

## Article

# TinoTranscriptDB: A Database of Transcripts and Microsatellite Markers of *Tinospora cordifolia*, an Important Medicinal Plant

Rakesh Singh <sup>1,\*</sup>,<sup>†</sup> , Ajay Kumar Mahato <sup>2,†</sup>, Akshay Singh <sup>1</sup> , Rajesh Kumar <sup>1</sup>, Amit K. Singh <sup>1</sup>, Sundeep Kumar <sup>1</sup>, Soma S. Marla <sup>1</sup>, Ashok Kumar <sup>1</sup> and Nagendra K. Singh <sup>3</sup> 

<sup>1</sup> ICAR-National Bureau of Plant Genetic Resources, Pusa, New Delhi 110012, India

<sup>2</sup> The Centre for DNA Fingerprinting and Diagnostics, Hyderabad 500039, India

<sup>3</sup> ICAR-National Institute for Plant Biotechnology, Pusa, New Delhi 110012, India

\* Correspondence: rakesh.singh2@icar.gov.in; Tel.: +91-011-25802791; Fax: +91-011-25842495

† These authors contributed equally to this work.

**Abstract:** *Tinospora cordifolia*, commonly known as “Giloe” in India, is a shrub belonging to the family Menispermaceae. It is an important medicinal plant known for its antipyretic, anti-inflammatory, antispasmodic, and antidiabetic properties and is used in the treatment of jaundice, gout, and rheumatism. Despite its economic importance, the limited information related to its genomic resources prohibits its judicious exploitation through molecular breeding or biotechnological approaches. In this study, we generated a meta-transcriptome assembly of 43,090 non-redundant transcripts by merging the RNASeq data obtained from Roche 454 GS-FLX, and Illumina platforms, and report the first transcriptome-based database for simple sequence repeats and transcription factors (“Tino-TranscriptDB” (*Tinospora cordifolia* Transcriptome Database)). We annotated 26,716 (62%) of the total transcripts successfully from National Center for Biotechnology Information non-redundant protein (NCBI-NR), gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-Prot, and Pfam databases. This database contains information of 2620 perfect simple sequence repeats (P-SSRs) with a relative abundance of 340.12 (loci/Mb), and relative density of 6309.29 (bp/Mb). Excluding mono-nucleotides, the most abundant SSR motifs were tri-nucleotides (54.31%), followed by di-nucleotides (37.51%), tetra-nucleotides (4.54%), penta-nucleotides (3.16%) and hexa-nucleotides (0.45%). Additionally, we also identified 4,311 transcription factors (TFs) and categorized them into 55 sub-families. This database is expected to fill the gap in genomic resource availability in *T. cordifolia* and thus accelerate molecular breeding and related functional and other applied studies aimed towards genetic improvements of *T. cordifolia* and related species.

**Keywords:** *Tinospora*; markers; simple sequence repeats (SSRs); transcriptome; gene ontology (GO); transcription factors (TFs)



**Citation:** Singh, R.; Mahato, A.K.; Singh, A.; Kumar, R.; Singh, A.K.; Kumar, S.; Marla, S.S.; Kumar, A.; Singh, N.K. TinoTranscriptDB: A Database of Transcripts and Microsatellite Markers of *Tinospora cordifolia*, an Important Medicinal Plant. *Genes* **2022**, *13*, 1433. <https://doi.org/10.3390/genes13081433>

Received: 22 June 2022

Accepted: 6 August 2022

Published: 12 August 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

*T. cordifolia* (Willd.) Miers., commonly named as “Giloe” or “Guduchi”, is a large, perennial, deciduous climbing shrub belonging to the family Menispermaceae. It is distributed throughout the tropical Indian sub-continent, Myanmar, Sri Lanka, and China [1]. A wide range of active ingredients such as alkaloids, steroids, diterpenoids, lactones, glycosides, sesquiterpenoids, phenolics, polysaccharides, and aliphatic compounds have been isolated from the different parts of the plant, including, stem, root, and whole plant [2,3]. *T. cordifolia* is known for its antispasmodic, anti-inflammatory, antidiabetic, antiallergic, antipyretic, and analgesic properties and is being used in the treatment of gout, jaundice, rheumatism, urinary disorder, diabetes, anemia, skin diseases and infections [4–7]. According to the Ayurveda (Indian medicine) literature, the phytoactive ingredients of *T. cordifolia* are a constituent of several formulations which is used for the treatment of dyspepsia, urinary, and general debility diseases [8]. The whole plant of *T. cordifolia* has medicinal

value; stems were reported useful in the treatments of fever, flatulence, diabetes, hypertension and leucorrhea [9,10]; roots were reported to be useful for spleen and gynecological disorders [11,12]. The starch of roots and stem was reported to be nutritious and used to cure chronic diarrhea and dysentery [13]. It is also known for its strong immunostimulant action due to the presence of a novel polysaccharide,  $\alpha$ -d-glucan [14,15]. In addition, *T. cordifolia* is also a rich source of various micronutrients such as calcium, phosphorus, iron, zinc, copper and manganese [16]. The leaves of this plant are rich in protein (11.2%) and also calcium and phosphorus [17]. *T. cordifolia* has been studied for its antioxidant properties. In diabetic rats, the oral administration of aqueous *T. cordifolia* root extract lowered blood glucose levels [18]. The root, stem, and leaf extracts of *T. cordifolia* also show antimicrobial [19,20], antitoxic [21,22], antistress [23], anticancer [24–26], anti-HIV [27], and anti-osteoporotic [28,29] activities.

The survey of public databases and literature revealed extremely limited information with regard to molecular markers and SSRs in particular. SSR markers are very useful because of their codominant nature, their ability to detect multiallelic variability, and reliable reproducibility. These SSR markers have various applications in the field of genomics and are thus useful for the generation of high-density marker-based genetic maps for gene/QTL discovery, phylogenetic studies, and comparative genomics [30,31].

We have reported information on the transcriptome sequencing of *T. cordifolia* and also validated nearly 96 SSR markers across different *T. cordifolia* species [32,33]. Such research encouraged us to develop a dedicated database of transcriptome-based SSR markers along with annotation details of transcripts for scaling up research by the scientific community engaged in the genetic improvement of this medicinal plant. In the present study, we report a database of SSR markers and transcription factors (TinoTranscriptDB) and, to the best of our knowledge, this is the first relational database of transcriptome-based microsatellite markers and transcription factors that contains information of primer sequences along with the functional annotation details of respective transcripts. We also incorporated records of transcript sequences, blast results, gene ontology, motif details, and enzyme commission number details in the database so that users can access annotated transcript details for their intended use.

## 2. Materials and Methods

### 2.1. Raw Data Processing and De Novo Meta-Transcriptome Assembly

Our in-house leaf tissue sequenced RNA-Seq data of *T. cordifolia* generated through Roche-454 GS-FLX platform (SRA Acc. No. SRP018583), and another publicly available Illumina platform (SRA Acc. No. SRR15221490) was used in this study. The transcriptome shotgun assembly (TSA) of Roche 454 dataset (25,406 transcripts) was downloaded from the NCBI TSA database and Illumina SRA reads accession no. SRR15221490 was downloaded from the NCBI SRA database. Illumina raw reads were quality-checked using FASTQC (accessed on 2 June 2022; <http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>), and the adapters and poor-quality bases were trimmed using Trimmomatic V3.39 [34]. Then, filtered reads were assembled using rnaSPAdes software with default parameters [35]. The Illumina assembled transcripts (34,027) were further merged with our in-house de novo assembled transcripts (25,406) to generate a non-redundant meta-transcriptome assembly (43,090 transcripts) using the CH-HIT program with 90% identity [36]. This meta-transcriptome assembly was further used for downstream analysis.

### 2.2. In Silico SSR Mining and Primer Designing

The Krait v1.3.3 software (accessed on 17 June 2022; <https://github.com/lmdu/krait>) [37] was used to identify the microsatellite markers (SSRs) in the non-redundant assembled transcripts dataset using the search criteria of six repeat units for di-nucleotide and five repeat units for tri-, tetra-, penta- and hexa-nucleotide repeats and the maximal number of bases interrupting 2 SSRs in a compound microsatellite to be 100 bp. The output results of Krait contained information such as repeat number, motif sequence, length of

SSR repeat, type of repeats, start and end positions, T<sub>m</sub>, and GC content. The primer pairs were designed from the 200 bp flanking regions of SSR motifs using Primer3 v1.0 (<https://github.com/primer3-org/primer3>, accessed on 17 June 2022) [38] with parameters (primer length = 20–25 bp; size of PCR product = 100–250 bp, with optimum of 280 bp; annealing temperature = 65 °C; GC content of 40–60% with optimum of 50%).

### 2.3. Functional Annotation of Transcripts

The non-redundant transcript dataset was BLAST-searched against different databases including NCBI NR, Swiss-Prot, Pfam, gene ontology (GO), and KEGG using BLASTX program with a cut-off E-value of  $1 \times 10^{-10}$  [39]. The functional descriptions were assigned to non-redundant transcripts with BLAST search against NCBI non-redundant and Swiss-Prot protein databases using BLAST2GO 6.0.3 [40] followed by a search for gene ontology terms (GO), to carry out pathway analysis (using KEGG), and assign enzyme commission (EC) numbers.

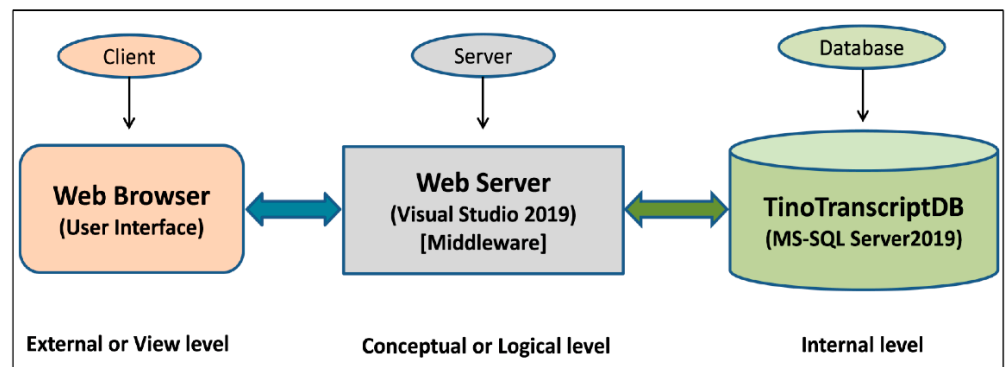
### 2.4. Identification of Transcription Factors and Gene Families

The transcription factors (TFs), working as molecular switches, were identified using the entire non-redundant assembled transcript set, and were BLAST-search against the Plant Transcription Factor Databases (PlantTFDB v4.0) [41] with a parameters bit score > 100 and e-value of  $1 \times 10^{-3}$  [42]. The annotation of the identified TFs was performed using InterProScan ver.5.56 [43], which contains various inbuilt functional databases including Pfam (<http://pfam.xfam.org/>, accessed on 17 June 2022), PRINTS (<http://www.bioinf.manchester.ac.uk/dbbrowser/PRINTS/>, accessed on 17 June 2022), SMART (<http://smart.embl-heidelberg.de/>, accessed on 17 June 2022), PrositePatterns (<http://prosite.expasy.org/>, accessed on 17 June 2022), PrositeProfiles (<http://prosite.expasy.org/>, accessed on 17 June 2022), SUPERFAMILY (<http://supfam.cs.bris.ac.uk/>, accessed on 17 June 2022), Panther (<http://www.pantherdb.org/panther/>, accessed on 17 June 2022) and gene ontology (GO) (<http://amigo.geneontology.org/>, accessed on 17 June 2022), to predict the functional domains and structural motifs along with their annotation details.

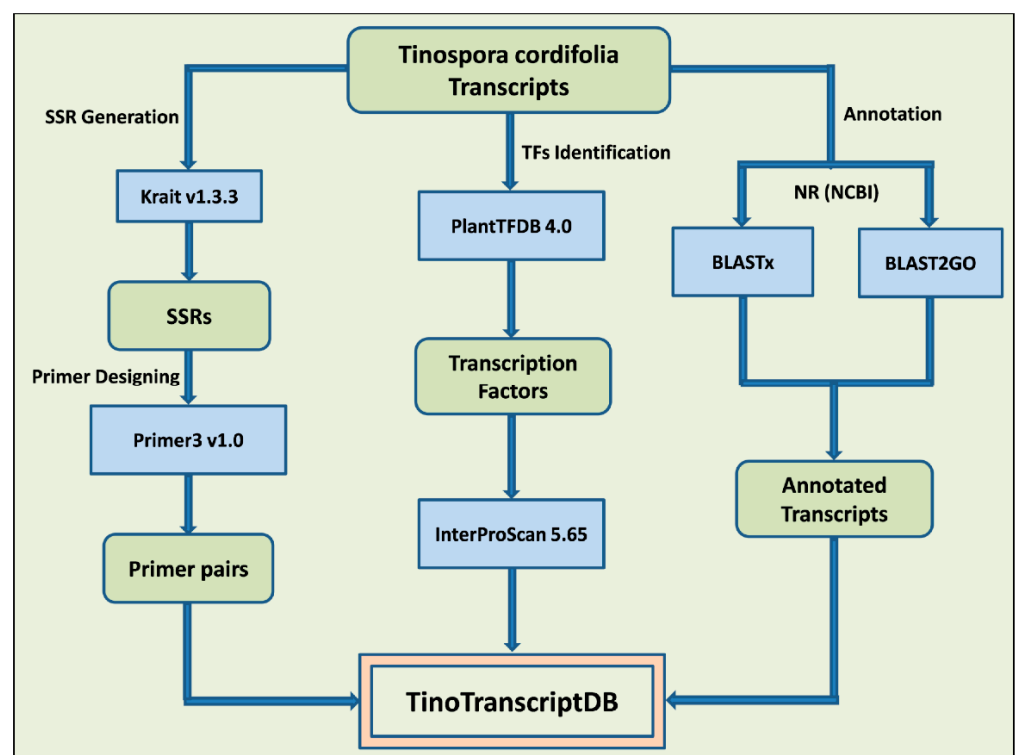
The gene families related to the biosynthesis of biochemical compounds in *T. cordifolia* were also analyzed based on annotation of transcripts.

### 2.5. Database Architecture and Web Interface

*T. cordifolia* transcripts and SSR marker database (TinoTranscriptDB) is a relational and interactive database that contains detailed information on annotated transcripts, SSR markers with three pairs of primers, and TFs with their functional annotation details. Additionally, it also contains transcripts' sequence information from which SSR markers were designed. This database is based on "three level schema architecture" with client tier, server tier, and database tier (Figure 1). The interactive user interface of the database was designed and developed using a server-side web programming language (ASP.NET), and database tables were stored in MSSQL Server 2019. The schematic of the data generation for the present database creation is shown in Figure 2.



**Figure 1.** Three-level database schema architecture used for the construction of TinoTranscriptDB.



**Figure 2.** Process flow diagram and tools used for the transcriptome SSR marker identification, primer designing, annotation and transcription factor categorization.

### 3. Results

#### 3.1. Microsatellite Markers and Primer Pairs

From the non-redundant assembled transcripts dataset, a total of 7980 perfect SSRs, 360 compound SSRs, and 32,551 imperfect SSRs were identified (Table S1). Here, our main focus was perfect SSRs because perfect SSRs are more useful as compared to compound and imperfect SSRs. Out of the total perfect SSRs, tri-nucleotides were the most abundant at 50.13% (32,007), followed by di-nucleotides at 41.95% (24,158), tetra-nucleotides at 4.50% (3588), penta-nucleotides at 2.98% (3075), and hexa-nucleotides at 0.44% (528) (Figure S1). Mono-nucleotides were excluded from our analysis. Among the different types of repeat motifs, it was observed that in each motif category, one particular motif category was dominant. Among all the identified SSR motifs of the total di-nucleotide repeats, 9.52% were “AT”, 8.86% of the total tri-nucleotide repeats were “AAG”, 0.83% of the total tetra-nucleotide repeats were “AAAT”, 0.81% of the total penta-nucleotide repeats were “AATGG”, and 0.01% of the total hexa-nucleotide repeats were “AAAAAC” (Table S2). Out

of the total identified SSRs, three primer pair sequences were generated for 2620 perfect, 136 compound, and 18,272 imperfect SSRs from 2200, 135, and 12,099 transcripts, respectively. Among the three primer pairs of 2620 perfect SSRs, tri-nucleotide repeat primer pairs were the most abundant (54.31%), followed by di-(37.51%), tetra-(4.54%), penta-(3.16%), and hexa-nucleotides (0.45%) (Table 1).

**Table 1.** Distribution of SSR motifs (di to hexa) with three primer pairs, their relative abundance (loci/Mb) and relative density (bp/Mb).

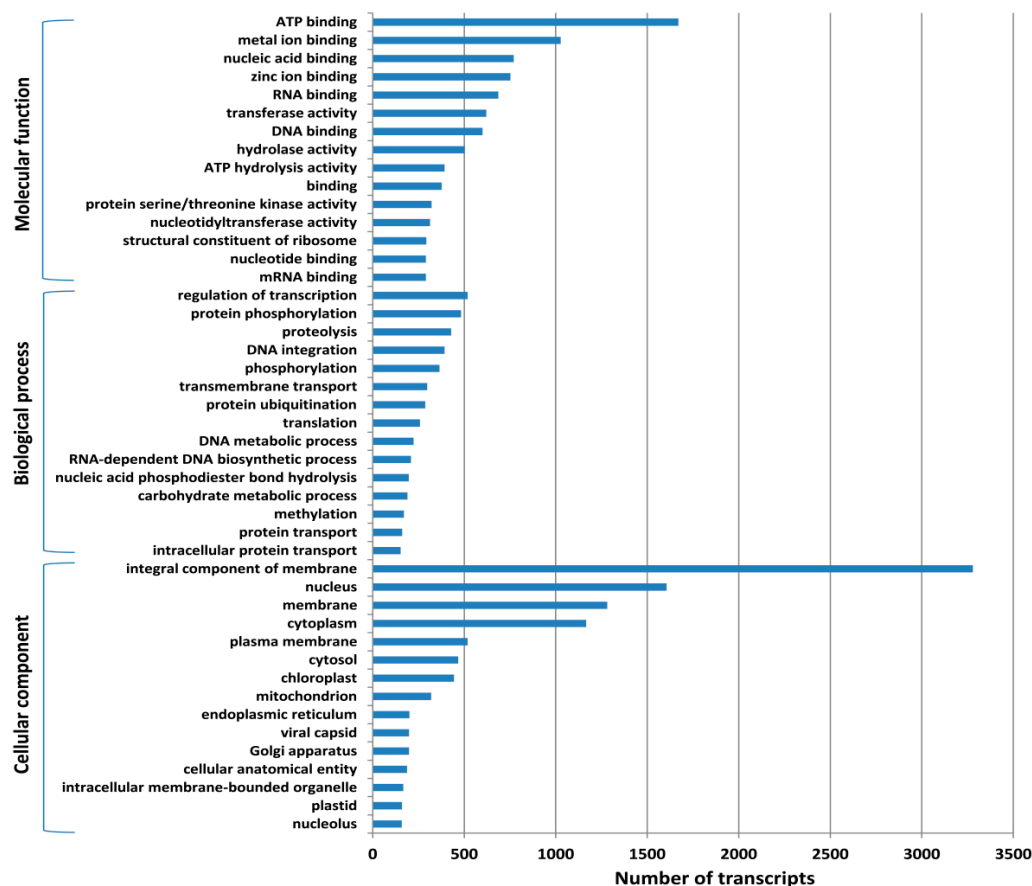
SSR Type	No. of SSRs	Length (bp)	Percentage (%)	Relative Abundance (loci/Mb)	Relative Density (bp/Mb)
Di	983	15,394	37.51	41.90	656.2
Tri	1423	26,262	54.31	60.66	1119.4
Tetra	119	2744	4.54	5.07	117.0
Penta	83	2500	3.16	3.54	106.6
Hexa	12	408	0.45	0.51	17.4

### 3.2. Transcripts Annotation

In order to determine the function of transcripts, we blasted all 43,090 transcript sequences against the non-redundant (nr) protein database (NCBI) using the BLASTx program with an E-value cut off of  $1 \times 10^{-10}$ . A total of 26,487 (61.47%) transcripts showed significant similarity with the NCBI-NR protein database. The similarity distribution of the BLAST hits showed that more than 50% of the transcript sequence had similarities in the range of 40% to 100% (Figure S2). The species distribution graph of the BLAST results showed the highest similarity of *T. cordifolia* transcripts against *Vitis vinifera*, followed by *Nelumbo nucifera*, *Aquilegia coerulea*, and other crop species (Figure S3). Further, 16,614 (38.55%) of the total transcripts showed significant matches with the Swiss-prot database (Table S3). Out of the 26,487 transcripts, 20,364 transcripts were assigned GO terms. A total of 84,172 GO terms were assigned to 20,364 transcripts and their distribution was broadly categorized into molecular functions, biological processes and cellular components (Table S4). Based on the assigned GO terms, the maximum number of transcripts were found to be in “ATP binding” and “metal ion binding” under molecular function category; “regulation of transcription” and “protein phosphorylation” under the biological process category; and “integral component of membrane” and “nucleus” under the cellular component category (Figure 3). A total of 9,259 transcripts were assigned enzyme codes and enzyme names. Out of these, the maximum number of transcripts encode translocases followed by transferases, hydrolases, and oxidoreductases (Table S5).

### 3.3. Transcription Factors

A total of 4311 TFs were identified and were distributed among 55 TF families (Figure 4). Among the TF families, the maximum number of TFs were included in bHLH (448), followed by NAC (287), ERF (285), MYB (281), and FAR1 (210) (Table S6). Interestingly, among various TFs families, the length of the assembled transcript of different TFs, ranged between ~200 and 5865 bp (Figure 5). As the lengths of the majority of the assembled transcripts of TFs are in a workable range in terms of designing primer pairs, we can explore the spatial/temporal expression of candidate TFs along with other assembled transcripts of comparable length. The annotation of 4311 TFs using Inter ProScan showed significant matches with different databases such as pfam, PRINTS, Prosite Pattern, Prosite Profiles, Gene3D, and Panther (Table S7).



**Figure 3.** Functional classification of transcripts based on GO terms, distributed in three major categories: molecular function, biological process, and cellular component.

### 3.4. Gene Families Identified for Biochemical Synthesis

The annotated transcripts were analyzed to identify the gene families involved in the biochemical synthesis of compounds having medicinal value such as, cytochrome p50, kinases, HSPs, and transporters. The cytochrome p50 represents one of the largest enzyme families which play a crucial role the biosynthesis of secondary metabolites, antioxidants biosynthesis, fatty acid metabolism, xenobiotic metabolism, hormone regulation, and plant defense in higher plants [44]. In the present database, 122 transcripts have been identified to be related to cytochrome p50 (Table S8). Similarly, 510 transcripts have been identified as related to protein kinases, which play a fundamental role in signal transduction pathways, regulating a number of cellular functions such as cell growth, differentiation, and cell death. A number of plant extracts and their isolated secondary metabolites such as flavonoids, phenolics, terpenoids, and alkaloids have exhibited activities against various kinases [45]. A total of 30 transcripts related HSPs were found and are known to function as molecular chaperones, which are involved in the therapeutic treatments of many diseases [46]. Transporters are involved in the transport of plant secondary metabolites and a total of 257 transcripts related to transporter genes were found (Table S8). The transporters are used for the transport of secondary metabolites, such as alkaloids, terpenoids, phenolics, steroids, glycosides, aliphatic compounds and polysaccharides [47].

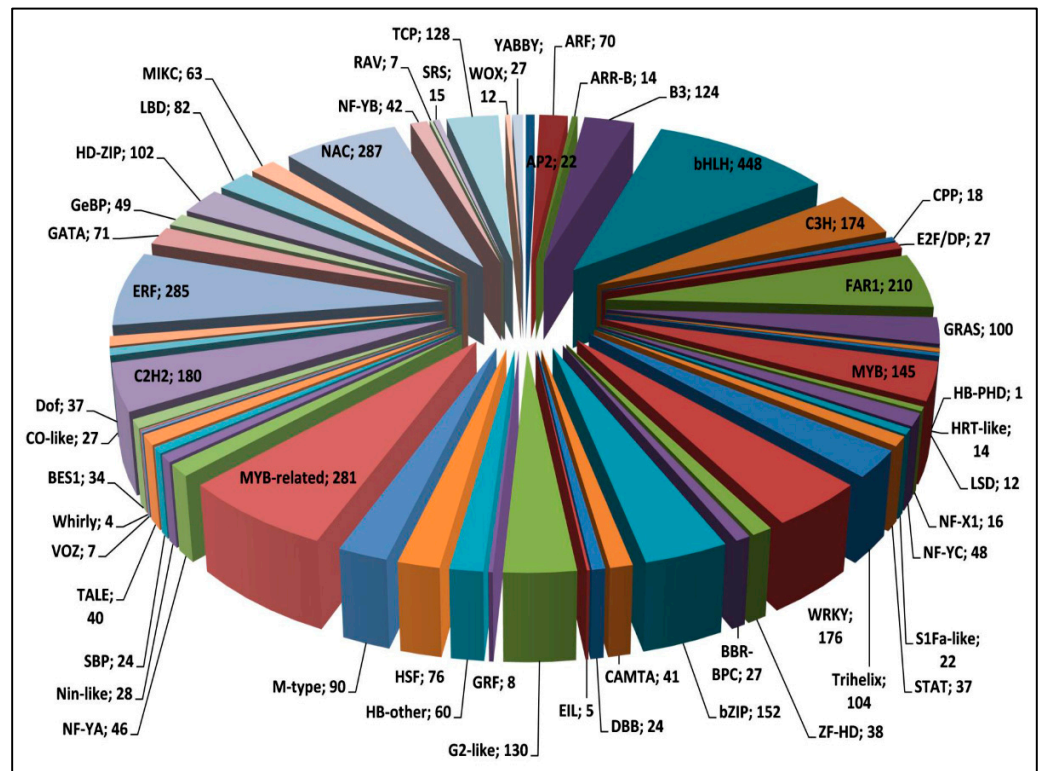


Figure 4. Pie-chart distribution of *T. cordifolia* 4311 transcripts in to 55 TF categories based on the BlastX search against the PlantTFDB.

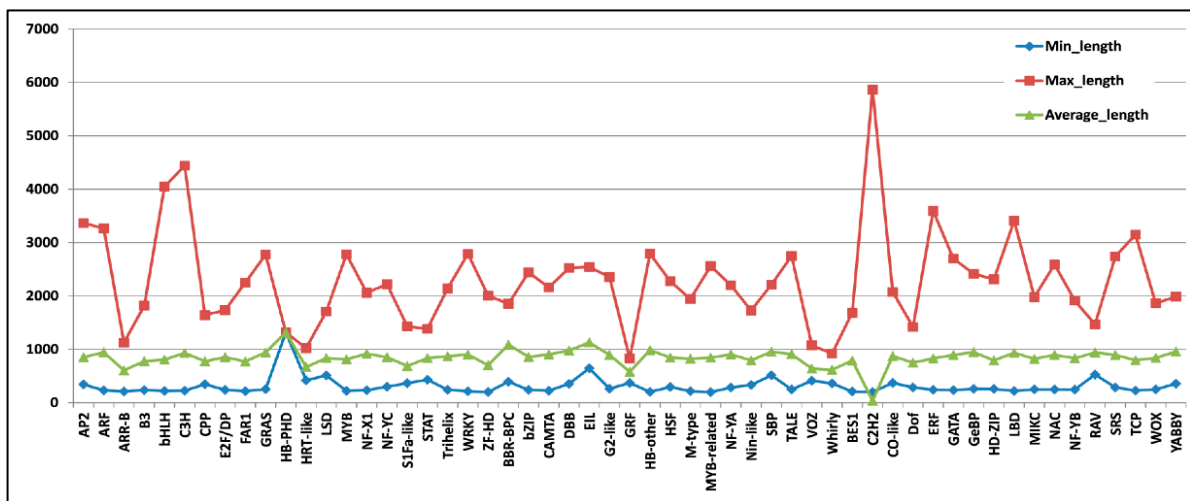


Figure 5. Line graph coverage graph of identified 4311 *T. cordifolia* transcripts categorized into 55 different TF categories based on minimum, maximum and average transcript length of each TFs category.

### 3.5. Database Utilities

We designed an interactive and user-friendly online database of transcript sequences, which contains information of various categories of TFs, SSR markers and GO components. The TinoTranscriptDB was developed to provide uninterrupted public access and deliver useful information about an important medicinal plant, *T. cordifolia*. TinoTranscriptDB has a user-friendly interface for efficient access, visualization and retrieval of information based on queries of interest with a simple or combination of search criteria along with a useful data download option. In the TinoTranscriptDB, data are organized into three

major sections: (1) SSR markers with their three pairs of primer details, (2) transcription factors with their annotation information, and (3) annotated transcripts categories based on GO numbers. Using various search options, users can access the database to browse and download various structural (SSR details) and functional (annotation details) features of the *T. cordifolia* transcriptome.

TinoTranscriptDB has nine different tabs with six sub-tabs under database tab (About, Database, BLAST, Tutorial, Download, Feedback, Links, Team and Contact Us) (Figure 6). The analysis summary of the transcript dataset is shown under the SSR generation tab and the tutorial page contains detailed information regarding the usage of the database. In the TinoTranscriptDB, we have highlighted perfect SSRs because perfect SSRs are more useful as compared to compound and imperfect SSRs. An SSR search of *T. cordifolia* is based on SSR ID, SSRs type, motif sequence, annotation keyword or both motif type and repeats (Figure 6). The search result page provides basic information about the SSR markers, such as SSR ID, SSR type, SSR motif, motif repeat, and SSR length, along with two hyperlinked fields, one for primer details and another for annotation details. When users click on “primer details”, the link takes them to the next web page which displays information related to three sets of primer pairs (Figure 6). Similarly, clicking the “annotation details” takes users to the next web page which contains the full annotation details, including information on gene ontology, Pfam annotation, enzyme commission (EC) number, and transcript sequence from which the SSR has been identified (Figure 6). This annotation result will help researchers in selecting transcripts of their interest for designing primer pairs. The TF categories tab can be used to obtain category-wise TF-related transcripts in Fasta format (Figure 6). Similarly, from the GO categories tab, users can obtain GO category-wise transcripts in Fasta format (Figure 6). In GO categories, entire assembled transcripts were grouped into three major categories as cellular components, biological processes, and molecular function. Each category has subgroups as well as respective transcript information. These Fasta sequences of TF-related as well as GO-related transcripts may be used for further downstream analysis. TinoTranscriptDB also has a separate search function for BLASTx as well as gene ontology (GO) annotation, from which users can view annotations, accession or ID, GO class, and transcript sequences as per the query fed in functional keywords or GO number search functions (Figure 7). From this page, users can also download the transcript data, categorized into three major GO categories (cellular component, molecular function and biological process) in separate Fasta files. The local BLAST feature of TinoTranscriptDB enables the user to compare query sequence with the *T. cordifolia* transcripts.



The screenshot displays the TinoTranscriptDB website interface. On the left, the 'Home page' section includes 'About', 'Database', 'SSRs Information', 'Perfect SSR Search', 'Get TF Categories', 'Tutorial', 'Get GO Categories', 'Download', 'Feedback', 'Annotation Search', 'Links', and 'Contact Us'. Below this is the 'Custom search options' section with search filters for 'Type of SSRs', 'Type of Motif Sequence', 'Motif length and Repeat type', 'Search by Annotation', and 'Search by Sequence ID / SSR ID'. The 'Search results' section shows a table of SSRs with columns for SNo, SSR ID, SSR Type, Motif, Motif Repeat, SSR Length, and Get links. On the right, the 'Primer's information' section shows primer details for three SSRs. Below that, the 'TF categories' section lists transcription factors like AP2, MYB, and MYB-related. The 'GO categories' section lists biological processes like 'Cellulose synthase complex' and 'Chloroplast inner Membrane'. Red arrows (1-6) highlight specific search criteria and links across these sections.

**Figure 6.** Screenshot of the TinoTranscriptDB showing SSR and primer searching using different search options, (1) arrow shows perfect SSR search criteria on SSR search page; (2) arrow shows search for SSR di to hexa types; (3) arrow shows Primer Details link of particular SSR; (4) arrow shows Annotation Details link of particular SSR with annotation information; (5) arrow shows Sequences link of that TF category in a fasta sequence file; (6) arrow shows GO Id in fasta sequence file.

The screenshot shows the transcript annotation search interface. It includes a 'Gene Ontology Search' section with a search bar and 'Search GO' button. Below it is a 'BLASTx Search' section with a search bar and 'Search' button. The 'BLASTx search result' section displays a table with columns for S.No, Transcript Id, Accession, BLASTx Annotation, and Sequence. The 'GO Annotation Search Result' section shows a table with columns for S.No, Transcript Id, GO Class, GO Id, GO Term, Annotation, and Sequence. The 'Amigo information' section shows detailed information for a specific GO term, 'nucleotide binding'. Red arrows (1-3) highlight search options and results across these sections.

**Figure 7.** Screenshot of the TinoTranscriptDB showing transcript annotation searching with the download option for several sequences, (1) arrow shows Gene ontology search by the GO Id; (2) arrow shows by click on the particular GO Id detailed information on a separate page; (3) arrow shows BLASTX search by the annotation keyword on a separate page.

#### 4. Discussion

For the strategic improvement of the genetic traits of germplasm resources, more specifically, for enhancing their economic potential, there is a prior requirement of information related to genomic resources. Such information includes not only sequences pertaining to the genetic determinants that regulate the outcome of the phenotype (functional resources), but also information related to molecular markers (structural resources). Molecular markers serve as mileposts for the genome as they are distributed randomly genome-wide but with fixed locations. Among several molecular markers, the microsatellites, also known as simple sequence repeats (SSR), have become more essential for speeding up research not only through molecular breeding approaches but also in the area of the diversity assessment of the germplasms of plant species. Moreover, if such SSRs are developed based on the sequence information of transcripts, the utilities of such markers can be expanded further. Such gene-based SSR markers are also useful for determining candidate genes, exploring allelic variation, etc. In the development of a database of gene-based SSR markers, utilizing the sequence information of transcripts is researchers a promising method to expand research in the area of marker-assisted selection aimed toward the genetic enhancement of *T. cordifolia*. Additionally, such markers can also be exploited to assess diversity in the available germplasm of *T. cordifolia* [32]. The SSRs can also be employed in related species of *T. cordifolia* through transferability studies [33]. We have already tested the efficacy of transferability studies in related species. The annotation details of transcripts, from which SSRs have been developed, are useful for researchers to help them select markers. The allelic variation with respect to any SSR marker can also provide information related to changes in phenotype, if any. In addition to information on SSR markers, the database also includes annotation details of assembled transcripts, including those of TFs and gene families related to biochemical compound synthesis with medicinal values, thus offering researchers a choice to clone their favorite genes associated with the phenotype of interest.

#### 5. Conclusions

TinoTranscriptDB is a user-friendly database, accessible freely via the address <http://www.nbpg.ernet.in:8080/Tinospora/> (accessed on 17 June 2022), contains a total of 2620 SSR markers with three primer pairs for each identified SSR marker. This database provides users with the flexibility to leverage not only information on SSRs, but also information about transcripts of *T. cordifolia*. To facilitate the usage of transcript-related information, the entire transcripts were categorized into three GO functional groups, TF categories, and gene families. This database is expected to be of practical significance to researchers working on the genetic improvement of traits of economic importance in *T. cordifolia*. In addition, this database will facilitate molecular breeding programs and other applied studies in *T. cordifolia* and related medicinal plant species.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13081433/s1>, Figure S1: Percentage distribution of di to hexa type of perfect SSRs; Table S2: Abundance of each type of SSR motifs in di to hexa categories; Figure S2: Sequence similarity distribution graph of BLAST hits; Figure S3: Sequence similarity distribution of *T. cordifolia* transcripts with other species; Table S1: The summary of Perfect, Compound, and Imperfect Microsatellite marker types; Table S2: Abundance of each type of SSR motif category; Table S3: Transcripts similarity distribution with SwissProt database; Table S4: Gene Ontology (GO) assignment of the transcripts; Table S5: Enzyme codes and Enzyme name assignments of the transcripts; Table S6: TF categories their maximum, minimum and average sequence lengths; Table S7: Transcripts annotation with InterProScan; Table S8: Transcripts categorized into gene families related to the biosynthesis of biochemical compounds in *T. cordifolia*.

**Author Contributions:** R.S. conceived and designed the experiments; A.K.M., A.S. and R.K. performed the experiments; A.S. designed and developed the database; A.K.S. and S.K. analyzed the data; S.S.M. and A.K. contributed reagents/materials/analysis tools; N.K.S., R.S., A.K.M. and A.S. writing—review and editing, A.S. visualization.; R.S. supervision and project administration, R.S. funding acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Indian Council of Agricultural Research (ICAR), New Delhi, India, under project PGR/GRD-BUR-DEL-01.05. The funder has no role in the design of the study and collection, analysis and interpretation of data and in writing the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All the data and Supplementary Materials are available at database online.

**Acknowledgments:** We are thankful to the Director, ICAR-NBPGR, New Delhi, for providing facilities for this work and Head, Division of Germplasm Evaluation, ICAR-NBPGR for providing the plant material.

**Conflicts of Interest:** The authors declare no conflict of interest. The funder has no role in the design of the study and collection, analysis and interpretation of data and in writing the manuscript. There is no clinical evidence for the effectiveness of the *Tinospora cordifolia* in the treatment of any disease.

## References

1. Akbar, S. *Tinospora cordifolia* (Willd.) Miers ex Hook. F. & Thoms (Menispermaceae). In *Handbook of 200 Medicinal Plants*; Springer: Cham, Switzerland, 2020; pp. 1811–1824.
2. Saha, S.; Ghosh, S. *Tinospora cordifolia*: One plant, many roles. *Anc. Sci. Life* **2012**, *31*, 151–159. [[PubMed](#)]
3. Upadhyay, A.K.; Kumar, K.; Kumar, A.; Mishra, H.S. *Tinospora cordifolia* (Willd.) Hook. f. and Thoms. (Guduchi)—Validation of the ayurvedic pharmacology through experimental and clinical studies. *Int. J. Ayurveda Res.* **2010**, *1*, 112–121. [[CrossRef](#)] [[PubMed](#)]
4. Arora, R.; Malhotra, P.; Sundriyal, S.; Yashavanth, H.S.; Pai, R.J.; Baliga, M.S. Medicinal Plants as Remedies for Gastrointestinal Ailments and Diseases. In *Bioactive Food as Dietary Interventions for Liver and Gastrointestinal Disease*; Elsevier: Amsterdam, The Netherlands, 2013; pp. 301–311.
5. Sharma, P.; Dwivedee, B.P.; Bisht, D.; Dash, A.K.; Kumar, D. The chemical constituents and diverse pharmacological importance of *Tinospora cordifolia*. *Heliyon* **2019**, *5*, e02437. [[CrossRef](#)] [[PubMed](#)]
6. Sharma, R.; Amin, H.; Galib; Prajapati, P.K. Antidiabetic claims of *Tinospora cordifolia* (Willd.) Miers: Critical appraisal and role in therapy. *Asian Pac. J. Trop. Biomed.* **2015**, *5*, 68–78. [[CrossRef](#)]
7. Sharma, U.; Bala, M.; Kumar, N.; Singh, B.; Munshi, R.K.; Bhalerao, S. Immunomodulatory active compounds from *Tinospora cordifolia*. *J. Ethnopharmacol.* **2012**, *141*, 918–926. [[CrossRef](#)] [[PubMed](#)]
8. Devasagayam, T.P.; Sainis, K.B. Immune system and antioxidants, especially those derived from Indian medicinal plants. *Ind. J. Exp. Biol.* **2002**, *40*, 639–655.
9. Tiwari, P.; Nayak, P.; Prusty, S.K.; Sahu, P.K. Phytochemistry and pharmacology of *Tinospora cordifolia*. *Syst. Rev. Pharm.* **2018**, *9*, 70–78. [[CrossRef](#)]
10. Rajalakshmi, M.; Eliza, J.; Priya, C.E.; Nirmala, A.; Daisy, P. Anti-diabetic properties of *Tinospora cordifolia* stem extracts on streptozotocin-induced diabetic rats. *Afr. J. Pharm. Pharmacol.* **2009**, *3*, 171–180.
11. Singh, S.; Pandey, S.; Srivastava, S.; Gupta, V.K.; Patro, B.; Ghosh, A.C. Chemistry and medicinal properties of *Tinospora cordifolia* (GUDUCHI). *Indian J. Pharmacol.* **2003**, *35*, 83–91.
12. Kirtikar, K.R.; Basu, B.D.; Blatter, E.; Causis, J.R.; Mhaskar, K.S. *Indian Medicinal Plants*; Lalit Mohan Basu: Allahabad, India, 1933; Volume 1, p. 77.
13. Khare, C.P. *Tinospora cordifolia* (Willd.) and Miers ex Hook.f. and Thoms. In *Indian Medicinal Plants*; Springer: New York, NY, USA, 2007; p. 1. [[CrossRef](#)]
14. Thatte, U.M.; Rao, S.G.; Dahanukar, S.A. *Tinospora cordifolia* induces colony stimulating activity in serum. *J. Postgrad. Med.* **1994**, *40*, 202. [[PubMed](#)]
15. Raveendran Nair, P.K.; Rodriguez, S.; Ramachandran, R.; Alamo, A.; Melnick, S.J.; Escalon, E.; Garcia, P.I.; Wnuk, S.F.; Ramachandran, C. Immune stimulating properties of a novel polysaccharide from the medicinal plant *Tinospora cordifolia*. *Int. Immunopharmacol.* **2004**, *4*, 1645–1659. [[CrossRef](#)] [[PubMed](#)]
16. Saeed, M.; Naveed, M.; Leskovec, J.; Ali kamboh, A.; Kakar, I.; Ullah, K.; Ahmad, F.; Sharif, M.; Javaid, A.; Rauf, M.; et al. Using Guduchi (*Tinospora cordifolia*) as an eco-friendly feed supplement in human and poultry nutrition. *Poult. Sci.* **2020**, *99*, 801–811. [[CrossRef](#)] [[PubMed](#)]
17. Joshi, B.C.; Niyal, S. Pharmacognostical Review of *Tinospora cordifolia*. *Inventi Rapid Planta Act.* **2017**, *1*, 1–10.

18. Prince, P.S.; Menon, V.P. Antioxidant activity of *Tinospora cordifolia* roots in experimental diabetes. *J. Ethnopharmacol.* **1999**, *65*, 277–281. [[CrossRef](#)]
19. Jeyachandran, R.; Xavier, T.F.; Anand, S.P. Antibacterial activity of stem extracts of *Tinospora cordifolia* (willd) hook. *Anc. Sci. Life Res.* **2003**, *25*, 40–43.
20. Agarwal, S.; Ramamurthy, P.H.; Fernandes, B.; Rath, A.; Sidhu, P. Assessment of antimicrobial activity of different concentrations of *Tinospora cordifolia* against *Streptococcus mutans*: An in-vitro study. *Dent. Res. J.* **2019**, *16*, 24–28.
21. Gupta, R.; Sharma, V.A. Meliorative effects of *Tinospora cordifolia* root extract on histopathological and biochemical changes induced by aflatoxin-b in mice kidney. *Toxicol. Int.* **2011**, *18*, 94–98.
22. Sharma, V.; Pandey, D. Beneficial effects of *Tinospora cordifolia* on blood profiles in male mice exposed to lead. *Toxicol. Int.* **2010**, *17*, 8–11. [[CrossRef](#)]
23. Sarma, D.N.K.; Khosa, R.L.; Chansauria, J.P.N.; Sahai, M. Antistress Activity of *Tinospora cordifolia* and *Centella asiatica* Extracts. *Phytother. Res.* **1996**, *10*, 181–183. [[CrossRef](#)]
24. Mishra, R.; Kaur, G. Aqueous ethanolic extract of *Tinospora cordifolia* as a potential candidate for differentiation-based therapy of glioblastomas. *PLoS ONE* **2013**, *8*, e78764. [[CrossRef](#)]
25. Jagetia, G.C. Anticancer activity of Giloe, *Tinospora Cordifolia* (Willd.) Miers Ex Hook F & Thoms. *Int. J. Complement. Altern. Med.* **2019**, *12*, 79–85.
26. Ali, H.; Dixit, S. Extraction optimization of *Tinospora cordifolia* and assessment of the anticancer activity of its alkaloid palmatine. *Sci. World J.* **2013**, *28*, 376216.
27. Estari, M.; Venkanna, L.; Reddy, A.S. In vitro anti-HIV activity of crude extracts from *Tinospora cordifolia*. *BMC Infect. Dis.* **2012**, *12*, 10. [[CrossRef](#)]
28. Abiramasundari, G.; Sumalatha, K.R.; Sreepriya, M. Effects of *Tinospora cordifolia* (Menispermaceae) on the proliferation, osteogenic differentiation and mineralization of osteoblast model systems in vitro. *J. Ethnopharmacol.* **2012**, *141*, 474–480. [[CrossRef](#)]
29. Kapur, P.; Jarry, H.; Wuttke, W.; Pereira, B.M.; Seidlova-Wuttke, D. Evaluation of the antiosteoporotic potential of *Tinospora cordifolia* in female rats. *Maturitas* **2008**, *59*, 329–338. [[CrossRef](#)]
30. Delseny, M.; Laroche, M.; Penon, P. Detection of sequences with Z-DNA forming potential in higher plants. *Biochem. Biophys. Res. Commun.* **1983**, *116*, 113–120. [[CrossRef](#)]
31. Gupta, S.; Kumari, K.; Sahu, P.P.; Vidapu, S.; Prasad, M. Sequence based novel genomic microsatellite markers for robust genotyping purposes in foxtail millet [*Setaria italica* (L) P Beauv]. *Plant Cell Rep.* **2012**, *31*, 323–337. [[CrossRef](#)] [[PubMed](#)]
32. Singh, R.; Kumar, R.; Mahato, A.K.; Paliwal, R.; Singh, A.K.; Kumar, S.; Marla, S.S.; Kumar, A.; Singh, N.K. De novo transcriptome sequencing facilitates genomic resource generation in *Tinospora cordifolia*. *Funct. Integr. Genom.* **2016**, *16*, 581. [[CrossRef](#)] [[PubMed](#)]
33. Paliwal, R.; Kumar, R.; Choudhury, D.R.; Singh, A.K.; Kumar, S.; Kumar, A.; Bhatt, K.C.; Singh, R.; Mahato, A.K.; Singh, N.K.; et al. Development of genomic simple sequence repeats (g-SSR) markers in *Tinospora cordifolia* and their application in diversity analyses. *Plant Gene* **2016**, *5*, 118–125. [[CrossRef](#)]
34. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
35. Bushmanova, E.; Antipov, D.; Lapidus, A.; Prjibelski, A.D. rnaSPAdes: A de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience* **2019**, *8*, giz100. [[CrossRef](#)] [[PubMed](#)]
36. Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. [[CrossRef](#)] [[PubMed](#)]
37. Du, L.; Zhang, C.; Liu, Q.; Zhang, X.; Yue, B. Krait: An ultrafast tool for genome-wide survey of microsatellites and primer design. *Bioinformatics* **2018**, *34*, 681–683. [[CrossRef](#)]
38. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **2012**, *40*, e115. [[CrossRef](#)]
39. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
40. Conesa, A.; Gotz, S.; Garcia-Gomez, J.M.; Terol, J.; Talon, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [[CrossRef](#)] [[PubMed](#)]
41. Jin, J.; Tian, F.; Yang, D.C.; Meng, Y.Q.; Kong, L.; Luo, J.; Gao, G. PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* **2016**, *45*, D1040–D1045. [[CrossRef](#)]
42. Singh, A.; Sharma, A.K.; Singh, N.K.; Sharma, T.R. PpTFDB: A pigeonpea transcription factor database for exploring functional genomics in legumes. *PLoS ONE* **2017**, *12*, e0179736. [[CrossRef](#)]
43. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]
44. Pandian, B.A.; Sathishraj, R.; Djanaguiraman, M.; Prasad, P.V.V.; Jugulam, M. Role of cytochrome P450 enzymes in plant stress response. *Antioxidants* **2020**, *9*, 454. [[CrossRef](#)] [[PubMed](#)]
45. Gill, M.S.; Saleem, H.; Ahemad, N. Plant extracts and their secondary metabolites as modulators of kinases. *Curr. Top. Med. Chem.* **2020**, *20*, 1093–1104. [[CrossRef](#)] [[PubMed](#)]

46. Ahmed, K.; Furusawa, Y.; Tabuchi, Y.; Emam, H.F.; Piao, J.L.; Hassan, M.A.; Yamamoto, T.; Kondo, T.; Kadowaki, M. Chemical inducers of heat shock proteins derived from medicinal plants and cytoprotective genes response. *Int. J. Hyperth.* **2012**, *28*, 1–8. [[CrossRef](#)] [[PubMed](#)]
47. Yazaki, K. ABC transporters involved in the transport of plant secondary metabolites. *FEBS Lett.* **2005**, *580*, 1183–1191. [[CrossRef](#)] [[PubMed](#)]