(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 April 2010 (15.04.2010)

PCT

(10) International Publication Number
WO 2010/042862 A2

(51) **International Patent Classification**:
A01H 5/00 (2006.01)

(21) **International Application Number**:
PCT/US2009/060232

(22) **International Filing Date**:
9 October 2009 (09.10.2009)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/195,819 10 October 2008 (10.10.2008) US

(71) **Applicant** *(for all designated States except US)*: **THE SALK INSTITUTE FOR BIOLOGICAL STUDIES** [US/US]; 10010 North Torrey Pines Road, La Jolla, California 92037 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** *(for US only)*: **LOUDET, Olivier** [FR/FR]; Bazainville (FR). **MICHAEL, Todd P.** [US/US]; Piscataway, New Jersey (US). **WEIGEL, Detlef** [US/DE]; Tuebingen (DE). **CHORY, Joanne** [US/US]; Del Mar, California (US).

(74) **Agents**: **HINSCH, Matthew E.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th Floor, San Francisco, California 94111 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: ZINC KNUCKLE PROTEINS

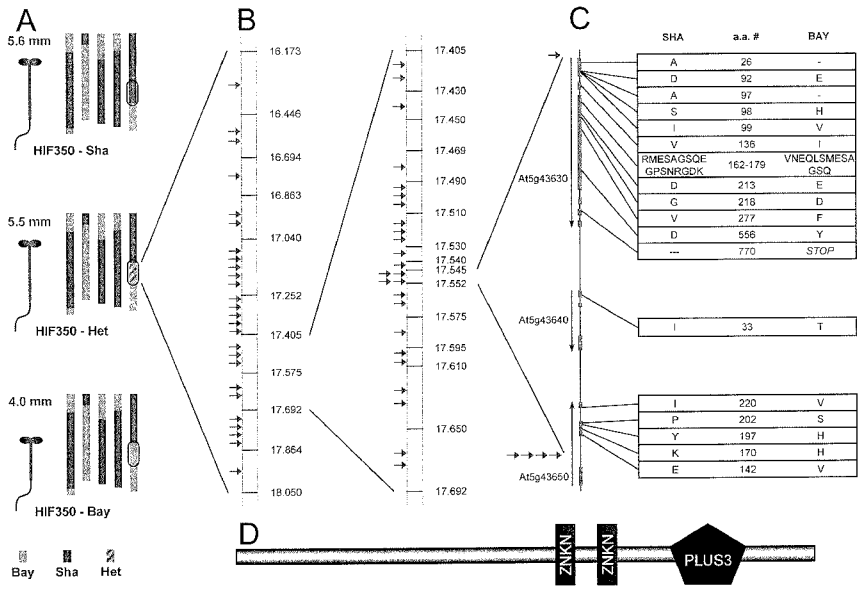

FIGURE 1

(57) **Abstract**: TZP proteins and method of their use for improving plant characteristics are disclosed.

WO 2010/042862 A2

ZINC KNUCKLE PROTEINS

CROSS-REFERENCE TO RELATED PATENT APPLICATIONS

[0001] The present application claims benefit of priority to US Provisional Patent Application No. 61/195,819, filed October 10, 2008, which is incorporated by reference for all purposes.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

[0002] This invention was made with government support under (identify the contract) awarded by the National Institutes of Health (NIH Grant No. GM62932). The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0003] The embryonic stem or hypocotyl is an excellent model for studying both internal and external factors controlling growth in plants (Nozue, K. et al., *Nature*, 448:358-361 (2007)). Genetic screens in common laboratory accessions have yielded direct molecular insight into how light and hormone dependent signaling pathways interact with the circadian clock to regulate the final length of the hypocotyl (Nozue, K. et al., *Nature*, 448:358-361 (2007)). The power of the hypocotyl assay is its simplicity, as well as its obvious meaningfulness. When germinating seeds are exposed to low levels of light, such as those caused by a covering layer of debris, the hypocotyl has to grow for a while. Only after the surface has been broken by the tip of the hypocotyls can the embryonic leaves, the cotyledons, unfold. Conversely, if a seed has fallen on open ground, there is no need for the hypocotyls to be particularly long. Because of the ease and reproducibility with which hypocotyl length can be measured in thousands of individuals, it has also been a powerful model in mapping genes with more subtle effects on light and hormone regulated growth, using methods of quantitative genetics (Borevitz, J.O. et al., *Genetics*, 160:683-696 (2002)). Multiple light signaling genes controlling hypocotyl length have been characterized in quantitative trait locus (QTL) studies (Aukerman, M.J. et al., *Plant Cell*, 9:1317-1326 (1997); Balasubramanian, S. et al., *Nat Genet*, 38:711-715 (2006); Filiault, D.L. et al., *Proc Natl Acad Sci USA*, 105:3157-3162 (2008); Maloof, J.N. et al., *Nat Genet*, 29:441-446 (2001)).

BRIEF SUMMARY OF THE INVENTION

[0004] The present invention provides for plants comprising a heterologous expression cassette comprising a promoter operably linked to a polynucleotide encoding a TZP polypeptide. In some embodiments, the promoter is heterologous to the polynucleotide. In some embodiments, the plant is characterized by increased size, speed of growth, elongation of organs or delay or extension of time of seed and/or fruit production compared to a plant lacking the heterologous expression cassette. In some embodiments, the promoter is a tissue-specific promoter. In some embodiments, the TZP polypeptide is at least 50% (e.g., 60%, 70%, 80%, 90%, 95%, 98%, 99%) identical to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:17, or SEQ ID NO:18. In addition to the above described plants, the present invention also provides for processed plant material, including but not limited to processed plant parts (e.g., seeds, fruit, nuts, tubers, leaves, etc.). The present invention further provides a method of making processed plant material, the method comprising processing a plant or plant part of the present invention, thereby generating processed plant material.

[0005] The present invention also provides for isolated nucleic acids comprising a promoter operably linked to a polynucleotide encoding a TZP polypeptide wherein the promoter is heterologous to the polynucleotide. In some embodiments, the TZP polypeptide is at least 50% (e.g., 60%, 70%, 80%, 90%, 95%, 98%, 99%) identical to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:17, or SEQ ID NO:18.

[0006] The present invention also provides for methods for increasing size, speed of growth or biomass in a plant. In some embodiments, the method comprises, introducing an expression cassette into one or more plants, the expression cassette comprising a promoter operably linked to a polynucleotide encoding a TZP polypeptide, and selecting a plant that has increased size, speed of growth, elongation of organs or delay or extension of time of seed and/or fruit production compared to a plant lacking the heterologous expression cassette. In some embodiments, the TZP polypeptide is at least 50% (e.g., 60%, 70%, 80%, 90%, 95%, 98%, 99%) identical to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID

NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:17, or SEQ ID NO:18.

DEFINITIONS

[0007] The term "promoter" or "regulatory element" refers to a region or sequence determinants located upstream or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. Promoters need not be of plant origin, for example, promoters derived from plant viruses, such as the CaMV35S promoter, can be used in the present invention.

[0008] The term "plant" includes whole plants, shoot vegetative organs/structures (*e.g.* leaves, stems and tubers), roots, flowers and floral organs/structures (*e.g.* bracts, sepals, petals, stamens, carpels, anthers and ovules), seed (including embryo, endosperm, and seed coat) and fruit (the mature ovary), plant tissue (*e.g.* vascular tissue, ground tissue, and the like) and cells (*e.g.* guard cells, egg cells, trichomes and the like), and progeny of same. The class of plants that can be used in the method of the invention is generally as broad as the class of higher and lower plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), gymnosperms, ferns, and multicellular algae. It includes plants of a variety of ploidy levels, including aneuploid, polyploid, diploid, haploid and hemizygous.

[0009] A polynucleotide sequence is "heterologous to" a second polynucleotide sequence if it originates from a foreign species, or, if from the same species, is modified by human action from its original form. For example, a promoter operably linked to a heterologous coding sequence refers to a coding sequence from a species different from that from which the promoter was derived, or, if from the same species, a coding sequence which is different from naturally occurring allelic variants.

[0010] An "expression cassette" refers to a nucleic acid construct, which when introduced into a host cell, results in transcription and/or translation of a RNA or polypeptide, respectively. Antisense constructs or sense constructs that are not or cannot be translated are expressly included by this definition.

[0011] An "TZP polypeptide" is a polypeptide substantially identical to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:17, or SEQ ID NO:18 or as

otherwise described herein. TZP polypeptides comprise a tandem (e.g., two) zinc knuckle domains ($CX_2CX_4HX_4C$; SEQ ID NO:13) as well as a PLUS3 domain. Exemplary TZP genomic and coding sequences are shown in SEQ ID NO:1 and SEQ ID NO:3 , respectively.

[0012] Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman *Add. APL. Math.* 2:482 (1981), by the homology alignment algorithm of Needle man and Wunsch *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson and Lipman *Proc. Natl. Acad. Sci. (U.S.A.)* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

[0013] "Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

[0014] The term "substantial identity" of polypeptide sequences means that a polypeptide comprises a sequence that has at least 25% sequence identity. Alternatively, percent identity can be any integer from 25% to 100%. Exemplary embodiments include at least: 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99%. compared to a reference sequence using the programs described herein; preferably BLAST using standard parameters, as described below. Accordingly, TZP sequences of the invention include nucleic acid sequences encoding a polypeptide that has substantial identity to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, or SEQ ID NO:12. TZP sequences of the invention also include polypeptide sequences having substantial identify to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, or SEQ ID NO:12. One of skill will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking

into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Polypeptides which are "substantially similar" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, aspartic acid-glutamic acid, and asparagine-glutamine.

[0015] Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other, or a third nucleic acid, under stringent conditions. Stringent conditions are sequence dependent and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent conditions will be those in which the salt concentration is about 0.02 molar at pH 7 and the temperature is at least about 60°C.

[0016] The term “isolated”, when applied to a nucleic acid or protein, denotes that the nucleic acid or protein is essentially free of other cellular components with which it is associated in the natural state. It can be, for example,in a homogeneous state and may be in either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is substantially purified.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Figure 1. Confirmation and fine-mapping of the LIGHT5 QTL to three genes, among which At5g43630 is highly polymorphic between Bay-0 and Shahdara. (A) Confirmation of LIGHT5 using heterogeneous inbred families (HIFs). The Sha allele is fully dominant over the Bay allele. Two rounds of recombinant screening from HIF350 followed. (B) Horizontal marks on the chromosomes are markers with physical position in Mb indicated to the right. Red arrows indicate the approximate position of different recombination events that were individually tested in rHIF to establish the QTL position. Fine-mapping identified a 7kb-region containing three genes. (C) Grey boxes along the vertical axes represent exons from three genes highlighted by vertical arrows. Horizontal red arrows indicate the exact physical position of the last 5 recombinants defining the QTL candidate region. Amino-acid changes between Bay-0 and Shahdara within the interval are presented in the table (RMESAGSQEGPSNRGDK = SEQ ID NO:14; VNEQLSMESAGSQ = SEQ ID NO:15) and linked to their respective physical position along the gene model. (D) Model of the protein structure of TZP (At5g43630).

[0018] **Figure 2.** LIGHT5/TZP controls growth throughout development. (**A**) Increased TZP activity results in longer hypocotyls under blue light. Seedlings were grown at the fluence indicated and measured on the sixth day. *TZP-OX(3)* and *TZP-OX(5)* are two independent transgenic lines overexpressing *TZP* in rHIF138-8 background. rHIF138-8 contains the Bay allele with the premature STOP in *TZP* and rHIF138-13 contains the functional Sha allele. (**B**) Sha allele of *TZP* or overexpression result in longer hypocotyls. Plants were grown under light/dark cycles (12 hrs/12 hrs) at 22°C and hypocotyls were measured 7 DAG. Representative seedlings were used to make the images in (C). Measurements represent three independent experiments of twenty seedlings each. (**C**) *TZP-OX* plants have long hypocotyls 7 DAG under light/dark cycles (12 hrs/12 hrs) compared to its background (rHIF138-8). (**D**) *TZP-OX* petioles are elongated 24 DAG compared to rHIF138-8. (**E**) Mature *TZP-OX* plants are almost twice as tall as background plants 50 DAG. Plants were grown under long days (light/dark: 16 hrs/8 hrs) for 50 days. All plants pictured are representative of at least two independent transgenic lines (lines 3 and 5) and two independent experiments. The vertical bar represents 20cm. (**F**) TZP::YFP is localized to speckles in the nucleus. T2 plants carrying the *35S::TZP:YFP* fusion were imaged to detect

TZP localization. TZP:YFP localizes to the nucleus in guard cells of the stomata, in addition to all other tissues tested. Grey lines highlight the cell walls.

[0019] Figure 3. LIGHT5/*TZP* controls morning-specific growth pathways. (**A**) *TZP* displays dawn-specific transcript abundance under light/dark cycles (12 hrs/12 hrs) and constant 22°C (six time points). The second day of data is copied from the first (double plotted) for visualization purposes. Expression was determined by qPCR with primers specific to *TZP* and SyberGreen. (**B**) *TZP* transcript abundance is overexpressed in *TZP-OX*. Five independent lines overexpressing *TZP* were characterized. Two lines are shown here. Data were collected and plotted as in (A). (**C**) The genes that are misexpressed ($P < 0.01$) in long hypocotyl genotypes (*TZP-OX* or rHIF138-13 vs. rHIF138-8) under blue light short day photocycles are expressed at dawn as determined with PHASER. (**D**) Long hypocotyls of the *TZP-OX* mutant are due in part to overexpression of cell wall genes. As an example of the expression pattern of the cell wall genes that are overexpressed in the *TZP-OX* mutant, *IRX1* continues to cycle with peak expression at dawn, but its peak expression is 3 fold higher in *TZP-OX*. Other genes that are overexpressed in *TZP-OX* are listed in table S3 and S10 from our website data at www.inra.fr/vast/Files/Loudet_PNAS_SITables.xls. Results are from array data.

[0020] Figure 4. Phenotypic variation and significant QTL detected in the Bay-0 x Shahdara RIL set. Distribution of phenotypic values for hypocotyl elongation among 164 Bay-0 x Shahdara RILs across four different environments: two temperature x two white-light conditions (**A**). 'B' and 'S' above bars indicate phenotypic values obtained for the parents, Bay-0 and Shahdara respectively. Position and effect of the five significant QTL controlling this variation (**B**). Each QTL is depicted by a triangle located at the most probable QTL position on one of the 5 chromosomes. Upward- and downward-pointing triangles represent QTLs with a positive or negative allelic effect, respectively ('2a' in table S1 from www.inra.fr/vast/Files/Loudet_PNAS_SITables.xls, which represents the mean effect of the replacement of both Shahdara alleles by Bay-0 alleles at the QTL). The framework genetic map (horizontal marks indicate marker positions) is from Loudet *et al.* (2002). QTLs denoted as "LIGHT" were detected in all four environments. Other QTLs are named according to the specific environment in which they had a significant effect.

[0021] Figure 5. *LIGHT5* is At5g43630 and the causal polymorphism is the 8bp-insertion in Bay-0. Two rHIFs segregating for *LIGHT5* and defining the limits of the candidate

interval were crossed to generate an advanced rHIF (arHIF47) segregating solely for the 7kb candidate region. Phenotypes are shown for the arHIF47 progeny fixed for either the Bay or Sha 7-kb region. Bay-0[41AV] lacks the 8 bp insertion causing the early stop in At5g43630, with the rest of the genome being the same as in Bay-0. Phenotypes are shown from F2 plants between the two isogenic parents. Different letters on bars indicate significantly different means ($P < 0.01$; least significant difference test).

[0022] Figure 6. The PLUS3 domain is conserved across species. (A) Alignment of the PLUS3 domain in multiple species (SEQ ID NOS:19-22 and 24-36) and the consensus of 96 *Arabidopsis thaliana* accessions (AtTZPaccessions = SEQ ID NO:23).

(B) Phylogenetic tree of the PLUS3 domain across species and with other PLUS3 domains of *Arabidopsis thaliana*. The *TZP* PLUS3 domain is more closely related to the PLUS3 domain from *TZP* orthologs in other species, than other PLUS3 domains in *Arabidopsis thaliana*. *Arabidopsis lyrata (Al)*, *Carica papaya (Cp)*, *Oryza sativa (Os)*, *Ricinus communis (Rc)*, *Populus trichocarpa (Pt)*, *Sorghum bicolor (Sb)*, *Brachypodium distachyon (Bd)*, *Glycine max* (*Gm*), *Vitis vinifera* (Vv) and *Selaginella moellendorffii (Sm)*.

[0023] Figure 7. *TZP* expression is correlated with increased hypocotyl growth. Plants were grown under continuous blue light for five days and tissue was collected at dawn. Expression was measured by qPCR. Expression of TZP was higher in lines with longer hypocotyls, including *TZP-OX (3)*, *rHIF138-13*, *arHIF47-5* and Shahdara. The bar for *TZP-OX (3)* was truncated due to the expression being on a different scale than the rest of the lines.

[0024] Figure 8. *TZP* transcript abundance peaks at dawn in the rHIF and the accession Bay-0 and Shahdara. Plants were grown for seven days under light/dark cycles and sampled over one day. Expression was measured using qPCR.

[0025] Figure 9. *TZP* expression is disrupted in core circadian clock mutants. The expression of *TZP* was measured in two circadian clock mutants, *late elongated hypocotyl* (*lhy*) and *early flowering 3* (*elf3-7*) under short day photocycles (8 hrs light/ 16 hrs dark at 22°C).

[0026] Figure 10. Core circadian clock gene expression is not disrupted in *TZP-OX*. Both *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *GIGANTEA* (*GI*) display completely wild type expression in *TZP-OX* as compared to HIF138-8.

[0027] **Figure 11.** Genes upregulated under light/dark cycles in *TZP-OX* are expressed at dawn. Normally, the 117 genes that are upregulated by *TZP-OX* under light/dark cycles are phased to dawn. Overrepresentation plot created using the PHASER tool.

[0028] **Figure 12.** *TZP* overexpression specifically disrupts light-specific, time-of-day growth pathways.

(A) *PEROXIDASE30*

(B) *DFL1*

(C) *HAT4*

(D) *CATALASE3* (*CAT3*)

(E) *PW9* (MATH/TRAF)

(F) *SWI/PLUS3* (At2g16480)

(G) *PIF5*

(H) *PIF4*

(I) *HFR1*

(J) *PKS1*

(K) *PHYB*

(L) *PHYD*

[0029] **Figure 13.** TZP is conserved among species. A. Alignment of full-length TZP sequences from different plant species obtained by BLAST searches using Phytozome (SEQ ID NOS:37, 2, 10, 9, 8, 6, 7, 11, 5 and 12, respectively). The conservation scoring is performed by PRALINE. Increased sequence similarity and identity is observed near the C-terminus of the protein, where the PLUS3 domain resides. B. Phylogenetic tree of TZP protein sequence across species was performed based on Clustal W alignment using MegAlign: *Arabidopsis thaliana* (At), *Arabidopsis lyrata* (Al), *Oryza sativa* (Os), *Populus trichocarpa* (Pt), *Glycine max* (Gm), *Brachypodium distachium* (Bd), *Vitis vinifera* (Vv), *Carica papaya* (Cp), *Selanginella moellondorfii* (Sm), *Ricinus communis* (Rc).

DETAILED DESCRIPTION

I. Introduction

[0030] The present invention provides for methods of improving plant traits including but not limited to increasing plant organ elongation, increasing plant height, delaying seed or fruit set, increasing the time of fruit or seed set, and/or increasing plant size. In some

embodiments of the invention, a TZP polypeptide is ectopically or otherwise overexpressed. In some embodiments, for example, the invention provides for transgenic plants comprising a heterologous expression cassette for expressing a TZP polypeptide in the plant.

[0031] Alternatively, the present invention provides for plants with shortened stature. In these embodiments, endogenous TZP expression is reduced, e.g., by use of siRNA, antisense, or other gene expression reduction technology.

II. Use of nucleic acids of the invention to enhance gene expression

[0032] Isolated nucleic acid sequences encoding all or part of a TZP polypeptide can be used to prepare expression cassettes that enhance, or increase endogenous, *TZP* gene expression. Where overexpression of a gene is desired, the desired gene from a different species may be used to decrease potential sense suppression effects.

[0033] Any of a number of means well known in the art can be used to increase TZP activity in plants. Any organ can be targeted, such as shoot vegetative organs/structures (*e.g.* leaves, stems and tubers), roots, flowers and floral organs/structures (*e.g.* bracts, sepals, petals, stamens, carpels, anthers and ovules), seed (including embryo, endosperm, and seed coat) and fruit. Alternatively, one or several TZP genes can be expressed constitutively (*e.g.*, using the CaMV 35S promoter).

[0034] One of skill will recognize that the polypeptides encoded by the genes of the invention, like other proteins, have different domains which perform different functions. Thus, the gene sequences need not be full length, so long as the desired functional domain of the protein is expressed.

III. Preparation of recombinant vectors

[0035] To use isolated sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. *See*, for example, Weising *et al. Ann. Rev. Genet.* 22:421-477 (1988). A DNA sequence coding for the desired polypeptide, for example a cDNA sequence encoding a full length protein, will preferably be combined with transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the gene in the intended tissues of the transformed plant.

[0036] For example, for overexpression, a plant promoter fragment may be employed which will direct expression of the gene in all tissues of a regenerated plant. Such promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of Agrobacterium tumafaciens, and other transcription initiation regions from various plant genes known to those of skill.

[0037] Alternatively, the plant promoter may direct expression of the polynucleotide of the invention in a specific tissue (tissue-specific promoters) or may be otherwise under more precise environmental control (inducible promoters). Examples of tissue-specific promoters under developmental control include promoters that initiate transcription only in certain tissues, such as fruit, seeds, or flowers. Examples of environmental conditions that may affect transcription by inducible promoters include anaerobic conditions, elevated temperature, or the presence of light.

[0038] If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA.

[0039] The vector comprising the sequences (e.g., promoters or coding regions) from genes of the invention can optionally comprise a marker gene that confers a selectable phenotype on plant cells. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosluforon or Basta.

[0040] In some embodiments, *TZP* nucleic acid sequences of the invention are expressed recombinantly in plant cells to enhance and increase levels of *TZP* polypeptides. Alternatively, antisense or other *TZP* constructs are used to suppress *TZP* levels of expression. A variety of different expression constructs, such as expression cassettes and vectors suitable for transformation of plant cells can be prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, e.g., Weising *et al. Ann. Rev. Genet.* 22:421-477 (1988). A *TZP* sequence coding for a TZP polypeptide, e.g., a cDNA sequence encoding a full length protein, can be combined with cis-acting (promoter) and trans-acting (enhancer) transcriptional regulatory sequences to direct the timing, tissue type and levels of

transcription in the intended tissues of the transformed plant. Translational control elements can also be used.

[0041] The invention provides an TZP nucleic acid operably linked to a promoter that, in some embodiments, is capable of driving the transcription of the TZP coding sequence in plants. The promoter can be, e.g., derived from plant or viral sources. The promoter can be, e.g., constitutively active, inducible, or tissue specific. In construction of recombinant expression cassettes, vectors, transgenics, of the invention, a different promoters can be chosen and employed to differentially direct gene expression, e.g., in some or all tissues of a plant or animal. Typically, as discussed above, desired promoters are identified by analyzing the 5' sequences of a genomic clone corresponding to the TZP genes described here.

A. Constitutive Promoters

[0042] A promoter fragment can be employed that will direct expression of TZP nucleic acid in all transformed cells or tissues, e.g. as those of a regenerated plant. The term "constitutive regulatory element" means a regulatory element that confers a level of expression upon an operatively linked nucleic molecule that is relatively independent of the cell or tissue type in which the constitutive regulatory element is expressed. A constitutive regulatory element that is expressed in a plant generally is widely expressed in a large number of cell and tissue types. Promoters that drive expression continuously under physiological conditions are referred to as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation.

[0043] A variety of constitutive regulatory elements useful for ectopic expression in a transgenic plant are well known in the art. The cauliflower mosaic virus 35S (CaMV 35S) promoter, for example, is a well-characterized constitutive regulatory element that produces a high level of expression in all plant tissues (Odell e*t al.*, *Nature* 313:810-812 (1985)). The CaMV 35S promoter can be particularly useful due to its activity in numerous diverse plant species (Benfey and Chua, *Science* 250:959-966 (1990); Futterer *et al.*, *Physiol. Plant* 79:154 (1990); Odell *et al.*, *supra*, 1985). A tandem 35S promoter, in which the intrinsic promoter element has been duplicated, confers higher expression levels in comparison to the unmodified 35S promoter (Kay *et al.*, *Science* 236:1299 (1987)). Other useful constitutive regulatory elements include, for example, the cauliflower mosaic virus 19S promoter; the

Figwort mosaic virus promoter; and the nopaline synthase (nos) gene promoter (Singer *et al.*, *Plant Mol. Biol.* 14:433 (1990); An, *Plant Physiol.* 81:86 (1986)).

[0044] Additional constitutive regulatory elements including those for efficient expression in monocots also are known in the art, for example, the pEmu promoter and promoters based on the rice Actin-1 5' region (Last et al., *Theor. Appl. Genet.* 81:581 (1991); Mcelroy *et al.*, *Mol. Gen. Genet.* 231:150 (1991); Mcelroy *et al.*, *Plant Cell* 2:163 (1990)). Chimeric regulatory elements, which combine elements from different genes, also can be useful for ectopically expressing a nucleic acid molecule encoding an TZP polynucleotide (Comai *et al.*, *Plant Mol. Biol.* 15:373 (1990)).

[0045] Other examples of constitutive promoters include the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens* (see, e.g., Mengiste (1997) supra; O'Grady (1995) *Plant Mol. Biol.* 29:99-108); actin promoters, such as the *Arabidopsis* actin gene promoter (see, e.g., Huang (1997) *Plant Mol. Biol.* 1997 33:125-139); alcohol dehydrogenase (Adh) gene promoters (see, e.g., Millar (1996) *Plant Mol. Biol.* 31:897-904); *ACT11* from *Arabidopsis* (Huang *et al. Plant Mol. Biol.* 33:125-139 (1996)), *Cat3* from *Arabidopsis* (GenBank No. U43147, Zhong *et al.*, *Mol. Gen. Genet.* 251:196-203 (1996)), the gene encoding stearoyl-acyl carrier protein desaturase from *Brassica napus* (Genbank No. X74782, Solocombe *et al. Plant Physiol.* 104:1167-1176 (1994)), *GPc1* from maize (GenBank No. X15596, Martinez *et al. J. Mol. Biol* 208:551-565 (1989)), *Gpc2* from maize (GenBank No. U45855, Manjunath *et al.*, *Plant Mol. Biol.* 33:97-112 (1997)), other transcription initiation regions from various plant genes known to those of skill. *See also* Holtorf *Plant Mol. Biol.* 29:637-646 (1995).

B. Inducible Promoters

[0046] Alternatively, a promoter may direct expression of the *TZP* nucleic acid of the invention under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters. For example, the invention incorporates the drought-inducible promoter of maize (Busk (1997) supra); the cold, drought, and high salt inducible promoter from potato (Kirch (1997) *Plant Mol. Biol.* 33:897-909).

[0047] Alternatively, plant promoters which are inducible upon exposure to plant hormones, such as auxins, are used to express the nucleic acids of the invention. For example, the invention can use the auxin-response elements E1 promoter fragment (AuxREs) in the soybean (*Glycine max* L.) (Liu (1997) *Plant Physiol.* 115:397-407); the auxin-responsive *Arabidopsis* GST6 promoter (also responsive to salicylic acid and hydrogen peroxide) (Chen (1996) *Plant J.* 10: 955-966); the auxin-inducible parC promoter from tobacco (Sakai (1996) 37:906-913); a plant biotin response element (Streit (1997) *Mol. Plant Microbe Interact.* 10:933-937); and, the promoter responsive to the stress hormone abscisic acid (Sheen (1996) *Science* 274:1900-1902).

[0048] Promoters that are inducible upon exposure to chemicals reagents applied to the plant, such as herbicides or antibiotics, can also be used to express the nucleic acids of the invention. For example, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, can be used (De Veylder (1997) *Plant Cell Physiol.* 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. *TZP* coding sequence can also be under the control of, e.g., a tetracycline-inducible promoter, e.g., as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) *Plant J.* 11:465-473); or, a salicylic acid-responsive element (Stange (1997) *Plant J.* 11:1315-1324; Uknes *et al.*, *Plant Cell* 5:159-169 (1993); Bi *et a*l., *Plant J.* 8:235-245 (1995)).

[0049] Other inducible regulatory elements include but are not limited to copper-inducible regulatory elements (Mett *et al.*, *Proc. Natl. Acad. Sci. USA* 90:4567-4571 (1993); Furst *et al.*, *Cell* 55:705-717 (1988)); tetracycline and chlor-tetracycline-inducible regulatory elements (Gatz *et al.*, *Plant J.* 2:397-404 (1992); Röder *et al.*, *Mol. Gen. Genet.* 243:32-38 (1994); Gatz, *Meth. Cell Biol.* 50:411-424 (1995)); ecdysone inducible regulatory elements (Christopherson *et al.*, *Proc. Natl. Acad. Sci. USA* 89:6314-6318 (1992); Kreutzweiser *et al.*, *Ecotoxicol. Environ. Safety* 28:14-24 (1994)); heat shock inducible regulatory elements (Takahashi *et al.*, *Plant Physiol.* 99:383-390 (1992); Yabe *et al.*, *Plant Cell Physiol.* 35:1207-1219 (1994); Ueda *et al.*, *Mol. Gen. Genet.* 250:533-539 (1996)); and *lac* operon elements, which are used in combination with a constitutively expressed lac repressor to confer, for example, IPTG-inducible expression (Wilde *et al.*, *EMBO J.* 11:1251-1259 (1992)). An inducible regulatory element useful in the transgenic plants of the invention also can be, for example, a nitrate-inducible promoter derived from the spinach nitrite reductase gene (Back

et al., *Plant Mol. Biol.* 17:9 (1991)) or a light-inducible promoter, such as that associated with the small subunit of RuBP carboxylase or the LHCP gene families (Feinbaum *et al.*, *Mol. Gen. Genet.* 226:449 (1991); Lam and Chua, *Science* 248:471 (1990)).

C. Tissue-Specific Promoters

[0050] Alternatively, the plant promoter may direct expression of the polynucleotide of the invention in a specific tissue (tissue-specific promoters). Tissue specific promoters are transcriptional control elements that are only active in particular cells or tissues at specific times during plant development, such as in vegetative tissues or reproductive tissues.

[0051] Examples of tissue-specific promoters under developmental control include promoters that initiate transcription only (or primarily only) in certain tissues, such as vegetative tissues, e.g., roots or leaves, or reproductive tissues, such as fruit, ovules, seeds, pollen, pistols, flowers, or any embryonic tissue. Reproductive tissue-specific promoters may be, e.g., ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed and seed coat-specific, pollen-specific, petal-specific, sepal-specific, or some combination thereof.

[0052] Other tissue-specific promoters include seed promoters. Suitable seed-specific promoters are derived from the following genes: *MAC1* from maize (Sheridan (1996) *Genetics* 142:1009-1020); *Cat3* from maize (GenBank No. L05934, Abler (1993) *Plant Mol. Biol.* 22:10131-1038); *vivparous-1* from *Arabidopsis* (Genbank No. U93215); *atmyc1* from *Arabidopsis* (Urao (1996) *Plant Mol. Biol.* 32:571-57; Conceicao (1994) *Plant* 5:493-505); *napA* from *Brassica napus* (GenBank No. J02798, Josefsson (1987) JBL 26:12196-1301); and the napin gene family from *Brassica napus* (Sjodahl (1995) *Planta* 197:264-271).

[0053] A variety of promoters specifically active in vegetative tissues, such as leaves, stems, roots and tubers, can also be used to express the *TZP* nucleic acids of the invention. For example, promoters controlling patatin, the major storage protein of the potato tuber, can be used, see, e.g., Kim (1994) *Plant Mol. Biol.* 26:603-615; Martin (1997) *Plant J.* 11:53-62. The ORF13 promoter from Agrobacterium rhizogenes which exhibits high activity in roots can also be used (Hansen (1997) *Mol. Gen. Genet.* 254:337-343. Other useful vegetative tissue-specific promoters include: the tarin promoter of the gene encoding a globulin from a major taro (*Colocasia esculenta* L. Schott) corm protein family, tarin (Bezerra (1995) *Plant Mol. Biol.* 28:137-144); the curculin promoter active during taro corm development (de

Castro (1992) *Plant Cell* 4:1549-1559) and the promoter for the tobacco root-specific gene TobRB7, whose expression is localized to root meristem and immature central cylinder regions (Yamamoto (1991) *Plant Cell* 3:371-382).

[0054] Leaf-specific promoters, such as the ribulose biphosphate carboxylase (RBCS) promoters can be used. For example, the tomato RBCS1, RBCS2 and RBCS3A genes are expressed in leaves and light-grown seedlings, only RBCS1 and RBCS2 are expressed in developing tomato fruits (Meier (1997) *FEBS Lett.* 415:91-95). A ribulose bisphosphate carboxylase promoters expressed almost exclusively in mesophyll cells in leaf blades and leaf sheaths at high levels, described by Matsuoka (1994) *Plant J.* 6:311-319, can be used. Another leaf-specific promoter is the light harvesting chlorophyll a/b binding protein gene promoter, see, e.g., Shiina (1997) *Plant Physiol.* 115:477-483; Casal (1998) *Plant Physiol.* 116:1533-1538. The Arabidopsis thaliana myb-related gene promoter (Atmyb5) described by Li (1996) *FEBS Lett.* 379:117-121, is leaf-specific. The Atmyb5 promoter is expressed in developing leaf trichomes, stipules, and epidermal cells on the margins of young rosette and cauline leaves, and in immature seeds. Atmyb5 mRNA appears between fertilization and the 16 cell stage of embryo development and persists beyond the heart stage. A leaf promoter identified in maize by Busk (1997) *Plant J.* 11:1285-1295, can also be used.

[0055] Another class of useful vegetative tissue-specific promoters are meristematic (root tip and shoot apex) promoters. For example, the "*SHOOTMERISTEMLESS*" and "*SCARECROW*" promoters, which are active in the developing shoot or root apical meristems and are described by Di Laurenzio (1996) *Cell* 86:423-433; and, Long (1996) *Nature* 379:66-69, can be used. Another promoter is the 3-hydroxy-3- methylglutaryl coenzyme A reductase HMG2 gene promoter, whose expression is restricted to meristematic and floral (secretory zone of the stigma, mature pollen grains, gynoecium vascular tissue, and fertilized ovules) tissues (see, e.g., Enjuto (1995) *Plant Cell.* 7:517-527). Additional promoter examples include the kn1-related gene promoters from maize and other species that show meristem-specific expression, see, e.g., Granger (1996) *Plant Mol. Biol.* 31:373-378; Kerstetter (1994) *Plant Cell* 6:1877-1887; Hake (1995) *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 350:45-51. One such example is the *Arabidopsis thaliana* KNAT1 promoter. In the shoot apex, KNAT1 transcript is localized primarily to the shoot apical meristem; the expression of KNAT1 in the shoot meristem decreases during the floral transition and is restricted to the cortex of the inflorescence stem (see, e.g., Lincoln (1994) *Plant* Cell 6:1859-1876).

[0056] One of skill will recognize that a tissue-specific promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, as used herein a tissue-specific promoter is one that drives expression preferentially in the target tissue, but may also lead to some expression in other tissues as well.

[0057] In another embodiment, a *TZP* nucleic acid is expressed through a transposable element. This allows for constitutive, yet periodic and infrequent expression of the constitutively active polypeptide. The invention also provides for use of tissue-specific promoters derived from viruses which can include, e.g., the tobamovirus subgenomic promoter (Kumagai (1995) *Proc. Natl. Acad. Sci. USA* 92:1679-1683; the rice tungro bacilliform virus (RTBV), which replicates only in phloem cells in infected rice plants, with its promoter which drives strong phloem-specific reporter gene expression; the cassava vein mosaic virus (CVMV) promoter, with highest activity in vascular elements, in leaf mesophyll cells, and in root tips (Verdaguer (1996) *Plant Mol. Biol.* 31:1129-1139).

V. Production of transgenic plants

[0058] DNA constructs of the invention may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional Agrobacterium tumefaciens host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria.

[0059] Microinjection techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al. *EMBO J.* 3:2717-2722 (1984). Electroporation techniques are described in Fromm et *al. Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein et al. *Nature* 327:70-73 (1987).

[0060] *Agrobacterium tumefaciens*-mediated transformation techniques, including disarming and use of binary vectors, are well described in the scientific literature. See, for

example Horsch *et al. Science* 233:496-498 (1984), and Fraley *et al. Proc. Natl. Acad. Sci. USA* 80:4803 (1983).

[0061] Transformed plant cells that are derived from any transformation technique can be cultured to regenerate a whole plant that possesses the transformed genotype and thus the desired phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, optionally relying on a biocide and/or herbicide marker that has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans et al., Protoplasts Isolation and Culture, Handbook of Plant Cell Culture, pp. 124-176, MacMillilan Publishing Company, New York, 1983; and Binding, Regeneration of Plants, Plant Protoplasts, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al. Ann. Rev. of Plant Phys.* 38:467-486 (1987).

[0062] The nucleic acids of the invention can be used to confer desired traits on essentially any plant. Thus, the invention has use over a broad range of plants, including species from the genera *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucumis*, *Cucurbita*, *Daucus*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Oryza*, *Panieum*, *Pannesetum*, *Persea*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Trigonella*, *Triticum*, *Vitis*, *Vigna*, and, *Zea*.

EXAMPLES

[0063] The following examples are offered to illustrate, but not to limit the claimed invention.

[0064] In this example, we utilize the hypocotyl assay to identify QTL controlling growth in two light and two temperature conditions. We identified a recessive large effect QTL on chromosome five controlling 40% of the growth variation segregating in Recombinant Inbred Lines (RILs) derived from the Bay-0 and Shahdara accessions of *A. thaliana*. The QTL was fine-mapped and the causal factor shown to be a mutation affecting a tandem zinc knuckle/PLUS3 protein (At5g43630; *TZP*), which is encoded by a single-copy gene in all completely sequenced plant genomes. We show that TZP acts downstream of the circadian clock and light signaling, directly regulating blue light-dependent, morning (dawn) specific

growth, during seedling development and beyond. Based on its nuclear localization and its novel domain structure, we argue that TZP functions at the transcriptional level to control growth-promoting pathways. TZP represents a new component of the growth pathway that was not previously identified using traditional genetic screens.

Mapping QTL for hypocotyl elongation in the Bay-0 x Shahdara RIL population

[0065] Intraspecific variation provides a fertile source of genetic combinations that can be utilized to map new genes (or new alleles) involved in complex traits such as growth. To investigate natural variation for hypocotyl elongation response to light and temperature, we phenotyped a core set of 164 RILs from the Bay-0 x Shahdara cross in four different environments combining two white-light (17 $\mu mol/m^{-2}s^{-1}$ [L1] and 10 $\mu mol/m^{-2}s^{-1}$ [L2]) and two temperature (22°C and 26°C) conditions (supporting information [SI] Fig. 4A). The parental phenotypes reveal that in our conditions Shahdara responds poorly to temperatures above 22°C or light below 17 $\mu mol.m^{-2}.s^{-1}$, or a combination of both. In contrast, Bay-0 responds strongly to both temperature and light, with a synergistic interaction between both factors. Variation among the RILs seems to follow parental variation with signs of bimodality (especially at 26°C / L1 and 26°C / L2) suggesting the segregation of some large-effect QTL. Transgression was also prevalent in all conditions and in both directions. Overall, genotypic variation was significant in each environment and broad-sense heritability of the trait was accordingly high, above 70% (see table S1 from our website data at www.inra.fr/vast/Files/Loudet_PNAS_SITables.xls). While RIL response to contrasted temperature and light treatments was significant ($P<0.001$), RIL x light interactions were not significant at either 22°C or 26°C, and the RIL x temperature interaction was only significant under L1 (data not shown). This indicates that most of the phenotypic variation between RILs is stable and conserved across environments.

[0066] The genetic architecture of variation in hypocotyl elongation under these environmental conditions is presented in SI Fig. 4 B (see also table S1 from our website data). Two loci with major effects are detected across all environments and called *LIGHT1* and *LIGHT5*, whereas the remaining loci called '*HYP*' are specific to a single environment with more subtle phenotypic contributions (only 4% each). The identification of these two major effect QTL is in accordance with the meager RIL x environment interactions found. The *LIGHT5* locus explains over 40% of the variance, with no *LIGHT5* x environment interaction found in any condition (data not shown). Its negative allelic effect is predicted to represent a combined 2.1 to 2.5 mm increase in hypocotyl elongation contributed by

Shahdara alleles (Sha) relative to the Bay-0 alleles (Bay). In contrast, Sha alleles at *LIGHT1* are responsible for a decrease in hypocotyl elongation compared to Bay alleles, but with a relatively smaller phenotypic contribution (explaining 25 to 30% of the variance). The opposite allelic effects of *LIGHT1* and *LIGHT5* are responsible for most of the transgression observed in each environment. *LIGHT1* interacts with temperature under either L1 or L2 conditions and with light at either 22°C or 26°C (data not shown). There is no significant epistatic relationship between *LIGHT1* and *LIGHT5*, or among any other pair of loci.

Confirmation and fine-mapping of *LIGHT5* to three candidate genes

[0067] Identifying the Quantitative Trait Gene (QTG) underlying a QTL is a challenging task that requires several independent lines of proof that a gene is linked to a trait of interest and that variation in this gene explains the trait (Weigel, D. and Nordborg, M., *Plant Physiol*, 138:567-568 (2005)). The most general approach is fine-mapping to a very small physical candidate interval, which in the best case allows immediate identification of candidate polymorphisms (Quantitative Trait Nucleotide, QTN) within the causal gene or regulatory regions (Clark, R.M. et al., *Science*, 317:338-342 (2007)). For most QTL and situations in *A. thaliana*, phenotyping for a QTL effect remains much more limiting than genotyping many individuals. Therefore, an efficient strategy for fine-mapping is to first isolate recombinants within a segregating nearly-isogenic line based on genotype alone and then to phenotypically interrogate only the informative ones (in successive rounds) to reduce the candidate interval to the gene level (Peleman, J.D. et al., *Genetics*, 171:1341-1352 (2005)).

[0068] We followed the HIF strategy to build nearly-isogenic lines from a RIL (RIL350) that was segregating solely for the *LIGHT5* region. Comparing plants homozygous for the Sha allele with plants homozygous for the Bay allele at the QTL region (in an otherwise identical genetic background) confirmed the phenotypic impact of *LIGHT5* on hypocotyl elongation (Fig. 1A). HIF350-Sha hypocotyls are consistently 1.6 mm longer than those of HIF350-Bay (slightly less than predicted by the QTL analysis). The analysis of heterozygous plants showed that the Sha allele of *LIGHT5* is fully dominant over the Bay allele (Fig. 1A). The phenotypic effect observed was identical when first fixing alternate genotypes at the QTL region and then comparing the phenotypes of the descendants produced by those homozygous plants ("fixed progeny") or when directly studying the segregating descendants of a heterozygous plant ("progeny testing"). This precludes any maternal phenotypic effect and demonstrates a direct control of the phenotype expressed in seedlings by the *LIGHT5* alleles.

[0069] Screening for recombinants by genotyping 600 plants descended from the initial RIL350 individual (heterozygous over the whole QTL region) allowed us to identify 80 recombinants (recombined HIF, or rHIF) over the ~1.9 Mb heterozygous region. Twenty-six of these rHIF were individually interrogated in successive rounds of progeny testing to score for the presence (or absence) of the QTL effect caused by the remaining interval. This first screen allowed us to narrow the candidate region to less than 300 kb between markers at 17.405 and 17.692 Mb (Fig. 1B). Screening 4,000 descendants from one of the positive rHIFs found in the previous step allowed us to identify 73 new rHIFs within the 300 kb candidate interval; 25 of these were individually progeny tested to define a smaller interval containing the QTL (Fig. 1B). Results were consistent among rHIFs, and the last five recombinants delimited a 7 kb interval, from 17.545 Mb (one recombinant) to 17.552 Mb (four independent recombinants; Fig. 1C). Three predicted genes, At5g43630, At5g43640, and At5g43650, are at least partially included in the 7 kb candidate region. We crossed two independent rHIFs with appropriate genotypes (see SI Fig. 5) in order to generate a line that was segregating only for the candidate region (and fixed to the north and south), following a strategy suggested by Kroymann and Mitchell-Olds (Kroymann, J. and Mitchell-Olds, T., *Nature*, 435:95-98 (2005)) that we named 'advanced rHIF (arHIF) cross'. This approach confirmed that the 17.545 – 17.552 Mb interval was sufficient to recapitulate the *LIGHT5* phenotype (SI Fig. 5).

[0070] Sequencing the 7 kb interval in Bay-0 and Shahdara revealed dozens of SNPs and several indels, many of which resulted in non-silent changes in the coding regions of the three candidate genes. Eight, one and five SNPs caused amino-acid changed in At5g43630, At5g43640 and At5g43650, respectively. Two single amino-acid deletions and one larger deletion were discovered in At5g43630. Finally, an 8 bp insertion in Bay-0 caused a frameshift and a premature stop within the coding region for the predicted PLUS3 domain of At5g43630 (Figs. 1C and D).

Identification of the *LIGHT5* causal gene as TANDEM ZINC KNUCKLE/PLUS3 (TZP)

[0071] The three genes in the *LIGHT5* interval are annotated as encoding a tandem zinc knuckle/PLUS3 (*TZP*, At5g43630), a 40S ribosomal protein (*RPS15E*, At5g43640), and a basic helix-loop-helix protein (*bHLH092*, At5g43650). *TZP* has no close homologs, but does share similarity with other Arabidopsis proteins that only have the zinc knuckle or PLUS3 domains. *bHLH092* is part of the large bHLH family of transcription factors (Toledo-Ortiz, G. et al., *Plant Cell*, 15:1749-1770 (2003)), and a close homolog is TRANSPARENT

TESTA 8 (*TT8*; At4g09820). There are four closely related *RPS15E* genes (At1g04270, At5g09500, At5g09510, and At5g09490) (Barakat, A. et al., *Plant Physiol*, 127:398-415 (2001)). The transcript/expression of all three annotated genes in the Columbia background was confirmed by tiling array data and full length cDNAs. Quantitative RT-PCR (qPCR) revealed that all three genes are expressed in Bay-0 and Shahdara (data not shown).

[0072] T-DNA-insertions in *RPS15E* or *bHLH092* did not result in hypocotyl elongation phenotypes (data not shown). No insertion mutants were available for *TZP* (an insertion GABI-KAT line could not be recovered). To determine the association of the stop codon in *TZP* with the hypocotyl phenotype, we sequenced the PLUS3 domain in other *A. thaliana* accessions. Although we identified 10 synonymous changes and 13 non-synonymous changes (including two changes affecting residues at least partially conserved in other species), we could not detect the Bay-0 premature stop codon polymorphism in a panel of ~300 additional accessions (data not shown). We discovered that a Bay-0 single seed descent (SSD) line (Bay-0[41AV]), which was derived independently of the parent for the RIL set, did not have the 8 bp-insertion in *TZP*. Sequencing the entire 7 kb-candidate region revealed that this was the only sequence difference between Bay-0 and Bay-0[41AV] at *LIGHT5*. Genotyping additional markers in Bay-0[41AV] (as well as a careful phenotypic observation of the lines) showed that it is from the same genetic background as Bay-0 (data not shown). We crossed the two lines, with and without the *TZP* stop codon, to determine whether the mutation was responsible for the *LIGHT5* phenotype. As shown in SI Fig. 5, phenotypes of F2 plants homozygous for either allele were very similar to the phenotypes of the arHIF fixed for either the Bay (stop) or Sha allele, respectively. This demonstrates that the 8 bp-insertion in *TZP* is sufficient to explain the *LIGHT5* phenotype, and allows us to conclude that *TZP* controls hypocotyl growth. Bay-0[41AV] is the only Bay-0 stock we have which does not carry the 8 bp-insertion. The question remains as to when this causative polymorphism appeared in the Bay-0 lineage (before or after collection) and whether it really exists in nature. Unfortunately, indications about the exact collection site of Bay-0 are very poor and make it nearly impossible to locate the original natural population and answer this question.

[0073] Additionally, we employed a transgenic approach to confirm the role of *TZP* in hypocotyl growth. The Sha alleles of all three genes were overexpressed in the rHIF containing the Bay allele (rHIF138-8). Overexpression of *TZP* (*TZP-OX*) caused plants to have very long hypocotyls (Figs. 2B-C), with increased growth throughout development and extended duration of the reproductive phase (Figs. 2D-E). All linearly-elongating organs

were more extended, including petioles, internodes and peduncles and the main floral stem of *TZP-OX* plants was usually twice as long as in its background. In contrast, *RPS15E* or *bHLH092* over-expressing plants were indistinguishable from the parental line in terms of growth (data not shown). Over-expression of *bHLH092* resulted in white seeds, similar to a phenotype found in mutants of its closest homolog TT8 (Nesi, N. et al., *Plant Cell*, 12:1863-1878 (2000)). From these data we conclude that *TZP* is the QTG explaining the *LIGHT5* QTL, and that the 8 bp indel leading to a premature stop codon in the Bay-0 allele is the causative polymorphism.

TZP is a large, nuclear-localized protein encoded by a single copy gene

[0074] *TZP* is a single copy gene in *A. thaliana*. *TZP* orthologs are also single copy in the fully sequenced plant genomes (SI Fig. 6). Interestingly, *TZP* is not found in the moss *Physcomitrella patens*, or in *Chlamydomonas reinhardtii* or any earlier algal lineages queried, suggesting that it is specific to vascular plants.

[0075] The zinc knuckle (znkn; $CX_2CX_4HX_4C$; SEQ ID NO:13) domain has been shown to be important for protein-protein interactions, as well as for binding single-stranded DNA (Vo, L.T. et al., *Mol Cell Biol*, 21:8346-8356 (2001)). There are at least 24 znkn-containing proteins in *Arabidopsis*, five of which are closely related to each other, but not to *TZP* (table S2 from our website data). Mutations in one of them, *SWELLMAP 1* (*SMP1*), result in small plants with small cells because of a dysfunction in the commitment to the cell cycle (Clay, N.K. and Nelson, T., *Plant Cell*, 17:1994-2008 (2005)). Another znkn protein, RSZ33, regulates interactions between splicing factors (Kalyna, M. et al., *Mol Biol Cell*, 14:3565-3577 (2003); Lopato, S. et al., *Plant Mol Biol*, 39:761-773 (1999)). In yeast, the znkn of MPE1 has been found to control 3' end processing of pre-mRNA, and its human ortholog RBBP6 has been shown to interact with tumor suppressor pRB1 (Vo, L.T. et al., *Mol Cell Biol*, 21:8346-8356 (2001)). Znkn domains are also found in retroviral gag proteins (nucleocapsid), including that of HIV (De Guzman, R.N. et al., *Science*, 279:384-388 (1998)).

[0076] The PLUS3 domain is thought to play a role in nucleic acid binding through three conserved positive amino acids (de Jong, R.N. et al., *Structure*, 16:149-159 (2008)). There are five proteins in *Arabidopsis* with the PLUS3 domain (SI Fig. 6 and our website data). Three of these also have a SWIB domain, two have either a CCCH or a C3H3C4 zinc finger, and VERNALIZATION INDEPENDENT 5 (VIP5) contains only the PLUS3 domain (Oh, S.

et al., *Plant Cell*, 16:2940-2953 (2004)). VIP5 is a homolog of the yeast RTF1 protein, which is part of the yeast Paf1 complex that regulates histone H2B ubiquitination, histone H3 methylation, RNA polymerase II carboxy-terminal Ser2 phosphorylation and RNA 3' end processing (de Jong, R.N. et al., *Structure*, 16:149-159 (2008)). The SWIB domain-containing proteins are part of the SWI/SNF chromatin-remodeling complex (Mlynarova, L. et al., *Plant J*, 51:874-885 (2007)). The combination of tandem zinc knuckles and a PLUS3 domain is unique to *TZP*-type genes in plants. Based on GFP fusions, TZP is localized to the nucleus in small punctate structures (Fig. 2F). Considering the domain structure and the localization, it seems most likely that TZP has a role in transcriptional control, perhaps at the level of chromatin remodeling.

TZP regulates light quality-dependent growth

[0077] Our initial QTL study suggested that *LIGHT5* is involved in light-regulated hypocotyl growth. To determine whether the growth defects in *LIGHT5* are specific to certain light environments, we measured hypocotyl lengths for both the rHIF and *TZP-OX* lines under different fluences of monochromatic red, blue and far-red light, and in continuous dark. We found that both loss and gain of *TZP* activity had a significant effect on hypocotyl length under a range of blue or white fluences, but not in red or far-red light, or in the dark (Figs. 2A-B; data not shown). These results are consistent with *TZP* playing a role in light-dependent hypocotyl elongation, and suggest that TZP is involved in blue-light signaling.

[0078] We measured transcript abundance in Shahdara, Bay-0, rHIF138-8, rHIF138-13, arHIF47-2, arHIF47-5, and *TZP-OX* line #3 seedlings grown in constant blue light (15 $\mu mol/m^2s$) for five days and then harvested at subjective dawn (relative to the time when they were moved from stratification to light). The extent of growth paralleled *TZP* expression levels (SI Fig. 7), consistent with *TZP* having a direct effect on growth. To identify potential downstream transcriptional targets of TZP, we analyzed genome-wide expression patterns of five genotypes (rHIF138-8 and rHIF138-13; arHIF47-2 and arHIF47-5; *TZP-OX* line#3) using Affymetrix *Arabidopsis* ATH1 GeneChip arrays. In pair-wise comparisons, we found 135, 12 and 769 genes to be differentially expressed between lines with contrasting alleles, rHIF138-8 vs. rHIF138-13, arHIF47-2 vs. arHIF47-5, and *TZP-OX* vs. rHIF138-8, respectively ($P < 0.01$; tables S3-S6 from our website data; Materials and Methods in SI Text). The top four significant gene ontology categories in the comparison of *TZP-OX* and rHIF138-8 (769 genes) are cytosol, ribosome, structural molecule activity, and cell wall (table S7 from our website data), consistent with TZP playing a

specific role in modulating growth. We found similar results with the rHIF138-8 vs. rHIF138-13 comparison (data not shown).

TZP controls morning-specific growth through an auxin-related pathway

[0079] It is well established that hypocotyl elongation is controlled by the circadian clock (Nozue, K. et al., *Nature*, 448:358-361 (2007);Dowson-Day, M.J. and Millar, A.J., *Plant J*, 17:63-71 (1999)). Therefore, we tested whether *TZP* is clock regulated and if the circadian clock is altered in *TZP-OX* plants and the rHIF carrying the Bay-0 allele. We found that *TZP* cycles under both diurnal and circadian conditions with peak expression at dawn (transition from dark to light; Fig. 3A; SI Fig. 8). Hypocotyl growth is maximal at dawn and many genes that regulate growth, such as cell wall and phytohormone genes, have dawn-specific transcript abundance (Nozue, K. et al., *Nature*, 448:358-361 (2007); Michael, T.P. et al., *PLoS Biology*, (in press) (2008)). The dawn-specific *TZP* transcript peak suggests that it may be part of a circadian-controlled growth mechanism. To confirm that the clock controls *TZP*, we asked if circadian mutants disrupt expression of *TZP*. Indeed, *TZP* expression was changed in both *early flowering3* (*elf3*) and *late elongated hypocotyl* (*lhy*) mutants (SI Fig. 9). However, the circadian clock is not disrupted in either the rHIF or *TZP-OX* under light/dark cycles (SI Fig. 10), consistent with no feedback of *TZP* into the circadian clock. Based on these results we propose that TZP functions to control growth downstream of the circadian clock.

[0080] Since *TZP* is regulated by light/dark cycles and the circadian clock, we asked if the genes disrupted by *TZP-OX* were time-of-day specific. We used the PHASER time-of-day analysis tool (http://phaser.cgrb.oregonstate.edu/), which determines if there is a pattern of time-of-day co-expression in a given gene list compared to a background model. The peak expression of genes that were disrupted in *TZP-OX*, and differentially expressed in rHIF138-8 vs. rHIF138-13 is biased towards dawn (Fig. 3C). This is consistent with the dawn-specific expression of *TZP*. In addition, morning-specific response elements such as the morning element (CCACA), Gbox (CACGTG) and HUD (CACATG) [table S8 from our website data; (Michael, T.P. et al., *PLoS Biology*, (in press) (2008); Michael, T.P. et al., *PLoS Genet*, 4:e14 (2008))] were overrepresented in the promoters (500 bp) of these genes as determined using the ELEMENT motif-searching tool (Mockler, T.C. et al., *Cold Spring Harb Symp Quant Biol*, 72:353-363 (2007)); http://element.cgrb.oregonstate.edu/). These results support a role of *TZP* in the transcriptional activity of dawn-specific genes.

[0081] In an effort to capture the entire effect of *TZP-OX*, we carried out a time course under light/dark cycles (12 hrs white light/12 hrs dark) in seven-day-old seedlings, sampling every four hours over one day in the same genotypes as above. We validated the overexpression of *TZP* using qPCR, which revealed that *TZP* continued to cycle despite overexpression (Fig. 3B), suggesting that *TZP* is also controlled post-transcriptionally. Global expression changes were assessed using Affymetrix ATH1 GeneChip arrays in *TZP-OX*, rHIF138-8 and rHIF138-13. The resulting time courses were analyzed for differentially expressed genes and for cycling genes [(Michael, T.P. et al., *PLoS Genet*, 4:e14 (2008)); tables S9-S12 from our website data; Materials and Methods in SI text]. Using the time points as replicates, we identified 117 upregulated and 40 downregulated genes in *TZP-OX* ($P < 0.01$).

[0082] Similar to the results in blue light, the genes that were upregulated in *TZP-OX* were dawn-specific (SI Fig. 11) and cell wall genes were overrepresented (table S13 from our website data). *IRX1* is one example of the eight cell wall genes that were upregulated in *TZP-OX* (Fig. 3D). Like *IRX1*, many of the upregulated genes were specifically overexpressed at dawn, similar to the overexpression pattern of *TZP* itself (Fig. 3A; SI Fig. 12). Peroxidases, which can function to polymerize cell wall compounds, were also upregulated in *TZP-OX*, consistent with their role in growth and cell wall expansion (Passardi, F. et al., *Planta*, 223:965-974 (2006)). Peroxidases *PER27*, *PER30*, and *PER64* have been shown to be part of the cell wall proteome (Bayer, E.M. et al., *Proteomics*, 6:301-311 (2006)) (6 PER genes total; SI Fig. 12A). However, *CATALASE 3* (*CAT3*; *SEN2*) was one of the most downregulated genes (SI Fig. 12D). *CAT3* cycles under all diurnal and circadian conditions and may play a role in senescence and stress responses (Zimmermann, P. et al., *Plant Cell Environ*, 29:1049-1060 (2006); Michael, T.P. and McClung, C.R., *Plant Physiol*, 130:627-638 (2002)). Two additional genes of note were downregulated in *TZP-OX*: *PW9* (encoding a MATH/TRAF domain protein) and one of the PLUS3 homologs that encodes also a SWI/SNF domain (SI Figs. 12E-F). In general, the genes that were misregulated in *TZP-OX* are involved in cell expansion, consistent with TZP being intimately related to regulation of hypocotyl elongation and downstream of the circadian clock.

[0083] Auxin-response genes, including *WES1* [GH3.5; (Park, J.E. et al., *Plant Cell Physiol*, 48:1236-1241 (2007))], *DFL1* [GH3.6; (Nakazawa, M. et al., *Plant J*, 25:213-221 (2001))] (SI Fig. 12B), and *AXR5* [IAA1; (Yang, X. et al., *Plant J*, 40:772-782 (2004))], all of which have been shown to control hypocotyl growth, were also upregulated upon *TZP*

overexpression. It has been shown recently that auxin controls growth in a time-of-day fashion (Covington, M.F. and Harmer, S.L., *PLoS Biol*, 5:e222 (2007)), which fits with TZP controlling morning-specific growth through these genes. In addition, two homeobox-leucine zipper genes induced in *TZP-OX* (*HAT4* and *HAT52*) are upregulated under low light conditions and in response to auxin (SI Fig. 12C). Mutations in these genes lead to growth defects (Sawa, S. et al., *Plant J*, 32:1011-1022 (2002); Sessa, G. et al., *Genes Dev*, 19:2811-2815 (2005)). Together, these results support the notion that TZP plays a broad role in the regulation of phytohormone-dependent gene transcription (Michael, T.P. et al., *PLoS Biology*, (in press) (2008)).

[0084] A similar number of genes were found to cycle across all three genotypes as reported for this condition previously [(Michael, T.P. et al., *PLoS Genet*, 4:e14 (2008)); ~7,700 genes], and there was no significant difference between the number of genes cycling in rHIF lines and *TZP-OX*, although there were genes specific to each genotype. However, the peak transcript abundance of 362 genes was shifted by six hours or more in *TZP-OX* compared to rHIF138-8 (only genes that cycled in both genotypes were considered). Phytohormone-related (*SIR1, ETO1, GH3.3, RGA1, ABF4*), homeobox-leucine zipper (*HAT2, HAT3*), chromatin remodeling (*HD2B, HDT4, SUVH9, CHR4, TAF1*), leaf polarity (*KAN3, AS1*) and ribosomal genes were misphased in *TZP-OX*. *PHYTOCHROME D* (*PHYD*) was also phased eight hours earlier in *TZP-OX* (SI Fig. 12L). The first *phyD* mutant was originally identified as a natural allele, and subsequently was shown to affect shade-avoidance associated growth and flowering (Aukerman, M.J. et al., *Plant Cell*, 9:1317-1326 (1997); Devlin, P.F. et al., *Plant Physiol*, 119:909-915 (1999)).

[0085] Consistent with TZP controlling blue light dependent growth, *LONG HYPOCOTYL IN FARRED 1* (*HFR1*) is overexpressed at dawn more than 100 fold with the same pattern as *IRX1* (SI Fig. 12I). *HFR1* encodes a bHLH transcription factor that is required for both phytochrome A-mediated far-red and cryptochrome 1-mediated blue light signaling (Fairchild, C.D. et al., *Genes Dev*, 14:2377-2391 (2000)). *HFR1* expression is high under low light conditions such as shade and continuous dark conditions (Sessa, G. et al., *Genes Dev*, 19:2811-2815 (2005)), and is elevated in circadian and light signaling mutants, much like in *TZP-OX* [SI Fig. 12I; (Michael, T.P. et al., *PLoS Biology*, (in press) (2008))]. Recently it has been shown that two other bHLH transcription factor genes related to *HFR1* —*PIF4 and PIF5*— control morning-specific hypocotyl growth. Disruption of the circadian clock gene *CCA1* results in the overexpression of *PIF4* and *PIF5* leading to uncontrolled

elongation (Nozue, K. et al., *Nature*, 448:358-361 (2007)). However, *PIF4* and *PIF5* are expressed at control or slightly lower levels in *TZP-OX* (SI Figs. 12G-H). These results support *TZP* acting in parallel with (or downstream of) *PIF4* and *PIF5* in growth control.

[0086] Conclusions: Despite extensive forward genetics screens in *A. thaliana*, natural variation has recently made important contributions to the identification of genes not previously known to impact several different traits [e.g., (Bentsink, L. et al., *Proc Natl Acad Sci USA*, 103:17042-17047 (2006); Mouchel, C.F. et al., *Genes Dev*, 18:700-714 (2004); Macquet, A. et al., *Plant Cell*, 19:3990-4006 (2007); Baxter, I. et al., *PLoS Genetics*, 4:e100000438-41 (2008))]. Apart from being able to exploit allelic variation (in multiple genetic backgrounds) that cannot be generated by conventional mutagenesis, the success of these studies has often been due to the use of quantitative phenotyping, as opposed to the qualitative gauges employed in typical mutant screens. We have demonstrated here the power of QTL analysis to reveal a new component of the hypocotyl growth pathway in *A. thaliana*, TZP, a unique, tandem zinc knuckle/PLUS3 domain protein encoded by a single copy gene in the vascular plant lineage. TZP provides a direct link between light signaling and the pathways that control growth in an environmentally independent fashion.

Materials and Methods

[0087] A detailed and referenced version of this section is available online (SI Text).

[0088] Plant material and phenotyping. The Core-Population of 164 RILs from the Bay-0 x Shahdara set (http://dbsgap.versailles.inra.fr/vnat/) was phenotyped in four different light and temperature environments to map QTL affecting hypocotyl elongation. Complete phenotypic data from RILs is available in table S14 from our website data at www.inra.fr/vast/Files/Loudet_PNAS_SITables.xls. HIF350 was developed from an F7 line (RIL350) that still segregated for a single and limited genomic region around *LIGHT5* locus. Plants still heterozygous for the QTL region were screened with adequate markers to isolate recombinants (rHIF) used in the fine-mapping process. Advanced rHIF crosses were generated from two different rHIFs recombined immediately to the north or immediately to the south of the *LIGHT5* interval giving rise to lines arHIF47. Distinct Bay-0 lines from the stock center were used to find variants at *LIGHT5*. rHIF138-8[Bay] was complemented by over-expressing each of the three positional candidate genes cloned from rHIF138-13[Sha].

[0089] QTL mapping. Analyses used hypocotyl length mean values of an average of 16 seedlings (from two distinct experiments) per genotype per environment. QTL analyses were

performed using QTL Cartographer, with classical parameters for interval mapping and composite interval mapping.

[0090] Microarray analysis. Microarray experiments were carried out per Affymetrix protocols (ATH1 GeneChip), on seven-day-old tissue harvested under either continuous blue at subjective dawn, or every four hours (starting at dawn) under 12 hrs white light / 12 hrs dark cycles over one day (six time points). Hybridization intensities from all microarrays were normalized together using gcRMA implemented in the R statistical package. The blue dataset was then separated and differentially expressed genes were identified using linear modeling with the limma bioconductor package in R.

[0091] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

SEQUENCE LISTING

SEQ ID NO:1 - Arabidopsis LIGHT5 coding sequence

ATGGGAGATGGAGATGAGCAAAGTAAAGAGTTAGGAGGAGTTTCGTCTTCGAGTCGTCGT
TGTTCTTCCGGTACTGCTGGTGCTGCAAATGCAGAAGCAAGAATGAAGTTTGCTGCTGTT
GATGCTATAACAGAATTAGTGTGGTCTCCTAGCAACGGTCTAAGCTTAAGATGTGCAGAT
ATCAGTTTCACGGGAAAGGCAAAATTACTCTCTCCTAACTTCTTTGATATAGGACTAACT
AACATGGCGATTCACTCAAATAGCACAAGTATTGAAGACCAAGAAGATCACGTAGATGTC
GAATTGAGAAACCGAGATCAAGTGAACCAAGCAATGATTGGTGGTTCTGTTGAGGATATG
AAACCTGAAATGGTTGAGGATAAGGTTGAAACCAATGATGATATAAAGAATGAAGAAGCA
GGATGTTCCAAAAGATCTTCAGATAGTCCGAAAGCTATGGAAGGAGAAACCAGAGACTTG
TTGGTTAATGAACAGTTGAGGATGGAAAGTGCTGGATCTCAAGAAGAAGGGGATAAGGCG
CATAATAGAGTTGATCGTTTAGAATCAATGGATGAGAACAACCTCGCAACTCTTGCTGTT
GTAGCATGTGAAGGTAAAGGGGACTATTTGCCGGAGGGTGAGGCTGGACCTAGTGGTTCC
TACAGGCGTCGTGAGAAAGCAAAAGGGAAAGAGAAAGCTTTATCTGATGAAAACTTTGGT
GGTGATGGTGAAGATGAGGATGAGGAGAGTTTTGGTAGTGTAGAAAGTTGCAATAGTGCG
GGTTTGCTCTCAAGGGGAAAGAAAAGACCAGGCTTTGAGGAGCAGCTGATTTTTGGAAGC
AAACGGCTCAAAACGTTAAATCAAGAATGTTTGGGGTCAACTTCAAAGCTCAAGCAAGAT
AGTTCCTTCATGAATTGGATATCAAATATGACTAAAGGAATCTGGAAAGGTAATGAAGAA
GACAATTCTCCTTTTGTAGCTCTTACCACCACCTCAAATGCTAATGGTCATGGCCAAGTC
AATGCCATCGTTGACCAACAACAATTGAGTCCATGCTGTGTGAAGGAAAATAGTGGATGC
AGAAACACCGGTTTCCAGTCTTTTTTCCAGTCTATATATTGTCCGAAGAAACAAAGTCAA
GATGTCGTCGACATGGATTTCCCGAATGATGTTAATGCTGCTCCTTTGCAGGAACTTCCC
TGGATTCCTGAACATTGCGACATTTCCAAAGGGGATGACTTGTCTTCATCTGGTAATGAG
ATTGGTCCAGTGGCTGAACCCAACATTTCATCAGGAAAAGTTGTATTTAATCAGACAAGC
AAAACACAGTCTTCAGAGAACAAACGTGAGGACAAAGAGCCAAACATCTCGTTGATGTCT
CTTAGTAAGTCTAAGCCAAATGAAGAGCCAAAAACATGCGGTGAAGCTGATGGAAAGGTT
AGTCCATGTTTAACCAACAGAAACTCTGGTCTTAAAAGTTTGTGGATTAGTCGGTTTTCT
TCTAAAGGTTCGTTCCCTCAGAAAAAGGCAAGTGAAACGGCTAAAGAGGCCAATGCTTCA
GCCTCAGACGCAGCTAAAACCCGTGATTCACGGAAAATGCTGGCAGACAAGAATGTTATA
AGACCCAGTATCAGCTCGGTAGATGGGCCCGATAAACCGGACACTGTCCTTCCTATAGTT
TCGTCAATGAGAATAGAGTCTTCAGAAGCAATGGCTTCTTTGTTTGCTAGAAGATTAGAA
GCAATGAAAAGCATCATGCCCTCAGGTTCTTTAGCTGAGAATGCAGAGGAGGAACAGAGG
GATCTAATTTGTTTCTACTGCGGTAAAAAGGGTCATTGTTTACGGGATTGTTTGGAAGTA
ACTGATACTGAGCTCAGAGATCTAGTACAGAACATAAGTGTGCGGAATGGAAGAGAAGAA
GCGTCAAGTCTATGCATTAGATGTTTTCAGCTTAGTCACTGGGCTGCAACATGCCCGAAT
GCTCCCCTGTATGGTTCAGGAGCTGAAGGTAGAGCTATGAAGAACGCTTTAGCTTCTACC
TCTGGCATGAAGCTGCCCATAAGTGGTTTCACAGACGTACCAAGGGCGGTCTTTGATGCT
GTTCAAGTGCTTCGCCTGAGCCGCACAGACGTTCTCAAATGGATAAACACCAAGAAGTCT
GTATCGGGTCTCGAAGGGTTCTTCTTGCGGTTAAGGCTCGGGAAATGGGAAGAAGGGCTT
GGAGGAACAGGGTATTATGTAGCTCGCATAGATGGGGACACAGAGGGACAAAGCTCAAGG
AGACATTCAGAGAAAAGCTTAATTTCAGTTAAGGTTAAAGGTGTGACTTGCCTTGTTGAG
TCTCAGTTTATCTCAAACCAAGACTTTCTTGAGGAGGAGTTAAAGGCCTGGTGGCAGAGC
GCGGGGAAAAGTGCCAGAACAAGCGGCTACGACGGCATCCCGTCGGCGGAGGAACTTAGT
CGGAAAATTCAGCAGAGAAAAATGTTAGGCTTTTAG

SEQ ID NO:2 - Arabidopsis LIGHT5 amino acid sequence

MGDGDEQSKELGGVSSSSRRCSSGTAGAANAEARMKFAAVDAITELVWSPSNGLSLRCAD
ISFTGKAKLLSPNFFDIGLTNMAIHSNSTSIEDQEDHVDVELRNRDQVNQAMIGGSVEDM
KPEMVEDKVETNDDIKNEEAGCSKRSSDSPKAMEGETRDLLVNEQLRMESAGSQEEGDKA
HNRVDRLESMDENNLATLAVVACEGKGDYLPEGEAGPSGSYRRREKAKGKEKALSDENFG
GDGEDEDEESFGSVESCNSAGLLSRGKKRPGFEEQLIFGSKRLKTLNQECLGSTSKLKQD
SSFMNWISNMTKGIWKGNEEDNSPFVALTTTSNANGHGQVNAIVDQQQLSPCCVKENSGC
RNTGFQSFFQSIYCPKKQSQDVVDMDFPNDVNAAPLQELPWIPEHCDISKGDDLSSSGNE
IGPVAEPNISSGKVVFNQTSKTQSSENKREDKEPNISLMSLSKSKPNEEPKTCGEADGKV
SPCLTNRNSGLKSLWISRFSSKGSFPQKKASETAKEANASASDAAKTRDSRKMLADKNVI
RPSISSVDGPDKPDTVLPIVSSMRIESSEAMASLFARRLEAMKSIMPSGSLAENAEEEQR
DLICFYCGKKGHCLRDCLEVTDTELRDLVQNISVRNGREEASSLCIRCFQLSHWAATCPN
APLYGSGAEGRAMKNALASTSGMKLPISGFTDVPRAVFDAVQVLRLSRTDVLKWINTKKS
VSGLEGFFLRLRLGKWEEGLGGTGYYVARIDGDTEGQSSRRHSEKSLISVKVKGVTCLVE
SQFISNQDFLEEELKAWWQSAGKSARTSGYDGIPSAEELSRKIQQRKMLGF*

SEQ ID NO:3 Arabidopsis LIGHT5 genomic sequence

GGCGCTGTGTCGTCTTCGTCTTCTTCTTCCTTCTTCTTCTTCATCTCTCTCTTTCGTTTT
CTCAGTAAGTGAAGGATTTGCAGAAGTAGATGATTGAATATCCATGGAGGTAATTTATGG
TTTCTAATTTGGTTTTAATTCTTTGAAAGTTTGATTTTTTTATGTTATGAGATTTCTGGG
TTTTTGATTCTCAAGTTTTTATGGATTTTAGGGTTTTCAGTAATTGCAAAACTTTATTCT
GAAGATTGAGTGATTGCTTGAGATATAGGTTATCAATTTCCTTTATCCCTAATATTCCAA
ATCTCGAAATTTCAAATTACTTCTCTAAGTTGTTCTATGAGTTTATGATCTTAGAAAATA
AATCTAGTGTTTTGATTACTCTGTTTTGCATATTTTTGATGGGTTTGTGTTACTTGTTTT
GTAAGCTAAGTGATATTCATTTTTGTGGTTGAGGTCAATAGTGTATGGATTCTGGTTCTT
TATGTCTCAGGGTGATTGGTTTCAGGAGGTATGGGAGATGGAGATGAGCAAAGTAAAGAG
TTAGGAGGAGTTTCGTCTTCGAGTCGTCGTTGTTCTTCCGGTACTGCTGGTGCTGCAAAT
GCAGAAGCAAGAATGAAGTTTGCTGCTGTTGATGCTATAACAGAATTAGTGTGGTCTCCT
AGCAACGGTCTAAGCTTAAGATGTGCAGATATCAGTTTCACGGGAAAGGCAAAATTACTC
TCTCCTAACTTCTTTGATATAGGACTAACTAACATGGCGATTCACTCAAATAGCACAAGT
ATTGAAGACCAAGAAGATCACGTAGATGTCGAATTGAGAAACCGAGATCAAGTGAACCAA
GCAAGTAATGAAATGAATAGACCATTTTCAGGTTGCTATATGATAAGTTATCTTTTTTGA
TAATGAGAGTTCTTGTGTGTTTGCAGTGATTGGTGGTTCTGTTGAGGATATGAAACCTGA
AATGGTTGAGGATAAGGTTGAAACCAATGATGATATAAAGAATGAAGAAGCAGGATGTTC
CAAAAGATCTTCAGGTAAAATGGTTATATCAAGTTAATAATGTTTGCGTCTCTTCCATTT
ATGACAACTATGGATTTTGAGTTGGTTTCTTCGTCTTTTGTGCTTGTTTTGGTTATCGAG
TAGATAGTCCGAAAGCTATGGAAGGAGAAACCAGAGACTTGTTGGTTAATGAACAGTTGA
GGATGGAAAGTGCTGGATCTCAAGAAGAAGGGGATAAGGCGCATAATAGAGTTGATCGTT
TAGAATCAATGGATGAGAACAACCTCGCAACTCTTGCTGTTGTAGCATGTGAAGGTAAAG
GGGACTATTTGCCGGAGGGTGAGGCTGGACCTAGTGGTTCCTACAGGCGTCGTGAGAAAG
CAAAAGGGAAAGAGAAAGCTTTATCTGATGAAAACTTTGGTGGTGATGGTGAAGATGAGG
ATGAGGAGAGTTTTGGTAGTGTAGAAAGTTGCAATAGTGCGGGTTTGCTCTCAAGGGGAA
AGAAAAGACCAGGCTTTGAGGAGCAGCTGATTTTTGGAAGCAAACGGCTCAAAACGTTAA
ATCAAGAATGTTTGGGGTCAACTTCAAAGCTCAAGCAAGATAGTTCCTTCATGAATTGGA
TATCAAATATGACTAAAGGAATCTGGAAAGGTAATGAAGAAGACAATTCTCCTTTTGTAG
CTCTTACCACCACCTCAAATGCTAATGGTCATGGCCAAGTCAATGCCATCGTTGACCAAC
AACAATTGAGTCCATGCTGTGTGAAGGAAAATAGTGGATGCAGAAACACCGGTTTCCAGT
CTTTTTTCCAGTCTATATATTGTCCGAAGAAACAAAGTCAAGATGTCGTCGACATGGATT
TCCCGAATGATGTTAATGCTGCTCCTTTGCAGGAACTTCCCTGGATTCCTGAACATTGCG
ACATTTCCAAAGGGGATGACTTGTCTTCATCTGGTAATGAGATTGGTCCAGTGGCTGAAC
CCAACATTTCATCAGGAAAAGTTGTATTTAATCAGACAAGCAAAACACAGTCTTCAGAGA
ACAAACGTGAGGACAAAGAGCCAAACATCTCGTTGATGTCTCTTAGTAAGTCTAAGCCAA
ATGAAGAGCCAAAAACATGCGGTGAAGCTGATGGAAAGGTTAGTCCATGTTTAACCAACA
GAAACTCTGGTCTTAAAAGTTTGTGGATTAGTCGGTTTTCTTCTAAAGGTTCGTTCCCTC
AGAAAAAGGCAAGTGAAACGGCTAAAGAGGCCAATGCTTCAGCCTCAGACGCAGCTAAAA
CCCGTGATTCACGGAAAATGCTGGCAGACAAGAATGTTATAAGACCCAGTATCAGCTCGG
TAGATGGGCCCGATAAACCGGACACTGTCCTTCCTATAGTTTCGTCAATGAGAATAGAGT
CTTCAGAAGCAATGGCTTCTTTGTTTGCTAGAAGATTAGAAGCAATGAAAAGCATCATGC
CCTCAGGTTCTTTAGCTGAGAATGCAGAGGAGGAACAGAGGGATCTAATTTGTTTCTACT
GCGGTAAAAAGGGTCATTGTTTACGGGATTGTTTGGAAGTAACTGATACTGAGCTCAGAG
ATCTAGTACAGAACATAAGTGTGCGGAATGGAAGAGAAGAAGCGTCAAGTCTATGCATTA
GATGTTTTCAGCTTAGTCACTGGGCTGCAACATGCCCGAATGCTCCCCTGTATGGTTCAG
GAGCTGAAGGTAGAGCTATGAAGAACGCTTTAGCTTCTACCTCTGGCATGAAGCTGCCCA
TAAGTGGTTTCACAGACGTACCAAGGGCGGTCTTTGATGCTGTTCAAGTGCTTCGCCTGA
GCCGCACAGACGTTCTCAAGTAAGATCTAATTCACACCTAAAAGACTGTTTCCCTTGTCA
CATTATGAGGCTTTAACATCTTTTGACATTTCATTTCAGATGGATAAACACCAAGAAGTC
TGTATCGGGTCTCGAAGGGTTCTTCTTGCGGTTAAGGCTCGGGAAATGGGAAGAAGGGCT
TGGAGGAACAGGGTATTATGTAGCTCGCATAGATGGTTGGTACACTTTCATATAAATCAA
TACTCAAACTCGCGTATTTCTCGAATATTTGCAACTGAAATATGTTTGTTTTAACCAGGG
GACACAGAGGGACAAAGCTCAAGGAGACATTCAGAGAAAAGCTTAATTTCAGTTAAGGTT
AAAGGTGTGACTTGCCTTGTTGAGTCTCAGTTTATCTCAAACCAAGACTTTCTTGAGGTA
TATAATAAAATTTGACCAATTATGTCAAAAGTTTTGAAATTTGCTAATTTGCTAAGTTTT
CATCTGAATATTTTGCATTAATTTATCTGAGCAGGAGGAGTTAAAGGCCTGGTGGCAGAG
CGCGGGGAAAAGTGCCAGAACAAGCGGCTACGACGGCATCCCGTCGGCGGAGGAACTTAG

TCGGAAAATTCAGCAGAGAAAAATGTTAGGCTTTTAGGACATTTGTTCAAGTAGATAAAT
GTTTTTAAGTTATAGGTTTTACGATTACCCGTGTATTTTTTCGGTTGAAAAGGAGAGTGT
AAATTTCTTAGTTCAGAGGGCATCGAATGTCTCAAACCATACAACAATCTAAAACTGGAT
GTCCAATCTATTTAGATTTGCGAGTTTTATTCGGATCTTTATGTCTCGATGACTATTTAT
TGTGATTGTTAAGCTTGTATAGTCTATGTGTGAAGGATCAAAGTATCATGGAAGAACAAA
GTACATTGATATCAAGAGAGATATTGTGGTGTCTTGCTCATTTTTC

SEQ ID NO:4
>GmTZPa_soybean
MMLPCDSRIHMAINKGKEKSLSDGDANVILSREENDSHSSVESCNSAGFFSTGKKRRNFQQQLIIGSKRVKKQIE
ESSGFKSYVKQDSSFMNWISNMVKGLQQSIQNDSNTLALTLTNPDHHNLLPDEKLFTCNMNQDPEPKNTGFKSFF
QSIYCPSLKNGGTRMSHQEGKSSDDLEPGNMEHGIDATPITYCAENNSLSKLRLQSNKFEVSIGGNDAGPSSQPK
VKPLNFFNCQESSKNNPVETKNYSILGHSKDKEEVASHSSSTKQNTDDNDNIDSNALPDRKEEENICHRRDNLGS
LWITRFSPKFTAPLREQPANDTEASTDLKEDKGNNDHKSMYMFKPLSSSPGLRNLEPMASMFARRFSAIKHIIPT
NATDTTTQVNMLCLFCGTKGHQLSDCSAIAENKLEDLQKNIDSYGGLEEHSCLCIKCFQPNHWAISCPTSISTRK
HELKANALVNDCGKQKHLIPSNEESARLLTDEDDRVLSGGSINDETDQRTGQNINLKLKSNEIITHKVGCNASFQ
KYCGSSLEENKFRENPISSPSKLTERQISHVPKKIFDAVKKLQLSRTDILKCINTHGSISQLDGFFLRLRLGKWE
EGLGGTGYHVAYINETQSQRQCPEQNTRKCLSVKVGSIKCMVESQYISNHDFLEEEITEWWSNTSEAGAEIPSEE
YLIEKFKKKEMLGL*

SEQ ID NO:5
>GmTZPb_soybean
MLPCDKILPVLHSPCHSRIHMAINKGKEKSLSDGDANVMLSREENDSHSSVESCNSTGFFSTGKKRRNFQQQLII
GSKRVKKQIEESSGPKSYVKQDNSFMNWISNMVKGLSPSIQNDSNTLALTLANPDHHNLQPDEKLIACNMNQDPE
PKNTGFKSIFQSICCPSLKNVGTRMSHQEGKSSQDLVPGNMEHGIDATPITCWAENNSLSKLCLQSNKFEVSTGG
NDAGLSSQPKIKPLNFFNCHESSKNNPVETKNYSILGHSKDKEEVASHSSSTKQNTDNNDNIDSNVLCDRKEEEN
ICHRRDNLGSLWITRFSPKFTAPLREQPANDTEVSTDLKEDKGNNDHKSMYLFKPLSSSPGFRNLEPTASMFGRR
FGAIKQIIPTNATDTTTQVNMLCFFCGTRGHQLSDCLAIAENKLEDLQKNIDSYGGLEEHHCLCIKCFQPNHWAI
SCPTSISTRKHELKANALVNDCGKHLISSNEGSARLLTDEDDRVLSGGSINDETDQRAGQNINLKWKSNEIITHK
VGCNASFKKYRGLSSEENKFRENPTLSPSKLAERQISQVPKEIFEAVKKLQLSRTDILKWINTRGSISQLDGFFL
RLRLGKWEEGPGGTGYHVAYINETQSQRQCSEQNTRKSLSVKVGSIKCMVESQYISNHDFLEEEIMEWWSNTSQA
GAEISSEEYIIEKFKKKEMLG

SEQ ID NO:6
>OSTZP_rice
MDVMNEGKETSDNFCVDKLEKEDEVGSCPTRYCNDTSHSLGSASRKEVMPIIAEKQAFCATTVHDERSWAANAWR
ARLVKAISQKDSVLPKNADNIHSTSAFGSIGNTENMPGKLTSMLGNRNDSSQDQAMQEKHKDGLIVARCESVSAV
NPVARCELASGVNPLARHESTSGCNPRKLEKGKEKLIYDMSNCVSNTNEGDDSNESIESCPSTKAPKRKHGQFSA
AEMTSGNKRCRREDNESSCSGLFHKNDSSFFNWMSTLTNGVKVFDETTAVPLNQKFSAATGEEFPTNPVPLQNNC
GVPLQSVGFNSLFQSLYSQNVMITSRNTCHQSESSYTANRLTLGFKSSKPVSMGRETLNVATETLAAGRIQMDSY
GDRGAFQNQMGIFPLRAERNQNGFHGSSSNAASGHKDDFSESLWVSRLLPKTPMKVMDTTRCDEETDFCSANPKG
LGDSSSPQDFNVEKELNNSQYFTSKGSDNETTSSKCAAPQDENKPSETMASIFAKRLDALRHANTSAVHVAITCD
HGTPKGRNHKTSSFVVSYNSHDEQESGQKTHKSSGGEGRIVLWTGDKGKEQLSPGNDKELGEKVLSKHENQNCEG
SSDGKVVPPKCNLETNTYIEEIDRKRLQNKEGAPNSMENQPDNKQMVPYGIVPNDVYDEASVVFGALQRLRLSRS
DIIRWMRSPVMHTTLDGFFLRLRFGKWEEALGGTGYHVARINGVLDKNRLSVTIRNSTCQVDSRFVSNHDFHQDE
LKAWWSAAMKSGWKLPSQEELNTKLRERELLRF

SEQ ID NO:7
>PtTZP_poplar
MDTNDKNIEPVIDLGFSLGYSNQCIQRRLKNDSGAGANAASSVDMTFVATNALSELVWSPKKGLSLKCAD
GTFSNQKPSLLRGAGPSDMVSGSNADKAIGKKVFMTPPEESDVRSEVAGRDNPTKFVTSDTGLFPLLSES
MHKVKIGNYEFLAATDDHKEEMKTAVGLPFLQKMEDARNNKAEDIYDPINLQVDEISRTWETKFPSLSDE
TKLDVAQNGPTSKEPNVRIGGVGDASHTLQTEIVSASQVCSVEECESYDTNMQKAPLGREHFESPSCMEK
ERENNMGTGPYICPLEKLESTAENDFKTPHSENVCAVATEIVGSQNAKEVRSSSQQDDEILPKDNDCAIK
QSPTYSRTRRYQMKGKAKALSDGNLNERMLDMDDDSHESVESCNSVGLFSTGKRQRNFDPHSYVGSKSIK
TKIQESPGSSSFVKHDGSFMNWISNMMKGFLKSNEDEAPSLALTLANHKHGHEDRDKNLISCNRNQDQGC
KTMGFHSLFQSLYCPKTKAQETVALNANTQTEGSKELGLDNKICDSNATPIPCRMVTDNVYKRFLQPNEK
LNESTSGNGTASPALTKLLSTNIASGQEISGSNSAEKKNSCNMATDKEKNGTSSNSSRGKRKMNDAEQPS
EGKATNTSGYRSDPLTSLWITRLSPKTSGPLSNRDLCHRRTGEALDGFTDFIRLKAQWQNHPSSYQDKNI
VGAREEEHFTEDPVCMHNCANSTEVSFSINKVNGHHDEKSMCKMNSTLPFSRFRNSEAMASVFARRLDAL
MHIMPSYGTDDSSHGNLTCFFCGIKCHHVRDCPEIIDSELADILRNANSFNGANEFPCVCIRCFQSNHWA
VACPSASSRTRHQAEYGASLVHESSPCKILLNPRNEDDAKQSDGKDSQLQAADAPTVRNGKLHEASASGK
INMNMKPFERDTASSSGEKKLKENQVMPLSNFINSQIADVPKGIFDAVKRLRLSRTIILK

SEQ ID NO:8
>CpTZP_papaya
MNVDKDNLEPHTDLGLTVNYSNHCIQRRLNNHLGAGANAGSSRGMTFVTTDPLSELVWSP
QKGPSLKCADCSFSDKKSSLLWDAGPSNVVFSSAPPMISAVRCSNDKLMNENNNMIKSHM

GFYVKSDVGDRHTSARCPTNNAAIKPVSGPSPEQKTGTGGHMEETSTILQLSVQHANKND
DLSRPKGEVTCDLNTIQADEAAEAIEDNLKNFLDVLRPNVAQTEPLFENPAGGYCDANSE
NQTSKIEINLMSDIHAKNKTEACDDALKSQTVPPKRREEPASFTEGENDRKIKSTDSSNI
RCMEKLESTAENDLQIHIGENGCDAASKSVATEFAHEGKRCSQQEKARFREKLASGKYSA
KNSGIQKYERKGKEKALSDGNLSMMSKEADDSYESVESCNGTGLFSTGKRRWNFEQQMIA
GSKRAKMQIRETLESTSFVKQDSSFVNWIGNMMKGFLKLKEDEAPSVSLTLAYPTNGGPD
HDTIATTRNKNPGCRTVGFQSIFQSIYCPQRRIQDVTTSNDKCHKEVELPNDIFNPNPVN
CHGNFTKQFTISDERNTESTFGNGVLLETQPKISSVNFAASQRTRIPTLQRQMVVQFAQY
KGKEGTSSSSSLGKYKLNIAENIDSEPLSEGESGQNFGNGSDHLGSLWIARLAPKTSGFS
LNLNQQNRSTDVDLEYSADCAKLPHSPQYHVSSPSEAKIVEARQHSLDDQKIAIGNDLPK
CAGETECKITCNSAKDHNDHKSTYNMNPILSSPIENSEAMASLFARRLDTLRHIMPSDLV
ENTASATVSCFFCGRKGHHLRDCPEMTDEDLEDLLRTVNKYNGTEEFSSFCVRCFRLNHW
AVTCPNASSKGNRQRECGASLAGEFSPSSGKHNSRIEENPKFLNGNDSQFQVAEVHSLLD
RNDSGIRASLKLNSTEKLVASSSGKNKLRGNEIVPLSKYITRESSDVPKGIFDAVRMLRL
SRTDILKWMNSQMAGSHLDGCFLRLRVGKWEKGLGGTGYYVACITEAQRQSSPQSSKNSI
FVVVRGIRCLVESQYISNHDFLEDELNAWWFTIEKNGGTIPMEEVLRLKDRHKYNLYIGV
VCYVDKKF

SEQ ID NO:9
>RcTZP_castorbean
MNVDDKNIEPSTNLSLALGYSNQCIQRNLSNDPGAGANAASTADITFVATDPLSELVWSP
HKGLSLRCADGSFIDKKPSLLPGVGPTYMASGSSSDKPISNTGKLFDNEICIASLPACKL
ASEISGDNSTTFLTSNVGIMPLSGTGLDKTATGDQVVEMKNAVNYFLQKEDLRNDKAEDE
TKLDVAQNYRTFEEPIVRATDVNDDHELGMEIVLVSDFHTVKGREDYGIKIQNAACSGKE
NEEPPSVREKERKNKMVIGRPGIFSLDKLESTAENDLETPFGENSCSMRNKNLASESADR
VENNTQHELIPIEYALGYNQSPTSSRLQNIQRQGQSKALSDGDAKERMLNEEDGSHESVE
SCNSTELFSTGKQRWNFDQQLIVGSKRVKRQIQDSPGSSSLGKQDSSFVNWISNMMKGFL
KSSEGEAPFLSSALSNPNYGHENPSQDVFTCNRKEDPACDTRGFQSVFQSLYCRKTKGQE
TVTLNVNHQTEGSKECDQDNKICDLNAAPIACRMVTGNVYKRFLPSNEKHNEPTSGYHAG
MTVHSRDISMSFPVIPESNGSVSTENKNSCNLAIGKEKDGTDSNFSHGKHKTSSAGKIDP
ELPSEDKTAHGFGYKGDPLGSLWIARFSPKTSGAPFNHYPSNKSTGEAFNCSADSMGLIP
QVQNPLGSSSEHEIVEVRNKNFQEPLPIQNYSTANRAPFDFYNVKGNIDNDSGNKLNPIL
SSARVKTSEAMASVSPRRLDAPKYITPSDDADNSDRASMTCFFCGIKGHDLRECSEVTDT
ELEDLLRNINIYGGIKELPCVCIRCFQLNHWAVACPSTCPRVRSKAECHASSVSHAGPSK
SQLHVINEDDTKAKNVTGSGHAICYGNDYGMDKDMNSWKSNEAATSGKMKLNIRLFEKNI
SSTSREKELKENQIIPLYGFVNGLISDVPNGIFDAVRSLRLTRTNILKWMNSSASLSIDG
YFVRLRLGKWEEGLGGTGYYVARITGMKSKKSIAVNVGGIQCVIESQFVSNHDFLEDELK
AWWSATSKVGGKLPSEKELRLKVEEKNXXGGGGGNKIVIENHIVQYQ

SEQ ID NO:10
>BdTZP_brachypodium
MCPLSELVWSPDDGLSIKIAASSLSTRKASLRWNADTLNIVISSPQQSSGRGKSGDNIDATIADAGEMPSQPRTR
CDSSVRLSMASPNRIRNLDAQQSTSVRSQEQDSKCCGGISVMNKGKEVSQSCFVYNADKGEVGSCPTRCCKDVSD
GSASRKGVIPSISEKQVYCATTVHNERPWADNAWRARLVKAISQKDYVLPNNAMNAQSASSFEKFGNAEKVAGKL
AGFLGKENDNHQDQVMQENHHGNHQDRIVQENHNDSHQYHVMQENHMDEPVLGRCGSASGGNPVSRCESASGVNS
AARYESSSGVNPTKLEKGKEKVMHDQSNCVSNIKEGDDSNESMESCQSMKAQKREHAQCSIAEMSSRTKRCRREF
NESSCSGFLQRNGSSFFNWVSSLTNGLTVFDKSTTDVSLDQKFSVSTVHEFAEQSGPLQNNSSVPVQSVGFNSLF
QSLYRHNVMITSTDTCHQSEKKCTEHEADRVALALNDSNSMLGKQIGTSRKTLDVPTETLAADSLLMDFGGGRGN
FQNQIGVFPLRAGRNLMMPNSSKSCSRSLEEKQNEAHAGSLKASVGNKGGFRESLWVSRLLPKTSMKLTDATPCN
IKSDFCAVNPKGAADKLYCSSQQNFSVEKEFNNSQYYTSTGSDNGTTSSKCPAIPPEENKQSETMASILAKRLDA
LRHAKTSAVRLGNSCDQTISKECNHGKSPFVVSYSGHDVQEARHETQKSSSGDGKLVLWLGDKGKGQLCTGSDEE
VRVKFLSGGDRQHCGGSMAGKAAAPHDNLEANTSAEYVQRRGVKIKEVLSNSMESLPDNKQIVPYGIMSSDEYDR
SSAVFGALERLRLSRSDIIRWLTSPVRHTTLDGFFLRLRFGKWEEALGGTGYHVARINGALDRNRLSVTIRNSTC
QVDSRFVSNHEFHEDELKAWWSAAMKGGWKLPSNEELSKKLRERELLHPQNRTGQHNDT

SEQ ID NO:11
>SbTZP_sorghum
MSPLSELVWSPDEGLSIKIAASSLSTRKASLRWNADTLSILISSPQHSGSGPGVKSGDTIYDNLEVSEKM
PSQLRIRSDSSVRVTMDSPNRVTNVDALQSTSIRSQEQDSSSASRKEVMPSISENQVCCATTVRNERSWA
TNAWRARLVKAVCQKEPMLSMNTENPMLPSSLGISCDAVEVSGKLVGSQGNRNVQSLGSDNNVISNAPEI
PSHNNCQDPVLQESHKDEPVVARGESASGVNAVARCESVPGVNSRKLVKDKEKVLYNDSNYGNNTMEGDD
SNESIESCTSTKAPKRKHAQFCAATMPSSGNKRFRREDNESSCSGLLQKCGTSFFNWMSSLTNGLPMLDG
ATAAIPLDQKFSASTGEGSAAPSQPLQNNSSVPMQSVGFNSLFQSLYTHNVMITSRDNCHQPESNYAGHV
FNRLTLELNDSNSMLDKQIGMGRETLDSGCKKNQCTAACSNDATQNKGGLRESLWVTRLLPKASVELMEA
TPCNVENAVNPQAVGDKLCCPPLQNFNLERGSSDGVTSSKCPATPPEEPKQSETMASVFAKRLDALRHAN
TSAVRLAIACDRGSPKLRNHKTNSFVVSYSSHDKVEAGHENHKSSSRNGRIVLWLGDKGKEQLCLSNKES
RGNFISEHEHQHHGGNTAGKSATRPNLELNTLAEDTDRSQVKLKGGVSDFMAGPSDNKQIVPYGTMPNDV
CDESSVVFGALHRLRLSRSDIIRWLRSPIMHTTLDGFFVRLRFGKWEEALGGTGYHVARLNGALDRSRLS
VTIRNSTCQVDSRFVSNHDFHEDELKAWWSAAMKSEWKLPSKEELSVKLRERELLRS

SEQ ID NO:12
>VvTZP_grape
MYRHRTKGKGKALSDGDRSGRKSNKEDDSDESVESCNSAALFSTGKKRWGYEQQLITGSK
RIRKQINGSPGSTSFVRQDSSFMSWISNMMKGLSKSNQDETPSLALTLARPNHDNYDQKL
VTCNKNQDPGCRNIGFQSIFQSLYCPTTKVQESRTLNADNQTGEGSKEFCLANKLCDFNQ
STFGNRAGPSTQPKVLSAKFAVSQENYKTSSVENRSASNPVSSSSSLGKRKANSAENNDS
DPPSEGKTIHNFGYKSDLLGSLWVTRFSPKTSSPTCKVDHCNQNTGGATELSTDCMGLIP
YSQNRTEVSFGFKKNNAHNNQNSIYKLNPISPSQRFKSSEAMASLFARRLDALKNIITLN
QTDTEARATPTCFFCGIRGHSIHDCSEIKETELEDLLRNNNLYPGAEEPPCFCIRCFQLN
HWAVACPSVLKRQNQSECGASLVNRCSSESQIIPLCNFVNPQISDVPKGIFDAIKRLRLS
RGDILKWMNSVFPFSHLNGFFLRLRLGKWEEGLGGTGYYVACISGAQKERPSQSSKNPIA
VNIGGVKCLVQSQYISNHDFLEDELMAWWGATTRAGGKIPSEEDLKVKLEERKKFGF

SEQ ID NO:16
>Populus2TZP
MEKERDNYMETGPYICPLEKLESTAENDFKTPHSENVCDVATEIVGSQNAKEVRSSSQQDDE
ILPKDNDC
AIKQSPTYSRTRRYQMKGKVKALSDGNLNERMLDMDDDSHEKNDFKTPHSENVCAVATEIVG
SQNAKEVR
SSSQQDDEILPKDNDCAIKQSPTYSRTRRYQMKGKAKALSDGNLNERMLDMDDDSHESVESC
NSVGLFST
GKRQRNFDPHSYVGSKSIKTKIQESPGSSSFVKHDGSFMNWISNMMKGFLKSNEDEAPSLAL
TLANHKHG
HEDRDKNLISCNRNQDQGCKTMGFHSLFQSLYCPKTKAQETVALNANTQTEGSKELGLDNKI
CDSNATPI
TCPMVTDNVYKRFLQPNEKLNESTSGNGAASPALTKLLSTNIASSQEISGSNSAEKKNSCNM
ATDKEKNG
TSSNSSPGKRKMNDAEQPSEGKATNTSGYRSDPLTSLWITRLSPKTSGPLSNRDLCHRRTGE
ALDGFTDF
IRLKAQWQNHPSSYQDKNIVGAREEEHFTEDPVCMHNCANSTEVSFSINKVNGHHDEKSMCK
MNSTLPFS
RFRNSEAMASVFARRLDALMHIMPSYGTDDSSHGNLTCFFCGIKCHHVRDCPEIIDSELADI
LRNANSFN
GANEFPCVCIRCFQSNHWAVACPSASSRTRHQAEYGASLVHESSPCKILLNPRNEDDAKQSD
GKDSQLQA
ADAPTVRNGKLHEASASGKINMNMKPFERDTASSSGEKKLKENQVMPLSNFINSQIADVPKG
IFDAVKRL
RLSRTIILK*

SEQ ID NO:17
>Gm3TZP
MSAENEKIKPKTDIELFLNNANQCIWKKLNNDSGAGANAASRADMTLAATDPLSEIVWSPDK
GLSLKCADSSFAHKNSSL
LRDVGTSCMVFAPPQNFTGGSSTTDKPLDDDFLKPIAVVCAKSDIAEADAPTMPPTGDSGVK
AKCKAYEEDDIGSVGNKE
KVNTAATAPNLPNEQNGNLTNNWEKITGDQANSGTDKVSGIEGNRISAISEFFFCCSGPTCD
VSFSVHRHSFVNSNGEGQ
ADQGPFDHLLLQSDENKPSMDQNPSPGRHSDGSVNIGLEKKAVVTDDDLHTAVEPIIEYRGS
GAHETNLASSSKNPLEKL
EYSAENDLQTFNCEAACAGTSRVNVSETENKFQDTEMMLPCDKILPVLHSPCHSRIHMAINK
GKEKSLSDGDANVMLSRE
ENDSHSSVESCNSTGFFSTGKKRRNFQQQLIIGSKRVKKQIEESSGPKSYVKQDNSFMNWIS
NMVKGLSPSIQNDSNTLA
LTLANPDHHNLQPDEKLIACNMNQDPEPKNTGFKSIFQSICCPSLKNVGTRMSHQEGKSSQD
LVPGNMEHGIDATPITCW
AENNSLSKLCLQSNKFEVSTGGNDAGLSSQPKIKPLNFFNCHESSKNNPVETKNYSILGHSK
DKEEVASHSSSTKQNTDN
NDNIDSNVLCDRKEEENICHRRDNLGSLWITRFSPKFTAPLREQPANDTEVSTDLKEDKGNN

DHKSMYLFKPLSSSPGFR
NLEPTASMFGRRFGAIKQIIPTNATDTTTQVNMLCFFCGTRGHQLSDCLAIAENKLEDLQKN
IDSYGGLEEHHCLCIKCF
QPNHWAISCPTSISTRKHELKANALVNDCGKHLISSNEGSARLLTDEDDRVLSGGSINDETD
QRAGQNINLKWKSNEIIT
HKVGCNASFKKYRGLSSEENKFRENPTLSPSKLAERQISQVPKEIFEAVKKLQLSRTDILK*

SEQ ID NO:18
>CpTZP2
MNVDKDNLEPHTDLGLTVNYSNHCIQRRLNNHLGAGANAGSSRGMTFVTTDPLSELVWSPQK
GPSLKCADCSFSDKKSSL
LWDAGPSNVVFSSAPPMISAVRCSNDKLMNENNNNMIKSHMGFYVKSDVGDRHTSARCPTNNA
AIKPVSGPSPEQKTGTGG
HMEETSTILQLSVQHANKNDDLSRPKGEVTCDLNTIQADEAAEAIEDNLKNFLDVLRPNVAQ
TEPLFENPAGGYCDANSE
NQTSKIEINLMSDIHAKNKTEACDDALKSQTVPPKRREEPASFTEGENDRKIKSTDSSNIRC
MEKLESTAENDLQIHIGE
NGCDAASKSVATEFAHEGKRCSQQEKARFREKLASGKYSAKNSGIQKYERKGKEKALSDGNL
SMMSKEADDSYESVESCN
GTGLFSTGKRRWNFEQQMIAGSKRAKMQIRETLESTSFVKQDSSFVNWIGNMMKGFLKLKED
EAPSVSLTLAYPTNGGPD
HDTIATTRNKNPGCRTVGFQSIFQSIYCPQRRIQDVTTSNDKCHKEVELPNDIFNPNPVNCH
GNFTKQFTISDERNTEST
FGNVREQEYQLFRDKWLCNLHNSKGKEGTSSSSSLGKYKLNIAENIDSEPLSEGESGQNFGN
GSDHLGSLWIARLAPKTS
GFSLNLNQQNRSTDVDLEYSADCAKLPHSPQYHVSSPSEAKIVEARQHSLDDQKIAIGNDLP
KCAGETECKITCNSAKDH
NDHKSTYNMNPILSSPIENSEAMASLFARRLDTLRHIMPSDLVENTASATVSCFFCGRKGHH
LRDCPEMTDEDLEDLLRT
VNKYNGTEEFSSFCVRCFRLNHWAVTCPNASSKGNRQRECGASLAGEFSPSSGKHNSRIEEN
PKFLNGNDSQFQVAEVHS
LLDRNDSGIRASLKLNSTEKLVASSSGKNKLRGNEIVPLSKYITRESSDVPKGIFDAVRMLR
LSRTDILKWMNSQMAGSH
LDGCFLRLRVGKWEKGLGGTGYYVACITEAQRQSSPQSSKNSIFVVVRGIRCLVESQYISNH
DFLEDELNAWWFTIEKNG
GTIPMEEVLRLKDRHKYNLYIGVVCYVDKKF*

WHAT IS CLAIMED IS:

1. A plant comprising a heterologous expression cassette comprising a promoter operably linked to a polynucleotide encoding a TZP polypeptide.

2. The plant of claim 1, wherein the promoter is heterologous to the polynucleotide.

3. The transgenic plant of claim 1, wherein the plant is characterized by increased size, speed of growth, elongation of organs or delay or extension of time of seed and/or fruit production compared to a plant lacking the heterologous expression cassette.

4. The transgenic plant of claim 1, wherein the promoter is a tissue-specific promoter.

5. The transgenic plant of claim 1, wherein the TZP polypeptide is at least 50% identical to SEQ ID NO: 2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:17, or SEQ ID NO:18.

6. An isolated nucleic acid comprising a promoter operably linked to a polynucleotide encoding a TZP polypeptide wherein the promoter is heterologous to the polynucleotide.

7. The isolated nucleic acid of claim 6, wherein the TZP polypeptide is at least 50% identical to SEQ ID NO: 2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:17, or SEQ ID NO:18.

8. A method for increasing size, speed of growth or biomass in a plant, the method comprising,

introducing an expression cassette into one or more plants, the expression cassette comprising a promoter operably linked to a polynucleotide encoding a TZP polypeptide, and

selecting a plant that has increased size, speed of growth, elongation of organs or delay or extension of time of seed and/or fruit production compared to a plant lacking the heterologous expression cassette.

9. The method of claim 8, wherein the TZP polypeptide is at least 50% identical to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:17, or SEQ ID NO:18. .

10. A method of producing processed plant material, the method comprising providing the plant of any of claims 1-4, or a part thereof; and

processing the plant or part thereof, thereby generating processed plant material.

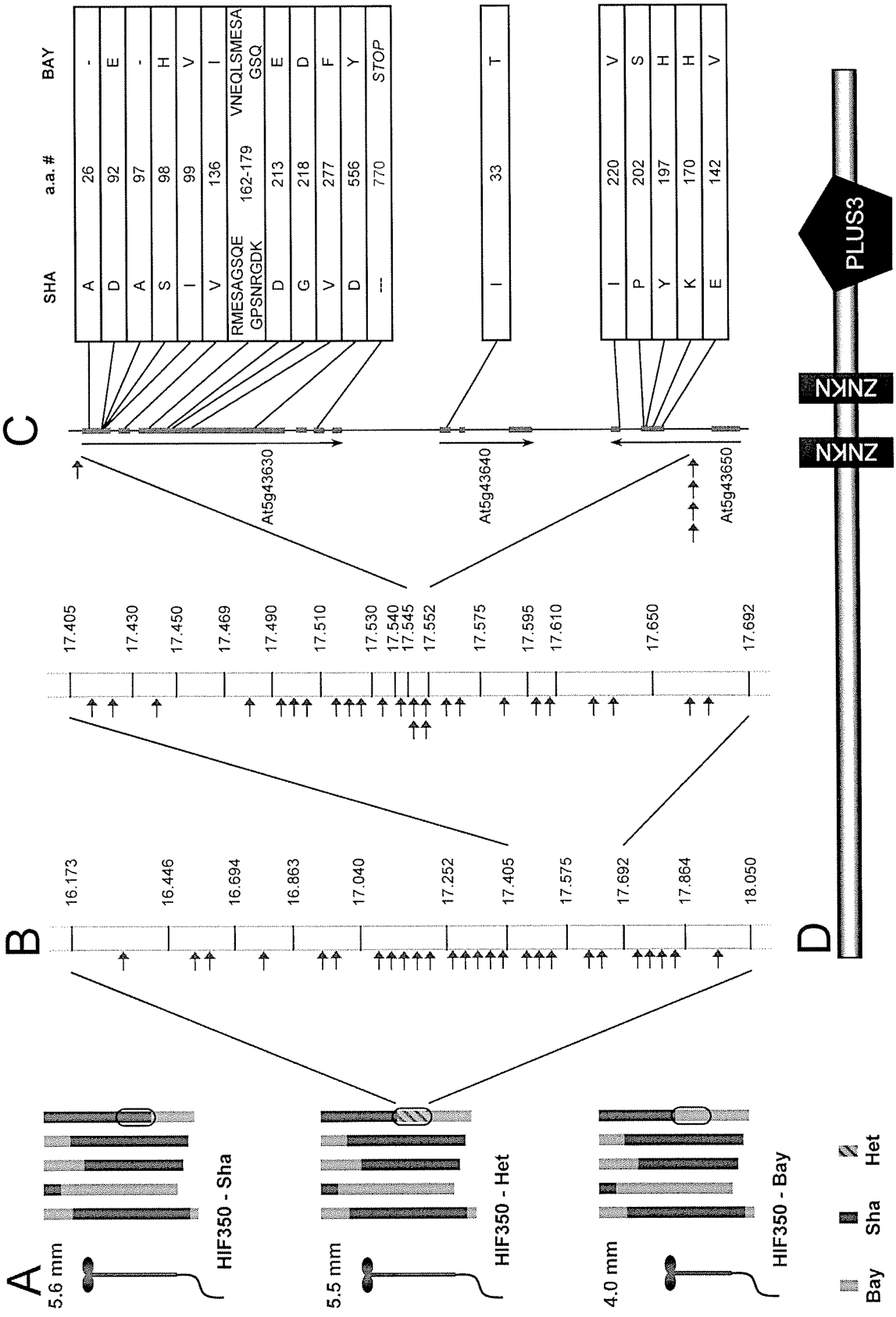
A
5.6 mm
HIF350 - Sha
5.5 mm
HIF350 - Het
4.0 mm
HIF350 - Bay
Bay
Sha
Het
B
16.173
16.446
16.694
16.863
17.040
17.252
17.405
17.575
17.692
17.864
18.050
17.405
17.430
17.450
17.469
17.490
17.510
17.530
17.540
17.545
17.552
17.575
17.595
17.610
17.650
17.692
C
SHA
a.a. #
BAY
A 26 -
D 92 E
A 97 -
S 98 H
I 99 V
V 136 I
RMESAGSQE 162-179 VNEQLSMESA
GPSNRGDK GSQ
D 213 E
G 218 D
V 277 F
D 556 Y
--- 770 STOP
I 33 T
I 220 V
P 202 S
Y 197 H
K 170 H
E 142 V
At5g43630
At5g43640
At5g43650
D
PLUS3
ZNKN
ZNKN

FIGURE 1

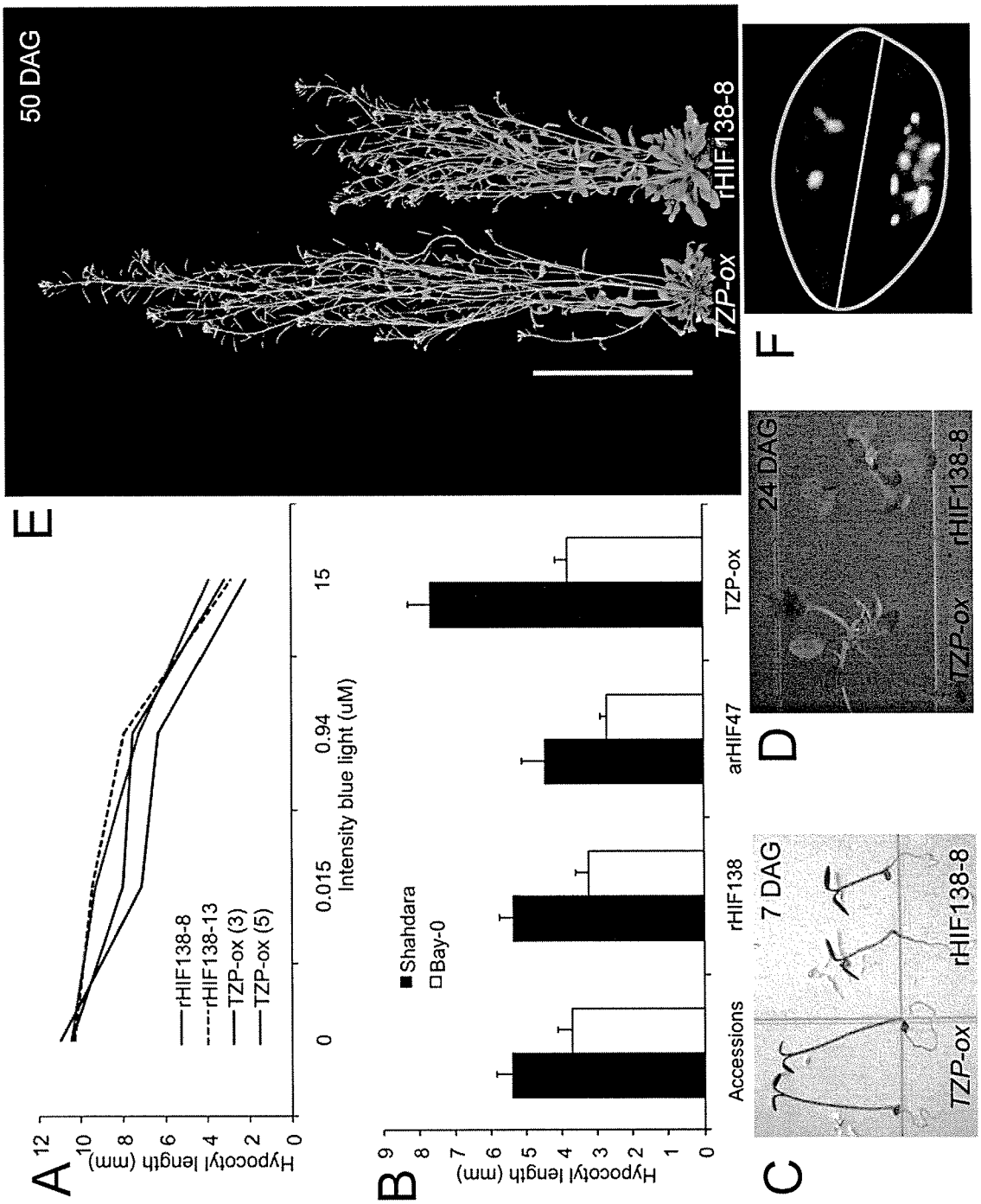
A
Hypocotyl length (mm)
12
10
8
6
4
2
0
0
0.015
0.94
15
Intensity blue light (uM)
rHIF138-8
rHIF138-13
TZP-ox (3)
TZP-ox (5)
B
Hypocotyl length (mm)
9
8
7
6
5
4
3
2
1
0
Shahdara
Bay-0
Accessions
rHIF138
arHIF47
TZP-ox
C
7 DAG
TZP-ox
rHIF138-8
D
24 DAG
TZP-ox
rHIF138-8
E
50 DAG
TZP-ox
rHIF138-8
F

FIGURE 2

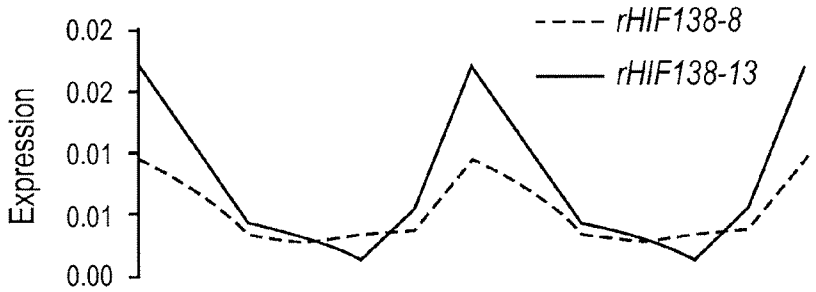

FIGURE 3A

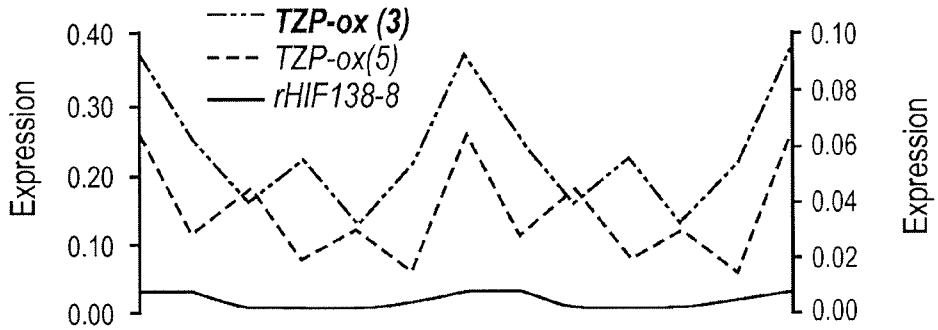

FIGURE 3B

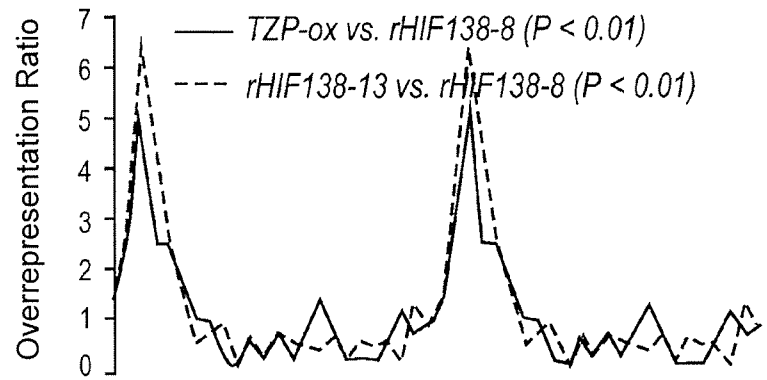

FIGURE 3C

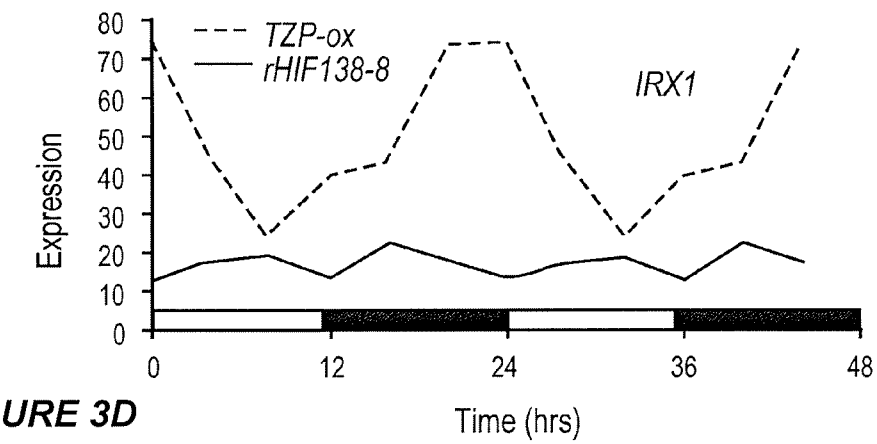

FIGURE 3D

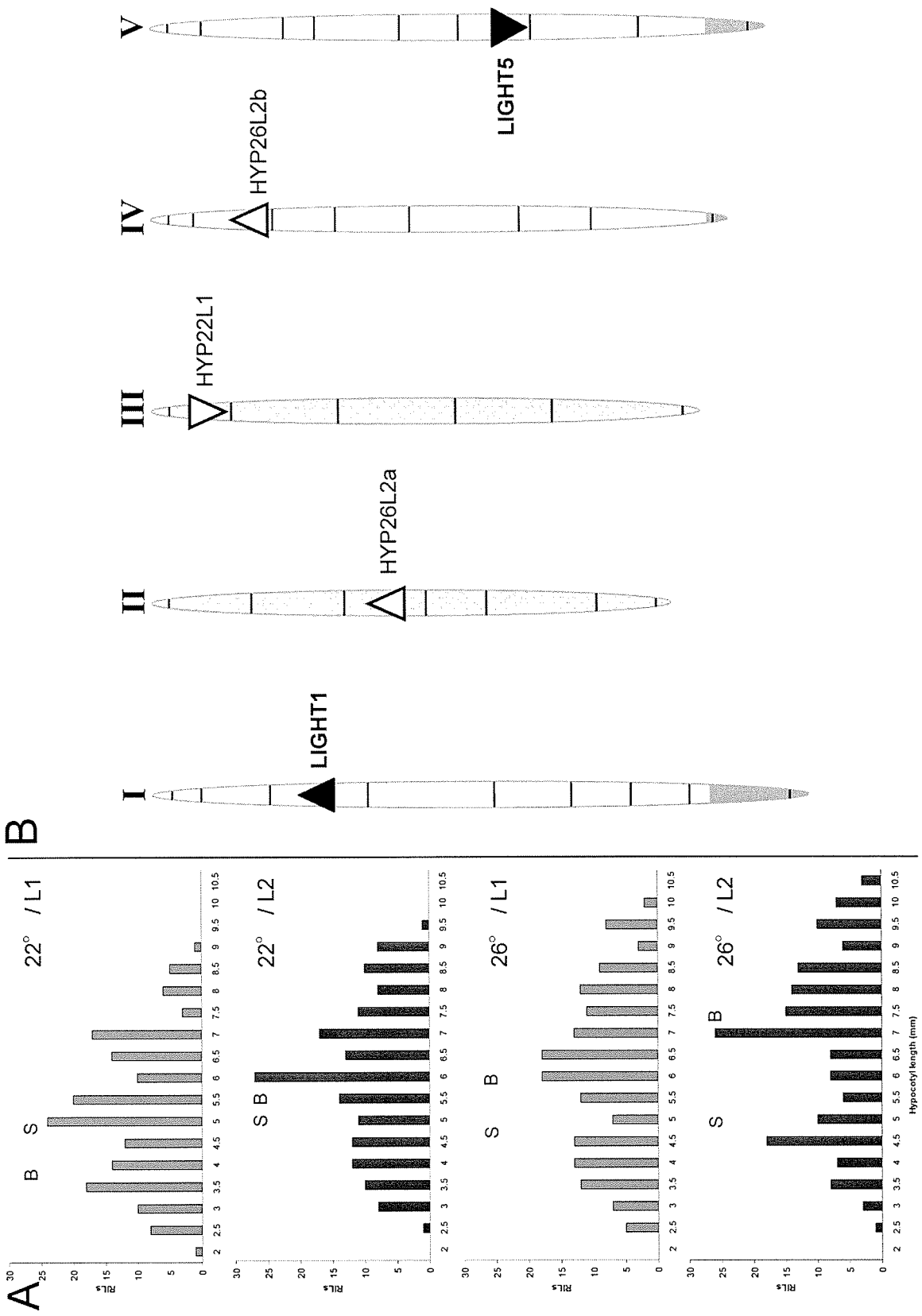
A
B
S
22° / L1
22° / L2
26° / L1
26° / L2
RILs
Hypocotyl length (mm)
B
I
II
III
IV
V
LIGHT1
HYP26L2a
HYP22L1
HYP26L2b
LIGHT5

FIGURE 4

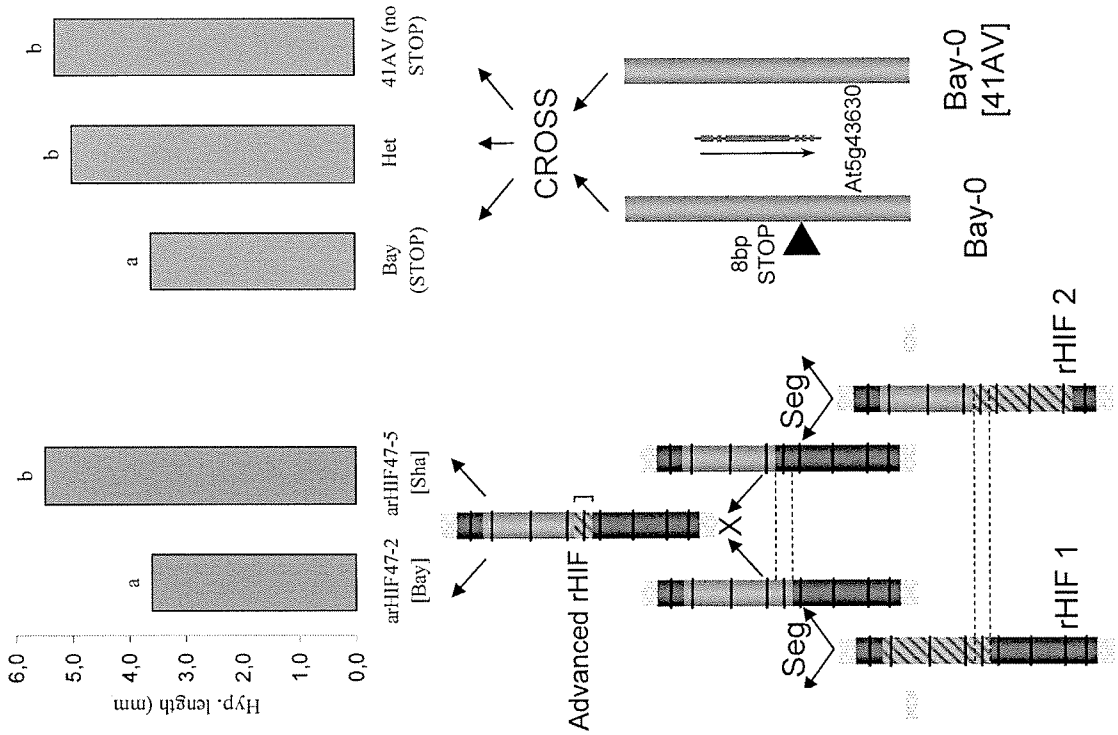
Hyp. length (mm
6,0
5,0
4,0
3,0
2,0
1,0
0,0
a
b
arHIF47-2 [Bay]
arHIF47-5 [Sha]
Advanced rHIF [ ]
X
Seg
Seg
rHIF 1
rHIF 2
Bay-0
Bay-0 [41AV]
8bp STOP
At5g43630
CROSS
Bay (STOP)
Het
41AV (no STOP)

FIGURE 5

```
                  :. : : :  :  :         : . :::: .  .         .              : :
BdTZP          --SAVFGALERLRLSRSDIIRWLTS-PVRHTTLDGFFLRLRFGKWEEALG--GTGYHVARIN---------GALDRNRLSVTIRNSTCQVDSRFVSNHEFHED--ELKAWWSAAMK  100
SbTZP          --SVVFGALHRLRLSRSDIIRWLRS-PIMHTTLDGFFVRLRFGKWEEALG--GTGYHVARLN---------GALDRSRLSVTIRNSTCQVDSRFVSNHDFHED--ELKAWWSAAMK  100
OsTZP          --SVVFGALQRLRLSRSDIIRWMRS-PVMHTTLDGFFLRLRFGKWEEALG--GTGYHVARIN---------GVLDKNRLSVTIRNSTCQVDSRFVSNHDFHQD--ELKAWWSAAMK  100
AtTZP          --RAVFDAVQVLRLSRTDVLKWINT-KKSVSGLEGFFLRLRLGKWEEGLG--GTGYYVARIDGDTEGQSSRRHSEKSLISVKVKGVTCLVESQFISNQDFLEE--ELKAWWQSAGK  109
AtTZPaccessions --KAVFDAVQGLRLSRTDVLKWINT-KKSVSGLKGFFLRLRLGKWEEWLG--GTGYYVARIDGDTEGQSSRRHSERSLISVKVKGVTCLVESQFISNQHFLEE--ELKAWWQSAGK  109
AlTZP          --KAVFEAVQVLRLSRTDVLKWINT-KKSVSGLEGFFLRLRLGKWEEGLG--GTGYYVARIDGDTEGQSSRRPSEKSSISVKVKGVTCLVESQFISNHDFLEE--ELKAWWRSAGK  109
PtTZP          -----FDAVKMLRLSRAVILKLMDS-HAAPSRLDGFFLRLRLGKWEQGLG--GTGYYVACI--TGVESESSTQKSKNSIAVNVGGIKYLVESQCISNHDFTED--ELMAWWRATLK  104
VvTZP          -----FDAIKRLRLSRGDILKWMNS-VFPFSHLNGFFLRLRLGKWEEGLG--GTGYYVACI--SGAQKERPSQSSKNPIAVNIGGVKCLVQSQYISNHDFLED--ELMAWWGATTR  104
CpTZP          -----FDAVRMLRLSRTDILKWMNS-QMAGSHLDGCFLRLRVGKWEKGLG--GTGYYVACI--TEAQRQSSPQSSKNSIFVVVRGIRCLVESQYISNHDFLED--ELNAWWFTIEK  104
RcTZP          VPNGIFDAVRSLRLTRTNILKWMNS-SASLS-IDGYFVRLRLGKWEEGLG--GTGYYVARI--TGMK-------SKKSIAVNVGGIQCVIESQFVSNHDFLED--ELKAWWSATSK  101
GmTZPa         --KKIFDAVKKLQLSRTDILKCINT-HGSISQLDGFFLRLRLGKWEEGLG--GTGYHVAYINETQSQRQCPEQNTRKCLSVKVGSIKCMVESQYISNHDFLEE--EITEWWSNTSE  109
GmTZPb         --KEIFEAVKKLQLSRTDILKWINT-RGSISQLDGFFLRLRLGKWEEGPG--GTGYHVAYINETQSQRQCSEQNTRKSLSVKVGSIKCMVESQYISNHDFLEE--EIMEWWSNTSQ  109
SmTZP          --KDMLAAIESVRLSRSELWRWMDS-PDFNTNVRGFFLRLRFGRWEKDLG--GTGYRMARIR--GAIRKLGGDDKNMSVSVDLGDMECTVDSQYVSNHSFTEGSKDLSSLWIVLCS  109
At3g51120      --AIDVHNINLIYLRRKFLESLLDDINKVDEKVVGTILRIKVSGSDQKLD-IHRLVQVVGTSKAIASYQLGAKTTDVMLEILNLDKREVISIDQLSDQNITED--ECKRLRQSIKC  111
AT2G16480      --AVDMHNINLIYLRRSLVEDLLEDSTAFEEKVASAFVRLRISG-NQKQD-LYRLVQVVGTSKAPEPYKVGKKTTDYVLEILNLDKTEVISIDIISNQDFTEV--TY--------C  102
At5g08430      --AIVSDNIKLLYLRKSLVQELLKSPDTFEGKMLGSFVRIKSDPNDYLQKYPYQLVQVTGVKKEHG-------TDDFLLQVTNYVKD--VSISVLSDDNFSQE--ECEDLHQRIKN  103
At5g63700      --TINADNLKLVYLRKSLVLELLKQNDSFVDKVVGSFVKVKNGPRDFMA---YQILQVTGIKNADDQ------SEGVLLHVSGMASG--VSISKLDDSDIREE--EIKDLKQKVMN  101
AtVIP5         --VPTFEDVKEVTIRRSKLAKWLME-PFFEELIVGCFVRVGIGRSKSGPIYRLCWVKNVDATDPDKTYKLENKTTHKYLNVVWGNETSAARWQMAMISDGHPLEEEYRQWIREVER  113
ruler          1.......10........20........30........40........50........60........70........80........90.......100.......110......
```

FIGURE 6A

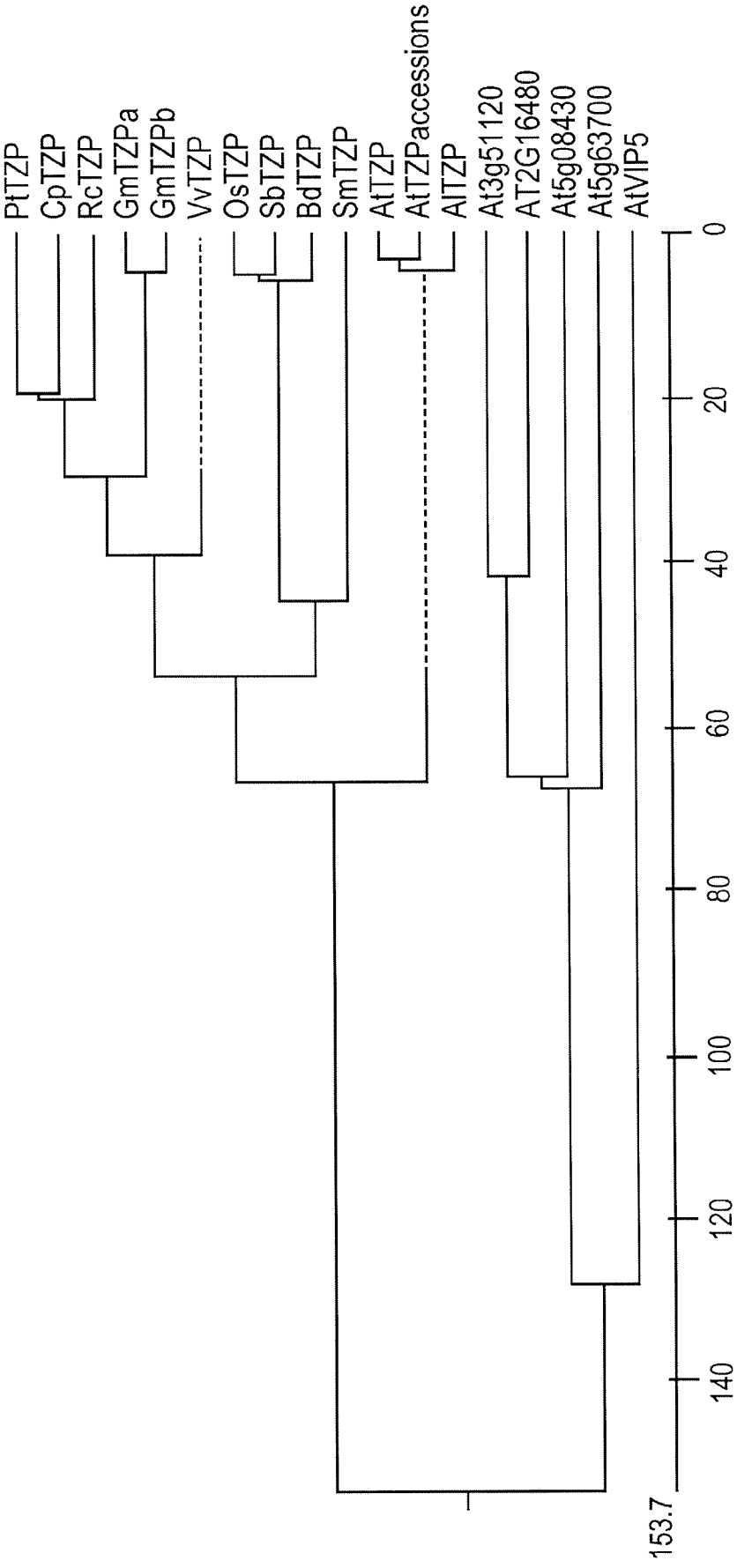
PtTZP
CpTZP
RcTZP
GmTZPa
GmTZPb
VvTZP
OsTZP
SbTZP
BdTZP
SmTZP
AtTZP
AtTZPaccessions
AlTZP
At3g51120
AT2G16480
At5g08430
At5g63700
AtVIP5
0
20
40
60
80
100
120
140
153.7
Amino Acid Substitutions (x100)

FIGURE 6B

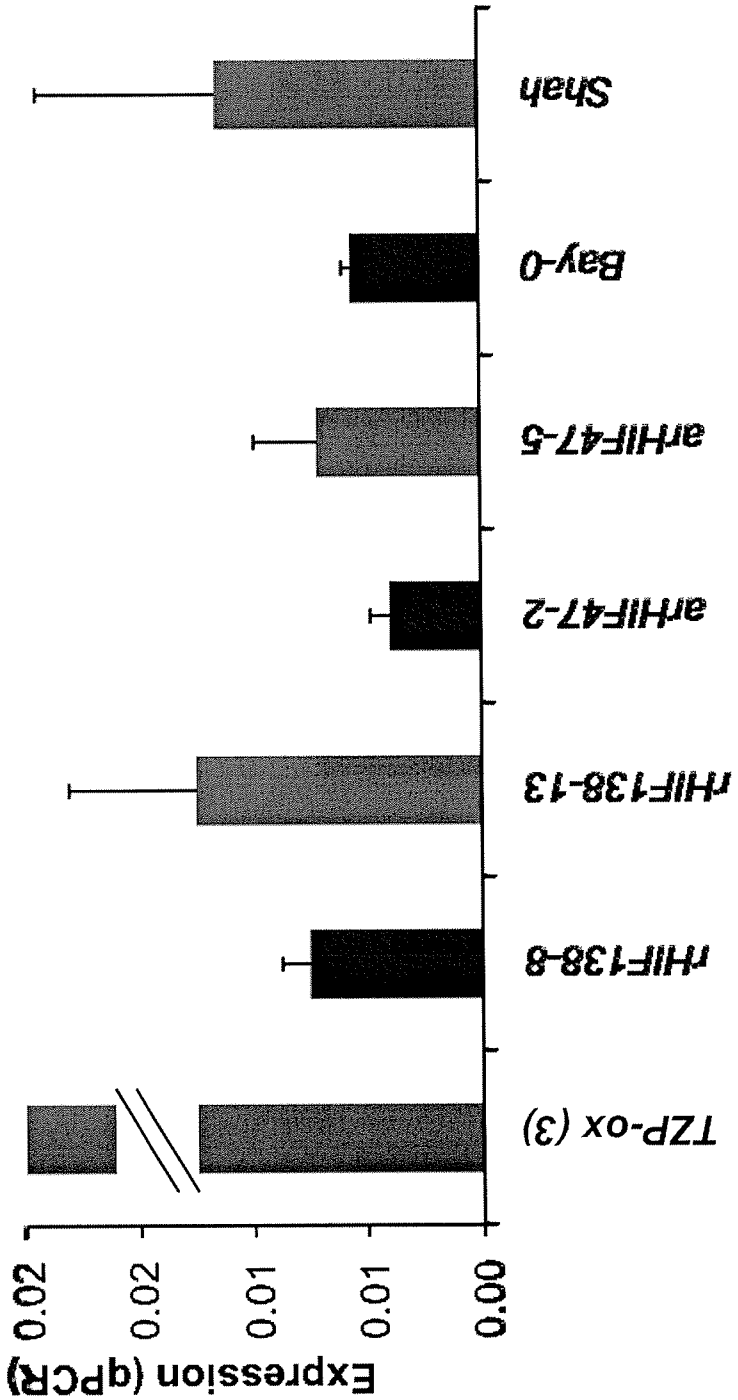
Expression (qPCR)
0.02
0.02
0.01
0.01
0.00
TZP-ox (3)
rHIF138-8
rHIF138-13
arHIF47-2
arHIF47-5
Bay-0
Shah

FIGURE 7

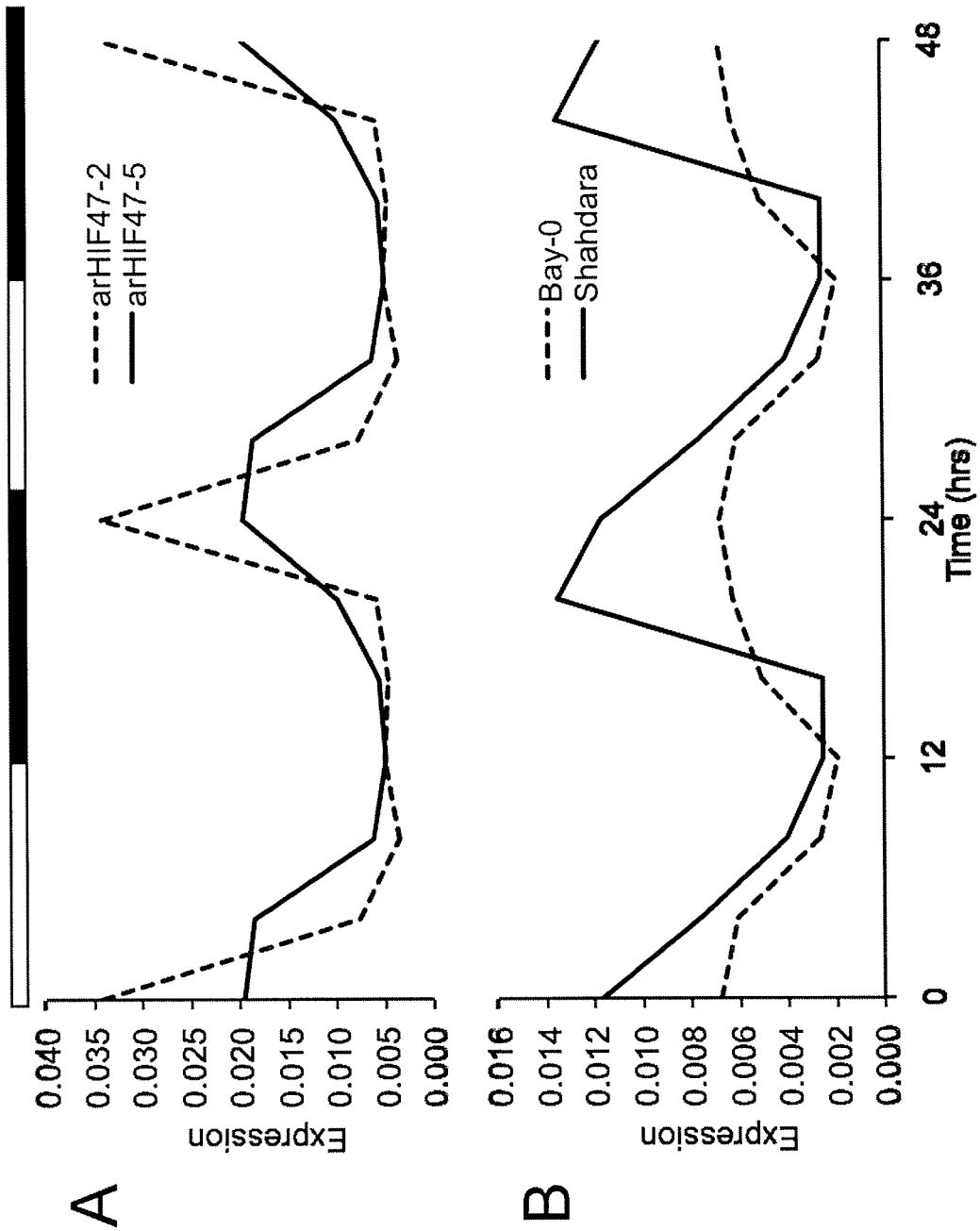
A
Expression
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
arHIF47-2
arHIF47-5
B
Expression
0.016
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
Bay-0
Shahdara
0
12
24
36
48
Time (hrs)

FIGURE 8

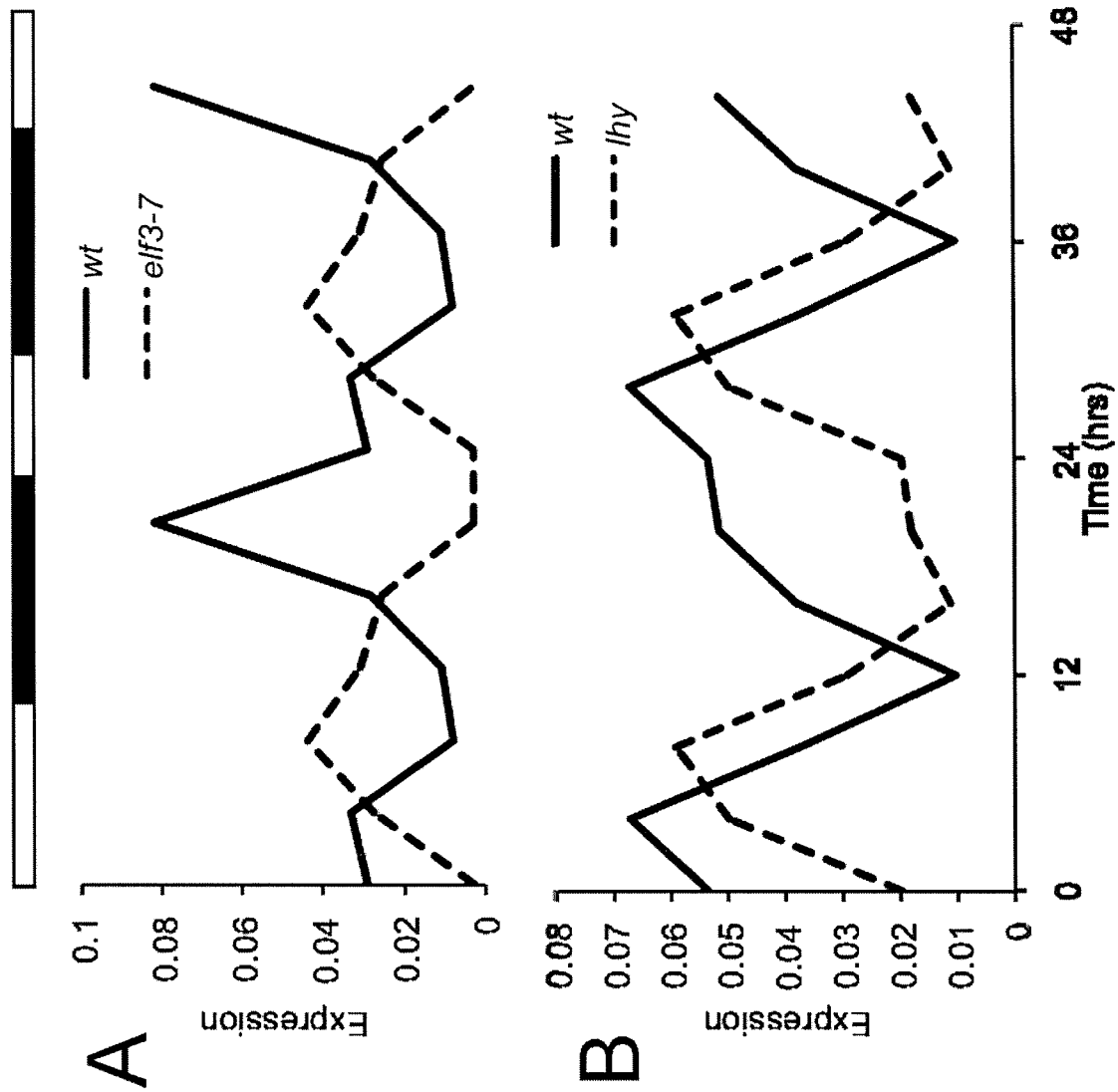
A
wt
elf3-7
Expression
0.1
0.08
0.06
0.04
0.02
0
B
wt
lhy
Expression
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
12
24
36
48
Time (hrs)

FIGURE 9

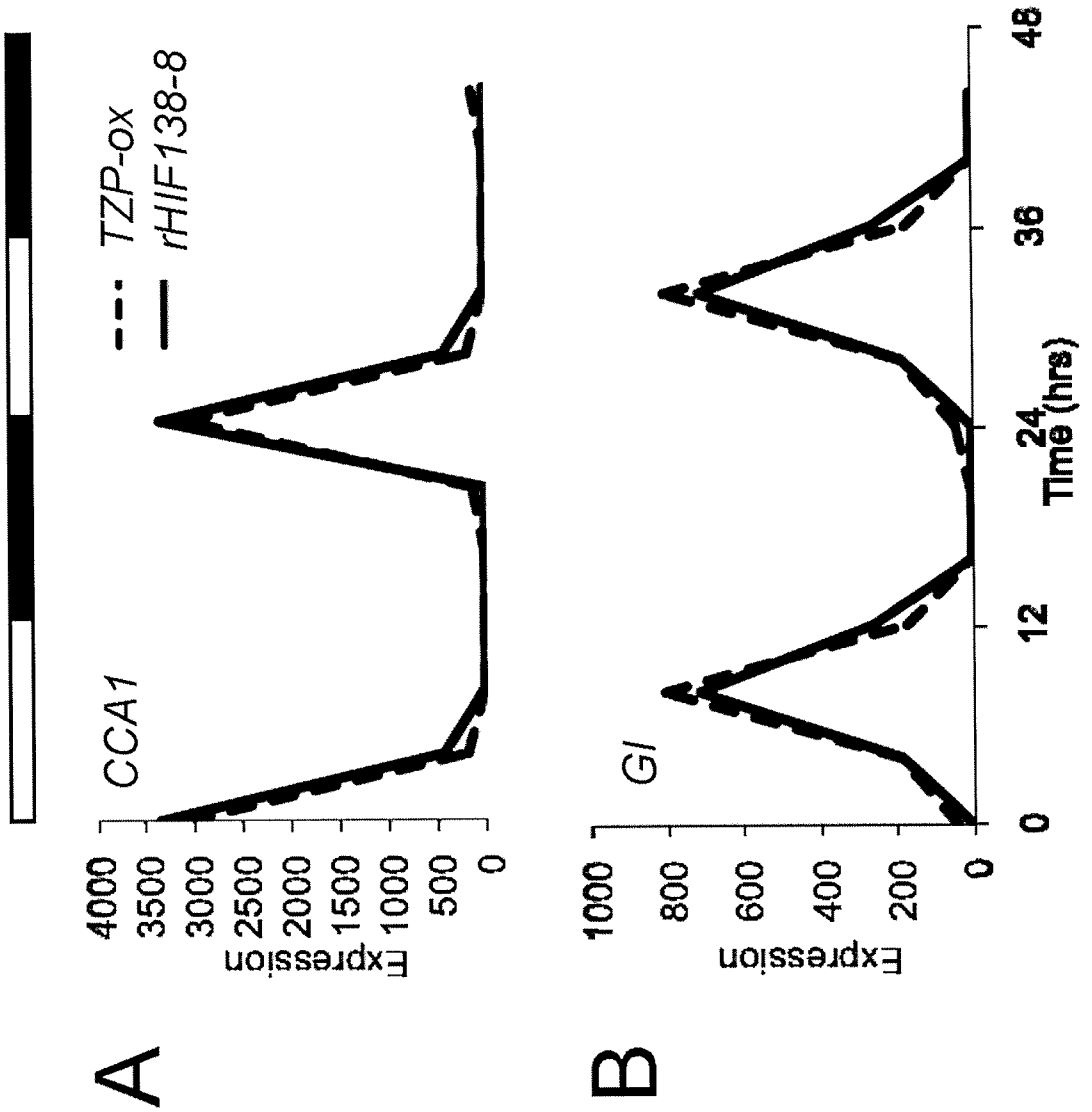
A
CCA1
TZP-ox
rHIF138-8
Expression
4000
3500
3000
2500
2000
1500
1000
500
0
B
GI
Expression
1000
800
600
400
200
0
0
12
24
36
48
Time (hrs)

FIGURE 10

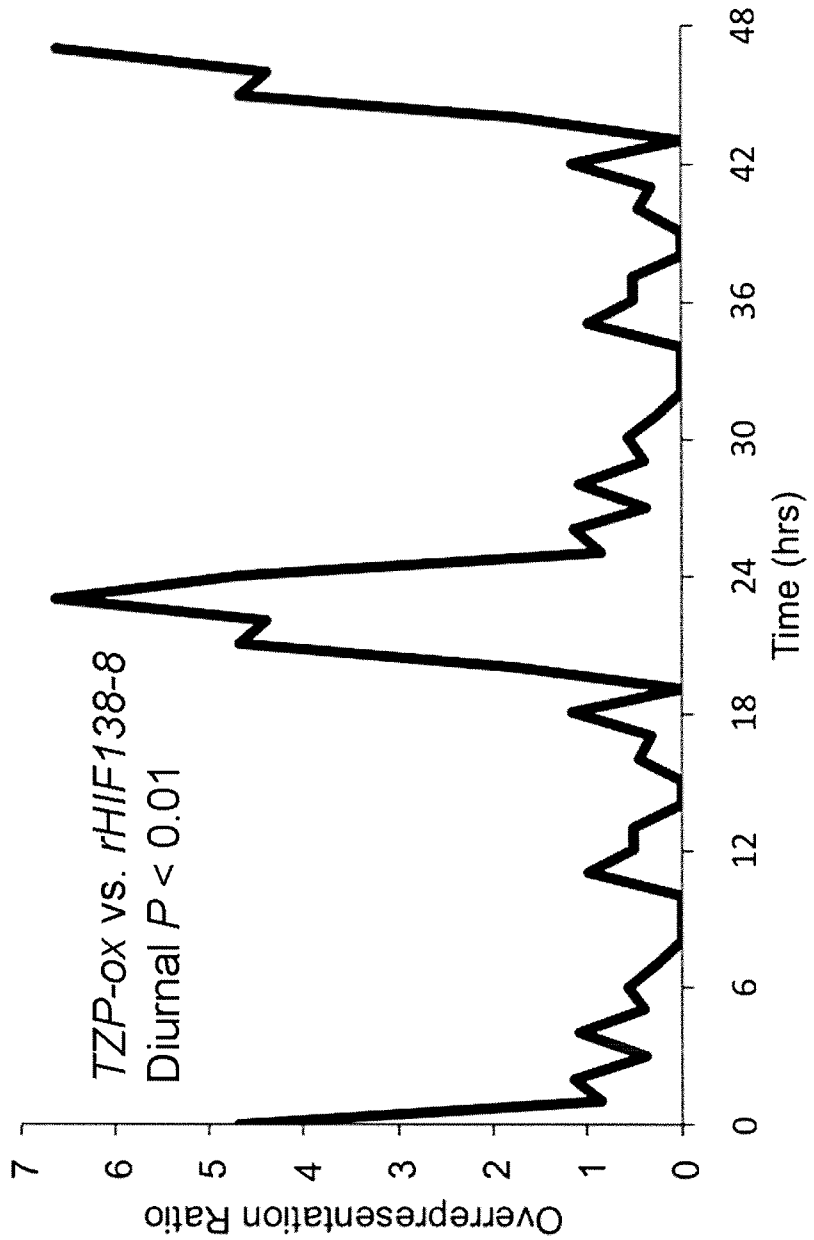
TZP-ox vs. rHIF138-8
Diurnal P < 0.01
Overrepresentation Ratio
Time (hrs)
0
1
2
3
4
5
6
7
0
6
12
18
24
30
36
42
48

FIGURE 11

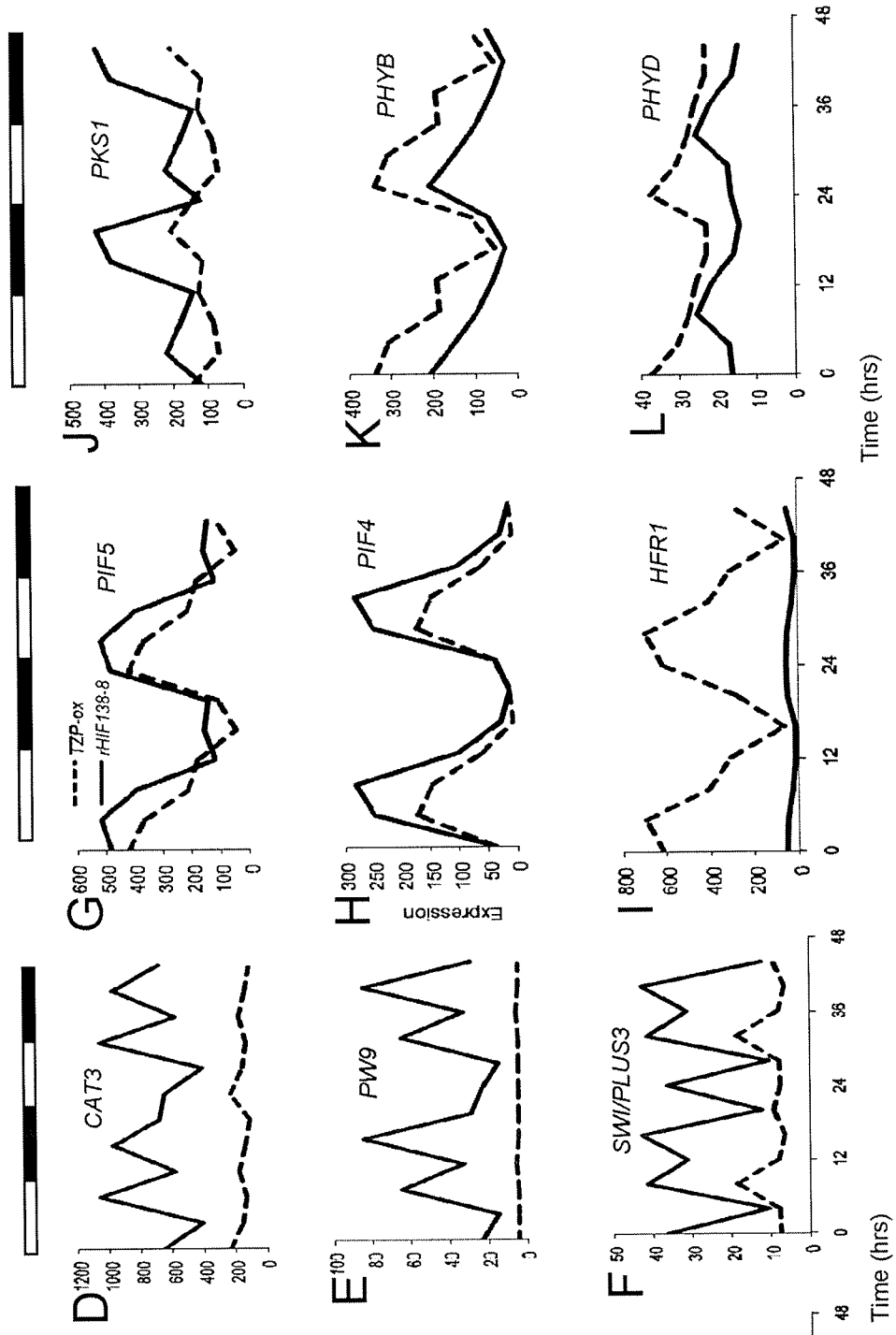
PEROXIDASE30
TZP-ox
rHIF138-8
CAT3
PIF5
PKS1
DFL1
PW9
PIF4
PHYB
HAT4
SWI/PLUS3
HFR1
PHYD
Expression
Time (hrs)

FIGURE 12

```
Al_pro_Created_  MGDGD...EQ SKELE...GL SSS..SRRSS GP..S..SAN AEARMKFASV
At_pro_Created_  MGDGD...EQ SKELG...GV SSS..SRRCS SG..TAGAAN AEARMKFAAV
BdTZP_brachypod  ..........  ..........  ..........  ..........  ..........M
RcTZP_castorbea  MNVDDKNIEP STNLSLALGY SNQCIQRNLS NDPGAGANAA STADITFVAT
CpTZP_papaya_pr  MNVDKDNLEP HTDLGLTVNY SNHCIQRRLN NHLGAGANAG SSRGMTFVTT
OSTZP_rice_pro_  ..........  ..........  ..........  ..........  ..........
PtTZP_poplar_pr  MDTNDKNIEP VIDLGFSLGY SNQCIQRRLK NDSGAGANAA SSVDMTFVAT
SbTZP_sorghum_p  ..........  ..........  ..........  ..........  ..........M
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created_  DAITELVWSP GNGLSLRCAD ISFTGKAKLV SPNFFDIGLT NMAIH....S
At_pro_Created_  DAITELVWSP SNGLSLRCAD ISFTGKAKLL SPNFFDIGLT NMAIH....S
BdTZP_brachypod  CPLSELVWSP DDGLSIKIAA SSLSTRKASL RWN...ADTL NIVISS..PQ
RcTZP_castorbea  DPLSELVWSP HKGLSLRCAD GSFIDKKPSL LPG...VGPT YMASGS..SS
CpTZP_papaya_pr  DPLSELVWSP QKGPSLKCAD CSFSDKKSSL LWD...AGPS NVVFSSAPPM
OSTZP_rice_pro_  ..........  ..........  ..........  ..........  ..........
PtTZP_poplar_pr  NALSELVWSP KKGLSLKCAD GTFSNQKPSL LRG...AGPS DMVSGS...N
SbTZP_sorghum_p  SPLSELVWSP DEGLSIKIAA SSLSTRKASL RWN...ADTL SILISS...P
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created_  NST.......  .........SI EHQE.....  ..DVELRSR.  ..........
At_pro_Created_  NST.......  .........SI EDQEDHV...  ..DVELRNR.  ..........
BdTZP_brachypod  QSSGRG...K SGDNID.ATI ADAGEM....  ..PSQPRTR.  CD..SSVRLS
RcTZP_castorbea  DK.PIS...N TGKLFD.NEI CIASLP....  ..ACKLASEI  SG..DNSTTF
CpTZP_papaya_pr  ISAVRC...S NDKLMNENNN MIKSHM....  ..GFYVKSDV  GDRHTSARCP
OSTZP_rice_pro_  ..........  ..........  ..........  ..........  ..........
PtTZP_poplar_pr  ADKAIG...K KVFMTP.PEE SDVRSEVAGR  DNPTKFVTS.  DT..GLFPLL
SbTZP_sorghum_p  QHSGSGPGVK SGDTIY.DNL EVSEKM....  ..PSQLRIR.  SD..SSVRVT
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created   ..........  ..........  ..........  ..........  ..........
```

Figure 13(A)

```
At_pro_Created_  ..........  ..........  ..........  ..........  ..........
BdTZP_brachypod  MASPN.R.IR  NLDAQQSTSV  RSQEQDSKCC  GGISVM....  ....NKGK..
RcTZP_castorbea  LTSNV.G.I.  .......MPL  SGTGLDKTAT  G..DQV....  ....VEMK..
CpTZP_papaya_pr  TNNAAIKPVS  GPSPEQKTGT  GGHMEETSTI  LQLSVQ....  ....HANKND
OSTZP_rice_pro_  ..........  ..........  ..........  ..MDVM....  .....NEGK..
PtTZP_poplar_pr  SESMH.K.VK  IGNYEPLAAT  DDHKEEMKTA  VGLPFLQKME  DARNNKAE..
SbTZP_sorghum_p  MDSPN.R.VT  NVDALQSTSI  RSQEQDS...  ..........  ..........
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created_  ..........  ..........  ..........  ...DQVNQ..  .........E
At_pro_Created_  ..........  ..........  ..........  ...DQVNQ..  .........A
BdTZP_brachypod  EVS....QSC  F.VYNADK..  ..........  ...GEVGSCP  T.RCCKDV..
RcTZP_castorbea  NAV....NYF  ..LQK.E...  ..........  ...DLRND..  ..KAEDETKL
CpTZP_papaya_pr  DLSRPKGEVT  CDLNTIQA..  ..........  ...DEAAE..  ...AIEDNLK
OSTZP_rice_pro_  ETS....DNF  C.VDKLEKE.  ..........  ...DEVGSCP  T.RYCNDTSH
PtTZP_poplar_pr  DIYDP..INL  Q.VDEISRTW  ETKEPSLSDE  TKLDVAQNGP  TSKEPNVRIG
SbTZP_sorghum_p  ..........  ..........  ..........  ..........  ..........
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created_  RIGG......  ..........  ........SV  ED........  MKPEMVEDKV
At_pro_Created_  MIGG......  ..........  ........SV  ED........  MKPEMVEDKV
BdTZP_brachypod  SDGSASRKGV  I...PSISEK  Q.VYC.ATTV  HNERPWADNA  WRARLVK.AI
RcTZP_castorbea  DVAQNYRTFE  E...PIVR..  ......ATDV  NDDHELGME.  .....IV.LV
CpTZP_papaya_pr  MFLDVLRPNV  AQTEPLFENP  AGGYCDANSE  NQTSKIEINL  MSDIHAKNKT
OSTZP_rice_pro_  SLGSASRKEV  M...PIIAEK  Q.AFC.ATTV  HDERSWAANA  WRARLVK.AI
PtTZP_poplar_pr  GVGDASHTLQ  T...EIVSAS  Q.....VCSV  EECESYDTNM  QRAPLGR.EH
SbTZP_sorghum_p  ..SSASRKEV  M...PSISEN  Q.VCC.ATTV  RNERSWATNA  WRARLVK.AV
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created_  ETDDDIKNEV  AGSSKRSSDS  PKVMEGETRD  L.........  ..........
At_pro_Created_  ETNDDIKNEE  AGCSKRSSDS  PKAMEGETRD  L.........  ..........
BdTZP_brachypod  SQKDYVLPNN  A..MNAQSAS  SFEKFGNAEK  VA........  .....GKLAG
RcTZP_castorbea  SDFHTVKGRE  D..YGIKIQN  AACSGKENEE  PPSVREKERK  NKMVIGRPGI
CpTZP_papaya_pr  EACDDALKSQ  TVPPKRREEP  ASFTEGENDR  KIKST.....  .....DSSNI
OSTZP_rice_pro_  SQKDSVLPKN  A..DNIHSTS  AFGSIGNTEN  MP........  .....GKLTS
PtTZP_poplar_pr  FESPSCMEKE  R..ENNMGTG  PY..ICPLEK  L.........  ..........
SbTZP_sorghum_p  CQKEPMLSMN  T..ENPMLPS  SLGISCDAVE  VS........  .....GKLVG
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created_  .LVNEQLRME  SAGS......  ..........  ..........  QEG.......
At_pro_Created_  .LVNEQLRME  SAGS......  ..........  ..........  QEEGDKAHNR
BdTZP_brachypod  FLGKENDNHQ  DQVMQENHHG  NHQDRIVQEN  HNDSHQYHVM  QEN.HMDEPV
RcTZP_castorbea  FSLDRLESTA  ENDL......  ..........  ..........  ETP.FGENSC
CpTZP_papaya_pr  RCMEKLESTA  ENDL......  ..........  .......QIH  GEN.GCDAAS
OSTZP_rice_pro_  MLGNRNDSSQ  DQAM......  ..........  ..........  QEK.HKDGLI
PtTZP_poplar_pr  .......ESTA  ENDF......  ..........  ..........  KTP.HSENVC
SbTZP_sorghum_p  SQGNRNVQSL  GSDNNVI...  ...SNAPEIP  SHNNCQDPVL  QES.HKDEPV
GmTZPb_soybean_  ..........  ..........  ..........  ..........  ..........
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........
```

Figure 13(A)

```
Al_pro_Created_  ..........  ..........  ..........  ..........  .....P....  .....TNRGD
At_pro_Created_  VDR.....LE  SMDENNLATL  AVVA...CEG  KGDYLPEGE.  AG..PSGSYR
BdTZP_brachypod  LGR......CG  SASGGNFV.S  RC......ES  ASGVNSAAR.  YE.SSSGVNP
RcTZP_castorbea  SMRNKNLASE  SADRVENN.T  QH......EL  IPIEYALGY.  NQ.SPTSSRL
CpTZP_papaya_pr  KSVATEFAHE  GKRCS.....  QQ......EK  ARFREKLAS.  GKYSAKNSGI
OSTZP_rice_pro_  VAR.....CE  SVSAVNFV.A  RC......EL  ASGVNPLAR.  HE.STSCCNP
PtTZP_poplar_pr  AVA.....TE  IVGSQNAKEV  RSSSQQDDEI  LPKDNDCAI.  KQ.SPTYSRT
SbTZP_sorghum_p  VAR.....GE  SASG......  ..........  ...VNAVAR.  CE.SVPGVNS
GmTZPb_soybean_  ..........  ..........  ..........  ...MLPCDKI  LPVLHSPCHS
VvTZP_grape_pro  ..........  ..........  ..........  ..........  ..........

Al_pro_Created_  K..DEGDKAN  NRIDRL....  ..........  .ESMDENNLA  TLAVVACEGQ
At_pro_Created_  R..REKAKGK  EKALSDENFG  GDGEDEDEES  FGSVESCNSA  GLLSRG...K
BdTZP_brachypod  T..KLE.KGK  EKVMHDQSNC  VSNIKEGDDS  NESMESCQSM  K..AQK.REH
RcTZP_castorbea  Q..NIQRQGQ  SKALSDGDAK  ERMLNEEDGS  HESVESCNST  ELFSTG.KQR
CpTZP_papaya_pr  Q...KYERKGK  EKALSDG.NL  SMMSKEADDS  YESVESCNGT  GLFSTG.KRR
OSTZP_rice_pro_  R..KLE.KGK  EKLIYDMSNC  VSNTNEGDDS  NESIESCPST  K..APK.RKH
PtTZP_poplar_pr  R..RYQHKGK  AKALSDGNLN  ERMLDMDDDS  HESVESCNSV  GLFSTG.KRQ
SbTZP_sorghum_p  R..KLV.KDK  EKVLYNDSNY  GNNTMEGDDS  NESIESCTST  K..APK.RKH
GmTZPb_soybean_  RIHMAINKGK  EKSLSDGDAN  VMLSRENDS  HSSVESCNST  GFFSTG...K
VvTZP_grape_pro  .MYRHRTKGK  GKALSDGDRS  GRKSNKEDDS  DESVESCNSA  ALFSTG...K

Al_pro_Created_  GEYSLENEA.  ..........  .GPSGSYRRP  KQDSSFMNWI  SNMTKGIWKG
At_pro_Created_  KRPGFEEQLI  FGSKRLKTLN  QECLGSTSKL  KQDSSFMNWI  SNMTKGIWKG
BdTZP_brachypod  AQCSIAEM.S  SRTKRCKREF  NESSCSGFLQ  RMGSSFFNWV  SSLTNGLTV.
RcTZP_castorbea  WNF.DQQL.I  VGSKRVKRQI  QDSPGSSSLG  KQDSSFVNWI  SNMMKGFLK.
CpTZP_papaya_pr  WNF.EQQM.I  AGSKRAKMQI  RETLESTSFV  KQDSSFVNWI  GNMMKGFLK.
OSTZP_rice_pro_  GQFSAAEM.T  SGNKRCRRED  NESSCSGLFH  KNDSSFFNWM  STLTNGVKV.
PtTZP_poplar_pr  RNF.DPHS.Y  VGSKSIKTKI  QESPGSSSFV  KHDGSFMNWI  SNMMKGFLK.
SbTZP_sorghum_p  AQFCAATMPS  SGNKRFRRED  NESSCSGLLQ  KCGTSFFNWM  SSLTNGLPM.
GmTZPb_soybean_  KRRNFQQQLI  IGSKRVKKQI  EESSGPKSVV  KQDNSFMNWI  SNNVKGLSP.
VvTZP_grape_pro  KRWGYEQQLI  TGSKRIRKQI  NGSPGSTSFV  KQDSSFMSWI  SNMMKGLSK.

Al_pro_Created_  NEEDDSPFAA  LTTTSDANGH  GQVNAIVDQQ  QLSPCCVKEN  SGCRNTGFQS
At_pro_Created_  NEEDNSPFVA  LTTTSNANGH  GQVNAIVDQQ  QLSPCCVKEN  SGCRNTGFQS
BdTZP_brachypod  F.DKSTTDVS  LDQKFSVST.  VHEFA...EQ  SGFLQNNSSV  P.VQSVGFNS
RcTZP_castorbea  S.SE.GEAPF  LSSALSNPNY  GHEMP...SQ  DVFTCNRKED  PACDTRGFQS
CpTZP_papaya_pr  LKEDEAPSVS  LTLAYPTNG.  GPDHD...TI  ATTRNKNPG.  ..CRTVGFQS
OSTZP_rice_pro_  F..DETTAVP  LNQKFSAAT.  GEEFP...TN  PVPLQNNCGV  P.LQSVGFNS
PtTZP_poplar_pr  SNEDEAPSLA  LTLANHKHG.  HEDRD...KN  LISCNRNQDQ  G.CKTMGFHS
SbTZP_sorghum_p  L.DGATAAIP  LDQKFSAST.  GEGSA...AP  SQPLQNNSSV  P.MQSVGFNS
GmTZPb_soybean_  SIQNDSNTLA  LTLANPDHHN  LQPD....EK  LIACNMNQDP  E.PKNTGFKS
VvTZP_grape_pro  SNQDETPSLA  LTLARPNHDN  Y..D....QK  LVTCNKNQDP  G.CRNIGFQS

Al_pro_Created_  LFQSIYCPKK  RSQDAVEMDF  PNDANATSLQ  ELPW......  ..........
At_pro_Created_  FFQSIYCPKK  QSQDVVDMDF  PNDVNAAPLQ  ELPW......  ..........
BdTZP_brachypod  LFQSLYRHN.  ..VM...ITS  .TDTCHQSEK  KCTEH.....  ..........
RcTZP_castorbea  VFQSLYCRK.  ..TKGQETVT  .LNVNHQTEG  SKECDQDNKI  CDLMAAPIAC
CpTZP_papaya_pr  IFQSIYCPQ.  ..RRIQDVTT  SNDKCHKEVE  LPNDIFNPNP  VNCHG.....
OSTZP_rice_pro_  LFQSLYSQN.  ..VM...ITS  .RNTCHQSES  SY..........  ..........
PtTZP_poplar_pr  LFQSLYCPKT  KAQETVALNA  .NTQTEGSKE  LG..........  ..........
```

Figure 13(A)

```
CpTZP_papaya_pr LDDQKIAIGN DLPKCAGETE CKI..TCNSA KDHNDHKSTY NWNPILSSPI
OSTZP_rice_pro_ VEKELNNSQY FTSKGSDNET TSS..KC.AA PQDEN..... ..........
PtTZP_poplar_pr REEEHFTEDP VCMHNCANST EVSFSINKVN GHHDEKSMCK MNSTLPFSRF
SbTZP_sorghum_p LERGSSD... .........GV TSS..KCPAT PPEEP..... ..........
GmTZPb_soybean_ LKEDKGNND. .......HKSM YLFK...PLS SSPGF..... ..........
VvTZP_grape_pro FKKNNAHNN. .......QNSI YKLK...PIS PSQRF..... ..........

Al_pro_Created_ ESSEAMASLF ARRLEAMKHI IPAGSLAENA EEEQPNLICF YCGKKGHCLQ
At_pro_Created_ ESSEAMASLF ARRLEAMKSI MPSGSLAENA EEEQRDLICF YCGKKGHCLR
BdTZP_brachypod KQSETMASIL AKRLDALRHA K...TS.... .AVRLGNSCD QTISKECNHG
RcTZP_castorbea KTSEAMASVS PRRLDAPKYI T...PSDDAD NSDRASMTCF FCGIKGHDLR
CpTZP_papaya_pr ENSEAMASLF ARRLDTLRHI M...PSDLVE NTASATVSCF FCGRKGHHLR
OSTZP_rice_pro_ KPSETMASIF AKRLDALRHA N...TS.... .AVHVAITCD HGTPKGRNHK
PtTZP_poplar_pr RNSEAMASVF ARRLDALMHI M...PSYGTD DSSHGNLTCF FCGIKCHHVR
SbTZP_sorghum_p KQSETMASVF AKRLDALRHA N...TS.... .AVRLAIACD RGSPKLRNHK
GmTZPb_soybean_ RNLEFTASMF GRRFGAIKQI I...PTNATD TTTQVNMLCF FCGTRGHQLS
VvTZP_grape_pro KSSEAMASLF ARRLDALKNI I...TLNQTD TEARATFTCF FCGIRGHSIH

Al_pro_Created_ DCLEVTDTEL RDLVQNISSR NGREEASSLC IRCFQLSHWA ATCP......
At_pro_Created_ DCLEVTDTEL RDLVQNISVR NGREEASSLC IRCFQLSHWA ATCP......
BdTZP_brachypod KSPFVVSYSG .......... .......... .......HDV QEARHETQ.K
RcTZP_castorbea ECSEVTDTEL EDLLRNINIY GGIKELPCVC IRCFQLNHWA VACPSTCPRV
CpTZP_papaya_pr DCPEMTDEDL EDLLRTVNKY NGTEEFSSFC VRCFRLNHWA VTCP......
OSTZP_rice_pro_ TSSFVVSYNS HDEQES.... .......... ...GQKTHKS SG........
PtTZP_poplar_pr DCPEIIDSEL ADILRNANSF NGANEFPCVC IRCFQSNHWA VAC.......
SbTZP_sorghum_p TNSFVVSYSS HDKVEAGHEN HKSSSRNGRI VLWLGDKGK EQL.......
GmTZPb_soybean_ DCLAIAENKL EDLQKNIDSY GGLEEHHCLC IKCFQPNHWA ISCPTSISTR
VvTZP_grape_pro DCSEIKETEL EDLLRNNNLY PGAEEFPCFC IRCFQLNHWA VAC.......

Al_pro_Created_ .......... .......... .......... .......... ..........
At_pro_Created_ .......... .......... .......... .......... ..........
BdTZP_brachypod SSSGDGKLVL WL.GDKGKGQ LCTGSDEEVR VKFLSGGDRQ HCGGSM..AG
RcTZP_castorbea RSKAECHASS VSHAGPSKSQ LHVINEDDTK AKNVTGSGHA ICYGNDYGMD
CpTZP_papaya_pr NASSKGNRQR EC.GASLAGE FSPSSGKHNS RIEENPKFLN GNDSQ...FQ
OSTZP_rice_pro_ ...GEGRIVL WT.GDKGKEQ LSPGNDKELG EKVLSKHENQ NCEGSS..DG
PtTZP_poplar_pr .......... .......... .......... .......... ..........
SbTZP_sorghum_p .......... .........CL SNKESRGNFI SEHEHQHHGG N.......TA
GmTZPb_soybean_ KHELKANALV NDCGKHLISS NEGSARLLTD EDDRVLSGGS INDETD.QRA
VvTZP_grape_pro .PSVLKRQNQ SECGASLVNR CSSESQIIPL CN...FVNPQ ISDVPK.GI.

Al_pro_Created_ ...NGPLYSS GAEDRAMKHT LASTSGMKLP VSGFT..... ..........
At_pro_Created_ ...NAPLYGS GAEGRAMKNA LASTSGMKLP ISGFT..... ..........
BdTZP_brachypod KAAAPHDNLE ANTSAEYVQR RGVKIKEVLS NSMSSLPDNK QIVP.YGIMS
RcTZP_castorbea KDMNSWKSNE AATSGKMKLN IRLFEKNISS TSREKELKEN QIIPLYGFVN
CpTZP_papaya_pr VAEVHSLLDR NDSGIRASLK LNSTEKLVAS SSGKNKLRGN EIVPLSKYIT
OSTZP_rice_pro_ KVVPPKCNLE TNTYIEEIDR KRLQNKEGAP NSMENQPDNK QMVP.YGIVP
PtTZP_poplar_pr .......... ..PSASSRTR HQAEYGASLV HESSPCKIL. ..........
SbTZP_sorghum_p GKSATRPNLE LNTLAEDTDR SQVKLKGGVS DFMAGPSDNK QIVP.YGTMP
GmTZPb_soybean_ GQNINLKWKS NEITTHKVGC NASFKKYRGL SSEENKFREN PTLSPSKLAE
VvTZP_grape_pro .......... .......... .......... .......... ..........

Al_pro_Created_ ....DVPKAV FEAVQVLRLS RTDVLKWINT KKSVSGLEGF FLRLRLGKWE
```

Figure 13(A)

```
At_pro_Created_  ....DVPRAV FDAVQVLRLS RTDVLKWINT KKSVSGLEGF FLRLRLGKWE
BdTZP_brachypod  SDEVDRSSAV FGALERLRLS RSDIIRWLTS PVRHTTLDGF FLRLRFGKWE
RcTZP_castorbea  GLISDVPNGI FDAVRSLRLT RTNILKWMNS S.ASLSIDGY FVRLRLGKWE
CpTZP_papaya_pr  RESSDVPKGI FDAVRMLRLS RTDILKWMNS QMAGSHLDGC FLRLRVGKWE
OSTZP_rice_pro_  NDVYDEASVV FGALQRLRLS RSDIIRWMKS PVMHTTLDGF FLRLRFGKWE
PtTZP_poplar_pr  LNFRNEDDAK QSDGKDSQLQ AADAPTVRNG KLHEASASG. ..KIN.....
SbTZP_sorghum_p  NDVCDESSVV FGALHRLRLS RSDIIRWLRS PIMHTTLDGF FVRLRFGKWE
GmTZPb_soybean_  RQISQVPKEI FEAVKKLQLS RTDILKWINT RGSISQLDGF FLRLRLGKWE
VvTZP_grape_pro  .......... FDAIKRLRLS RGDILKWMNS VFPFSHLNGF FLRLRLGKWE

Al_pro_Created_  EGLGGTGYYV ARIDEGQSSR RPS...EKSS ISVKVKGVTC LVESQFISNH
At_pro_Created_  EGLGGTGYYV ARIDGDTEGQ SSRRHSEKSL ISVKVKGVTC LVESQFISNQ
BdTZP_brachypod  EALGGTGYHV ARINGALDRN ..R..... LSVTIRNSTC QVDSRFVSNH
RcTZP_castorbea  EGLGGTGYYV ARITGMKSKK ..S..... IAVNVGGIQC VIESQFVSNH
CpTZP_papaya_pr  KGLGGTGYYV ACITEAQRQS ..SPQSSKNS IFVVVRGIRC LVESQYISNH
OSTZP_rice_pro_  EALGGTGYHV ARINGVLDKN ..R..... LSVTIRNSTC QVDSRFVSNH
PtTZP_poplar_pr  .......MNM KPFER..... ..R..... DTASSSGEKK LKENQVNPLS
SbTZP_sorghum_p  EALGGTGYHV ARLNGALDRS ..R..... LSVTIRNSTC QVDSRFVSNH
GmTZPb_soybean_  EGPGGTGYHV AYINETQSQR QCSEQNTRKS LSVKVGSIKC MVESQYISNH
VvTZP_grape_pro  EGLGGTGYYV ACISGAQKER P..SQSSKNP IAVNIGGVKC LVQSQYISNH

Al_pro_Created_  DFLEEELKAW WRSAGKSAER SGCEGIPSAE ELSRKIQQRK MLNR.......
At_pro_Created_  DFLEEELKAW WQSAGKSART SGYDGIPSAE ELSRKIQQRK MLGF.......
BdTZP_brachypod  EFHEDELKAW WSAAMKGGWK .....LPSWE ELSKKLRERE LLHPQNRTGQ
RcTZP_castorbea  DFLEDELKAW WSATSKVGGK .....LPSEK ELRLKVEEKN XXGGGGGNKI
CpTZP_papaya_pr  DFLEDELNAW WFTIEKNGGT .....IPMEE VLRLKDRHKY NLY.....I
OSTZP_rice_pro_  DFHQDELKAW WSRAMKSGWK .....LPSQE ELNTKLRERE LLR....F...
PtTZP_poplar_pr  NFINSQIADV PKGIFDA.... .....VK RLRLSRTIIL K.........
SbTZP_sorghum_p  DFHEDELKAW WSAAMKSEWK .....LPSKE ELSVKLRERE LLRS......
GmTZPb_soybean_  DFLEEEIMEW WSNTSQAGAE .....ISSEE YIIEKFKKKE MLG.......
VvTZP_grape_pro  DFLEDELMAW WGATTRAGGK .....IPSEE DLKVKLEERK KFG....F...

Al_pro_Created_  ..........
At_pro_Created_  ..........
BdTZP_brachypod  HNDT......
RcTZP_castorbea  VIENHIVQYQ
CpTZP_papaya_pr  GVVCYVDKKF
OSTZP_rice_pro_  ..........
PtTZP_poplar_pr  ..........
SbTZP_sorghum_p  ..........
GmTZPb_soybean_  ..........
VvTZP_grape_pro  ..........
```

Figure 13(A)

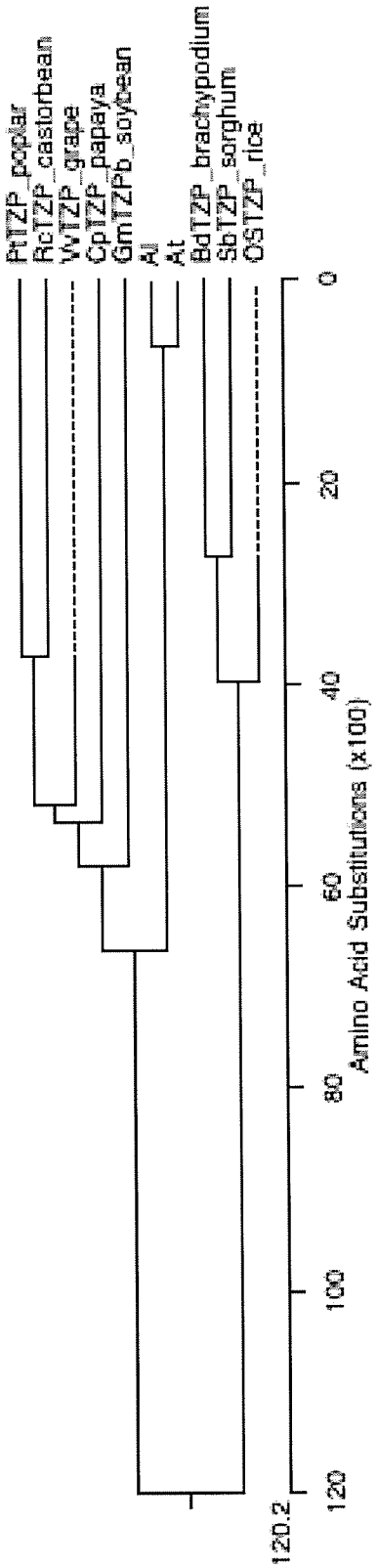
PtTZP_poplar
RcTZP_castorbean
VvTZP_grape
CpTZP_papaya
GmTZPb_soybean
Al
At
BdTZP_brachypodium
SbTZP_sorghum
OsTZP_rice
0
20
40
60
80
100
120
120.2
Amino Acid Substitutions (x100)

Figure 13(B)