Identification and Characterization of TUBBY genes in developmental stages of *Zea mays*

Sendi Mejia  
Soil & Nutrition Laboratory Research Unit

Suhani Shah  
Soil & Nutrition Laboratory Research Unit

Yara Abdelsalam  
Soil & Nutrition Laboratory Research Unit

Ali Nimra  
Soil & Nutrition Laboratory Research Unit

Munzir Bhatt  
Soil & Nutrition Laboratory Research Unit

Daniela Borges  
Soil & Nutrition Laboratory Research Unit

Joey Chen  
Soil & Nutrition Laboratory Research Unit

George Dwamena  
Soil & Nutrition Laboratory Research Unit

Angel Elcock  
Soil & Nutrition Laboratory Research Unit

Jason Firrello  
Soil & Nutrition Laboratory Research Unit

Anordia Foster  
Soil & Nutrition Laboratory Research Unit

Josephine James  
Soil & Nutrition Laboratory Research Unit

Delivrance Jenkins  
Soil & Nutrition Laboratory Research Unit

Maheen Khan  
Soil & Nutrition Laboratory Research Unit

Adwoa Lewis  
Soil & Nutrition Laboratory Research Unit

Ana Manzanilla  
Soil & Nutrition Laboratory Research Unit

Melisa Morataya
Abstract

In this paper, we studied the organization and function of TUBBY Transcription Factor gene family in maize. Initially, using comparative approach, we discovered the Arabidopsis thaliana orthologs in Zea mays. We found in total 13 genes, 12 of which are orthologs and a unique paralog that exhibits the highest activity in maize. We studied the role of TUBBY gene family across different developmental stages using existing expression data, and discovered the binding motifs present in the promoter region of the genes.

Introduction

Maize is one of the most important agricultural crops worldwide, while United States of America is the top maize producer. Maize prospers in a range of climatic zones and it fits the niche between rice and wheat in the phylogenetic tree (Kellogg 2001). It is used as livestock feed and for industrial production of starch and oil extraction. For human consumption, maize is an excellent source for carbohydrates, minerals, protein and iron as well as for vitamin B.

Transcription Factors (TFs) are trans-regulatory proteins controlling the expression of other genes by recognizing their promoter regions (Zheng, Jiao et al. 2016, Tian 2020). Specific domains in the proteins recognize specific DNA motifs that bind to, thereby inducing the transcription of a particular gene. TFs are clustered in different groups of genes depending on the motifs and the domains they contain (Boggon, Shan et al. 1999), as well as the developmental stages and environmental conditions they have been found to be associated with (Zhang, Lubberstedt et al. 2013, Gapper, Giovannoni et al. 2014, Mickelbart, Hasegawa et al. 2015, Xu and Zhang 2015).

TUBBY genes were first identified to be expressed in obese in mice (Kleyn, Fan et al. 1996), and they were characterized as TFs (Boggon, Shan et al. 1999). Nevertheless, they have been identified in many eukaryotic species ranging from animals to plants implying that they are involved in conserved functions (Wang, Xu et al. 2018). The TUBBY gene family was suggested to expand via segmental duplication and random translocation and insertion (Yang, Zhou et al. 2008). Members of the Tubby-Like Protein family (TLP) are characterized by a conserved C-terminal TUB domain, a domain hypothesized to bind to specific membrane phosphoinositide phosphatases (PIPs), to form a stable but flexible complex (Meng Wang, et. al., 2018). In regards to their subcellular localization, TUBBY proteins are expressed mainly in plasma membrane, cytoplasm and nucleus (Wang, Xu et al. 2018).

In plants, TUBBY genes were first characterized systematically in Arabidopsis thaliana, where it was found that they are involved in ABA signaling pathway, in both biotic and abiotic stress responses. In the vast majority of plants, TUBBY genes contain an N-terminal F-box domain. Moreover, the ten aminoacid residues between the F-box domain and TUB domain is suggested to be evolutionarily conserved in plants. However, TUBBY genes in some plant species may not contain the F-box domain. For instance, in Arabidopsis thaliana, all AtTLP genes except AtTLP8 possess an F-box domain (Lai, Lee et al. 2004), and in Oryza sativa, OsTLP13 is the only gene lacking the F-box domain (Liu 2008). Furthermore, within the C
terminal region of the TUB domain, two segmental patterns that are defined as TUB1 and TUB2 motifs consisting of 14 and 16 amino acids, respectively (Lai, Lee et al. 2004).

In this paper, we used comparative genomics to identify the *Zea mays* TUBBY homologue genes using the *Arabidopsis thaliana* TUBBY genes. After manual annotation and support of further analysis, we ended up with twelve orthologs and a novel maize-specific TUBBY gene. We used available expression data at different developmental stages and performed manual curation of the genes using long and short reads. We also looked for the motifs present in each of the ZmTLP and their alternative-splicing products. We identified *cis*-regulatory elements for each TUBBY gene, to improve our understanding of the processes each gene is involved. Finally, we used RT-PCR to confirm the expression of the genes in the different tissues. Our results offer insight into possible functions of a largely unknown transcription factor family.

**Material And Methods**

**Discovering the TUB Homologs**

Initially, we have used as our reference the TUBBY annotated genes from *Arabidopsis thaliana* genome using The Arabidopsis Information Resource (www.arabidopsis.org/browse/genefamily/TUB.jsp) on www.arabidopsis.org Feb 5, 2019. Following the identification of the gene names associated to the TUBBY family, each of the genes were then searched in the Gramene website, a comparative resource site for plants (Tello-Ruiz, Naithani et al. 2018), under the Arabidopsis genome. After searching for a particular gene name, we were directed to the Ensembl database to view gene trees. The gene trees displayed the orthologs of the TUBBY gene family for various plant species. In order to find the orthologs of *Zea mays* to *Arabidopsis thaliana* genome, we searched for *Zea mays*, current genome version (v4), under the tab Plant Compare. Once the orthologs were identified, we selected the compared regions to obtain the coordinates needed to locate the genes in the *Zea mays* genome. In addition to finding the orthologs, we interested in finding paralogs within *Zea mays*, if any, using the paralog tab. The overall workflow of our analysis can be found in Fig. 1.

**Annotating the TUB genes**

For the manual annotation of the of *Arabidopsis thaliana* TUBBY gene family orthologs in *Zea mays*, we used the Apollo website utility (Fig. 2). Apollo is an interactive browser, which grants one entry to modify and refine the exact location and structure of genomic features (Lee, Helt et al. 2013). The chromosomal region of interest was located using the coordinates found from the Gramene database as described in the previous section. Before annotating the genes, we selected the following evidence tracks from the annotator panels: est2genome.iso, est2genome-flc, Maker_updated, and v_model_mapped to_v4. These tracks provided us with various gene predictions, which accurately reflect the structure of a given gene. By utilizing the tracks, we were able to produce reasonable gene models. Afterwards, we selected the gene models that most resembled the structure of the gene, using the gene as an archetype to refine the gene models. When necessary, editing functions were used to add exons and modify splice sites in transcript
alignments to correspond the selected gene model to the given gene. We continued this process for each ortholog found. To strengthen the validation, we have mapped Illumina BAM files from different tissues within Apollo (Fig. 2).

**Construction of phylogenetic trees**

To identify the levels of conservation among plant species of the TUBBY genes, we used the Gramene tab resource on Homology. Then, we followed the link for the Ensembl Gene Tree view, under which we checked for the presence of each of the TUBBY genes in available genomes in Gramene. Using the Genomic regions of each of the TUBBY genes, we used the online clustalw (Larkin, Blackshields et al. 2007) algorithm in GenomeNet (http://www.genome.jp) to construct the alignment of all TUBBY genes in maize followed by the construction of a phylogenetic tree using the PhyML (Guindon, Lethiec et al. 2005) and bootstrap 10,000.

**Searching for domains and motifs**

We used Prosite web resources (https://prosite.expasy.org) with the default parameters, (Sigrist, de Castro et al. 2013) to search for domains of TUBBY proteins. We used The Plant Promoter Analysis Navigator plantPAN 3.0 to search 2,000 bp upstream of the transcription start site for each of the TUBBY genes to find binding sites for other transcription factors (Chow, Lee et al. 2019).

**Digital expression**

We used the Gramene nomenclature to explore the digital expression of the gene family using the free version of Genevestigator platform (Hruz, Laule et al. 2008). Within this platform, we explored clustering for both gene and conditions or developmental stages, using the Euclidian clustering method.

**Designing primers to validate gene structures**

To validate the gene modules, primers were designed to amplify the whole length mRNA of the genes. RT-PCR were performed on maize cDNA from leaf, stem and root tissues. Amplified products were visualized on 1% electrophoresis gel made with TBE buffer. The workflow is graphically illustrated in Fig. 1.

**Results**

**Discovering the TUB Homologs**

In *Arabidopsis thaliana*, there are 10 TUBBY (AtTLP1–7 and AtTLP9–11) genes. For each of the genes, we searched for *Zea mays* orthologs. Out of the 10 genes in *Arabidopsis* thaliana, we found five to have homologs in *Zea mays*. Those are the AtTLP2 and AtTLP3 that have two homologs each in *Zea mays*, namely ZmTLP6 and ZmTLP9, and ZmTLP10 and ZmTLP11, respectively. AtTLP5 has five homologs in *Zea mays*, namely, ZmTL2, ZmTLP5, ZmTLP7, ZmTLP8 and ZmTLP14. Also, AtTLP7 has three homologs ZmTLP3, ZmTLP4 and ZmTLP12 and finally AtTLP8 with one homolog, ZmTLP13. As far as ZmTLP10 concerns, in the process of manual curation, we did not find enough evidence supporting the gene module
and therefore it was eliminated. In total, we came up with thirteen unique *Zea mays* TUBBY TFs genes homologs to *Arabidopsis thaliana* named as ZmTLP. Three ZmTLP genes are located in chromosome four, two genes in each of chromosomes three and five, and one gene in each of chromosomes six, eight, nine and ten (Table 1; Fig. 3). Moreover, we found one paralog gene -ZmTLP1- unique to maize that is located in chromosome 8.

### Table 1

Comparative analysis of TUBBY Transcription Factor Gene Family between *A. thaliana* and *Z. mays*. The first column refers to the general TUBBY gene module. The second and third columns refer to the *Zea mays* gene, and the *Arabidopsis thaliana* gene module, respectively. In the next two columns the alternative-splicing events in *Zea mays* are presented: more specifically the Before manual annotation refers to the alternative-spliced modules found for each locus initially, and the After manual annotation column refers to the modules we kept after performing the manual annotations. In the last three columns under the *Z. mays* Genomic location part, the chromosome number, start and end coordinates of each gene locus in *Zea mays* is shown.

| Gene name | Z.mays genes | Homolog A.thaliana | Alternative-splicing events in Z.mays | Z.mays Genomic location |
|-----------|---------------|--------------------|-------------------------------------|--------------------------|
|           |               |                    | Before manual annotation | After manual annotation | Chr  | Start  | End    |
| TUBBY14   | ZmTLP1        | -                  | 4 | 2 | 10 | 8711082 | 8725025 |
| TUBBY13   | ZmTLP10       | AtTLP3             | - | - | 9  | 115435278 | 115437104 |
| TUBBY11   | ZmTLP11       | AtTLP3             | 4 | 1 | 1  | 60879438 | 60888390 |
| TUBBY5    | ZmTLP12       | AtTLP7             | 5 | 2 | 6  | 162624149 | 162628635 |
| TUBBY8    | ZmTLP13       | AtTLP8             | 2 | 1 | 5  | 93095266 | 93097731 |
| TUBBY2    | ZmTLP14       | AtTLP5             | 6 | 2 | 3  | 178236350 | 178240117 |
| TUBBY12   | ZmTLP15       | -                  | - | - | -  | -        | -        |
| TUBBY3    | ZmTLP2        | AtTLP5             | 5 | 3 | 4  | 218421396 | 218425679 |
| TUBBY15   | ZmTLP3        | AtTLP7             | 7 | 5 | 3  | 213249521 | 213254004 |
| TUBBY1    | ZmTLP4        | AtTLP7             | 8 | 1 | 8  | 154851163 | 154856045 |
| TUBBY4    | ZmTLP5        | AtTLP5             | 4 | 2 | 2  | 26117809 | 26122010 |
| TUBBY9    | ZmTLP6        | AtTLP2             | 6 | 6 | 5  | 206380608 | 206384416 |
| TUBBY10   | ZmTLP7        | AtTLP5             | 4 | 3 | 6  | 154166396 | 154172068 |
| TUBBY7    | ZmTLP8        | AtTLP5             | 4 | 4 | 4  | 200038358 | 200042350 |
| TUBBY6    | ZmTLP9        | AtTLP2             | 4 | 3 | 4  | 162975759 | 162979667 |

We also searched for alternative-splicing events. We found in total 63 alternative-splicing events that per ZmTLP gene ranged from two to eight as shown in Table 1. Specifically, the maximum of eight splicing
events were found in ZmTLP4, and the least of two splicing events in ZmTLP13. After manual annotation, we ended up with 35 alternative-splicing events.

**Annotating the TUBBY genes**

Once we identified the *Zea mays* TUBBY transcription factors, we used a web version of Apollo to validate the annotations (Table 1; Fig. 1). In this process, we took into consideration expression data (both PacBio and Illumina) from different tissues. We validated 12 TUBBY genes (Table 1; Fig. 1) from which 35 alternative-spliced models were confirmed by expression data (Table 1). By going over with the manual annotation of the TUBBY genes, we deleted 44.4% of the automatically predicted alternative-spliced transcripts.

**Phylogenetic analysis**

We found that all ZmTLP genes have homologs in all vascular eukaryotic plants but not in non-vascular plants (Supplementary Fig. 1). ZmTLP11 has no homologs in *Hordeum vulgare*, while ZmTLP13 has no homologs in any of the *Corchorus capsularis, Trifolium pretense* and *Lupinus angustifolius* vascular plants. In *Selaginella moellendorfii* and ancient vascular seedless plant species, most of the ZmTLPs genes have homologs with the exception of ZmTLP1, ZmTLP3 and ZmTLP4. Finally, ZmTLP1 is not present in most of the vascular plants, except of several Oryza species (Supplementary Fig. 1).

To find the evolutionary relationship of the TLP genes in maize, we constructed the phylogenetic tree of the *Zea mays* genes (Fig. 4). We observed four sub-groups to be formed. One group consisted ZmTLP11 and 13, another group with ZmTLP3, 4 and 12, a third group with ZmTLP6 and 9, and finally, the last and largest subgroup contained ZmTLP2, 5, 7, 8, 14 and ZmTLP1. These results are in agreement with the findings in Table 1.

**Discovering motifs and domains**

We were further interested in characterizing the annotated products/proteins (Fig. 5). The majority of the ZmTLP loci including the alternative-splicing products contained the TUB1 and TUB2 motifs. The ZmTLP8.1, 8.3 and 8.4 have a prokaryotic lipoprotein (ID PS51257) motif at the N-terminus of the sequences, while ZmTLP8.2 follows the pattern of the rest of the proteins having both TUB1 and TUB2 motifs. The ZmTLP7 locus has three final products as a result of alternative splicing, but ZmTLP7.3 has no predicted TUB1 and TUB2 motif. From locus ZmTLP6, all alternative spliced products have the TUB1 and TUB2 motifs with the exception of ZmTLP6.5 that has only TUB2 motif. The ZmTLP3 locus has five alternative spliced products and have only TUB1 motif. Interestingly, the maize-specific gene ZmTLP1 has X alternative-spliced products and consists of both TUB1 and TUB2 motifs. Finally, ZmTLP13 has no TUB1 or TUB2 predicted motif.

Further, we searched the maize TLP genes for cis-regulatory elements using the PlantPan 3.0 database and selecting all available plant genomes. In total, 55 elements in the 13 genes were identified (Supplementary table 1), out of which 22 cis-regulatory elements were found present in all genes. On our
set of genes, we identified motifs for binding site of other transcription factors that are responsible for organ development. Those are motifs recognized by B3, SRS, AT-hook, Trihelix. We also found motifs that other regulatory transcription factors involved in both stress and development bind to, such as bHLH, bZIP, C2H2, HD-ZIP, ZF-HD transcription factors. Moreover, additional motifs for binding of transcription factors involved only in the stress responses such as WRKY, MADF, MADS and Homeodomain were identified. Finally, there are binding sites for transcription factors such as the C3H involved in embryo development, the MYB involved in organ identity, the NAC involved in leaf senescence, the SBP involved in flower development, and the VOZ involved in pollen maturation.

Digital Expression across developmental stages

The next step was to use available data to study gene expression of the different homologs and across different developmental stages. We clustered the results for both genes and developmental stages (Fig. 6). Two major groups are formed regarding the developmental stages, the dough stage and fruit formation called Group 1, where we do not observe high expression of the genes. Group 2 is sub-divided into two subgroups SG1 subgroup with relative higher expression values, and the SG2 subgroup with even highest expression than SG1. The developmental stages in SG1 subgroup are the germination, stem elongation and the seedling stage, while the SG2 subgroup is the inflorescence formation and anthesis. Specifically, the group composed of ZmTLP2, 4, 5, 6, 7, and 9 displayed higher expression than the second group composed of ZmTLP3, 8, 11, 12 and 14. Finally, the ZmTLP1 has the highest expression and the ZmTLP13 has the lowest expression of all (Fig. 6).

In the process of the gene annotation, we took into consideration the tissues that genes were expressed. We found all genes expressed in tassel, ear, and embryo tissues, while ZmTLP1 and ZmTLP13 are additionally expressed in root tissues. In endosperm tissue, the ZmTLP2, ZmTLP4, 5, and 6 as well the ZmTLP9, 11 and 14 are expressed. In pollen tissues, only ZmTLP1, 3 and 14 are expressed (example gene, TUBBY 11, is shown in Fig. 2b). To verify the expression of the genes we used cDNA from roots, leaves and stems. Here, we confirmed the expression of ZmTLP6.2, ZmTLP14.1, ZmTLP3.1, ZmTLP3.2 and ZmTLP11.1 in stem tissue (Fig. 7).

Discussion

In this paper, we described the 13 TUBBY transcription factor genes in maize genome (version 4). Twelve of them were orthologs to five out of ten Arabidopsis thaliana TUBBY transcription factors (AtTLP2, AtTLP3, AtTLP5, AtTLP7 and AtTLP8), and one of them, ZmTLP1, was unique to Zea mays (Table 1, Fig. 3). We observed from simple to complex evolutionary relationships. For example AtTLP7 has a one to many relationships since, AtTLP7 is a homolog to three Zea mays genes, while AtTLP8 has a one to one relationship to the homolog ZmTLP13(Gabaldon and Koonin 2013).

TLP genes (Chen Yulong 2016, Xu, Xing et al. 2016) belong to a conservative transcription factor family found in many eukaryotic organisms in both animals and plants (Wang, Xu et al. 2018). When we explored in detail the phylogeny of the genes in plants, we noticed that they are predominately found in
vascular plants (Supplementary Fig. 1), with the exception of ZmTLP1 that is found in few vascular plants. Moreover, by performing motif prediction, we noticed that for some loci with alternative-spliced final products, there were no characteristic TUB1 and TUB2 motifs present. This indicates that those proteins might be involved in different functions (Fig. 5).

Based on the digital expression analysis, several groups and sub-groups were formed at the developmental stages and at gene level. We found that the unique maize TLP, ZmTLP1, has the highest expression (Fig. 6) and it is expressed in most of the developmental stages, suggesting it plays a major regulatory role. The similarity analysis indicates ZmTLP1 gene does not have a homolog, but it is close to the pair of ZmTLP2 and 5 (Fig. 4). ZmTLP13, on the other hand, is the least expressed gene with the exception of seedling stage, indicating it might have a specialized role as a transcription factor to this particular stage. From our analysis, dough stage, fruit formation, anthesis with inflorescence formation, and stem elongation with seedling stages might be regulated by the same set of transcription factors. The manner in which the genes are clustered is an indication that co-clustered genes might be involved in the same developmental stages. For example, the ZmTLP11 and ZmTLP14 have the highest expression in the germination stage and, therefore, play a major regulatory role in this developmental stage.

Our search for cis-regulatory elements (Supplementary table 1) resulted in many binding sites for transcription factors involved mainly in developmental stages and organ identity (Kortschak, Tucker et al. 2000, Peng and Weselake 2013, Gu, Guo et al. 2017), followed by binding sites for transcription factors involved in development and stress (Jakoby, Weisshaar et al. 2002, Huang, Wang et al. 2004, Zanetti, Ripodas et al. 2017, Zhang, Lv et al. 2018), and binding sites motifs for transcription factors involved only in stress only (Sasaki and Imai 2011, Kim, Park et al. 2013, Grimplet, Pimentel et al. 2017, Yu, Wang et al. 2018). This is an indication the TLPs is an important family involved in various pathways and have different functions.

**Declarations**

**Acknowledgements**

I thank Dr. Elias Lolis for stimulating discussions and comments on the manuscript.

**References**

1. Boggon TJ, Shan WS, Santagata S, Myers SC, Shapiro L. "Implication of tubby proteins as transcription factors by structure-based functional analysis." Science. 1999;286(5447):2119–25.

2. Chen Yulong DW, Baoming S. Zhao Yang & Ma Qing. (2016). "Genome-wide identification and comparative analysis of the TUBBY-like protein gene family in maize." Genes Genomics 38: 25–36.

3. Chow CN, Lee TY, Hung YC, Li GZ, Tseng KC, Liu YH, Kuo PL, Zheng HQ, Chang WC. "PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants.". Nucleic Acids Res. 2019;47(D1):D1155–63.
4. Gabaldon T, Koonin EV. "Functional and evolutionary implications of gene orthology.". Nat Rev Genet. 2013;14(5):360–6.

5. Gapper NE, Giovannoni JJ, Watkins CB. Understanding development and ripening of fruit crops in an 'omics' era. Hortic Res. 2014;1:14034.

6. Grimplet J, Pimentel D, Agudelo-Romero P, Martínez-Zapater JM, Fortes AM. "The LATERAL ORGAN BOUNDARIES Domain gene family in grapevine: genome-wide characterization and expression analyses during developmental processes and stress responses.". Sci Rep. 2017;7(1):15968.

7. Gu C, Guo ZH, Hao PP, Wang GM, Jin ZM, Zhang SL. "Multiple regulatory roles of AP2/ERF transcription factor in angiosperm.". Bot Stud. 2017;58(1):6.

8. Guindon S, Lethiec F, Duroux P, Gascuel O. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 2005;33:W557–9. (Web Server issue): ). "".

9. Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. Adv Bioinformatics. 2008;2008:420747.

10. Huang J, Wang JF, Zhang HS. "[Structure and function of plant C2H2 zinc finger protein]. " Yi Chuan. 2004;26(3):414–8.

11. Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T F, Parcy and Z. I. P. R. G. b (2002). "bZIP transcription factors in Arabidopsis." Trends Plant Sci 7(3): 106–111.

12. Kellogg EA. "Evolutionary history of the grasses" Plant Physiol. 2001;125(3):1198–205.

13. Kim Y, Park S, Gilmour SJ, Thomashow MF. "Roles of CAMTA transcription factors and salicylic acid in configuring the low-temperature transcriptome and freezing tolerance of Arabidopsis.". Plant J. 2013;75(3):364–76.

14. Kleyn PW, Fan W, Kovats SG, Lee JJ, Pulido JC, Wu Y, Berkemeier LR, Misumi DJ, Holmgren L, Charlat O, Woolf EA, Tayber O, Brody T, Shu P, Hawkins F, Kennedy B, Baldini L, Ebeling C, Alperin GD, Deeds J, Lakey ND, Culpepper J, Chen H, Glucksman-Kuis MA, Carlson GA, Duyk GM, Moore KJ. Identification and characterization of the mouse obesity gene tubby: a member of a novel gene family. Cell. 1996;85(2):281–90.

15. Kortschak RD, Tucker PW, Saint R. ARID proteins come in from the desert. Trends Biochem Sci. 2000;25(6):294–9.

16. Lai CP, Lee CL, Chen PH, Wu SH, Yang CC, Shaw JF. "Molecular analyses of the Arabidopsis TUBBY-like protein gene family.". Plant Physiol. 2004;134(4):1586–97.

17. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. "Clustal W and Clustal X version 2.0.". Bioinformatics. 2007;23(21):2947–8.

18. Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, Stein L, Holmes IH, Elsik CG, Lewis SE. "Web Apollo: a web-based genomic annotation editing platform.". Genome Biol. 2013;14(8):R93.

19. Liu Q. "Identification of rice TUBBY-like genes and their evolution.". FEBS J. 2008;275(1):163–71.
20. Mickelbart MV, Hasegawa PM, Bailey-Serres J. "Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability.". Nat Rev Genet. 2015;16(4):237–51.

21. Peng FY, Weselake RJ. "Genome-wide identification and analysis of the B3 superfamily of transcription factors in Brassicaceae and major crop plants.". Theor Appl Genet. 2013;126(5):1305–19.

22. Sasaki K, Imai R. Pleiotropic roles of cold shock domain proteins in plants. Front Plant Sci. 2011;2:116.

23. Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. "New and continuing developments at PROSITE.". Nucleic Acids Res. 2013;41(Database issue):D344–7.

24. Tello-Ruiz MK, Naithani S, Stein JC, Gupta P, Campbell M, Olson A, Wei S, Preece J, Geniza MJ, Jiao Y, Lee YK, Wang B, Mulvaney J, Chougule K, Elser J, Al-Bader N, Kumari S, Thomason J, Kumar V, Bolser DM, Naamati G, Tapanari E, Fonseca N, Huerta L, Iqbal H, Keays M, Munoz-Pomer A, Fuentes A, Tang A, Fabregat P, D'Eustachio J, Weiser LD, Stein R, Petryszak I, Papatheodorou PJ, Kersey P, Lockhart C, Taylor P, Jaiswal, Ware D. "Gramene 2018: unifying comparative genomics and pathway resources for plant research.". Nucleic Acids Res. 2018;46(D1):D1181–9.

25. Tian F, Yang DC, Meng YQ, Jin J, Gao G. "PlantRegMap: charting functional regulatory maps in plants.". Nucleic Acids Res. 2020;48:D1104–13.

26. Wang M, Xu Z, Kong Y. "The tubby-like proteins kingdom in animals and plants.". Gene. 2018;642:16–25.

27. Xu J, Zhang S. "Mitogen-activated protein kinase cascades in signaling plant growth and development.". Trends Plant Sci. 2015;20(1):56–64.

28. Xu JN, Xing SS, Zhang ZR, Chen XS, Wang XY. "Genome-Wide Identification and Expression Analysis of the Tubby-Like Protein Family in the Malus domestica Genome.". Front Plant Sci. 2016;7:1693.

29. Yang Z, Zhou Y, Wang X, Gu S, Yu J, Liang G, Yan C, Xu C. Genomewide comparative phylogenetic and molecular evolutionary analysis of tubby-like protein family in Arabidopsis, rice, and poplar. Genomics. 2008;92(4):246–53.

30. Yu Z, Wang X, Zhang L. (2018). "Structural and Functional Dynamics of Dehydrins: A Plant Protector Protein under Abiotic Stress." Int J Mol Sci 19(11).

31. Zanetti ME, Ripodas C, Niebel A. Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. Biochim Biophys Acta Gene Regul Mech. 2017;1860(5):645–54.

32. Zhang T, Lv W, Zhang H, Ma L, Li P, Ge L, Li G. "Genome-wide analysis of the basic Helix-Loop-Helix (bHLH) transcription factor family in maize.". BMC Plant Biol. 2018;18(1):235.

33. Zhang Y, Lubberstedt T, Xu M. The genetic and molecular basis of plant resistance to pathogens. J Genet Genomics. 2013;40(1):23–35.

34. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ, Zhao PX, Rhee SY, Fei Z. "iTAK: A Program for Genome-wide Prediction and Classification of Plant
Transcription Factors, Transcriptional Regulators, and Protein Kinases. " Mol Plant. 2016;9(12):1667–70.

Figures

**Figure 1**

A flowchart describing the methodology we have developed and applied for the characterization of the TUBBY Transcription Factor proteins.
Figure 2

Web Apollo screenshot of a gene annotation. The gene annotated is TUBBY 11. In a) we see on top the gene structure, with the Iso-Seq data and in b) we see the BAM mapped reads. In this particular case the reads are from root tissue.
Figure 3

Distribution of the TUBBY genes across the ten maize chromosomes. No TUBBY gene was detected in chromosome 7.
Figure 4

Phylogenetic tree of the Zea mays TUBBY genes
Figure 5

a) The general TUBBY model consists of genes with an F-box domain (grey color) and without an F-box domain. b) Computational prediction of the two TUB motifs in each of the alternative spliced proteins. In this prediction the TUB1 (green color) and TUB2 (grey color) motifs are shown as well.
Seven developmental stages were used to examine the expression of each of the thirteen TUBBY transcription factors. The developmental stages are divided into Group 1 (G1) and Group 2 (G2). Group 2 is further divided into Subgroup 1 (SG1) and Subgroup 2 (SG2). The samples refer to the number of samples per developmental stage each of the genes is expressed. As an example, there are 138 experiments performed for the dough stage.

Figure 6
Figure 7

PCR verification of some TUBBY genes in maize stems

Supplementary Files

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

- suptable1.docx
- suppicture1.png