

RESEARCH ARTICLE

# CTDB: An Integrated Chickpea Transcriptome Database for Functional and Applied Genomics

Mohit Verma<sup>☉</sup>, Vinay Kumar<sup>☉</sup>, Ravi K. Patel, Rohini Garg, Mukesh Jain\*

Functional and Applied Genomics Laboratory, National Institute of Plant Genome Research (NIPGR), New Delhi, India

☉ These authors contributed equally to this work.

\* [mjain@nipgr.ac.in](mailto:mjain@nipgr.ac.in)



**OPEN ACCESS**

**Citation:** Verma M, Kumar V, Patel RK, Garg R, Jain M (2015) CTDB: An Integrated Chickpea Transcriptome Database for Functional and Applied Genomics. PLoS ONE 10(8): e0136880. doi:10.1371/journal.pone.0136880

**Editor:** Nicholas James Provart, University of Toronto, CANADA

**Received:** June 11, 2015

**Accepted:** August 10, 2015

**Published:** August 31, 2015

**Copyright:** © 2015 Verma et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** The financial support was received from the Department of Biotechnology, Government of India, New Delhi, under the Next Generation Challenge Program on Chickpea Genomics (grant number BT/PR12919/AGR/02/676/2009 from 2009-2014).

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Chickpea is an important grain legume used as a rich source of protein in human diet. The narrow genetic diversity and limited availability of genomic resources are the major constraints in implementing breeding strategies and biotechnological interventions for genetic enhancement of chickpea. We developed an integrated Chickpea Transcriptome Database (CTDB), which provides the comprehensive web interface for visualization and easy retrieval of transcriptome data in chickpea. The database features many tools for similarity search, functional annotation (putative function, PFAM domain and gene ontology) search and comparative gene expression analysis. The current release of CTDB (v2.0) hosts transcriptome datasets with high quality functional annotation from cultivated (desi and kabuli types) and wild chickpea. A catalog of transcription factor families and their expression profiles in chickpea are available in the database. The gene expression data have been integrated to study the expression profiles of chickpea transcripts in major tissues/organs and various stages of flower development. The utilities, such as similarity search, ortholog identification and comparative gene expression have also been implemented in the database to facilitate comparative genomic studies among different legumes and Arabidopsis. Furthermore, the CTDB represents a resource for the discovery of functional molecular markers (microsatellites and single nucleotide polymorphisms) between different chickpea types. We anticipate that integrated information content of this database will accelerate the functional and applied genomic research for improvement of chickpea. The CTDB web service is freely available at <http://nipgr.res.in/ctdb.html>.

## Introduction

Chickpea is an agriculturally important legume crop, which ranks second in cultivated area and third in production among legumes worldwide. It is a good source of proteins and dietary fibers in human diet. Similar to other members of leguminosae family, chickpea can restore the soil fertility by fixing atmospheric nitrogen. India ranks first in production as well as

consumption of chickpea (FAOSTAT, 2012; <http://faostat.fao.org>). The chickpea productivity is substantially low worldwide and affected by several biotic and abiotic factors. Hence, the development of high-yielding chickpea varieties is a pressing need in order to meet demands of an overgrowing population. However, due to limited availability of genomic and genetic resources, the progress in the improvement of chickpea yield has been limited in the past years.

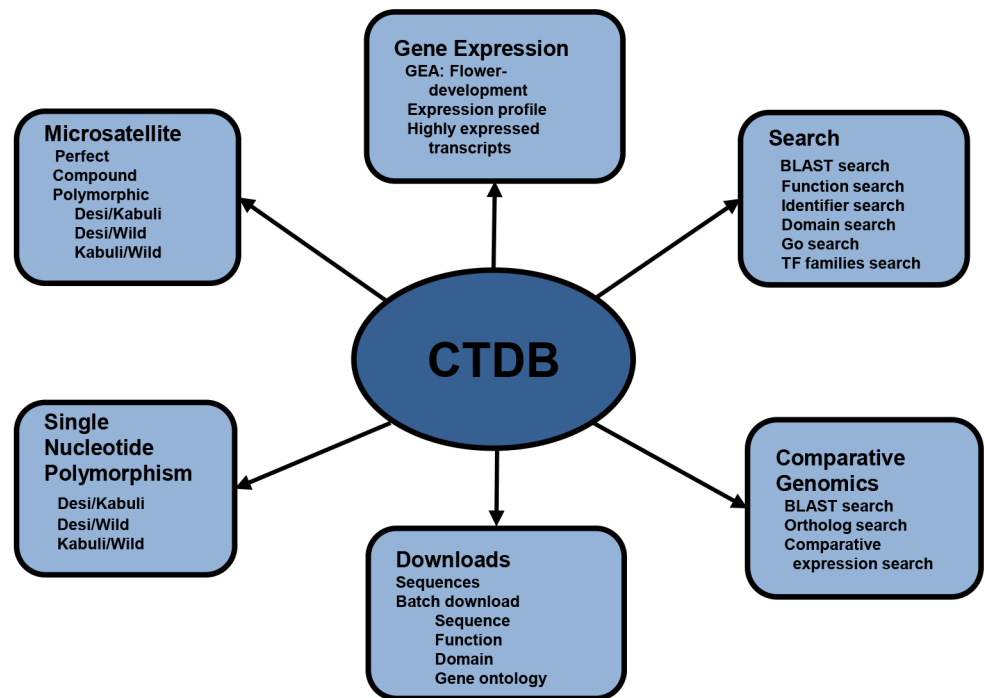
Over the past three decades, extensive research has been carried out for the chickpea improvement. However, the narrow genetic diversity in cultivated and wild accessions restricts the genetic improvement by limiting introduction of diverse germplasm using conventional/molecular-breeding approaches. The cultivated chickpeas are categorized into kabuli and desi types based on phenotypic appearance of the seed (size, shape and color). Wild chickpea (*Cicer reticulatum*) is the progenitor of domesticated chickpea (*C. arietinum*) and can serve as a potential source of genes responsible for several agronomically important traits [1]. The availability of next generation sequencing technologies has catalyzed the genomics research in chickpea. The draft genome sequences of kabuli (CDC frontier) and desi (ICC4958) chickpea genotypes have been released [2,3]. The transcriptome datasets of wild, kabuli and desi type chickpea have been analyzed to provide resources for gene discovery and development of functional molecular markers [4–9]. Furthermore, gene expression studies from different tissues/organs and/or environmental conditions have also been performed to reveal the putative functions of chickpea transcripts [4,6,10,11]. Recently, two microsatellite databases, CicArMiSatDB [12] and CMsDB (<http://nipgr.res.in/CMsDB.html>), have been developed to identify functional microsatellite markers in the chickpea genome for various applications. The availability of these high-throughput data has opened several avenues for genetics and genomics research in chickpea. In addition, a few web resources, such as SoySeq, SoyPlex, MtGEA, LjGEA and LegumeIP, are available for functional and/or comparative genomics in other legumes [13–18]. Most of these web resources/databases host mainly gene expression profile datasets.

Despite the extensive research in the development of genomics resources for chickpea, it is still difficult for a biologist to mine these data manually to parse out the relevant biological information. To make the information retrieval easy, we developed the Chickpea Transcriptome Database (CTDB) as an integrated web resource to provide comprehensive information on various aspects of transcriptome data from different (desi, kabuli and wild) chickpea genotypes. The database provides several user-friendly utilities for similarity search, functional annotation and gene expression data from different tissues/organs in chickpea and an option to conduct comparative genomics studies with/among different legumes. Furthermore, the database provides access to the functional molecular markers (microsatellites and single nucleotide polymorphisms) in the transcriptomes of different chickpea types. We envisage that the CTDB would serve as a potential resource for functional and applied genomics research in chickpea.

## Results and Discussion

### Database overview

In view of establishment of an integrated platform for comprehensive transcriptome analysis in chickpea, we have developed CTDB to provide public access to various aspects of transcriptome for wild, desi and kabuli chickpea genotypes [5, 7–8]. The CTDB has a user-friendly interface for the efficient retrieval and visualization of information of interest with the aid of multiple search options. The data retrieval system of CTDB has been hierarchically organized into five major sections (Search, Gene Expression, Microsatellite, SNPs and Comparative Genomics), with several utilities as shown in Fig 1. Using these options, users can sail through the database and visualize/browse and download various functional (gene annotation and TF



**Fig 1. The complete schema of the CTDB (v2.0) database.** Different modules with various utilities integrated in the database have been highlighted.

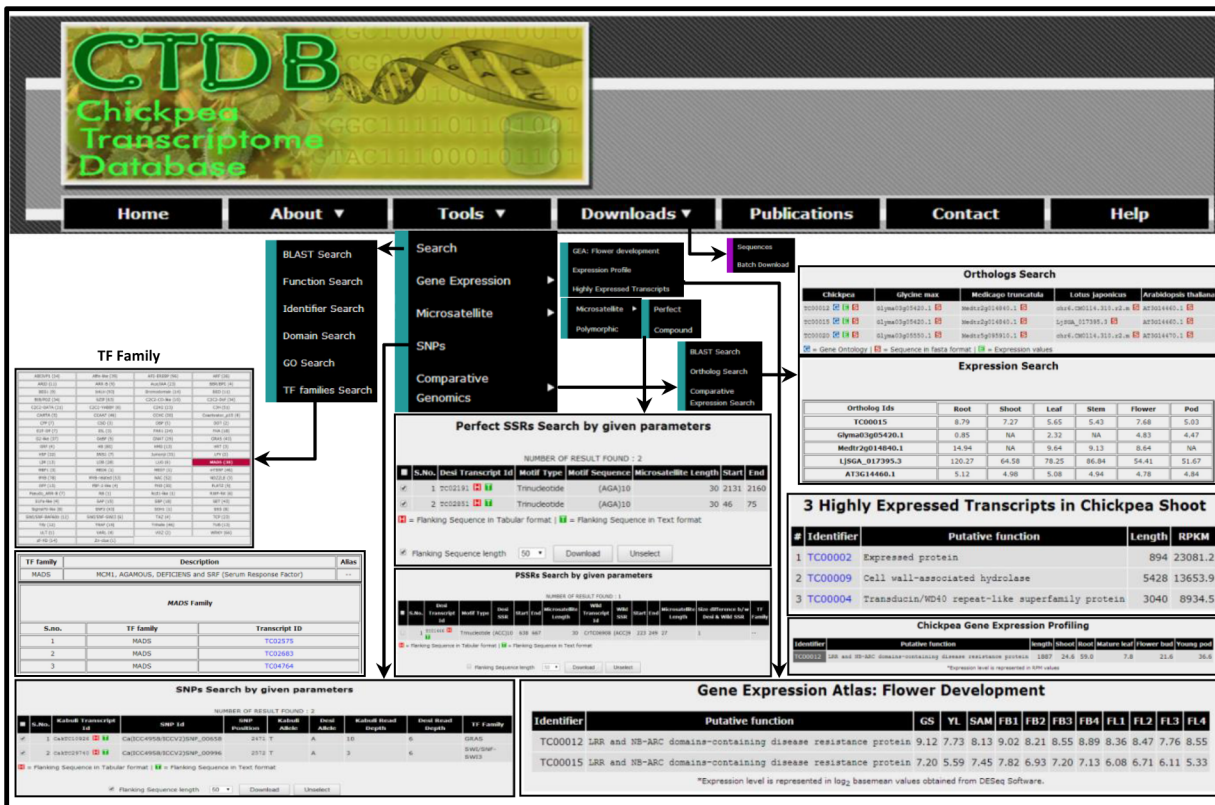
doi:10.1371/journal.pone.0136880.g001

families), behavioral (gene expression), and structural (microsatellite and SNP) features of the chickpea transcriptome (Fig 2). Further, CTDB allows comparative analysis of chickpea transcriptome with other legumes (*Glycine max*, *Medicago truncatula* and *Lotus japonicus*) and model plant *Arabidopsis*.

## Key features of the CTDB

**Similarity and function search.** The database provides various utilities under “Search” menu for data retrieval, similarity search and functional annotation of query transcript sequence(s).

- BLAST search allows user to find homologous transcript or protein sequences for the sequence(s) of interest in the selected chickpea genotype.
- Function search utility facilitates the retrieval of chickpea transcripts for a given keyword query, for instance proteases or RNA polymerases.
- The identifier search enables users to retrieve the putative function and gene ontology (GO) term(s) associated with a given transcript.
- Domain search offers the identification of a set of transcripts harboring the Pfam domain of interest. The interface also allows users to access the putative function, gene ontology (GO) and sequences in fasta format.
- GO search utility provides an option to explore the GO associated with a given transcript. Users can use either the transcript identifier as a query to retrieve the GO terms associated with the transcript or the GO identifier to list all transcripts linked to a GO term.



**Fig 2. Screenshots of the CTDB (v2.0) for searching, exploring and visualization for different queries.** The new/updated features/utilities, including TF family search, GEA: Flower development, Microsatellite, SNPs and Comparative Genomics are illustrated.

doi:10.1371/journal.pone.0136880.g002

- Transcription Factor (TF) family search provides access to various transcription factor families in chickpea. Each TF family links to a list of members and their expression profiles.

These search utilities can help to quickly query the chickpea transcriptome data for numerous functional genomics related applications.

**Gene expression.** The gene expression analysis can provide insights into the putative function of individual genes. Such analysis for a set of tissues/developmental stages has been performed in chickpea [4,11]. The CTDB offers three options for the gene expression search in chickpea; “GEA (gene expression atlas): flower development”, “Expression Profile” and “Highly Expressed Transcripts”.

The “GEA: Flower Development” option presents a resource to analyze the expression of gene(s)/transcript(s) during different developmental stages, including germinating seedling, young leaf, shoot apical meristem and eight stages of flower development (early flower bud to senescing flower) in chickpea [11]. Single or multiple transcript identifiers can be queried to study their expression profile(s) (Fig 2). The second gene expression analysis option, “Expression Profile”, provides expression information in major tissue samples (shoot, root, mature leaf, flower bud and young pod) for single/multiple transcript(s). Third option of “Highly Expressed Transcripts” has been provided to enlist most abundant transcripts in a particular tissue of chickpea. A hyperlink to the expression profile of each transcript has also been provided in the output. Overall, CTDB offers an opportunity to systematically identify expression patterns of individual or a set of transcripts in different tissues of chickpea. This tool provides an opportunity to the user(s) to choose appropriate gene(s) for functional analysis via reverse

genetics approaches. The expression data provided in the database have been used in previous studies for characterization of individual gene [19] and gene family [20], which demonstrate the confidence of scientific community in our database.

**Microsatellite search.** Microsatellites (simple sequence repeats, SSRs) are important molecular markers for the study of genetic variations among different cultivars using marker-assisted selection and mapping of quantitative trait loci [21–23]. The screening of polymorphic SSR markers using transcriptomes is one of the approach to investigate potential functional markers in chickpea [4, 8]. The CTDB provides an opportunity for the screening of microsatellites along with their physical location on the transcript(s) in different (desi, kabuli and wild) chickpea genotypes (Fig 2). The database provides separate retrieval systems for different microsatellite repeat types, including perfect, compound and polymorphic microsatellites. The interface for the perfect microsatellite search provides various parameters, such as chickpea genotype, repeat type, repeat sequence, repeat number, and length and position of the microsatellite(s). The compound microsatellite is a combination of more than one type of perfect repeats, which makes their search challenging. The CTDB offers an efficient query builder to search the compound microsatellites using different parameters, including repeat type, repeat sequence and number of repeat units. In addition, CTDB provides an access to the polymorphic SSRs between different chickpea genotypes. The database hosts a total of 130, 561 and 493 polymorphic SSRs between kabuli/desi, desi/wild and kabuli/wild, respectively. The user interface for all types of microsatellites offers an option to download downstream and upstream flanking sequences to the microsatellite, which can be used for designing PCR primers for genotyping applications. The utility of these data to discover the microsatellites for construction of the marker-trait linkages and the identification of genes associated with important traits has already been demonstrated [24]. Agarwal et al. [25] also exploited the CTDB resource for identification of novel genic SSRs and confirmed their applicability in a select set of 44 chickpea genotypes. Thus, the present resource of microsatellites is expected to accelerate chickpea improvement programmes via molecular breeding approaches.

**SNP search.** A search option to facilitate the discovery of SNPs between different genotypes of chickpea has also been integrated in the CTDB (Fig 2). The CTDB contains 1986, 36446 and 37954 high-quality SNPs between kabuli/desi, desi/wild and kabuli/wild genotypes, respectively using data from previous studies [7,8]. This information can assist accessing diversity between various genotypes. The search system is also equipped with various parameters to make the retrieval system user-friendly, including a specific base change (e.g. A/T, G/A), read depth, location of SNP in the transcript and SNPs in the members of a given TF family. The use of these search parameters makes the data mining flexible and more relevant. The output also provides an option to download the downstream and upstream flanking sequences to the SNPs for designing PCR primers as per user requirement. This module will prove to be an effective tool to enable marker design for associated traits in chickpea crop improvement programs.

**Comparative genomics analysis.** An interface has been integrated in the database to conduct a comparative analysis of chickpea transcripts with other plant species, including legumes (*G. max*, *M. truncatula*, *L. japonicas*) and model plant *Arabidopsis* (Fig 2). This option provides several utilities, such as BLAST search, ortholog search and comparative expression search modules. The “BLAST Search” page offers opportunity to conduct similarity search for protein/nucleotide sequence(s) of chickpea against other legumes (*G. max*, *M. truncatula* and *L. japonicus*) and *Arabidopsis*. The putative orthologs of chickpea transcripts in legumes and *Arabidopsis* can be retrieved using “Ortholog Search” along with an option to download their corresponding protein sequences. Another search module, “Comparative Expression Search” has been provided to conduct comparative expression analysis of chickpea transcripts and



their orthologs in legumes and Arabidopsis during different developmental stages. The CTDB data has been utilized for analyzing the influence of gene duplication and gene loss on evolution of heat shock transcription factor genes in legumes including chickpea [26]. This feature of CTDB will be a useful resource for the search of ortholog(s) and the comparative genomics analysis.

**Download and help.** The user can download the complete transcriptome of wild (PI 489777), kabuli (ICCV 2) and desi (ICC 4958) chickpea genotypes in fasta format. An additional tool, “Batch Download”, enables downloading of transcript data including the sequence, function, domain and GO for a given set of transcripts. In order to demonstrate the possible fundamental utilities of this database, we have contrived help option. This page provides tutorial and easy to follow step-wise instructions.

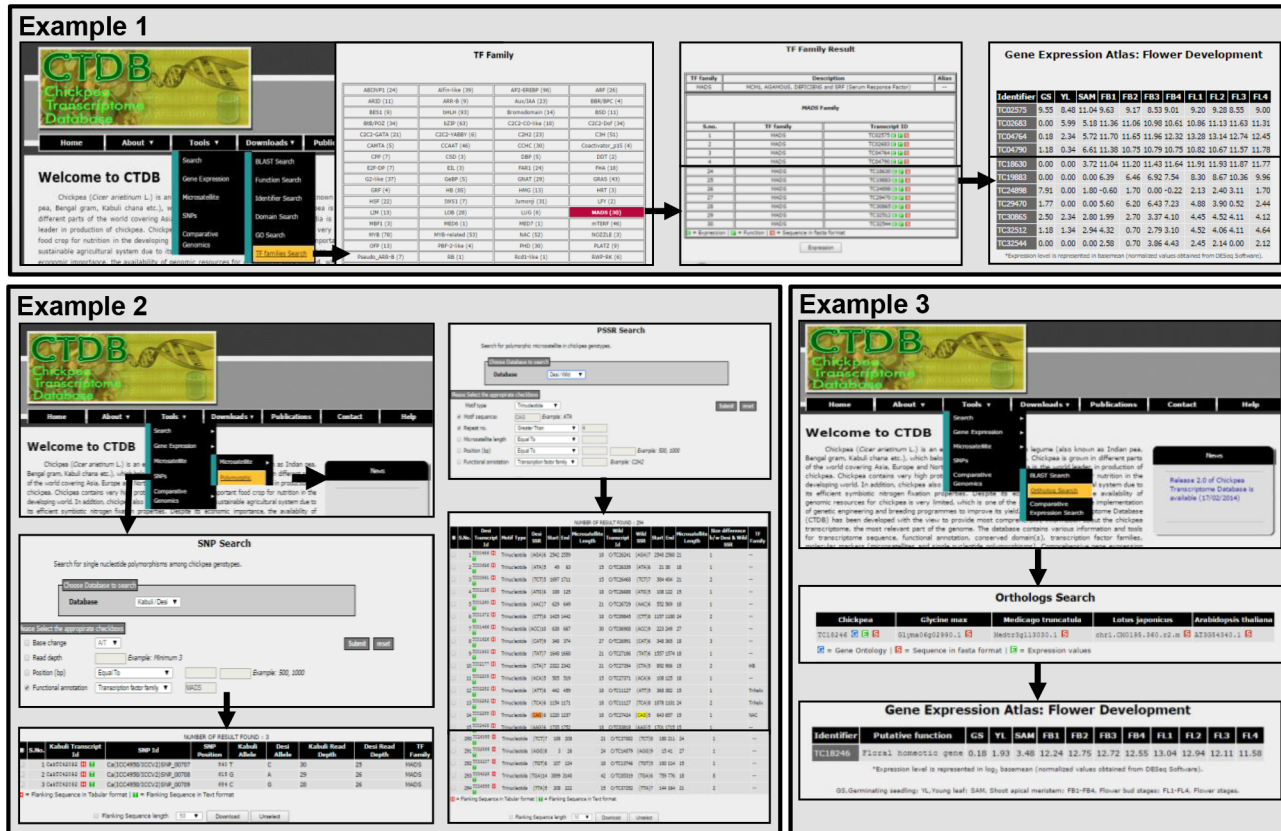
## Examples highlighting major applications of the database

To demonstrate the utility of the CTDB, we employed various features of the database for studying specific transcripts, discovery of polymorphic SSRs and SNPs lying in the transcripts and for performing comparative genomics analysis with other legumes and Arabidopsis (Fig 3).

**Example 1. Analysis of gene expression patterns of MADS-box transcription factor gene family members.** The expression analysis of candidate transcripts/genes during different tissues/developmental stages of chickpea is an important feature of the CTDB to facilitate their functional analysis. The demonstration of this utility has been presented by analyzing expression patterns of transcripts encoding MADS-box family transcription factors. Two steps are required to investigate the gene expression pattern of the transcripts of interest. In the first step, the CTDB provides a list of all TF families in chickpea with number of members under the “TF families search” tab. In the second step, a list of transcripts belonging to MADS-box family could be viewed by clicking on the hyperlink provided with name of MADS-box family in list of TF families. A hyperlink for each transcript ID, has been provided to access the respective expression profile during different tissues/developmental stages. An additional option at the bottom of table has also been provided to retrieve the expression profiles of all transcripts belonging to MADS-box family together during different tissues/developmental stages (Fig 3).

**Example 2. Discovery of polymorphic SSRs and SNPs between desi and wild chickpea genotypes.** This application has been presented with two examples to illustrate the utility of the CTDB in breeding approaches. The identification of polymorphic tri-nucleotide SSRs between desi and wild type genotypes is given as an example (Fig 3). A hyperlink of “Polymorphic” tab under “Microsatellite” enables screening of polymorphic SSRs with various search options. After selecting the “Desi and Wild” database, the customization of various parameters (motif type, trinucleotide; motif sequence, CAG; repeat number, greater than 4) can be carried out to retrieve the list of polymorphic SSRs of interest. The resulting page displays four polymorphic SSRs from desi transcript database with descriptions including their respective unique transcript identifier (TC02288, TC05305, TC05490 and TC06246), motif-type, repeat length, starting and end positions, length in desi and wild type and their related TF family (if any). The option to download the flanking sequence(s) of the selected SSRs of a particular length can be used to design PCR primers for genotyping applications.

The second example involves the identification of SNPs between desi and kabuli genotypes located within transcripts/genes of MADS-box family. The hyperlink of “SNPs” tab enables to access the search page for customization of the query. The selection of “Desi and Kabuli” as a database option along with relevant search options (function annotation; MADS), the result



## Future developments

The database will be updated regularly as and when the new data are available to expand utilities of CTDB. Additional features may also be included in the database to cater the demands of researchers in future.

## Availability and requirements

The CTDB 2.0 web service is freely available at <http://nipgr.res.in/ctdb.html>.

## Conclusions

In summary, the CTDB (v2.0) provides a user-friendly platform for many useful features/data to study the transcriptome of various chickpea genotypes via multiple search modules, expression analysis and comparative transcriptome analysis. The transcriptome datasets integrated in this database are potential source for the discovery of novel genes/transcripts in chickpea. In addition, this database presents the functional molecular markers (microsatellites and SNPs) to facilitate the molecular breeding programs. Altogether, this database is a significant advancement towards establishment of transcriptome resource for chickpea and is expected to accelerate functional and applied studies in chickpea and related legumes.

## Methods

### Database construction

The CTDB (v2.0) is presently hosted on a Sun Microsystem Workstation with two Intel Xeon quad core processors and 12 GB of random access memory which runs Linux operating system (Centos v5.5). MySQL relational database management system (v5.0.77) has been used to store and manage the processed transcriptome data [4]. The client web interface queries the Apache HTTP server (v5.5.29) to access the data from MySQL databases. Perl (v5.8.8), JavaScript (v1.6.0) and HTML (v4.0.1) based technologies have been employed to create the query-builder module for connecting user queries to the database.

### Data source, processing and implementation

The transcriptome datasets of wild [8], and kabuli [7] and desi [4] chickpea have been utilized for the construction of this database. The resources for the identification of putative function, similarity searching and prediction of transcripts encoding for transcript family have been utilized from previous study [4]. The functional molecular markers (microsatellites and SNPs) data have been utilized from previous studies [7,8]. The normalized expression data for chickpea transcripts from major tissues (reads per kilobase per million) and various stages of flower development ( $\log_2$  basemean value) were obtained from the previous studies [4,11] for integration into CTDB.

The genomic information for other legumes, *G. max*, *M. truncatula* and *L. japonicus* was retrieved from Phytozome v9.1 and Plant GDB databases. The Arabidopsis information resource (TAIR) was utilized to access genomic information of Arabidopsis. For comparative transcriptome analysis, orthologs of chickpea transcripts in other legumes and Arabidopsis were identified using reciprocal blast search ( $E$ -value  $\leq 1e-5$ ). The orthologs from different legumes and Arabidopsis were compiled and stored in MySQL. The normalized gene expression data of the selected tissues/organs for Arabidopsis, *G. max*, *M. truncatula* and *L. japonicus* were retrieved from TAIR ([ftp://ftp.arabidopsis.org/home/tair/Microarrays/analyzed\\_data/affy\\_data\\_1436\\_10132005.zip](ftp://ftp.arabidopsis.org/home/tair/Microarrays/analyzed_data/affy_data_1436_10132005.zip)), Soybean Atlas [14], MtGEA [13] and LjGEA [18], respectively. These data were tabulated and stored in the MySQL database.



## Acknowledgments

RG acknowledges the INSPIRE Faculty Award from the Department of Science and Technology, New Delhi.

## Author Contributions

Conceived and designed the experiments: MJ RG. Performed the experiments: MV RKP. Analyzed the data: MV VK RKP RG MJ. Contributed reagents/materials/analysis tools: MV RKP MJ. Wrote the paper: VK RKP RG MJ.

## References

1. Singh KB, Ocampo B. Exploitation of wild *Cicer* species for yield improvement in chickpea. *Theor Appl Genet* 1997; 95: 418–423.
2. Varshney RK, Song C, Saxcena RK, Azam S, Tu S, Sharpe AG, et al. Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat Biotechnol* 2013; 31: 240–246. doi: [10.1038/nbt.2491](https://doi.org/10.1038/nbt.2491) PMID: [23354103](https://pubmed.ncbi.nlm.nih.gov/23354103/)
3. Jain M, Misra G, Patel RK, Priya P, Jhanwar S, Khan AW, et al. A draft sequence of the pulse crop chickpea (*Cicer arietinum* L.). 2013; *Plant J* 74: 715–729. doi: [10.1111/tpj.12173](https://doi.org/10.1111/tpj.12173) PMID: [23489434](https://pubmed.ncbi.nlm.nih.gov/23489434/)
4. Garg R, Patel RK, Jhanwar S, Priya P, Bhattacharjee A, Yadav G, et al. Gene discovery and tissue-specific transcription analysis in chickpea with massively parallel pyrosequencing and web resource development. *Plant Physiol* 2011; 156: 1661–1678. doi: [10.1104/pp.111.178616](https://doi.org/10.1104/pp.111.178616) PMID: [21653784](https://pubmed.ncbi.nlm.nih.gov/21653784/)
5. Garg R, Patel RK, Tyagi AK, Jain M. *De novo* assembly of chickpea transcriptome using short reads for gene discovery and marker identification. *DNA Res* 2011; 18: 53–63. doi: [10.1093/dnares/dsq028](https://doi.org/10.1093/dnares/dsq028) PMID: [21217129](https://pubmed.ncbi.nlm.nih.gov/21217129/)
6. Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H, Tujeja R. Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Plant Biotechnol J* 2011; 9: 922–931. doi: [10.1111/j.1467-7652.2011.00625.x](https://doi.org/10.1111/j.1467-7652.2011.00625.x) PMID: [21615673](https://pubmed.ncbi.nlm.nih.gov/21615673/)
7. Agarwal G, Jhanwar S, Priya P, Singh VK, Saxena MS, Parida SK, et al. Comparative analysis of kabuli chickpea transcription with desi and wild chickpea provided a rich resource for development of functional markers. *PLoS ONE* 2012; 7: e52443. doi: [10.1371/journal.pone.0052443](https://doi.org/10.1371/journal.pone.0052443) PMID: [23300670](https://pubmed.ncbi.nlm.nih.gov/23300670/)
8. Jhanwar S, Priya P, Garg R, Prida SK, Tyagi AK, Jain M. Transcription sequencing of wild chickpea as a rich resource for marker development. *Plant Biotechnol J* 2012; 10: 690–702. doi: [10.1111/j.1467-7652.2012.00712.x](https://doi.org/10.1111/j.1467-7652.2012.00712.x) PMID: [22672127](https://pubmed.ncbi.nlm.nih.gov/22672127/)
9. Kudapa H, Azam S, Sharpe AG, Tara B, Li B, Deonovi B, et al. Comprehensive transcriptome assembly of chickpea (*Cicer arietinum* L.) using sanger and next generation sequencing platforms: development and applications. *PLoS ONE* 2014; 9: e86039. doi: [10.1371/journal.pone.0086039](https://doi.org/10.1371/journal.pone.0086039) PMID: [24465857](https://pubmed.ncbi.nlm.nih.gov/24465857/)
10. Garg R, Jain M. Pyrosequencing data reveals tissue-specific expression of lineage-specific transcription in chickpea. *Plant Signal Behav* 2011; 6: 1868–1870. doi: [10.4161/psb.6.11.17879](https://doi.org/10.4161/psb.6.11.17879) PMID: [22057340](https://pubmed.ncbi.nlm.nih.gov/22057340/)
11. Singh VK, Garg R, Jain M. A global view of transcription dynamics during flower development in chickpea by deep sequencing. *Plant Biotechnol J* 2013; 11: 691–701. doi: [10.1111/pbi.12059](https://doi.org/10.1111/pbi.12059) PMID: [23551980](https://pubmed.ncbi.nlm.nih.gov/23551980/)
12. Doddamani D, Katta MAVSK, Khan AW, Agarwal G, Shah TM, Varshney RK. CicArMiSatDB: the chickpea microsatellite database. *BMC Bioinformatics* 2014; 15: 212. doi: [10.1186/1471-2105-15-212](https://doi.org/10.1186/1471-2105-15-212) PMID: [24952649](https://pubmed.ncbi.nlm.nih.gov/24952649/)
13. He J, Benedito VA, Wang M, Murray JD, Zhao PX, Tang Y, et al. The *Medicago truncatula* gene expression atlas web server. *BMC Bioinformatics* 2009; 10:441. doi: [10.1186/1471-2105-10-441](https://doi.org/10.1186/1471-2105-10-441) PMID: [20028527](https://pubmed.ncbi.nlm.nih.gov/20028527/)
14. Libault M, Farmer A, Joshi T, Takahashi K, Langley RJ, Franklin LD, et al. An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J* 2010; 63:86–99. doi: [10.1111/j.1365-3113X.2010.04222.x](https://doi.org/10.1111/j.1365-3113X.2010.04222.x) PMID: [20408999](https://pubmed.ncbi.nlm.nih.gov/20408999/)
15. Severin AJ, Woody JL, Bolon YT, Joseph B, Diers BW, Farmer AD, et al. RNA-Seq atlas of *Glycine max*: A guide to the soybean transcriptome. *BMC Plant Biol* 2010; 10:160. doi: [10.1186/1471-2229-10-160](https://doi.org/10.1186/1471-2229-10-160) PMID: [20687943](https://pubmed.ncbi.nlm.nih.gov/20687943/)
16. Li J, Dai X, Liu T, Zhao PX Legume IP. An integrative database for comparative genomics and transcriptomics of model legumes. *Nucleic Acids Res* 2012; 40: D1221–D1229 doi: [10.1093/nar/gkr939](https://doi.org/10.1093/nar/gkr939) PMID: [22110036](https://pubmed.ncbi.nlm.nih.gov/22110036/)

17. Garg R, Jain M. Transcriptome analyses in legumes: A resource for functional genomics. *Plant Genome* 2013; 6, doi: [10.3835/plantgenome2013.04.0011](https://doi.org/10.3835/plantgenome2013.04.0011)
18. Verdier J, Torres-Jerez I, Wang M, Andriankaja A, Allen SN, He J, et al. Establishment of the *Lotus japonicus* Gene Expression Atlas (LjGEA) and its use to explore legume seed maturation. *Plant J* 2013; 74:351–362. doi: [10.1111/tbj.12119](https://doi.org/10.1111/tbj.12119) PMID: [23452239](https://pubmed.ncbi.nlm.nih.gov/23452239/)
19. Jaiswal DK, Mishra P, Subha P, Rathi D, Chakraborty S, Chakraborty N. Membrane-associated proteomics of chickpea identifies Sad1/UNC-84 protein (CaSUN1), a novel component of dehydration signaling. *Sci Rep* 2013; 4: 4177.
20. Ha CV, Esfahani MN, Watanabe Y, Tran UT, Sulieman S, Mochida K, et al. Genome-wide identification and expression analysis of the CaNAC family members in chickpea during development, dehydration and ABA treatments. *PLoS ONE* 2014; 9: e114107. doi: [10.1371/journal.pone.0114107](https://doi.org/10.1371/journal.pone.0114107) PMID: [25479253](https://pubmed.ncbi.nlm.nih.gov/25479253/)
21. Millan T, Clarke HJ, Siddique KHM, Buhariwalla HK, Gaur PM, Kumar J. Chickpea molecular breeding: new tools and concepts. *Euphytica* 2006; 147: 81–103.
22. Choudary P, Khanna SM, Jain PK, Bharadwaj C, Kumar J, Lakhera PC, et al. Genetic structure and diversity analysis of the primary gene pool of chickpea using SSR markers. *Genet Mol Res* 2012; 11: 891–905 doi: [10.4238/2012.April.10.5](https://doi.org/10.4238/2012.April.10.5) PMID: [22576917](https://pubmed.ncbi.nlm.nih.gov/22576917/)
23. Gaur PM, Jukanti AK, Varshney RK. Impact of genomic technologies on chickpea breeding strategies. *Agronomy* 2012; 2: 199–221.
24. Kujur A, Bajaj D, Saxena MS, Tripathi S, Upadhyaya D, Gowda CLL, et al. Functionally relevant microsatellite markers from chickpea transcription factor genes for efficient genotyping applications and trait association mapping. *DNA Res* 2013; 20: 355–373. doi: [10.1093/dnares/dst015](https://doi.org/10.1093/dnares/dst015) PMID: [23633531](https://pubmed.ncbi.nlm.nih.gov/23633531/)
25. Agarwal G, Sabbavarapu MM, Singh VK, Thudi M, Sheelamary S, Gaur PM et al. Identification of a non-redundant set of 202 in silico SSR markers and applicability of a select set in chickpea (*Cicer arietinum* L.). *Euphytica* 2015; doi: [10.1007/s10681-015-1394-3](https://doi.org/10.1007/s10681-015-1394-3)
26. Lin Y, Cheng Y, Jin J, Jin X, Jiang H, Yan H, et al. Genome duplication and gene loss affect the evolution of heat shock transcription factor genes in legumes. *PLoS ONE* 2014; 9: e102825. doi: [10.1371/journal.pone.0102825](https://doi.org/10.1371/journal.pone.0102825) PMID: [25047803](https://pubmed.ncbi.nlm.nih.gov/25047803/)
27. Roque E, Serwatowska J, Rochina MC, Wen J, Mysore K, Yenush L, et al. Functional specialization of duplicated AP3-like genes in *Medicago truncatula*. *Plant J* 2013; 73: 663–675. doi: [10.1111/tbj.12068](https://doi.org/10.1111/tbj.12068) PMID: [23146152](https://pubmed.ncbi.nlm.nih.gov/23146152/)