

Identification and Expression Analysis of *GARP* Superfamily Genes in Response to Nitrogen and Phosphorus Stress in *Spirodela Polyrhiza*

Xuyao zhao

Institute of Hydrobiology

Jingjing Yang

Institute of Hydrobiology

Xiaozhe Li

Institute of Hydrobiology

Gaojie Li

Institute of Hydrobiology

Zuoliang Sun

Institute of Hydrobiology

Yan Chen

Institute of Hydrobiology

Yimeng Chen

Institute of Hydrobiology

Manli Xia

Institute of Hydrobiology

Yixian Li

Institute of Hydrobiology

Lunguang Yao

Nanyang Normal University

Hongwei Hou (✉ Houhw@ihb.ac.cn)

Institute of Hydrobiology

Research Article

Keywords: Spirodela polyrhiza, GARP superfamily, NIGT1/HRS1/HHO transcription factor, PHR transcription factor, Expression pattern

Posted Date: March 9th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1384808/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: GARP transcription factors perform the critical roles in plant development and response to environmental stimulate, especially in the phosphorus (P) and nitrogen (N) sensing and uptake. *Spirodela polyrhiza* (giant duckweed) is widely used for phytoremediation and biomass production for its efficient N and P removal capacity and rapid growth. However, there has not yet been a comprehensive analysis of the *GRAP* gene family in *S. polyrhiza*.

Results: We conducted a comprehensive study of the *GARP* genes in *S. polyrhiza*. First, we investigated 35 *SpGARP* genes in *S. polyrhiza*, which have been classified into three groups on the basis of their gene structure, conserved motifs and phylogenetic relationship. Then, we identified the duplication events, performed the synteny analysis, calculated the K_a/K_s ratio in *SpGARP* genes. Next, the regulation and co-expression networks of *SpGARPs* were constructed using *cis*-acting element analysis and weighted correlation network analysis (WGCNA). Last, the expression pattern of *SpGARP* genes were performed used RNA-seq data and qRT-PCR, several members of group IIId were involved in both N and P starvation response.

Conclusions: The study provide insight into the evolution and functional annotation of *GARP* gene family in *S. polyrhiza*, and lays the foundation for the further functional verification of *SpGARP* genes.

Background

GARP (G: GOLDEN2 in *Zea mays*, AR: ARR-B in *Arabidopsis thaliana*, P: Psr1 in *Chlamydomonas reinhardtii*) is a plant specific transcription factor (TF) superfamily with diverse functions: nutrient (N and P) sensing, chloroplast development, circadian clock oscillation and hormonal signaling [1, 2]. GARP superfamily TFs are made of type-B authentic response regulators (ARR-B) and GOLDEN2-like (G2-like, GLK) TFs which both contain GARP motifs (also named B motifs), first be named by Riechmann in *Arabidopsis thaliana* [1, 2]. B-motif is a domain of about 60 amino acids which forms a three- α -helices 3D structure containing a helix-turn-helix (HTH) motif, it also be the multifunctional domain responsible for both nuclear localization and DNA binding [2, 3].

GOLDEN2 gene was first reported in maize which functioned in chloroplast biogenesis, then, the orthologues of *GOLDEN2* have been identified in *A. thaliana*, *Physcomitrella patens*, *Capsicum annuum*, *Oryza sativa* and *Solanum lycopersicum* [4–8]. *AtGLK1/2* are the orthologous genes of *ZmGOFEN2* in *A. thaliana*, they are expressed in photosynthetic tissue and exhibit functional redundancy, however, the double mutants reduced the accumulation of photosynthetic gene products and thylakoids in chloroplasts [7]. *ZmGOLDEN2* and their orthologues in *O. sativa* and *A. thaliana* are regulated by light [6, 7]. Subsequently, *GOLDEN2* and its orthologues was classified into the GARP family with ARR-B and Psr1-like genes [1]. A gene can be classified in the GARP family if the derived protein contain B-motif [2]. ARR-B proteins contain a B-motif in the C-terminal, act as positive regulators in the two-component cytokinin signaling pathway and play pivotal roles in plant development, including vascular development, light sensitivity, chlorophyll production, hypocotyl elongation, and cell division in the root and shoot [9]. *Phosphorus Starvation Response 1 (Psr1)* gene in *Chlamydomonas* and *Phosphate starvation Response 1 (PHR1)* in *Arabidopsis* both shown to be the central regulatory module in inorganic phosphate (Pi) sensing system [10, 11].

The duckweed family (*Lemnaceae*) is a group of rapidly reproducing free-floating aquatic plants, and is distributed in various fresh water environments throughout the world in all but the most extreme habitats [12]. There are 36 recognized duckweed species belonging to five genera (*Spirodela* (2), *Landoltia* (1), *Lemna* (12), *Wolffiella* (10), and *Wolffia* (11)) based on the morphological characteristics and molecular taxonomy [13, 14]. Duckweed is supposed to be an attractive model for plant research for the convenient laboratory culture, clear genetic background, and robust transformation methods [15]. They are some of the fastest growing flowering plants (doubling time < 30 h under the optimal growth conditions) in the world, with the high productivity of dry mass (80–100 tons per hectare per year) which is more than five

times that of maize [16–18]. Duckweed exhibits efficient N and P removal capacity from wastewater that can reach 1.3 g/m²/d and 0.18 g/m²/d, respectively [19]. Thus, duckweed is also considered as the ideal plant for the phytoremediation of eutrophic water to recover nutrients (N and P) from wastewaters. *S. polyrhiza* occupies the ancestral phylogenetic position among duckweeds, possesses the largest individual and the smallest genome size in the *Lemnaceae* [20]. The prominent performance of *S. polyrhiza* in nutrient remove from wastewater was observed in previous studies, for its environmental adaptability and economically attractive application of duckweed biomass [21, 22]. The whole genome sequencing and the robust transformation method of *S. polyrhiza* providing theoretical and technical support to the research of molecular mechanism and germplasm improvement [23–28]. Thus, *S. polyrhiza* is one of the best prospective species in Lemnaceae for phytoremediation and biomass production.

N and P are two major essential elements required for plant growth. Nitrate (NO₃[−]) and ammonium (NH₄⁺) are the main forms of N source in soils, inorganic phosphate (Pi) is the main form of P source for plant uptake in the environment [29, 30]. In plants, it crucial to adapt itself to many sorts of nutritmental environments, however, little did we know about the molecular mechanism. The GARPs TFs are involved in the responses to nutrients and include probable nutrient sensors, CrPsr1 is the first reported GARP TF shown to be involved in nutritional responses, it is critical for acclimation of *Chlamydomonas* to phosphorus starvation [10]. AtPHR1 was proved to be the central regulator in the downstream of Pi starvation signaling pathway [11, 31]. Then, Nitrate-Inducible Garp-Type Transcriptional Repressor 1 (NIGT1)/Hypersensitivity to Low Phosphate-Elicited Primary Root Shortening 1 (HRS1)/HRS1 Homolog (HHO) were found to be the most robustly and quickly NO₃[−] regulated genes [32–34]. Even more interesting is that NIGT1 TFs are considered as potentially at the crossroads of N and P signals through interact with PHR1, Nodule Inception (NIN)-like protein (NLP), and SYG1-Pho81-XPR1 (SPX) domain proteins [35].

The *GARP* superfamily or *GLK* subfamily has been identified and characterized in *A. thaliana*, rice, maize, tobacco, and *Gossypium hirsutum* [1, 36–39]. Landup Bhutia found that some GARP members (OsGLK10, OsGLK15, OsGLK22 and OsGLK30) were functioned as the Pi starvation tolerance in rice [37]. However, the *GARP* gene family has not been thoroughly examined in *S. polyrhiza*, to the best of our knowledge. Herein, we perform a comprehensive analysis of GARP superfamily in giant duckweed, including chromosomal locations, evolutionary perspectives, structural arrangement, and their functional role through gene expression analysis. The results presented here offer a robust platform for functional studies of the candidate *SpGARP* genes to understand the molecular mechanism of N/P sensing, acquisition and balance in plants, and allow their efficient use in the germplasm improvement to enhance the phytoremediation potential of *S. polyrhiza*.

Results

Genome-wide identification of GARP superfamily genes in *S. polyrhiza*

The GARP proteins in *Arabidopsis* and rice were used as queries to performed the BLASTP research in *S. polyrhiza* 7498 genome databases. To mining all the potential GARP proteins that harbored B-box domain in giant duckweed, the Hidden Markov Model (HMM) profile of B-box domain was built using HMMER 3.3.2 based on the identified SpGARPs [40], and was used for the further search of SpGARP proteins. A total of 35 *GARP* genes were identified in giant duckweed, including 7 *SpARR-B* and 28 *SpGLK* genes. The detailed information of *SpGARP* genes were listed in Table 1, including gene ID, genomic position, gene length, exon number, amino acid number, molecular weight, isoelectric point (pI), instability index, GRAVY and subcellular localization. The *GARP* genes were named as *SpARR-B1-7* and *SpGLK1-28* according to their chromosomal location. The length of SpGARP proteins ranges from 229 aa (SpGLK17) to 693 aa (SpARR-B2), with the average of 386 aa, and the Molecular Weight (MV) is ranging from 25.47 (SpGLK17) to 74.33 kDa (SpARR-B3). The pIs are around 5.0 in SpARR-B family, and ranging from 5.2 to 9.6 in GLK family. The instability index is range from 39.61 (SpGLK21) to 75.66 (SpGLK13) The GRAVY is ranging from −0.327 (SpARR-B2) to −0.964 (SpGLK1). SpARR-B6 was

predicted to localized to extracellular space for the 14 aa signal peptides (SP) (MAASILSLFPGGLG) in the N-terminal while other 34 SpGARP proteins were predicted to localized to nucleus. To examine the evolutionary relationships of *SpGARP* genes, GARPs in two close related species *Colocasia esculenta* (10 *CeARR-B* and 35 *CeGLK* genes) and *Wolffia australiana* (5 *WaARR-B* and 23 *WaGLK* genes) were identified (Table S2).

Table 1

Characteristic features of SpGARPs.

S. No.	Gene ID	Protein length (aa)	exons	Molecular weight (kDa)	Theoretical pI	Instability index	Grand average of hydropathicity (GRAVY)	Subcellular location
SpARR-B1	Spo004408	693	5	73.98	6.24	53.21	-0.327	nucleus
SpARR-B2	Spo001985	541	5	60.87	5.66	44.75	-0.523	nucleus
SpARR-B3	Spo006367	689	6	74.33	5.41	46.65	-0.439	nucleus
SpARR-B4	Spo007463	468	6	52.79	6.06	51.1	-0.537	nucleus
SpARR-B5	Spo008950	575	6	63.58	4.63	58.94	-0.504	extracellular space
SpARR-B6	Spo008951	527	6	58.10	5.18	56.33	-0.518	nucleus
SpARR-B7	Spo012555	526	11	57.50	5.85	64.41	-0.588	nucleus
SpGLK1	Spo000053	352	5	39.32	9.53	59.96	-0.964	nucleus
SpGLK2	Spo000614	356	1	37.80	6.57	64.92	-0.517	nucleus
SpGLK3	Spo003067	459	7	49.72	5.2	70.6	-0.628	nucleus
SpGLK4	Spo003413	294	6	32.47	6.67	49.39	-0.714	nucleus
SpGLK5	Spo005377	297	6	32.66	7.01	47.1	-0.665	nucleus
SpGLK6	Spo005389	376	5	40.46	6.49	73.16	-0.779	nucleus
SpGLK7	Spo001652	367	4	40.24	9.55	54.31	-0.402	nucleus
SpGLK8	Spo008782	322	1	34.54	6.43	54.53	-0.437	nucleus
SpGLK9	Spo010181	242	4	26.12	9.32	65.94	-0.617	nucleus
SpGLK10	Spo007174	433	6	46.84	5.66	60.43	-0.718	nucleus
SpGLK11	Spo007312	503	6	55.74	7.05	59.69	-0.808	nucleus
SpGLK12	Spo010995	411	7	44.39	6.13	57.49	-0.721	nucleus
SpGLK13	Spo011612	351	4	38.67	7.09	75.66	-0.649	nucleus
SpGLK14	Spo011622	282	6	30.84	8.98	59.36	-0.678	nucleus
SpGLK15	Spo008930	283	6	31.65	7.8	60.46	-0.622	nucleus
SpGLK16	Spo009589	239	6	25.73	9.59	75.36	-0.753	nucleus
SpGLK17	Spo012149	229	6	25.47	9.27	60.33	-0.669	nucleus
SpGLK18	Spo012150	246	6	27.39	8.48	45.72	-0.574	nucleus
SpGLK19	Spo012405	277	6	30.73	9.07	41.16	-0.565	nucleus
SpGLK20	Spo012640	306	4	33.94	8.53	61.63	-0.73	nucleus
SpGLK21	Spo012709	315	4	34.12	9.6	39.61	-0.555	nucleus

S. No.	Gene ID	Protein length (aa)	exons	Molecular weight (kDa)	Theoretical pI	Instability index	Grand average of hydropathicity (GRAVY)	Subcellular location
SpGLK22	Spo012971	307	6	32.82	8.68	55.89	-0.581	nucleus
SpGLK23	Spo013101	310	6	34.83	9.32	67	-0.904	nucleus
SpGLK24	Spo013406	481	8	52.79	5.63	68.13	-0.74	nucleus
SpGLK25	Spo014991	428	5	46.96	6.45	73.48	-0.84	nucleus
SpGLK26	Spo017076	269	6	29.81	8.83	56.22	-0.599	nucleus
SpGLK27	Spo018029	318	5	35.72	5.92	62.48	-0.815	nucleus
SpGLK28	Spo018160	432	4	47.01	5.71	55.43	-0.716	nucleus

Phylogenetic Analysis Of Garp Superfamily Genes In Five Species

The full-length amino acid sequences of SpGARPs, AtGARPs (14 *AtARR-B* and 41 *AtGLK* genes), OsGARPs (10 *OsARR-B* and 47 *OsGLK* genes), CeGARPs and WaGARPs were subjected to multiple sequence alignments using ClustalW [41]. The unrooted NJ phylogenetic tree was divided into eight branch which belonged to three big groups (I, IIa-f and III). As shown in Fig.1, the majority of ARR-B proteins were clustered in the group I, besides six members (*At4G12020.1*, *CeARR-B8*, *OsGLK27*, *SpARR-B7*, *WaARR-B3*, and *WaARR-B4*) clustered into the clade of IIa. The orthologous genes of *ZmGOLDEN2* of the five species were clustered into the subgroup IIb, Arabidopsis and rice both possess two members, while taro and giant duckweed have the single copy in their genome, however, three orthologous genes present in the smallest and simplest flowering plant *W. australiana*. Indicating the unusual evolutionary pathways of *ZmGOLDEN2* orthologous genes in Araceae plants, especially in the duckweed plants. A total of 20 *GLK* genes (5 *AtGLKs*, 5 *CeGLKs*, 4 *OsGLKs*, 2 *SpGLKs* and 4 *WaGLKs*) were clustered into subgroup IIc. The function of the members in subgroup IIc were involved in circadian oscillation, including *LUX* (*AT3G46640.3*), *BOA* (*AT5G59570.2*), *MYBC1* (*AT2G40970.1*), *OsPCL* (*OsGLK5*), *OsPCL-like* (*OsGLK21*) genes in Arabidopsis and rice [37]. The members of 25 NIGT1/HRS1/HHO (also named NIGT1) TFs clustered in to subgroup IId. NIGT1 TFs have been found coordinate nitrate and Pi responses in Arabidopsis [35, 42–45]. Six *SpGLK* genes belonged to NIGT1 subfamily and named based on their topological location in the phylogenetic tree: *SpHHO1* (*SpGLK6*), *SpHHO2* (*SpGLK9*), *SpHHO2* (*SpGLK13*), *SpHHO4* (*SpGLK14*), *SpNIGT1.1* (*SpGLK25*) and *SpNIGT1.2* (*SpGLK27*). Subgroup IIe harbored 23 *KANADI* (*KAN*) genes which play the key roles in organ positioning, cell type patterning and organ morphogenesis of *A. thaliana* SAM [46, 47]. Subgroup II f contained 35 members which functioned to leaf development [37], including 7 *AtGLKs*, 11 *OsGLKs*, 6 *CeGLKs*, 3 *SpGLKs* and 4 *WaGLKs*. All the Phosphate Starvation Response (PHR)/PHR1-like (PHL) TFs were clustered in group III, *SpGLK3* (*SpPHR1*) and *SpGLK12* (*SpPHR2*) exhibited a closer relationship to *AtPHR1* and *OsPHR2*, which are the center regulator of phosphate starvation response in Arabidopsis and rice, respectively [11, 48, 49].

Figure 1. **Phylogenetic trees and subgroup classification of GARP proteins in *A. thaliana*, *C. esculenta*, *O. sativa*, *S. polyrhiza*, and *W. australiana*.** The green rectangle, blue rectangle, red rectangle, red cycle and green cycle represent GARP proteins from *A. thaliana*, *C. esculenta*, *O. sativa*, *S. polyrhiza*, and *W. australiana*.

Chromosomal locations, duplication analysis, and synteny analysis of SpGARP genes

As shown in Fig. 2, a total of 35 *SpGARP* genes distributed in 15 chromosomes and one contig (tig00010334_1, *SpGLK28*), most chromosome contain 1–3 GARP genes except chr10 (4 *SpGARP* genes, *SpGLK15/16* and *SpARR-B5/6*) and chr11 (6 *SpGARP* genes, *SpGLK17-21* and *SpARR-B7*). The duplication events of *SpGARP* genes were analyzed using McScanX [50],

four segmental duplication events (*SpGLK3-SpGLK12*, *SpGLK5-SpGLK14*, *SpGLK6-SpGLK13* and *SpGLK17-SpGLK22*) and two tandem duplication events (*SpARR-B5-SpARR-B6*, *SpGLK17-SpGLK18*) were identified, all the duplicated gene pairs belonged to the same group.

A total of 10 *SpGARP* paralogs gene pairs were identified based on the stringent amino acid sequence homology analysis (Table S3). We also identified the orthologous relationship of *SpGARP* genes with *A. thaliana* (8 gene pairs), *C. esculenta* (42 gene pairs), *O. sativa* (49 gene pairs), and *W. australiana* (27 gene pairs) (Table S3). The K_a/K_s ratio of the paralogs and ortholog gene pairs was analyzed to estimate the evolutionary pressure on *SpGARPs*. The K_a/K_s ratio ranged from 0.15 to 0.51 (median 0.32) among the *SpGARP* paralogs, suggesting that these genes are under a purifying selection pressure. Similarly, the K_a/K_s ratio of *SpGARP* orthologs in *A. thaliana*, *C. esculenta*, *O. sativa* and *W. australiana* ranged from 0.07 to 0.61 (median 0.21), 0.01 to 0.38 (median 0.23), 0.06 to 0.55 (median 0.29), and 0.08 to 0.40 (median 0.19), respectively, indicating that all of them were under the influence of strong purifying or negative selection pressure (Fig. 3A and Table S3). Gene duplication time based on K_s substitution rate in the case of *SpGARP* paralogs was observed in the range of 22.3–88.4 MYA, with a median age of 39.7 MYA. The divergence time estimated for the *SpGARP* orthologs in *A. thaliana*, *C. esculenta*, *O. sativa* and *W. australiana* ranged from 57.47 to 119.08 (median 80.75), 22.67 to 75.74 (median 36.54), 30.49 to 166.16 (median 60.78), and 22.68 to 67.65 (median 33.70), respectively (Fig. 3B and Table S3). The Syntenic relationship of *SpGARP* orthologs in *C. esculenta* (Fig. 3C) and *O. sativa* (Fig. 3D) showed that IId (green) and III (brown) groups shared the majority of orthologs gene pairs.

Conserved motifs, gene structure and phylogenetic analysis of SpGARPs in *S. polyrhiza*

An unrooted phylogenetic tree was constructed to visualize the evolutionary relationships between *GARP* members, using 35 *SpGARP* protein sequences (Fig. 4A). As shown in Fig. 4B, 15 conserved motifs were identified using MEME, motif 1 and 2 were distributed in all the *SpGARP* proteins, motif 5 in the I and IIa cluster, motif 6/8/14/15 only present in the cluster I, motif 7/9/10/13 present in the cluster of IId, motif 3/4/11 present in group III. The amino acid sequences of B-motif in *SpGARP* proteins were obtained and aligned ClustalW using displayed in GeneDoc [41]. The B-motif only have one tryptophan (W) residue while the MYB domain have three and MYB-like domain have two W residues [2]. As shown in Fig. 4C, there is only one W residue in the B-motif and a consensus sequence ((A/K)SHLQ(K/M)) in the third helix in *SpGARP* TFs, as well as the B-motifs of *GARP* TFs in *Arabidopsis*, rice and tobacco [2, 37, 38].

The structure of exons/introns were determined through aligning the genomic DNA sequences and the full-length cDNA of *SpGARP* genes (Fig. 5A). There are 1 (*SpGLK2/8*) to 11 (*SpARR-B7*) exons in *SpGARP* genes, and most of the *SpGARP* genes containing 4–8 exons. The members belong to the same clade of sharing the gene structure and conserved motifs, such as I clade have 5–6 exons, *SpGLK2* and *SpGLK8* which belong to IIa group only have 1 exon, and the *SpNIGT1s* contain 4–5 exons.

cis-regulatory element analysis of the *SpGARP* gene family

GARP TFs perform diverse functions in plants, such as nutrient sensing, chloroplast development and circadian clock oscillation. The *cis*-regulatory elements of the promoter regions play important roles in the expression of *GARP* genes for the environmental signals, such as nutrient stress and abiotic stress [36–38]. As shown in Table S4, 246 potential *cis*-acting elements were identified in the promoter region of *SpGARP* genes, including the *cis*-elements involved in stress and hormone responses. To explore the possible expression patterns of *SpGARP* genes under the environmental signals, such as P and N starvation, the distribution of phosphorus starvation related and nitrogen starvation related *cis*-elements in *SpGARP* genes were exhibited in Fig. 5B. Most of the members of *SpGARP* genes contained two or more type of phosphorus starvation related and nitrogen starvation related *cis*-elements in their promoter region. WRKY71OS is the most abundant phosphorus related *cis*-elements, as well as NODCON2GM in nitrogen starvation related *cis*-elements of *SpGARPs*.

Figure 5. **Gene structures and cis-elements distribution of SpGARP genes. A, the intron–exon structures of SpGARP genes. B, Frequency of phosphorus and nitrogen starvation related cis-elements in the 2-kb promoter region of SpGARP genes.**

We also predicted the TFs that may be involved in regulating the expression of *SpGARP* genes. As shown in Fig. 6, most of the *SpGARP* genes were interact with multiple TFs, suggesting that they may be involved in many physiological processes. Among the predicted transcription factors, *Apetala2* (AP2), *Barley B Recombinant/Basic Pentacysteine* (BBR-BPC), *C2H2* zinc finger (C2H2), *DNA binding with one finger* (Dof), *Ethylene Response Factor* (ERF), *Gibberellic-acid Insensitive* (GAI)/*Repressor of Gai* (RGA)/*Scarecrow* (SCR) (GRAS), *G2-like*, *Three Amino-acid Loop Extension* (TALE), and *MIKC-type MADS-box* (MIKC_MADS) were the most abundant (Supplementary Table 2 and Fig. 6). Several *SpGARP* genes were regulated by GARP superfamily TFs, six members (*SpARR-B1/3* and *SpGLK1/5/6/20*) contained the ARR-B TF binding site in their promoter regions, and 17 members (3 *SpARR-B* genes and 14 *SpGLK* genes) were regulated by G2-like TFs.

Transcriptional patterns of GARP genes in *S. polyrhiza*

The RNA-seq data of *S. polyrhiza* were analyzed to explore the expression pattern of *SpGARP* genes. As shown in Fig. 7, the expression of three members of group I (*SpARR-B4/5/6*) haven't been detected in all the transcriptome data. *SpARR-B7*, *SpGLK10* and several members of group III (*SpGLK3/13/18/24/22*) were robust expressing in both fronds and roots as well as under the various conditions. The members of I group (*SpSRR-B1/2/3*), IIc (*SpGLK2/8*), IId (*SpGLK9/25/28*), IIe (*SpGLK1/11*), IIl (*SpGLK5/7/16/20*) and III (*SpGLK12/21/26/23*) were differential expression in fronds and roots, most of them were upregulated in roots. NIGT1s were identified as the critical regulators in the plant nutrient sensing, especially in the N, P starvation response, all the members of IId subgroup were regulated expression under N or/and P deprivation. Four out of six members (*SpGLK9/13/25/27*) of group IId were upregulated in the condition of P deprivation (-P/+N), while downregulated in the condition of N deprivation (+P/-N, -P/-N and H₂O). PHRs/PHLs were a class of MYB-CC TFs which were central regulators of phosphate starvation response, only one member (*SpGLK23*) of group III shown the regulatory expression under the nutrient stresses. *SpARR-B2*, *SpGLK2*, and *SpGLK25* were response to salt stress.

To further understand the function of *SpGARP* genes under different N and P conditions. The correlations between the expression patterns of *SpGARP* genes was analyzed (Fig.8). A total of 69 *SpGARP* gene pairs showed correlation expression patterns, including 54 positive correlation and 15 negative correlation *SpGARP* gene pairs. The expression of three IId members (*SpGLK9*, *SpGLK25*, and *SpGLK27*) were positive correlation, and were both negative correlation with *SpGLK6* and *SpGLK21*. Some *SpGARP* genes that belonged to different groups showed correlation expression patterns, such as *SpGLK7/10/14/17/18/20*, *SpARR-B1* and *SpGLK2/6/11/21/28*, *SpARR-B3* and *SpGLK14/23/24*.

Figure 8. **The correlations between the expression patterns of SpGARP genes.**

The gene regulation network of SpGARPs and N/P response genes

To identify the potential cooperative genes of *SpGARPs*. WGCNA was performed using 15 RNA-seq data (PRJNA724886) of *S. polyrhiza* under different N and P conditions. After removed the low-expression genes (averaged TPM < 1), a total of 14,424 genes were used for network construction. Those genes clustered into 18 modules (labeled with different colors), 32 *SpGARP* genes were assigned to 9 modules: 17 *SpGARPs* (*SpARR-B1-3* and *SpGLK1/2/6/8/9/11/14/17/18/21/24/25/27/28*) assigned to turquoise module, 5 members (*SpGLK7/10/12/20/23*) assigned to brown module, *SpGLK16* and *SpARR-B7* assigned to blue module, *SpGLK15* and *SpGLK22* assigned to red module, and other six modules each contained one *SpGARP* member (*SpGLK26*-cyan, *SpGLK4*-green, *SpGLK5*-magenta, *SpGLK13*-salmon, *SpGLK19*-tan, and *SpGLK3*-yellow). The gene co-expression network of *SpGARPs* and N/P response genes in the turquoise module were constructed (Fig. 9). There were 46 genes were in the network, including 5 members of NIGT1 subfamily (*SpGLK6/9/25/27/28*), 10 *phosphate transporter* (*PHT*) genes, 3 *ammonium transporter* (*AMT*) genes, 11 *nitrate transporter 1* (*NRT1*) genes, 3 *nitrate transporter 2* (*NRT2*) genes, *NLP1/2*, *nitrate reductase 2* (*NR2*), and *nitrite reductase 1* (*NIR1*).

Expression analysis of PHR and NIGT1 genes in response to P or/and N deprivation by qRT-PCR

To further clarify the potential ability of the *GARP* genes to respond to the P and N stress, the expression profiles of six *NIGT1* subfamily genes (*SpGLK6*, *SpGLK9*, *SpGLK13*, *SpGLK25*, *SpGLK27*, *SpGLK28*) and two *PHR* genes (*SpGLK3*, *SpGLK12*) were verified using qRT-PCR. *SpGLK13* have not been detected in the cDNA samples from both CK and nutrient stress treatments. As showed in Fig. 10, *SpGLK3* was upregulated under nitrogen starvation (NS) treatment (NS5 and NS7) and low phosphorus (LP) treatment (LP1, LP5 and LP7). The expression of *SpGLK12* was strongly induced by NS treatment. *SpGLK9/25/27* showed the *similar* expression profiles under PS and NS treatments, they both upregulated under PS and downregulated under NS treatment, with the exception of *SpGLK27* in the NS7. *SpGLK9* and *SpGLK27* were also upregulated under LP and LN treatments. The expression of *SpGLK6/28* were induced under PS, NS, LP and LN treatments, especially under the NS treatment.

Discussion

GARP is a plant specific TF superfamily with diverse functions, and have been identified and in several species, including *Arabidopsis*, rice, maize, tobacco, cotton. However, it had not been reported in giant duckweed and limited our comprehending of the molecular mechanism of N/P responses. In this research, we identified 35 SpGARPs which belonged to three groups, and further classified into eight subgroups. The number of *SpGARP* (35) and *WaGARP* genes (28) in duckweeds was lower than *A. thaliana* (56), *C. esculenta* (46), *O. sativa* (56), in consistent with the previous report of WRKY TF family, which may be related to the minimal gene set of duckweed genomes (about 19,000 genes in *S. polyrhiza*, 15,000 genes in *W. australiana*) that represents a nonredundant catalog of core plant proteins [23, 51]. However, *S. polyrhiza* had more IId members than *C. esculenta*, *O. sativa* and *W. australiana*, suggesting that NIGT1 subfamily have undergone expanded in the adopting to various aquatic environments, and it should be responsible for the high capability of nutrient uptake in *S. polyrhiza*. The SpGARPs gathered in the same subfamily had the similar gene and protein structures, suggesting the strong evolutionary conservation of *SpGARP* genes [52]. ARR-B family was classified into I and IIa group, however, IIa and IIb showed the closer relationship in the phylogenetic tree, as well as the more similar gene and protein structures, indicating the diverse origins of ARR-B proteins.

GARP members play important roles in many physiological processes as transcriptional regulators [2]. The earlier research was focused on chloroplast development, and their function of nutrient sensing and uptake gaining increasing attention in the plant science [2, 45, 53]. PHRs are the center regulator in the response of phosphate starvation, however, the transcriptional level expression of *AtPHR1*, *OsPHR2* haven't been improved under the Pi starvation condition [11, 31, 49]. SPX protein withheld PHRs in the cytoplasm under Pi sufficient condition to avoid the toxicity of high Pi, the repression be relaxed to activate phosphate starvation response under Pi deficient condition [54, 55]. NIGT1 could regulate the expression of *SPX* genes in transcriptional level, and NIGT1–SPX–PHR cascade mediates regulation of phosphate uptake and starvation signal [35, 42].

TFs regulate the expression of target genes at the transcriptional level through binding to the *cis*-element in the promoter region. Many types of *cis*-acting elements are present in the promoter of *SpGARP* genes, including abundant of phosphorus starvation related and nitrogen starvation related *cis*-element, indicating that most *SpGARPs* are involved in hormone and stress response, especially phosphorus starvation and nitrogen starvation. The expression pattern of *SpGARP* genes were analyzed in this study, 18 out of 35 *SpGARP* genes were differential expressing between fronds and roots, and most members of group IId and III were regulated in roots, indicating that the root might be playing important roles in mineral sensing and uptake in *S. polyrhiza*. The N or/and P starvation alter the expression of group I and IId genes, especially *SpGARP6/13/9/25/27*, which is in accord with previous research in *Arabidopsis* [33, 34]. Using WGCNA, the co-expression network showed that regulatory relationship between *SpGARP* and N, P response genes, which were involved in N, P sensing, uptake and assimilation. In qRT-PCR analysis, most of the *SpPHRs* and *SpNIGT1* genes were differentially expressed under phosphorus starvation (PS), nitrogen starvation (NS), low phosphorus (LP), and low nitrogen (LN)

treatment. Interestingly, *NIGT1s* showed different expression patterns, such as *SpGLK6/28* were strong induced by NS while *SpGLK9/25/27* were downregulated under NS. The different expression pattern between PS and LP, NS and LN suggesting that these genes were involved in N/P sensing.

Giant duckweed is distributed throughout the world, and is spotlighted in the plant science, environment renovation, biomass energy, and food security [15]. The discharge of agricultural, industrial and domestic wastewater led to the accumulation of nitrogen, phosphorus in water, and has resulted in dramatic water eutrophication and algae blooms. It destroys the ecological balance of the water bodies, lowers the content of dissolved oxygen, deteriorates water quality, and even causes the death of aquatic creatures. Phytoremediation has been recommended as an alternative solution to treat eutrophic wastewater due to its cost-effective, environment friendly and sustainable characteristics. Duckweed is widely used in the phytoremediation for several outstanding characteristics: fast growth rate, adaptability to a wide range of environmental conditions, high nutrient uptake capacity, and higher proportion of carbohydrate [56]. Duckweed adapt itself to various nutrimental environments, and balanced the N and P uptake and assimilation, our comprehensive analysis of *GARP* gene superfamily in giant duckweed showed that the members of *SpNIGT1* subfamily should be the hub regulator of N and P sensing and acquisition.

Conclusion

In this study, we identified 35 *GARP* genes in *S. polyrhiza* genome, including 7 *ARR-B* subfamily genes and 28 *GLK* subfamily genes. The gene structures and phylogenetic analysis suggest a complex evolution history of this gene family in *S. polyrhiza*. As the central regulator in the P and N nutrient, the *PHR/PHL* subfamily genes were transcriptional induced by NS and LN treatments, while the *NIGT1* subfamily genes were response to P and N stresses, especially *SpGLK9/25/27* which upregulated under PS treatment and downregulated under NS treatment. This study provide insight into the evolution and functional annotation of *GARP* gene family in *S. polyrhiza*, and facilitate the further functional verification of *SpGARP* genes.

Materials And Methods

Genome-wide identification of GARP proteins in *S. polyrhiza*

The genome data *S. polyrhiza* 7498 (55878) was downloaded from the Comparative Genomics Platform (<https://genomeevolution.org/CoGe>) [23]. The genome of *C. esculenta* (CNP0001082) was downloaded from China National GeneBank DataBase (CNGDBdb, <https://db.cngb.org/>), and the genome of *W. australiana* 8730 was downloaded from Wolffia DB (<https://duckweeds.plantprofile.net/>) [57]. The published *GARP* superfamily genes (G2-like and *ARR-B*) of *A. thaliana* [2], *O. sativa* [37] were downloaded from The Arabidopsis Information Resource (<http://www.arabidopsis.org/>), and the Rice Genome Annotation Project (<http://rice.plantbiology.msu.edu/>) respectively [58]. The peptide sequences of *GARP* from Arabidopsis, and rice were used as queries for protein basic local alignment search tool (BLASTp) analysis against whole-genome sequences in the *S. polyrhiza* 7498 v3 genomes with an e-value cutoff set as $1e^{-10}$. The HMM profile of B-box domain was built using HMMER 3.3.2 based on the identified SpGARPs and used as query for HMMER search (e-value < $1e^{-5}$). All putative *GARP* were further checked using Pfam database (<http://pfam.xfam.org/>), Conserved Domains Database (CDD, <https://www.ncbi.nlm.nih.gov/cdd/>), and Simple Modular Architecture Research Tool database (SMART, <http://smart.embl.de/smart/batch.pl>). The obtained sequences were aligned with PacBio isoform sequencing data (SRX5321175) of *S. polyrhiza* 7498 to identify the full-length proteins. The *GARP* genes of *C. esculenta* and *W. australiana* were identified using the same methods.

Basic Physicochemical Properties And Phylogenetic Analysis Of Spgarp

The molecular weight (MW), isoelectric point (pI) and grand average of hydropathicity (GRAVY) of SpGARP proteins were calculated for each gene using ExPASy (<http://www.expasy.org/tools/>) [59]. The subcellular localization of proteins was determined by analysis of the WoLF PSORT (<https://wolfsort.hgc.jp/>), CELLO (<http://cello.life.nctu.edu.tw/>), and Bologna Unified Subcellular Component Annotator (BUSCA, <http://busca.biocomp.unibo.it>), and decided based on consensus localization for two or more algorithms [60–62].

The protein sequences of *GARP* TFs from *A. thaliana*, *C. esculenta*, *O. sativa*, *S. polyrhiza* and *W. australiana* were aligned using ClustalW [41]. A neighbor-joining (NJ) phylogenetic tree was constructed in MEGAX (<http://www.megasoftware.net/>) based the multiple sequence alignment with 1000 bootstrap replicates, and displayed using Interactive Tree of Life (<https://itol.embl.de/>) [63, 64].

Gene duplication and K_a/K_s analysis

The information regarding chromosome length and gene locations of *GARP* family genes in *S. polyrhiza* was extracted from the Generic Feature Format files. The duplication events were defined based on the collinearity analysis of candidate gene pairs using MCScanX [50]. The synteny analysis of *SpGARP* family genes with *CeGARPs* and *OsGARPs* was performed using MCScanX and visualized using Circos [65]. The non-synonymous substitution rate (K_a) and synonymous substitution rate (K_s) of the duplication and orthologous gene pairs were calculated using PAMLX [66].

Gene Structure Analysis And Identification Of Conserved Motifs

Multiple Expectation Maximizations for Motif Elicitation (MEME, <http://meme-suite.org>) was employed for the analysis of conserved motifs in GARP proteins with the following parameters: maximum number of motifs, 15; motif length, 6 to 50 amino acids [67]. The structure of *GARP* genes, including intron and exon information, was visualized using the online tools Gene Structure Display Server 2.0 (<http://gsds.gao-lab.org/index.php>) [68].

Promoter analysis of SpGARP genes

Cis-acting elements in the promoter region of *GARP* genes (2000 bp upstream of the start codon) were predicted and analyzed in New PLACE (<https://www.dna.affrc.go.jp/PLACE/?action=newplace>) [69]. The subset of data representing phosphorus and nitrogen starvation related *cis* elements was visualized using TBtools [70]. To discover the transcription factors which involved in the regulatory expression of *SpGARP* genes, the online tool PlantRegMAP (http://plantregmap.gao-lab.org/binding_site_prediction.php) was used to predicted the potential binding sites of TFs in the promoter region of *SpGARP* genes, the regulation networks between *SpGARP* genes and potential transcription factors was presented using Cytoscape 3.7.0 [71, 72].

Rna-seq Atlas Analysis

The temporal and spatial expression profiles of *SpGARPs* in different tissues/organs (leaves, roots, and stipule, Bio-Project PRJNA557001), under various nutrient stresses (supply with P and N (+ P/+N), N deprivation (+ P/-N), P deprivation (- P/+N), P and N deprivation (- P/-N), and none nutrient supply (H_2O) for seven days, PRJNA724886), under salt stress (salt-0 h, salt-6 h, salt-12 h, salt-24 h, PRJNA563960) were obtained using the publicly available transcriptome data from NCBI [23, 73]. The heat map showing the correlations between the expression patterns of *SpGARP* genes was generated with the gplots package in R (version 4.1.2, <https://www.R-project.org/>). The significant correlation gene pairs should satisfy both of the following criteria: the absolute of the correlation value > 0.85, and *p*-value < 0.05. The genes whose coverage TPM > 1 were filtered and used for WGCNA. Co-expression network modules were identified with the WGCNA

package 1.63 in R to generate the signed weighted correlation network, and the network of genes was visualized in Cytoscape 3.7.0 [72, 74].

Rna Isolation And Qrt-pcr Analysis

S. polyrhiza strain 7498 was cultivated using the same methods in previous research [51]. Then, duckweed was treated in half-strength MS solution (Control, CK), half-strength MS solution without P (phosphorus starvation (PS) treatment), half-strength MS solution without N (nitrogen starvation (NS) treatment), half-strength MS solution with 1 μM KH_2PO_4 (low phosphorus (LP) treatment), half-strength MS solution with 1 μM NH_4NO_3 (low nitrogen (LN) treatment). The samples were harvested respectively at a set time points (0, 1, 3, 5, 7 and 10 days) after the treatment, and were immediately frozen in liquid nitrogen and then stored at -80°C for RNA isolation. The oligonucleotide primers of *SpGARP* genes were shown in Table S1. Total RNA, first-strand cDNA was synthesized, and qRT-PCR program was performed using the same methods in previous research [51]. Each reaction was analyzed in triplicate and the $2^{-\Delta\Delta\text{CT}}$ method was used to analyze the data [75].

Abbreviations

GARP

GOLDEN2-ARR-B-Psr1

ARR-B

Type-B authentic response regulators

GLK

GOLDEN2-like

Psr1

Phosphorus Starvation Response 1

PHR1

Phosphate Starvation Response 1

NIGT1

Nitrate-Inducible Garp-Type Transcriptional Repressor 1

HRS1

Hypersensitivity to Low Phosphate-Elicited Primary Root Shortening 1

HHO

HRS1 Homolog

Pi

Inorganic phosphate

HMM

Hidden Markov Model

MW

Protein molecular weight

pI

Isoelectric point

GRAVY

Grand average of hydropathicity

K_a/K_s

The ratio of the number of nonsynonymous substitutions per nonsynonymous site (K_a) to the number of synonymous substitutions per synonymous site (K_s)

TPM

Transcripts Per Million
qRT-PCR
Quantitative Real-time PCR
WGCNA
Weighted correlation network analysis

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

RNA-seq data of *Spirodela polyrhiza* under nutrient (Nitrogen and Phosphorus) deficiency have been deposited in NCBI SRA with the accession number PRJNA724886. RNA-seq data of *Spirodela polyrhiza* tissues/organs are available in NCBI under accession number PRJNA557001. RNA-seq data of *Spirodela polyrhiza* under salt stress are available in NCBI under accession number PRJNA563960. The expression data of *SpPHR* and *SpNIGT1* genes which generated by qRT-PCR were available from the corresponding authors when needed. All other data generated or analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China [Grant number, 32001107, 31870384]; National Key R & D Program [Grant numbers, 2020YFD0900305, 2018YFD0900801, 2017YFE0128800]; International Partnership Program of Chinese Academy of Sciences [152342KYSB20200021], Open project of State Key Laboratory of Freshwater Ecology and Biotechnology [2021FB04].

Author contribution statement

Xuyao Zhao: Conceptualization, Investigation, Writing-original draft. **Jingjing Yang:** Investigation, Data Curation, Preparation of materials. **XiaoZhe Li:** Investigation, Data Curation. **Gaojie Li** and **Zuoliang Sun:** Methodology, Formal analysis. **Yan Chen** and **Yimeng Chen:** Methodology, Conceptualization. **Manli Xia** and **Yixian Li:** Methodology, Data curation. **Lunguang Yao:** Writing-review & editing. **Hongwei Hou:** Writing-review & editing, Funding acquisition, Supervision, Project administration. All authors read and approved the manuscript.

References

1. Riechmann JL, Heard J, Martin G, Reuber L, Jiang CZ, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR *et al.* **Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes.** *Science* 2000, **290**(5499):2105-2110.
2. Safi A, Medici A, Szponarski W, Ruffel S, Lacombe B, Krouk G: **The world according to GARP transcription factors.** *Curr Opin Plant Biol* 2017, **39**:159-167.

3. Hosoda K, Imamura A, Katoh E, Hatta T, Tachiki M, Yamada H, Mizuno T, Yamazaki T: **Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the Arabidopsis response regulators.** *Plant Cell* 2002, **14**(9):2015-2029.
4. Hall LN, Rossini L, Cribb L, Langdale JA: **GOLDEN 2: A novel transcriptional regulator of cellular differentiation in the maize leaf.** *Plant Cell* 1998, **10**(6):925-936.
5. Brand A, Borovsky Y, Hill T, Rahman KAA, Bellalou A, Van Deynze A, Paran I: **CaGLK2 regulates natural variation of chlorophyll content and fruit color in pepper fruit.** *Theor Appl Genet* 2014, **127**(10):2139-2148.
6. Rossini L, Cribb L, Martin DJ, Langdale JA: **The maize Golden2 gene defines a novel class of transcriptional regulators in plants.** *Plant Cell* 2001, **13**(5):1231-1244.
7. Fitter DW, Martin DJ, Copley MJ, Scotland RW, Langdale JA: **GLK gene pairs regulate chloroplast development in diverse plant species.** *Plant J* 2002, **31**(6):713-727.
8. Bravo-Garcia A, Yasumura Y, Langdale JA: **Specialization of the Golden2-like regulatory pathway during land plant evolution.** *New Phytol* 2009, **183**(1):133-141.
9. Argyros RD, Mathews DE, Chiang YH, Palmer CM, Thibault DM, Etheridge N, Argyros DA, Mason MG, Kieber JJ, Schaller GE: **Type B response regulators of Arabidopsis play key roles in cytokinin signaling and plant development.** *Plant Cell* 2008, **20**(8):2102-2116.
10. Wykoff DD, Grossman AR, Weeks DP, Usuda H, Shimogawara K: **Psr1, a nuclear localized protein that regulates phosphorus metabolism in Chlamydomonas.** *Proc Natl Acad Sci U S A* 1999, **96**(26):15336-15341.
11. Bari R, Pant BD, Stitt M, Scheible WR: **PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants.** *Plant Physiol* 2006, **141**(3):988-999.
12. Xu YL, Ma S, Huang M, Peng M, Bog M, Sree KS, Appenroth KJ, Zhang JM: **Species distribution, genetic diversity and barcoding in the duckweed family (Lemnaceae).** *Hydrobiologia* 2015, **743**(1):75-87.
13. Bog M, Appenroth KJ, Sree KS: **Duckweed (Lemnaceae): Its molecular taxonomy.** *Front Sustain Food S* 2019, **3**:117.
14. Les DH, Crawford DJ, Landolt E, Gabel JD, Kembell RT: **Phylogeny and systematics of Lemnaceae, the duckweed family.** *Syst Bot* 2002, **27**(2):221-240.
15. Acosta K, Appenroth KJ, Borisjuk L, Edelman M, Heinig U, Jansen MAK, Oyama T, Pasaribu B, Schubert I, Sorrels S *et al*: **Return of the Lemnaceae: Duckweed as a model plant system in the genomics and post-genomics era.** *Plant cell* 2021, **33**:3207-3234.
16. Lam E, Appenroth KJ, Michael T, Mori K, Fakhoorian T: **Duckweed in bloom: the 2nd International Conference on Duckweed Research and Applications heralds the return of a plant model for plant biology.** *Plant Mol Biol* 2014, **84**(6):737-742.
17. Michael TP, Bryant D, Gutierrez R, Borisjuk N, Chu P, Zhang HZ, Xia J, Zhou JF, Peng H, El Baidouri M *et al*: **Comprehensive definition of genome features in Spirodela polyrhiza by high-depth physical mapping and short-read DNA sequencing strategies.** *Plant J* 2017, **89**(3):617-635.
18. Ziegler P, Adelmann K, Zimmer S, Schmidt C, Appenroth KJ: **Relative in vitro growth rates of duckweeds (Lemnaceae)-the most rapidly growing higher plants.** *Plant Biol* 2015, **17**:33-41.
19. Cheng JJ, Stomp AM: **Growing duckweed to recover nutrients from wastewaters and for production of fuel ethanol and animal feed.** *Clean-Soil Air Water* 2009, **37**(1):17-26.
20. Wang W, Haberer G, Gundlach H, Glasser C, Nussbaumer T, Luo MC, Lomsadze A, Borodovsky M, Kerstetter RA, Shanklin J *et al*: **The Spirodela polyrhiza genome reveals insights into its neotenus reduction fast growth and aquatic lifestyle.** *Nature Communications* 2014, **5**:3311.
21. Ng YS, Chan DJC: **Phytoremediation capabilities of Spirodela polyrhiza, Salvinia molesta and Lemna sp. in synthetic wastewater: A comparative study.** *Int J Phytoremediation* 2018, **20**(12):1179-1186.

22. Ng YS, Chan DJC: **The enhancement of treatment capacity and the performance of phytoremediation system by fed batch and periodic harvesting.** *Rsc Adv* 2021, **11**(11):6049-6059.
23. An D, Zhou Y, Li CS, Xiao Q, Wang T, Zhang YT, Wu YR, Li YB, Chao DY, Messing J *et al*: **Plant evolution and environmental adaptation unveiled by long-read whole-genome sequencing of *Spirodela*.** *Proc Natl Acad Sci U S A* 2019, **116**(38):18893-18899.
24. Xu SQ, Stapley J, Gablenz S, Boyer J, Appenroth KJ, Sree KS, Gershenzon J, Widmer A, Huber M: **Low genetic variation is associated with low mutation rate in the giant duckweed.** *Nature Communications* 2019, **10**:1243.
25. Harkess A, McLoughlin F, Bilkey N, Elliott K, Emenecker R, Mattoon E, Miller K, Czymmek K, Vierstra RD, Meyers BC *et al*: **Improved *Spirodela polyrhiza* genome and proteomic analyses reveal a conserved chromosomal structure with high abundance of chloroplastic proteins favoring energy production.** *J Exp Bot* 2021, **72**(7):2491-2500.
26. Hoang PNT, Michael TP, Gilbert S, Chu P, Motley ST, Appenroth KJ, Schubert I, Lam E: **Generating a high-confidence reference genome map of the greater duckweed by integration of cytogenomic, optical mapping, and Oxford Nanopore technologies.** *Plant J* 2018, **96**(3):670-684.
27. Yang JJ, Lia GJ, Hua SQ, Bishopp A, Heenatigala PPM, Kumar S, Duan PF, Yao LG, Hou HW: **A protocol for efficient callus induction and stable transformation of *Spirodela polyrhiza* (L.) Schleiden using *Agrobacterium tumefaciens*.** *Aquat Bot* 2018, **151**:80-86.
28. Chen Y, Li G, Yang J, Zhao X, Sun Z, Hou H: **Role of Nramp transporter genes of *Spirodela polyrhiza* in cadmium accumulation.** *Ecotoxicol Environ Saf* 2021, **227**:112907-112907.
29. Fageria VD: **Nutrient interactions in crop plants.** *J Plant Nutr* 2001, **24**(8):1269-1290.
30. Elser JJ, Bracken MES, Cleland EE, Gruner DS, Harpole WS, Hillebrand H, Ngai JT, Seabloom EW, Shurin JB, Smith JE: **Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems.** *Ecol Lett* 2007, **10**(12):1135-1142.
31. Rubio V, Linhares F, Solano R, Martin AC, Iglesias J, Leyva A, Paz-Ares J: **A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae.** *Genes Dev* 2001, **15**(16):2122-2133.
32. Krouk G, Mirowski P, LeCun Y, Shasha DE, Coruzzi GM: **Predictive network modeling of the high-resolution dynamic plant transcriptome in response to nitrate.** *Genome Biol* 2010, **11**(12):R123.
33. Canales J, Moyano TC, Villarroel E, Gutierrez RA: **Systems analysis of transcriptome data provides new hypotheses about *Arabidopsis* root response to nitrate treatments.** *Front Plant Sci* 2014, **5**:22.
34. Krouk G, Tranchina D, Lejay L, Cruikshank AA, Shasha D, Coruzzi GM, Gutierrez RA: **A systems approach uncovers restrictions for signal interactions regulating genome-wide responses to nutritional cues in *Arabidopsis*.** *PLoS Comp Biol* 2009, **5**(3):e1000326.
35. Ueda Y, Kiba T, Yanagisawa S: **Nitrate-inducible NIGT1 proteins modulate phosphate uptake and starvation signalling via transcriptional regulation of SPX genes.** *Plant J* 2020, **102**(3):448-466.
36. Zhao Z, Shuang J, Li Z, Xiao H, Liu Y, Wang T, Wei Y, Hu S, Wan S, Peng R: **Identification of the Golden-2-like transcription factors gene family in *Gossypium hirsutum*.** *Peerj* 2021, **9**:e12484.
37. Bhutia KL, Nongbri EL, Gympad E, Rai M, Tyagi W: **In silico characterization, and expression analysis of rice golden 2-like (OsGLK) members in response to low phosphorous.** *Mol Biol Rep* 2020, **47**(4):2529-2549.
38. Qin MY, Zhang BH, Gu G, Yuan JZ, Yang XS, Yang JH, Xie XF: **Genome-wide analysis of the G2-like transcription factor genes and their expression in different senescence stages of tobacco (*Nicotiana tabacum* L.).** *Front Genet* 2021, **12**:626352.
39. Liu F, Xu YJ, Han GM, Zhou LY, Ali A, Zhu SW, Li XY: **Molecular evolution and genetic variation of G2-like transcription factor genes in maize.** *PLoS One* 2016, **11**(8):e0161763.

40. Karplus K, Barrett C, Hughey R: **Hidden Markov models for detecting remote protein homologies.** *Bioinformatics* 1998, **14**(10):846-856.
41. Li KB: **ClustalW-MPI: ClustalW analysis using distributed and parallel computing.** *Bioinformatics* 2003, **19**(12):1585-1586.
42. Ludewig U, Vatov E, Hedderich D, Neuhauser B: **Adjusting plant nutrient acquisition to fluctuating availability: transcriptional co-regulation of the nitrate and phosphate deprivation responses in roots.** *J Exp Bot* 2021, **72**(10):3500-3503.
43. Kiba T, Inaba J, Kudo T, Ueda N, Konishi M, Mitsuda N, Takiguchi Y, Kondou Y, Yoshizumi T, Ohme-Takagi M *et al*: **Repression of nitrogen starvation responses by members of the Arabidopsis GARP-type transcription factor NIGT1/HRS1 subfamily.** *Plant Cell* 2018, **30**(4):925-945.
44. Maeda Y, Konishi M, Kiba T, Sakuraba Y, Sawaki N, Kurai T, Ueda Y, Sakakibara H, Yanagisawa S: **A NIGT1-centred transcriptional cascade regulates nitrate signalling and incorporates phosphorus starvation signals in Arabidopsis.** *Nature Communications* 2018, **9**:1376.
45. Wang X, Wang HF, Chen Y, Sun MM, Wang Y, Chen YF: **The transcription factor nigt1.2 modulates both phosphate uptake and nitrate influx during phosphate starvation in Arabidopsis and maize.** *Plant Cell* 2020, **32**(11):3519-3534.
46. Ram H, Sahadevan S, Gale N, Caggiano MP, Yu XL, Ohno C, Heisler MG: **An integrated analysis of cell-type specific gene expression reveals genes regulated by REVOLUTA and KANADI1 in the Arabidopsis shoot apical meristem.** *PLoS Genet* 2020, **16**(4):e1008661.
47. Caggiano MP, Yu XL, Bhatia N, Larsson A, Ram H, Ohno C, Sappl P, Meyerowitz EM, Jonsson H, Heisler MG: **Cell type boundaries organize plant development.** *Elife* 2017, **6**:e27421.
48. Sun LC, Song L, Zhang Y, Zheng Z, Liu D: **Arabidopsis PHL2 and PHR1 act redundantly as the key components of the central regulatory system controlling transcriptional responses to phosphate starvation.** *Plant Physiol* 2016, **170**(1):499-514.
49. Zhou J, Jiao FC, Wu ZC, Li YY, Wang XM, He XW, Zhong WQ, Wu P: **OsPHR2 is involved in phosphate-starvation signaling and excessive phosphate accumulation in shoots of plants.** *Plant Physiol* 2008, **146**(4):1673-1686.
50. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H *et al*: **MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity.** *Nucleic Acids Res* 2012, **40**(7):e49.
51. Zhao X, Yang J, Li G, Sun Z, Hu S, Chen Y, Guo W, Hou H: **Genome-wide identification and comparative analysis of the WRKY gene family in aquatic plants and their response to abiotic stresses in giant duckweed (Spirodela polyrhiza).** *Genomics* 2021, **113**(4):1761-1777.
52. Ohta T: **Gene conversion and evolution of gene families: An overview.** *Genes* 2010, **1**(3):349-356.
53. Guo MN, Ruan WY, Li CY, Huang FL, Zeng M, Liu YY, Yu YN, Ding XM, Wu YR, Wu ZC *et al*: **Integrative comparison of the role of the Phosphate Response1 subfamily in phosphate signaling and homeostasis in rice.** *Plant Physiol* 2015, **168**(4):1762-U1134.
54. Qi W, Manfield IW, Muench SP, Baker A: **AtSPX1 affects the AtPHR1-DNA-binding equilibrium by binding monomeric AtPHR1 in solution.** *Biochem J* 2017, **474**:3675-3687.
55. Puga MI, Mateos I, Charukesi R, Wang Z, Franco-Zorrilla JM, de Lorenzo L, Irigoye ML, Masiero S, Bustos R, Rodriguez J *et al*: **SPX1 is a phosphate-dependent inhibitor of Phosphate Starvation Response 1 in Arabidopsis.** *Proc Natl Acad Sci U S A* 2014, **111**(41):14947-14952.
56. Liu Y, Xu H, Yu CJ, Zhou GK: **Multifaceted roles of duckweed in aquatic phytoremediation and bioproducts synthesis.** *Gcb Bioenergy* 2021, **13**(1):70-82.
57. Park H, Park JH, Lee Y, Woo DU, Jeon HH, Sung YW, Shim S, Kim SH, Lee KO, Kim JY *et al*: **Genome of the world's smallest flowering plant, Wolffia australiana, helps explain its specialized physiology and unique morphology.** *Communications Biology* 2021, **4**(1):900.

58. Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu JZ, Zhou SG *et al*: **Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data.** *Rice* 2013, **6**:4.
59. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A: **ExPASy: the proteomics server for in-depth protein knowledge and analysis.** *Nucleic Acids Res* 2003, **31**(13):3784-3788.
60. Yu CS, Lin CJ, Hwang JK: **Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions.** *Protein Sci* 2004, **13**(5):1402-1406.
61. Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K: **WoLF PSORT: protein localization predictor.** *Nucleic Acids Res* 2007, **35**:W585-W587.
62. Wu Z-C, Xiao X, Chou K-C: **iLoc-Plant: a multi-label classifier for predicting the subcellular localization of plant proteins with both single and multiple sites.** *Mol Biosyst* 2011, **7**(12):3287-3297.
63. Kumar S, Stecher G, Li M, Knyaz C, Tamura K: **MEGA X: Molecular evolutionary genetics analysis across computing platforms.** *Mol Biol Evol* 2018, **35**(6):1547-1549.
64. Letunic I, Bork P: **Interactive Tree Of Life (iTOL) v4: recent updates and new developments.** *Nucleic Acids Res* 2019, **47**(W1):W256-W259.
65. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA: **Circos: An information aesthetic for comparative genomics.** *Genome Res* 2009, **19**(9):1639-1645.
66. Xu B, Yang ZH: **pamlX: A graphical user interface for PAML.** *Mol Biol Evol* 2013, **30**(12):2723-2724.
67. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: **MEME SUITE: tools for motif discovery and searching.** *Nucleic Acids Res* 2009, **37**:W202-W208.
68. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G: **GSDS 2.0: an upgraded gene feature visualization server.** *Bioinformatics* 2015, **31**(8):1296-1297.
69. Higo K, Ugawa Y, Iwamoto M, Korenaga T: **Plant cis-acting regulatory DNA elements (PLACE) database: 1999.** *Nucleic Acids Res* 1999, **27**(1):297-300.
70. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R: **TBtools: An integrative toolkit developed for interactive analyses of big biological data.** *Molecular Plant* 2020, **13**(8):1194-1202.
71. Tian F, Yang DC, Meng YQ, Jin JP, Gao G: **PlantRegMap: Charting functional regulatory maps in plants.** *Nucleic Acids Res* 2020, **48**(D1):D1104-D1113.
72. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: A software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13**(11):2498-2504.
73. Fu LL, Ding ZH, Sun XP, Zhang JM: **Physiological and transcriptomic analysis reveals distorted ion homeostasis and responses in the freshwater plant *Spirodela polyrhiza* L. under salt stress.** *Genes* 2019, **10**(10):743.
74. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis.** *BMC Bioinformatics* 2008, **9**:559.
75. Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR.** *Nucleic Acids Res* 2001, **29**(9):e45.

Figures

Figure 1

Phylogenetic trees and subgroup classification of GARP proteins in *A. thaliana*, *C. esculenta*, *O. sativa*, *S. polyrhiza*, and *W. australiana*. The green rectangle, blue rectangle, red rectangle, red cycle and green cycle represent GARP proteins from *A. thaliana*, *C. esculenta*, *O. sativa*, *S. polyrhiza*, and *W. Australian*.

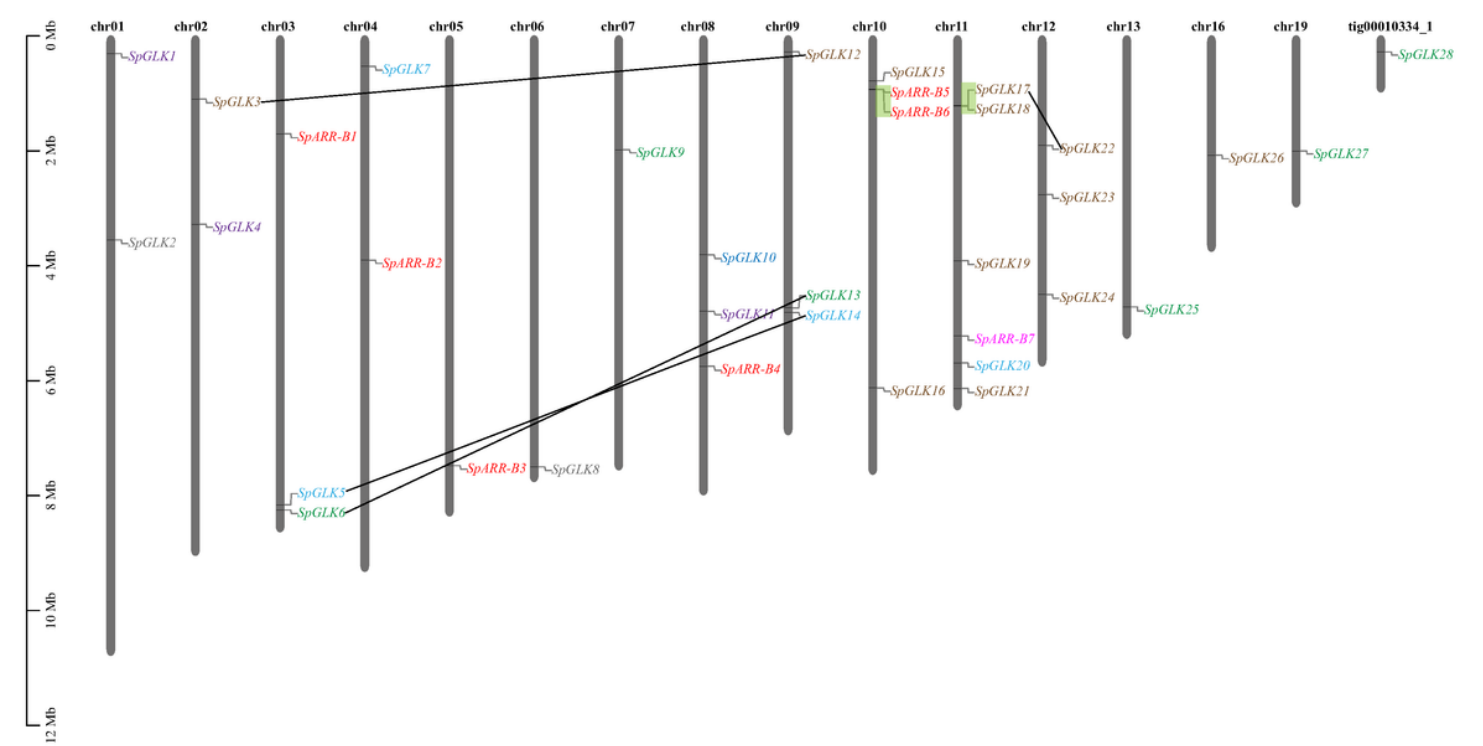


Figure 2

Chromosomal distribution of *SpGARP*. Segmental duplications are marked with Black solid line, tandem duplications are marked with green background.

Figure 3

Inference of *SpGARP* gene duplications. A) Box and whisker plots representing the distribution of K_a/K_s ratio in *SpGARP* paralogs and orthologs; B) Estimated gene duplication time of *SpGARP* paralogs and orthologs in MYA (MYA: million years ago); C) Syntenic relationship of *SpGARP* orthologs in chromosomes of *C. esculenta*; D) Syntenic relationship of *SpGARP* orthologs in chromosomes of *O. sativa*.

Figure 4

Phylogenetic analysis, conserved motifs (1-15) and the alignment of B-motif of *SpGARP* proteins. A, a neighbor-joining (NJ) phylogenetic tree based on the full-length protein sequences. B, the motifs were ordered manually based on the results of the MEME analysis. C, Multiple sequence alignment of the domain of B-motif in *SpGARPs*.

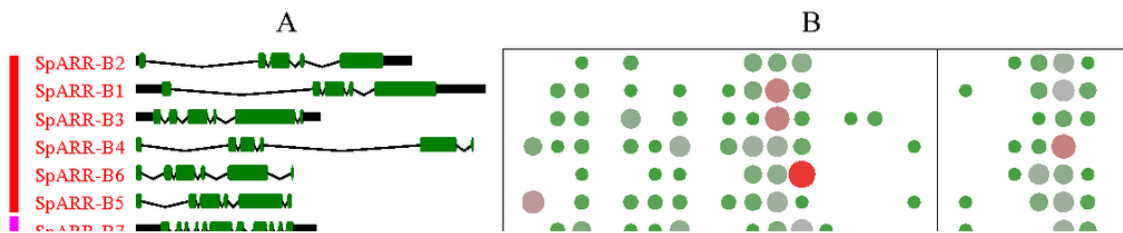


Figure 5

Gene structures and *cis*-elements distribution of *SpGARP* genes. A, the intron–exon structures of *SpGARP* genes. B, Frequency of phosphorus and nitrogen starvation related *cis*-elements in the 2-kb promoter region of *SpGARP* genes.

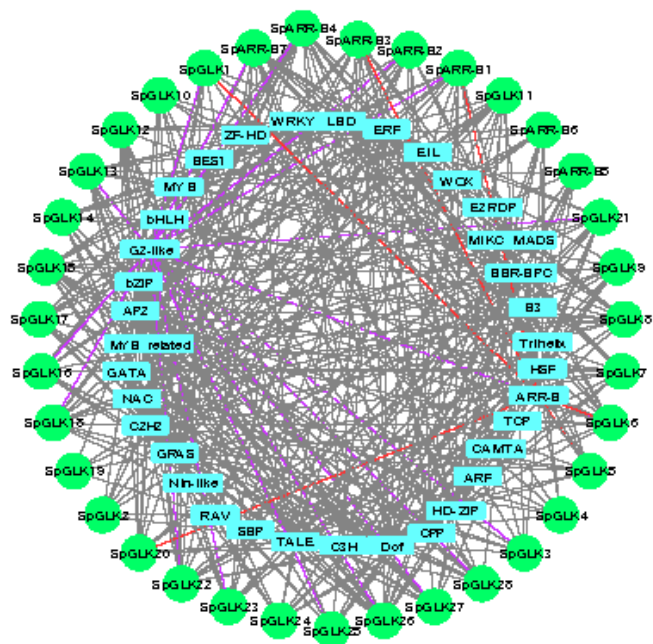


Figure 6
 Regulation networks between *SpGLK* and potential transcription factors.

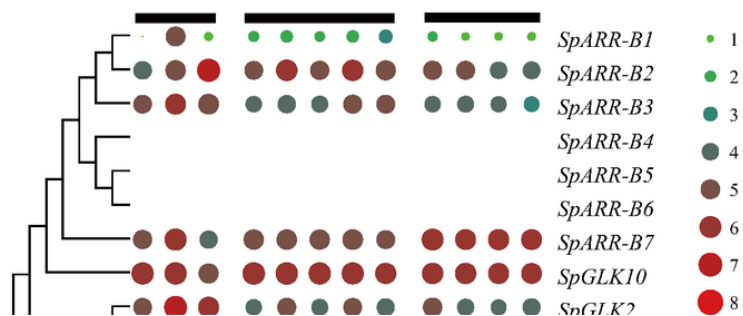


Figure 7

Expression profiles of *SpGARP* genes. The expression data were obtained from RNA-seq data and shown as log2 values based on TPM values (Transcripts Per Kilobase of exon model per Million mapped reads).

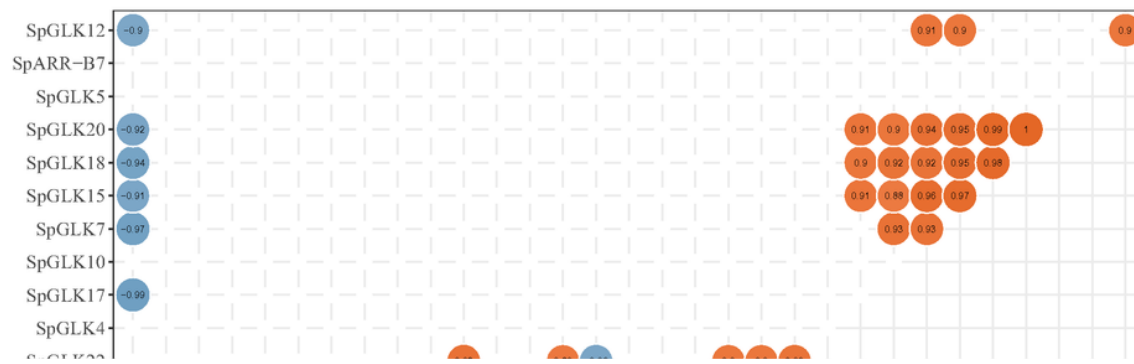


Figure 8

The correlations between the expression patterns of *SpGARP* genes.

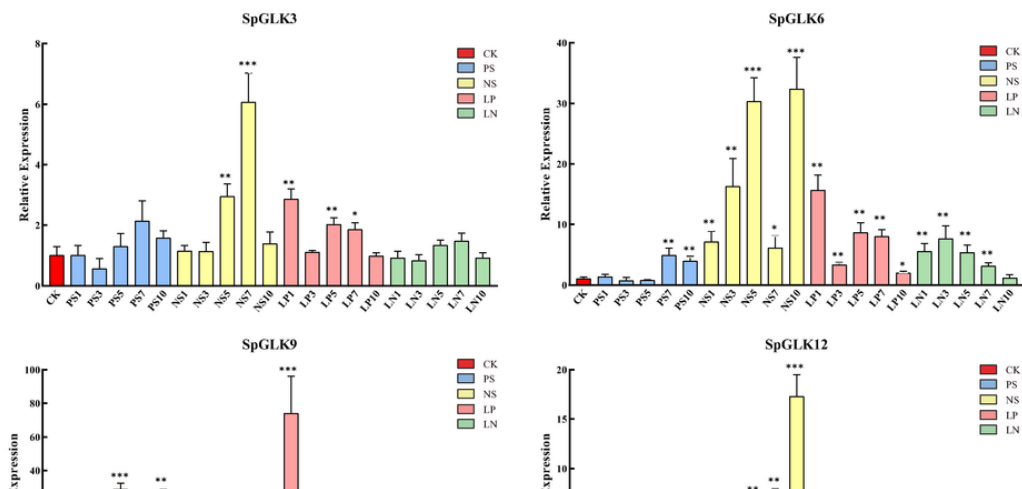


Figure 10

Relative expression of two *PHR* genes and five *NIGT1* genes under nutrient stress conditions. Y axis indicates the relative expression level and X axis represents the samples from different timepoints (0, 1, 3, 5, 7 and 10 day) under four treatments (PS, NS, LP and LN). Each data point represents mean value $\pm$ standard deviation (SD) ($n = 3$). Error bars indicate standard deviation. Asterisks indicate the significant degree of expression level compared to the value of the control (* < 0.05, ** < 0.01, *** < 0.001).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupportingMaterial.zip](#)