.158	0.034	0.03
S3871	TDB	WD	3.11E-04	4	325.57	C	A	0.137	0.038	0.03
	TDB	WL	2.33E-04						0.043	0.04
S4026	TDB	WD	2.33E-04	Unk	290.00	A	T	0.169	0.043	0.05
S0558	TDB	WL	3.89E-04	1	306.46	C	T	0.064	0.061	0.04
S0809	TDB	WL	4.28E-04	4	10.35	G	A	0.066	0.054	0.03
S1463	TDB	WL	3.50E-04	5	694.02	G	C	0.080	0.052	0.03
S2242	TDB	WL	5.45E-04	6	35.53	T	C	0.063	0.053	0.03
S3619	TDB	WL	2.72E-04	1	37.29	T	C	0.092	0.048	0.03
S4108	TDB	WL	4.67E-04	Unk	294.50	C	A	0.197	0.033	0.03
S4764	TDB	WL	3.11E-04	Unk	349.50	C	G	0.190	0.034	0.03
S5238	TDB	WL	5.06E-04	6	377.96	C	A	0.083	-0.049	0.03
GWAS phenotypic input: delta values obtained from subtracting BLUE _{stress} —BLUE _{WVW} in each trait										
Marker	Trait	Delta	adj.p-BH ^a	Grass Pea genome location (chr ^b position (Mpb))	Ref.Allele ^c	Var.Allele ^d	Frequency ^e	Effect ^f	V _{QTL} /V _G (%) ^g	
S2299	A ₈₆₉	ΔWL	3.89E-05	6	467.04	T	A	0.076	-1.838	0.05
S1312	√E347	ΔWL	3.89E-05	4	646.15	C	T	0.099	-0.258	0.07
	√gs347	ΔWL	3.89E-05						-0.078	0.05
S1318	√E347	ΔWL	7.78E-05	4	646.15	T	G	0.097	-0.257	0.07
	√gs347	ΔWL	7.78E-05						-0.078	0.05
S4163	√E347	ΔWL	1.17E-04	Unk	114.00	C	A	0.236	-0.156	0.05
	E869	ΔWL	7.78E-05						-0.598	0.05
S2693	E ₈₆₉	ΔWL	3.89E-05	7	714.31	T	C	0.107	0.858	0.06
S3023	E ₈₆₉	ΔWL	1.17E-04	7	533.59	T	C	0.461	0.576	0.07
S0258	√WUE ₈₆₉	ΔWD	3.50E-04	1	149.60	T	G	0.108	0.259	0.04
S0511	√WUE ₈₆₉	ΔWD	5.83E-04	2	459.29	G	A	0.053	0.352	0.04
S1051	√WUE ₈₆₉	ΔWD	1.17E-04	1	871.77	A	C	0.096	0.326	0.06
S1513	√WUE ₈₆₉	ΔWD	1.56E-04	Unk	23.00	C	G	0.102	0.281	0.04
S1839	√WUE ₈₆₉	ΔWD	7.78E-05	5	410.28	G	A	0.089	0.359	0.06
S1840	√WUE ₈₆₉	ΔWD	3.89E-05	5	410.28	G	C	0.083	0.396	0.07
S1930	√WUE ₈₆₉	ΔWD	2.72E-04	5	48.53	G	A	0.052	0.384	0.05
S2269	√WUE ₈₆₉	ΔWD	4.28E-04	7	22.70	T	G	0.071	0.331	0.05
S2316	√WUE ₈₆₉	ΔWD	3.89E-04	6	449.83	G	A	0.109	0.268	0.04
S2318	√WUE ₈₆₉	ΔWD	4.67E-04	6	449.83	A	C	0.108	0.268	0.04
S2319	√WUE ₈₆₉	ΔWD	4.67E-04	6	449.83	G	T	0.108	0.268	0.04
S2369	√WUE ₈₆₉	ΔWD	3.11E-04	6	62.24	A	G	0.250	0.199	0.05
S2742	√WUE ₈₆₉	ΔWD	1.94E-04	7	684.22	C	T	0.084	0.330	0.05
S5235	√WUE ₈₆₉	ΔWD	5.45E-04	4	681.16	T	C	0.061	0.336	0.04
S5478	√WUE ₈₆₉	ΔWD	2.33E-04	6	190.05	T	G	0.158	0.249	0.05
S5360	RWC	ΔWD	1.94E-04	1	47.51	A	T	0.084	-0.033	0.05
S5474	RWC	ΔWD	3.89E-05	5	94.45	C	A	0.152	-0.030	0.07
S2878	RWC	ΔWL	3.89E-05	7	611.78	G	A	0.085	0.020	0.06
S1073	√C _{ab} C _{cx}	ΔWD	3.89E-05	4	640.00	C	T	0.462	-0.060	0.08
S0353	√√Cb	ΔWD	7.78E-05	1	35.07	T	C	0.117	0.101	0.08
S0821	√√Cb	ΔWD	1.56E-04	3	634.87	T	C	0.064	0.108	0.05
S0822	√√Cb	ΔWD	3.89E-05	3	634.87	A	G	0.070	0.116	0.06
S1616	√√Cb	ΔWD	1.17E-04	5	581.27	A	G	0.106	0.096	0.06
S5270	√R/S	ΔWD	7.78E-05	1	307.23	C	T	0.050	-0.131	0.06
S0853	TDB	ΔWD	3.89E-05	3	592.85	A	C	0.074	-0.062	0.06
S1462	TDB	ΔWD	1.17E-04	4	54.10	A	G	0.085	0.052	0.04
S5446	TDB	ΔWD	7.78E-05	7	315.61	C	T	0.097	-0.053	0.05
S0002	TDB	ΔWL	1.94E-04	1	772.72	G	A	0.087	0.057	0.04

Table 1 (continued)

GWAS phenotypic input: BLUEs from all the accessions in each trait and treatment combination										
S0179	TDB	Δ WL	2.72E-04	1	575.70	C	T	0.054	-0.071	0.04
S0180	TDB	Δ WL	3.89E-04	1	575.70	T	C	0.066	-0.061	0.04
S0676	TDB	Δ WL	2.33E-04	6	17.29	T	C	0.089	-0.054	0.04
S1361	TDB	Δ WL	3.11E-04	4	23.17	A	G	0.070	-0.057	0.04
S1872	TDB	Δ WL	7.78E-05	5	364.10	T	C	0.101	-0.051	0.04
S2166	TDB	Δ WL	3.50E-04	6	6.01	T	C	0.155	0.045	0.05
GWAS phenotypic input: PCA coordinates (first four PC's) from all the accessions and treatments										
Marker	PC	Treat	adj.p-BH ^a	Grass Pea genome location (chr ^b position (Mpb))	Ref.Allele ^c	Var.Allele ^d	Frequency ^e	Effect ^f	V _{QTL} /V _g (%) ^g	
S4328	PC-3	WL	3.89E-05	3	679.93	G	A	0.340	0.550	0.08
S1254	PC-4	WD	7.78E-05	4	439.15	G	A	0.090	-0.666	0.05
S1255	PC-4	WD	3.89E-05	4	439.15	T	G	0.083	-0.729	0.06

^aadj.p-BH, adjusted *p*-values < 0.01 (with an adapted Benjamini–Hochberg method, $\alpha=0.2$, $k=3007$) across 115 independent GWAS, ^bChr., chromosome in the grass pea genome, ^creference allele, ^dvariant allele, ^efrequency and ^feffect of the allelic variant, ^gthe proportion of genotypic variance explained by each SNP-trait association (V_{QTL}/V_g) using 194 grass pea accessions subject to treatments

the genome (plus unmapped SNPs), with significant associations highlighted in red. Each concentric circle represents a trait-treatment combination.

In a second approach, GWAS were performed using the phenotypic input of Δ WD and Δ WL values (calculated as BLUE_{stress} - BLUE_{WW}), accounting for the relative effect of water stress on each trait and accession, with more negative Δ values indicating greater trait sensitivity to stress [82]. This method identified 26 and 14 SNPs associated with Δ WD and Δ WL traits, respectively, including three SNPs simultaneously associated with two traits each (Fig. 1d-e). Of the total 40 SNPs associated with Δ traits, 16 overlapped with those identified in the individual trait-treatment GWAS. Thirty-eight of the Δ -associated SNPs were successfully mapped to the grass pea reference genome (Table 1).

Finally, fifteen GWAS were performed across the three water treatments using eigenvalues from the five first principal components (PCs), which together represented 81.15% of the total phenotypic variance in the previous multivariate analysis [82]. Under WD, two SNPs were associated with PC-4, which captured variability in relative water content (RWC), total dry biomass (TDB), root-to-shoot ratio (R/S), and the greenness ratio (C_{ab}/C_{cx}) [82]. Under WL, one SNP was associated with PC-3, primarily driven by chlorophyll *a* fluorescence related traits [82] and was commonly associated with water use efficiency at 347 $\mu\text{mol}/\text{m}^2/\text{s}$ light intensity (WUE₃₄₇) in WL (Table 1, Fig. 1b-c). All three SNPs had known positions in the grass pea genome. No significant associations were detected for any PCs under WW conditions.

Of the 160 marker-trait associations identified (114 from trait-treatment, 43 from Δ traits, and three from PC-treatment GWAS), 103 showed a positive effect of the variant allele relative to the most frequent allele. In the trait-treatment GWAS, a positive effect of the variant allele indicated an increase in the measured trait value.

In the Δ trait GWAS, it corresponded to a less negative (or positive) Δ , reflecting reduced susceptibility to WD or WL. In the multivariate GWAS, positive effects represented increased values in traits with positive loadings within the corresponding PC. For example, PC-3 included positive loadings for carbon assimilation rates at 347 and 869 $\mu\text{mol}/\text{m}^2/\text{s}$ light intensities (A₃₄₇, A₈₆₉), stomatal conductance and water use efficiency at 869 $\mu\text{mol}/\text{m}^2/\text{s}$ light intensity (gs₈₆₉, WUE₈₆₉), WUE₃₄₇, maximum and actual quantum yields of photosystem II (Fv/Fm, Fv/Fo), performance index expressed on absorption basis (PI_{ABS}), C_{ab}/C_{cx}, R/S, and TDB, whereas PC-4 included positive loadings for transpiration rates at 347 and 869 $\mu\text{mol}/\text{m}^2/\text{s}$ light intensities (E₃₄₇, E₈₆₉), stomatal conductance and water use efficiency at 347 $\mu\text{mol}/\text{m}^2/\text{s}$ light intensity (gs₃₄₇), gs₈₆₉, Fv/Fm, Fv/Fo, PI_{ABS}, chlorophyll *b* and chlorophylls *a* and *b* contents (C_b, C_{ab}), C_{ab}/C_{cx}, and R/S [82].

As shown in Table 1, RWC had the highest number of associated markers (38 in total): one SNP under WW, two under WD, 32 under WL, two in Δ WD and one with Δ WL. One SNP (S5474) was common to both WD and Δ WD, but no SNPs were shared across treatments. It was followed by WUE₈₆₉ with 31 associations: two under WW, 28 under WD (14 of which also associated with Δ WD), and one exclusive to Δ WD. No associations were detected for WUE₈₆₉ under WL or Δ WL. TDB was associated with 28 markers: eight under WD, ten under WL (S3871 and S3489 shared), plus three in Δ WD and seven in Δ WL. No associations were found under WW.

In most cases where a single SNP was associated with multiple trait-treatment combinations, the traits were highly correlated [82], Supplementary Figure S1). For instance, S0870 and S1252 were associated with chlorophyll *a* fluorescence traits Fv/Fm and Fv/Fo under WW and WD, respectively; S2102 with pigment content-related traits C_{ab} and carotenoids + xanthophylls content

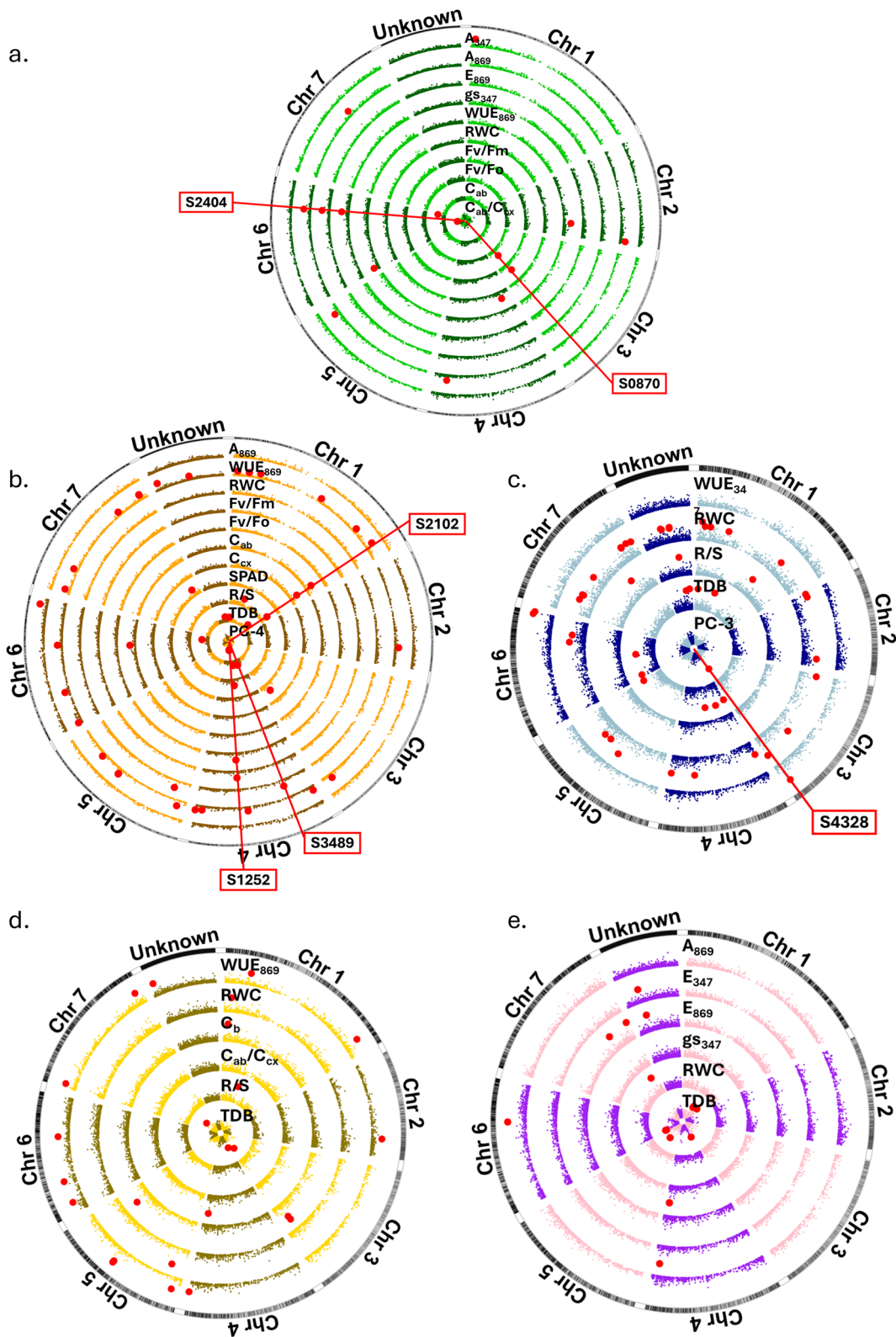


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Circular Manhattan plots representing the significant marker-trait associations found in the genome-wide association studies using 194 grass pea accessions, performed with i) BLUEs and ii) eigen values of each accession retrieved from Sanches et al. [82], integrating all the 20 measured traits, under **a**) WW (green), **b**) WD (orange) and **c**) WL (blue) conditions, and the significant marker-trait associations found in the genome-wide association studies performed with iii) delta values calculated for each phenotypic trait and accession in **d**) Δ WD (yellow) and **e**) Δ WL (purple). Each concentric circle depicts the Manhattan plot of a trait, as labelled (A_{347} : CO₂ assimilation rate at growth light intensity (347 $\mu\text{mol}/\text{m}^2/\text{s}$); A_{869} : CO₂ assimilation rate at Amax light intensity (869 $\mu\text{mol}/\text{m}^2/\text{s}$); E_{347} : transpiration rate at growth light intensity; E_{869} : transpiration rate at Amax light intensity; g_{s347} : stomatal conductance at growth light intensity; WUE_{347} : instantaneous water use efficiency at growth light intensity (A/E); WUE_{869} : instantaneous water use efficiency at Amax light intensity; RWC: leaf relative water content; Fv/Fm: maximum quantum yield of photosystem II; Fv/Fo: actual quantum yield of photosystem II; C_b : chlorophyll b content (spectrophotometrically measured); C_{ab} : chlorophylls a + b content (spectrophotometrically measured); C_{cx} : xanthophyll + carotene content (spectrophotometrically measured); C_{ab}/C_{cx} : greenness ratio based on spectrophotometrically measured pigment contents; SPAD: greenness index given by SPAD-502 instrument; R/S: root to shoot dry biomass ratio; TDB: total dry biomass; PC-3 and PC-4: eigen values of principal components 3 and 4 obtained for each accession x treatment combination in the multivariate analysis previously performed by Sanches et al. [82]. WW: well-watered conditions (21 days with constant soil moisture of about 95% field capacity); WD: water deficit treatment consisting of 16 days of regular watering followed by 5 days of no watering (corresponding to a final soil water content of about 55% FC); WL: waterlogging treatment consisting of well-watered conditions during the first 7 days, followed by 14 days of partial submergence (water reaching approximately 1 cm height of the stem) until measurement day; Δ WD: difference between the Best Linear Unbiased Estimators (BLUEs) of each accession under WD vs WW; Δ WL: difference between the BLUEs of each accession under WL vs WW). The y-axis represents the $-\log_{10}(\text{adj.}p)$ of 5651 SNP markers, and the x-axis exhibits their chromosome position on the grass pea reference genome [100]. Red dots highlight the significantly associated SNPs (FDR method, $\alpha=0.2$). Some SNPs with multiple trait associations are further highlighted with their names in labels in a), b) and c). In the outer band of each circular plot, the mapping of SNPs in each grass pea chromosome (+‘Unknown’ position) is depicted by black bands

(C_{cx}) under WD; and S2404 with A_{869} , E_{869} and g_{s347} under WW. S0363 was linked to A_{347} under WW and WUE_{869} under WD; S1312 and S1318 with E_{347} and g_{s347} in Δ WL; and S4163 with E_{347} and E_{869} (all gas-exchange related traits). Notably, S2404 was also associated with TDB under WD, making it the marker with the highest number of associations (four). Additionally, S4328 under WL was associated with both WUE_{347} and PC-3, the PC primarily capturing variation in chlorophyll *a* fluorescence-related traits.

Each SNP-trait association accounted for only a small proportion of the observed genotypic variance. The highest individual contributions were from S5567 and S5574, both associated with RWC under WL, explaining 23.41% and 19.33% of the variance, respectively. Regarding association strength, we have observed that there are five cases where the presence of the alternative allele (i.e. the least frequent in the studied collection) conduced to an estimated increase or decrease of more than 80% of the respective trait value: S2661, S2102 and S4294 positive effects over C_{ab} in WW (3.595) and WD (2.801 and 1.971) respectively; and S2299 and S0837 strong negative effects (−1.838 and −1.828) over A_{869} in Δ WL and SPAD in WD respectively. All marker-trait effects are reported in Table 1.

The 4,871 mapped SNPs were evenly distributed across the seven chromosomes of grass pea. The number of markers per chromosome ranged from 476 to 837 in chromosomes 3 and 7, respectively. Under WW conditions, chromosome 6 had the highest number of unique associated markers (four), whereas under WD conditions, chromosomes 1 and 4 had 14 and 10 markers, respectively and under WL conditions, chromosomes 1 and 7 eight and ten associated markers, respectively. These patterns remained consistent after normalising for SNP density across chromosomes. Associations from

GWAS using Δ trait values were more evenly distributed. Five and three Δ WD-associated markers were primarily found on chromosomes 1, 5 and 6, and on chromosomes 3, 4 and 7, respectively. The three Δ WL-associated markers were distributed across chromosomes 1, 4, 6, and 7.

We also analysed the chromosomal distribution of associated SNPs by trait group. For gas-exchange-related traits (i.e. A, E, *g_s* and WUE), most of the associated markers mapped on chromosomes 6 and 7 (12 and eight out of the 47 total associations, respectively). For RWC, 27% of all associated markers (10 out of 37) mapped to chromosome 7. Among the 11 markers associated with pigment-related traits (C_a , C_b , C_{ab} , C_{cx} , C_{ab}/C_{cx} and SPAD), the greatest number of associations were found on chromosomes 3 and 6 with three markers each. Interestingly, chromosome 3 exhibited the lowest overall SNP density in the dataset.

Differences in allelic frequencies of associated SNPs among grass pea ecotypes

To compare the frequency of allelic variants of the marker-traits associations across grass pea ecotypes, we defined favourable allele as one causing increase on the trait value. For the associations involving PC-3 and PC-4 or Δ traits, definition of the ‘favourable allele’ required further thought. For example, in Δ WD, the ‘favourable allele’ would be the one that contributes to a less negative (or more positive) Δ value by reducing the impact of the stress relative to WW condition. In contrast, for PC-3-based associations, the ‘favourable’ allele was considered the one that increases the value of the traits with strong positive loadings in the respective component. However, this same allele could be unfavourable for traits contributing negatively for the PC-3 component.

Figure 2 illustrates the traits of greatest physiological relevance under each treatment, focusing on those

showing the most pronounced phenotypic differences between seed type clusters [82]. For clarity, seed colour and size were combined to define four accession clusters: dark and small, dark and large, light and small, and light and large seeds.

A comparison of favourable allele frequencies across these grass pea clusters revealed that, in most cases (109 out of 160 associations), seed colour generally had a stronger influence than seed size. However, the effects of seed morphology traits were significant only in a very few cases. Three markers (S1073, S1323, and S2566 associated with C_{ab}/C_{cx} in Δ WD, RWC in WW, and RWC in WL, respectively) had significantly higher favourable allele frequency in light, large-seeded accessions (adjusted p -value < 0.05); two markers (S3023 and S4424, associated with E_{869} in Δ WL and RWC in WL, respectively) in dark, small-seeded accessions; and one (S4328, associated with WUE_{347} and PC-3 in WL) in large-seeded accessions. Full details are provided in Supplementary Table S3 and Fig. 2.

For markers S2616, S5474, and S2369, associated respectively with A_{869} , RWC and WUE_{869} traits in WD and Δ WD, the favourable allele was present at a higher frequency in dark-seeded accessions. In contrast, for several other markers (S1255, S1254, S1930, S0303, S1840, S2269, S3657, S1839, S2316, S2318, S2319, S1051, S1513, S3240, S3395, S5038, and S0363) associated with WUE traits (WUE_{347} , WUE_{869} , and the negative PC-4) under the same treatments, the frequency of the favourable allele was higher in light-seeded than in dark-seeded accessions of the same size (Fig. 2a). Regarding seed size, for marker S2616, which has a strong effect on A_{869} under WD conditions, large-seeded accessions had a higher frequency of the favourable allele than small-seeded ones of the same seed colour (Fig. 2a, leftmost panel). A similar pattern was observed for markers S1668, S1429, S3395, S5038, and S0363, all associated with WUE_{869} under WD conditions (with S0363 also linked to A347 under WW conditions). However, for markers S1255 and S1254 (associated with PC-4) and S2742 and S1513 (associated with WUE_{869} under WD and Δ WD), the favourable allele was more prevalent in small-seeded accessions than in large-seeded ones of the same colour.

For almost all markers associated with PC-3, R/S and gas-exchange-related traits under WL and Δ WL (except for S3107 and S3623), the frequency of favourable allele was higher in the dark-seeded accessions than in the light-seeded ones of the same size (Fig. 2b). Regarding seed size, the favourable allele of marker S4163 (linked to E_{347} and E_{869} in Δ WL) was more frequent in small-seeded accessions, whereas marker S4328 (associated with PC-3 in WL and primarily driven by chlorophyll a fluorescence traits) exhibited a significantly higher frequency of the favourable allele in large-seeded accessions.

Considering the four groups of accessions defined by seed colour and size, and noting the underrepresentation of 'light and small' (yellow triangles) and 'dark and large' (purple circles) accessions compared to 'light and large' (yellow circles) and 'dark and small' (purple triangles) accessions in the germplasm collection, we established the following. For 22 out of 29 markers associated with WUE in WD/ Δ WD (Fig. 2a, right-most panel), and for marker S4329 associated with PC-3 under WL (Fig. 2b, central panel), the light and large-seeded accessions exhibit higher favourable allele frequencies than dark and small-seeded ones. Similarly, across the 28 markers associated with TDB under WD/ Δ WD and WL/ Δ WL, the favourable allele was more often found at higher frequencies in light and large seeded accessions (Supplementary Table S3). In contrast, for markers associated with A_{869} and RWC under WD/ Δ WD (Fig. 2a, leftmost panels), and for most markers associated with gas-exchange traits under WL/ Δ WL (six out of seven; Fig. 2b, left panel), the frequencies of the favourable allele was higher in the dark and small-seeded accessions than in the light and large-seeded ones.

A comparison of allelic frequencies among the geographic origin clusters of grass pea, with a focus on the TDB trait, revealed that markers responsible for a small proportion of the total variance (i.e. those positioned to the left side of the x -axis, Fig. 3), showed the highest favourable allele frequencies in East European and Mediterranean accessions. In contrast, for markers accounting for a large proportion of the variance, South Asian accessions occasionally exhibited a higher favourable allele frequency than the Mediterranean ones. Accessions from Sub-Saharan Africa and North Asia generally displayed frequencies close to the global average, across all markers. Due to the limited number of accessions from South and North America groups (two each), no meaningful patterns could be inferred for these regions.

Identification of candidate genes and respective functions

To search for putative candidate genes related to WD and WL tolerance, we used the location of the SNPs significantly associated with at least one trait per treatment in the grass pea reference genome by means of Mercator and MapMan analysis for gene identification. Candidate genes were listed only if the SNP mapped within the gene. Of the total of 130 distinct associated markers (namely of the 160 associations, 25, one, and one were common between two, three, and four traits per treatment combinations, respectively), 119 were mapped to a known chromosome position in the grass pea reference genome, of which 41 markers were located in intergenic regions of chromosomes 1 to 7 and several mapped within the same coding DNA sequence, resulting in 66 unique candidate genes.

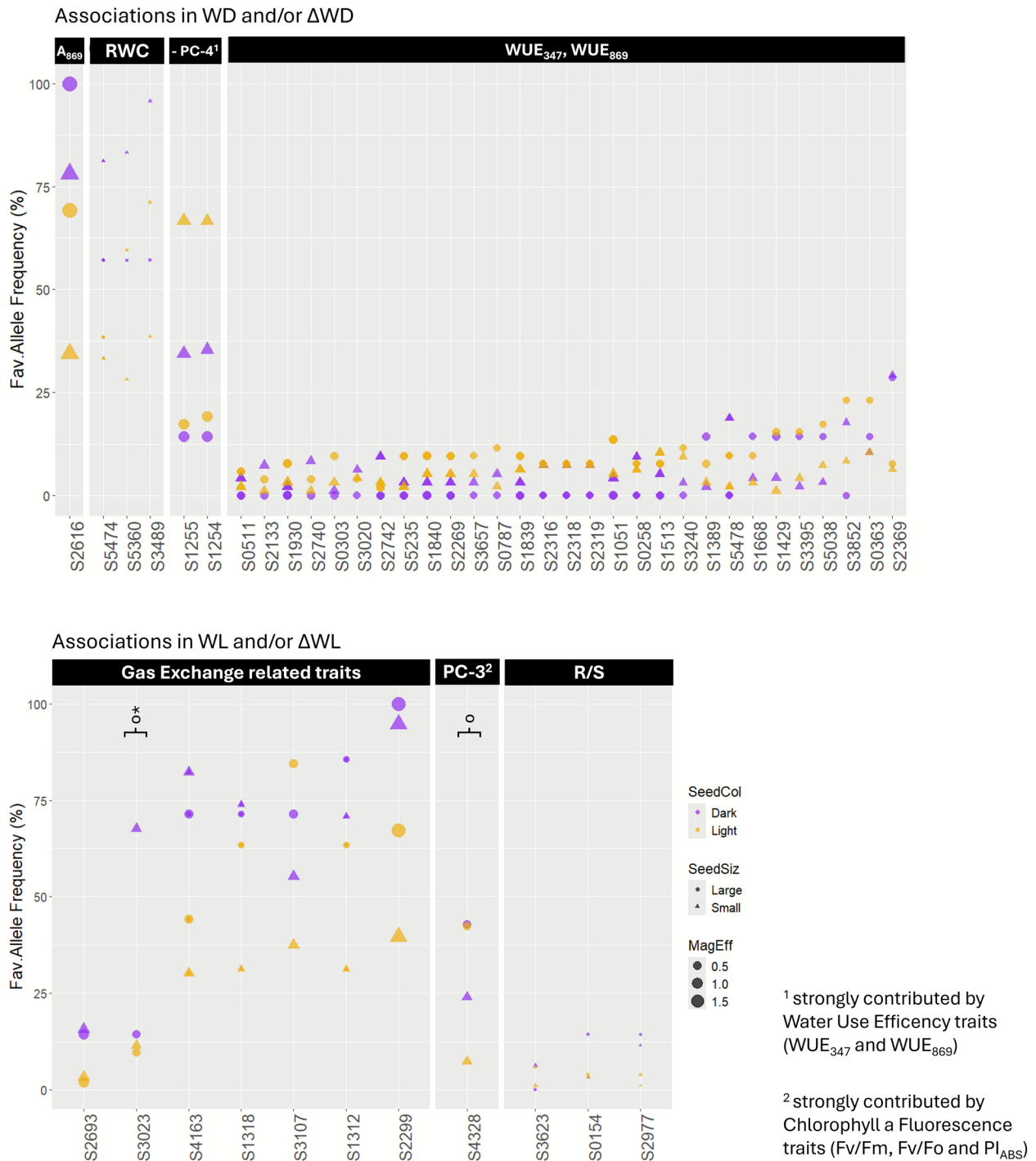


Fig. 2 (See legend on next page.)

We examined the functions, gene families, and interaction networks of each candidate gene using the Swiss-Prot and InterPro–UniProt databases and classified them into broad biological functional categories. Candidate genes identified through markers associated with distinct uncorrelated traits (Supplementary Figure S1) and/or stress treatments were prioritized for discussion.

Additional genes were also considered where their inclusion was supported by functional interconnectedness and/or recurring biological roles. For the complete list of candidate genes, see Supplementary Table S4. Several of the identified candidate genes seem to be related to carbohydrate metabolism. One example, the chloroplastic starch debranching enzyme isoamylase 1 (Uniprot:

(See figure on previous page.)

Fig. 2 Frequency of the favourable (Fav.) allele in different grass pea seed type clusters, for SNP markers associated with a) A_{869} , RWC, PC-4, WUE_{347} and WUE_{869} traits in either WD and/or Δ WD; and b) gas-exchange related traits (specifically gs_{347} , E_{347} , E_{869} and A_{869}), PC-3 and R/S traits (A_{869} : CO₂ assimilation rate at Amax light intensity (869 $\mu\text{mol}/\text{m}^2/\text{s}$); E_{347} : transpiration rate at growth light intensity; E_{869} : transpiration rate at Amax light intensity; gs_{347} : stomatal conductance at growth light intensity; WUE_{347} : instantaneous water use efficiency at growth light intensity (A/E); WUE_{869} : instantaneous water use efficiency at Amax light intensity; R/S: root to shoot dry biomass ratio; PC-3: eigen values of principal component 3 obtained for each accession x treatment combination in the multivariate analysis previously performed by Sanches et al. [82]. WD: water deficit treatment consisting of 16 days of regular watering followed by 5 days of no watering (corresponding to a final soil water content of about 55% FC); WL: waterlogging treatment consisting of well-watered conditions during the first 7 days, followed by 14 days of partial submergence (water reaching approximately 1 cm height of the stem) until measurement day; Δ WD: difference between the Best Linear Unbiased Estimators (BLUEs) of each accession under WD vs well-watered conditions (WW); Δ WL: difference between the BLUEs of each accession under WL vs WW, in either WL and/or Δ WL, as detected by GWAS using 194 grass pea accessions screened with 5651 SNP markers. Size of dots corresponds to the magnitude of the effect (MagEff) of the favourable allele over the trait, MagEff = 1 meaning that the presence of the favourable allele will increase the trait value with 1 × the average BLUE for that trait (i.e., it will double the average trait value). Shape and colour of dots correspond respectively to seed size and seed colour categories of accessions—purple triangles: dark and small seeded accessions; purple circles: dark and large seeded accessions; yellow triangles: light and small seeded accessions; yellow circles: light and large seeded accessions. Symbols ° and * stand for significant differences between favourable allele frequencies (employing Penalized Generalized Linear Mixed Models for binomial data, with $\alpha=0.05$ with Bonferroni correction) in seed size and seed colour clusters, respectively

O04196), is worth mentioning because three SNPs are located within it and are associated to contrasting water stress conditions, namely S2718 and S2719 (associated with RWC in WL) and S2742 (associated with WUE_{869} in both WD and Δ WD). Another starch-related protein was identified: granule-bound starch synthase 2, chloroplastic/amyloplastic (Uniprot: Q43093), associated with R/S in WL (S0154), whereas S1051, strongly associated with WUE_{869} in both WD and Δ WD, was located within the gene coding for the core subunit of NADH dehydrogenase [ubiquinone] flavoprotein 1, which catalyses the transfer of electrons from NADH through the mitochondrial respiratory chain using ubiquinone as an electron acceptor (UniProt: Q91YT0).

Two candidate genes encoding proteins with redox activity were identified under contrasting water stress conditions: Gfo/Idh/MocA-like oxidoreductase (InterPro: IPR004104), which is related to S3553 associated with R/S in WD, and an aldehyde dehydrogenase family 3 member H1, which is involved in detoxification of lipid peroxidation products (Uniprot: Q70DU8) and related to S1708 associated with RWC in WL. Other genes related to redox signalling mechanisms were found, such as ABC transporter C family member 2. This gene colocalized with S0339 associated with RWC in WL and codes for a pump that transports glutathione S-conjugates (GSH and GSSG, reduced and oxidised forms, respectively).

Candidate genes related to hormonal and transcriptional regulation were also identified. For example, two auxin-responsive genes, auxin response factor 3 (UniProt: O23661) and auxin-responsive protein IAA9 (UniProt: D3K0E6), were associated to RWC in WL (S0354) and R/S in WD (S1286), respectively. S2535 is significantly associated with WUE_{869} in WL and Δ WL and colocalizes with the gene for the PHD finger protein alfin 1 (UniProt: Q9FFF5) gene, a transcription factor involved in plant development and abiotic stress responses.

Finally, some markers linked to the waterlogging context were located in genes involved in cell wall plasticity.

For instance, S1312 and S1318, which associated with E_{347} and gs_{347} in Δ WL and S4328, linked to WUE_{347} and PC-3 in WL, localised within two probable pectinesterase/pectinesterase inhibitors (InterPro: IPR000070). Additionally, S0837 with a strong (negative) effect over SPAD trait in WD locates within a gene coding for NETWORKED 1B protein (NET1B, InterPro: F4JIF4), expressed mainly in root meristems with probable developmental- and cytoskeletal-structural related functions. Table 2 lists the abovementioned and other highlighted candidate genes, organised by involvement in similar mechanisms of the water stress response.

Discussion

Lathyrus sativus is a grain legume that has attracted interest as a promising climate-change proof crop due to its phenotypic and physiological superiority in more marginal environments [5, 30, 43, 55, 98]. However, the genetic and molecular basis of its abiotic stress tolerance has been less studied than that of other grain legumes (Dixit et al. 2016), resulting in limited understanding and hindering efforts in precision breeding for multi-stress resistance.

In this study, a geographically diverse panel of 194 grass pea accessions, comprising broad morphological and genetic variation, enabled a robust GWAS approach to dissect the genetic basis of the species water stress tolerance. In total, 160 significant associations were detected across 20 physiologically relevant traits and five principal components derived from individual water treatments or differentials of stress-control treatment. These results confirmed that tolerance to water deficit and waterlogging in grass pea is a quantitative trait, governed by numerous genes, each contributing modestly to phenotypic variation. Several SNPs were shared among traits, but rarely among treatments, hinting that drought and flood tolerances in grass pea rely on independent genetic architectures, although sometimes related to common molecular processes (namely redox regulation or

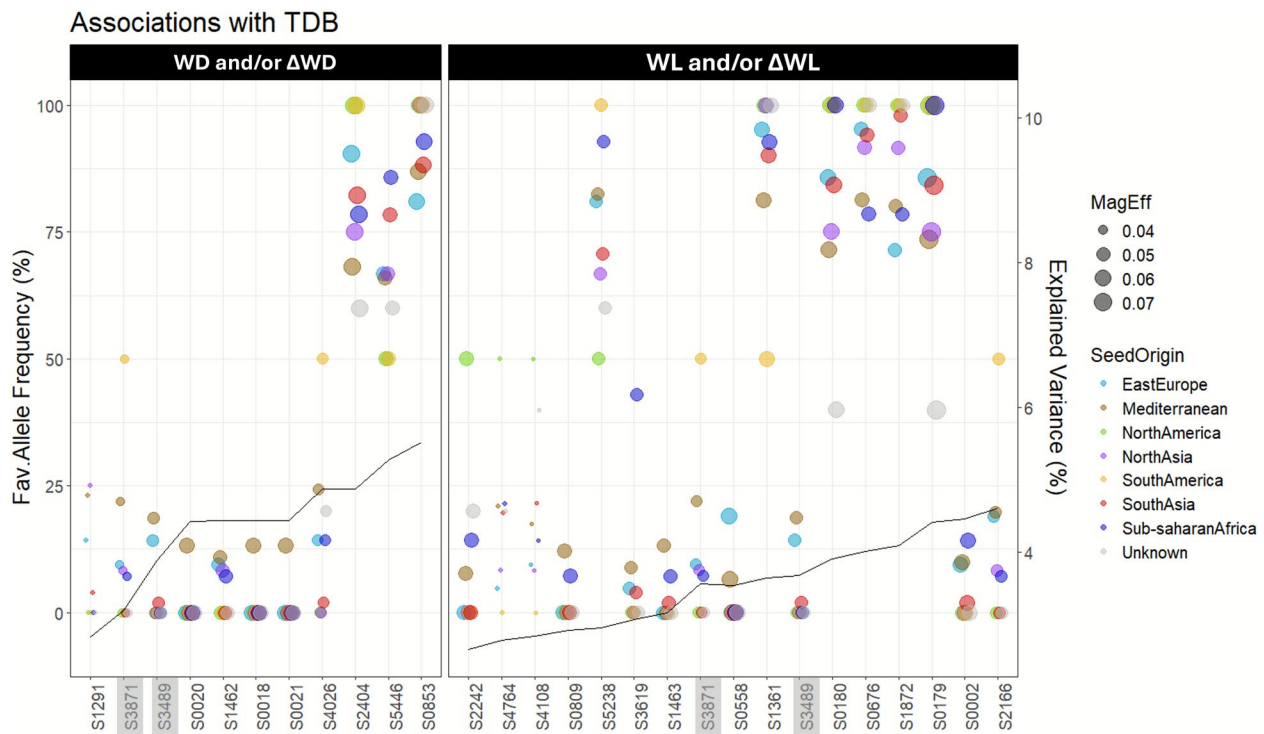


Fig. 3 Frequency of the favourable (Fav.) allele in different grass pea geographical origin clusters (eight colours corresponding to seven origins plus one ‘unknown’ category), for markers associated with total TDB in (left panel) WD and/or Δ WD and (right panel) WL and/or Δ WL (WD: water deficit treatment consisting of 16 days of regular watering followed by 5 days of no watering (corresponding to a final soil water content of about 55% FC); WL: waterlogging treatment consisting of well-watered conditions during the first 7 days, followed by 14 days of partial submergence (water reaching approximately 1 cm height of the stem) until measurement day; Δ WD: difference between the Best Linear Unbiased Estimators (BLUEs) of each accession under WD vs well-watered conditions (WW); Δ WL: difference between the BLUEs of each accession under WL vs WW), as detected by GWAS using 194 accessions screened with 5651 SNP markers. Size of dots corresponds to the magnitude of the effect (MagEff) of the favourable allele over the trait. Markers arranged in each panel in ascending order of the fraction of the trait variance explained by the QTL (a function of total frequency and effect of the favourable allele for each marker), with a black line marking the values (secondary y-axis)

carbohydrate metabolism), as evidenced by the 66 identified candidate genes. The most strongly associated SNPs with the largest phenotypic effects represent promising targets for the development of molecular tools to support precision breeding for multi-stress tolerance in grass pea and related species.

The mostly non-concomitant grass pea’s tolerance to WD and WL is achieved by various independent mechanisms

Among the 160 significant SNP-trait associations identified, 24 SNPs were associated in more than one GWAS, resulting in a final set of 130 unique associated SNPs (Table 1, Supplementary Table S1). These multiple associations involved either highly correlated traits or related stress conditions. For example, marker S2404 was simultaneously associated with A_{869} , E_{869} , and gs_{347} under WW conditions (all gas-exchange-related traits), whereas marker S5474 was associated with RWC in both WD and Δ WD.

Only two SNPs (S3489 and S3871) were associated with the same trait, TDB, in both WD and WL and S3489 also with RWC in WD. As TDB exhibited the highest

heritability and was minimally affected by water treatment [82], a great overlap of associated markers across conditions was expected. Another case (see below) involves a candidate gene containing three SNPs, associated with traits under contrasting water stresses. The scarcity of such double-stress associations highlights the different genetic architectures underlying the drought and flood tolerances in grass pea, in line with previous phenotypic data [82] that showed that only two accessions exhibited similar physiological response to both stresses (PI283546, double tolerant and PTL1001, double susceptible to WD and WL). Overall, our research supports the conclusion that drought and flood tolerance of grass pea do not necessarily coincide on the same germplasm. This interpretation is consistent with the literature, in which natural tolerance to both extremes is rare. Nonetheless, once the genetic basis of individual stress tolerance has been clarified, stress tolerances might be combined, as has been demonstrated in rice (*Oryza sativa*), where marker-assisted selection enabled the development of varieties tolerant to salinity, drought, and

Table 2 Short list of potentially interesting candidate genes associated with grass pea physiological responses under well-watered (WW), water-deficit (WD), water logging (WL) conditions or with the comparison between water stress and the well water condition (Δ WD and Δ WL co-located with markers associated to at least one trait.treatment found though GWAS using 194 accessions screened with 5651 SNP markers

Marker(s)	Trait(s)	Associated candidate gene and location (bp) ^a	Gene annotation (Unitprot) ^b	Function ^c	Protein involvement	Reason for prioritization
S2718, S2719, S2742	RWC in WL; WUE ₈₆₉ in WD and Δ WD	Lsat_g27235 chr7:697,921,143..697952797(bp)	Isoamylase 1, chloroplastic (O04196)	Starch-debranching enzyme involved in amylopectin biosynthesis in endosperm	carbohydrate metabolism	Multiple-trait association
S0154	R/S in WL	Lsat_g2817 chr1:627,912,928..627916929(bp)	Granule-bound starch synthase 2, chloroplastic/amyloplastic (Q43093)	Required for the synthesis of amylose in endosperm	carbohydrate metabolism	Functional interconnectedness and recurring biological roles
S2977	R/S in WL	Lsat_g29796 chr7:500,859,595..500873539(bp)	Fructose-bisphosphate aldolase class-I (P58314)	Essential glycolytic enzyme that catalyses reversible carbon-carbon bond formation	carbohydrate metabolism	Functional interconnectedness and recurring biological roles
S1051	WUE ₈₆₉ in WD and Δ WD	Lsat_g855 chr1:871,770,402..871774062(bp)	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial (Q91YT0)	Catalyzes electron transfer from NADH through the respiratory chain, using ubiquinone as an electron acceptor	carbohydrate metabolism	Functional interconnectedness and recurring biological roles
S0853	TDB in Δ WD	Lsat_g11108 chr3:592,849,973..592854243(bp)	Raffinose synthase or seed imbibition protein Sip1 (Q84VX0)	Involved in the synthesis of raffinose, a major soluble carbohydrate in seeds, roots and tubers	carbohydrate synthesis and osmoprotection	Functional interconnectedness and recurring biological roles
S3553	R/S in WD	Lsat_g517 chr1:898,674,015..898678373(bp)	Gfo/ldh/MocA-like oxidoreductase, C-terminal (InterPro: IPR004104)	Glucose-fructose oxidoreductase?	redox activity and signalling	Recurring biological roles across environments
S1708	RWC in WL	Lsat_g20489 chr5:509,011,641..509014203(bp)	Aldehyde dehydrogenase family 3 member H1 (Q70DU8)	Involved in oxidative stress tolerance by detoxifying reactive aldehydes derived from lipid peroxidation	redox activity and signalling	Recurring biological roles across environments
S0339	RWC in WL	Lsat_g5249 chr1:58,390,157..58407639(bp)	ABC transporter C family member 2 (NCBI GeneID: 125,603,599)	Pump for glutathione S-conjugates such as GSH and GSSG (glutathione oxidized and reduced forms)	redox activity and signalling	Recurring biological roles across environments
S0870	Fv/Fo and Fv/Fm in WW	Lsat_g11301 chr3:570,738,448..570745588(bp)	Calmodulin-interacting protein 111 (Q9LET7)	Enables ATP binding and hydrolysis, locates in cytoplasm and chloroplast	redox activity and signalling	Recurring biological roles across environments

Table 2 (continued)

Marker(s)	Trait(s)	Associated candidate gene and location (bp) ^a	Gene annotation (Unitprot) ^b	Function ^c	Protein involvement	Reason for prioritization
S2102	C _{ab} and C _{cx} in WD	Lsat_g782 chr1:877,166,910..877,173,278(bp)	ABC transporter Pleiotropic Drug Resistance protein 1 (Q949G3)	Roles on constitutive and jasmonate-driven pathogen defense, by means of antifungal compounds excretion	pathogen defense	Strong effect; multiple-trait association
S0303	WUE ₈₆₉ in WD	Lsat_g4983 chr1:85,895,317..85918705(bp)	YT521-B-like domain, 30-kDa cleavage and polyadenylation specificity factor	Probable processing endonuclease. Nucleus-localized, its RNA-binding activity is inhibited by calmodulin in a calcium-dependent fashion	transcriptional/pos-transcriptional regulation	Functional interconnectedness
S5235	WUE ₈₆₉ in WD and ΔWD	Lsat_g17909 chr4:681,162,821..681167002(bp)	PHD finger protein Alfin1 (Q9FFF5)	Transcriptional regulator of the salt-inducible PRP2 gene (among others). Plays a role in salinity tolerance in Medicago sativa	transcriptional/pos-transcriptional regulation	Functional interconnectedness
S1361	TDB in ΔWL	Lsat_g14223 chr4:23,170,358..23170802(bp)	FCS-Like Zinc finger 18 (P0D012)	May act as adapter to facilitate the interaction of SnRK1 complex with effector proteins	transcriptional/pos-transcriptional regulation	Functional interconnectedness
S1286	R/S in WD	Lsat_g15369 chr4:377,207,828..377209628(bp)	Auxin-responsive protein IAA9 (D3K0E6)	Represses early auxin response genes at low auxin concentrations through interaction with auxin response factors	hormonal and transcriptional regulation	Recurring biological roles across environments
S0354	RWC in WL	Lsat_g5491 chr1:35,499,016..35502501(bp)	Auxin response factor 3 (Q23661)	Could act as transcriptional activator or repressor. Formation of heterodimers with Aux/IAA proteins may alter their ability to modulate early auxin response genes expression	hormonal and transcriptional regulation	Recurring biological roles across environments
S1312, S1318	E ₃₄₇ and g ₃₄₇ in ΔWL	Lsat_g14828 chr4:111,351,305..111354460(bp)	Probable pectinesterase/pectinesterase inhibitor 47 (Q9FF77; InterPro: IPR000070)	Acts in the modification of cell walls via demethylesterification of cell wall pectin	cell structure and dynamics	Functional interconnectedness and recurring biological roles

Table 2 (continued)

Marker(s)	Trait(s)	Associated candidate gene and location (bp) ^a	Gene annotation (Unitprot) ^b	Function ^c	Protein involvement	Reason for prioritization
S4328	WUE ₃₄₇ and PC-3 in WL	Lsat_g10076 chr3:679,926,871..679929017(bp)	Probable pectinesterase/pectinesterase inhibitor 41 (Q8RXX7; InterPro: IPR000070)	Acts in the modification of cell walls via demethylesterification of cell wall pectin	cell structure and dynamics	Multiple-trait association; functional interconnectedness and recurring biological roles
S0837	SPAD in WD	Lsat_g10884 chr3:611,217,767..611223852(bp)	NETWORKED 1B (F4JIF4)	Expressed mainly in root meristems with probable developmental- and cytoskeletal-structural related functions	cell structure and dynamics	Strong effect (C _{ab}); multiple-trait association

^aFrom grass pea L007 reference genome [100]

^bFrom Mercator4 v2.0 (<https://www.plabipd.de/portal/mercator4>) or Swissprot (<https://www.uniprot.org/>)

^cFrom Interpro-UniProtKB (<https://www.uniprot.org/>)

submergence by pyramiding beneficial alleles from different inbred lines [63].

Although PI283546 had been classified as tolerant to both WD and WL [82], analysis of this accession's alleles at stress response-associated markers did not reveal the expected clear enrichment of favourable alleles, or a depletion of unfavourable ones (Supplementary Table S3). This finding may be due in part to the low frequency of alternative alleles at almost all the screened markers (the presence of numerous 'rare' alleles makes comparisons of favourable allele counts across accessions difficult). Additionally, the absence of obvious 'patterns' of favourable/unfavourable alleles suggest that grass pea tolerance to WD and WL may arise from various different mechanisms, operating independently or in parallel.

Most SNP-trait associations were identified under water stress conditions, with 50 in WD, 49 in WL, and only 15 in WW. Such complexity was anticipated because abiotic stress tolerance typically involves multiple pathways and intricate gene interactions (for a review, see [47] and specifically for grain legumes, [5]), and is generally more complex at the molecular and genetic levels than biotic stress resistance [94, 97]. For instance, we detected many more marker-trait associations than in previous GWAS studies conducted with the same grass pea collection, which focused on biotic stresses [58, 59, 80].

Consistent with the polygenic nature of abiotic stress tolerance, each of the 160 significant SNP-trait associations explained only a small proportion of the phenotypic variance, ranging from 2.67% (S2242 associated with TDB in WL) to 23.41% (S5567 associated with RWC in WL), with a mean of 5.42% (Table 1). Consequently, validating the detected associations is complicated and genomic regions with an increased association density,

such as chromosome 7, should be prioritized for QTL inspection [106, 107].

Identification of preferential genetic targets for assisted breeding, despite the genetic complexity of grass pea water stress tolerance

Stomatal closure is one of the earliest plant responses to water shortage [104], quickly impairing photosynthesis [12, 23]. The 29 SNP markers significantly associated with WUE₈₆₉ (calculated as the ratio of carbon assimilation and transpiration) in WD or ΔWD (Table 1) demonstrate the physiological complexity of the trait and the central role of gas exchange in regulating available water. Gas exchange and photosynthetic efficiency traits are widely used to assess drought responses and have been central of numerous QTL mapping studies aimed at identifying genetic determinants of stress tolerance [21, 26, 96, 106]. Marker S0363 is particularly interesting because it was associated both with WUE₈₆₉ under WD and A₃₄₇ under WW. This dual association suggests its potential utility in identifying accessions that exhibit drought tolerance and high photosynthetic performance under non-stress conditions concomitantly.

Root-related traits are also frequently evaluated in abiotic stress tolerance studies due to the crucial positive impact of a deep, dense root system in the crops' capacity to avoid drought under low-moisture conditions [36, 54, 77]. Similarly, leaf pigment contents and chlorophyll *a* fluorescence parameters are widely used to assess the physiological performance of other grain legumes under drought and combined stresses [31, 62, 83, 93]. Markers S5474, S1252, S2102, and S5270 were each associated with multiple traits under WD and/or ΔWD (namely RWC, chlorophyll *a* fluorescence, pigment content, and

R/S) and explained 6–8% of phenotypic variance. Given their simultaneous involvement in distinct traits, these markers represent promising targets for use in marker-assisted selection of drought-tolerant accessions.

Furthermore, S2102 was also among the markers with the strongest effect on the C_{ab} trait (effect size = 2.801; Table 1). This suggests that selecting accessions carrying the alternative allele at this locus, particularly when combined with the alternative allele at marker S4194 (which also shows a large effect of 1.971 on the same trait; Table 1), has the potential to result in more than a three-fold increase in chlorophyll content under mild drought conditions. Whether such an effect is desirable, however, depends on the breeder's objectives. An additional interesting feature of S2102 marker is that it locates within *Lsat_g782* gene, encoding the ABC transporter Pleiotropic Drug Resistance protein 1 (PDR1), known to play a major role in pathogen defence in *Nicotiana plumbaginifolia* and other plant species [91]. This makes the locus potentially attractive for breeding aimed at improving combined biotic and abiotic stress tolerance.

Under waterlogging conditions, pigment biosynthesis can be impaired, as observed in tomato (*Solanum lycopersicum*) [74] and mung bean [44], possibly because of deficient nutrient absorption [55]. Most crops also exhibit morphological changes in roots and shoots in response to waterlogging [49]. Consequently, traits such as pigment content, greenness index, shoot growth, and elongation rate are commonly employed in association or linkage studies targeting waterlogging tolerance [9, 45, 57, 88]. Markers S0154, S2977, and S3623 were all significantly associated with R/S under WL, each explaining between 3.8% and 4.6% of the phenotypic variance. Therefore, they represent potential candidates for further validation to differentiate grass pea accessions with distinct responses to WL. However, caution is required when interpreting the biological relevance of 'favourable' and 'unfavourable' alleles. A statistical association between the 'favourable' allele with an increased R/S does not necessarily indicate enhanced root depth or density; these patterns were actually rarely observed in the previous phenotyping [82]. Instead, an increased R/S ratio under WL may result from reduced shoot elongation, which is not inherently advantageous, or from root thickening, possibly due to aerenchyma formation [110]. Further anatomical studies of grass pea roots under WL conditions are needed to clarify these mechanisms.

The trait with the highest number of distinct SNP associations under WL was RWC, totalling 33. This high number of associations was somehow unexpected, given the relatively low phenotypic variability observed for this trait [82]. Nevertheless, the substantial genotypic variability suggests that different molecular mechanisms may be responsible for maintaining RWC under WL

conditions, i.e. accessions may achieve similar RWC values by activating different cellular responses driven by allelic variation at various loci.

A total of 28 associations were found for TDB: eight in WD, three in Δ WD, ten in WL, and seven in Δ WL. Interestingly, markers S3489 and S3871 were associated both to WD and WL, reflecting intrinsic differences in biomass productivity rather than stress-specific responses. Marker S2404 is particularly noteworthy: in addition to its association with TDB under WD, it is also linked with A_{869} (with a relatively strong positive effect of 1.643, Table 1), E_{869} and gs_{347} in WW, explaining between 4.9% and 6.6% of phenotypic variation of these traits, respectively. This dual association suggests that S2404 could be useful in selecting accessions with more efficient gas exchange and photosynthesis under optimal watering conditions.

Pigment content and greenness ratio are two commonly used phenotypic traits in association or linkage studies of waterlogging tolerance [9, 16, 45], but no significant associations with pigment-related traits were identified in WL or Δ WL. Nevertheless, PC-3, which integrates a linear combination of chlorophyll *a* fluorescence-related traits [82], was significantly associated with marker S4328, which was also associated with WUE_{347} . Despite these associations, the relatively low proportion of phenotypic variance explained (5.6% for WUE_{347} and 7.8% for PC-3) limits the usefulness of this marker for marker-assisted selection.

Genetic confirmation of previously reported differences in physiological responses among grass pea ecotypes under contrasting water stress conditions

We detected interesting ecotypic response differences to WW, WD, and WL treatments based on seed colour, size, and geographical origin. For instance, favourable alleles for WUE traits relevant to drought response were generally more prevalent in accessions with light-coloured large seeds across the Δ WD and WD marker-trait associations. This observation is in contrast with previous findings, in which dark-coloured, small-seeded accessions exhibited higher WUE_{869} under WD [82]. Strikingly, the highest frequency of favourable alleles was found in dark and/or dark-coloured small-seeded accessions for markers explaining the greatest phenotypic variance (e.g., S1389, S5478, S1668, S1429, S3395, S5038, S3852, S0363, and S2369), suggesting a more subtle allele-ecotype interaction.

For markers associated with A_{869} and RWC in WD and Δ WD, as well as for most markers associated with gas exchange-related traits in WL and Δ WL, dark and small-seeded accessions exhibited a higher favourable allele frequency than light and large-seeded accessions. This finding is consistent with previous phenotypic

observations, which showed that dark and small-seeded accessions had significantly higher RWC under WD than light and large-seeded ones, less negative Δ WD of RWC and gas exchange parameters, and positive Δ WL responses in A, E, and gs [82].

Regarding TDB, the typically light and large-seeded accessions of Mediterranean and East European origin were previously reported to show increased TDB across water treatments [82]. We also observed a tendency of higher favourable allele frequencies for TDB-associated markers in these origin groups, although most evident in markers explaining a smaller proportion of phenotypic variance. Conversely, for markers explaining a larger fraction of the total variance, the typically dark and small-seeded accessions of the South Asian origin group occasionally surpassed the Mediterranean one in favourable allele frequencies. Among the 28 TDB associations identified under WD, Δ WD, WL, and Δ WL, favourable alleles were more frequently found in light and large-seeded accessions, supporting early phenotypic findings [82].

Functional prediction of candidate genes indicates an array of common and specific pathways

Significant advancements have recently been made in the annotation of the grass pea genome. The published chromosome-scale reference genome [100] was used here to map candidate genes for the 119 selected SNP markers. Whereas most of these candidate genes encode proteins of unknown function, we identified a set of genes with possible involvement in responses to water stress (Table 2).

Among the most interesting and exemplary candidate genes, several were involved in carbohydrate metabolism, such as *Lsat_g27235*, associated to WUE_{869} in WD and Δ WD and to RWC in WL; and *Lsat_29796*, and *Lsat_g2817*, associated to R/S in WL. Additionally, *Lsat_g855*, associated to WUE_{869} in WD and Δ WD, encodes the core subunit of a NADH dehydrogenase in the mitochondrial electron transport chain and is therefore related to carbohydrate anabolism. These associations are consistent with previous findings in mung bean, where carbohydrate metabolism was linked to adventitious root development during transient waterlogging [45]. More generally, carbohydrate metabolism plays a critical role during water stress. During transient waterlogging, the breakdown of sugars helps to sustain energy production, compensating for the shift in aerobic-to-anaerobic metabolism that often occurs due to hypoxia [101]. Under water deficit, degradation of large polysaccharides is a key mechanism for plants in several aspects: the increased soluble sugar concentration in stomatal guard cells drives stomata closure, the soluble sugars in the mesophyll or in the root cells allow for osmotic regulation, helping water

retention; and broken-down sugars provide energy and carbon when photosynthesis is potentially limited by water stress [87, 92].

Furthermore, the accumulation of soluble sugars and proline play an important role in osmoprotection and has been positively associated with abiotic stress tolerance in other legume species, such as faba bean [1] and chickpea (*Cicer arietinum*) [6]. We identified the candidate gene, *Lsat_g11108*, which is associated to TDB in Δ WD and coding for a protein involved in the synthesis of raffinose, a major soluble carbohydrate in seeds, roots, and tubers. Mutant maize (*Zea mays*) plants lacking the protein raffinose synthase (RAFS) or seed imbibition protein Sip1 hyperaccumulate galactinol and are more sensitive to drought stress, whereas *ZmRAFS* overexpression in *Arabidopsis thaliana* improved drought tolerance via increased myo-inositol [51]. Myo-inositol is a sugar alcohol and product of galactinol hydrolysis catalysed by RAFS, which can play an effective role in signal transduction, cell turgor maintenance, and regulation of ROS [109], all key mechanisms during drought responses.

This leads us to the discussion of three other genes: *Lsat_g517*, associated with R/S in WD, and *Lsat_g20489* and *Lsat_g5249*, associated with RWC in WL. These genes are probably involved in redox mechanisms, which are crucial for plant responses to abiotic stresses. Under both drought and waterlogging, ROS accumulate due to stomatal closure, impaired gas exchange, and over-reduction of the photosynthetic electron transport chain (ETC). Potential ROS generation [33, 46, 65] also occurs at mitochondrial ETC complex III [69]. If the plant's antioxidant defences are unable to respond adequately, oxidative damage can occur, including lipid peroxidation, protein and DNA degradation, chlorophyll loss, membrane disruption, and cell death [28, 29, 66].

Increased antioxidant activity has been suggested to be a possible mechanism involved in grass pea's tolerance to drought [37, 95], waterlogging [110], and other general stresses [10, 76]. A recent GWAS in cowpea identified a gene encoding an oxygenase family protein (*Vigun04g161000*) associated with pod weight under normal and water stress conditions [83]. In the same study, three candidate genes were associated to drought response for being involved in photoprotection (avoidance of excessive ROS production in the chloroplast under conditions where the dark reaction of photosynthesis is impaired or restricted): *Vigun07g133400*, *Vigun07g132700*, and *Vigun07g258000*. Specifically, the first gene encodes a protein with organic solute transporter-related functions; the second encodes a germin-like protein (GLS2-1) with superoxide dismutase activity; and the third encodes a UDP-glycosyltransferase 84a1-related protein (UGT84A1), implicated in the glycosylation of various plant metabolites, such as

flavonoids and other antioxidants, thus playing a role in mitigating oxidative stress [17]. Together, these findings reinforce the pivotal role of redox homeostasis and antioxidant defence in conferring resilience to water stress in grass pea and other related legumes.

The oxidative burst triggered at the onset of a stress is not only damaging but might also play an important signalling role [13, 70]. One interesting mechanism is the so-called ROS-wave, in which a Ca^{2+} -dependent activation of membrane NADPH oxidases under low O_2 leads to transient H_2O_2 accumulation in adjacent cells and facilitates a rapid, 'overall' response [60, 61]. The identified *Lsat_g11301* gene, associated with chlorophyll *a* fluorescence traits in WW, encodes a calmodulin-interacting protein and may therefore be involved in this ROS- and Ca^{2+} -mediated signalling.

Another candidate gene involved in transcriptional or post-transcriptional regulation is *Lsat_g4983*, which is associated with WUE_{869} in WD and encodes a YT521-B-like domain from a probable processing endonuclease 30-kDa subunit, binding N6-methyladenosine-modified RNA. A similar YT521-B homology domain-containing protein was detected in association with drought in common bean, as well as a zinc finger protein encoded by the *Phvul.010G032700* gene that was implicated in ABA-mediated drought tolerance [50]. Likewise in alfalfa (*Medicago sativa*), the zinc finger protein gene *MtPHD6* [72] and six others in cowpea [106] were associated with drought-responsive QTLs. Zinc finger proteins, particularly those of the plant homeodomain (PHD) class, are widespread in eukaryotic genomes and frequently involved in abiotic stress responses [71]. We identified two such proteins: the PHD finger protein Alfin1 encoded by *Lsat_g17909*, was associated to WUE_{869} in WD and ΔWD and is known to regulate *MsPRP2* expression in alfalfa roots and to be involved in salt stress tolerance [11], hinting at a broader role in abiotic stress tolerance.

The second zinc finger protein found is FCS-Like Zinc finger 18, which is encoded by *Lsat_g14223* and associated to TDB in ΔWL . This protein is thought to function as an adapter, facilitating the interaction between the SnRK1 complex and effector proteins under heat stress [53]. *SnRK1* genes, particularly *SnRK1.1* and *SnRK1.2*, are central regulators of plant hypoxia responses [7, 52]. Consequently, a QTL associated with germination under flooding was found linked with the SnRK1.1-homologous gene *Phvul.008G039400* in common bean [90]. Flooding is also known to induce *SnRK1* expression in an ABA-dependent manner [7, 52], and the interaction between ABA and SnRK1 has been shown to regulate the miR156/SPL module, mediating adaptive responses to flooding in alfalfa [27]. In both WD and WL we characterised marker associations linked to candidate genes encoding

auxin-regulated transcription factors: genes *Lsat_g15369* (associated with R/S) and *Lsat_g5491* (associated with RWC), respectively. Auxin signalling is well known to regulate root growth and architecture, both of which are key traits in both drought and waterlogging tolerance in legumes (reviewed by Kathun et al. 2021 and [108], respectively).

Additionally, two candidate genes (*Lsat_g10076* and *Lsat_g14828*) encoding pectinesterase/pectinesterase-inhibitor proteins were associated with several WL tolerance-related traits, including WUE, PC-3, E, and gs. These associations imply a possible role for cell wall plasticity in the response to WL, potentially linked with the shoot elongation commonly observed in this grass pea collection [82], and the known partial submergence escape strategy [8]. Moreover, cell wall degradation is a critical process during ethylene-mediated aerenchyma formation, another flood-adaptive trait [85]. Two cellulose synthase and two pectinesterase inhibitor contigs have previously been reported to be downregulated in rust-susceptible grass pea plants upon fungal inoculation [3], highlighting a broader role in cell wall remodelling in grass pea responses to both abiotic and biotic stresses. At the marker loci within the *Lsat_g10076* and *Lsat_g14828* (the two pectinesterase/pectinesterase-inhibitor genes found in our study), S1312, S1318, and S4328, two grass pea rust-susceptible accessions (NC024582 and NC050432) carried all favourable alleles for E_{347} and gs_{347} under ΔWL , and WUE_{347} and PC-3 under WL. Interestingly, one of these accessions (NC024582) was also previously identified as susceptible under WD [82], suggesting possible links between biotic and abiotic stress susceptibility.

Another protein with a putative structural role was identified associated with SPAD in WD: NET1B, encoded by *Lsat_g10884*. This association had one of the strongest marker-trait effects in our study (S0837 effect = -1.828 , Table 1). NET1B belongs to a superfamily of actin-binding proteins and, in *Arabidopsis*, is primarily expressed in the root meristem and early elongation zones, functioning as a membrane-cytoskeleton adaptor. A *net1a net1b* double knockout in *Arabidopsis* shows a long-root phenotype, suggesting that NET1 proteins contribute to root growth control, likely via effects on plasmodesmatal transport, though the transported signals have not been identified [22]. The majority of the identified candidate genes either encode uncharacterised proteins or contain broadly defined annotation domains, limiting the ability to infer specific roles in responses to water stress. Importantly, novel genes and proteins potentially involved in drought and flood tolerance in grass pea need to be functionally characterised. Overall, the proposed candidate genes should be viewed as hypotheses that help guide future investigations by highlighting recurring functional

roles and interconnected pathways likely involved in water-stress responses in grass pea and possibly other grain legumes. Given the highly polygenic nature of abiotic stress tolerance and the large number of associated genomic regions identified, any attempt to pinpoint specific causal genes for drought or waterlogging tolerance in grass pea cannot be robustly supported at this stage.

Conclusions

This study provides the initial clarification of the genetic basis underlying contrasting water stress tolerance (in terms of drought and waterlogging) in grass pea, making use of the recently published reference genome [100]. A total of 160 significant marker-trait associations were identified across 115 independent GWAS analyses (one per trait-treatment combination), corresponding to 130 unique associated regions due to common SNPs across traits.

The largely non-overlapping loci associated with drought and waterlogging (WD/ Δ WD and WL/ Δ WL) tolerance, imply that these stress tolerances in grass pea are governed by a distinct and highly polygenic genetic basis. However, the identified candidate genes span a range of biological pathways, which, in some cases, suggest some shared or at least interconnected mechanisms, such as those interlinking with reactive oxygen species, carbohydrate metabolism, and auxin signalling. Nevertheless, some associated loci encode uncharacterised proteins, indicating that novel genes associated with drought and flood tolerance in grass pea remain to be discovered.

The SNP markers identified here are a valuable resource for developing molecular tools to support precision breeding for multi-stress tolerance in grass pea and related species. Furthermore, this study paves the way for the future functional validation and expression analysis of candidate genes, certainly possible thanks to the recent development of a genetic transformation protocol via hairy root in grass pea [99]. These advances bring us closer to unlocking the genetic mechanisms behind water stress resilience in this underutilised yet promising crop.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-026-08229-y>.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Supplementary Material 5.

Acknowledgements

The authors thank Prof. Dr. Fred van Eeuwijk (Biometris, Wageningen University & Research) for statistical and support, Stephane Rombouts (VIB-Ghent University) for the integration of grass pea genomic data into the ORCAE genome annotation resource, Martine De Cock (Ghent University) for help in preparing the manuscript, Susana Leitão (ITQB NOVA) for the help in retrieving relevant information about the studied plant material.

Authors' contributions

Conceptualization: M.S. and M.C.V.P.; Methodology: M.S. and M.C.V.P.; Formal analysis and investigation: M.S.; Visualization: M.S.; Writing-original draft preparation: M.S.; GWAS processing: M.V.; Mercator and Mapman analysis: C.S.; Writing—review and editing: M.S., C.S., A.M., S.A., F.V.B., and M.C.V.P.; Funding acquisition: M.C.V.P., Supervision: F.V.B. and M.C.V.P.

Funding

This research was funded by the Fundação para a Ciência e Tecnologia, Portugal through PhD grant to MS (PD/BD/150296/2019), GREEN-IT Bioresources for Sustainability R&D Unit base (UIDB/04551/2020 and UIDP/04551/2020), and LS4FUTURE Associated Laboratory (LA/P/0087/2020), and by the European Union through the DIVIN FOOD project (Horizon 2020 Research and Innovation program, Grant Agreement nº 101000383). The work was also supported by the Fernand Lambein Fund.

Data availability

Data will be made available on request.

Declarations

Ethics approval and consent to participate

The Portuguese traditional landraces seeds were voluntarily provided for this study by the farmers who own and cultivate them.

Competing interests

The authors declare no competing interests.

Author details

¹Instituto de Tecnologia Química e Biológica António Xavier, ITQB NOVA, Universidade Nova de Lisboa, Oeiras 2780-157, Portugal

²Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent 9052, Belgium

³Center of Plant Systems Biology, VIB, Ghent 9052, Belgium

⁴Gnomixx, Melle 9090, Belgium

⁵MORE – Laboratório Colaborativo Montanhas de Investigação, Bragança, Portugal

⁶CIMO, LA SusTEC, Instituto Politécnico de Bragança, Campus de Santa Apolónia, Bragança 5300-253, Portugal

Received: 18 August 2025 / Accepted: 20 January 2026

Published online: 05 February 2026

References

1. Abid G, Saidi MN, Ouertani RN, Muhovski Y, Jebara SH, Ghouili E, et al. Differential gene expression reveals candidate genes for osmotic stress response in faba bean (*Vicia faba* L.) involved in different molecular pathways. *Acta Physiol Plant*. 2021;43:40.
2. Aleem M, Razaq MK, Aleem M, Yan W, Sharif I, Siddiqui MH, et al. Genome-wide association study provides new insight into the underlying mechanism of drought tolerance during seed germination stage in soybean. *Sci Rep*. 2024;14:20765.
3. Almeida NF, Leitão ST, Krezdorn N, Rotter B, Winter P, Rubiales D, et al. Allelic diversity in the transcriptomes of contrasting rust-infected genotypes of *Lathyrus sativus*, a lasting resource for smart breeding. *BMC Plant Biol*. 2014;14:376.
4. Aloui K, Choukri H, El Haddad N, Gupta P, El Bouhmedi K, Emmrich PMF, et al. Impact of heat and drought stress on grasspea and its wild relatives. *Plants*. 2023;12:3501. <https://doi.org/10.3390/plants12193501>.

5. Araujo SS, Beebe S, Crespi M, Delbreil B, González EM, Gruber V, et al. Abiotic stress responses in legumes: strategies used to cope with environmental challenges. *Crit Rev Plant Sci*. 2015;34:237–80.
6. Awasthi R, Gaur P, Turner NC, Vadez V, Siddique KHM, Nayyar H. Effects of individual and combined heat and drought stress during seed filling on the oxidative metabolism and yield of chickpea (*Cicer arietinum*) genotypes differing in heat and drought tolerance. *Crop Pasture Sci*. 2017;68:823–41.
7. Baena-González E, Rolland F, Thevelein JM, Sheen J. A central integrator of transcription networks in plant stress and energy signalling. *Nature*. 2007;448:938–42.
8. Bailey-Serres J, Lee SC, Brinton E. Waterproofing crops: effective flooding survival strategies. *Plant Physiol*. 2012;160:1698–709. <https://doi.org/10.1104/pp.112.208173>.
9. Ballesteros DC, Mason RE, Addison CK, Andrea Acuña M, Nelly Arguello M, Subramanian N, et al. Tolerance of wheat to vegetative stage soil waterlogging is conditioned by both constitutive and adaptive QTL. *Euphytica*. 2015;201:329–43. <https://doi.org/10.1007/s10681-014-1184-3>.
10. Banerjee J, Das A, Parihar AK, Sharma R, Pramanik K, Barpete S. Genomic designing towards development of abiotic stress tolerant grass pea for food and nutritional security. In: Kole C (ed) *Genomic Designing for Abiotic Stress Resistant Pulse Crops*. Cham: Springer International Publishing; 2022. pp 345–381. https://doi.org/10.1007/978-3-030-91039-6_9.
11. Bastola DR, Pethe VV, Winicov I, Alfin1, a novel zinc-finger protein in alfalfa roots that binds to promoter elements in the salt-inducible *MsPRP2* gene. *Plant Mol Biol*. 1998;38:1123–35. <https://doi.org/10.1023/A:1006081926699>.
12. Basu S, Ramegowda V, Kumar A, Pereira A. Plant adaptation to drought stress [version 1; peer review: 3 approved]. *F1000Res*. 2016;5:1554. <https://doi.org/10.12688/f1000research.7678.1>.
13. Baxter-Burrell A, Yang Z, Springer PS, Bailey-Serres J. RopGAP4-dependent Rop GTPase rheostat control of *Arabidopsis* oxygen deprivation tolerance. *Science*. 2002;296:2026–8. <https://doi.org/10.1126/science.1071505>.
14. Bhat KV, Mondal TK, Gaikwad AB, Kole PR, Chandel G, Mohapatra T. Genome-wide identification of drought-responsive miRNAs in grass pea (*Lathyrus sativus* L.). *Plant Gene*. 2020;21:100210. <https://doi.org/10.1016/j.plgene.2019.100210>.
15. Biederbeck VO, Bouman OT. Water use by annual green manure legumes in dryland cropping systems. *Agron J*. 1994;86:543–9. <https://doi.org/10.2134/ agronj1994.00021962008600030016x>.
16. Borrego-Benjumea A, Carter A, Zhu M, Tucker JR, Zhou M, Badea A. Genome-wide association study of waterlogging tolerance in barley (*Hordeum vulgare* L.) under controlled field conditions. *Front Plant Sci*. 2021;12:711654. <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2021.711654>.
17. Bowles D, Lim E-K, Poppenberger B, Vaistij FE. Glycosyltransferases of lipophilic small molecules. *Annu Rev Plant Biol*. 2006;57:567–97.
18. Brzyski D, Peterson CB, Sobczyk P, Candès EJ, Bogdan M, Sabatti C. Controlling the rate of GWAS false discoveries. *Genetics*. 2017;205:61–75. <https://doi.org/10.1534/genetics.116.193987>.
19. Campbell CG. Grass pea, *Lathyrus sativus* L. Promoting the conservation and use of underutilized and neglected crops, Vol. 18. Rome: Bioversity International; 1997.
20. Chattopadhyay A, Subba P, Pandey A, Bhushan D, Kumar R, Datta A, et al. Analysis of the grasspea proteome and identification of stress-responsive proteins upon exposure to high salinity, low temperature, and abscisic acid treatment. *Phytochemistry*. 2011;72:1293–307. <https://doi.org/10.1016/j.phytochem.2011.01.024>.
21. Damerum A, Smith HK, Clarkson GJJ, Truco MJ, Michelmore RW, Taylor G. The genetic basis of water-use efficiency and yield in lettuce. *BMC Plant Biol*. 2021;21:237. <https://doi.org/10.1186/s12870-021-02987-7>.
22. Deeks MJ, Calcutt JR, Ingle EKS, Hawkins TJ, Chapman S, Richardson AC, Mentlak DA, Dixon MR, Cartwright F, Smertenko AP, Oparika K, Hussey PJ. A superfamily of actin-binding proteins at the actin-membrane nexus of higher plants. *Curr Biol*. 2012;22(17):1595–600. <https://doi.org/10.1016/j.cub.2012.06.041>. Dixit GP, Parihar AK, Bohra A, Singh NP. Achievements and prospects of grass pea (*Lathyrus sativus* L.) improvement for sustainable food production. *Crop J*. 2016;4:407–416. <https://doi.org/10.1016/j.cj.2016.06.008>.
23. Dubois M, Inzé D. Plant growth under suboptimal water conditions: early responses and methods to study them. *J Exp Bot*. 2020;71:1706–22. <https://doi.org/10.1093/jxb/eraa037>.
24. Duc G, Agrama H, Bao S, Berger J, Bourion V, De Ron AM, et al. Breeding annual grain legumes for sustainable agriculture: new methods to approach complex traits and target new cultivar ideotypes. *Crit Rev Plant Sci*. 2015;34:381–411. <https://doi.org/10.1080/07352689.2014.898469>.
25. Edwards A, Njaci I, Sarkar A, Jiang Z, Kaithakkottil GG, Moore C, et al. Genomics and biochemical analyses reveal a metabolon key to β -L-ODAP biosynthesis in *Lathyrus sativus*. *Nat Commun*. 2023;14:876. <https://doi.org/10.1038/s41467-023-36503-2>.
26. Ferguson JN, Fernandes SB, Monier B, Miller ND, Allen D, Dmitrieva A, et al. Machine learning-enabled phenotyping for GWAS and TWAS of WUE traits in 869 field-grown sorghum accessions. *Plant Physiol*. 2021;187:1481–500. <https://doi.org/10.1093/plphys/kiab346>.
27. Feyissa BA, Amyot L, Nasrollahi V, Papadopoulou Y, Kohalmi SE, Hannoufa A. Involvement of the miR156/SPL module in flooding response in *Medicago sativa*. *Sci Rep*. 2021;11:3243.
28. Foyer CH, Ruban AV, Noctor G. Viewing oxidative stress through the lens of oxidative signalling rather than damage. *Biochem J*. 2017;474:877–83. <https://doi.org/10.1042/BCJ20160814>.
29. Gill SS, Tuteja N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol Biochem*. 2010;48:909–30. <https://doi.org/10.1016/j.plaphy.2010.08.016>.
30. Gonçalves L, Rubiales D, Lourenço M, Vaz Patto MC. Exploring grass pea (*Lathyrus sativus* L.) genetic diversity in Mediterranean changing climate conditions. *Eur J Agron*. 2024;156:127142. <https://linkinghub.elsevier.com/retrieve/pii/S1161030124000637>.
31. Goufo P, Moutinho-Pereira JM, Jorge TF, Correia CM, Oliveira MR, Rosa EAS, et al. Cowpea (*Vigna unguiculata* L. Walp.) metabolomics: osmoprotection as a physiological strategy for drought stress resistance and improved yield. *Front Plant Sci*. 2017;8:586.
32. Gutiérrez N, Pégard M, Balko C, Torres AM. Genome-wide association analysis for drought tolerance and associated traits in faba bean (*Vicia faba* L.). *Front Plant Sci*. 2023;14:1091875. <https://doi.org/10.3389/fpls.2023.1091875>.
33. Hasanuzzaman M, Al Mahmud J, Nahar K, Anee TI, Inafuku M, Oku H, Fujita M. Responses, adaptation, and ROS metabolism in plants exposed to waterlogging stress. In: Khan MIR, Khan NA (eds) *Reactive Oxygen Species and Antioxidant Systems in Plants: Role and Regulation under Abiotic Stress*. Singapore: Springer Singapore; 2017. pp. 257–281. https://doi.org/10.1007/978-981-10-5254-5_10.
34. Hoyos-Villegas V, Song Q, Kelly JD. Genome-wide association analysis for drought tolerance and associated traits in common bean. *Plant Genome*. 2017;10:2015.2012.0122. <https://doi.org/10.3835/plantgenome2015.12.0122>.
35. Jafarinasab A, Azari A, Siddique KHM, Madahhosseini S. Variation of yield and physiological characteristics of *Lathyrus sativus* L. populations under terminal drought. *Agric Water Manage*. 2022;273:107886. <https://doi.org/10.1016/j.agwat.2022.107886>.
36. Jaganathan D, Thudi M, Kale S, Azam S, Roorkiwal M, Gaur PM, et al. Genotyping-by-sequencing based intra-specific genetic map refines a “QTL-hotspot” region for drought tolerance in chickpea. *Mol Genet Genomics*. 2015;290:559–71. <https://doi.org/10.1007/s00438-014-0932-3>.
37. Jiang J, Su M, Chen Y, Gao N, Jiao C, Sun Z, et al. Correlation of drought resistance in grass pea (*Lathyrus sativus*) with reactive oxygen species scavenging and osmotic adjustment. *Biologia*. 2013;68:231–40. <https://doi.org/10.2478/s11756-013-0003-y>.
38. Kamenya SN, Mikwa EO, Song B, Odeny DA. Genetics and breeding for climate change in orphan crops. *Theor Appl Genet*. 2021;134:1787–815. <https://doi.org/10.1007/s00122-020-03755-1>.
39. Karadağ Y, İptas S, Yavuz M. Agronomic potential of grasspea (*Lathyrus sativus* L.) under rainfed condition in semi-arid regions of Turkey. *Asian J Plant Sci*. 2004;3:151–155. <https://scialert.net/abstract/?doi=ajps.2004.151.155>.
40. Khan MA, Tong F, Wang W, He J, Zhao T, Gai J. Analysis of QTL-allele system conferring drought tolerance at seedling stage in a nested association mapping population of soybean [*Glycine max* (L.) Merr.] using a novel GWAS procedure. *Planta*. 2018;248:947–62. <https://doi.org/10.1007/s00425-018-2952-4>.
41. Khatun M, Sarkar S, Era FM, Islam AKMAKMM, Anwar MP, Fahad S, et al. Drought stress in grain legumes: effects, tolerance mechanisms and management. *Agronomy*. 2021;11:1–35.
42. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*. 2013;9:29. <https://doi.org/10.1186/1746-4811-9-29>.
43. Kumar S, Bejiga G, Ahmed S, Nakkoul H, Sarker A. Genetic improvement of grass pea for low neurotoxin (β -ODAP) content. *Food Chem Toxicol*. 2011;49:589–600. <https://doi.org/10.1016/j.fct.2010.06.051>.
44. Kyu KL, Malik AI, Colmer TD, Siddique KHM, Erskine W. Response of mungbean (cvs. Celera II-AU and Jade-AU) and blackgram (cv. Onyx-AU) to

- transient waterlogging. *Front Plant Sci.* 2021;12:709102. <https://doi.org/10.3389/fpls.2021.709102>.
45. Kyu KL, Taylor CM, Douglas CA, Malik AI, Colmer TD, Siddique KHM, et al. Genetic diversity and candidate genes for transient waterlogging tolerance in mungbean at the germination and seedling stages. *Front Plant Sci.* 2024;15:1297096. <https://doi.org/10.3389/fpls.2024.1297096>.
 46. Lal M, Kumari A, Pooja, Sheokand S. Reactive oxygen species, reactive nitrogen species and oxidative metabolism under waterlogging stress. In: Hasanuzzaman M, Fotopoulos V, Nahar K, Fujita M (eds) *Reactive Oxygen, Nitrogen and Sulfur Species in Plants*. Hoboken, NJ: John Wiley & Sons; 2019. pp. 777–812. <https://doi.org/10.1002/9781119468677.ch34>.
 47. Lal MK, Tiwari RK, Altaf MA, Kumar A, Kumar R. Editorial: Abiotic and biotic stress in horticultural crops: insight into recent advances in the underlying tolerance mechanism. *Front Plant Sci.* 2023;14:1212982. <https://doi.org/10.3389/fpls.2023.1212982>.
 48. Lambein F, Travella S, Kuo Y-H, Van Montagu M, Heijde M. Grass pea (*Lathyrus sativus* L.): orphan crop, nutraceutical or just plain food? *Planta.* 2019;250:821–38. <https://doi.org/10.1007/s00425-018-03084-0>.
 49. Langan P, Bernád V, Walsh J, Henchy J, Khodaeiaminjan M, Mangina E, et al. Phenotyping for waterlogging tolerance in crops: current trends and future prospects. *J Exp Bot.* 2022;73:5149–69. <https://doi.org/10.1093/jxb/erac243>.
 50. Leitão ST, Alves ML, Pereira F, Zerrouk A, Godinho B, Barradas A, et al. Towards a trait-based approach to potentiate yield under drought in legume-rich annual forage mixtures. *Plants.* 2021;21:1763. <https://doi.org/10.3390/plants10091763>.
 51. Li Y, Ai Y, Zhang J, Fei J, Liu B, Wang J, et al. A novel expression vector for *Corynebacterium glutamicum* with an auxotrophy complementation system. *Plasmid.* 2020;107:102476. <https://doi.org/10.1016/j.plasmid.2019.102476>.
 52. Loreti E, Valeri MC, Novi G, Perata P. Gene regulation and survival under hypoxia requires starch availability and metabolism. *Plant Physiol.* 2018;176:1286–98. <https://doi.org/10.1104/pp.17.01002>.
 53. Ma F, Liang Y, Wang Q, Li W, Khan A, Li B, et al. Genome-wide identification of *FCS-like Zinc finger (FLZ)* genes in four Solanaceae plant species and functional characterization of *SIFLZ2* and *SIFLZ18* in tomato under heat stress. *Sci Hortic.* 2023;317:112015. <https://doi.org/10.1016/j.scienta.2023.112015>.
 54. Ma J, Zhao D, Tang X, Yuan M, Zhang D, Xu M, et al. Genome-wide association study on root system architecture and identification of candidate genes in wheat (*Triticum aestivum* L.). *Int J Mol Sci.* 2022;23:1843. <https://doi.org/10.3390/ijms23031843>.
 55. Malik AI, Ailewe TI, Erskine W. Tolerance of three grain legume species to transient waterlogging. *AoB Plants.* 2015;7:plv040. <https://doi.org/10.1093/aobpla/plv040>.
 56. Malosetti M, van der Linden CG, Vosman B, van Eeuwijk FA. A mixed-model approach to association mapping using pedigree information with an illustration of resistance to *Phytophthora infestans* in potato. *Genetics.* 2007;175:879–89. <https://doi.org/10.1534/genetics.105.054932>.
 57. Manik SMN, Quamruzzaman M, Zhao C, Johnson P, Hunt I, Shabala S, et al. Genome-wide association study reveals marker trait associations (MTA) for waterlogging-triggered adventitious roots and aerenchyma formation in barley. *Int J Mol Sci.* 2022;23:3341. <https://doi.org/10.3390/ijms23063341>.
 58. Martins D, Santos C, Sampaio AM, Rubiales D, Vaz Patto MC. *Lathyrus sativus* resistance against the existing and emerging pathogens *Erysiphe pisi* and *E. trifolii*: a case of commonalities or total discrepancy? *Phytopathology.* 2023;113:866–72. <https://doi.org/10.1094/PHYTO-06-22-0227-FI>.
 59. Martins DC, Rubiales D, Vaz Patto MC. Association mapping of *Lathyrus sativus* disease response to *Uromyces pisi* reveals novel loci underlying partial resistance. *Front Plant Sci.* 2022;13:842545. <https://doi.org/10.3389/fpls.2022.842545>.
 60. Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoel K, et al. ROS signaling: The new wave? *Trends Plant Sci.* 2011;16:300–9. <https://doi.org/10.1371/journal.pone.0278500>.
 61. Mittler R, Zandalinas SJ, Fichman Y, Van Breusegem F. Reactive oxygen species signalling in plant stress responses. *Nat Rev Mol Cell Biol.* 2022;23:663–79. <https://doi.org/10.1038/s41580-022-00499-2>.
 62. Mutari B, Sibiya J, Shayanowako A, Chidzanga C, Matova PM, Gasura E. Genome-wide association mapping for component traits of drought tolerance in dry beans (*Phaseolus vulgaris* L.). *PLoS One.* 2023;18:e0278500. <https://doi.org/10.1371/journal.pone.0278500>.
 63. Muthu V, Abbai R, Nallathambi J, Rahman H, Ramasamy S, Kambale R, et al. Pyramiding QTLs controlling tolerance against drought, salinity, and submergence in rice through marker assisted breeding. *PLoS One.* 2020;15:e0227421. <https://doi.org/10.1371/journal.pone.0227421>.
 64. Narayana PK, von Wettberg EJB. GWAS and genomic approaches in legumes, an expanding toolkit for examining responses to abiotic stresses. In: Hasanuzzaman M, Araújo S, Gill SS (eds) *The Plant Family Fabaceae: Biology and Physiological Responses to Environmental Stresses*. Singapore: Springer Singapore; 2020. pp. 161–180. https://doi.org/10.1007/978-981-15-4752-2_6.
 65. Noctor G, Mhamdi A, Foyer CH. The roles of reactive oxygen metabolism in drought: not so cut and dried. *Plant Physiol.* 2014;164:1636–48. <https://doi.org/10.1104/pp.113.233478>.
 66. Noctor G, Mhamdi A, Foyer CH. Oxidative stress and antioxidative systems: recipes for successful data collection and interpretation. *Plant Cell Environ.* 2016;39:1140–60. <https://doi.org/10.1111/pce.12726>.
 67. Piowarczyk B, Tokarz K, Makowski W, Łukasiewicz A. Different acclimatization mechanisms of two grass pea cultivars to osmotic stress in *in vitro* culture. *Acta Physiol Plant.* 2017;39:96. <https://doi.org/10.1007/s11738-017-2389-6>.
 68. Przybylska J, Zimniak-Przybylska Z, Krajewski P. Diversity of seed globulins in *Lathyrus sativus* L. and some related species. *Genet Resour Crop Evol.* 2000;47:239–46. <https://doi.org/10.1023/A:1008750423357>.
 69. Pucciariello C, Perata P. New insights into reactive oxygen species and nitric oxide signalling under low oxygen in plants. *Plant Cell Environ.* 2017;40:473–82. <https://doi.org/10.1111/pce.12715>.
 70. Pucciariello C, Perata P. Flooding tolerance in plants. In: Shabala S (ed) *Plant Stress Physiology*, 2nd edn. Ascot: CABI Digital Library; 2022. pp. 148–170. <https://doi.org/10.1079/9781845939953.0148>.
 71. Quan W, Chan Z, Wei P, Mao Y, Bartels D, Liu X. Phd finger proteins function in plant development and abiotic stress responses: an overview. *Front Plant Sci.* 2023;14:1297607. <https://doi.org/10.3389/fpls.2023.1297607>.
 72. Quan W, Liu X, Wang L, Yin M, Yang L, Chan Z. Ectopic expression of *Medicago truncatula* homeodomain finger protein, *MtPHD6*, enhances drought tolerance in *Arabidopsis*. *BMC Genomics.* 2019;20:982. <https://doi.org/10.1186/s12864-019-6350-5>.
 73. R Core Team. *A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2024.
 74. Rasheed R, Iqbal M, Ashraf MA, Hussain I, Shafiq F, Yousaf A, et al. Glycine betaine counteracts the inhibitory effects of waterlogging on growth, photosynthetic pigments, oxidative defence system, nutrient composition, and fruit quality in tomato. *J Horticult Sci Biotechnol.* 2018;93:385–91. <https://doi.org/10.1080/14620316.2017.1373037>.
 75. Rath D, Preeek A, Gayali S, Chakraborty S, Chakraborty N. Variety-specific nutrient acquisition and dehydration-induced proteomic landscape of grasspea (*Lathyrus sativus* L.). *J Proteomics.* 2018;183:45–57. <https://doi.org/10.1016/j.jpro.2018.05.013>.
 76. Rath D, Verma JK, Chakraborty S, Chakraborty N. Suspension cell secretome of the grain legume *Lathyrus sativus* (grasspea) reveals roles in plant development and defense responses. *Phytochemistry.* 2022;202:113296. <https://doi.org/10.1016/j.phytochem.2022.113296>.
 77. Rathore P, Shivashakarappa K, Ghimire N, Dumenyo K, Yadegari Z, Taheri A. Genome-wide association study for root system architecture traits in field soybean [*Glycine max* (L.) Merr.]. *Sci Rep.* 2024;14:25075. <https://doi.org/10.1038/s41598-024-76515-6>.
 78. Rubiales D, Annicchiarico P, Vaz Patto MC, Julier B. Legume breeding for the agroecological transition of global agri-food systems: a European perspective. *Front Plant Sci.* 2021;12:782574. <https://doi.org/10.3389/fpls.2021.782574>.
 79. Saleem A, Roldán-Ruiz I, Aper J, Muelle H. Genetic control of tolerance to drought stress in soybean. *BMC Plant Biol.* 2022;22:615. <https://doi.org/10.1186/s12870-022-03996-w>.
 80. Sampaio AM, Alves ML, Pereira P, Valliollahi E, Santos C, Šatović Z, Rubiales D, de Sousa Araújo S, van Eeuwijk F, Vaz Patto MC. Grass pea natural variation reveals oligogenic resistance to *Fusarium oxysporum* f. sp. *pisii*. *Plant Genome.* 2021;14:e20154. <https://doi.org/10.1002/tpg2.20154>.
 81. Sampaio AM, Vitale S, Turrá D, Di Pietro A, Rubiales D, van Eeuwijk F, et al. A diversity of resistance sources to *Fusarium oxysporum* f. sp. *pisii* found within grass pea germplasm. *Plant Soil.* 2021;463:19–38. <https://doi.org/10.1007/s1104-021-04895-z>.
 82. Sanchez M, Sampaio AM, Araújo S, van Eeuwijk F, Van Breusegem F, Vaz Patto MC. Grass pea (*Lathyrus sativus*) interesting panoply of mechanisms to cope with contrasting water stress conditions – a controlled study of sub-population differences in a worldwide collection of accessions. *Agric Water Manage.* 2024;292:108664. <https://doi.org/10.1016/j.agwat.2023.108664>.

83. Sansa O, Abberton MT, Ariyo J, Paliwal R, Ige A, Dieng I, et al. Genome-wide association studies of photosynthetic and agronomic traits in cowpea collection. *G3: Genes Genomes Genet.* 2024;14:jkae233.
84. Santos C, Polanco C, Rubiales D, Vaz Patto MC. The *MLO1* powdery mildew susceptibility gene in *Lathyrus species*: the power of high-density linkage maps in comparative mapping and synteny analysis. *Plant Genome.* 2021;14:e20090. <https://doi.org/10.1002/tpg2.20090>.
85. Sasidharan R, Voeselek LACJ. Ethylene-mediated acclimations to flooding stress. *Plant Physiol.* 2015;169:3–12. <https://doi.org/10.1104/pp.15.00387>.
86. Schwacke R, Ponce-Soto GY, Krause K, Bolger AM, Arsova B, Hallab A, et al. Mapman4: a refined protein classification and annotation framework applicable to multi-omics data analysis. *Mol Plant.* 2019;12:879–92. <https://doi.org/10.1016/j.molp.2019.01.003>.
87. Seki M, Umezawa T, Urano K, Shinozaki K. Regulatory metabolic networks in drought stress responses. *Curr Opin Plant Biol.* 2007;10:296–302. <https://doi.org/10.1016/j.pbi.2007.04.014>.
88. Sharmin RA, Karikari B, Chang F, Al Amin GM, Bhuiyan MR, Hina A, et al. Genome-wide association study uncovers major genetic loci associated with seed flooding tolerance in soybean. *BMC Plant Biol.* 2021;21:497. <https://doi.org/10.1186/s12870-021-03268-z>.
89. Siddique KHM, Regan KL, Tennant D, Thomson BD. Water use and water use efficiency of cool season grain legumes in low rainfall Mediterranean-type environments. *Eur J Agron.* 2001;15:267–80. [https://doi.org/10.1016/S1161-0301\(01\)00106-X](https://doi.org/10.1016/S1161-0301(01)00106-X).
90. Soltani A, Mafii Moghaddam S, Oladzad-Abbasbadi A, Walter K, Kearns PJ, Vasquez-Guzman J, et al. Genetic analysis of flooding tolerance in an andean diversity panel of dry bean (*Phaseolus vulgaris* L.). *Front Plant Sci.* 2018;9:767. <https://doi.org/10.3389/fpls.2018.00767>.
91. Stukkens Y, Bultreys A, Grec S, Trombik T, Vanham D, Boutry M. NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana glauca*, plays a major role in plant pathogen defense. *Plant Physiol.* 2005;139(1):341–52. <https://doi.org/10.1104/pp.105.062372>.
92. Szabados L, Savouré A. Proline: a multifunctional amino acid. *Trends Plant Sci.* 2010;15:P89–97. <https://doi.org/10.1016/j.tplants.2009.11.009>.
93. Tafesse EG, Gali KK, Lachagari VBR, Bueckert R, Warkentin TD. Genome-wide association mapping for heat and drought adaptive traits in pea. *Genes.* 2021;12:1897. <https://doi.org/10.3390/genes12121897>.
94. Thoen MPM, Davila Olivas NH, Kloth KJ, Coolen S, Huang PP, Aarts MGM, et al. Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytol.* 2017;213:1346–62. <https://doi.org/10.1111/nph.14220>.
95. Tokarz B, Wójtowicz T, Makowski W, Jędrzejczyk RJ, Tokarz KM. What is the difference between the response of grass pea (*Lathyrus sativus* L.) to salinity and drought stress?—A physiological study. *Agronomy.* 2020;10:833. <https://doi.org/10.3390/agronomy10060833>.
96. Trenti M, Lorenzi S, Bianchedi PL, Grossi D, Failla O, Grando MS, et al. Candidate genes and SNPs associated with stomatal conductance under drought stress in *Vitis*. *BMC Plant Biol.* 2021;21:7. <https://doi.org/10.1186/s12870-020-02739-z>.
97. Varadharajan V, Rajendran R, Muthuramalingam P, Runthala A, Madhesh V, Swaminathan G, et al. Multi-omics approaches against abiotic and biotic stress—a review. *Plants.* 2025;14:865. <https://doi.org/10.3390/plants14060865>.
98. Vaz Patto MC, Skiba B, Pang ECK, Ochatt SJ, Lambein F, Rubiales D. *Lathyrus* improvement for resistance against biotic and abiotic stresses: from classical breeding to marker assisted selection. *Euphytica.* 2006;147:133–47. <https://doi.org/10.1007/s10681-006-3607-2>.
99. Verma A, Kaur L, Kaur N, Bhardwaj A, Pandey AK, Kandath PK. Genome editing of an oxalyl-CoA synthetase gene in *Lathyrus sativus* reveals its role in oxalate metabolism. *Plant Cell Rep.* 2024;43:280. <https://doi.org/10.1007/s00299-024-03368-8>.
100. Vigouroux M, Novák P, Oliveira LC, Santos C, Cheema J, Wouters RHM, et al. A chromosome-scale reference genome of grasspea (*Lathyrus sativus*). *Sci Data.* 2024;11:1035. <https://doi.org/10.1038/s41597-024-03868-y>.
101. Voeselek LACJ, Bailey-Serres J. Flood adaptive traits and processes: an overview. *New Phytol.* 2015;206:57–73. <https://doi.org/10.1111/nph.13209>.
102. VSN International. Genstat for Windows, 22nd edn. Hemel Hempstead, UK: VSN International; 2022.
103. Wang S, Xie F, Xu S. Estimating genetic variance contributed by a quantitative trait locus: a random model approach. *PLoS Comput Biol.* 2022;18:e1009923. <https://doi.org/10.1371/journal.pcbi.1009923>.
104. Wilkinson S, Davies WJ. ABA-based chemical signalling: the co-ordination of responses to stress in plants. *Plant Cell Environ.* 2002;25:195–210. <https://doi.org/10.1046/j.0016-8025.2001.00824.x>.
105. Wiraguna E, Malik AI, Colmer TD, Erskine W. Waterlogging tolerance of grass pea (*Lathyrus sativus* L.) at germination related to country of origin. *Exp Agric.* 2020;56:837–50. <https://doi.org/10.1017/S0014479720000356>.
106. Wu X, Sun T, Xu W, Sun Y, Wang B, Wang Y, et al. Unraveling the genetic architecture of two complex, stomata-related drought-responsive traits by high-throughput physiological phenotyping and GWAS in cowpea (*Vigna unguiculata* L. Walp). *Front Genet.* 2021;12:743758. <https://doi.org/10.3389/fgene.2021.743758>.
107. Xu P, Wu X, Muñoz-Amatrián M, Wang B, Wu X, Hu Y, et al. Genomic regions, cellular components and gene regulatory basis underlying pod length variations in cowpea (*V. unguiculata* L. Walp). *Plant Biotechnol J.* 2016;15:547–57. <https://doi.org/10.1111/pbi.12639>.
108. Ye H, Roorkiwal M, Valliyodan B, Zhou L, Chen P, Varshney RK, et al. Genetic diversity of root system architecture in response to drought stress in grain legumes. *J Exp Bot.* 2018;69:3267–77. <https://doi.org/10.1093/jxb/ery082>.
109. Yildizli A, Çevik S, Ünyayar S. Effects of exogenous myo-inositol on leaf water status and oxidative stress of *Capsicum annuum* under drought stress. *Acta Physiol Plant.* 2018;40:122. <https://doi.org/10.1007/s11738-018-2690-z>.
110. Zhou L, Cheng W, Hou H, Peng R, Hai N, Bian Z, et al. Antioxidative responses and morpho-anatomical alterations for coping with flood-induced hypoxic stress in grass pea (*Lathyrus sativus* L.) in comparison with pea (*Pisum sativum*). *J Plant Growth Regul.* 2016;35:690–700. <https://doi.org/10.1007/s00344-016-9572-7>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.