International Journal of Computational Bioinformatics and In Silico Modeling
Vol. 7, No. 3 (2018): 991-1001
Research Article
Open Access

ISSN: 2320-0634

A bioinformatics characterization of maize (*Zea mays* ssp. *mays*) phytoglobins

Itzel Rojas-Sánchez and Raúl Arredondo-Peter*

Laboratorio de Biofísica y Biología Molecular, Centro de Investigación en Dinámica Celular, Instituto de Investigación en Ciencias Básicas y Aplicadas, Universidad Autónoma del Estado de Morelos, Avenida Universidad 1001, Colonia Chamilpa, 62210 Cuernavaca, Morelos, México.

*Corresponding author: Raúl Arredondo-Peter; email: ra@uaem.mx

Received: 30 April 2018 Accepted: 26 May 2018 Online: 02 June 2018

ABSTRACT

Phytoglobins (Phytogbs) are globins widely distributed in plants, ranging from algae to angiosperms. Maize (*Zea mays* ssp. *mays*) Phytogbs (Zeamaymphytogbs) have been partially characterized and consequently knowledge on Zeamaymphytogbs was limited. Thus, we used a bioinformatics approach to identify and characterize Zeamaymphytogb sequences and complement our knowledge on Zeamaymphytogbs. Main results from this work indicated that only *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* exist in the maize genome, *zeamaymphytogbs* could differentially express according to the maize physiology and growing conditions, Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 transcripts occur in embryonic and vegetative organs/tissues from normal maize, Zeamaymphytogb1.1 and Zeamaymphytogb1.2 transcripts occur at higher levels than those of Zeamaymphytogb3 in maize subjected to hormonal treatments and stress conditions, predicted Zeamaymphytogb3 and *Arabidopsis thaliana* Phytogb3 exhibit a highly similar structure and could function similarly, and Zeamaymphytogb1.1 and Zeamaymphytogb1.2 exhibit structural differences between each other which could result in different biochemical properties and function in maize plants..

Keywords: Function; gene expression; *in silico*; modeling; promoter; structure.

Abbreviations: Arathaphytogb, *Arabidopsis thaliana* Phytoglobin; Glb, globin; Mb, myoglobin; mya, millions of years ago; Orysatjphytogb, rice (*Oryza sativa*) var. japonica Phytoglobin; Phytogb, phytoglobin; RMS, root mean square; Zeamaymphytogb, maize Phytoglobin.

1. INTRODUCTION

The globin (Glb) superfamily comprises heme-containing proteins that bind a variety of ligands such as O_2 and nitric oxide (NO). A major function for Glbs in organisms is O_2-transport, however they exhibit diverse functions, including NO-detoxification [1, 2]. Structurally, Glbs are mostly formed by α-helices (designated with letters A to H) that fold into either a 3/3-folding (also known as the myoglobin [Mb]-fold) or 2/2-folding where helices A, E and F overlap to helices B, G and H and helices B and E overlap to helices G and H, respectively. It is believed that Glbs originated early in the evolution of life (*i.e.* ~3500 millions of years ago [mya]) and diversified with organisms as they have been detected from archaea and eubacteria to vertebrates [3, 4].

Globins are widely distributed in the Plant Kingdom, ranging from algae to angiosperms [5, 6]. Plant Glbs are named Phytoglobins (Phytogbs) and are classified into Phytogbs0, which exist in algae, bryophytes and gymnosperms, Phytogbs1 and Phytogbs2, which exist in angiosperms, symbiotic Phytogbs (Symphytogbs), which exist in non-legume N_2-fixing plants, leghemoglobins (Lbs), which exist in N_2-fixing legumes, and Phytogbs3, which exist in algae and land plants [7]. Structurally, Phytogbs0, 1 and 2, Symphytogbs and Lbs fold into the 3/3-folding whereas Phytogbs3 fold into

the 2/2-folding [8]. The best characterized Phytogb is Lb, which is specifically localized within nodules of N_2-fixing legumes. The apparent function of Lbs is to facilitate the diffusion of O_2 to the respiring bacteroids [9]. Although the function of Symphytogbs is poorly understood, it is possible that these proteins function similarly to Lbs within root nodules of non-legume N_2-fixing plants [8]. Land plant Phytogbs0 and Phytogbs1 and 2 are synthesized in diverse plant organs [10-14]. It has been proposed that they regulate a variety of plant processes by modulating the levels of NO (*i.e*. they exhibit an NO dioxygenase activity) [15-17]. A distinctive characteristic among Phytogbs0, 1 and 2 is that heme-Fe in Phytogbs0 and 2 and Phytogbs1 is mostly penta- and hexacoordinate, respectively [18-21], resulting in that Phytogbs1 exhibit a very high affinity for O_2 because in these proteins the O_2-dissociation rate constant is extremely low (*e.g*. the k_{off} of rice Phytogb1 is 0.038 s^{-1} [18]). Expression of the *phytogb3* genes has been detected in diverse plant organs [22], however little is known about the function of Phytogbs3. Because of their structural similarity to some bacterial truncated Glbs (tHbs) it is likely that Phytogbs3 and bacterial tHbs function similarly by detoxifying NO [8].

Monocot Phytogbs have been characterized with detail. For example, the kinetic rate constants of rice [18] and barley [19] Phytogb1 for O_2 have been determined, the crystal structure of rice Phytogb1 was elucidated [23] and the localization of Phytogbs in rice [13] and wheat [24] organs were reported. Also, a phylogenomic analysis revealed that only the *phytogb1* gene evolved within monocots which duplicated ~140 mya originating clade I and clade II *phytogb*s (*phytogbI* and *phytogbII*, respectively) [25].

Maize (*Zea mays* ssp. *mays*) genome contains at least one copy of the *phytogbI* (*zeamaymphytogb1.1*), *phytogbII* (*zeamaymphytogb1.2*) [25] and *phytogb3* (*zeamaymphytogb3*) [26] genes. A recent report using bioinformatics tools revealed that regulatory elements responsive to plant hormones and stress conditions could exist upstream to the *zeamaymphytogb1.1* and *zeamaymphytogb1.2* genes. This observation suggested that Zeamaymphytogb1.1 and Zeamaymphytogb1.2 function in physiological processes regulated by plant hormones and in the plant response to stress conditions [16]. Also, a Western blot analysis using a polyclonal anti-rice Phytogb1 antibody revealed that Phytogb proteins are localized in embryonic and vegetative organs from maize growing in normal conditions, thus suggesting that Zeamaymphytogbs also function in normal plants [27]. However, the above antibodies were not specific to Zeamaymphytogb1.1, Zeamaymphytogb1.2 or Zeamaymphytogb3 and therefore it was not known which Zeamaymphytogb (*i.e*. either Zeamaymphytogb1.1, Zeamaymphytogb1.2 or Zeamaymphytogb3) is functional in maize. Structurally, the tertiary structure of Zeamaymphytogb1.1 [28] and Zeamaymphytogb1.2 [25] was predicted (the predicted structure of Zeamaymphytogb1.1 has been verified by X-ray crystallography, PDB ID 2R50) showing that Zeamaymphytogb1.1 and Zeamaymphytogb1.2 are mostly hexacoordinate and fold into the 3/3-folding and that Zeamaymphytogb1.2 exhibits a disorganized and extended post-helix H region.

Despite the above information knowledge on Zeamaymphytogbs is limited. For example, it is not known the exact number of the *zeamaymphytogb* gene copies, the existence of regulatory boxes upstream to *zeamaymphytogb*s and the expression of *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*. Also, the tertiary structure of Zeamaymphytogb3 has not been elucidated. Hence, in this work we report a bioinformatics characterization that complements our knowledge on Zeamaymphytogbs, which includes the detection of gene copy number using homologous probes, gene characterization and mapping, promoter identification and analysis, gene expression in normal plants and in plants subjected to hormonal treatments and stress conditions and protein structure analysis.

2. MATERIALS AND METHODS

2.1. Search of Zeamaymphytogb sequences in databases

Putative maize Phytogb (Zeamaymphytogb) sequences were searched in the maize (*Zea mays* ssp. *mays*) B73 genome using the Phytozome database (http://phytozome.jgi.doe.gov/pz/portal.html) and *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* as query sequences [25, 26] (Phytozome accession number GRMZM2G067402, GRMZM2G168898 and GRMZM2G155868, respectively). Resulting sequences were subjected to a FUGUE analysis (http://www-cryst.bioc.cam.ac.uk) [29] to determine the most similar Phytogb structure and presence of proximal H at the Mb-fold position F8. Putative Phytogbs had to satisfy the following criteria: length higher than or ~100 amino acids, a FUGUE Z score higher than 6 (which corresponds to 99% specificity [29]) with known Glb structures, and the presence of proximal H at position F8.

2.2 Zeamaymphytogbs sequence alignment

Sequence alignments were performed using the ClustalX program [30]. Alignments were manually verified using the procedure described by Kapp *et al*. [31] based on the Mb-fold [32].

2.3. *zeamaymphytogb* gene mapping and detection of promoter sequences

Scaffolds containing copies of the *zeamaymphytogb* gene were used for mapping *zeamaymphytogb*s within the maize genome. This included the detection of open reading frames (ORFs) 25 kb up- and downstream to *zeamaymphytogb*s and ORF length, transcription orientation and localization in the +/- strand. CAAT-, TATA- and GC-boxes and *cis*-acting regulatory elements responsive to plant hormones (auxins, cytokinins, methyl jasmonate and abscisic acid), stress (light,

circadian cycle, temperature and anaerobiosis) conditions and specific expression in meristems and endosperm (http://www.cambia.org/daisy/promoters/239/g1/240.html) were searched within 1kb upstream to *zeamaymphytogb*s using the consensus sequence of the CAAT-, TATA- and GC-boxes and the search tool of MS Word2007® and the PlantCARE server (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), respectively.

2.4. Detection of Zeamaymphytogb transcripts in maize organs

The Zeamaymphytogb transcripts were identified in maize organs from normal plants using the Maize eFP Browser (http://bar.utoronto.ca/efp_maize/cgi-bin/efpWeb.cgi) [33] and the Zeamaymphytogb accession numbers GRMZM2G067402, GRMZM2G168898 and GRMZM2G155868 corresponding to *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, respectively (see subsection 2.1). The Zeamaymphytogb transcripts were identified in maize subjected to either hormonal treatments or stress conditions using the Maize 18k and Microarray platform options and default parameters of the PLEXdb database (http://www.plexdb.org/plex.php?database=Corn) [34] and the probe sets Zm.485.1.A1_at, Zm.11985.1.A1_at and Zm.5040.1.S1_at corresponding to Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3, respectively.

2.5. Modeling of Zeamaymphytogb3 and analysis of the Zeamaymphytogbs tertiary structure

The tertiary structure of Zeamaymphytogb3 was modeled using the I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) [35-37] and the crystal structure of *Arabidopsis thaliana* Phytogb3 (Arathaphytogb3) (PDB ID 4c0n) as template. Reliability of the selected model was assessed using the Veryfy-3D server (http://services.mbi.ucla.edu/Verify_3D/) [38]. Model was edited using the VMD program [39] and Adobe Photoshop® software. The structural variability between Zeamaymphytogb1.1 and Zeamaymphytogb1.2 and Zeamaymphytogb3 and Arathaphytogb3 was evaluated using the Magic fit and RMS (Root Mean Square) tools of the SwissPDBViewer program [40] and the predicted structure of Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 (Caspur PMDB ID PM0075007, PM0078174 and PM0080450, respectively) and the crystal structure of Arathaphytogb3 (PDB ID 4c0n). Distance of amino acids at the heme pocket to heme-Fe was calculated using the distance tool of the SwissPDBViewer program as described by Sáenz-Rivera *et al*. [28] and Gopalasubramaniam *et al*. [41].

3. RESULTS AND DISCUSSION

3.1. Detection of Zeamaymphytogb sequences in the maize genome

Previously, Rodríguez-Alonso and Arredondo-Peter [25, 26] identified a single copy of the *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* genes in the maize genome using the rice (*Oryza sativa*) Phytogb1.1 and Phytogb3 sequences (GenBank accession number AAK72229.1 and AAK72230.1, respectively) as heterologous probes. To increase the sensitivity for gene detection, we searched for *zeamaymphytogb* and *zeamaymphytogb*-like sequences in the maize genome using the *zeamaymphytogb* sequences identified by the above authors as homologous probes. As result we detected sequences with the Phytozome accession number GRMZM2G067402_T02, GRMZM2G168898_T01 and GRMZM2G155868_T01 which satisfied the criteria to be considered as Glbs, *i.e*. length higher than or ~100 amino acids, FUGUE Z score higher than 6 with known Glb structures and the existence of proximal H at position F8. Also, sequence alignment showed that GRMZM2G067402_T02, GRMZM2G168898_T01 and GRMZM2G155868_T01 are identical to *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, respectively. No *zeamaymphytogb*-like sequences (*e.g*. pseudogenes) were detected in the maize genome. Hence, apparently only the *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* genes exist in maize.

The *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* genes are 874, 957 and 2,051 nucleotides in length and code for predicted Zeamaymphytogb1.1, Zeamaymphytogb1.2 y Zeamaymphytogb3 proteins 166, 192 and 172 amino acids in length, respectively. Also, coding sequences for *zeamaymphytogb*s are interrupted by 3 introns (IVS-I, IVS-II and IVS-III) (supplementary Figure 1) located at positions B11.2, E13.0 and G6.0 for *zeamaymphytogb1.1* and *zeamaymphytogb1.2* and B13.0, F/F´2.2 and H13.0 for *zeamaymphytogb3* (Figure 1) similarly to other land plant *phytogb*s [3, 6, 42, 43]. The main difference among the *zeamaymphytogb*s length is due to the existence of a long 1,304 nucleotides IVS-II in *zeamaymphytogb3*. Additionally, nucleotide sequence among the *zeamaymphytogb*s exon/intron limits is variable but similar to that in other land plant *phytogb*s [5, 18, 27, 42] (supplementary Figure 1).

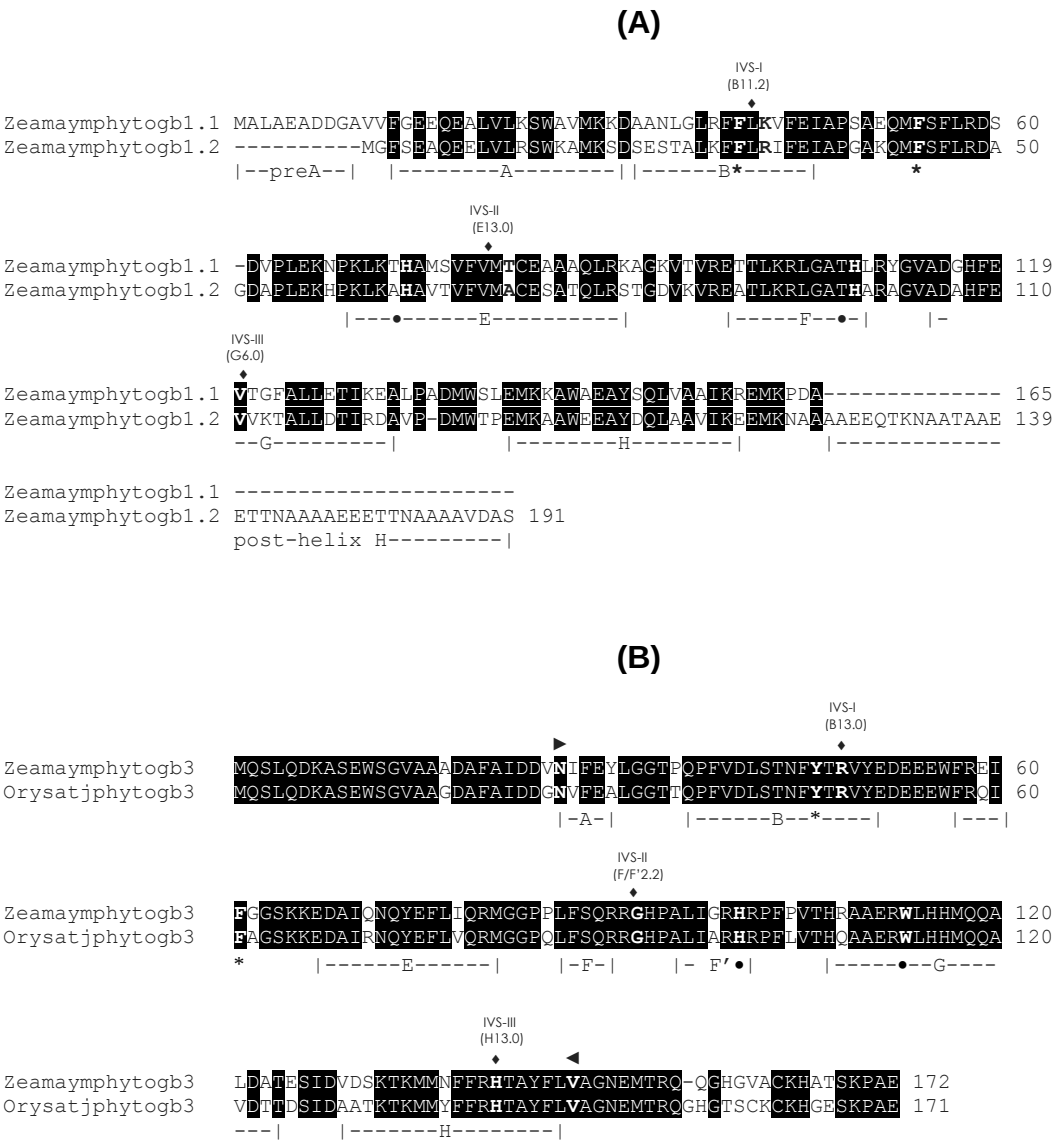

Figure 1. (A) Sequence alignment of Zeamaymphytogb1.1 and Zeamaymphytogb1.2. Pre-helix A, helices A to H and post-helix H are indicated based on Hargrove *et al*. [23] and Rodríguez-Alonso and Arredondo-Peter [25]. •, distal (H73/64) and proximal (H108/99) His; *, Phe B10 (F54/44) and CD1 (F40/30). (B) Sequence alignment of Zeamaymphytogb3 and *Oryza sativa* var. japonica Phytogb3 (Orysatjphytogb3) [47]. Helices A to H are indicated based on Reeder and Hough [46] and Arredondo-Peter *et al*. [47]. •, proximal H (H100) and proposed distal Trp (W113); *, Tyr B10 (Y46) and Phe CD1 (F61). Right- and left-oriented arrows indicate the limits for the globin domain. Intron (IVS) position within the Mb-fold is indicated in parenthesis and with the ♦ symbol. Highly conserved amino acids are indicated with black background.

3.2. Sequence alignment of the Zeamaymphytogb proteins

Pairwise sequence alignment showed that Zeamaymphytogb1.1 and Zeamaymphytogb1.2 are 52.87% identical and that they are different between each other mostly because Zeamaymphytogb1.1 contains a pre-helix A and Zeamaymphytogb1.2 contains a post-helix H. The Zeamaymphytogbs1 contain distal (H73/64) and proximal (H108/99) His and Phe CD1 (F40/30) and B10 (F54/44.) (Figure 1A). This analysis also showed that Zeamaymphytogb1.1 and Zeamaymphytogb1.2 are 9.35 and 14.61% identical to Zeamaymphytogb3, respectively, and that Zeamaymphytogb3 and *Oryza sativa* var. japonica Phytogb3 (Orysatjphytogb3) are highly conserved (*i.e*. they are 84% identical). In turn, the Zeamaymphytogb3 globin domain contains proximal His (H100), proposed distal Trp (W113), Phe CD1 (F61) and Tyr B10 (Y46) and is flanked by 25 and 23 amino acid extensions at the N- and C-terminal, respectively (Figure 1B). With the exception of the existence of a post-helix H in Zeamaymphytogb1.2, the above observations showed that the main characteristics that define to land plant Phytogb sequences (such as the existence of His, Phe and Tyr at the protein´s heme pocket) were conserved during the evolution of the Zeamaymphytogb sequences.

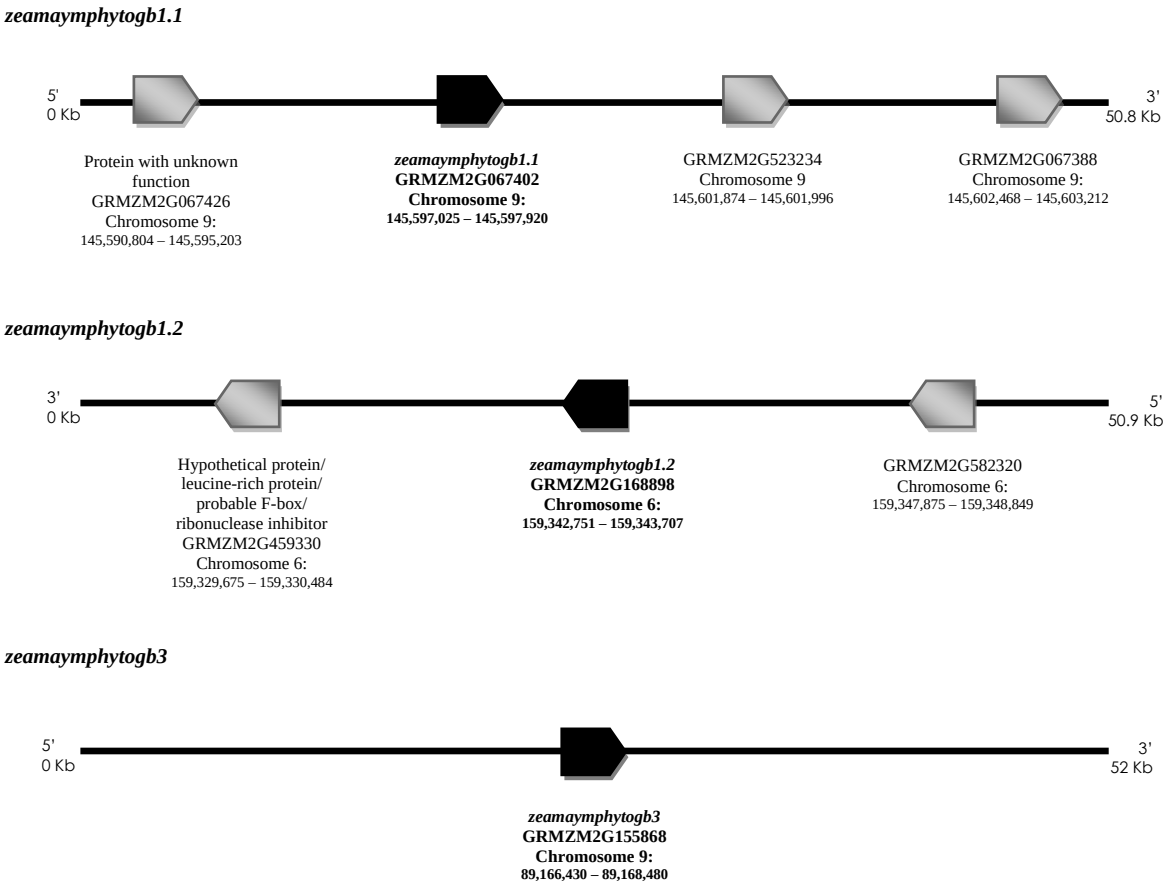

Figure 2. Mapping of *zeamaymphytogb* genes into the maize genome. The *zeamaymphytogb* (black) and flanking (gray) genes were mapped within ~50 kb fragments of the maize genome using the Phytozome server and the *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* sequences as probe. Arrows indicate the transcription orientation. Information for each gene corresponds to predicted protein (following the Phytozome nomenclature), locus name into the maize genome and position at the maize chromosome. Gene sizes and distance between genes are not shown at scale.

3.3. Mapping of the *zeamaymphytogb* genes within the maize genome

Mapping analysis revealed that *zeamaymphytogb1.1* and *zeamaymphytogb3* are localized at positions 145,597,025 to 145,597,920 and 89,166,430 to 89,168,480 of maize chromosome 9, respectively, whereas *zeamaymphytogb1.2* is localized at position 159,342,751 to 159,343,707 of maize chromosome 6 (Figure 2). This observation revealed that *zeamaymphytogb*s do not form clusters but are rather dispersed within the maize genome. This analysis also revealed that *zeamaymphytogb1.1* and *zeamaymphytogb3* are localized in the + strand with a 5′ → 3′orientation and *zeamaymphytogb1.2* is localized in the – strand with a 5′ → 3′ orientation. With the exception of few ORFs mostly coding for unknown or hypothetical proteins located far from *zeamaymphytogbs1*, no ORFs coding for known proteins were detected up- and downstream to *zeamaymphytogb*s (Figure 2). Hence, the above observations suggested that *zeamaymphytogb*s might not coexpress with each other and with genes different to *zeamaymphytogb*s (*i.e*. their expression could be regulated by separate).

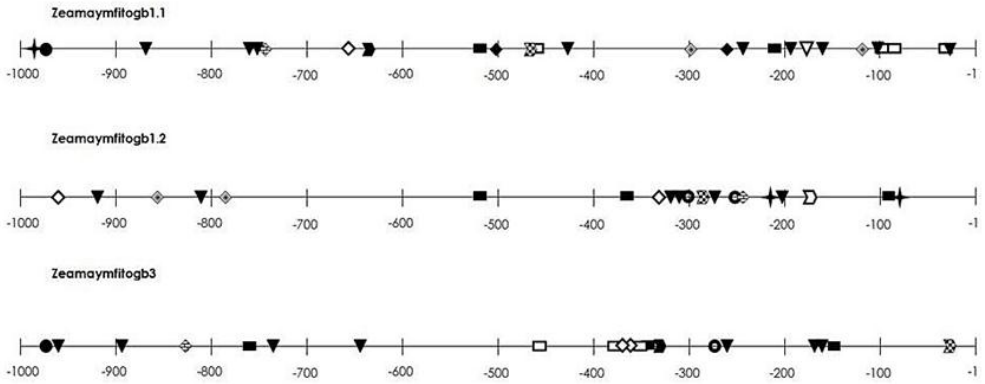

Figure 3. Regulatory sequences located upstream to the *zeamaymphytogb* genes. ■, TATA-box; □, CAAT-box. *Cis*-acting regulatory elements: ◆, Auxins; ◇, Citokinins; ◈, Abscisic acid; ◈, Methyl jasmonate; ▼, Light; ▽, Circadian cycle; ◗, Heat; ◨, Low temperature; ◪, Anaerobiosis; ●, Specific activation in meristems; ◉, Specific expression in the endosperm

3.4. Detection of promoter sequences upstream to the *zeamaymphytogb* genes

Mapping analysis (see subsection 3.3) suggested that expression of *zeamaymphytogb*s is regulated by different promoters. Previously, Hill *et al*. [16] identified some putative regulatory elements responsive to plant hormones and stress conditions upstream to *zeamaymphytogb1.1* and *zeamaymphytogb1.2* using bioinformatics tools. However, the existence of regulatory boxes and number and position of *cis*-acting regulatory elements upstream to *zeamaymphytogb*s was not reported by the above authors. Thus, we identified the number and position of TATA-, CAAT- and GC-boxes and *cis*-acting regulatory elements responsive to the plant hormones auxins, cytokinins, methyl jasmonate and abscisic acid, the stress conditions light, circadian cycle, temperature and anaerobiosis and the specific expression in meristems and endosperm within 1kb upstream to *zeamaymphytogb*s (supplementary Table 1). Our results showed that 2 and 3 TATA-boxes exist upstream to *zeamaymphytogb1.1* and *zeamaymphytogb1.2* and *zeamaymphytogb3*, 4 and 3 CAAT-boxes exist upstream to *zeamaymphytogb1.1* and *zeamaymphytogb3*, and CAAT- and GC-boxes do not exist 1kb upstream to *zeamaymphytogb1.2* and *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, respectively (Figure 3 and supplementary Table 1). Also, *cis*-acting regulatory elements responsive to auxins were only detected upstream to *zeamaymphytogb1.1*, *cis*-acting regulatory elements responsive to cytokinins were detected upstream to *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, *cis*-acting regulatory elements responsive to abscisic acid were only detected upstream to *zeamaymphytogb1.1* and *zeamaymphytogb1.2*, a *cis*-acting regulatory element responsive to methyl jasmonate was detected upstream to *zeamaymphyogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, *cis*-acting regulatory elements responsive to light were detected upstream to *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, a *cis*-acting regulatory element responsive to the circadian cycle was only detected upstream to *zeamaymphytogb1.1*, a *cis*-acting regulatory element responsive to heat was only detected upstream to *zeamaymphytogb1.1* and *zeamaymphytogb3*, a *cis*-acting regulatory element responsive to low temperature was only detected upstream to *zeamaymphytogb1.2*, a *cis*-acting regulatory element responsive to anaerobiosis was detected upstream to *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, a *cis*-acting regulatory element that is specifically activated in meristems was detected upstream to *zeamaymphytogb1.1* and *zeamaymphytogb3*, and *cis*-acting regulatory elements that are specifically activated in the endosperm were detected upstream to *zeamaymphytogb1.2* and *zeamaymphytogb3* (Figure 3 and supplementary Table 1).

The above observations suggested that the expression of *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* in normal plants is regulated by CAAT- and/or TATA-boxes. Hence, it is possible that Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 are synthesized and function in diverse organs from normal maize (see subsection 3.5). Likewise, apparently auxins only regulate the expression of *zeamaymphytogb1.1*, cytokinins and methyl jasmonate regulate the expression of *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3*, and abscisic acid only regulates the expression of *zeamaymphytogb1.1* and *zeamaymphytogb1.2*. Also, it is possible that *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* express in maize growing under light-stress conditions and anaerobiosis, only *zeamaymphytogb1.1* expresses during the circadian cycle, *zeamaymphytogb1.1* and *zeamaymphytogb3* and *zeamaymphytogb1.2* express in maize growing under low and high temperature, respectively, only *zeamaymphytogb1.1* and *zeamaymphytogb3* express in the maize meristem and only *zeamaymphytogb1.2* and *zeamaymphytogb3* express in the maize endosperm. Hence, a possible scenario derived from these observations is that *zeamaymphytogb*s differentially express according to the maize physiology and growing conditions.

Table 1. Absolute levels of the Zeamaymphytogb transcripts in maize organs/tissues from plants grown under normal conditions. Approximate values correspond to the color scale reported by the Maize eFP Browser. Numeric scale for seed organs/tissues, mature leaves and (young and mature) roots correspond to 0.0-6079.77, 0.0-47.84 and 0.0-3141.43, respectively. Note that values are different among organs and thus are not comparable among seeds, leaves and roots but among Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 within the same organ. ND, non-detected.

MAIZE ORGAN/TISSUE	TRANSCRIPT LEVELS		
	Zeamaymphytogb1.1	Zeamaymphytogb1.2	Zeamaymphytogb3
Seeds (2 to 24 days after pollination)	0.0 to 607.97	0.0 to 607.97	1215.95 to 1823.93
Embryo	0.0 to 607.97	2431.9 to 3039.88	0.0 to 607.97
Endosperm	0.0 to 607.97	0.0 to 607.97	1215.95 to 1823.93
Germinating seeds (24 h)	607.97 to 1215.95	5471.79 to 6079.77	607.97 to 1215.95
Coleoptile	0.0 to 607.97	0.0 to 607.97	0.0 to 607.97
Seminal roots	0.0 to 607.97	0.0 to 607.97	0.0 to 607.97
Mature leaves:			
Proximal (0 to 1 cm) region	43.05 to 47.84	0.0 to 4.78	4.78 to 9.56
Middle (4 to 10 cm) region	4.78 to 9.56	33.48 to 43.05	19.13 to 33.48
Distal (12 to 15 cm) region	4.78 to 9.56	43.05 to 47.84	43.05 to 47.84
Roots:			
Young	ND	314.14 to 2199.0	314.14 to 2199.0
Mature	ND	0.0 to 1570.71	1570.71 to 2827.28

Table 2. Relative levels of the Zeamaymphytogb transcripts in maize organs/tissues induced by plant hormones and stress conditions. See subsection 2.4 for experimental details.

	MAIZE ORGAN/ TISSUE	TRANSCRIPT LEVELS (%)		
		Zeamaymphytogb1.1	Zeamaymphytogb1.2	Zeamaymphytogb3
Hormone induction				
Auxins	Mesocotyls	12	2	2
Giberellins	Mesocotyls	11	1	2
Abiotic stress				
Drought	Plantlets	13	18	6
Acid (pH 4.1) soil	Root tips	12	13	3
	Leaves	7	11	6
Biotic stress				
Symbiosis with arbuscular mycorhizae	Seminal roots	23	11	1
Pathogen infection by				
Ustilago maydis	Leaves	9	15	8
Fusarium moniliforme	Whole plant	15	7	4
Phytophtora cinnamomi	Seminal roots	9	4	1
Colletotrichum graminicola	Leaves	16	5	1
Sesamia nonagroides.	Husk	1	3	2
Sporisoriu reilianum	Leaves	10	8	5
Meloidogyne incognita	Roots	7	17	5

3.5 Detection of Zeamaymphytogb transcripts in maize organs

Transcripts [44, 45] and proteins [27] for Zeamaymphytogbs have been detected in organs from normal maize and in maize subjected to either hormonal treatments or stress conditions. However, the above reports were not specific to Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 and therefore it was not known which Zeamaymphytogb is synthesized in maize organs. Thus, we identified Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3

transcripts in maize organs from normal plants and in maize subjected to either hormonal treatments or stress conditions using databases (see subsection 2.4). Our results showed that Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 transcripts exist in embryonic and vegetative organs/tissues from maize grown in normal conditions (Table 1). Specifically, Zeamaymphytogb1.1 transcripts exist at low (0.0 to 607.97) to moderate (607.97 to 1215.95) levels in embryonic organs/tissues and moderate (4.78 to 9.56) to high (43.05 to 47.84) levels in mature leaves. Transcripts for Zeamaymphytogb1.1 were not detected in young and mature roots. Transcripts for Zeamaymphytogb1.2 exist at low to high levels in embryonic organs/tissues (0.0 to 6079.77), mature leaves (0.0 to 47.84) and young and mature roots (0.0 to 2199.0). Transcripts for Zeamaymphytogb3 exist at low to moderate levels in embryonic organs/tissues (0.0 to 1823.93), low to high levels in mature leaves (4.78 to 47.84) and moderate to high levels in young and mature roots (314.14 to 2827.28). Furthermore, table 2 shows that transcripts for Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 exist in maize organs subjected to either hormonal treatments or stress conditions. The highest (≥11%) levels of transcripts corresponded to Zeamaymphytogb1.1 from maize organs treated with the plant hormones auxins and giberellins and subjected to the abiotic and biotic stress drought and acid soil and infection with arbuscular mycorhizae and the pathogens *Fusarium moniliforme* and *Colletotrichum graminicola*, and to Zeamaymphytogb1.2 from maize organs subjected to the abiotic and biotic stress drought and acid soil and infection with arbuscular mycorhizae and the pathogens *Ustilago maydis* and *Meloidogyne incognita*, respectively. The lowest (1 to 8%) levels of transcripts corresponded to Zeamaymphytogb3 from maize organs subjected to either hormonal treatments or abiotic and biotic stress.

The above observations showed that transcripts for Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 widely occur in embryonic and vegetative organs/tissues from normal maize and Zeamaymphytogb1.1 and Zeamaymphytogb1.2 transcripts exist at higher levels than those for Zeamaymphytogb3 in maize subjected to either hormonal treatments or stress conditions. Hence, a possible scenario derived from these observations is that Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 function in multiple organs from normal maize and Zeamaymphytogb1.1 and Zeamaymphytogb1.2 function over Zeamaymphytogb3 in the maize response to hormonal treatments and stress conditions.

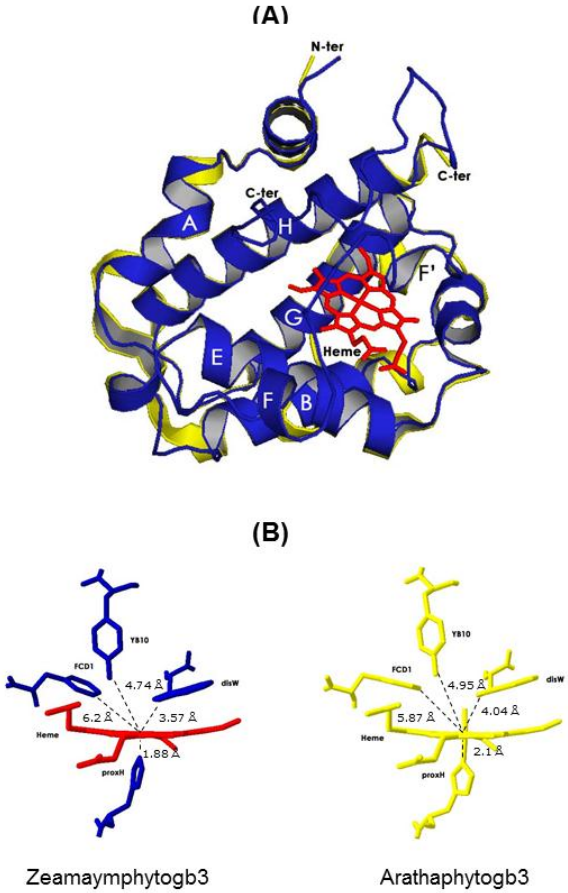

Figure 4. (A) Overlapping of the Zeamaymphytogb3 predicted structure (blue) and Arathaphytogb3 crystal structure (yellow). Helices are indicated with letters A to H. (B) Distance of proximal H and proposed distal W, FB10 and FCD1 to the Zeamaymphytogb3 and Arathaphytogb3 heme-Fe.

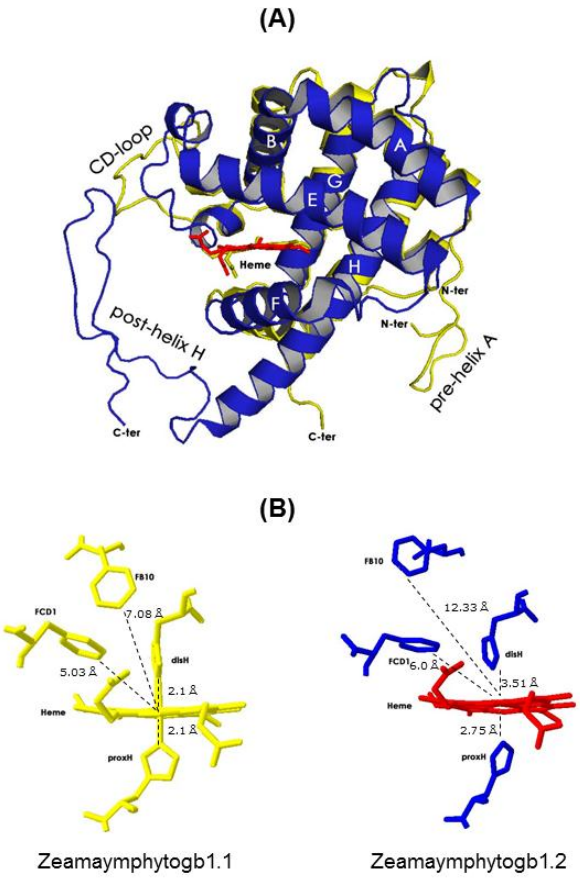

Figure 5. (A) Overlapping of the Zeamaymphytogb1.1 (yellow) and Zeamaymphytogb1.2 (blue) predicted structures. Helices are indicated with letters A to H. (B) Distance of proximal and distal H, FB10 and FCD1 to the Zeamaymphytogb1.1 and Zeamaymphytogb1.2 heme-Fe.

3.6 Modeling of the predicted Zeamaymphytogb3 and analysis of the Zeamaymphytogbs predicted tertiary structure

Land plant Phytogbs3 were discovered in 2001 by Watts *et al*. [22] but have been poorly characterized from a structural and functional perspective. The first crystal structure for a land plant Phytogb3 corresponding to that of Arathaphytogb3 was elucidated in 2014 by Reeder and Hough [46], however no other Phytogb3 tertiary structure has been reported. Thus, we modeled and analyzed the tertiary structure of Zeamaymphytogb3 using bioinformatics tools and the crystal structure of Arathaphytogb3 as template (see subsection 2.5). Analysis using the Verify3D server showed that most (92.98%) of the amino acids from the selected model were positive indicating that the predicted structure for Zeamaymphytogb3 was reliable. The predicted structure of Zeamaymphytogb3 was deposited in the Caspur PMDB database with the ID number PM0080450. Figure 4A shows the predicted structure of Zeamaymphytogb3 overlapped to the crystal structure of Arathaphytogb3. Folding of Zeamaymphytogb3 is highly similar to that of Arathaphytogb3 (RMS = 0.56Å): Zeamaymphytogb3 folds into a 2/2-folding globin domain consisting of helices A to H flanked by N- and C-terminal extensions (see also Figure 1B). The N-terminal extension from Zeamaymphytogb3 folds into a short (2 to 3 turns) helix similarly to that from Arathaphytogb3. Heme orientation and distance of amino acids at the heme pocket to the heme-Fe is also highly similar in Zeamaymphytogb3 and Arathaphytogb3 (Figure 4B). These observations showed that the Phytogb3 structure conserved in monocots and dicots (represented here by Zeamaymphytogb3 and Arathaphytogb3, respectively) and suggested that biochemical properties (*e.g*. heme-Fe coordination, O_2-affinity and NO dioxygenase activity) and function of Zeamaymphytogb3 and Arathaphytogb3 are similar.

The crystal and predicted structure of Zeamaymphytogb1.1 (PDB ID 2R50) [28] and Zeamaymphytogb1.2 [25] have been elucidated. Our comparative analysis between the Zeamaymphytogb1.1 and Zeamaymphytogb1.2 structures (Figure 5A) showed that these proteins are highly similar (RMS = 1.35Å) and fold into a 3/3-folding consisting of helices A to H flanked by unfolded pre-helix A and post-helix H regions, respectively (see also Figure 1A). A major difference between the Zeamaymphytogb1.1 and Zeamaymphytogb1.2 structure is that CD-loop, which apparently provides flexibility to helix E where distal His is located [23, 28], is extended and disorganized in Zeamaymphytogb1.1 and rather compact and folded into a short (1 to 2 turns) helix in Zeamaymphytogb1.2.

Also, heme orientation and distance of proximal and distal His and FCD1 to the heme-Fe is highly similar in Zeamaymphytogb1.1 and Zeamaymphytogb1.2, however FB10 is located 7.08 and 12.33Å from the heme-Fe in Zeamaymphytogb1.1 and Zeamaymphytogb1.2, respectively (Figure 5B). These observations suggested that helix E is less flexible in Zeamaymphytogb1.2 than in Zeamaymphytogb1.1, movement of distal His is restricted in Zeamaymphytogb1.2 and heme-pocket conformation is rather open in Zeamaymphytogb1.2 compared to that from Zeamaymphytogb1.1. Hence, despite folding similarly into a 3/3-folding apparently Zeamaymphytogb1.1 and Zeamaymphytogb1.2 exhibit a number of structural differences between each other which could result in different biochemical properties and function in maize.

4. CONCLUSION

Maize Phytogbs had been characterized previously to this work, however knowledge on Zeamaymphytogbs was limited. Thus, we used bioinformatics tools to complement the knowledge on Zeamaymphytogbs. In summary, our results indicated that (i) only *zeamaymphytogb1.1*, *zeamaymphytogb1.2* and *zeamaymphytogb3* exist in the maize genome, (ii) *zeamaymphytogb*s are interrupted by 3 introns similarly to other land plant *phytogb*s, (iii) the main characteristics that define to land plant Phytogb sequences were conserved during the evolution of the Zeamaymphytogb sequences, (iv) the expression of *zeamaymphytogb*s could be regulated by separate, (v) *zeamaymphytogb*s could differentially express according to the maize physiology and growing conditions, (vi) Zeamaymphytogb1.1, Zeamaymphytogb1.2 and Zeamaymphytogb3 transcripts occur in embryonic and vegetative organs/tissues from normal maize, (vii) Zeamaymphytogb1.1 and Zeamaymphytogb1.2 transcripts occur at higher levels than those of Zeamaymphytogb3 in maize subjected to hormonal treatments and stress conditions, (viii) predicted Zeamaymphytogb3 and Arathaphytogb3 exhibit a highly similar structure and could function similarly, and (ix) Zeamaymphytogb1.1 and Zeamaymphytogb1.2 exhibit structural differences between each other which could result in different biochemical properties and function in maize plants. The information reported here allows a better understanding of Zeamaymphytogbs and might be useful for carrying out further experimental work.

Acknowledgements

This work was partially financed by SEP-PROMEP (grant number UAEMor-PTC-01-01/PTC23) and Consejo Nacional de Ciencia y Tecnología (CoNaCyT grant numbers 25229N and 42873Q), México

5. REFERENCES

1. Giardina, B., Messana, I., Scatena, R. and Castagnola, M. (1995) The multiple functions of hemoglobin, Crit. Rev. Biochem. Mol. Biol. 30, 165-196.
2. Weber, R. and Vinogradov, S. N. (2001) Nonvertebrate hemoglobins: functions and molecular adaptations, Physiol. Rev. 81, 569-628.
3. Vinogradov, S. N., Hoogewijs, D., Bailly, X., Arredondo-Peter, R., Gough, J., Dewilde, S., Moens, L. and Vanfleteren, J. R. (2006) A phylogenomic profile of globins, BMC Evol. Biol. 6, 31-47.
4. Vinogradov, S. N., Hoogewijs, D., Bailly, X., Arredondo-Peter, R., Guertin, M., Gough, J., Dewilde, S., Moens, L. and Vanfleteren, J. R. (2005) Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life, Proc. Natl. Acad. Sci. USA. 102, 11385-11389.
5. Vinogradov, S. N., Fernández, I., Hoogewijs, D. and Arredondo-Peter, R. (2011) Phylogenetic relationships of 3/3 and 2/2 hemoglobins in Archaeplastida genomes to bacterial and other eukaryote hemoglobins, Mol. Plant. 4, 42-58.
6. Vázquez-Limón, C., Hoogewijs, D., Vinogradov, S. N. and Arredondo-Peter, R. (2012) The evolution of land plant hemoglobins, Plant Sci. 191-192, 71-81.
7. Hill, R., Hargrove, M. S. and Arredondo-Peter, R. (2016) Phytoglobin: a novel nomenclature for plant globins accepted by the globin community at the 2014 XVIII conference on Oxygen-Binding and Sensing Proteins, F1000Research. 5, 212.
8. Garrocho-Villegas, V., Gopalasubramaniam, S. K. and Arredondo-Peter, R. (2007) Plant hemoglobins: what we know six decades after their discovery, Gene: Funct. Evol. Genom. 398, 78-85.
9. Appleby, C. A. (1992) The origin and functions of haemoglobin in plants, Sci. Progress 76, 365-398.
10. Garrocho-Villegas, V. and Arredondo-Peter, R. (2008) Molecular cloning and characterization of a moss (*Ceratodon purpureus*) non-symbiotic hemoglobin provides insight into the early evolution of plant non-symbiotic hemoglobins, Mol. Biol. Evol. 25, 1482-1487.
11. Lira-Ruan, V., Ruiz-Kubli, M. and Arredondo-Peter, R. (2011) Expression of non-symbiotic hemoglobin 1 and 2 genes in rice (*Oryza sativa*) embryonic organs, Commun. Integrat. Biol. 4, 457-458.
12. Lira-Ruan, V., Sarath, G., Klucas, R. V. and Arredondo-Peter, R. (2001) Synthesis of hemoglobins in rice (*Oryza sativa* var. Jackson) plants growing in normal and stress conditions, Plant Sci. 161, 279-287.
13. Ross, E. J. H., Shearman, L., Mathiesen, M., Zhou, J., Arredondo-Peter, R., Sarath, G. and Klucas, R. V. (2001) Non-symbiotic hemoglobins are synthesized during germination and in differentiating cell types, Protoplasma 218, 125-133.
14. Trevaskis, B., Watts, R. A., Andersson, S. R., Llewellyn, D. J., Hargrove, M. S., Olson, J. S., Dennis, E. S. and Peacock, W. J. (1997) Two hemoglobin genes in *Arabidopsis thaliana*: the evolutionary origins of leghemoglobins, Proc. Natl. Acad. Sci. USA. 94, 12230-12234.
15. Hill, R. D. (2012) Non-symbiotic haemoglobins-What´s happening beyond nitric oxide scavenging?, AoB Plants. Pls004, doi: 10.1093/aobpla/pls004.
16. Hill, R. D., Huang, S. and Stasolla, C. (2013) Hemoglobins, programmed cell death and somatic embryogenesis, Plant Sci. 211, 35-41.
17. Stasolla, C. and Hill, R. D. (2017) Determining cellular responses: phytoglobins may direct the traffic, Trends Plant Sci. 22, 820-822.
18. Arredondo-Peter, R., Hargrove, M. S., Sarath, G., Moran, J. F., Lohrman, J., Olson, J. S. and Klucas, R. V. (1997) Rice hemoglobins: gene cloning, analysis and oxygen-binding kinetics of a recombinant protein synthesized in *Escherichia coli*, Plant Physiol. 115, 1259-1266.
19. Duff, S. M. G., Wittenberg, J. B. and Hill, R. D. (1997) Expression, purification and properties of recombinant barley (*Hordeum* sp.) hemoglobin: optical spectra and reactions with gaseous ligands, J. Biol. Chem. 272, 16746-16752.
20. Kumar, N., Astegno, A., Chen, J., Giorgetti, A. and Dominici, P. (2016) Residues in the distal heme pocket of Arabidopsis non-symbiotic hemoglobins: Implication for

nitrite reductase activity, Int. J. Mol. Sci. 17, doi:10.3390/ijms17050640.
21. Vázquez-Limón, C., Castro-Bustos, S. and Arredondo-Peter, R. (2012) Spectroscopic analysis of moss (*Ceratodon purpureus* and *Physcomitrella patens*) recombinant non-symbiotic hemoglobins, Comm. Integ. Biol. 5, 527-530.
22. Watts, R. A., Hunt, P. W., Hvitved, A. N., Hargrove, M. S., Peacock, W. J. and Dennis, E. S. (2001) A hemoglobin from plants homologous to truncated hemoglobins of microorganisms, Proc. Natl. Acad. Sci. USA. 98, 10119-10124.
23. Hargrove, M., Brucker, E. A., Stec, B., Sarath, G., Arredondo-Peter, R., Klucas, R. V., Olson, J. S. and PhilipsJr., G. N. (2000) Crystal structure of a non-symbiotic hemoglobin, Structure. 8, 1005-1014.
24. Larsen, K. (2003) Molecular cloning and characterization of cDNAs encoding hemoglobin from wheat (*Triticum aestivum*) and potato (*Solanum tuberosum*), Biochim. Biophys. Acta. 1621, 299-305.
25. Rodríguez-Alonso, G. and Arredondo-Peter, R. (2012) Phylogenetic analysis reveals an apparent duplication of the non-symbiotic hemoglobin 1 gene early in the evolution of monocotyledonous plants, ScienceJet. 1, 27.
26. Rodríguez-Alonso, G. and Arredondo-Peter, R. (2013) Variability of non-symbiotic and truncated hemoglobin genes from the genome of cultivated monocots, Comm. Integr. Biol. 6, e27496.
27. Aréchaga-Ocampo, E., Sáenz-Rivera, J., Sarath, G., Klucas, R. V. and Arredondo-Peter, R. (2001) Cloning and expression analysis of hemoglobin genes from maize (*Zea mays* ssp. *mays*) and teosinte (*Zea mays* ssp. *parviglumis*), Biochim. Biophys. Acta: Gene Struct. Expr. 1522, 1-8.
28. Sáenz-Rivera, J., Sarath, G. and Arredondo-Peter, R. (2004) Modeling the tertiary structure of a maize (*Zea mays* ssp. *mays*) non-symbiotic hemoglobin, Plant Physiol. Biochem. 42, 891-897.
29. Shi, J., Blundell, T. and Miziguchi, K. (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties, J. Mol. Biol. 310, 243-257.
30. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucl. Acids Res. 24, 4876-4882.
31. Kapp, O. H., Moens, L., Vanfleteren, J., Trotman, C. N. A., Suzuki, T. and Vinogradov, S. N. (1995) Alignment of 700 globin sequences: extent of amino acid substitution and its correlation with variation in volume, Prot. Sci. 4, 2179-2190.
32. Lesk, A. M. and Chothia, C. (1980) How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins, J. Mol. Biol. 136, 225-270.
33. Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S. L., Kebrom, T. H., Provart, N., Patel, R., Myers, C. R., Reidel, E. J., Turgeon, R., Liu, P., Sun, Q., Nelson, T. and Brutnell, T. P. (2010) The developmental dynamics of the maize leaf transcriptome, Nature Genetics. 42, 1060-1067.
34. Dash, S., Hemert, J. V., Hong, L., Wise, R. P. and Dickerson, J. A. (2012) PLEXdb: gene expression resources for plants and plant pathogens, Nucl. Acids Res. 40, D1194-D1201.
35. Roy, A., Kucukural, A. and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction, Nature Protocols. 5, 725-737.
36. Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2011) A protocol for computer-based protein structure and function prediction, J. Visual. Exp. 57, e3259.
37. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction, BMC Bioinformatics. 9, 40.
38. Eisenberg, D., Lüthy, R. and Bowie, J. U. (1997) VERIFY3D: assessment of protein models with three-dimensional profiles, Meth. Enzymol. 277, 396-404.
39. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD-Visual molecular dynamics, J. Mol. Graph. 14, 33-38.
40. Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and Swiss-PdbViewer: An environment for comparative protein modeling, Electrophoresis. 18, 2714-2723.
41. Gopalasubramaniam, S. K., Garrocho-Villegas, V., Bustos, G., Pastor, N. and Arredondo-Peter, R. (2008) Use of *in silico* (computer) methods to predict and analyze the tertiary structure of plant hemoglobins, Meth. Enzymol. 436, 393-410.
42. Appleby, C. A., Dennis, E. S. and Peacock, W. J. (1990) A primaeval origin for plant and animal hemoglobins?, Aust. Syst. Bot. 3, 81-89.
43. Martínez-Ocampo, F., Vázquez-Limón, C. and Arredondo-Peter, R. (2012) Detection, PCR-amplification and characterization of a 4 intron-containing non-symbiotic hemoglobin gene from the moss *Physcomitrella patens*, Global J. Biochem. 3, 12.
44. Huang, S., Hill, R. D. and Stasolla, C. (2014) Plant hemoglobin participation in cell fate determination, Plant Signal Behav., e29485-1.
45. Zhao, L., Gu, R., Gao, P. and Wang, G. (2008) A nonsymbiotic hemoglobin gene from maize, *ZmHb*, is involved in response to submergence, high-salt and osmotic stresses, Plant Cell Tiss. Organ Cult. 95, 227-237.
46. Reeder, B. J. and Hough, M. A. (2014) The structure of a class 3 nonsymbiotic plant haemoglobin from *Arabidopsis thaliana* reveals a novel N-terminal helical extension, Acta Cryst. D70, 1411-1418.
47. Arredondo-Peter, R., Moran, J. F. and Sarath, G. (2014) Rice (*Oryza*) hemoglobins, F1000Research. 3, 253.

Access supplementary Table-1 and Figure-1 at
http://bioinfo.aizeonpublishers.net/content/2018/3/bioinfo991-1001_suppl_table-1.pdf and
http://bioinfo.aizeonpublishers.net/content/2018/3/bioinfo991-1001_suppl_fig-1.pdf

*****