Genome-wide Analysis of NPF, NRT2, CLC and SLAC1/SLAH Nitrate Transporters in Hexaploid Wheat (*Triticum Aestivum*)

Aman Kumar

Punjab Agricultural University

Nitika Sandhu (✉ nitikasandhu@pau.edu)

Punjab Agricultural University

Pankaj Kumar

Punjab Agricultural University

Gomsie Pruthi

Punjab Agricultural University

Jasneet Singh

Punjab Agricultural University

Satinder Kaur

Punjab Agricultural University

Parveen Chhuneja

Punjab Agricultural University

Research Article

Keywords: Nitrate transporters, TaNPF, TaNRT2, TaCLC, TaSLAC1/TaSLAH, Wheat, Candidate genes

Posted Date: September 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-907654/v1>

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

Abstract

Nitrogen transport is one of the most important processes in plants mediated by specialized transmembrane proteins. Plants have two main systems for nitrogen uptake from soil and its transport within the system - a low-affinity transport system and a high-affinity transport system. Nitrate transporters are of special interest in cereal crops because large amount of money is spent on N fertilizers every year to enhance the crop productivity. Till date four gene families of nitrate transporter proteins; *NPF* (nitrate transporter 1/peptide transporter family), *NRT2* (nitrate transporter 2 family), the *CLC* (chloride channel family), and the *SLAC/SLAH* (slow anion channel-associated homologues) have been reported in plants. In our study, *in silico* mining of nitrate transporter genes along with their detailed structure, phylogenetic and expression analysis was carried out. A total of 412 nitrate transporter genes were identified in hexaploid wheat genome using HMMER based homology searches in IWGSC Refseq v2.0. Out of those twenty genes were root specific, 11 leaf/shoot specific and 17 genes were grain/spike specific. The identification of nitrate transporter genes in the close proximity to the previously identified 67 marker-traits associations associated with the nitrogen use efficiency related traits in nested synthetic hexaploid wheat introgression library indicated the robustness of the reported transporter genes. The detailed crosstalk between the genome and proteome and the validation of identified putative candidate genes through expression and gene editing studies may lay down the foundation to improve nitrogen use efficiency of cereal crops.

Key Message:

In-silico mining, phylogenetic analysis and expression studies reported a total of 412 nitrate transporter genes including 20 root specific, 11 leaf/shoot specific and 17 grain/spike specific genes.

1 Introduction

Nitrogen is one of the essential elements required by plants. It is a constituent of nucleic acids, amino acids and proteins and therefore is of great importance in plant physiology and metabolic processes. Though N_2 is abundant in atmosphere, only legumes are able to fix atmospheric N_2 with the help of *Rhizobium* bacteria. All other plants mainly absorb N in the form of inorganic ions (ammonium (NH_4^+) and nitrate (NO_3^-)) from soil. Nitrate is mostly absorbed in aerobic soils, while ammonium is mostly absorbed in acidic soils and wet lands. After uptake, NO_3^- and NH_4^+ are assimilated, transformed and mobilized through various processes within plant system.

The agricultural systems focussed on the high-yield crop production remove nitrogen from the soil and depends mostly on the application of large quantities of nitrogenous fertilizers such as urea for the sustained productivity over time. Unfortunately, a large fraction of the applied nitrogen is not directly absorbed by the plants and is lost by the leaching (Hirel et al., 2011). Despite significant efforts made by the scientific community in the last 50 years, the nitrogen-use efficiency for the cereal crops has not been improved (Cassman et al., 2002). Beyond this, the economic losses and detrimental environmental consequences caused by the use of large quantities of fertilizers in agriculture are critical issues to be considered (Hirel et al. 2007; Galloway et al. 2008). Unravelling the genomic regions or the putative candidate genes improving nitrogen-use efficiency will be the first step toward developing nutrient-efficient crop varieties.

To transport N from soil to roots and to other parts of plants, plasma membrane localized proteins known as transporters are essential. They are involved in regulation of N root uptake, root to shoot and leaf to sink transport (Tegeder 2014, Fan et al 2017). Plants have evolved two systems for N uptake to cope with changes in N availability. These two systems are the low-affinity transport system (LATS) and high-affinity transport system (HATS). A low-affinity transport system (LATS) is involved where adequate amounts of nitrogen levels are present. A high-affinity transport system (HATS) is involved where limited amounts of N are present. Plants have two low-affinity and two high affinity N transport systems; for nitrate (*NRT1*- low-affinity NO_3^- transporters and *NRT2*-high-affinity NO_3^- transporters) and ammonium (*AMT1*- low-affinity NH_4^+ transporters and *AMT2*-high affinity NH_4^+ transporters). Majority of N in cereal crops such as wheat is taken up in form of nitrate (NO_3^-). Therefore, nitrate transporters are of great importance.

In plants four families of NO_3^- transporters have been identified named *NPF* (*NRT1/PTR*), *NRT2*, *CLC* (chloride channel) and *SLAC1/SLAH* (slow type anion channel associated homologs) (Wang et al 2012). *NRT1.1* was first NO_3^- transporter to be identified in *Arabidopsis* (Tsay et al 1993). The *NRT1* transporter family which has been renamed as *NPF* family is the largest family of nitrate transporters and can further be classified into eight subfamilies (Leran et al 2014). In *Arabidopsis* *NPF* transporters have been well characterized and contain 53 members divided into eight subfamilies (Leran et al 2014). In rice (*Oryza sativa*) *NPF* transporters contain 93 members (von Wittgenstein et al 2014). The majority of *NPF* transporters are involved in LATS with few exceptions of *NRT1.1/NPF6.3* in *Arabidopsis* and *MtNRT1.3* in *Medicago truncatula*, which are involved in both HATS and LATS (Liu and Tsay 2003, Morere-Le Paven et al 2011). Although majority of *NPFs* are involved in nitrate transport, several studies have suggested their role in transport of other substrates such as nitrite (Sugiura et al 2007), peptides (Komarova et al 2008), amino acids (Zhou et al 1998) and several plant hormones (Krouk et al 2010, Chiba et al 2015, David et al 2016, Saito et al 2015, Tal et al 2016). The second family known as *NRT2* contains high affinity nitrate transporters. A total of seven *NRT2* transporters in *Arabidopsis* (Miller et al 2007) and five *NRT2* transporters in rice have been reported (Feng et al 2011). Most of *NRT2* transporters require a partner protein - *NAR2* (nitrate assimilation related protein) to function as high affinity nitrate transporters (Kotur et al 2012, Feng et al 2011, Yan et al 2011). Third family of nitrate transporters, *CLC* (chloride channel) family is mainly associated with vacuolar transport of NO_3^- (Zifareli et al 2009). In *Arabidopsis*, six *CLC* genes have been reported and are responsible for nitrate and chloride homeostasis, thereby regulating stomatal movement and salt tolerance (Harada et al 2004, Jossier et al 2010, Zifareli et al 2009). The fourth family - *SLAC1/SLAH* (slow type anion channel associated homologs) is anion channel family. In *Arabidopsis* this family contains four members-*SLAC1*, *SLAH1*, *SLAH2* and *SLAH3* which are involved in the nitrate transport in guard cells and roots and in chloride acquisition (Hedrich et al 2017). Together these four transporter families are involved in efficient nitrate uptake and utilization in plants.

To the best of our knowledge, the nitrate transporters in hexaploid wheat have not been characterized and explored completely. There are some studies conducted to access the effect of different nitrogen conditions on some of *NPF* and *NRT2* genes (Buchner and Hawkesford 2014). Most of the studies in

wheat have been conducted on members of *TaNRT2* gene family. Overexpression of *TaNRT2.5* has been associated with increased grain nitrate uptake and yield (Li et al 2019). *TaNRT2.1* has been associated with post flowering nitrate uptake in wheat (Taulemesse et al 2015). Expression of *TaNRT2.1* can be induced by nitrogen starvation and abscisic acid (ABA) (Guo et al 2014, Yin et al 2007, Wang et al 2020). Some phylogenetic studies and expression-based studies have been conducted on *NPF* and *NRT2* genes recently (Wang et al 2020, Bajgain et al 2018) but *CLC* and *SLAC1/SLAH* genes still remain uncharacterized. Structure of proteins play very important role in the functionality of transporter proteins but still no studies have been conducted on structure prediction of any of *NPF*, *NRT2*, *CLC* and *SLAC1/SLAH* genes in wheat. In our study we have identified and characterized genes belonging to all the four families of nitrate transporters. Our analysis includes gene composition, chromosomal location, phylogenetic relations with members of rice and *Arabidopsis* and expression analysis. We adopted a new nomenclature for identified genes as the earlier nomenclature systems do not include complete information about subgenome and homoeologs. We have classified the genes based on phylogeny and identified homoeologous pairs of the gene. Expression profiles of all the genes were studied for different developmental stages and different tissues. Further the structures of all the members of gene families were investigated.

2 Materials And Methods

2.1 Sequence search and annotation of nitrate transporter genes

Two methods were used for the identification of *NRT1*, *NRT2* genes in wheat. In the first method, the CDD IDs (conserved domain database IDs) specific to *TaNPF*, *TaCLC*, *TaSLAC1/TaSLAH* and *TaNRT2* genes (Table 1) were used as identifiers to retrieve genes from the wheat reference genome (IWGSC RefSeq V2.0) from the Ensembl Plants (<https://plants.ensembl.org/index.html>). In the second method, protein sequences were downloaded from the NCBI database using Nitrate/Nitrogen transporters, and *NRT* as queries. Incomplete, partial sequences, hypothetical, and predicted protein sequences were filtered out. The downloaded sequences were manually curated to remove duplicate sequences and incomplete sequences. The remaining protein sequences (1687 genes) were aligned using Clustal Omega, and the output Stockholm file was used to create the HMMER profile. The HMMER profile was used to search similar protein sequences in the wheat protein database downloaded from IWGSC. A total of 403 high confidence and 38 low confidence proteins were obtained. Separate searches were performed for *TaCLC* and *TaSLAC1/TaSLAH* genes using the same method. A total of 41 *TaCLC* and 43 *TaSLAC1/TaSLAH* high confidence genes and 10 *TaCLC* and 7 *TaSLAC1/TaSLAH* low confidence genes were obtained. The sequences from both the methods were combined, followed by the removal of low confidence proteins and duplicate sequences, and after manual curation, a final set of 412 genes belonging to all four nitrate transporter families were selected. The same methodology was used to identify sequences for *Triticum dicoccoides*(AABB), *T. turgidum*(AABB), *T. urartu*(AA), and *Aegilops tauschii*(DD) for comparative analysis.

2.2 Maximum likelihood phylogeny of nitrate transporter genes

The alignments of *TaNRT1/TaNPF*, *TaCLC*, *TaSLAC1/TaSLAH* and *TaNRT2* sequences were created separately using wheat, rice, and *Arabidopsis* sequences by MAFT (E-INS-I algorithm). The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model and then selecting the topology having superior log-likelihood value. Evolutionary studies were conducted in MEGA X. The consistency of the phylogenetic estimate was evaluated by bootstraps (1000 replicates). The resulting tree was visualized using FIGTREE v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

2.3 Gene structure prediction and identification of homoeologs

The genomic and CDS sequences of genes were downloaded from the Ensembl plants database. The sequence information was utilized to predict the intron/exon positions by using the GSDS server (Gene Structure Display Server, <http://gsds.cbi.pku.edu.cn>). Separate phylogenies were generated for members of each subfamily to resolve the relationship between them. The analysis was performed in MEGA X by the method described previously. Homoeologous genes were identified based on the phylogenetic relationship between the members of subfamilies. The information regarding physical positions of genes were obtained from Ensembl Plants database. Genome wide distribution map of nitrate transporter genes was developed by web based online visualization tool PhenoGram (<http://visualization.ritchielab.org/phenograms/plot>).

2.4 Naming of *TaNPF*, *TaNRT2*, *TaCLC* and *TaSLAC1/SLAH* genes

We adopted the method proposed by Schilling et al. (2020) for the naming of *NRT* genes. The genes were named based on their phylogenetic relationships and subgenome location (A, B, or D). Each gene name started with the abbreviation for the species name *Triticum aestivum* (*Ta*), followed by the most closely related *Arabidopsis* gene name (i.e., *NPF1-NPF8*, *NRT2*), which was followed by the subgenome identifier (A, B, and D). Putative homoeologs were given identical gene names except for the subgenome identifier (*TaNPF4-A1*, *TaNPF4-B1* and *TaNPF4-D1*). The genes belonging to the same subfamily in the same subgenome were consecutively numbered (Table 3).

2.5 Structure prediction of nitrate transporter proteins

Due to the unavailability of crystal structures, gene homology modelling was carried out to predict their three-dimensional (3D) structure. The sequences of *TaNRT1*, *TaCLC*, *TaSLAC1/TaSLAH* and *TaNRT2* genes were submitted to web-based server Phyre2 (Kelly and Sternberg, 2009). Briefly, Phyre2 used PSI-BLAST to detect sequence homologues which was followed by Psi-pred and Diso-pred to predict secondary structure and disorder. Then Hidden Markov

models (HMM) of sequences were generated based on homologues detected before. HMMs of query proteins were scanned against library of HMMs of proteins with experimentally solved structures to construct 3D models of query proteins. Transmembrane helix and topology prediction was carried by memsat-svm (Kelly and Sternberg, 2009).

2.6 Expression analysis of nitrate transporter genes

The RNAseq data of *TaNPF*, *TaNRT2*, *TaCLC* and *TaSLAC1/TaSLAH* genes of various tissues (root, shoot/leaf, spike, grain) at three developmental stages (seedling, vegetative and reproductive) for Chinese spring and Azhurnaya (cv) was downloaded from the wheat expression database (www.wheat-expression.com). Expression levels were downloaded as $\log_2$ (transcripts per million) ($\log_2$ tpm) for different tissues at different time points. Several tissue-specific (root, shoot, leaf, grain) genes were identified based on expression patterns. For triad expression analysis, a method described by Ramirez-Gonzalez et al. (2018) was used. Briefly, the expression data from spring wheat (CS) and Azhurnaya was downloaded from the wheat expression database as TPM for root, leave, shoot spike and grain. For analysis, the triads with expression below one tpm were excluded. Expression values were normalized, triads were assigned balanced, A/B/D suppressed or A/B/D dominant profiles.

2.7 Development of validation panel to check the efficacy of the identified nitrate transporter genes

The nested synthetic hexaploid wheat (N-SHW) introgression library constituting a set of 352 breeding lines derived from four sub-populations (Pop1: 75 lines from PDW233/*Ae. tauschii* acc. pau 14135 amphiploid //2*BWL4444; Pop2: 106 lines from PDW233/*Ae. tauschii* acc. pau 14135 amphiploid //2*BWL3531; Pop3: 88 lines from PBW114/*Ae. tauschii* acc. pau 14170 amphiploid //2*BWL4444; Pop4: 83 lines from PBW114/*Ae. tauschii* acc. pau 14170 amphiploid //2*BWL3531) were developed (Sandhu et al 2021). These N-SHW library, six parents and two synthetic hexaploid wheats were assessed over 2 years in 2018 and 2019 at 3 nitrogen levels [i.e., zero N (0 Kg ha⁻¹), half N (60 Kg ha⁻¹) and full N (recommended, 120 Kg ha⁻¹)]. The detailed phenotyping of the N-SHW introgression libraries for the nitrogen-use efficiency related traits was carried out across years and treatments (Sandhu et al 2021). High-density genotyping was performed using the 35K Axiom® Wheat Breeder's Array (Affymetrix UK Ltd., United Kingdom). Significant marker-trait associations were identified using CMLM (compressed mixed linear model)/P3D (population parameters previously defined) in GAPIT (Genome Association and Prediction Integrated Tool) executed in R. Over 322 marker trait associations for NUE were compared to nitrate transporter genes.

3 Results

3.1 The wheat genome consists of 412 nitrate transporter genes belonging to four different families

A total of 412 nitrate transporter sequences excluding splice variants were identified in IWGSC wheat genome assembly (RefSeq V2.0). The wheat genome consists of 292 *TaNPF* genes, 34 *TaCLC* genes, 40 *TaSLAC1/TaSLAH* genes and 46 *TaNRT2* genes. The *TaNPF* genes could be divided into eight subgroups (*TaNPF1* to *TaNPF8*) based on the presence of conserved domains (Table 1). *TaNPF5* subgroup was the largest group consisting of 97 genes followed by *TaNPF8* (70 genes), *TaNPF2* (41 genes), *TaNPF4* (33 genes), *TaNPF6* (22 genes), *TaNPF3* (12 genes) and *TaNPF7* (11 genes). The *NPF1* subgroup was the smallest one consisting of 6 genes present on homoeologous group chromosomes 3A, 3B and 3D. *TaNRT1/TaNPF* genes were present throughout the genome (Figure 1). The location of genes across chromosomes varied according to the size of the subfamily. The genes belonging to larger subfamilies (e.g., *TaNPF5*, *TaNPF8*, *TaNPF2*) were predominantly located in tandem positions on the distal region of chromosomes. The genes belonging to smaller subfamilies (*TaNPF1*, *TaNPF7*, *TaNPF3*) were located on proximal regions of chromosomes. The genes present near distal ends of chromosomes were found to be in the form of clusters in close vicinity to each other. The majority of *TaNRT2* genes were present in the clusters on the distal end of homoeologous chromosomes 6A, 6B and 6D. *TaCLC* genes were distributed across the wheat genome. *TaSLAC1/TaSLAH* genes were only distributed on homoeologous chromosomes 1A, 1B, 1D, 2A, 2B, 2D, 3A, 3B and 3D. The predicted gene structures contained several intron regions (Supplementary Figures 1 a-k) for many genes in *TaNPF*, *TaCLC* and *TaSLAC1/TaSLAH* families. All the *TaNRT* genes were intron less. The size of predicted genes ranged between 1 Kb to 25 Kb. Several truncated and duplicated genes were also predicted.

3.2 Phylogenetic relationships among nitrate transporter genes

The maximum likelihood phylogenetic tree of all the nitrate transporter genes predicted that wheat contains all the major subfamilies present in *Arabidopsis* and rice (*Oryza sativa*) (Figure 2a). The *TaNRT1/TaNPF* and *TaNRT2* genes could be classified into five subclades. The subclades in the phylogenetic tree followed species phylogeny with *Arabidopsis* genes displaying sister group relationship with wheat genes. Based on the phylogenetic relationship, *TaNRT1/TaNPF* genes fitted well into eight subfamilies (*TaNPF1* to *TaNPF8*) following the *Arabidopsis* model. The topology of larger subclades (*TaNPF5*, *TaNPF8*, *TaNPF2*) was more complex than smaller subclades as they were more expanded in wheat than *Arabidopsis* and rice (Figure 2a, Supplementary Figure 2). *TaNRT2* genes were present as a separate subclade and were closely related to the *TaNPF2* subfamily. The phylogenetic analysis of *TaCLC* and *TaSLAC1/TaSLAH* genes was carried out separately. The results showed *TaCLC* genes could be classified into 6 groups according to phylogenetic relation with *Arabidopsis* and rice genes (Figure 2b). *TaSLAC1/TaSLAH* genes were divided into 4 subclades. The largest subclade in *TaSLAC1/TaSLAH* genes showed close relationship with rice *SLAC1/SLAH* genes but not with *Arabidopsis* genes (Figure 2c).

3.3 Homoeologs retention and gene duplication in nitrate transporter genes

The number of nitrate transporter genes in each family were significantly higher than those in *Arabidopsis* and rice (Table 1, supplementary table 3). The comparison with *T. dicoccoides* (AABB), *T. turgidum* (AABB), *T. urartu* (AA) and *Ae. tauschii* (DD) suggested that most of the homoeologs in hexaploid wheat were retained during evolution (Figure 6 supplementary table 3). There was also evidence of gene duplications in tetraploids and hexaploid wheat, reflected in gene number and phylogenetic data (Figure 2, Supplementary Figure 1 a-k). Most duplicated genes were present in subfamilies with a larger number of genes (*TaNPF5*, *TaNPF8*, *TaNPF2* and *TaNRT2*). Nitrate transporters could be grouped into 13 triads, 26 diads, 2 tetrads and 48 singleton genes based on phylogeny (Table 2). Out of a total of 292 *TaNPF* genes, about 74% of *TaNPF* genes could be grouped into 72 triads of homoeologous genes (A, B, D) based on phylogenetic relationships. Similarly, 71% of *TaNRT2* genes, 97% of *TaCLC* genes and 80% of *TaSLAC1/TaSLAH* genes could be grouped into homoeologous triads.

3.4 Nitrate transporter proteins contain multiple transmembrane helices

To study the structural features of nitrate transporters, we predicted the 3D structures of all 412 protein sequences. All nitrogen transporters were predicted to be transmembrane proteins containing multiple transmembrane segments (Figure 3 i). The majority of proteins comprised of 12-14 transmembrane helices (TMs) with some variation. The basic structure of *TaNRT/TaNPF* proteins included N and C terminal segments followed by multiple transmembrane helices (TMs). The transmembrane helices were connected by alternating cytoplasmic and extracellular loop segments (Figure 3 ii). In *TaNRT1/TaNPF* family, approximately 67% of the proteins contained 14 TMs, 21% contained 13 TMs, 7% of proteins contained 12 TMs while 4% of proteins contained less than 12 TMs (supplementary table 1). Subfamily wise studies showed *TaNPF1* proteins contained only 13 TMs and *TaNPF7* contained only 14 TMs. In rest of subfamilies (*TaNPF2-6*, *TaNPF8*) majority of proteins contained 14 TMs but variation existed. Proteins with even number of TMs had both C and N terminals in cytoplasmic side of membrane. Proteins with odd number of TMs had one end in cytoplasmic side and other in extracellular side (Figure 3 ii). All *TaNRT2* family members contained only 12 TMs (supplementary table 1) (Figure 3 ii). Both C and N terminals of *TaNRT2* proteins were present in cytoplasmic side of the membrane. Both *TaCLC* and *TaSLAC1/SLAH* proteins contained 10 TMs with both N and C terminals in cytoplasmic side of membrane. *TaCLC* genes were characterized by presence of a 30-40 amino acids long re-entrant helix in cytoplasmic side (Figure 3 ii) which was not observed in the proteins of other nitrate transporter gene families.

3.5 Expression patterns of nitrate transporter genes in development stages of wheat

To elucidate the expression patterns of nitrate transporter genes, we studied and compared the expression data of Chinese spring and Azhurnaya for different developmental stages. Approximately 77% of *TaNPF* genes, 30% of *TaNRT2*, 85% of *TaCLC* genes and 36% of *TaSLAC1/TaSLAH* genes were expressed at least at one developmental stage in wheat with a wide expression range of 1-103 tpm (Supplementary table 2). The remaining genes showed very low or no expression (tpm<1). Overall, we identified 20 triads in which 48 genes were showing tissue specific expression, out of which 8 triads were root specific, 5 triads were leaf/shoot specific and 7 triads were showing grain/ spike specific expression (Supplementary table 4). Tissue and developmental stage-specific expression were observed in *TaNPF1* genes, which were only expressed in spike and grain at the reproductive stage (Figure 4A). Similarly, *TaNRT2* genes were predominantly expressed in roots in both vegetative and reproductive stages (Figure 4A). *TaSLAC1/TaSLAH* genes were predominately expressed in roots and leaves with some genes showing expression in spikes also (Figure 4B). *TaCLC* genes showed mostly ubiquitous expression (Figure 4B). For the rest of the subfamilies, the genes within one subfamily differed considerably in their expression patterns. In *TaNPF2* genes, spike/grain specific (3 genes), leaf, spike and grain specific (5 genes) and ubiquitous expression (6 genes) were observed (Figure 4A). *TaNPF3* genes showed spike/grain, leaf specific expression, *TaNPF4* genes showed leaf/root-specific (4 genes) and ubiquitous expression (10 genes) (Figure 5A). *TaNPF5* and *TaNPF8* genes mostly showed ubiquitous expression though the root-specific expression was observed in a few genes (Figure 4A). *TaNPF6* showed ubiquitous (6 genes), leaf and root-specific (6 genes), spike specific (3 genes) and root-specific expression (Figure 4A). *TaNPF7* showed ubiquitous expression in three genes, grain specific expression in two genes and root-specific expression in one gene (Figure 4A).

To find out up to what extent homoeologs differ in the expression patterns, triad expression analysis was performed. Most of the triads showed balanced expression ranging from 55.6% to 65.2% in all the tissues (Figure 5A). In roots, a total of 54 triads were showing expression out of total 83 triads. Out of which 55.6% showed balanced expression, 18.5% showed A suppressed, 11.1% showed D suppressed, 9.3 % showed B suppressed expression. Three triads showed A, B and D dominant expression (1 each) (Figure 5B). In leaf/shoot out of 51 triads, 64.7% showed balanced expression, 9.8% showed A suppressed and B suppressed each, 3.9% triads showed D suppressed expression. 5.8% triads showed A and D dominant expression each while no B dominant expression was observed (Figure 5B). In spikes, 61.9% triads out of 42 triads showed balanced expression. Only D dominant expression was observed in 9.5% of triads while A suppressed, B suppressed, and D suppressed expressions were in about 16.7%, 7.1% 4.7% triads (Figure 5B). Only 23 triads were expressing in grains at the reproductive stage, out of which 65.2% showed balanced expression, 8.7% triads showed A, B, and D suppressed each and 4.3% triads showed B and D dominant expression (Figure 5B).

3.6 Nitrate transporter genes are located in close proximity to the NUE associated SNPs

In a parallel study in our laboratory, the nested synthetic wheat introgression libraries capturing novel genetic variation from wild wheat for the nitrogen use efficiency related traits were developed and genotyped using a high-density SNP array (Sandhu et al 2021). These libraries were phenotypically assessed for the root traits and agronomic performance under three nitrogen input conditions (N: 0 kg ha⁻¹; N: 60 kg ha⁻¹ and N:120 kg ha⁻¹) in the field over two years in 2018 and 2019. Genome-wide association mapping was used to identify marker-trait associations for the root and agronomic traits to identify the marker-trait associations for traits improving nitrogen use efficiency in wheat. We compared 322 marker trait associations for NUE identified by Sandhu et al (2021) to

nitrate transporter genes identified in the present study. We identified 67 SNPs, which were in close proximity to nitrate transporter genes in the wheat genome. A total of 93 nitrate transporter genes could be located near NUE linked SNPs, out of which, 63 genes belonged to *TaNPF* family, 15 genes belonged to *TaNRT2* family, 11 genes belonged to *TaCLC* and 4 genes belonged to *TaSLAC1/TaSLAH* family (Table 4, Supplementary figure 4).

The existing genetic variability for 48 tissue specific genes and 93 genes in close proximity to NUE associated SNPs identified in the present study in different wild and cultivated wheat accessions/varieties may be further utilized in genomics-assisted breeding programs targeting improved nitrogen-use efficiency in wheat. The identification of improved breeding lines or the wild accessions possessing the potential nitrate transporters may serve as novel donors to be used in genomics-assisted introgression program developing nitrogen-efficient wheat varieties. The identified nitrate transporters may have potential for efficient nitrogen uptake and its transport from source to sink.

4 Discussion

The main aim of this study was to identify and analyse nitrate transporters belonging to all the four families and study their dynamics in wheat. The number of nitrate transporter genes detected in wheat was higher as compared to other plant species. This could be explained by a large genome (~ 18Gb) and hexaploid nature of wheat. Presence of three homoeologous sub-genomes in wheat could allow multiple copies of nitrate transporters resulting in higher number of transporter genes. When comparing with diploid progenitors (*Ae. tauschii* and *T. urartu*) and tetraploid wheats (*T. dicoccoides* and *T. turgidum*) the number of genes in each subfamily were approximately proportional (Table 1). The genes were distributed randomly in the genome except for *TaNRT2* genes which were predominantly present on group 6 homoeologous chromosome. Many genes were present in form of clusters and showed high percentage of similarity indicating gene-duplication events. There were genes with deleted segments present in the genome. The phylogenetic relationships with orthologues in other plants could be used to classify the genes in subfamilies. All the major subclades were conserved in wheat in comparison to other plant species indicating biological importance of the subfamilies. Based on phylogeny the genes could be grouped in homoeologous triads. Almost 73% of the genes could be assigned to 1:1:1 homoeologous groups which is very much above the average homoeologous retention rate (35.8%) in wheat (IWGSC, 2018). Many genes were also grouped into tetrads and diads based on homology indicating gene duplication and deletion events in the genome. The overall results revealed that wheat nitrogen transporter families are much more complex than in other plant species. This complexity arises mostly due to presence of three sub-genomes (A B D) and gene duplication and deletion events.

The complexity of wheat genome also affects the expression patterns of genes. Due to presence of multiple sets of homoeologs on A, B and D genomes the buffering effects are observed in expression of genes. To study up to what extent these interactions affect the expression of nitrate transporters, triad expression analysis was performed. More than 55 % of genes showed balanced expression in all the tissues which is comparable to genome-wide assessment of all transcripts in wheat (Ramirez-Gonzalez *et al* 2018). The expression profiles of the genes identified in this study were in accordance to the previous studies in other plants. The expression patterns of nitrate transporter genes were similar to expression patterns of close orthologs in rice and *Arabidopsis* indicating the conservation of gene functions. *TaCLC* genes in previous studies in *Arabidopsis* showed ubiquitous expression which was observed in this study for wheat as well. Several tissue specific nitrate transporter genes were identified which can be targeted for gene manipulation for wheat improvement. Several *TaNRT2* and *TaSLAC1/TaSLAH* genes showed root specific expression suggesting their role in root nitrate uptake. Root specific expression of *NRT2* and *TaSLAC1/TaSLAH* genes has already been reported in rice and *Arabidopsis*. *TaNPF1* genes and some *TaSLAC1/SLAH* genes showed grain and spike specific expression suggesting their role in nitrate transfer in developing seeds.

Structure plays a very important role in the function of transporter proteins. X-ray crystallographic structures of eukaryotic nitrate transporters have been elucidated (Yan *et al* 2013). All the nitrate transporter families belong to a much larger major facilitator superfamily (MFS) according to transporter classification database (Sair *et al* 2014). All the nitrate transporter proteins were predicted to have a typical MFS protein structure with multiple TMs. To the best of our knowledge our study is the first one to report homology-based models of nitrate transporter proteins belonging to all four families in wheat. The number of transmembrane segments play very important role in the optimal functioning MFS transporter proteins (Perland *et al* 2017). For an MFS transporter protein to have optimal transport properties pseudosymmetry is important which is provided by even number of TMs (Yan *et al* 2013). According to previous studies most of MFS proteins required 12 TMs to have optimal function (Reddy *et al* 2012). In our study we predicted nitrate transporter families having variation in the number of TMs. *TaNPF* family being the largest of all showed most variation in the number of TMs with number ranging from 12 to 14. Several proteins with odd number of TMs were also observed. For example, all the members of *TaNPF1* subfamily contain 13 TMs. All *TaNRT2* proteins were highly conserved and contained 12 TMs. Most of the *TaCLC* and *TaSLAC1/TaSLAH* genes contained only 10 TMs. The variation in number of TMs between and within subfamilies and presence of odd number of TMs could not be correlated with expression data suggesting that a much more flexible criteria exists for the function of nitrate transporter proteins. The structural information presented in this study offer foundation for future work to identify molecular mechanisms responsible for functioning of nitrate transporters in wheat.

Previously in many studies overexpression of nitrate transporter genes has been linked to improved nitrogen use efficiency and yield in many plants (Fan *et al* 2016, Chen *et al* 2016, Wang *et al* 2018, Allen *et al* 2016 and Fu *et al* 2015). Overexpression of *OsNRT2.1*, *OsNRT2.3b*, *OsNPF6.3* in rice and *ZmNRT1.1A* in maize has resulted in increased grain yield (Wang *et al* 2018, Wang *et al* 2020, Luo *et al* 2018, Wang *et al* 2018, Yan *et al* 2011). In wheat *TaNRT2.1* is reported to be involved in post-flowering N uptake (Taulemesse *et al* 2015) and is an important gene for improvement of nitrogen use efficiency. The *CLC* genes have been reported to be involved in nitrate accumulation in plants (Zifarelli *et al* 2009) and many *CLC* genes have been reported to have role in stress responses. *SLAC1* is a key player in regulation of stomatal closure. *SLAH* genes are involved in root nitrate and chloride acquisition and translocation to shoot. *SLAC1/SLAH* genes have also been reported to have important role in drought responses (Zhang *et al* 2016). The genome wide analysis of *TaCLC* and *TaSLAC1/TaSLAH* genes in this study is the first reported study of these genes in wheat to the best of our knowledge. Nitrate transporters identified in this study can be promising candidates for gene manipulation to enhance productivity and nitrogen use efficiency in wheat. The identification of nitrate transporter genes in the close proximity to the marker-traits associations indicated the robustness of genome wide association mapping studies and the reliability of the reported transporter genes. The identified nitrate transporters could deepen the understanding of genetic and molecular mechanism behind improving

nitrogen-use efficiency in wheat crop. The nutrient efficient improved breeding lines/accessions possessing identified potential nitrate transporters in the present study may have an effective and strong coordinated signal transduction network involving nitrate transceptor, nitrate response regulator and the master response regulator.

The *in-silico* mining of nitrate transporter genes along with their detailed structure, phylogenetic and expression studies reported a total of 412 nitrate transporter genes including 20 root specific, 11 leaf/shoot specific and 17 grain/spike specific putative candidate genes. The identification of nitrate transporter genes in the close proximity to the previously identified 67 marker-traits associations associated with the nitrogen use efficiency related traits in nested synthetic hexaploid wheat introgression library (Sandhu et al 2021) indicated the robustness of the reported transporter genes. The detailed crosstalk between the genome and proteome and the validation of identified putative candidate genes through expression and gene editing studies may lay down the foundation to improve nitrogen use efficiency of cereal crops. Once validated, the candidate genes may further be deployed in genomics-assisted breeding program to develop nutrient efficient wheat varieties. The present study provides important information on potential nitrate transporters that may lay foundation to develop a new breeding strategy for the sustainable agricultural development of cereal crops with less input - more output and the environmental protection. The identified nitrate transports may be of great significance both in the theory and in the genomics-assisted breeding application.

Declarations

Acknowledgments

We are thankful to IWGSC for the data availability. We are thankful to the Department of Biotechnology, Govt. of India for providing DBT RA fellowship.

Availability of data and materials

All data used in this research are included in this published article and its supplementary information files.

Funding

The work was compiled under projects funded by the Department of Biotechnology, Govt. of India (Grant No. BT/IN/UK-VNC/42/RG/2015-16 and BT/PR30871/BIC/101/1159/2018).

Author's contributions

N.S. and A.K. designed this study; A.K. conducted the *in-silico* studies and drafted the manuscript; N.S. conducted the field and genome wide association mapping studies and contributed to manuscript draft; P.K. helped in the *in-silico* analysis; G.P. and J.S. helped in visualisation; N.S. and P.C. provided resources and contributed to the critical revision of the manuscript.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Allen, S., Guo, Loussaert, D., Rupe, M. & Wang, H. (2016). Enhanced nitrate uptake and nitrate translocation by over-expressing maize functional low-affinity nitrate transporters in transgenic maize. *US Patent 20160010101 A1*.
2. Bajgain, P., Russell, B. & Mohammadi, M. (2018). Phylogenetic analyses and in-seedling expression of ammonium and nitrate transporters in wheat. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-25430-8>
3. Buchner, P. & Hawkesford, M. J. Complex phylogeny and gene expression patterns of members of the nitrate transporter 1/Peptide transporter family (NPF) in wheat. *Journal of Experimental Botany*, 65 (19), 5697–5710101093231 (2014).
4. Cassman, K. G., Dobermann, A. & Walters, D. T. Agroecosystems, nitrogen-use efficiency, and nitrogen management. *AMBIO: A Journal of the Human Environment*, 31 (2), 132–1401015790044 (2002).
5. Chen, J. *et al.* Agronomic nitrogen-use efficiency of rice can be increased by drivingOsNRT2.1expression with theOsNAR2.1promoter. *Plant Biotechnol. J.*, 14 (8), 1705–17151011112531 (2016).
6. Chiba, Y. *et al.* Identification of arabidopsis thaliana nrt1/Ptr family (NPF) proteins capable of transporting plant hormones. *Journal of Plant Research*, 128 (4), 679–68610100710265 (2015).
7. David, L. C. *et al.* N availability modulates the role of NPF3.1, a gibberellin transporter, in GA-mediated phenotypes in arabidopsis., 244 (6), 1315–132810100700425 (2016).
8. Fan, X. *et al.* Plant nitrate transporters: From gene function to application. *Journal of Experimental Botany*, 68 (10), 2463–2475101093011 (2017).

9. Fan, X. *et al.* (2016). Overexpression of a pH-sensitive nitrate transporter in rice increases crop yields. *Proceedings of the National Academy of Sciences*, **113**(26), 7118–7123. <https://doi.org/10.1073/pnas.1525184113>
10. Feng, H. *et al.* Multiple roles of nitrate transport accessory protein NAR2 in plants. *Plant Signal. Behav.*, **6** (9), 1286–1289 [10.1016/j.pbs.2011.06.007](https://doi.org/10.1016/j.pbs.2011.06.007) (2011).
11. Feng, H. *et al.* Spatial expression and regulation of rice high-affinity nitrate transporters by nitrogen and carbon status. *Journal of Experimental Botany*, **62** (7), 2319–2330 (2011).
12. Fu, Y., Yi, H., Bao, J. & Gong, J. LeNRT2.3 functions in nitrate acquisition and long-distance transport in tomato. *FEBS Letters*, **589** (10), 1072–1079 [10.1016/j.febslet.2015.03.016](https://doi.org/10.1016/j.febslet.2015.03.016) (2015).
13. Galloway, J. N. *et al.* Transformation of the nitrogen cycle: Recent trends, questions, and potential solutions., **320** (5878), 889–892 [10.1126/science.1156674](https://doi.org/10.1126/science.1156674) (2008).
14. Guo, T. *et al.* Transcription analysis of genes encoding the wheat root transporter NRT1 and NRT2 families during nitrogen starvation. *Journal of Plant Growth Regulation*, **33** (4), 837–848 [10.1007/s00344-014-0244-4](https://doi.org/10.1007/s00344-014-0244-4) (2014).
15. Harada, H., Kuromori, T., Hirayama, T., Shinozaki, K. & Leigh, R. H. Quantitative trait loci analysis of nitrate storage in arabidopsis leading to an investigation of the contribution of the anion channel gene, AtCLC-C, to variation in nitrate levels. *Journal of Experimental Botany*, **55** (405), 2005–2014 [10.1093/jxb/erj024](https://doi.org/10.1093/jxb/erj024) (2004).
16. Hedrich, R. & Geiger, D. Biology of SLAC 1-type anion channels – from nutrient uptake to stomatal closure. *New Phytol.*, **216** (1), 46–61 [10.1111/nph.14685](https://doi.org/10.1111/nph.14685) (2017).
17. Hirel, B., Le Gouis, J., Ney, B. & Gallais, A. The challenge of improving nitrogen use efficiency in crop plants: Towards a more central role for genetic variability and quantitative genetics within integrated approaches. *Journal of Experimental Botany*, **58** (9), 2369–2387 [10.1093/jxb/erj097](https://doi.org/10.1093/jxb/erj097) (2007).
18. Hirel, B., Tétu, T., Lea, P. J. & Dubois, F. Improving nitrogen use efficiency in crops for sustainable agriculture. *Sustainability*, **3** (9), 1452–1485 [10.3390/s3091452](https://doi.org/10.3390/s3091452) (2011).
19. Hu, B. *et al.* GSDS 2.0: An upgraded gene feature visualization server., **31** (8), 1296–1297 [10.1093/bioinformatics/btu154](https://doi.org/10.1093/bioinformatics/btu154) (2014).
20. Jossier, M. *et al.* The arabidopsis vacuolar anion transporter, AtCLC_c, is involved in the regulation of stomatal movements and contributes to salt tolerance. *Plant J.*, **64** (4), 563–576 [10.1111/j.1365-3113.2010.04610.x](https://doi.org/10.1111/j.1365-3113.2010.04610.x) (2010).
21. Komarova, N. Y. *et al.* AtPTR1 and AtPTR5 transport dipeptides in planta. *Plant Physiol.*, **148** (2), 856–869 [10.1093/plphys/ktn238](https://doi.org/10.1093/plphys/ktn238) (2008).
22. Kotur, Z. *et al.* Nitrate transport capacity of the arabidopsis thaliana NRT2 family members and their interactions with AtNAR2.1. *New Phytol.*, **194** (3), 724–731 [10.1111/j.1365-3113.2012.04699.x](https://doi.org/10.1111/j.1365-3113.2012.04699.x) (2012).
23. Krouk, G. *et al.* Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. *Dev. Cell.*, **18** (6), 927–937 [10.1016/j.devcel.2010.05.008](https://doi.org/10.1016/j.devcel.2010.05.008) (2010).
24. Li, W. *et al.* A wheat transcription factor positively sets seed vigour by regulating the grain nitrate signal. *New Phytol.*, **225** (4), 1667–1680 [10.1111/nph.16234](https://doi.org/10.1111/nph.16234) (2019).
25. Liu, K. & Tsay, Y. Switching between the two action modes of the dual-affinity nitrate transporter CHL1 by phosphorylation. *The EMBO Journal*, **22** (5), 1005–1013 [10.1093/emboj/cdg018](https://doi.org/10.1093/emboj/cdg018) (2003).
26. Luo, B. *et al.* (2018). Overexpression of a high-affinity nitrate transporter OsNRT2.1 increases yield and manganese accumulation in rice under alternating wet and dry condition. *Frontiers in Plant Science*, **9**. <https://doi.org/10.3389/fpls.2018.01192>
27. Lérans, S., Varala, K., Boyer, J., Chiurazzi, M., Crawford, N., Daniel-Vedele, F., David, L., Dickstein, R., Fernandez, E., Forde, B., Gassmann, W., Geiger, D., Gojon, A., Gong, J., Halkier, B. A., Harris, J. M., Hedrich, R., Limami, A. M., Rentsch, D., ... Lacombe, B. (2014). A unified nomenclature of nitrate transporter 1/Peptide transporter family members in plants. *Trends in Plant Science*, **19**(1), 5–9. <https://doi.org/10.1016/j.tplants.2013.08.008>
28. Miller, A., Fan, X., Shen, Q. & Smith, S. Expression and functional analysis of rice NRT2 nitrate transporters. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **146** (4), S241. <https://doi.org/10.1016/j.cbpa.2007.01.558> (2007).
29. Morere-Le Paven, M. *et al.* Characterization of a dual-affinity nitrate transporter MtNRT1.3 in the model legume *Medicago truncatula*. *Journal of Experimental Botany*, **62** (15), 5595–5605 [10.1093/jxb/erj3243](https://doi.org/10.1093/jxb/erj3243) (2011).
30. Perland, E., Bagchi, S., Klaesson, A. & Fredriksson, R. Characteristics of 29 novel atypical solute carriers of major facilitator superfamily type: Evolutionary conservation, predicted structure and neuronal Co-expression. *Open Biology*, **7** (9), 170142 [10.1098/rsob.170142](https://doi.org/10.1098/rsob.170142) (2017).
31. Ramírez-González, R. H. *et al.* The transcriptional landscape of polyploid wheat., **361** (6403), eaar6089. <https://doi.org/10.1126/science.aar6089> (2018).
32. Reddy, V. S., Shlykov, M. A., Castillo, R., Sun, E. I. & Saier, M. H. The major facilitator superfamily (MFS) revisited. *FEBS Journal*, **279** (11), 2022–2035 [10.1111/j.1365-2425.2012.04699.x](https://doi.org/10.1111/j.1365-2425.2012.04699.x) (2012).
33. Saier, M. H., Reddy, V. S., Tamang, D. G. & Västermark, Å. The transporter classification database. *Nucleic Acids Res.*, **42** (D1), D251–D258. <https://doi.org/10.1093/nar/gkt1097> (2013).
34. Saito, H. *et al.* (2015). The jasmonate-responsive GTR1 transporter is required for gibberellin-mediated stamen development in arabidopsis. *Nature Communications*, **6**(1). <https://doi.org/10.1038/ncomms7095>
35. Sandhu, N. *et al.* Genetic dissection uncovers genome wide marker trait associations for plant growth, yield and yield related traits under varying nitrogen levels in nested synthetic wheat introgression libraries (Research under review, 2021).
36. Schilling, S., Kennedy, A., Pan, S., Jermini, L. S. & Melzer, R. Genome-wide analysis of MIKC-type MADS-box genes in wheat: Pervasive duplications, functional conservation and putative neofunctionalization. *New Phytol.*, **225** (1), 511–529 [10.1111/nph.16122](https://doi.org/10.1111/nph.16122) (2020).

37. Sugiura, M., Georgescu, M. N. & Takahashi, M. A nitrite transporter associated with nitrite uptake by higher plant chloroplasts. *Plant and Cell Physiology*, **48** (7), 1022–1035101093073 (2007).
38. Tal, I. *et al.* M., & (2016). The arabidopsis NPF3 protein is a GA transporter. *Nature Communications*, **7**(1). <https://doi.org/10.1038/ncomms11486>
39. Taulemesse, F., Le Gouis, J., Gouache, D., Gibon, Y. & Allard, V. Post-flowering nitrate uptake in wheat is controlled by N status at flowering, with a putative major role of root nitrate transporter NRT2.1. *PLOS ONE*, **10** (3), e0120291. <https://doi.org/10.1371/journal.pone.0120291> (2015).
40. Tegeder, M. Transporters involved in source to sink partitioning of amino acids and ureides: Opportunities for crop improvement. *Journal of Experimental Botany*, **65** (7), 1865–1878101093012 (2014).
41. Tsay, Y., Schroeder, J. I., Feldmann, K. A. & Crawford, N. M. The herbicide sensitivity gene CHL1 of arabidopsis encodes a nitrate-inducible nitrate transporter., **72** (5), 705–7131010160092 (1993).
42. Von Wittgenstein, N. J., Le, C. H., Hawkins, B. J. & Ehrling, J. Evolutionary classification of ammonium, nitrate, and peptide transporters in land plants. *BMC Evolutionary Biology*, **14** (1), 111011861471–111011862148 (2014).
43. Wang, H. *et al.* Phylogeny and gene expression of the complete nitrate transporter 1/Peptide transporter family in triticum aestivum. *Journal of Experimental Botany*, **71** (15), 4531–4546101093210 (2020).
44. Wang, M. *et al.* Taanr1-tabg1 and tawabi5-tanr2s/NARs link ABA metabolism and nitrate acquisition in wheat roots. *Plant Physiol*, **182** (3), 1440–14531011041901482 (2020).
45. Wang, W., Hu, B., Yuan, D., Liu, Y., Che, R., Hu, Y., Ou, S., Liu, Y., Zhang, Z., Wang, H., Li, H., Jiang, Z., Zhang, Z., Gao, X., Qiu, Y., Meng, X., Liu, Y., Bai, Y., Liang, Y., ... Chu, C. (2018). Expression of the nitrate transporter gene OsNRT1.1A/OsNPF6.3 confers high yield and early maturation in rice. *The Plant Cell*, **30**(3), 638–651. <https://doi.org/10.1105/tpc.17.00809>
46. Wang, W., Hu, B., Li, A. & Chu, C. NRT1.1s in plants: Functions beyond nitrate transport. *Journal of Experimental Botany*, **71** (15), 4373–4379101093554 (2020).
47. Wang, Y., Hsu, P. & Tsay, Y. Uptake, allocation and signaling of nitrate. *Trends in Plant Science*, **17** (8), 458–467101016201204006 (2012).
48. Yan, H. *et al.* Structure and mechanism of a nitrate transporter. *Cell Reports*, **3** (3), 716–723101016201303007 (2013).
49. Yan, M. *et al.* Rice OsNAR2.1 interacts with OsNRT2.1, OsNRT2.2 and OsNRT2.3a nitrate transporters to provide uptake over high and low concentration ranges. *Plant. Cell. Environ*, **34** (8), 1360–13721011111365 (2011).
50. Yin, L., Li, P., Wen, B., Taylor, D. & Berry, J. O. Characterization and expression of a high-affinity nitrate system transporter gene (TaNRT2.1) from wheat roots, and its evolutionary relationship to other NTR2 genes. *Plant Sci*, **172** (3), 621–631101016200611014 (2007).
51. Zhang, A. *et al.* S-type anion channels SLAC1 and SLAH3 function as essential negative regulators of inward K⁺ channels and stomatal opening in arabidopsis., **28** (4), 949–9651011051501050 (2016).
52. Zhou, J., Theodoulou, F. L., Muldin, I., Ingemarsson, B. & Miller, A. J. Cloning and functional characterization of a brassica napus transporter that is able to transport nitrate and histidine. *Journal of Biological Chemistry*, **273** (20), 12017–120231010742732012017 (1998).
53. Zifarelli, G. & Pusch, M. CLC transport proteins in plants. *FEBS Letters*, **584** (10), 2122–2127101016200912042 (2009).

Tables

Table 1

Summary of nitrate transporter gene numbers in wheat, rice, *Arabidopsis* and wheat progenitors

Table 2

Number of triads, tetrads, diads and singletons detected in nitrate transporter families in hexaploid wheat genome

Conserved domain (CDD/ Pfam Id)	Gene Family	<i>Triticum aestivum</i> (AABBDD) (2n = 42)	<i>Arabidopsis thaliana</i> (2n = 10)	<i>Oryza sativa</i> (2n = 24)	<i>Triticum dicoccoides</i> (AABB) (2n = 28)	<i>Triticum turgidum</i> (AABB) (2n = 28)	<i>Triticum urartu</i> (AA) (2n = 14)	<i>Aegilops tauchii</i> (DD) (2n = 14)
cd17413	<i>NPF6</i>	22	3	6	15	15	7	7
cd17414	<i>NPF4</i>	33	7	12	23	22	9	9
cd17415	<i>NPF3</i>	12	1	5	8	8	3	3
cd17416	<i>NPF1 &2</i>	47	17	13	39	32	17	16
cd17417	<i>NPF5</i>	97	16	32	73	79	31	29
cd17418	<i>NPF8</i>	70	5	16	47	47	20	24
cd17419	<i>NPF7</i>	11	3	4	10	8	5	3
Cd17341	<i>NRT2</i>	46	7	4	20	20	15	18
PF00654	<i>CLC</i>	34	6	7	23	22	10	11
PF03595	<i>SLAC1/SLAH</i>	40	5	8	21	21	12	10
Total		412	71	107	253	239	121	120

Family/ Subfamily	No. of Triads	No. of diads			No. of Tetrads			Singletons			Un
	A:B:D = (1:1:1)	A:B:D = (1:1:0)	A:B:D = (1:0:1)	A:B:D = (0:1:1)	A:B:D = (1:1:2)	A:B:D = (1:2:1)	A:B:D = (2:1:1)	A:B:D = (1:0:0)	A:B:D = (0:0:1)	A:B:D = (0:1:0)	
<i>TaNPF1</i>	2	0	0	0	0	0	0	0	0	0	1
<i>TaNPF2</i>	9	0	2	2	0	0	0	5	1	0	0
<i>TaNPF3</i>	4	0	0	0	0	0	0	0	0	0	0
<i>TaNPF4</i>	10	1	0	0	0	0	0	1	0	0	0
<i>TaNPF5</i>	22	1	1	6				2	12	1	0
<i>TaNPF6</i>	6	0	0	0	0	0	0	1	2	1	0
<i>TaNPF7</i>	2	0	0	0	0	0	0	3	1	1	1
<i>TaNPF8</i>	17	1	1	3				4	2	1	2
<i>TaNRT2</i>	10	2	3	0	1	0	0	0	0	0	1
<i>TaCLC</i>	10	0	0	0	1	0	0	0	0	0	0
<i>TaSLAC1/</i> <i>TaSLAH</i>	11	1	1	1	0	0	0	0	0	1	4
Total	103	6	8	12	2	0	0	16	18	5	9

Table 3

Grouping and Naming of nitrate transporter genes identified in wheat genome Refeq v2.0

IWGSC RefSeq ID					Name		
Triad/Tetrad/Diad/ Singleton	A	B	D	Un	A	B	D
TaNPF1-T1	<i>TraesCS3A02G304400</i>	<i>TraesCS3B02G332100</i>	<i>TraesCS3D02G297600</i>		<i>TaNPF1-3A1</i>	<i>TaNPF1-3B1</i>	<i>TaNPF1-3D1</i>
TaNPF1-T2	<i>TraesCS3A02G304500</i>	<i>TraesCS3B02G332000</i>	<i>TraesCS3D02G297400</i>		<i>TaNPF1-3A2</i>	<i>TaNPF1-3B2</i>	<i>TaNPF1-3D2</i>
TaNPF2-T1	<i>TraesCS2A02G045500</i>	<i>TraesCS2B02G057700</i>	<i>TraesCS2D02G044200</i>		<i>TaNPF2-2A1</i>	<i>TaNPF2-2B1</i>	<i>TaNPF2-2D1</i>
TaNPF2-T2	<i>TraesCS3A02G418700</i>	<i>TraesCS3B02G454000</i>	<i>TraesCS3D02G414300</i>		<i>TaNPF2-3A1</i>	<i>TaNPF2-3B1</i>	<i>TaNPF2-3D1</i>
TaNPF2-T3	<i>TraesCS3A02G418800</i>	<i>TraesCS3B02G454100</i>	<i>TraesCS3D02G414400</i>		<i>TaNPF2-3A2</i>	<i>TaNPF2-3B2</i>	<i>TaNPF2-3D2</i>
TaNPF2-T4	<i>TraesCS4A02G283900</i>	<i>TraesCS4B02G029600</i>	<i>TraesCS4D02G026800</i>		<i>TaNPF2-4A1</i>	<i>TaNPF2-4B1</i>	<i>TaNPF2-4D1</i>
TaNPF2-S1	<i>TraesCS4A02G440300</i>				<i>TaNPF2-4A2</i>		
TaNPF2-S2	<i>TraesCS4A02G440400</i>				<i>TaNPF2-4A3</i>		
TaNPF2-S3	<i>TraesCS4A02G440500</i>				<i>TaNPF2-4A4</i>		
TaNPF2-S4	<i>TraesCS4A02G440600</i>				<i>TaNPF2-4A5</i>		
TaNPF2-S5	<i>TraesCS4A02G440700</i>				<i>TaNPF2-4A6</i>		
TaNPF2-T5	<i>TraesCS5A02G004400</i>	<i>TraesCS5B02G001100</i>	<i>TraesCS5D02G012500</i>		<i>TaNPF2-5A1</i>	<i>TaNPF2-5B1</i>	<i>TaNPF2-5D1</i>
TaNPF2-T6	<i>TraesCS5A02G037900</i>	<i>TraesCS5B02G039100</i>	<i>TraesCS5D02G045300</i>		<i>TaNPF2-5A2</i>	<i>TaNPF2-5B2</i>	<i>TaNPF2-5D2</i>
TaNPF2-T7	<i>TraesCS5A02G153200</i>	<i>TraesCS5B02G152000</i>	<i>TraesCS5D02G158500</i>		<i>TaNPF2-5A3</i>	<i>TaNPF2-5B3</i>	<i>TaNPF2-5D3</i>
TaNPF2-D1	<i>TraesCS7A02G054000</i>		<i>TraesCS7D02G049300</i>		<i>TaNPF2-7A1</i>		<i>TaNPF2-7D1</i>
TaNPF2-D2	<i>TraesCS7A02G054100</i>		<i>TraesCS7D02G049400</i>		<i>TaNPF2-7A2</i>		<i>TaNPF2-7D2</i>
TaNPF2-T8	<i>TraesCS7A02G121600</i>	<i>TraesCS7B02G020200</i>	<i>TraesCS7D02G119800</i>		<i>TaNPF2-7A3</i>	<i>TaNPF2-7B3</i>	<i>TaNPF2-7D3</i>
TaNPF2-T9	<i>TraesCS7A02G121700</i>	<i>TraesCS7B02G020500</i>	<i>TraesCS7D02G120200</i>		<i>TaNPF2-7A4</i>	<i>TaNPF2-7B4</i>	<i>TaNPF2-7D4</i>
TaNPF2-D3		<i>TraesCS2B02G057600</i>	<i>TraesCS2D02G044000</i>			<i>TaNPF2-7B5</i>	<i>TaNPF2-7D5</i>
TaNPF2-D4		<i>TraesCS7B02G020300</i>	<i>TraesCS7D02G119900</i>			<i>TaNPF2-7B6</i>	<i>TaNPF2-7D6</i>
TaNPF2-S6			<i>TraesCS7D02G076900</i>				<i>TaNPF2-7D7</i>
TaNPF3-T1	<i>TraesCS1A02G257400</i>	<i>TraesCS1B02G267900</i>	<i>TraesCS1D02G256700</i>		<i>TaNPF3-1A1</i>	<i>TaNPF3-1B1</i>	<i>TaNPF3-1D1</i>
TaNPF3-T2	<i>TraesCS1A02G257800</i>	<i>TraesCS1B02G268200</i>	<i>TraesCS1D02G257100</i>		<i>TaNPF3-1A2</i>	<i>TaNPF3-1B2</i>	<i>TaNPF3-1D2</i>
TaNPF3-T3	<i>TraesCS1A02G257900</i>	<i>TraesCS1B02G268300</i>	<i>TraesCS1D02G257200</i>		<i>TaNPF3-1A3</i>	<i>TaNPF3-1B3</i>	<i>TaNPF3-1D3</i>
TaNPF3-T4	<i>TraesCS7A02G206400</i>	<i>TraesCS7B02G113600</i>	<i>TraesCS7D02G209200</i>		<i>TaNPF3-7A1</i>	<i>TaNPF3-7B1</i>	<i>TaNPF3-7D1</i>
TaNPF4-T1	<i>TraesCS2A02G264500</i>	<i>TraesCS2B02G277600</i>	<i>TraesCS2D02G259400</i>		<i>TaNPF4-2A1</i>	<i>TaNPF4-2B1</i>	<i>TaNPF4-2D1</i>

TaNPF4-T2	TraesCS2A02G309100	TraesCS2B02G326200	TraesCS2D02G307400	TaNPF4-2A2	TaNPF4-2B2	TaNPF4-2D2
TaNPF4-T3	TraesCS2A02G350000	TraesCS2B02G368500	TraesCS2D02G348400	TaNPF4-2A3	TaNPF4-2B3	TaNPF4-2D3
TaNPF4-T4	TraesCS2A02G350100	TraesCS2B02G368600	TraesCS2D02G348500	TaNPF4-2A4	TaNPF4-2B4	TaNPF4-2D4
TaNPF4-D1	TraesCS2A02G350200	TraesCS2B02G368400		TaNPF4-2A5	TaNPF4-2B5	
TaNPF4-T5	TraesCS2A02G350300	TraesCS2B02G368700	TraesCS2D02G348600	TaNPF4-2A6	TaNPF4-2B6	TaNPF4-2D6
TaNPF4-S1	TraesCS3A02G272600			TaNPF4-3A1		
TaNPF4-T6	TraesCS4A02G225400	TraesCS4B02G090800	TraesCS4D02G087900	TaNPF4-4A1	TaNPF4-4B1	TaNPF4-4D1
TaNPF4-T7	TraesCS5A02G056100	TraesCS5B02G060800	TraesCS5D02G067100	TaNPF4-5A1	TaNPF4-5B1	TaNPF4-5D1
TaNPF4-T8	TraesCS5A02G056200	TraesCS5B02G060500	TraesCS5D02G067400	TaNPF4-5A2	TaNPF4-5B2	TaNPF4-5D2
TaNPF4-T9	TraesCS5A02G388000	TraesCS5B02G393100	TraesCS5D02G398000	TaNPF4-5A3	TaNPF4-5B3	TaNPF4-5D3
TaNPF4-T10	TraesCS7A02G365100	TraesCS7B02G262200	TraesCS7D02G357300	TaNPF4-7A1	TaNPF4-7B1	TaNPF4-7D1
TaNPF5-T1	TraesCS1A02G150200	TraesCS1B02G168000	TraesCS1D02G147200	TaNPF5-1A1	TaNPF5-1B1	TaNPF5-1D1
TaNPF5-T2	TraesCS1A02G150400	TraesCS1B02G168100	TraesCS1D02G147400	TaNPF5-1A2	TaNPF5-1B2	TaNPF5-1D2
TaNPF5-T3	TraesCS1A02G269400	TraesCS1B02G279900	TraesCS1D02G269500	TaNPF5-1A3	TaNPF5-1B3	TaNPF5-1D3
TaNPF5-T4	TraesCS1A02G269500	TraesCS1B02G280000	TraesCS1D02G269600	TaNPF5-1A4	TaNPF5-1B4	TaNPF5-1D4
TaNPF5-T5	TraesCS1A02G269600	TraesCS1B02G280100	TraesCS1D02G269700	TaNPF5-1A5	TaNPF5-1B5	TaNPF5-1D5
TaNPF5-T6	TraesCS2A02G565600	TraesCS2B02G626000	TraesCS2D02G576000	TaNPF5-2A1	TaNPF5-2B1	TaNPF5-2D1
TaNPF5-D1	TraesCS2A02G571800		TraesCS2D02G583300	TaNPF5-2A2		TaNPF5-2D2
TaNPF5-T7	TraesCS2A02G571900	TraesCS2B02G615500	TraesCS2D02G583400	TaNPF5-2A3	TaNPF5-2B3	TaNPF5-2D3
TaNPF5-D2	TraesCS2A02G572000	TraesCS2B02G615400		TaNPF5-2A4	TaNPF5-2B4	
TaNPF5-S1	TraesCS2A02G572100			TaNPF5-2A5		
TaNPF5-T8	TraesCS2A02G572200	TraesCS2B02G615300	TraesCS2D02G583500	TaNPF5-2A6	TaNPF5-2B6	TaNPF5-2D6
TaNPF5-T9	TraesCS2A02G572300	TraesCS2B02G615200	TraesCS2D02G583600	TaNPF5-2A7	TaNPF5-2B7	TaNPF5-2D7
TaNPF5-T10	TraesCS3A02G185600	TraesCS3B02G215200	TraesCS3D02G189500	TaNPF5-3A1	TaNPF5-3B1	TaNPF5-3D1
TaNPF5-T11	TraesCS3A02G382100	TraesCS3B02G414800	TraesCS3D02G375900	TaNPF5-3A2	TaNPF5-3B2	TaNPF5-3D2
TaNPF5-T12	TraesCS3A02G382200	TraesCS3B02G414900	TraesCS3D02G375800	TaNPF5-3A3	TaNPF5-3B3	TaNPF5-3D3
TaNPF5-T13	TraesCS3A02G382300	TraesCS3B02G415200	TraesCS3D02G375700	TaNPF5-3A4	TaNPF5-3B4	TaNPF5-3D4
TaNPF5-T14	TraesCS3A02G382400	TraesCS3B02G415300	TraesCS3D02G375600	TaNPF5-3A5	TaNPF5-3B5	TaNPF5-3D5
TaNPF5-D3	TraesCS3A02G382600		TraesCS3D02G375500	TaNPF5-3A6		TaNPF5-3D6

TaNPF5-D4	TraesCS3A02G382700		TraesCS3D02G375400	TaNPF5-3A7		TaNPF5-3D7
TaNPF5-D5	TraesCS3A02G382800		TraesCS3D02G375300	TaNPF5-3A8		TaNPF5-3D8
TaNPF5-D6	TraesCS3A02G382900		TraesCS3D02G375200	TaNPF5-3A9		TaNPF5-3D9
TaNPF5-T15	TraesCS3A02G383200	TraesCS3B02G415600	TraesCS3D02G376200	TaNPF5-3A10	TaNPF5-3B10	TaNPF5-3D10
TaNPF5-T16	TraesCS3A02G383300	TraesCS3B02G415700	TraesCS3D02G376300	TaNPF5-3A11	TaNPF5-3B11	TaNPF5-3D11
TaNPF5-T17	TraesCS5A02G485000	TraesCS5B02G498400	TraesCS5D02G498500	TaNPF5-5A1	TaNPF5-5B1	TaNPF5-5D1
TaNPF5-T18	TraesCS5A02G485200	TraesCS5B02G498500	TraesCS5D02G498700	TaNPF5-5A2	TaNPF5-5B2	TaNPF5-5D2
TaNPF5-T19	TraesCS5A02G485300	TraesCS5B02G498700	TraesCS5D02G498800	TaNPF5-5A3	TaNPF5-5B3	TaNPF5-5D3
TaNPF5-S2	TraesCS5A02G508500			TaNPF5-5A4		
TaNPF5-T20	TraesCS6A02G041300	TraesCS6B02G056500	TraesCS6D02G047600	TaNPF5-6A1	TaNPF5-6B1	TaNPF5-6D1
TaNPF5-T21	TraesCS7A02G196100	TraesCS7B02G101800	TraesCS7D02G197600	TaNPF5-7A1	TaNPF5-7B1	TaNPF5-7D1
TaNPF5-T22	TraesCS7A02G461200	TraesCS7B02G362700	TraesCS7D02G449400	TaNPF5-7A2	TaNPF5-7B2	TaNPF5-7D2
TaNPF5-D7	TraesCS7A02G504300		TraesCS7D02G491400	TaNPF5-7A3		TaNPF5-7D3
TaNPF5-S3		TraesCS2B02G013000			TaNPF5-2B8	
TaNPF5-S4		TraesCS2B02G248000			TaNPF5-2B9	
TaNPF5-S5		TraesCS2B02G401000			TaNPF5-2B10	
TaNPF5-S6		TraesCS2B02G626100			TaNPF5-2B11	
TaNPF5-S7		TraesCS2B02G626600			TaNPF5-2B12	
TaNPF5-S8		TraesCS2B02G626700			TaNPF5-2B13	
TaNPF5-S9		TraesCS3B02G304500			TaNPF5-3B12	
TaNPF5-S10		TraesCS3B02G415000			TaNPF5-3B13	
TaNPF5-S11		TraesCS3B02G415100			TaNPF5-3B14	
TaNPF5-S12		TraesCS4B02G057000			TaNPF5-4B1	
TaNPF5-S13		TraesCS4B02G338600			TaNPF5-4B2	
TaNPF5-S14			TraesCS4D02G335100			TaNPF5-4D1
TaNPF5-D8		TraesCS7B02G040100	TraesCS7D02G139600		TaNPF5-7B4	TaNPF5-7D4
TaNPF5-S15		TraesCS7B02G312500			TaNPF5-7B5	
TaNPF6-T1	TraesCS1A02G031300	TraesCS1B02G038700	TraesCS1D02G032700	TaNPF6-1A1	TaNPF6-1B1	TaNPF6-1D1
TaNPF6-T2	TraesCS1A02G210900	TraesCS1B02G224900	TraesCS1D02G214200	TaNPF6-1A2	TaNPF6-1B2	TaNPF6-1D2

TaNPF6-T3	TraesCS1A02G211000	TraesCS1B02G225000	TraesCS1D02G214300	TaNPF6-1A3	TaNPF6-1B3	TaNPF6-1D3
TaNPF6-T4	TraesCS2A02G335800	TraesCS2B02G346100	TraesCS2D02G327000	TaNPF6-2A1	TaNPF6-2B1	TaNPF6-2D1
TaNPF6-S1	TraesCS4B02G371000				TaNPF6-4B1	
TaNPF6-S2	TraesCS4B02G375800				TaNPF6-4B2	
TaNPF6-S3	TraesCS4D02G361500					TaNPF6-4D3
TaNPF6-T5	TraesCS5A02G409600	TraesCS5B02G414000	TraesCS5D02G419200	TaNPF6-5A1	TaNPF6-5B1	TaNPF6-5D1
TaNPF6-S4	TraesCS5A02G537100			TaNPF6-5A2		
TaNPF6-T6	TraesCS7A02G301700	TraesCS7B02G201900	TraesCS7D02G297000	TaNPF6-7A1	TaNPF6-7B1	TaNPF6-7D1
TaNPF7-S1	TraesCS4A02G284300			TaNPF7-4A1		
TaNPF7-S2	TraesCS5A02G546200			TaNPF7-5A1		
TaNPF7-T1	TraesCS6A02G263500	TraesCS6B02G290500	TraesCS6D02G251500	TaNPF7-6A1	TaNPF7-6B1	TaNPF7-6D1
TaNPF7-T2	TraesCS6A02G280200	TraesCS6B02G309200	TraesCS6D02G260500	TaNPF7-6A2	TaNPF7-6B2	TaNPF7-6D2
TaNPF7-S3	TraesCS7A02G413200			TaNPF7-7A1		
TaNPF7-S4	TraesCS4B02G380000				TaNPF7-4B2	
TaNPF7-S5	TraesCSU02G130200					
TaNPF8-T1	TraesCS2A02G416800	TraesCS2B02G000500	TraesCS2D02G413900	TaNPF8-2A1	TaNPF8-2B1	TaNPF8-2D1
TaNPF8-T2	TraesCS3A02G056400	TraesCS3B02G069100	TraesCS3D02G056300	TaNPF8-3A1	TaNPF8-3B1	TaNPF8-3D1
TaNPF8-T3	TraesCS3A02G057000	TraesCS3B02G070200	TraesCS3D02G056700	TaNPF8-3A2	TaNPF8-3B2	TaNPF8-3D2
TaNPF8-T4	TraesCS3A02G392800	TraesCS3B02G424700	TraesCS3D02G385600	TaNPF8-3A3	TaNPF8-3B3	TaNPF8-3D3
TaNPF8-T5	TraesCS3A02G392900	TraesCS3B02G424800	TraesCS3D02G385700	TaNPF8-3A4	TaNPF8-3B4	TaNPF8-3D4
TaNPF8-T6	TraesCS4A02G075700	TraesCS4B02G231500	TraesCS4D02G232900	TaNPF8-4A1	TaNPF8-4B1	TaNPF8-4D1
TaNPF8-T7	TraesCS4A02G075900	TraesCS4B02G231700	TraesCS4D02G233000	TaNPF8-4A2	TaNPF8-4B2	TaNPF8-4D2
TaNPF8-T8	TraesCS4A02G076000	TraesCS4B02G231800	TraesCS4D02G233100	TaNPF8-4A3	TaNPF8-4B3	TaNPF8-4D3
TaNPF8-T9	TraesCS4A02G076100	TraesCS4B02G232000	TraesCS4D02G233000	TaNPF8-4A4	TaNPF8-4B4	TaNPF8-4D4
TaNPF8-T10	TraesCS4A02G076200	TraesCS4B02G232100	TraesCS4D02G233400	TaNPF8-4A5	TaNPF8-4B5	TaNPF8-4D5
TaNPF8-T11	TraesCS4A02G262700	TraesCS4B02G052200	TraesCS4D02G052400	TaNPF8-4A6	TaNPF8-4B6	TaNPF8-4D6
TaNPF8-S1	TraesCS4A02G287300			TaNPF8-4A7		
TaNPF8-S2	TraesCS4A02G287900			TaNPF8-4A8		

TaNPF8-T12	TraesCS6A02G142600	TraesCS6B02G171000	TraesCS6D02G132100	TaNPF8-6A1	TaNPF8-6B1	TaNPF8-6D1
TaNPF8-D1	TraesCS7A02G095200		TraesCS7D02G091600	TaNPF8-7A1		TaNPF8-7D1
TaNPF8-T13	TraesCS7A02G381500	TraesCS7B02G283400	TraesCS7D02G377800	TaNPF8-7A2	TaNPF8-7B2	TaNPF8-7D2
TaNPF8-S3	TraesCS7A02G381600			TaNPF8-7A3		
TaNPF8-T14	TraesCS7A02G381700	TraesCS7B02G283800	TraesCS7D02G377900	TaNPF8-7A4	TaNPF8-7B4	TaNPF8-7D4
TaNPF8-T15	TraesCS7A02G381800	TraesCS7B02G284300	TraesCS7D02G378300	TaNPF8-7A5	TaNPF8-7B5	TaNPF8-7D5
TaNPF8-D2	TraesCS7A02G412100	TraesCS7B02G311400		TaNPF8-7A6	TaNPF8-7B6	
TaNPF8-T16	TraesCS7A02G413100	TraesCS7B02G312600	TraesCS7D02G406200	TaNPF8-7A7	TaNPF8-7B7	TaNPF8-7D7
TaNPF8-T17	TraesCS7A02G413300	TraesCS7B02G312700	TraesCS7D02G406400	TaNPF8-7A8	TaNPF8-7B8	TaNPF8-7D8
TaNPF8-S4	TraesCS7A02G531000			TaNPF8-7A9		
TaNPF8-D3		TraesCS3B02G069900	TraesCS3D02G057000		TaNPF8-3B5	TaNPF8-3D5
TaNPF8-D4		TraesCS4B02G026700	TraesCS4D02G024400		TaNPF8-4B9	TaNPF8-4D9
TaNPF8-S5		TraesCS4B02G398100			TaNPF8-4B10	
TaNPF8-S6		TraesCS5B02G245300			TaNPF8-5B1	
TaNPF8-D5		TraesCS6B02G406100	TraesCS6D02G353500		TaNPF8-6B2	TaNPF8-6D2
TaNPF8-S7			TraesCS7D02G518900			TaNPF8-7D9
TaNPF8-S8			TraesCSU02G207500			
TaNPF8-S9			TraesCSU02G115500			
TaNRT2-D1	TraesCS2A02G074800		TraesCS2D02G073500	TaNRT2-2A1		TaNRT2-2D1
TaNRT2-T1	TraesCS3A02G254000	TraesCS3B02G285900	TraesCS3D02G254900	TaNRT2-3A1	TaNRT2-3B1	TaNRT2-3D1
TaNRT2-D2	TraesCS6A02G030700	TraesCS6B02G044100		TaNRT2-6A1	TaNRT2-6B1	
TaNRT2-T2	TraesCS6A02G030800	TraesCS6B02G044400	TraesCS6D02G035900	TaNRT2-6A2	TaNRT2-6B2	TaNRT2-6D2
TaNRT2-T3	TraesCS6A02G030900	TraesCS6B02G044300	TraesCS6D02G035800	TaNRT2-6A3	TaNRT2-6B3	TaNRT2-6D3
TaNRT2-T4	TraesCS6A02G031000	TraesCS6B02G044200	TraesCS6D02G035700	TaNRT2-6A4	TaNRT2-6B4	TaNRT2-6D4
TaNRT2-D3	TraesCS6A02G031100	TraesCS6B02G044500		TaNRT2-6A5	TaNRT2-6B5	
TaNRT2-T5	TraesCS6A02G031200	TraesCS6B02G044000	TraesCS6D02G035600	TaNRT2-6A6	TaNRT2-6B6	TaNRT2-6D6
TaNRT2-T6	TraesCS6A02G032400	TraesCS6B02G045600	TraesCS6D02G037200	TaNRT2-6A7	TaNRT2-6B7	TaNRT2-6D7
TaNRT2-T7	TraesCS6A02G032500	TraesCS6B02G045700	TraesCS6D02G037300	TaNRT2-6A8	TaNRT2-6B8	TaNRT2-6D8
TaNRT2-T8	TraesCS6A02G032800	TraesCS6B02G046500	TraesCS6D02G037800	TaNRT2-6A9	TaNRT2-6B9	TaNRT2-6D9

TaNRT2-D4	TraesCS6A02G032900		TraesCS6D02G037900		TaNRT2-6A10		TaNRT2-6D10
TaNRT2-TT1	TraesCS6A02G033000	TraesCS6B02G046600	TraesCS6D02G038100	TraesCS6D02G038000	TaNRT2-6A11	TaNRT2-6B11	TaNRT2-6D11
TaNRT2-D5	TraesCS6A02G033100		TraesCS6D02G038300		TaNRT2-6A12		TaNRT2-6D12
TaNRT2-T9	TraesCS6A02G033200	TraesCS6B02G046700	TraesCS6D02G038200		TaNRT2-6A13	TaNRT2-6B13	TaNRT2-6D13
TaNRT2-T10	TraesCS7A02G428500	TraesCS7B02G328700	TraesCS7D02G420900		TaNRT2-7A1	TaNRT2-7B1	TaNRT2-7D1
TaNRT2-D6			TraesCS1D02G035700	TraesCSU02G002800			TaNRT2-1D1
TaCLC-T1	TraesCS2A02G309900	TraesCS2B02G326900	TraesCS2D02G308100		TaCLC-2A1	TaCLC-2B1	TaCLC-2D1
TaCLC-T2	TraesCS2A02G517500	TraesCS2B02G546000	TraesCS2D02G519000		TaCLC-2A2	TaCLC-2B2	TaCLC-2D2
TaCLC-T3	TraesCS3A02G253600	TraesCS3B02G285500	TraesCS3D02G254500		TaCLC-3A1	TaCLC-3B1	TaCLC-3D1
TaCLC-TT1	TraesCS3A02G125300	TraesCS3B02G144700	TraesCS3D02G126700	TraesCS3D02G126600	TaCLC-3A2	TaCLC-3B2	TaCLC-3D2x
TaCLC-T4	TraesCS3A02G390100	TraesCS3B02G418700	TraesCS3D02G379600		TaCLC-3A3	TaCLC-3B3	TaCLC-3D3
TaCLC-T5	TraesCS4A02G277600	TraesCS4B02G035500	TraesCS4D02G033500		TaCLC-4A1	TaCLC-4B1	TaCLC-4D1
TaCLC-T6	TraesCS5A02G449500	TraesCS5B02G457100	TraesCS5D02G456000		TaCLC-5A1	TaCLC-5B1	TaCLC-5D1
TaCLC-T7	TraesCS6A02G098500	TraesCS6B02G126400	TraesCS6D02G084300		TaCLC-6A1	TaCLC-6B1	TaCLC-6D1
TaCLC-T8	TraesCS6A02G098600	TraesCS6B02G126800	TraesCS6D02G084000		TaCLC-6A2	TaCLC-6B2	TaCLC-6D2
TaCLC-T9	TraesCS6A02G283600	TraesCS6B02G312100	TraesCS6D02G264100		TaCLC-6A3	TaCLC-6B3	TaCLC-6D3
TaCLC-T10	TraesCS7A02G240700	TraesCS7B02G136300	TraesCS7D02G239700		TaCLC-7A1	TaCLC-7B1	TaCLC-7D1
TaSLAC-T1	TraesCS1A02G127500	TraesCS1B02G147400	TraesCS1D02G126500		TaSLAC-1A1	TaSLAC-1B1	TaSLAC-1D1
TaSLAC-D1	TraesCS1A02G423000	TraesCS1B02G455100			TaSLAC-1A2	TaSLAC-1B2	
TaSLAC-T2	TraesCS1A02G423900	TraesCS1B02G456000	TraesCS1D02G432500		TaSLAC-1A3	TaSLAC-1B3	TaSLAC-1D3
TaSLAC-D2	TraesCS1A02G424400			TraesCSU02G204200	TaSLAC-1A4		
TaSLAC-T3	TraesCS1A02G424500	TraesCS1B02G456500	TraesCS1D02G433100		TaSLAC-1A5	TaSLAC-1B5	TaSLAC-1D5
TaSLAC-T4	TraesCS2A02G398000	TraesCS2B02G416100	TraesCS2D02G395700		TaSLAC-2A1	TaSLAC-2B1	TaSLAC-2D1
TaSLAC-T5	TraesCS3A02G028100	TraesCS3B02G018300	TraesCS3D02G017800		TaSLAC-3A1	TaSLAC-3B1	TaSLAC-3D1
TaSLAC-T6	TraesCS3A02G151400	TraesCS3B02G178600	TraesCS3D02G159600		TaSLAC-3A2	TaSLAC-3B2	TaSLAC-3D2
TaSLAC-T7	TraesCS3A02G167000	TraesCS3B02G199200	TraesCS3D02G174800		TaSLAC-3A3	TaSLAC-3B3	TaSLAC-3D3
TaSLAC-T8	TraesCS3A02G225100	TraesCS3B02G254700	TraesCS3D02G228400		TaSLAC-3A4	TaSLAC-3B4	TaSLAC-3D4
TaSLAC-D3		TraesCS1B02G456100	TraesCS1D02G432700		TaSLAC-1B6		TaSLAC-1D6
TaSLAC-S1		TraesCS1B02G388600				TaSLAC-	

				1B7	
TaSLAC-T9	<i>TraesCS1B02G456200</i>	<i>TraesCS1D02G432900</i>	<i>TraesCSU02G001500</i>	<i>TaSLAC-1B8</i>	<i>TaSL1D8</i>
TaSLAC-T10	<i>TraesCS1B02G456400</i>	<i>TraesCS1D02G432800</i>	<i>TraesCSU02G001400</i>	<i>TaSLAC-1B9</i>	<i>TaSL1D9</i>
TaSLAC-T11	<i>TraesCS1B02G456300</i>	<i>TraesCS1D02G433200</i>	<i>TraesCSU02G001600</i>	<i>TaSLAC-1B10</i>	<i>TaSL1D10</i>

Table 4

Proximity of nitrogen use efficiency (NUE) linked SNPs (Sandhu et al 2021) to nitrate transporters detected in present study

SNP related to NUE	Chromosome	SNP Position	Nearby nitrate transporters	nitrate transporter Position	Distance (in Mb)
AX94950355	1A	12918698	<i>TaNPF6-1A1</i>	14519757	1.601059
AX94815202	1A	14468156	<i>TaNPF6-1A1</i>	14519757	0.051601
AX94665912	1B	624080881	<i>TaSLAC-1B10</i>	622365197	1.715684
AX94923560	2A	729858424	<i>TaCLC-2A2</i>	740847366	10.988942
AX94906008	2A	737049474	<i>TaCLC-2A2</i>	740847366	3.797892
AX95162328	2A	745066946	<i>TaCLC-2A2</i>	740847366	4.21958
AX94601746	2B	745715147	<i>TaCLC-2B2</i>	742813858	2.901289
AX95203088	2B	748700718	<i>TaCLC-2B2</i>	742813858	5.88686
AX95190948	2B	752830609	<i>TaCLC-2B2</i>	742813858	10.016751
AX95189671	2D	394797805	<i>TaNPF4-2D2, TaCLC-2D1</i>	394118961, 395130092	0.332287, 0.678844
AX94829391	2D	601212191	<i>TaCLC-2D2</i>	608915455	7.703264
AX95142803	2D	601600533	<i>TaCLC-2D2</i>	608915455	7.314922
AX94799671	2D	608756380	<i>TaCLC-2D2</i>	608915455	0.159075
AX95142189	2D	609577225	<i>TaCLC-2D2</i>	608915455	0.66177
AX94786006	2D	610277424	<i>TaCLC-2D2</i>	608915455	1.361969
AX95148777	2D	641963392	<i>TaNPF5-2D1-TaNPF5-2D5</i>	639677529-643761743	1.798351-2.285863
AX95238274	3A	429463868	<i>TaSLAC-3A4</i>	421719078	7.74479
AX94593608	3A	671144035	<i>TaNPF2-3A1, TaNPF2-3A2</i>	660436466, 660507764	10.636271, 10.707569
AX95237615	3B	6378879	<i>TaSLAC-3B1</i>	7598907	1.220028
AX95259763	3B	229302401	<i>TaSLAC-3B3</i>	227663976	1.638425
AX95136655	3B	235865416	<i>TaSLAC-3B3</i>	227663976	8.20144
AX94723497	3B	236511642	<i>TaSLAC-3B3, TaNPF3B1</i>	227663976	8.847666
AX94561045	3B	642481079	<i>TaNPF5-3B3 - TaNPF5-3B10, TaCLC-3B3</i>	651425224-655435367	8.944145-12.954288
AX94539428	3B	657947249	<i>TaCLC-3B3, TaNPF5-3B3 - TaNPF5-3B10, TaNPF-3B4, TaNPF-3B5</i>	651425224-662795946	2.511882-6.522025
AX94386613	3B	658604225	<i>TaCLC-3B3, TaNPF5-3B3 - TaNPF5-3B10, TaNPF-3B4, TaNPF-3B5</i>	651425224-662795946	3.168858-7.179001
AX94418180	3B	659275308	<i>TaCLC-3B3, TaNPF5-3B3 - TaNPF5-3B10, TaNPF-3B4, TaNPF-3B5</i>	651425224-662795946	3.839941-7.850084
AX94429243	3B	659787974	<i>TaCLC-3B3, TaNPF5-3B3 - TaNPF5-3B10, TaNPF-3B4, TaNPF-3B5</i>	651425224-662795946	4.352607-8.36275
AX94910184	3D	352948426	<i>TaCLC-3D1, TaNRT2-3D1</i>	355885478, 356623041	2.937052, 3.674615
AX94514369	4A	544201715	<i>TaNPF4-4A1</i>	533257983	10.943732
AX94926692	4A	544202284	<i>TaNPF4-4A1</i>	533257983	10.944301
AX94766675	4A	575009572	<i>TaNPF8-4A6</i>	575006132	0.00344
AX94400142	4A	581754986	<i>TaCLC-4A1, TaNPF2-4A1, TaNPF7-4A1, TaNPF8-4A7, TaNPF8-4A8</i>	585431883-593113134	3.676897-11.358148
AX94414780	4B	25929732	<i>TaCLC-4B1, TaNPF2-4B1, TaNPF4B1</i>	20278828-25842359	5.650904
AX94478236	4B	28716503	<i>TaCLC-4B1, TaNPF2-4B1, TaNPF4B1</i>	20278828-25842359	2.874144-8.437675
AX94997694	4B	34789538	<i>TaCLC-4B1, TaNPF2-4B1, TaNPF4B1</i>	20278828-25842359	8.947179-14.51071
AX94517352	4D	21886662	<i>TaCLC-4D1, TaNPF2-4D1, TaNPF8-4D1</i>	10764927-15356868	6.529794-11.121735

AX94586364	4D	22947854	<i>TaCLC-4D1, TaNPF2-4D1, TaNPF8-4D1</i>	10764927-15356868	7.590986-12.182927
AX94914919	4D	28974006	<i>TaNPF8-4D2</i>	28481269	0.492737
AX94738199	5D	10899555	<i>TaNPF2-5D1</i>	6820550	4.079005
AX95110067	5D	467774783	<i>TaNPF4-5D3</i>	464415850	3.358933
AX95002541	5D	468689841	<i>TaNPF4-5D3</i>	464415850	4.273991
AX95132327	5D	472234562	<i>TaNPF4-5D3</i>	464415850	7.818712
AX94631745	5D	528728651	<i>TaNPF5-5D1 - TaNPF5-5D3</i>	528294425-528587054	0.141597-0.43423
AX94803288	6A	14353974	<i>TaNRT2-6A1- TaNRT2-6A13</i>	15727844-16408185	1.37387-2.054211
AX95017906	6A	23433182	<i>TaNRT2-6A14</i>	21634811	1.798371
AX94983341	6A	28412753	<i>TaNRT2-6A14</i>	21634811	6.777942
AX95210745	6A	29967076	<i>TaNRT2-6A14</i>	21634811	8.332265
AX94510892	6A	112585030	<i>TaNPF8-6A1</i>	117412062	4.827032
AX94534539	6A	497462168	<i>TaNPF7-6A1</i>	486547388	10.91478
AX94573487	6D	27978202	<i>TaNPF5-6D1</i>	22172543	5.805659
AX94415776	6D	28700804	<i>TaNPF5-6D1</i>	22172543	6.528261
AX94978974	6D	29876083	<i>TaNPF5-6D1</i>	22172543	7.70354
AX94737868	6D	29876631	<i>TaNPF5-6D1</i>	22172543	7.704088
AX95250225	6D	29928065	<i>TaNPF5-6D1</i>	22172543	7.755522
AX94461279	6D	451183032	<i>TaNPF8-6D2</i>	449226044	1.956988
AX94665619	7A	222939896	<i>TaCLC-7A1</i>	216343576	6.59632
AX94566038	7A	683488235	<i>TaNPF5-7A3</i>	692626752	9.138517
AX95178548	7B	112337703	<i>TaNPF5-7B2</i>	116046396	3.708693
AX94532247	7B	524619772	<i>TaNPF8-7B1- TaNPF8-7B3</i>	517623485-518338133	6.281639-6.996287
AX94424632	7B	562740463	<i>TaNPF8-7B4 - TaNPF8-7B6</i>	556686657-558639959	4.100504-6.053806
AX94880654	7B	592345313	<i>TaNRT2-7B1</i>	583923053	8.42226
AX94553632	7B	633318727	<i>TaNPF5-7B4</i>	624468356	8.850371
AX94781629	7D	206695834	<i>TaCLC-7D1</i>	204246408	2.449426
AX94678472	7D	487753483	<i>TaNPF8-7D2 - TaNPF8-7D4</i>	489153269 - 489673028	1.39978- 1.91954
AX95080011	7D	592191388	<i>TaNPF5-7D4</i>	600836846	8.645458

Figures

Figure 1

Genome wide distribution of TaNPF, TaNRT2, TaCLC and TaSLAC1/TaSLAH genes in hexaploid wheat

Figure 2
Phylogenetic tree depicting relationship between a) TaNPF and TaNRT2 genes in hexaploid wheat and Arabidopsis thaliana b) TaCLC genes in wheat, rice and Arabidopsis thaliana c) TaSLAC1/SLAH genes in hexaploid wheat, rice and Arabidopsis thaliana

Figure 3
Page 21/23

Protein structure prediction: i) representative structures of TaNPF genes (A-H), TaNRT2 genes (I) TaCLC genes (J) and TaSLAC1/TaSLAH genes (K). ii) Representative TMs structures of nitrate transporters containing A)14 TMs, B) 13 TMs C) 12 TMs and D) CLC proteins containing 10 TMs and a re-entrant helix.

Figure 4 Expression patterns of nitrate transporter gene triads in wheat a) Tissue and development stage specific expression profiles of TaNPF and TaNRT2 genes b) Tissue and development stage specific expression profiles of TaCLC and TaSLAC1/SLAH genes

Figure 5 Triad expression of nitrate transporters in wheat A) Overall triad expression of all nitrate transporter genes B) Tissue specific triad expression of nitrate transporter genes

Figure 6

Syntenic relationships of wheat nitrate transporter genes orthologous with A) *A. thaliana*, B) *O. sativa*, C) *T. urartu*, D) *Ae. tauschii*, E) *T. dicoccoides* and, F) *T. turgidum*

Supplementary Files

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

- [SupplementaryFigures.docx](#)
- [Supplementarytables.xlsx](#)