



OPEN

## Comprehensive sequence and expression profile analysis of the phosphate transporter gene family in soybean

Xiaoshuang Wei<sup>1</sup>, Yu Fu<sup>2</sup>, Renjie Yu<sup>2</sup>, Lei Wu<sup>2</sup>, Zhihai Wu<sup>1,3</sup>, Ping Tian<sup>1</sup>, Siyuan Li<sup>2</sup>, Xue Yang<sup>2</sup>✉ & Meiyang Yang<sup>2</sup>✉

The family of phosphate transporters (PHTs) mediates the uptake and translocation of Pi inside the plants. However, little is known about transporters in soybean. Therefore, Searched the Genome Database for Soybean, 57 *GmPHTs* family members were identified in soybean, Phylogenetic analysis suggested that members of the PHTs gene family can be divided into six clades. Collinearity analysis revealed that most of the *GmPHT* genes shared syntenic relationships with PHTs members in *Arabidopsis thaliana* and that large segment duplication played a major driving force for *GmPHTs* evolution in addition to tandem duplication. Further analysis of the promoter revealed that light-responsive elements and abiotic stress-responsive elements were widely distributed within the promoter regions of *GmPHT* genes. Based on RNA-seq data, *GmPHTs* showed different expression patterns in roots and leaves of soybean treated with long-term low phosphorus and short-term low phosphorus, in addition, the expression levels of *GmPHT* genes can be regulated by drought stresses, it was implied that the induced expression of *GmPHTs* could promote phosphorus uptake and transport in soybean and thus adapt to low phosphorus and drought stress, which is the first step dissection of Pi transport system and probably refers to new roles of PHTs genes in soybean.

Phosphorus (Pi) is a key nutrient for plant growth and development, is the second essential macronutrient required for plant growth and development alongside nitrogen<sup>1</sup> and is also an essential component of fertilizers used to sustain modern agriculture because it helps quality and yield. Pi is part of important biomolecules and, in the form of phosphate, pyrophosphate, adenosine triphosphate, adenosine diphosphate, or adenosine monophosphate, involved in various metabolic functions and many living cell regulatory processes are P-dependent<sup>2</sup>. Globally, many soils are deficient in phosphate (Pi)<sup>3</sup>, Pi concentrations within plant cells are typically 1000-times those outside<sup>4</sup>. Therefore, plants must have specialized transport vehicles to transfer Pi from the soil to the cells to counter the large concentration gradient at the root-soil interface. The active phosphorus in the soil and the phosphorus used in various metabolic processes ultimately require direct uptake and transport through the "phosphate transporter protein (PHT)", a highly efficient plant root uptake and transport system, which is responsible for the transport of inorganic phosphorus in plants under the regulation of transcription factors. Genome sequence analyses and experimental evidence have indicated that plants contain numerous Pi transporter families, including *PHT1*, *PHT2*, *PHT3*, *PHT4*, and *PHT5*, which were distinguished by their protein sequences, structures, locations, and functions<sup>5–11</sup>.

The *PHT1* family has been most extensively studied in plants, and the main pathway for Pi to enter plants from the soil is through the *PHT1* family located on the plasma membrane, PHT type 1 protein is also defined as a major contributor to the Pi uptake system<sup>12</sup> and the plant has several *PHT1* transporter members that play a key role in P acquisition, transport and remobilization. The *Arabidopsis AtPHT1* gene family contains nine members, The rice *OsPHT1* gene family has 13 members<sup>13</sup>. The first high H<sup>+</sup>/PI phosphate transporter identified in higher plants was *AtPT1* from *Arabidopsis*<sup>14</sup>. This gene has an important role in the uptake of phosphorus from the soil<sup>15</sup>. Analysis of expressed sequence tags (ESTs) and genomic sequences revealed that nine genes in *Arabidopsis* are similar to *AtPT1*, with *PHT1;1* transcript being the most abundant<sup>16</sup>. Its overexpression increases Pi uptake in

<sup>1</sup>College of Agronomy, Jilin Agricultural University, Changchun 130118, Jilin, China. <sup>2</sup>College of Life Sciences, Jilin Agricultural University, Changchun 130118, Jilin, China. <sup>3</sup>National Crop Variety Approval and Characterization Station, Jilin Agricultural University, Changchun 130118, Jilin, China. ✉email: xyang@jlau.edu.cn; jlaumeiyang@163.com

*Arabidopsis*<sup>17</sup>. The *PHT2* gene family contains one member each in *Arabidopsis thaliana* and rice<sup>13</sup>, and orthologous gene have been characterized in wheat, alfalfa rhizobia and several *Solanum* species<sup>14</sup>. *Arabidopsis PHT2;1* is the first member of the *PHT2* family to be identified, with *OsPHT2;1* in rice being the putative low-affinity phosphate transporter gene involved in the accumulation and transport of Pi in the plant<sup>18</sup>. The third family of plant Pi transporters is localized to the mitochondria, Three *PHT3* genes have been identified in *Arabidopsis*<sup>8</sup> and six members identified in rice<sup>13</sup>, and homologs have been cloned and partially characterized soybean, maize, and birch<sup>19,20</sup>. The *Arabidopsis* genome contains six *PHT4* genes (*PHT4;1* to *PHT4;6*), and chloroplast localization of *PHT4;1*, *PHT4;2*, *PHT4;4*, and *PHT4;5* has been analyzed using GFP fusions or immunoblotting<sup>21–23</sup>. In *Arabidopsis*, phosphate transporter proteins containing the SPX–MFS structural domain have been designated as members of the phosphate transporter protein family *PHT5* or VPT. The results also confirmed that *PHT5* could increase plant tolerance under low phosphorus stress. The results also confirmed that *PHT5* could increase plant tolerance under low phosphorus stress<sup>24</sup>. *GmPHT5* in soybean may be responsible for transporting phosphorus from root vascular tissues to nodules, especially under phosphorus deficiency stress, and therefore *GmPHT5* may be a key transporter in regulating nodule phosphorus homeostasis<sup>25</sup>.

Soybean (*Glycine max* (L.) Merr.) is an economical and agronomical crop, which has been adopted as a diet staple throughout many parts of the world<sup>6</sup>. However, the low availability of phosphorus (P) in the soil is the main nutrient limiting factor for soybean yield<sup>26,27</sup>. Previous studies have reported that phosphorus addition improves drought tolerance in soybean<sup>28,29</sup>. Phosphorus application increased nitrogen fixation rates in soybean<sup>30,31</sup> which is sensitive to drought stress<sup>32–34</sup>. In production areas within China, drought is the most challenging stress for soybean<sup>35</sup> and the non-renewable resource of P is predicted to have limited availability, soybean seed yield will be threatened by both water and P deficits in the future<sup>36</sup>. Therefore, it is urgent that researchers develop plants with enhanced efficiency of soil phosphorus use under such conditions, improve the capacity of soybean to absorb phosphorus. Because Pi is moved from the soil into plant cells in response to excess phosphate, genomic analyses have been conducted with Pi transporter families in *Arabidopsis*, rice and apple.

Currently, the study of phosphorus transport protein genes in *Arabidopsis*, rice and other plants has gained great progress, while PHTs have been rarely studied in soybean. Therefore, based on the published *Arabidopsis thaliana* genome, this study identified and characterized putative soybean PHTs family genes through phylogenetic analysis, collinearity analysis, motif, promoter element analysis, and investigated the expression profile of PHTs genes in response to low phosphorus and drought. Transcriptome sequencing was used to analyze the induction of expression of root and leaf phosphorus transporter-related genes in soybean by low phosphorus treatment. The objective was to present a foundation for further functional dissection of *GmPHTs* so that genetic engineering approaches can be applied to improve the efficiency of phosphate uptake by stressed soybean plants.

## Result

**Identification of the PHTs family genes in soybean.** Based on the homology with PHTs family members in *Arabidopsis*, a total of 57 PHTs genes were identified in the whole genome of soybean (Table 1). Each subfamily member also has its own characteristics. The differences in the isoelectric point and hydrophilicity of the subfamily proteins are relatively obvious, Except for the *GmPHT5* subfamily, the isoelectric points of all subfamily members are basic, the isoelectric point of the *GmPHT5* subfamily is acidic, at around 6. The *GmPHT1*, *GmPHT2*, *GmPHT3*, *GmPHT4*, and *GmPHT5* subfamilies had positive hydrophilicity values ranging from 0.101 to 0.659, while the *GmPHO1* subfamily had negative hydrophilicity values ranging from –0.255 to 0.064. In contrast, the hydrophilicity among the members of the *GmPHT4* subfamily varied considerably, with GRAVY values ranging from 0.2 to 0.6. The differences in the basic physicochemical properties of the proteins reflect the differences in their physiological functions. In particular, the differences in isoelectric point and hydrophilicity exhibited among the members within the *GmPHT4* subfamily herald the diverse functional differentiation among the members of the *GmPHT4* subfamily.

**Phylogenetic analysis of the GmPHTs genes.** To determine the evolutionary relationships of the members of the *GmPHTs* gene family genes with those of the soybean species, we constructed a phylogenetic tree comprising 92 PHTs proteins, from soybean (57), *Arabidopsis* (35), based on a multialignment via MEGA 6. Our results showed that all the PHTs homologs could be classified into six clades. *GmPHT1* subfamily (14 members), *GmPHT2* subfamily (2 members), *GmPHT3* subfamily (9 members), *GmPHT4* subfamily (12 members), *GmPHT5* subfamily (6 members) and *GmPHO1* subfamily (14 members), respectively, highly similar to each other. In addition, almost all of the soybean *PHTs* genes appeared as pairs with the PHTs members in *Arabidopsis* in terms of phylogenetic relationships (Fig. 1).

**Collinearity analysis among members of the GmPHTs gene family.** There are currently three ways in which duplicated genes are generated, including whole-genome duplication (polyploidy), large segmental duplication events and tandem duplication event<sup>37</sup>. Chromosomal distribution and Collinearity analysis of the *GmPHTs* gene (Fig. 2) shows that, 49% of the 57 *GmPHTs* gene family members were found to arise through large segmental duplication events and 14% through tandem duplication event. In the *GmPHT1* subfamily, the *GmPHT1.1* and *GmPHT1.4*, *GmPHT1.2* and *GmPHT1.11* genes are generated by a large segmental duplication event in the soybean chromosome; the *GmPHT1.6* and *GmPHT1.7*, *GmPHT1.13* and *GmPHT1.14* genes are generated by tandem duplication event, two of the *GmPHT2* members were formed by large segmental duplication events. *GmPHT 3.1* and *GmPHT3.5*, *GmPHT3.3* and *GmPHT3.4*, and *GmPHT3.9* in the *GmPHT3* subfamily have common ancestor genes with *GmPHT3.6* and *GmPHT3.7* which were generated by large segmental duplication events; ancestor genes of *GmPHT3.6* and *GmPHT3.7* genes Tandem duplications occurred. Both the *GmPHT4* and *GmPHT5* subfamilies expanded the gene family size with only large segmental dupli-

Name	Gene locus <sup>1</sup>	Gene identifier <sup>2</sup>	Chr	Locations	Length(bp)	Protein ID <sup>1</sup>	Length(aa)	pI	Mw(kDa)	GRAVY	TMHs	Predicted subcellular location	
												WoLF PSORT	Softberry
GmPHT1.1	LOC100780201	Glyma.02G005800	2	634,487..639193	1602	NP_001241164.1	533	8.31	58.5	0.309	12	vacu:7	Plasma membrane
GmPHT1.2	LOC100819445	Glyma.03G162800	3	37,772,793..37775073	1620	NP_001304639.2	539	8.67	59.3	0.291	11	plas:10	Plasma membrane
GmPHT1.3	LOC100797683	Glyma.07G222700	7	39,866,989..39873025	1536	NP_001241574.1	511	8.51	57.6	0.295	12	plas:11	Plasma membrane
GmPHT1.4	LOC100795623	Glyma.10G006700	10	662,561..665361	1602	NP_001341396.1	533	8.54	58.4	0.324	12	plas:7	Plasma membrane
GmPHT1.5	LOC100786638	Glyma.10G036800	10	3,233,051..3236286	1566	NP_001304588.2	521	8.63	57.3	0.363	11	plas:7	Plasma membrane
GmPHT1.6	LOC100802365	Glyma.10G186400	10	41,935,063..41939261	1584	NP_001239971.1	527	8.91	58.1	0.397	11	plas:9	Plasma membrane
GmPHT1.7	LOC100802890	Glyma.10G186500	10	41,946,209..41950274	1611	NP_001240032.1	536	8.34	58.7	0.334	11	plas:10	Plasma membrane
GmPHT1.8	LOC100802261	Glyma.13G040200	13	12,579,481..12581070	1590	NP_001241400.1	529	8.82	58.7	0.316	10	plas:12	Plasma membrane
GmPHT1.9	LOC100805284	Glyma.14G123500	14	18,527,015..18528592	1578	NP_001241127.1	525	8.33	58.2	0.330	11	plas:11	Plasma membrane
GmPHT1.10	LOC100820250	Glyma.14G188000	14	45,277,372..45278961	1590	NP_001239765.1	557	8.18	61.2	0.232	12	plas:12	Plasma membrane
GmPHT1.11	LOC100803626	Glyma.19G164300	19	42,521,537..42524774	1620	NP_001254802.1	539	8.52	59.3	0.291	11	plas:10	Plasma membrane
GmPHT1.12	LOC100803834	Glyma.20G021600	20	2,196,534..2205616	1521	NP_001345390.1	506	8.63	56.8	0.333	11	plas:10	Plasma membrane
GmPHT1.13	LOC100792177	Glyma.20G204000	20	44,094,584..44098582	1611	NP_001240239.1	536	7.63	58.6	0.341	12	plas:11	Plasma membrane
GmPHT1.14	LOC100792711	Glyma.20G204100	20	44,103,550..44107427	1584	NP_001240957.1	527	8.93	58.1	0.386	11	plas:9	Plasma membrane
GmPHT2.1	LOC100787472	Glyma.08G282100	8	38,707,723..38713794	1728	XP_003530627.1	575	9.28	60.7	0.519	12	plas:7	Plasma membrane
GmPHT2.2	LOC100787349	Glyma.18G144100	18	23,306,379..23312679	1731	XP_003551229.1	576	9.31	60.5	0.502	12	plas:7	Plasma membrane
GmPHT3.1	LOC100788561	Glyma.01G157100	1	49,479,697..49484208	921	XP_003517120.1	306	9.13	33.8	0.167	1	chlo:5	Mitochondrial
GmPHT3.2	-	Glyma.02G065800	2	5,845,126..5847769	873	-	347	8.69	37.8	0.105	0	vacu:5	Mitochondrial
GmPHT3.3	LOC100791467	Glyma.05G201400	5	38,507,610..38510859	1080	XP_003524479.2	359	9.50	38.3	0.135	0	chlo:9	Mitochondrial
GmPHT3.4	LOC100800530	Glyma.08G008900	8	693,228..696708	1068	XP_003532391.2	355	9.37	38.0	0.115	0	chlo:9	Mitochondrial
GmPHT3.5	LOC100788974	Glyma.11G087800	11	6,610,115..6614787	921	XP_003537697.1	306	9.22	33.7	0.186	1	nucl:5	Mitochondrial
GmPHT3.6	LOC547633	Glyma.16G050000	16	4,788,380..4801207	1119	NP_001235652.2	372	9.25	39.8	0.163	0	chlo:4	Mitochondrial
GmPHT3.7	LOC100801963	Glyma.16G050100	16	4,798,523..4800531	1020	XP_025981816.1	339	9.38	36.8	0.184	0	chlo:11	Mitochondrial
GmPHT3.8	LOC100804257	Glyma.16G146700	16	30,755,899..30759360	1050	XP_003548029.1	350	9.25	37.6	0.101	0	cyto:8	Mitochondrial
GmPHT3.9	LOC548006	Glyma.19G101100	19	34,833,393..34841869	1128	NP_001237304.1	375	9.35	39.8	0.193	0	chlo:4	Mitochondrial
GmPHT4.1	LOC100789297	Glyma.02G224200	2	41,175,313..41180091	1551	XP_003519242.1	516	10.01	55.7	0.479	12	plas:7	Membrane bound chloroplast
GmPHT4.2	LOC100782221	Glyma.03G008200	3	793,191..799880	1782	XP_003520968.1	593	9.36	65.5	0.220	10	plas:11	Membrane bound chloroplast
GmPHT4.3	LOC100793618	Glyma.07G069600	7	6,307,073..6313498	1779	XP_003528848.1	592	9.37	65.2	0.218	10	plas:10	Membrane bound chloroplast
GmPHT4.4	LOC100819182	Glyma.07G144700	7	17,231,556..17235971	1515	XP_003529140.1	504	9.88	55.0	0.509	11	chlo:6	Membrane bound chloroplast
GmPHT4.5	LOC100812608	Glyma.07G274200	7	44,592,193..44600634	1290	NP_001239989.1	429	9.58	46.6	0.659	10	plas:8	Membrane bound Golgi
GmPHT4.6	LOC100787020	Glyma.11G175800	11	20,331,840..20340143	1554	XP_006591210.1	517	9.41	57.3	0.223	10	plas:9	Membrane bound chloroplast
GmPHT4.7	LOC100806856	Glyma.13G162900	13	27,823,310..27833361	1587	XP_003541488.1	525	7.16	57.4	0.339	7	chlo:7	Membrane bound chloroplast
GmPHT4.8	LOC100785142	Glyma.14G190900	14	45,569,552..45574890	1548	XP_003544870.1	515	9.93	55.8	0.540	11	plas:7	Membrane bound chloroplast
GmPHT4.9	LOC100809973	Glyma.17G000300	17	17,407..211108	1290	XP_003550165.1	429	9.68	46.5	0.624	10	plas:8	Membrane bound Golgi
GmPHT4.10	LOC100792616	Glyma.17G108300	17	8,486,013..8495685	1575	XP_003549724.1	524	6.83	57.5	0.330	7	plas:7	Membrane bound chloroplast
GmPHT4.11	LOC100792104	Glyma.18G066000	18	6,041,786..6050082	1476	XP_003551308.1	491	9.09	53.8	0.411	10	plas:6	Membrane bound chloroplast
GmPHT4.12	LOC100784401	Glyma.20G002000	20	213,184..219283	1794	XP_003556597.1	597	9.12	65.7	0.295	10	plas:10	Membrane bound chloroplast
GmPHT5.1	LOC100808181	Glyma.09G128500	9	32,046,109..32052575	2088	XP_003533972.1	695	6.17	78.2	0.214	11	plas:10	Membrane bound vacuolar

Continued

Name	Gene locus <sup>1</sup>	Gene identifier <sup>2</sup>	Chr	Locations	Length(bp)	Protein ID <sup>1</sup>	Length(aa)	pI	Mw(kDa)	GRAVY	TMHs	Predicted subcellular location	
												WoLF PSORT	Softberry
<i>GmPHT5.2</i>	LOC100777530	Glyma.09G263400	9	48,091,634..48104775	2094	XP_003534584.1	697	6.14	78.6	0.209	10	plas:11	Membrane bound vacuolar
<i>GmPHT5.3</i>	LOC100777032	Glyma.10G229900	10	46,001,071..46006711	2094	XP_014618833.1	697	6.22	78.7	0.170	11	plas:11	Membrane bound vacuolar
<i>GmPHT5.4</i>	LOC100778938	Glyma.16G176100	16	33,736,100..33743654	2088	XP_003548146.1	695	6.36	78.2	0.200	11	plas:10	Membrane bound vacuolar
<i>GmPHT5.5</i>	LOC100793709	Glyma.18G228700	18	51,774,933..51789151	2094	XP_006602776.1	697	6.55	78.7	0.232	11	plas:12	Membrane bound vacuolar
<i>GmPHT5.6</i>	LOC100805436	Glyma.20G163400	20	40,093,249..40098816	2094	XP_003556130.1	697	6.52	78.9	0.158	11	plas:11	Membrane bound vacuolar
<i>GmPHO1.1</i>	LOC100786083	Glyma.01G091800	1	27,867,550..27874181	2376	XP_003516868.1	791	9.14	92.2	-0.212	6	plas:8	Plasma membrane
<i>GmPHO1.2</i>	LOC100796500	Glyma.02G003700	2	431,852..445384	2292	XP_006574509.1	763	9.24	88.4	-0.106	6	plas:10	Plasma membrane
<i>GmPHO1.3</i>	LOC100795093	Glyma.02G110600	2	10,670,899..10676578	2256	XP_006574913.1	751	8.85	88.5	-0.316	4	plas:8.5	Plasma membrane
<i>GmPHO1.4</i>	LOC100787164	Glyma.02G130200	2	13,340,016..13346701	2370	XP_003518826.1	789	9.12	92.0	-0.219	6	plas:8	Plasma membrane
<i>GmPHO1.5</i>	LOC100819185	Glyma.07G228400	7	40,691,412..40698177	2412	XP_006583960.1	803	9.45	93.4	-0.183	9	plas:12	Plasma membrane
<i>GmPHO1.6</i>	LOC100780548	Glyma.09G235200	9	45,765,397..45771516	2280	XP_006587738.1	759	9.28	87.7	-0.103	6	plas:11	Plasma membrane
<i>GmPHO1.7</i>	LOC100781658	Glyma.10G004800	10	457,173..468529	2295	XP_006588543.1	764	9.41	88.4	-0.111	5	plas:9	Plasma membrane
<i>GmPHO1.8</i>	LOC100794580	Glyma.10G183300	10	41,633,437..41645279	2325	XP_006589283.1	774	9.06	90.1	-0.117	6	plas:10	Plasma membrane
<i>GmPHO1.9</i>	LOC100795104	Glyma.18G261900	18	54,775,710..54781535	2331	XP_003552542.1	776	9.24	89.3	-0.114	6	plas:9	Plasma membrane
<i>GmPHO1.10</i>	LOC100805979	Glyma.20G031700	20	3,929,374..3936931	2409	XP_006605562.1	802	9.27	93.0	-0.163	8	plas:12	Plasma membrane
<i>GmPHO1.11</i>	LOC100811310	Glyma.20G032400	20	4,183,317..4191520	2388	XP_003556776.1	795	9.38	92.8	-0.233	8	plas:12	Plasma membrane
<i>GmPHO1.12</i>	LOC100813281	Glyma.20G032500	20	4,194,291..4202575	2391	XP_003556778.1	796	9.32	92.6	-0.195	8	plas:11	Plasma membrane
<i>GmPHO1.13</i>	LOC100813818	Glyma.20G032600	20	4,204,879..4212818	2361	XP_003556779.1	786	9.24	91.7	-0.121	8	plas:12	Plasma membrane
<i>GmPHO1.14</i>	LOC100802251	Glyma.20G206900	20	44,375,656..44385721	2325	XP_006606372.1	774	9.14	89.9	-0.122	8	plas:9	Plasma membrane

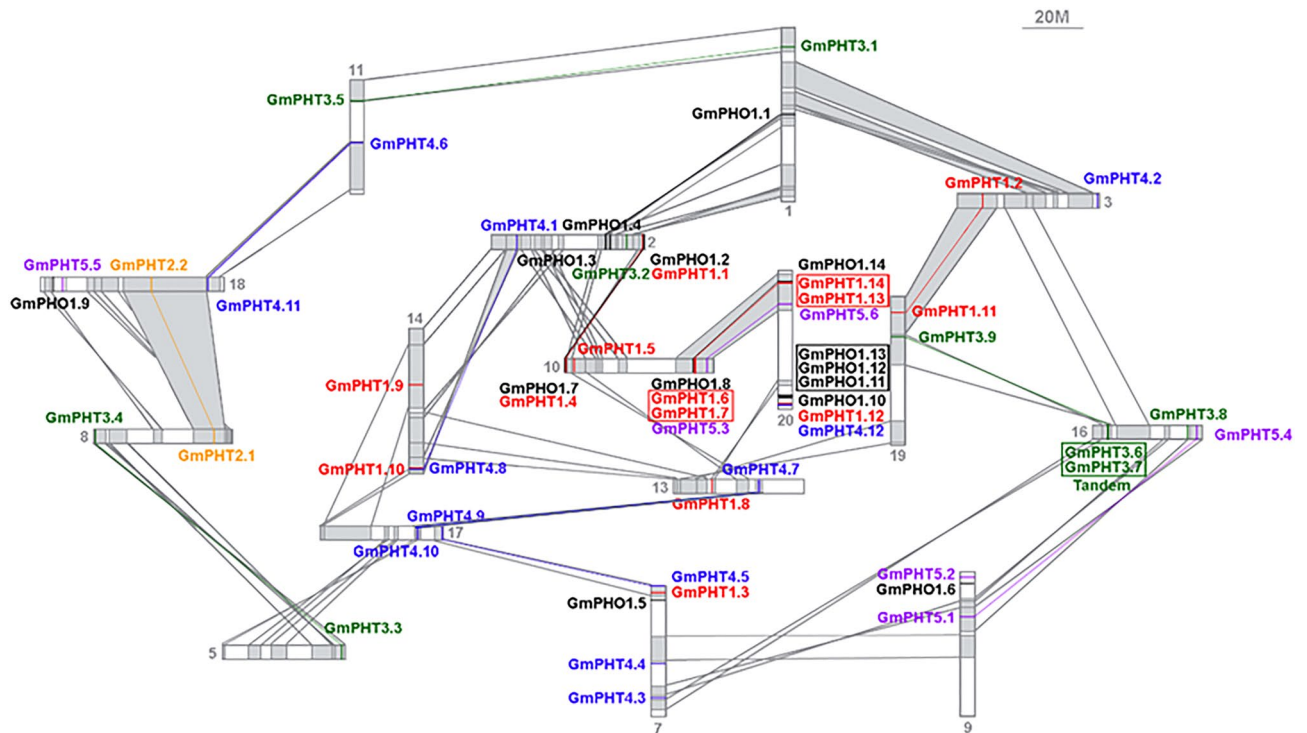
**Table 1.** Localization and physicochemical properties of members of the *GmPHTs* protein gene family. Indicates that the corresponding PHTs gene was not found in the NCBI soybean database. 1 represents the number of the *GmPHT* gene in the NCBI soybean database. 2 represents the number of the *GmPHT* gene in the Phytozome soybean database.

cation events (*GmPHT4.1* and *GmPHT4.8*, *GmPHT4.5* and *GmPHT4.9*, *GmPHT4.6* and *GmPHT4.11* genes; *GmPHT5.1* and *GmPHT5.4*, *GmPHT5.3* and *GmPHT5.6* genes). of the *GmPHO1* subfamily, *GmPHO1.1* and *GmPHO1.4*, *GmPHO1.2* and *GmPHO1.7*, *GmPHO1.8* and *GmPHO1.14* genes were generated by large segmental duplication events, and *GmPHO1.11*, *GmPHO1.12* and *GmPHO1.13* genes were formed by tandem duplication event. It suggested that large segmental duplication also played a major driving force for *GmPHTs* evolution in addition to tandem duplication.

**Gene structure and protein-specific motif analysis of the *GmPHTs* gene.** From the perspective of gene structure, the *GmPHO1* subfamily can be divided into 2 Groups, and Group I contain 8 genes including *GmPHO1.3*, *GmPHO1.5*, *GmPHO1.6*, *GmPHO1.9*, *GmPHO1.10*, *GmPHO1.11*, *GmPHO1.12* and *GmPHO1.13*. The three pairs of *GmPHO1* genes form Group II, which differs from the Group I genes mainly in the length of the first exon. Members of the *GmPHT3* subfamily have 5 or 6 exons, but the difference is mainly in whether the last exon is split or not. *GmPHT4.1* and *GmPHT4.8* both contain 8 exons, *GmPHT4.5* and *GmPHT4.9* genes contain only 1 exon, and *GmPHT4.7* and *GmPHT4.10* genes contain a maximum of 15 exons. The 14 members of the *GmPHT1* subfamily can also be divided into two categories based on the results of gene structure, with the exception of *GmPHT1.1* and *GmPHT1.4*, which have two exons, the other 10 genes in Group III have only one exon, and *GmPHT1.3* and *GmPHT1.12*, which have similar structures and three exons. *GmPHO1* subfamily has a maximum of 10 specific motifs and the number of motifs is the same for each member of the *GmPHT3* and *GmPHT5* subfamilies. The protein sequences of the *GmPHT1* subfamily members contain 7–9 specific motifs, notably, the four gene pairs generated by large block repeats or tandem repeats and the *GmPHT1.5* gene have exactly the same 9 motifs, whereas the protein encoded by the five genes generated by non-repetitive events has only 7 motifs except for 8 and 9. Only the number of specific motifs varied significantly among members of the *GmPHT4* subfamily, ranging from 1 to 4 (Fig. 3).







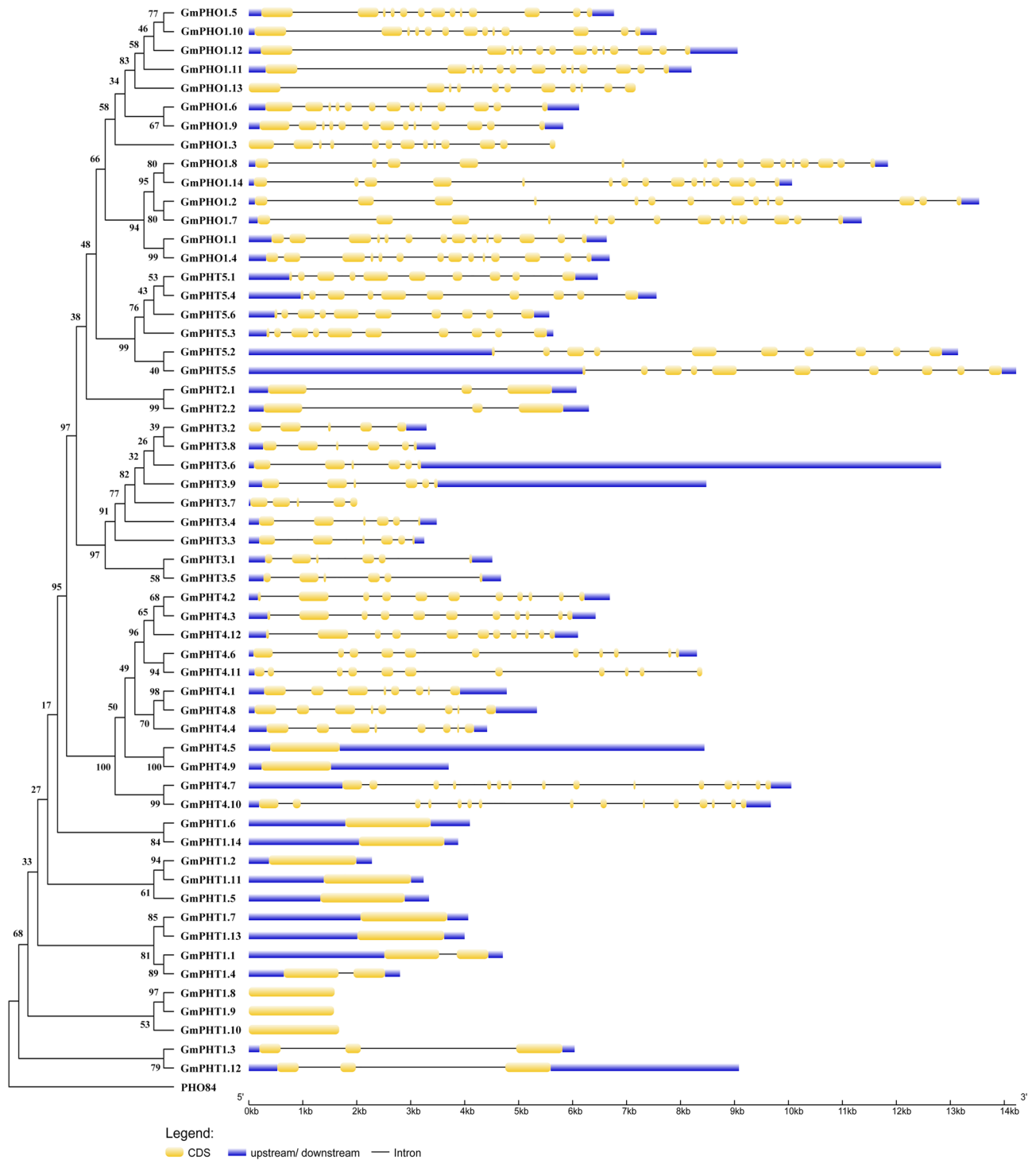
**Figure 2.** *GmPHTs* collinearity analysis results. *Note:* Members of the same subfamily are represented by the same color. Solid lines between genes indicate collinearity between the two. The genes enclosed in boxes are formed in the form of tandem repeats. The covariance data between chromosomes are cited Liu, et al. 2015<sup>38</sup>.

mainly responds to adverse environmental conditions such as heat, low temperature, anaerobic, drought, and low phosphorus. The phytohormone response elements mainly include abscisic acid, methyl jasmonate, gibberellin, ethylene, salicylic acid, and growth hormone response elements. Except for the *GmPHT3.1* gene, the promoters of all genes had phytohormone-like response elements, differing only in type and number. Biological stress response elements were the least, accounting for only about 5%, and none of the members of the *GmPHT* subfamily had biological stress response elements. It is thus hypothesized that each member of the *GmPHT* genes family has an important role in the growth and development of soybean and resistance to various external growth-adverse stimuli. The basic biological function of PHT proteins has been reported to be the transport of inorganic phosphorus, and the application of phosphorus has a mitigating effect on drought<sup>39</sup>, so among the promoter elements, the low phosphorus response element PIBS and the drought-inducing element MBS are the focus of our attention. Only 13 of the 57 genes have promoter regions containing both the low phosphorus response element PIBS and the drought-inducing element MBS. Therefore, further expression analysis was performed for 13 genes, including *GmPHT1.1*, *GmPHT1.7*, *GmPHT1.10*, *GmPHT1.14*, *GmPHT3.4*, *GmPHT3.5*, *GmPHT3.6*, *GmPHT4.7*, *GmPHT4.8*, *GmPHT4.10*, *GmPHO1.4*, *GmPHO1.5* and *GmPHO1.7*.

### Differential expression profiles of the *GmPHTs* genes under low phosphorus and drought stress.

So far, little is known about the specificity of tissues or organs of the soybean PHT genes, which may elucidate their functions in detail, because the *GmPHTs* family members in soybean have not been systemically examined. By the results of the preliminary qRT-PCR experiments (Data not shown) and the gene microarray data of soybean, *GmPHT1.10*, *GmPHT1.14*, *GmPHT3.4* and *GmPHT3.6* genes were not expressed among the 13 genes. Therefore, only the remaining nine genes were examined separately for their expression under simultaneous low phosphorus drought treatment conditions (Fig. 5). *GmPHT1.1* gene was highly expressed in roots only, *GmPHT3.5* gene was expressed in leaves only, *GmPHT1.7*, *GmPHT4.7*, *GmPHT4.8*, *GmPHT4.10*, *GmPHO1.4*, *GmPHO1.5* and *GmPHO1.7* genes were expressed in both roots and leaves. The *GmPHT1.7* and *GmPHT4.8* genes were expressed in both leaves and roots, and it was inferred that the *GmPHT1.7* and *GmPHT4.8* genes may play important roles in both roots and leaves, and the *GmPHT1.1* gene acts synergistically in the roots during the seedling stage of soybean. In contrast, *GmPHT4.10*, *GmPHO1.4*, *GmPHO1.5*, and *GmPHO1.7* genes were expressed in the roots but at low levels, and these three genes may not play a major role in the roots.

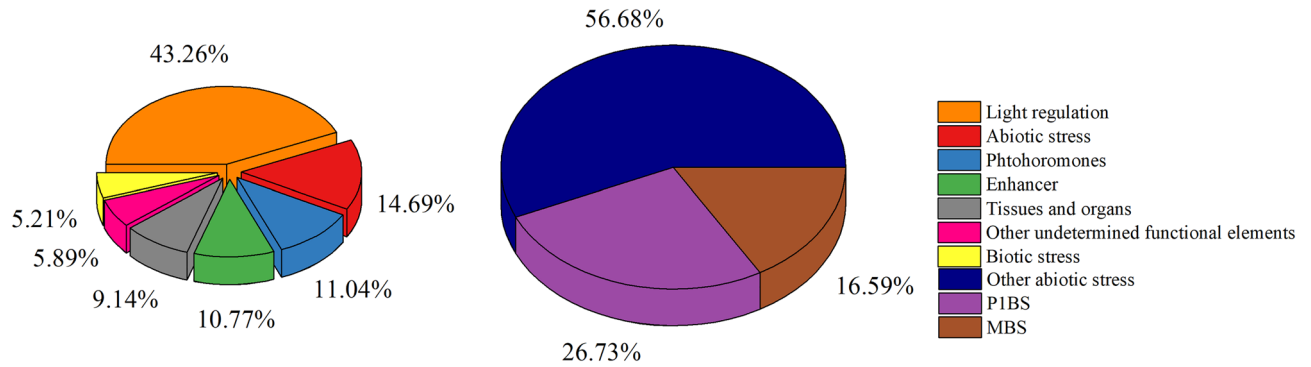
The expression pattern of these nine genes changed again when soybean was treated with both low phosphorus and drought for 12 h (Fig. 6). *GmPHT3.5*, *GmPHT4.7*, *GmPHT4.10*, *GmPHO1.5*, *GmPHO1.7*, and *GmPHT1.7* was significantly reduced. The expression of the *GmPHT1.1* gene in the roots increased gradually with time, and its expression increased significantly with low phosphorus and drought treatment for 12 h. The expression pattern of *GmPHO1.4* gene in leaves was similar to that of *GmPHT1.1* gene, both of which decreased first and then showed a significant increase with time. Only the expression of *GmPHT4.8* gene in leaves was not affected



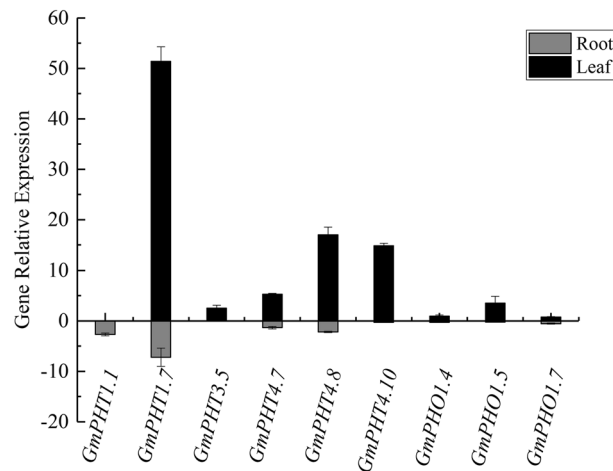
**Figure 3.** Analysis of the gene structure and protein-specific motifs of GmPHTs.

by low phosphorus and drought, and the expression was basically the same as that of the control, but the expression in roots was significantly lower.

**Coexpression networks of the PHTs family genes in soybean.** To further unravel the coexpression relationships between PHTs family genes and other genes, we calculated the interaction weight values of the target gene sets based on the FPKM values from the RNA-seq data. Figure 7 shows all the genes interacting with *GmPHTs* genes. In general, among the interacting genes, *GmPHT1:1* interacted with 31 genes, while *GmPHO1:4* interacted with 57 genes. Three genes interacted most strongly with *GmPHT1:1*, and six genes were co-expressed with *GmPHO1:4*. No other genes were found that interacted strongly with other *PHTs* genes under low phosphorus stress (Fig. 7). These include ubiquitin family genes, and ubiquitin-conjugating enzyme genes.



**Figure 4.** The promoter element composition of *GmPHTs* gene.



**Figure 5.** qRT-PCR results of *GmPHT* genes in soybean roots and leaves.

Ubiquitination modifications play a central regulatory role in the plant response to low phosphorus stress and may be involved in regulating the uptake of phosphorus.

## Discussion

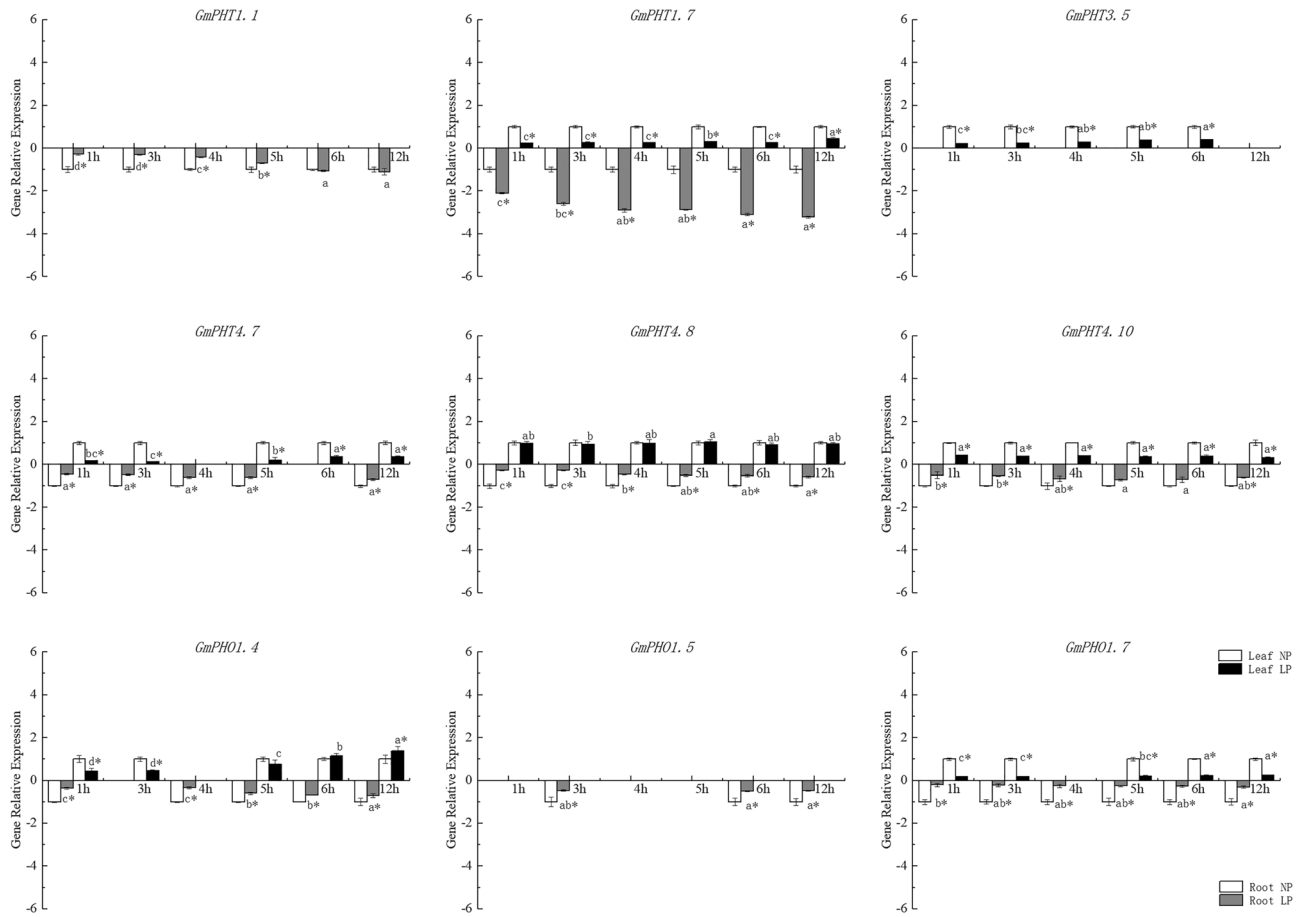
As an essential, nonsubstitutable element for plant growth, P plays a wide range of structural and biological roles<sup>40</sup>. P deficiency influences numerous biological processes in crop plants, such as rice, soybean and wheat, markedly reducing crop yield<sup>41–44</sup>. Plants need to obtain the phosphorus they need from the soil, but as the phosphorus concentration in the soil is lower than that in plant tissue, this sharp concentration gradient between the plant and the soil indicates the indispensable roles of PHTs, which can regulate Pi absorption<sup>45</sup>.

Pi transporters currently identified in *Arabidopsis*, rice, soybean, maize, wheat, poplar, *Medicago truncatula*, and the *Halophyte E. salsguineum*<sup>46</sup>. The concerted action of Pi transporters ensures Pi acquisition and distribution among tissues and cytosolic Pi homeostasis<sup>44</sup>. Therefore, engineered alterations of the expression of Pi transporters provide an opportunity to optimize uptake and distribution of Pi in crops to improve yield<sup>47</sup>. A systematic analysis of the whole PHT gene family in soybean was conducted in this study, which will provide an opportunity to find candidate genes that play roles in phosphate use efficiency and abiotic stress response. A total of 57 PHT genes are identified in soybean and show quite high similarity to their corresponding members of the PHT gene family in *Arabidopsis* (Fig. 1). This implies that these PHT genes are evolutionarily conserved and functionally similar in soybean and *Arabidopsis*.

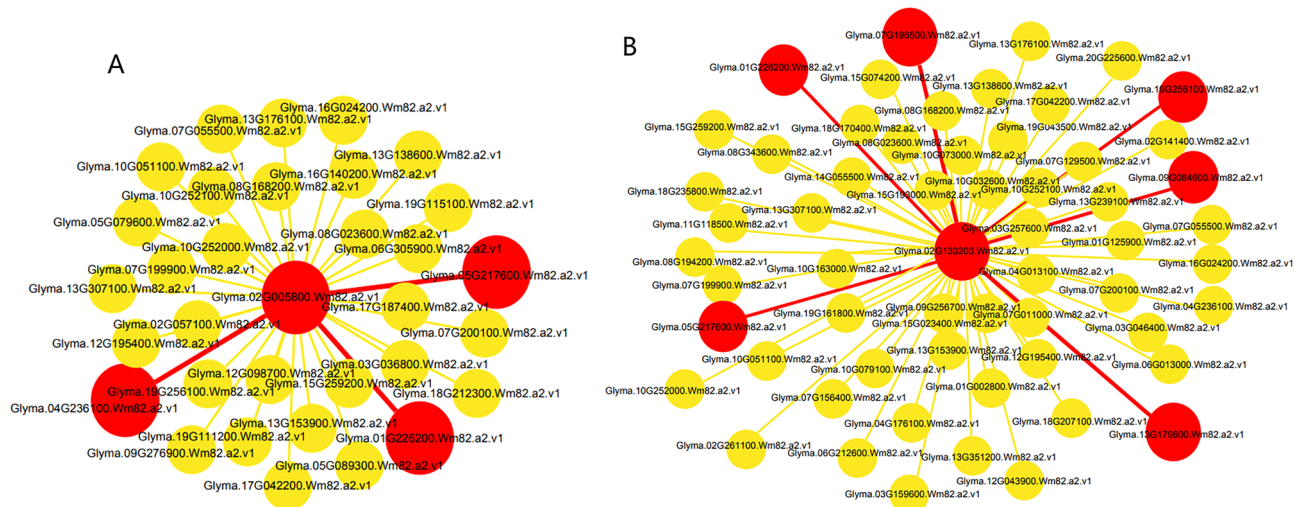
Gene duplication is considered one of the primary driving forces in the evolution of genomes and genetic systems. Of all the duplication patterns, segmental and tandem duplications have been purported to be the two main causes of gene family expansion in plants<sup>48</sup>. In the current research, 29 of the 57 *GmPHTs* (87.8%) in the soybean genome were associated with large segmental duplication events. Moreover, only 7 tandem duplication events were identified (Fig. 2). Taken together, our results indicate that large blockbuster duplication and tandem duplication events are the main force for the expansion of the PHT gene family in soybean. A relatively large number of members within a family suggests successful expansion and rearrangement of the genome by extensive duplication that occurred frequently during evolution<sup>49</sup>.

Cis elements may control the efficiency of promoters and thus regulate the expression of the genes that they control by interacting with the corresponding trans-regulatory factors<sup>50</sup>. Studies of cis elements might play a





**Figure 6.** qRT-PCR results of nine GmPHT genes in roots and leaves of soybean treated with low phosphorus as well as drought.



**Figure 7.** Coexpression networks of 2 PHTs family genes in soybean. (A), Glyma02G005800 (GmPHT1.1) coexpression network diagram; (B), Glyma02G130200 (GmPHO1.4) coexpression network diagram; The size of the node is proportional to the degree of the node, the more edges connected to the node, the greater its degree and the larger the node; the color indicates the strength of the interaction, the redder the color the stronger the interaction. The PHTs family genes located in the center of the network, while the most coexpressed genes were displayed in each network.

crucial role in dissecting the functions of genes. But the cis elements of all PHT gene family members in soybean have not yet been identified. In this study, the upstream regions of all the GmPHTs contain cis elements affecting light, hormone, defense and tissue-specific expression. These findings suggest that the expression of these *GmPHT* genes might be regulated by many factors, such as light-, hormone-, and defense-related factors. The presence of these cis elements in soybean *GmPHT* genes and their likely roles in regulating gene expression suggest that *GmPHT* genes may be involved in various stress responses in soybean. Some cis elements function under phosphate deficiency, for example, the P1BS element can be bound to PHR1 (a MYB transcription factor) to potentially regulate the phosphate starvation response in *Arabidopsis*<sup>51</sup>, PHR1 has been suggested as a central integrator to role in the transcriptional regulation of phosphate starvation responses<sup>52</sup>. In the current study, P1BS-like elements were found in 13 of the 57 GmPHT promoters examined suggesting that expression of these genes might be influenced by Pi starvation. Similarly, P1BS-like elements have also been detected in *OsPHT* genes in rice<sup>50</sup>. These elements are found to be in the promoters of phosphate regulated genes, not only in *Arabidopsis* but also in other plant species such as barley and rice<sup>50,53,54</sup>, which suggests that they may participate in a conserved signaling pathway for the phosphate starvation response in plants. The existence of P1BS-like elements in soybean, rice, and *Arabidopsis* suggests that they may serve a similar role in different species.

Drought and Pi deficiency stress often interfere with plant metabolism via various stress signals and hormonal changes that play essential roles in regulation processes under natural conditions<sup>55</sup>. So far, several studies have found that Pi transporters are involved in response to drought stress. Recently, several genes encoding Pi transporters from poplar were found to be regulated by drought stress, these genes especially those up-regulated by drought stress at low Pi level may contribute to drought tolerance of poplar plants in Pi-limited soils<sup>11</sup>. Under drought stress, 18 *PHT* genes were upregulated in the leaves, and 11 *PHT* genes were downregulated in *Brassica napus* roots<sup>56</sup>. The expression of 15 *MdPHT* genes was upregulated in apple under drought stress<sup>57</sup>. Cao et al. (2020) found that overexpression of *StPHT1;7* in potato influenced plant growth and tolerance to drought stress<sup>58</sup>. Therefore, it is urgent that researchers develop plants with enhanced efficiency of soil phosphorus use under such conditions. Secondly, using drought inducible promoters to control the expression of Pi transporters or their regulatory factors may be a potential strategy to improve crop PUE under drought stress conditions. Drought can prevent Pi uptake by reducing Pi distribution in the root system<sup>59</sup>. Pi can also increase tissue osmotic pressure and enhance tolerance to drought stress by adjusting the water content in the cell structure and soluble sugar content in the cells<sup>60</sup>. Plants fertilised with Pi show moderate growth under drought compared to unfertilised plants<sup>61</sup>. In addition, phosphorus fertilizers are often used to reduce Pi deficiency in the soil, improve drought tolerance and promote plant growth<sup>62</sup>. In *Alnus cremastogyne*, phosphorus fertilisation alleviates drought stress by regulating its antioxidant and osmotic potential<sup>63</sup>. Several recent studies have highlighted the role of phosphorus in alleviating drought stress in plants by promoting growth, especially of seedlings<sup>64</sup>. Plants primarily use phosphate transporters (PHT) located in the cell membranes of plant roots to take up Pi from the soil, which is then transported into the plant<sup>65</sup>. Increased phosphorus uptake enhances tolerance to low phosphorus<sup>66,67</sup> and drought stress<sup>62,68</sup>. Therefore, engineered alterations of the expression of Pi transporters provide an opportunity to optimize uptake and distribution of Pi in crops to improve yield.

The number of Pi transporters present may reflect the complexity and significance of the process of Pi transport<sup>5</sup>. Most PHT1 family genes expressed in roots show up-regulation in phosphate deprived *Arabidopsis* plants; In rice, high-affinity of *OsPHT9* and *OsPHT10* for Pi transport have been reported, as well as specific induction of their expression by Pi starvation<sup>69</sup>. The transcript levels of a few PHT genes in *Arabidopsis* were increased by Pi deficiency<sup>70</sup>; however, the expression of most PHO family members was not induced by Pi deficiency in poplar, suggesting that PHT genes might have different functions in woody species and herbs. In addition, some PHT4 family members might be low-affinity transporters, consistent with the relatively high Pi levels that have been reported in the cytosol and other subcellular compartments<sup>7</sup>. In this study, the expression patterns of *GmPHT* genes were analyzed under low phosphorus, drought and simultaneous low phosphorus drought stress conditions, and the results showed that the expression patterns of the genes differed under different treatment conditions. These results suggest that of crosstalk occurs among the P-starvation response and salt and drought stress responses, and Pi uptake in plants likely changes in association with the altered expression of the PHT1 genes under drought and salt stresses<sup>47</sup>. Our results show that under simultaneous drought and low phosphorus treatments, the expression of *GmPHT1.7* and *GmPHT1.1* genes in roots significantly increased, while the expression of *GmPHO1.4* gene in leaves showed a significant increase with the continuation of treatment time, and the expression of this gene in roots significantly decreased. The expression of *GmPHT4.8* gene in leaves was not affected by low phosphorus and drought, but the expression in roots was significantly reduced (Fig. 6). The identification of these genes is therefore of great significance for the development of transgenic crops to cope with low phosphorus and drought stress.

## Conclusions

Plant Pi transporters play an essential role in Pi acquisition and distribution. The family of phosphate transporters (PHTs) mediates the uptake and translocation of Pi inside the plants. In this study, a total of 57 PHT-domain-containing protein genes were identified in soybean genome, (*GmPHT1*, *GmPHT2*, *GmPHT3*, *GmPHT4*, *GmPHT5* and *GmPHO1*) had been found. Only 13 of the 57 genes have promoter regions containing both the low phosphorus response element P1BS and the drought-inducing element MBS, 4 genes were not expressed among the 13 genes. Therefore, only the remaining nine genes were examined separately for their expression under simultaneous low phosphorus drought treatment conditions. *GmPHT1.7*, *GmPHT1.1*, *GmPHO1.4* and *GmPHT4.8* genes displayed various expression patterns under low phosphorus and drought stress. The above results implied that they may play different roles in phosphate nutrition of Soybean different tissues and development stages. These results provide references for the further study of Soybean PHT-domain-containing proteins

family genes. Moreover, this study provides the selection of candidate genes for functional research and genome editing in Soybean phosphate nutrition.

## Materials and methods

**Identification of the GmPHT family genes in soybean.** The Arabidopsis (*Arabidopsis thaliana*) PHT protein sequence was used for TblastN homology alignment in the soybean genome database (Phytozome V12.1) to obtain sequences similar to members of the *AtPHT* gene family. The obtained sequences were then analyzed in the NCBI CDD database (v3.15) for conserved structural domains to identify whether the obtained similar sequences were GmPHTs proteins. The sequence of the protein finally identified as GmPHTs was sequence corrected in the NCBI soybean database. Data on the isoelectric point, molecular weight, and hydrophilicity of the proteins were obtained using the ProtParam tool in the ExPASy (<https://web.expasy.org/>) online analysis website.

**Phylogenetic, Multiple alignment and protein motif analysis.** The PHT protein sequences of *Arabidopsis* and soybean were selected for multiple sequence alignment using the ClustalW program in BioEdit (v7.0.5) software. The PHT protein sequences of yeast were used as outgroup to construct a phylogenetic tree by MEGA (v7.0.26) software. The method used was the Jones-Taylor-Thornton (JTT) model, and Bootstrap was 1000 replicates.

The PLAZA website (<https://bioinformatics.psb.ugent.be/plaza>) was used to analyze the collinearity of the soybean genome and finally to obtain the regions where large segments of duplication occurred in the soybean genome.

The number of introns and exons of genes was analyzed and counted using the online analysis website GSDS (v2.0) (<http://gsds.cbi.pku.edu.cn>). The unique or shared motifs in the proteins were done using the MEME line analysis website (<http://meme-suite.org/>).

**Promoter element analysis in GmPHT genes.** The gene sequence was obtained from the soybean genome database phytozome (V12.1) and the 2000 bp sequence upstream of the gene was used as the promoter region. The promoter element prediction was performed for all *GmPHT* members using the promoter online prediction website PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). The promoter elements other than the basic promoter elements (TATA box and CAAT box) were mapped using the GSDS online mapping tool for promoter elements.

**Plant materials and growth condition.** The experiments were conducted with the main cultivar Changnong 26 in Jilin Province (Validation No.: JI Audited Bean 2,010,004). To identify tissue-preferentially expressed genes, seedlings of Chang Nong 26 were grown under normal conditions and cultured in sand culture at 25 °C with 50% humidity and 13 h of light and 9 h of darkness, using 1/4 Hoagland (containing 0.5 mM KH<sub>2</sub>PO<sub>4</sub>) as the culture medium to provide essential nutrients for the soybean. The plants were converted to hydroponic culture when they reached the three-leaf stage, and the culture medium was still 1/4 Hoagland containing 0.5 mM KH<sub>2</sub>PO<sub>4</sub>. After 12 h of hydroponic culture, the plants were transferred to 1/4 Hoagland culture with 0.01 mM KH<sub>2</sub>PO<sub>4</sub>, 1/4 Hoagland culture with 2% PEG8000 (containing 0.5 mM KH<sub>2</sub>PO<sub>4</sub>) and 1/4 Hoagland culture with 2% PEG8000 containing 0.01 mM KH<sub>2</sub>PO<sub>4</sub> for Hydroponic culture. The plants were cultivated in a glasshouse (natural light, 75% relative humidity) and irrigated with 50 ml nutrient solution. Three treatments were set up with three replications, continuous normal phosphorus treatment, continuous low phosphorus, and stressed low phosphorus treatment. Samples were taken on time for 12 h in hydroponics, and the roots and leaves were stored separately at -80 °C for backup.

**Coexpression networks of the GmPHT family genes using RNA-seq data.** Gene coexpression network analysis was performed based on the RNA-seq data. For the RNA-seq experiment, fully expanded leaves and roots were taken when the plants were treated with low phosphorus nutrient solution until to 3 leaves separately for RNA extraction with three biological replicates. The transcript abundance (FPKM value) of each gene was calculated based on the length of the gene and the reads mapped to that gene. The interactions of the target gene sets were retrieved from the STRING protein database (<http://string-db.org/>), and the weight value of the target gene sets was calculated using the WGCNA R package based on the FPKM values. The gene coexpression networks were visualized by Cytoscape software<sup>71</sup>.

**Quantitative real-time PCR analysis.** RNA was extracted using Trizol (TAKARA, Beijing) and reverse transcribed using PrimeScript™ RT reagent Kit with gDNA Eraser kit (TAKARA, Beijing). Fluorescent quantitative PCR primers were designed according to the specific sequences of each gene, see Table S1. *Letm1* was used as the internal reference gene, and the fluorescent quantitative PCR was performed using the TAKARA SYBR Premix Ex Taq II kit (TAKARA, Beijing, China) with an ABI StepOne Plus real-time fluorescent quantitative PCR instrument (ABI, USA, model 7300). 25 µl of the reaction system contained 1 µl cDNA, 10 µl SYBR. The PCR program was: 95 °C for 10 min (pre-denaturation), 95 °C for 3 s (denaturation) and 60 °C for 1 min (annealing-extension) for 40 cycles. The lysis curve procedure was: 95 °C for 15 s, 60 °C for 1 min, and 95 °C for 15 s. Data were calculated using the 2<sup>-ΔΔCt</sup> algorithm for relative expression<sup>72</sup>. All experiments were performed in three replicates. Statistix (v 8.1) software was used to analyze the significance of the data.

**Statistical analysis of data.** Statistics was performed by Duncan's test or Student's t test. Significance of differences was defined as \* $P < 0.05$ , \*\* $P < 0.01$ .

### Data availability

The original contributions presented in the study are included in the article/Supplementary Material. The Supplementary Material for this article can be found online at: <https://www.ncbi.nlm.nih.gov/sra/PRJNA871448>. Accession number: PRJNA871448, NCBI SRA database.

Received: 15 August 2022; Accepted: 29 November 2022

Published online: 03 December 2022

### References

- Roch, G. V., Maharajan, T., Ceasar, S. A. & Ignacimuthu, S. The role of PHT1 family transporters in the acquisition and redistribution of phosphorus in plants. *CRC Crit. Rev. Plant Sci.* **38**, 171–198 (2019).
- Razaq, M., Zhang, P., Shen, H. & Salahuddin, S. Influence of nitrogen and phosphorus on the growth and root morphology of Acer mono. *PLoS ONE* **12**, e0171321 (2017).
- Hui, Y. & Dong, L. Signaling components involved in plant responses to phosphate starvation. *J. Chin. Bot. (Engl. Vers.)* **50**(7), 849–859 (2010).
- Raghothama, K. G. Phosphate transport and signaling. *Curr. Opin. Plant Biol.* **3**(3), 182–187 (2000).
- Tingting, S., Mingjun, Li., Yun, S., Lingyan, Yu. & Fengwang, Ma. Comprehensive genomic identification and expression analysis of the phosphate transporter (PHT) gene family in apple. *J. Front. Plant Sci.* **8**, 426 (2017).
- Schachtman, D. P., Reid, R. J. & Ayling, S. Phosphorus uptake by plants: From soil to cell. *Plant Physiol.* **116**, 447–453 (1998).
- Mimura, T. Regulation of phosphate transport and homeostasis in plant cells. *Int. Rev. Cytol.* **191**, 149–200 (1999).
- Rausch, C. & Bucher, M. Molecular mechanisms of phosphate transport in plants. *Planta* **216**, 23–37 (2002).
- Guo, B. *et al.* Functional analysis of the Arabidopsis PHT4 family of intracellular phosphate transporters. *New Phytol.* **177**, 889–898 (2008).
- Liu, T. Y. *et al.* Identification of plant vacuolar transporters mediating phosphate storage. *Nat. Commun.* **7**, 11095 (2016).
- Zhang, C., Meng, S., Li, M. & Zhao, Z. Genomic identification and expression analysis of the phosphate transporter gene family in poplar. *Front. Plant Sci.* **7**, 1398 (2016).
- Ceasar, S. A., Baker, A. & Ignacimuthu, S. Functional characterization of the PHT1 family transporters of foxtail millet with a novel Agrobacterium-mediated transformation procedure. *J. Sci. Rep.* <https://doi.org/10.1038/s41598-017-14447-0> (2017).
- Li, R. L., Wang, J. L., Xu, L., Sun, M. H. & Zhao, H. Y. Functional analysis of phosphate transporter OsPHT4 family members in rice. *J. Rice Sci.* **27**(06), 52–62 (2020).
- Muchhal, U. S., Pardo, J. M. & Raghothama, K. G. Phosphate transporters from the higher plant Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10519–10523 (1996).
- Lopez-Arredondo, D. L., Leyva-González, M. A., González-Morales, S. I., López-Bucio, J. & Herrera-Estrella, L. Phosphate nutrition: Improving low-phosphate tolerance in crops. *Plant Biol.* **65**, 95–123 (2014).
- Mudge, S. R., Rae, A. L., Diatloff, E. & Smith, F. W. Expression analysis suggests novel roles for members of the Pht1 family of phosphate transporters in Arabidopsis. *Plant J.* **31**, 341–353 (2002).
- Wang, H. *et al.* Arabidopsis WRKY45 transcription factor activates PHOSPHATE TRANSPORTER1; 1 expression in response to phosphate starvation. *Plant Physiol.* **164**, 2020–2029 (2014).
- Shi, S. L. *et al.* Function of phosphate transporter OsPHT2; 1 in improving phosphate utilization in rice. *Chin. J. Rice Sci.* **27**(5), 457–465 (2013) ((in Chinese with English abstract)).
- Kiiskinen, M., Korhonen, M. & Kangasjärvi, J. Isolation and characterization of cDNA for a plant mitochondrial phosphate transporter (Mpt1): Ozone stress induces Mpt1 mRNA accumulation in birch (*Betula pendula* Roth). *J. Plant Mol. Biol.* **35**(3), 271–277 (1997).
- Nakamori, K. *et al.* Cloning, functional expression, and mutational analysis of a cDNA for lotus Japonicus mitochondrial phosphate transporter. *J. Plant Cell Physiol.* **43**(10), 1250–1253 (2002).
- Miyaji, T. *et al.* AtPHT4; 4 is a chloroplast-localized ascorbate transporter in Arabidopsis. *Nat. Commun.* **6**, 5928 (2015).
- Irigoyen, S., Karlsson, P. M., Kuruvilla, J., Spetea, C. & Versaw, W. K. The sink-specific plastidic phosphate transporter PHT4; 2 influences starch accumulation and leaf size in Arabidopsis. *Plant Physiol.* **157**, 1765–1777 (2011).
- Wang, G. Y. *et al.* Circadian clock-regulated phosphate transporter PHT4; 1 plays an important role in Arabidopsis defense. *Mol. Plant.* **4**, 516–526 (2011).
- Wang, C. *et al.* OsSPX-MFS3, a vacuolar phosphate efflux transporter, is involved in maintaining Pi homeostasis in rice. *Plant Physiol.* **169**, 2822–2831 (2015).
- Qin, L. *et al.* The high-affinity phosphate transporter GmPT5 regulates phosphate transport to nodules and nodulation in soybean. *J. Plant Physiol.* **159**, 1634–1643 (2012).
- Wang, X., Yan, X. & Liao, H. Genetic improvement for phosphorus efficiency in soybean: A radical approach. *Ann. Bot.* **106**, 215–222 (2010).
- He, J. *et al.* Genotypic variation in yield, yield components, root morphology and architecture, in soybean in relation to water and phosphorus supply. *Front. Plant Sci.* **8**, 1499 (2017).
- Gutierrez-Boem, F. H. & Thomas, G. W. Phosphorus nutrition and water deficits in field-grown soybeans. *Plant Soil* **207**, 87–96 (1999).
- Jin, J. *et al.* Interaction between phosphorus nutrition and drought on grain yield, and assimilation of phosphorus and nitrogen in two soybean cultivars differing in protein concentration in grains. *J. Plant Nutr.* **29**, 1433–1449 (2006).
- Ogoke, I. J., Carsky, R. J., Togun, A. O. & Dashiell, K. Effect of P fertilizer application on N balance of soybean crop in the guinea savanna of Nigeria. *Agr. Ecosyst. Environ.* **100**, 153–159 (2003).
- Rotaru, V. & Sinclair, T. R. Interactive influence of phosphorus and iron on nitrogen fixation by soybean. *Environ. Exp. Bot.* **66**, 94–99 (2009).
- Sinclair, T. R. & Serraj, R. Legume nitrogen-fixation and drought. *Nature* **378**, 344–344 (1995).
- Serraj, R., Sinclair, T. R. & Purcell, L. C. Symbiotic N<sub>2</sub>fixation response to drought. *J. Exp. Bot.* **50**, 143–155 (1999).
- Zheng, H. F. *et al.* Phosphorus control as an effective strategy to adapt soybean to drought at the reproductive stage: Evidence from field experiments across northeast China. *J. Soil Use Manag.* **31**(1), 19–28 (2015).
- Basal, O. & Szabó, A. Physiomorphology of soybean as affected by drought stress and nitrogen application. *Scientifica (Cairo)*. **2020**, 6093836 (2020).
- Feng, Y. Y., He, J., Turner, N. C., Siddique Kadambot, H. M. & Li, F. M. Phosphorus supply increases internode length and leaf characteristics, and increases dry matter accumulation and seed yield in soybean under water deficit. *J. Agron.* **11**, 5 (2021).



37. Zhu, Y. *et al.* Soybean (Glycine max) expansin gene superfamily origins: Segmental and tandem duplication events followed by divergent selection among subfamilies. *J. BMC Plant Biol.* **14**(1), 93 (2014).
38. Liu, H. J. *et al.* Divergence in enzymatic activities in the soybean GST supergene family provides new insight into the evolutionary dynamics of whole-genome duplicates. *Mol. Biol. Evol.* **11**, 2844–2859 (2015).
39. Sun, T. *et al.* Phosphate transporter MdPHT1;7 enhances phosphorus accumulation and improves low phosphorus and drought tolerance. *J. Plant Biol.* **64**(5), 403–416 (2021).
40. Hawkesford, M. *et al.* Chapter 6: Functions of macronutrients. In *Marschner's mineral nutrition of higher plants* 3rd edn 158–165 (Academic Press, 2012).
41. Gamuyao, R. *et al.* The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature* **488**(7412), 535–539 (2012).
42. Wu, W. W. *et al.* Association of extracellular dNTP utilization with a GmPAP1-like protein identified in cell wall proteomic analysis of soybean roots. *J. Exp. Bot.* **69**(3), 603–617 (2018).
43. Lin, Y., Chen, G., Hu, H., Yang, X. & Liu, Y. Phenotypic and genetic variation in phosphorus-deficiency-tolerance traits in Chinese wheat landraces. *BMC Plant Biol.* **20**(1), 330 (2020).
44. Remy, E. *et al.* The Pht1;9 and Pht1;8 transporters mediate inorganic phosphate acquisition by the Arabidopsis thaliana root during phosphorus starvation. *J. New Phytologist.* **195**(2), 356–371 (2012).
45. Nussaume, L. *et al.* Phosphate import in plants: Focus on the PHT1 transporters. *Front. Plant Sci.* **2**, 83 (2011).
46. Wang, D., Lv, S., Jiang, P. & Li, Y. Roles, regulation, and agricultural application of plant phosphate transporters. *J. Front. Plant Sci.* **8**, 817 (2017).
47. López-Arredondo, D. L., Leyva-González, M. A., González-Morales, S. I., López-Bucio, J. & Herrera-Estrella, L. Phosphate nutrition: Improving low-phosphate tolerance in crops. *J. Annu. Rev. Plant Biol.* **65**(1), 95–123 (2014).
48. Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D. & May, G. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *BMC Plant Biol.* **4**, 10 (2004).
49. Liu, F. *et al.* Comprehensive sequence and whole-life-cycle expression profile analysis of the phosphate transporter gene family in rice. *Mol. Plant.* **4**, 1105–1122 (2011).
50. Sobkowiak, L. *et al.* The role of the P1BS element containing promoter-driven genes in Pi transport and homeostasis in plants [J]. *Front. Plant Sci.* **3**(58), 58 (2012).
51. Rubio, V. *et al.* A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes Dev.* **15**, 2122–2133 (2001).
52. Regla, B. *et al.* A central regulatory system largely controls transcriptional activation and repression responses to phosphate starvation in Arabidopsis. *PLoS Genet.* **6**, e1001102 (2010).
53. Schünmann, P. H., Richardson, A. E., Smith, F. W. & Delhaize, E. Characterization of promoter expression patterns derived from the Pht1 phosphate transporter genes of barley (*Hordeum vulgare* L.). *J. Exp. Bot.* **55**, 855–865 (2004).
54. Schünmann, P. H., Richardson, A. E., Vickers, C. E. & Delhaize, E. Promoter analysis of the barley Pht1;1 phosphate transporter gene identifies regions controlling root expression and responsiveness to phosphate deprivation. *Plant Physiol.* **136**, 4205–4214 (2004).
55. Martin, J. *et al.* Comparative analysis of the combined effects of different water and phosphate levels on growth and biological nitrogen fixation of nine cowpea varieties. *J. Front. Plant Sci.* **8**, 2111 (2017).
56. Li, Y. *et al.* Molecular identification of the phosphate transporter family 1 (PHT1) genes and their expression profiles in response to phosphorus deprivation and other abiotic stresses in Brassica napus. *PLoS ONE* **14**(7), e0220374b (2019).
57. Sun, T., Li, M., Shao, Y., Yu, L. & Ma, F. Comprehensive genomic identification and expression analysis of the phosphate transporter (PHT) gene family in apple. *Front. Plant Sci.* **8**, 426 (2017).
58. Cao, M. X. *et al.* Functional analysis of StPHT1;7, a Solanum tuberosum L. phosphate transporter gene, in growth and drought tolerance. *Plants-Basel.* **9**(10), 1384 (2020).
59. Ackerson, R. C. Osmoregulation in cotton in response to water stress: III effects of phosphorus fertility. *Plant Physiol.* **77**(2), 309–312 (1985).
60. Liu, C., Rubæk, G. H., Liu, F. & Andersen, M. N. Effect of partial root zone drying and deficit irrigation on nitrogen and phosphorus uptake in potato. *Agric. Water Manag.* **159**, 66–76 (2015).
61. Graciano, C., Guiamét, J. J. & Goya, J. F. Impact of nitrogen and phosphorus fertilization on drought responses in eucalyptus grandis seedlings. *For Ecol. Manag.* **212**(1–3), 40–49 (2005).
62. Faustino, L. I., Bulfe, N. M., Pinazo, M. A., Monteoliva, S. E. & Graciano, C. Dry weight partitioning and hydraulic traits in young Pinus taeda trees fertilized with nitrogen and phosphorus in a subtropical area. *Tree Physiol.* **33**(3), 241–251 (2013).
63. Tariq, A., Pan, K., Olatunji, O. A., Graciano, C. & Zhang, A. Phosphorous fertilization alleviates drought effects on Alnus cremastogyne by regulating its antioxidant and osmotic potential. *Sci. Rep.* **8**(1), 5644 (2018).
64. Tng, D. Y., Janos, D. P., Jordan, G. J., Weber, E. & Bowman, D. M. Phosphorus limits eucalyptus grandis seedling growth in an unburnt rain forest soil. *Front. Plant Sci.* **5**, 527 (2014).
65. Mimura, T., Mitsuhashi, N., Sekiguchi, Y. & Ohnishi, M. Phosphate transport and intracellular distribution in plants. *Tanpakushitsu Kakusan Koso* **48**(15), 2044–2051 (2003).
66. Kakar, K. M., Muhammad, T., Taj, F. H. & Nawab, K. Phosphorus use efficiency of soybean as affected by phosphorus application and inoculation. *J. Agron.* **92**(1), 49–50 (2001).
67. Bucher, M., Rausch, C. & Daram, P. Molecular and biochemical mechanisms of phosphorus uptake into plants. *J. Plant Nutr. Soil Sci.* **164**, 209–217 (2015).
68. Begum, N., Ahanger, M. A. & Zhang, L. X. AMF inoculation and phosphorus supplementation alleviates drought induced growth and photosynthetic decline in Nicotiana tabacum by up regulating antioxidant metabolism and osmolyte accumulation. *Environ. Exp. Bot.* **176**, 104088 (2020).
69. Wang, X. *et al.* Phosphate transporters OsPHT1;9 and OsPHT1;10 are involved in phosphate uptake in rice. *Plant Cell Environ.* **37**, 1159–1170 (2014).
70. Ribot, C., Wang, Y. & Poirier, Y. Expression analyses of three members of the AtPHO1 family reveal differential interactions between signaling pathways involved in phosphate deficiency and the responses to auxin, cytokinin, and abscisic acid. *Planta* **227**, 1025–1036 (2008).
71. Saito, R. *et al.* A travel guide to cytoscape plugins. *Nat. Method.* **9**, 1069–1076 (2012).
72. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C (T)) method. *Method* **25**, 402–408. <https://doi.org/10.1006/meth.2001.1262> (2001).

## Acknowledgements

We thank to researcher of Soybean Research Institute of Changchun Academy of Agricultural Sciences for providing experimental materials. This study was supported by the following funding sources: Jilin Science and technology development plan project (20210101045JC, 20200403016SF). Thirteenth Five-Year Plan" Science and Technology Project of Jilin Provincial Education Department (JJKH20200335KJ).



### Author contributions

X.Y. and M.-Y. Y. designed and supervised the project, X.-S.W. and Z.-H.W. analyzed the data, and wrote the manuscript, Y.F., R.-J.Y., and L.W. participated in the material preparation, P.T., S.-Y.L. revised the manuscript, all authors discussed the results and commented on the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25378-w>.

**Correspondence** and requests for materials should be addressed to X.Y. or M.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022