



Review Article

NTDBPlant: Nutrient Transporters Database for Plants

Farrukh Azeem^{1*}, Arooj Mumtaz², Khushbakht Aliya¹, Faiza Maqbool¹, Tunjeena Rehan¹, Sumaira Rasul³ and Muhammad Amjad Ali^{4*}

¹Department of Bioinformatics and Biotechnology, GC University, Faisalabad, Pakistan

²Virtual University, Lahore, Punjab, Pakistan

³Institute of Molecular Biology and Biotechnology, Bahauddin Zakariya University, Multan, Pakistan

⁴Department of Plant Pathology and Centre of Agricultural Biochemistry and Biotechnology, University of Agriculture, Faisalabad, Pakistan

*For correspondence: azeuaf@hotmail.com; amjad.ali@uaf.edu.pk

Abstract

Ion transporters are the proteins localized in the plasma membrane and play a vital role in the transport of ions, maintenance of pH and cell expansion. NTDBPlant is a database that provides a web interface in order to access information related to ion transporter genes in *Arabidopsis thaliana*, a small flowering plant. This database contains data of genomic sequences, coding sequences, transcriptomic data, protein sequences, domain annotations, scientific literature and external links to other databases for 233 transport-related genes from 37 gene families in genus *Arabidopsis*. The current database is the very first of its kind for plants. It will provide significant information on a single platform related to the ion transporters and is publicly available through <http://ntdbplant.net/>. In future, more data related to other ion transport related gene families will be added and will be helpful for the researchers around the globe. © 2018 Friends Science Publishers

Keywords: Ion transport; *Arabidopsis thaliana*; Channels; Transporters; Database

Introduction

Nutrient transport proteins play an important role in the supply of energy, uptake and sequestration of ions into vacuoles, in cell expansion and maintenance of pH in the cell environment. Plants acquire mineral nutrients from the soil (Grennan, 2009). The plant cell plasma membrane (PM) has a selective permeability that confirms selective metabolites and ions to enter into the cell. Together with tonoplast, the selectively permeable plasma membrane of plant cell facilitates cytoplasm to sustain intracellular homeostasis. These membranes are comprised of phospholipid bilayers along with transmembrane proteins and enable the passage of metabolites, ions and water molecules in order to maintain the pH of cytosol. Channel proteins, ATP powered pumps, co-transporters are required for the transport of organic solutes, inorganic ions and protons across the plasma membrane (Sze, 1999).

Electrophysiological studies revealed the presence of various transport pathways like ion channels and ion carriers for the uptake and translocation of cations. Identification of transporters at gene level revealed involvement of a number of genes and gene families in uptake and distribution of specific ions (Rodríguez-Navarro, 2000; Nieves-Cordones *et al.*, 2016; Azeem *et al.*, 2018). Analysis of *Arabidopsis thaliana* genome has revealed that gene families encoding

for particular ion transporters are surprisingly large (Hawkesford, 2003) (Table 1).

Earlier several databases have been developed for multiple plant species such as PlnTFDB (Pérez-Rodríguez *et al.*, 2010), PlanTAPDB (Richardt *et al.*, 2007), GRASSIUS (Yilmaz *et al.*, 2009), LegumeTFDB (Mochida *et al.*, 2010), TreeTFDB (Mochida *et al.*, 2013) and DATFAP (Fredslund, 2008). Similarly, databases like AGRIS (Yilmaz *et al.*, 2011), TOBFAC (Rushton *et al.*, 2008) and SoyDB (Wang *et al.*, 2010) are available for individual organisms. Most of these databases are publicly available for identification and annotation of plant TFs. By keeping in mind the importance and vast use of these databases, we have developed a database named as Nutrient Transporters database for plants “NTDBPlant” (<http://ntdbplant.net>), that provides complete primary information related to ion transporter gene families in *Arabidopsis*. It is first of its kind database that provides basic data related to nutrient transporter genes at one place. It could be a promising resource of information for the understanding of ion transporter gene families at a single forum. The NTDBPlant contains 233 gene members from 37 gene families of *Arabidopsis thaliana*. To provide comprehensive information for each gene member involved in nutrient transport, we have provided its DNA sequence, cDNA sequence, protein sequence, thousand base pairs promoter sequence and the CDS sequence.

Table 1: Basic information of 37 nutrient transporters gene families of *Arabidopsis thaliana* currently present in NTDBPlant database

Ion Transporter Gene families	Members	Ion Transporter Gene families	Members
Shaker family	9	Putative zinc transporter ZIP2 - like family	2
KCO-like protein 3 family	1	Putative potassium transporter family	3
KCO5 protein family	5	Fe(II) transporter isolog family	11
Glutamate Receptor (ATGLR) Family**	20	High affinity nitrate transporter family	7
Cyclic nucleotide gated channel family	20	Nramp2 family	6
Anion channel protein family	5	Putative sulfate transporter family	12
Putative calcium channel	1	Putative sulfate transporter family	2
Cation-chloride co-transporter family	2	Zinc transporter (ZAT) family	4
Ammonium transporter family	5	Putative potassium transporter family	3
Inorganic phosphate transporter family	9	Copper transport protein-like family	5
Ammonium transporter	1	Sodium transporter	1
<i>Arabidopsis thaliana</i> Antiporter Superfamily	28	Putative phosphate permease	1
Phosphate/triose-phosphate translocatorfamily	8	Putative Na ⁺ /H ⁺ antiporter family	26
Low affinity calcium antiporter CAX2 family	6	Putative Na ⁺ /H ⁺ antiporter family	2
Sodium proton exchanger family	4	Putative potassium proton antiporter family	3
Putative anion exchange protein family	7	Na ⁺ /H ⁺ antiporter-Putative family	2
ADP, ATP carrier-like Protein family	4	2-oxoglutarate/malate translocator precursor-like	3
sodium-calcium exchanger like-family	3	Putative Na ⁺ /Ca ²⁺ antiporter	1
Putative carnitine/acylcarnitinetranslocase	1	Potassium-dependent sodium-calcium exchanger like-family	3

Both the sequences and detailed annotations for each gene are available at NTDBPlant. We have also identified functional domains in proteins to aid functional analysis and annotations of various nutrient transport proteins.

Data Sources and Methods

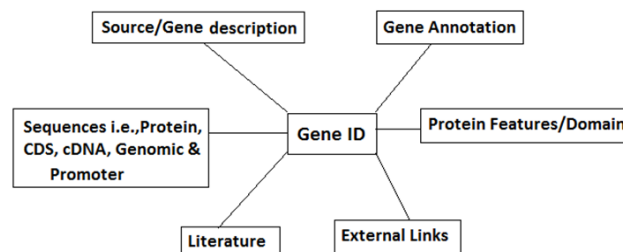
The main steps involved in the construction of NTDBPlant database include data integration, nutrient transporters sequence information, and gene annotations.

Sequence Data

The sequences for each gene member *i.e.*, DNA sequence, cDNA sequence, protein sequence, CDS sequence and a thousand base pair promoter sequence were downloaded from The Arabidopsis Information Resource (TAIR, <https://www.arabidopsis.org/>), Joint Genome Institute (JGI, <https://phytozome.jgi.doe.gov/pz/portal.html>) and National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>).

Gene Annotations

In order to offer comprehensive information for the different nutrient transport genes, we made extensive annotations at family and gene levels. On the page having list of ion transporter families, for each ion transporter family, a brief introduction and key references were listed. The annotations for each gene were made accessible and were gathered from multiple databases which include Kyoto Encyclopedia of Genes and Genomes (KEGG, <http://www.genome.jp/kegg/>), Gene Ontology Consortium (GO, <http://www.geneontology.org/>), Protein Analysis Through Evolutionary Relationships (Panther, <http://www.pantherdb.org/>) and PFAM (<https://pfam.xfam.org/>). The database also includes the detailed protein information. Gene location and protein lengths were also provided based on Phytozome and NCBI information.

**Fig. 1:** Flowchart shows the basic information required for each gene member in NTDBPlant database

The domains in each protein sequence are listed in NTDBPlant which were primarily gathered from Uniprot (www.uniprot.org/).

The gene descriptions (Fig. 1) were collected from NCBI (<https://www.ncbi.nlm.nih.gov/>), TAIR (<https://www.arabidopsis.org/>), Phytozome (<https://phytozome.jgi.doe.gov>) Ensembl plants (<https://plants.ensembl.org/index.html>) and PGSB (<http://pgsb.helmholtz-muenchen.de/plant/index.jsp>) databases. NTDBPlant database also contains related literature about each gene member of all the transporter families. Additionally, the database was linked to the other resources such as TAIR (<https://www.arabidopsis.org/>), phytozome (<https://phytozome.jgi.doe.gov>), NCBI (<https://www.ncbi.nlm.nih.gov/>), Ensembl plants (<https://plants.ensembl.org/index.html>), PGSB (<http://pgsb.helmholtz-muenchen.de/plant/index.jsp>), PGDD (<http://chibba.agtec.uga.edu/duplication>), e-FP browser (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>), Bio-grid (<https://thebiogrid.org/>), Atted-II (<http://atted.jp/>), Plaza (<https://bioinformatics.psb.ugent.be/plaza>), Gene visible (<https://genevisible.com/search>), plant genome database (<http://www.plantgdb.org>) and many other databases.

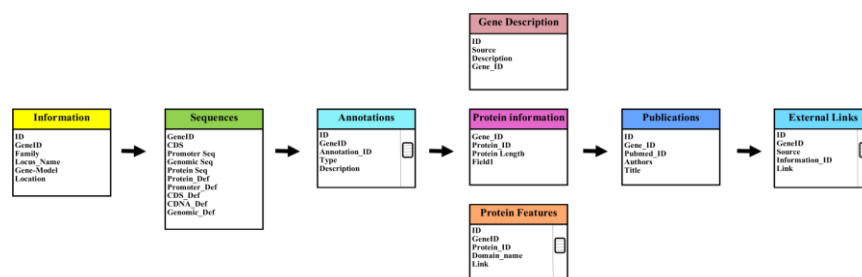


Fig. 2: The flowchart for construction of NTDBPlant

The screenshot shows the NTDB Plant website interface. At the top, there is a navigation bar with links: HOME, BROWSE, SEARCH, ABOUT, and CONTACT US. Below the navigation bar is a large image of a plant with white flowers. Under the image, there are two main sections: "BROWSE BY CHROMOSOME" and "BROWSE BY FAMILY". The "BROWSE BY CHROMOSOME" section lists Chromosome 1 (36), Chromosome 2 (29), Chromosome 3 (26), Chromosome 4 (20), and Chromosome 5 (48). The "BROWSE BY FAMILY" section lists several gene families, including Potassium-dependent sodium-calcium exchanger like-family (3), 2-oxoglutarate/malate translocator precursor-like (3), ADP, ATP carrier-like Protein family (4), Anion channel protein family (5), Arabidopsis thaliana Antiporter Superfamily (11), Copper transport protein-like family (4), Cyclic nucleotide gated channel family (20), and Fe(II) transporter isolog family (11). A green arrow points from the "Potassium-dependent sodium-calcium exchanger like-family (3)" entry to a detailed view of this gene family. The detailed view shows the following genes:

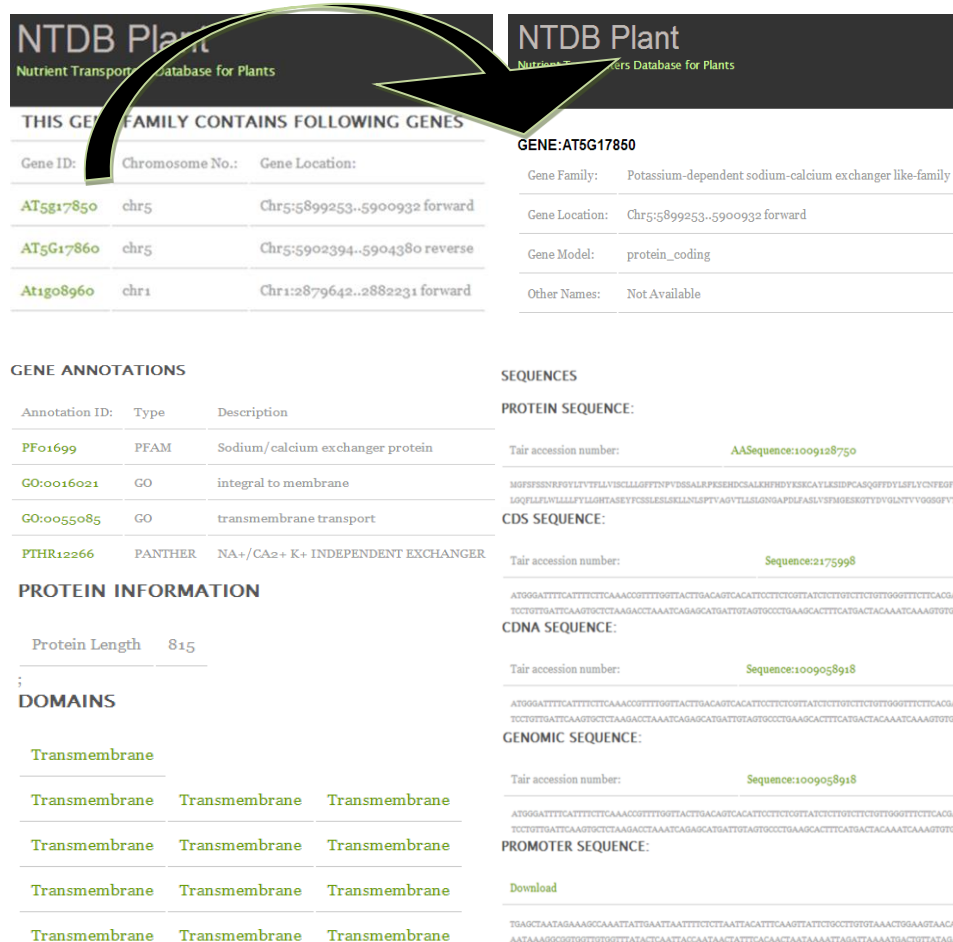
Gene ID:	Chromosome No.:	Gene Location:
AT5g17850	chr5	Chr5:5899253..5900932 forward
AT5G17860	chr5	Chr5:5902394..5904380 reverse
At1g08960	chr1	Chr1:2879642..2882231 forward

Fig. 3: Screenshot of a web page displaying information about nutrient transporters gene families in NTDBPlant. Web page in NTDBPlant displays basic information about gene families to search either by chromosome or by a gene family

Web Construction

We used relational database model for the construction of NTDBPlant database in Microsoft Access 2007. The data have been divided into 9 relations (Fig. 2) named as Gene_Description, Other names, Protein_information, Protein_Features, Annotations, Gene Accessions, Sequences,

Publications and External_links. Each of these 9 tables is connected with main relation "Information". For each gene, information relation stores Gene ID, chromosomal location, gene model, its family and family abbreviations for query purpose. Relation Gene_description stores description of each gene as reported in published biological databases with a source of information.



NTDB Plant
Nutrient Transporters Database for Plants

THIS GENE FAMILY CONTAINS FOLLOWING GENES

Gene ID:	Chromosome No.:	Gene Location:
AT5g17850	chr5	Chr5:5899253..5900932 forward
AT5G17860	chr5	Chr5:5902394..5904380 reverse
At1g08960	chr1	Chr1:2879642..2882231 forward

GENE: AT5G17850

Gene Family: Potassium-dependent sodium-calcium exchanger like-family

Gene Location: Chr5:5899253..5900932 forward

Gene Model: protein_coding

Other Names: Not Available

GENE ANNOTATIONS

Annotation ID:	Type	Description
PF01699	PFAM	Sodium/calcium exchanger protein
GO:0016021	GO	integral to membrane
GO:0055085	GO	transmembrane transport
PTHR12266	PANTHER	NA ⁺ /CA ²⁺ K ⁺ INDEPENDENT EXCHANGER

PROTEIN INFORMATION

Protein Length	815
;	

DOMAINS

Transmembrane		
Transmembrane	Transmembrane	Transmembrane
Transmembrane	Transmembrane	Transmembrane
Transmembrane	Transmembrane	Transmembrane
Transmembrane	Transmembrane	Transmembrane

SEQUENCES

PROTEIN SEQUENCE:

Tair accession number: AASquence:1009128750

M3FSSNNFQVLYTLLVYBCLLGGFTNTPYDSSALRPSKHEDCALQKHFDYKSEKATLSEIDPCASQFFDYLSFLYCNFEGPI
LQQLFLVYLLLYLGGTASEYFCSLESLKLLNLSPTVAGVITLLENGNAPDLFASLYTFMGESEKQTVDTGLNTVYGGSGFTYC

CDS SEQUENCE:

Tair accession number: Sequence:2175998

ATGGAATTTCATTCTTCAACCGTTTGGTACTGACAGTCACATTCCTCTGTTATCTCTGCTCTGTTGGTTCTTCACGAA
TCTGTTGATTCAGATTCCTCAAGACCTAAATCAGAGCATGATTGTAGTGCCTGAAGCATCTTCATGACTACAAATCAAGTGTCT

CDNA SEQUENCE:

Tair accession number: Sequence:1009058918

ATGGAATTTCATTCTTCAACCGTTTGGTACTGACAGTCACATTCCTCTGTTATCTCTGCTCTGTTGGTTCTTCACGAA
TCTGTTGATTCAGATTCCTCAAGACCTAAATCAGAGCATGATTGTAGTGCCTGAAGCATCTTCATGACTACAAATCAAGTGTCT

GENOMIC SEQUENCE:

Tair accession number: Sequence:1009058918

ATGGAATTTCATTCTTCAACCGTTTGGTACTGACAGTCACATTCCTCTGTTATCTCTGCTCTGTTGGTTCTTCACGAA
TCTGTTGATTCAGATTCCTCAAGACCTAAATCAGAGCATGATTGTAGTGCCTGAAGCATCTTCATGACTACAAATCAAGTGTCT

PROMOTER SEQUENCE:

Download

TGAGCTAATAGAAAGCAAAATTTGAATTAATTTCTCTAATTACATTCAGATTCCTGCTGTTTAACTGGAAATGACAG
AATAAAGGCGGTTGGTGTATCTCACTACCAATCAATCTTCACCACTAATAAATGATTAATAATGCTGTTATAGAT

Fig. 4: Every gene page in NTDBPlant displays the basic information for a typical nutrient transporter gene. The annotation information of a typical entry of the NTDBPlant displaying the rich annotation of a nutrient transporter gene (AT5g17850) of the Potassium-dependent sodium-calcium exchanger like-family. The annotation contains four major categories: (A) Basic information (B) Gene Annotations (C) Protein Information: domains are linked to the original domain databases and protein length is also provided for each gene member (D) Sequence Information: all sequence information with TAIR accession numbers is provided

Similarly, relation Annotation has been used to store different functional annotations regarding ion transporter gene with a brief description as mentioned in different sources (*i.e.*, PFAM, a database of protein families), KEGG (Kyoto Encyclopedia of Genes and Genomes), PANTHER (Protein Analysis Through Evolutionary Relationships), annotation ID and navigation links. Associated protein features such as protein length, domains and external links have been stored in two relations Protein_information and Protein_features. Two relations “Accessions” and “Sequences” contain protein, cDNA, CDS, genomic and promoter sequences and their unique accession numbers as specified in TAIR database. To provide references to original work relation “Publications” have been used which stores publication details (*i.e.*, paper title, journal, authors, PUBMED ID). However, we used relation External links to provide external database navigational links. This relation provides minimum 15 external sources for each gene.

Implementation and Web Interface

Apache Server (LAMP) has been used for the implementation and testing NTDBPlant. Web languages such as HTML, PHP, JAVA, C++, CSS and MySQL have been used to design freely accessible web interface. NTDBPlant provides two browsing facilities, browse by chromosome and browse by family (Fig. 3). Browse by chromosome allows the user to have a view of different gene families on a particular chromosome with a number of genes as recorded in the database. Browse by family provides an overview of all gene families in our database and a navigation link. This link provides all genes with chromosomal position in a respective gene family (Fig. 4). Search interface facilitates three types of inquiries, Search by chromosome, Search by Family and Search by Gene. All of these search options ultimately takes the user to a Gene information interface. This Gene information interface provides gene location, other

names used for gene, brief description of a gene, annotations, related protein domains and sequences (protein, cDNA, CDS, genomic and promoter) which can be downloaded, related publications and links to external databases such as TAIR, T-DNA Express, Plant Genome and Systems Biology, Ace view, Plant Proteome Database, Phytozome, Salk SNP Viewer, Plaza, NCBI-Entrez Gene, BioGRID, Araport, AtGDB View, e-FP Browser and EVEX.

Future Perspectives

There has been an increasing trend of transporter gene identification that are further isolated and then cloned to reveal understandings of the mechanism of ion transport across the membrane. The findings are handy for further research. The database we developed is just a beginning of an important resource. We intend to gradually expand the database with membrane transport gene families of other plant genomes as well.

References

- Azeem, F., B. Ahmad, R.M. Atif, M.A. Ali, H. Nadeem, S. Hussain, H. Manzoor, M. Azeem and M. Afzal, 2018. Genome-wide analysis of potassium transport-related genes in chickpea (*Cicer arietinum* L.) and their role in abiotic stress responses. *Plant Mol. Biol. Rep.*, 36: 451–468
- Fredslund, J., 2008. DATFAP: A database of primers and homology alignments for transcription factors from 13 plant species. *BMC Genomics*, 9: 140
- Grennan, A.K., 2009. Identification of genes involved in metal transport in plants. *Plant Physiol.*, 149: 1623–1624
- Hawkesford, M.J., 2003. Transporter gene families in plants: the sulphate transporter gene family - redundancy or specialization? *Physiol. Plantarum*, 117: 155–163
- Mochida, K., T. Yoshida, T. Sakurai, K. Yamaguchi-Shinozaki, K. Shinozaki and L.S.P. Tran, 2013. TreeTFDB: An integrative database of the transcription factors from six economically important tree crops for functional predictions and comparative and functional genomics. *DNA Res.*, 20: 151–162
- Mochida, K., T. Yoshida, T. Sakurai, K. Yamaguchi-Shinozaki, K. Shinozaki and L.S.P. Tran, 2010. LegumeTFDB: an integrative database of *Glycine max*, *Lotus japonicus* and *Medicago truncatula* transcription factors. *Bioinformatics*, 26: 290–291
- Nieves-Cordones, M., V. Martínez, B. Benito and F. Rubio, 2016. Comparison between arabidopsis and rice for main pathways of K⁺ and Na⁺ uptake by roots. *Front. Plant Sci.*, 7: 992
- Pérez-Rodríguez, P., D.M. Riaño-Pachón, L.G.G. Corrêa, S.A. Rensing, B. Kersten and Mueller-Roeber, 2010. PlnTFDB: updated content and new features of the plant transcription factor database. *Nucl. Acids Res.*, 38: 822–827
- Richardt, S., D. Lang, R. Reski, W. Frank and S.A. Rensing, 2007. PlanTAPDB, a phylogeny-based resource of plant transcription-associated proteins. *Plant Physiol.*, 143: 1452–1466
- Rodríguez-Navarro, A., 2000. Potassium transport in fungi and plants. *Biochim. Biophys. Acta*, 1469: 1–30
- Rushton, P.J., M.T. Bokowiec, T.W. Laudeman, J.F. Brannock, X. Chen and M.P. Timko, 2008. TOBFAC: the database of tobacco transcription factors. *BMC Bioinform.*, 9: 53
- Sze, H., 1999. Energization of plant cell membranes by H⁺-pumping ATPases: regulation and biosynthesis. *Plant Cell Online*, 11: 677–690
- Wang, Z., M. Libault, T. Joshi, B. Valliyodan, H.T. Nguyen, D. Xu, G. Stacey and J. Cheng, 2010. SoyDB: a knowledge database of soybean transcription factors. *BMC Plant Biol.*, 10: 14
- Yilmaz, A., M.K. Mejia-Guerra, K. Kurz, X. Liang, L. Welch and E. Grotewold, 2011. AGRIS: the arabidopsis gene regulatory information server, an update. *Nucl. Acids Res.*, 39: 1118–1122
- Yilmaz, A., M.Y. Nishiyama, B.G. Fuentes, G.M. Souza, D. Janies, J. Gray and E. Grotewold, 2009. GRASSIUS: a platform for comparative regulatory genomics across the grasses. *Plant Physiol.*, 149: 171–180

(Received 19 March 2018; Accepted 07 August 2018)